An optical distance sensor : tilt robust differential confocal measurement with mm range and nm uncertainty

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

An Optical Distance Sensor

tilt robust differential confocal measurement with mm range and nm uncertainty

Lennino Cacace

An Optical Distance Sensor

Tilt robust differential confocal measurement with mm range and nm uncertainty

PROEFONTWERP

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op dinsdag 1 december 2009 om 16.00 uur

door

Leonard Antonino Cacace

geboren te Middelburg

De documentatie van het proefontwerp is goedgekeurd door de promotor:

prof.dr.ir. M. Steinbuch

Copromotor:
dr.ir. P.C.J.N. Rosielle

SUMMARY

High-end optical systems that incorporate freeform optics can offer many advantages over systems that apply conventional optics only. The widespread use of freeforms is held back however, because a suitable measurement method is not available. The NANOMEFOS project aimed at realizing a universal freeform measurement machine to fill that void. The principle of operation of this machine required a novel sensor for surface distance measurement, the development and realization of which is the objective of the work presented in this thesis.

The sensor must enable non-contact, absolute distance measurement of surfaces with reflectivities from 3.5% to 99% over 5 mm range, with 1 nm resolution and a 2σ measurement uncertainty of 10 nm for surfaces perpendicular to the measurement direction and 35 nm for surfaces with tilts up to 5°. To meet these requirements, a dual-stage design is proposed: a primary measurement system tracks the surface under test by focusing its object lens, while the secondary measurement system measures the displacement of this lens. After an assessment of various measurement principles through comparison of characteristics inherent to their principle of operation and the potential for adaptation, differential confocal measurement has been selected as the primary measurement method. Dual-pass heterodyne interferometry is used as secondary measurement method. To allow for correction of tilt dependent error by calibration, a third system that measures through which part of the aperture the light returns is integrated.

An analytical model of the differential confocal measurement principle has been derived to enable optimization. To gain experience with differential confocal measurement, a demonstrator was built, which has resulted in insights and design rules for prototype development. The models show satisfactory agreement with the experimental results obtained with the demonstrator, thus building confidence that the models can be applied as design and optimization tools. Various properties that characterize the performance of a differential confocal measurement system are

identified. Their dependence on the design parameters has been studied through numerical simulations based on the analytical models. The results of this study are applied to optimize the sensor for use in NANOMEFOS.

A design is presented in which many of the optics of the interferometer and the differential confocal system are bonded to form one optical monolith. The benefits of this design include a reduction of ghost reflections, improved stability and reduced alignment effort compared to a conventional design. To obtain a system that fits the allotted volume envelope, many components are custom made and the optical path of the differential confocal system is folded using prisms and mirrors.

The optomechanical and mechatronic design incorporates a custom focusing unit to enable surface tracking. This unit consists of a rotationally symmetric, elastic guidance mechanism and a voice coil actuator. The lateral position of the guidance mechanism reproduces within 20 nm, and it is expected from the frequency response that a control bandwidth of at least 800 Hz can be realized. The power dissipated during measurement depends on the form of the freeform surface; for most surfaces anticipated, it is in the order of a few milliwatts.

For signal processing, and to drive the laser and the focusing unit, partly custom electronics are used. Control strategies for interferometer nulling, focus locking and surface tracking have been developed, implemented and tested.

The sensor realized has 5 mm range, -2.5 µm to 1.5 µm tracking range, sub-nanometer resolution, and a small-signal bandwidth of 150 kHz. Calibrations are performed to achieve the required measurement uncertainty. A new method is developed to calibrate the dependency of the sensor on surface tilt. This method does not rely on reference artifacts, and it can be employed to calibrate other types of optical distance sensors as well. Based on experiments, the 2σ measurement uncertainty after calibration is estimated to be 4.2 nm for measurement of rotationally symmetric surfaces, 21 nm for measurement of medium freeform surfaces and 34 nm for measurement of heavily freeform surfaces.

To test the performance of the machine with the sensor installed, measurements of a tilted flat have been carried out. In these measurements, a tilted flat serves as a reference freeform with known surface form. The measurement results demonstrate the reduction of tilt dependent error using the new calibration method.

A tilt robust, single point distance sensor with millimeter range and nanometer uncertainty was developed, realized and tested. It is installed in the freeform measurement machine for which it has been designed and is currently used for the measurement of optical surfaces. By applying the simulations based on the analytical models, the sensor can be optimized for other applications as well.

SAMENVATTING

Hoogwaardige optische systemen die freeforms bevatten kunnen veel voordelen hebben ten opzichte van systemen waarin alleen conventionele optieken worden toegepast. Een wijdverbreid gebruik van hoogwaardige freeforms in enkelstuks systemen wordt echter geremd door het ontbreken van een geschikte meetmethode. Het NANOMEFOS project had daarom als doel een universele freeform meetmachine te ontwikkelen. Het werkingsprincipe van deze machine vereiste het gebruik van een nieuwe afstandssensor, waarvan de ontwikkeling en realisatie het doel is van het in dit proefschrift gepresenteerde werk.

De sensor moet het mogelijk maken om contactloos de afstand te meten van oppervlakken met een reflectiviteit van 3,5% tot 99% over een bereik van 5 mm. De vereiste resolutie is 1 nm en de 2σ meetonzekerheid is 10 nm voor oppervlakken die loodrecht staan op de meetrichting en 35 nm voor oppervlakken die tot 5° zijn gekanteld. Om dit te bereiken is er een tweetraps meetsysteem ontwikkeld waarin een primair meetsysteem het te meten oppervlak volgt door een lens te focusseren, terwijl een secundair meetsysteem de axiale verplaatsing van deze lens meet. Diverse optische meetprincipes zijn geëvalueerd door vergelijking van de eigenschappen die inherent zijn aan de werkingsprincipes. Op basis hiervan is gekozen voor het gebruik van het differentieel confocale principe in het primaire meetsysteem. Het secundaire meetprincipe berust op dual-pass heterodyne interferometry. Om het mogelijk te maken om door middel van kalibratie de hoekafhankelijkheid van de afstandsmeting te corrigeren, is er een derde meetsysteem geïntegreerd dat meet door welk deel van de apertuur het licht terugkeert.

Er is een analytisch model van het differentieel confocale meetprincipe opgesteld dat het mogelijk maakt het systeem te optimaliseren. Bovendien is er een demonstrator gebouwd om ervaring op de doen met deze meetmethode. Dit heeft geresulteerd in inzichten en ontwerp regels welke zijn toegepast bij het

ontwikkelen van het prototype. Resultaten verkregen met de opgestelde modellen vertonen goede overeenkomsten met die van de demonstrator, hetgeen vertrouwen wekt in de toepasbaarheid van de modellen voor ontwerp en optimalisatie. Er zijn eigenschappen gedefinieerd die de prestaties van een differentieel confocaal systeem karakteriseren. De samenhang van deze eigenschappen met de relevante ontwerpparameters is vervolgens onderzocht door middel van simulaties gebaseerd op het analytische model. De resultaten van deze studie zijn gebruikt om de sensor te optimaliseren voor toepassing in NANOMEFOS.

Veel van de optische componenten van de interferometer en het differentieel confocale systeem zijn verlijmd met optische kit en vormen één monoliet. De voordelen van dit ontwerp zijn onder meer een vermindering van het aantal ongewilde reflecties aan glas-lucht overgangen, een verbeterde stabiliteit en een verminderde uitlijn inspanning in vergelijking met een traditioneel ontwerp. Om het systeem in de beschikbare ruimte te laten passen zijn veel onderdelen op maat gemaakt en is het optische pad gevouwen door middel van spiegels en prisma's.

Het optomechanische en mechatronische ontwerp bevat onder andere een focusseerunit bestaande uit een rotatiesymmetrische elastische rechtgeleiding en een voice coil actuator. De laterale reproduceerbaarheid van het geleidingsmechanisme is 20 nm, en gebaseerd op de frequentieresponsie is de verwachting dat een bandbreedte van op zijn minst 800 Hz haalbaar is. De dissipatie hangt af van de vorm en maat van het freeform oppervlak; voor de meeste verwachte oppervlakken is dit enkele milliwatts.

Er is elektronica ontworpen en gerealiseerd voor de signaal verwerking, het aansturen van de laser en de focusseerunit. Er zijn regelstrategieën ontwikkeld, toegepast en getest om de interferometer te nullen, om te focusseren en om het oppervlak te volgen.

De gerealiseerde sensor heeft 5 mm bereik, een volgbereik van -2,5 µm tot 1,5 µm, subnanometer resolutie en een bandbreedte van 150 kHz voor kleine signalen. Kalibraties zijn uitgevoerd om de gevraagde meetonzekerheid te halen. Een nieuwe methode is ontwikkeld om de hoekafhankelijkheid van de sensor te kalibreren. Deze methode maakt geen gebruik van een referentieartefact en is ook geschikt om andere typen sensoren mee te kalibreren. Gebaseerd op metingen is de 2σ meetonzekerheid na kalibratie geschat op 4,2 nm voor meting van rotatiesymmetrische oppervlakken, 21 nm voor meting van matig freeform oppervlakken en 34 nm voor meting van sterk freeform oppervlakken.

Om de prestaties van de machine met de geïnstalleerde sensor te testen zijn metingen van een schuin liggend vlakglas uitgevoerd. Bij deze metingen fungeert het vlakglas als referentie freeform waarvan de vorm met hoge nauwkeurigheid

bekend is. Deze metingen tonen de werking aan van de correctie van hoekafhankelijkheid door middel van de nieuwe kalibratiemethode.

Er is een hoekrobuuste, enkelpunts afstandssensor met millimeter bereik en nanometer onzekerheid ontwikkeld, gerealiseerd en getest. De sensor is geïnstalleerd in de meetmachine waarvoor hij is ontwikkeld en wordt momenteel gebruikt voor de meting van optische oppervlakken. Met behulp van de simulaties die gebaseerd zijn op de analytische modellen kan de sensor eenvoudig geoptimaliseerd worden voor andere toepassingen.

NOTATION AND ABBREVIATION

Coordinate systems
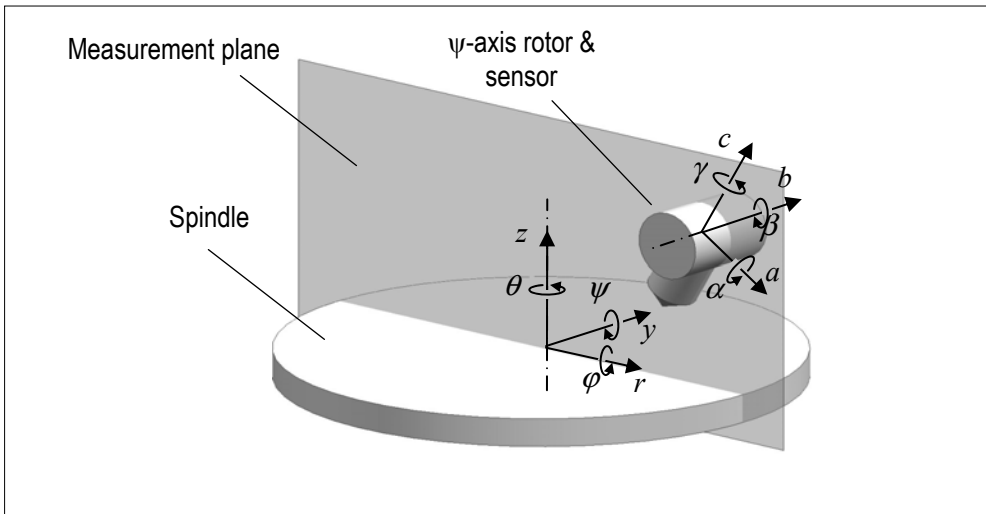
Figure 0.1: Definition of global coordinate system, relative to the machine base, and local coordinate system, relative to the sensor.

φ, ψ, θ	rotational axes of global coordinate system
r, y, z	translational axes of global coordinate system
α, β, γ	rotational axes of local coordinate system
a, b, c	translational axes of local coordinate system

The coordinate systems are discussed on page 7.

List of symbols

Symbol	Clarification	Unit
A	Local cross-sectional area of the beam	[m]
A_{pp}	peak-peak amplitude of focus unit trajectory	[m]
A_w	cross-sectional area of the wire	[m^2]
B	magnetic flux density	[T]
D	diameter at which the intensity is I_D	[m]
D_B	e^{-2}-diameter of the collimated laser beam	[m]
D_{ext}	diameter of the extended laser beam ($1.5 \cdot D_B$)	[m]
D_L	diameter of the aperture of the object lens	[m]
D_{ph}	pinhole diameter	[m]
$\overline{D}_{ph}$	pinhole diameter divided by D_0	[-]
D_z	e^{-2}-diameter of the beam at position z	[m]
D_0	e^{-2}-diameter of the beam at the waist	[m]
E	modulus of elasticity	[Pa]
FES	Focus Error Signal	[V] of [-]
$\overline{FES}$	dimensionless Focus Error Signal	[-]
$\mathbb{FES}$	normalized Focus Error Signal	[V] or [-]
$\overline{\mathbb{FES}}$	dimensionless normalized Focus Error Signal	[-]
F_{act}	actuator force	[F]
F_c	spring force in driving direction	[F]
F_{Imax}	maximum continuous force	[N]
F_{lz}	Lorentz force	[F]
F_{rms}	RMS value of the force exerted by the voice coil	[N]
F_{set}	force setpoint for focusing unit	[N]
FTP & $\overline{FTP}$	Fractionally Transferred Power	[-]
F_{vc}	force exerted by the voice coil	[N]
G_{el}	electrical processing gain	[V/A]
$\mathbb{G}_{el}$	electrical normalizing processing gain	[V]
I	current	[A]

I_D	intensity at diameter D in the beam	[W/m^2]
I_{max}	maximum continuous current	[I]
I_{ns1} & I_{ns2}	RMS value of shot noise in photocurrent of PDs	[I]
I_P	large quiescent photocurrent	[A]
I_{P1} & I_{P2}	large quiescent photocurrent of PD1 and PD2	[I]
I_{X1}, I_{X2}, I_{Y1} & I_{Y2}	Photocurrents through separate PSD x and y pins	[I]
I_{X1} & I_{X2}	large quiescent photocurrent of PD1 and PD2	[I]
I_0	intensity at the center of the beam	[W/m^2]
K	number of samples	[-]
K_f	force constant	[N/A]
K_m	motor constant	[N/√W]
L	number of track measurements	[-]
LSB_{rms}	RMS value of noise expressed in Least Significant Bit	[-]
M	magnification of optical system	[-]
M_1 & M_2	ψ-axis mass and moving mass, respectively	[kg]
M_m	moving mass	[kg]
NA_b	numerical aperture of the beam	[-]
NA_L	numerical aperture of the lens	[-]
P	pressure	[Pa]
P_b	power of the laser beam at the pinholes	[W]
P_{dis}	power dissipated in the coil	[W]
P_{enc}	encircled power	[W]
P_L	optical power of the laser	[W]
P_{mean}	average power dissipation in the focusing unit	[W]
R_{coil}	resistance of the coil	[Ω]
R_{pd}	responsivity of the photodiodes	[A/W]
R_{sut}	reflectivity of the SUT	[-]
S	sensitivity of the FES	[]
$\overline{S}$	dimensionless sensitivity of the FES	[-]
$\mathbb{S}$	sensitivity of the normalized FES	[]
$\overline{\mathbb{S}}$	dimensionless sensitivity of the normalized FES	[-]

T	temperature	[°C]
V_{amp}	Focusing unit amplifier setpoint	[V]
V_{coil}	conductor volume of the coil	[m³]
V_{FES}	Normalized FES output of processing electronics	[V]
V_{Imax}	maximum continuous voltage	[V]
V_{INT}	PSD intensity output of processing electronics	[V]
V_{na}	RMS value of ADC noise	[V]
V_{nc}	RMS value of the combined electrical noise	[V]
V_{ne}	RMS value of noise from processing electronics	[V]
V_{ns}	RMS value of shot noise remainder in the normalized FES voltage	[V]
$V_{on/off}$	Switch on, switch of signal for laser electronics unit	[V]
V_{pd1} & V_{pd2}	Normalized PD output of processing electronics	[V]
V_{setP}	Setpoint for laser power	[V]
V_{SUM}	PSS output of processing electronics	[V]
V_T, V_P & V_W	Environmental sensor electronics output for temperature, pressure and humidity	[V]
V_X & V_Y	PSD x and y output of processing electronics	[V]
W	partial pressure of water vapor	[Pa]
c_{ax}	axial stiffness of the guidance	[N/m]
d_{pt}	penetration depth	[m]
d_1	damping coefficient between M_1 and the fixed world	[Ns/m]
d_{12}	damping coefficient between M_1 and M_2	[Ns/m]
e_{pt}	measurement error due to penetration of the light in the PSD	[m]
e_1 & e_2	exponential terms associated with FTP at the pinholes	[-]
$\bar{e}_1$ & $\bar{e}_2$	exponential terms associated with FTP at the pinholes for dimensionless equation	[-]
f	frequency	[Hz]
f_{bw}	bandwidth of the measurement system	[Hz]
f_1	focal-length of the objective lens	[m]
f_2	focal-length of the image lens	[m]

i	integer counting variable	[-]
i_{ns1} & i_{ns2}	incremental small signal quantities of the photocurrent shot noise for PD1 and PD2	[I]
k_1	spring constant between M_1 and the fixed world	[N/m]
k_{12}	spring constant between M_1 and M_2	[N/m]
l	length of the conductor	[m]
l_w	length of the wire	[m]
$n_{[material\,name]}$	refractive index of material in subscript	[-]
q	electron charge	[C]
r_b	offset of the beam from the optical axis	[m]
r_r	offset of the ray from the optical axis	[m]
r_1	eccentricity of the beam at pinhole 1	[m]
u_{drift}	uncertainty due to drift after correction	[nm]
u_{ph}	nominal offset of the pinholes	[m]
$\bar{u}_{ph}$	nominal offset of the pinholes divided by z_r	[-]
u_{sut}	defocus of the SUT	[m]
$\bar{u}_{sut}$	defocus of the SUT divided by z_r	[-]
v_{FES}	instantaneous value of the FES voltage	[V]
Δx	offset between interferometer beams	[m]
z	distance from beam's waist along optical axis	[m]
z_{ph}	instantaneous pinhole position relative to waist	[m]
$\bar{z}_{ph}$	instantaneous pinhole position relative to waist divided by z_r	[-]
z_r	Rayleigh range	[m]
α_{bs}	angle that the beam splitter is rotated in α-direction	[°]
α_{ep}	angle the front surface of the entrance prism	[°]
α_{max}	maximum tilt to be corrected	[°]
α_t	tilt of the SUT in α-direction	[°]
α_w	wedge-angle in α-direction	[°]
β_w	wedge-angle in β-direction	[°]
η_{opt}	transmissivity of the optical train	[-]

θ_B	half-angle of the cone of the laser beam	[rad]
θ_b	divergence of the laser beam	[rad]
θ_{ext}	half-angle of the cone of the extended laser beam	[rad]
θ_f	full-angle of the e^{-2}-boundary of the laser beam	[rad]
θ_L	half-angle of the maximum cone of light passing through the lens	[rad]
λ	laser wavelength	[m]
ξ_{psd}	omnidirectional tilt of the PSD to the ry-plane	[°]
ρ	density	[kg/m^3]
ρ_w	resistivity of the wire material	[Ω·m]
σ	standard deviation	N.A.
σ_{ave}	standard deviation of the average	N.A.
$\sigma_{fatigue}$	fatigue stress at the required number of cycles	[Pa]
σ_{hm}	standard deviation of height measurement	N.A.
$\sigma_{p0.2}$	offset yield point	[Pa]
τ_e	electrical time constant of focusing unit	[s]
τ_m	mechanical time constant of focusing unit	[s]

List of abbreviations

Abbreviation	Clarification
ADC	Analog-to-Digital Converter
AR	AntiReflection
BE	Beam Expander
BNC	Bayonet Neill-Concelman
CAD	Computer Aided Design
CCD	Charge-Coupled Device
C_F	Force Controller
C_{FES}	Focus Error Signal Controller
C_{IF}	IF Controller
CMM	Coordinate-Measuring Machine
DAC	Digital-to-Analog Converter
DAQ	Data AcQuisition

DC	Direct Current
DCS	Differential Confocal System
DOF	Degree Of Freedom
EDM	Electric Discharge Machining
EDR	Electrical Dynamic Range
emf	electromotive force
EMI	ElectroMagnetic Interference
FES	Focus Error Signal
FSR	Full Signal Range
FTP	Fractionally Transferred Power
GTD	Gemeenschappelijke Technische Dienst (technical support at TU/e)
HeNe	Helium-Neon
HFF	Heavily FreeForm
i.e.	id est (that is)
IEEE	Institute of Electrical and Electronics Engineers
IF	InterFerometer
I/O	digital Input/Output
ISO	International Organization for Standardization
LDGU	Laser Detector Grating Unit
LED	Light Emitting Diode
LR	near-Linear Range
LSB	Least Significant Bit
M^2	ratio of actual beam parameter product to ideal Gaussian beam parameter product
MFF	Medium FreeForm
NA	Numerical Aperture
NANOMEFOS	Nanometer Accuracy NOn-contact MEasurement of Freeform Optical Surfaces
NPBS	Non-Polarizing Beam Splitter
OPD	Optical Path Difference
OSR	Optical Signal Range
PBS	Polarizing Beam Splitter
PD	PhotoDiode
PH	PinHole
PID	Proportional–Integral–Derivative

PIN	P-type, Intrinsic, N-type
PM	Polarization Maintaining
PSD	Position Sensing Detector
PSS	Photodiode Sum Signal
PU	PickUp
PV	Peak-Valley
PWM	Pulse-Width Modulation
QWP	Quarter Wave Plate
RC	Resistance-Capacitance
RMS	Root Mean Square
RS	Rotationally Symmetric
SC	Stiffness Compensation
SFR	Symmetric Full Range
SH	Sample and Hold
SI	Système International i.e. international system of units
SMB	SubMiniature version B
SUT	Surface Under Test
TDE	Tilt Dependent Error
TEM_{00}	Transverse ElectroMagnetic mode ($_{00}$ indicates the fundamental mode)
TES	Tracking Error Signal
TG	Trajectory Generator
TiAlV	Titanium Aluminium Vanadium alloy
TNO	Netherlands Organisation for Applied Scientific Research
UK	United Kingdom of Great Britain and Northern Ireland
UV	Ultra Violet

CONTENTS

1   INTRODUCTION[1]

The application of freeform optics in high-end optical systems offers many advantages, however, their widespread use is held back by the lack of a suitable measurement method. The NANOMEFOS project aims at realizing a universal freeform measurement machine to fill that void. The principle of operation of this machine requires a novel sensor for surface distance measurement, the development and realization of which is the objective of the work presented in this thesis. Characteristic for the sensor is the combination of 5 mm measurement range with 35 nm measurement uncertainty for surface tilt up to 5°.

After a brief introduction to freeform optics and the advantages offered by their use, the measurement of freeform surfaces as proposed in the NANOMEFOS project will be discussed. Next, sensor requirements are presented, which are specifically based on the goals of the project. The chapter is concluded with an outline of this thesis.

1.1   MOTIVATION AND BACKGROUND

Most current day high-end optical systems, such as scientific instruments for earth and space observation and lithographic systems, mainly apply spherical optics, which inherently introduce aberrations. To partly compensate for such aberrations, these systems often incorporate multiple spherical optics in series. Utilization of aspherical and freeform optics allows these aberrations to be reduced or eliminated using fewer components, while also offering various other advantages. In some low

---

[1] *Section 1.1 and 1.2 are based on, and figures 1.1, 1.3, 1.4 and 1.6 have been taken (with minor changes) from the work of R. Henselmans as described in (Henselmans, 2009).*

and moderate accuracy applications, such as camera objectives, illumination optics and spectacles, aspherical and freeform optics are already frequently encountered.

Various fabrication methods for freeforms are available, such as single point diamond turning with a slow- or fast-tool-servo, local polishing techniques, ion and plasma beam machining and precision grinding. For the measurement of single-piece high-end freeform optics, however, no universal fast measurement method is available. Since measurement is a critical link in the production chain, this forms a significant obstacle for their application.

1.1.1   FREEFORM OPTICS AND THEIR ADVANTAGES

Next to the possibility of optical designs that have minimal inherent aberrations or the correction of aberrations applying fewer components, aspherical and freeform optical components also offer many other advantages. Due to the reduced number of required components, the volume and mass of optical systems can decrease. This forms an advantage in many applications, especially in space instruments, where volume and mass are often critical design criteria. Although the use of aspherical or freeform optics generally increases the number of degrees of freedom for which alignment is critical on component level, for the system as a whole, the reduced number of components can lead to reduced alignment efforts and reduction of complexity of the optomechanics.

For conventional optical systems, the optics often determine the boundary conditions for the optomechanics. Freeforms and off-axis components allow for much greater design freedom, such as the possibility to deviate from rotational symmetry. This offers the opportunity to realize integral optomechatronic designs with superior performance and facilitates the design of optical systems within a fixed design space.

When freeform optics become more common, and the production and measurement of freeform surfaces becomes routine, the use of freeform optics can even lead to lower overall system costs, due to the decreased complexity and number of components.

Classification of surface types

Occasionally, freeform optics are referred to as aspheres, here however, a distinction is made between rotationally *symmetric* non-spherical surfaces, referred to as *aspheres*, and rotationally *non-symmetric* surfaces, referred to as *freeforms*. Another criterion for classification is whether a component is used on- or off-axis. Based on these two criteria, non-spherical surfaces can be divided into four categories, as is illustrated in Figure 1.1. Whether a surface is aspheric, freeform,

off-axis aspheric or off-axis freeform, influences the measurement, as will be clarified when the NANOMEFOS machine is treated in Section 1.2.
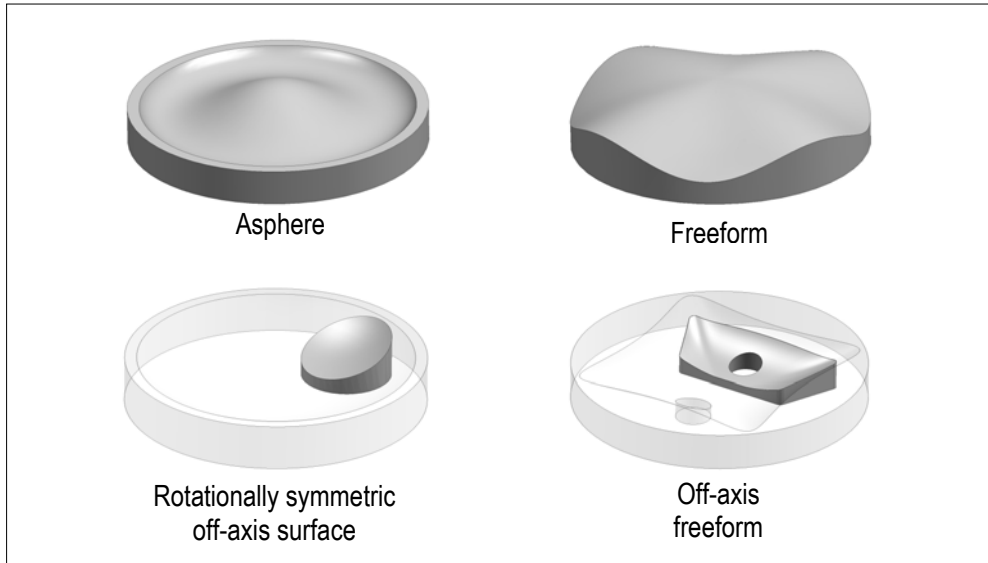
Figure 1.1: Examples of the four categories of non-spherical surfaces as used in this thesis; the deviation from spherical is magnified.

1.1.2 FABRICATION AND MEASUREMENT OF FREEFORMS

Figure 1.2 shows a diagram of a production chain for the fabrication of single-piece high-end freeform optics. Measurement plays a vital role as a source of feedback in this chain, as well as for validation of finished products.
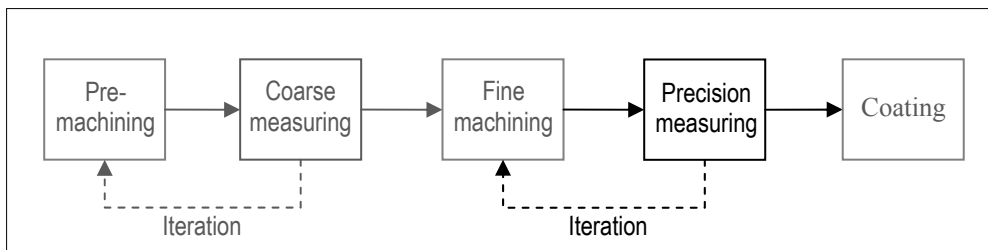
Figure 1.2: Optical production chain for freeform optics.

The pre-machining step usually involves conventional milling or grinding to bring a blank to within tens of micrometers of the desired form, but with a rough surface. The part is then inspected using a coarse measurement method: a Coordinate Measurement Machine (CMM) or spherometer, for example. When the pre-

machined part is close enough to the final form, a fine-machining operation is used to remove the sub-surface damage, after which it is inspected by means of a precision measuring method. Local fine-machining operations for freeform fabrication typically allow predictions of material removal rates with about 90% accuracy. Therefore, an iterative process of machining and measurement is needed to achieve the form requirement. When form and surface quality are within specification, a coating is often applied.

For high volume series, a specific measurement tool can be built; for single-piece production of freeforms, however, this is usually too expensive. It is desirable that a measurement tool for fabrication of single-piece, high-end freeform surfaces possesses the following five characteristics:

- high accuracy,

- universal applicability for high-end optical freeforms,

- low probability of surface damage (non-contact),

- large measurement volume, and

- short measurement time.

An overview of measurement methods for freeforms in general, is given in (Savio, et al., 2007). In (Henselmans, 2009), an evaluation is presented of various measurement methods, specifically with regard to the five aforementioned characteristics. The evaluated measurement methods are:

- interferometry, which for freeform measurement can be subdivided in:

  - conventional phase shifting techniques,

  - interferometry applying a null lens or computer generated hologram,

  - stitching interferometry,

- fringe projection,

- stylus profilometry,

- coordinate measurement using CMMs,

- swing arm profilometry,

- deflectometry,

- curvature measurement with a scanning miniature interferometer, and

- slope difference measurement with an autocollimator.

None of the evaluated measurement methods unifies the five aforementioned characteristics desirable for single-piece, high-end, freeform measurement.

1.2 NANOMEFOS PROJECT[2]

The aim of the NANOMEFOS project is to develop and realize a prototype freeform measurement machine that possesses the five aforementioned characteristics desirable for measurement of single-piece high-end freeform optics.

1.2.1 DESIGN GOALS FOR NANOMEFOS

The NANOMEFOS machine is designed for universal non-contact form measurement of flat, spherical, aspherical, freeform and off-axis surfaces that can be concave as well as convex. These surfaces can belong to transmission and reflection optics with product dimensions up to Ø500 mm x 100 mm. The 2σ measurement uncertainty should be 30 nm and measurements must be completed within 15 minutes.

Added to these criteria, the surface characteristics to be expected in future freeform designs are discussed. The surfaces to be measured are smoothly curved surfaces without steps in surface height; holes may however be present in the components, leading to surface discontinuity. For aspherical surfaces, the departure from spherical is not limited; on top of this, departure from a rotationally symmetric surface may be up to 5 mm peak to valley for freeforms. The surfaces may be concave or convex; the slope of the best-fit rotationally symmetric surface varies from -45° to 90°, and the local slope for freeforms may deviate up to 5° from that of the best-fit rotationally symmetric surface. The local curvature is expected to be limited to a minimal radius of some tens of millimeters.

The optics to be measured can consist of glass, ceramic or metal and, since they are in the stage of fine machining, are uncoated. When it is desirable to inspect coated products, this might be possible for simple reflective coatings. In practice, the reflectivity can therefore vary from 3.5% for fused silica to 99% for some silver coatings. Since the optics to be measured are high-end optics, their surfaces will have low roughness and few surface defects.

Because the machine is intended for measuring form of smoothly curved surfaces, a point spacing of about 0.5 mm to 2 mm is dense enough. It is desirable that in

[2] *NANOMEFOS is an acronym for Nanometer Accuracy NOn-contact MEasurement of Freeform Optical Surfaces. The project has been carried out within the scope of the M.Sc. and PhD projects of R. Henselmans, (Henselmans, 2005) and (Henselmans, 2009). It is a collaboration of Technische Universiteit Eindhoven, TNO Science & Industry, and the Netherlands Metrology institute Van Swinden Laboratory. Subsidy has been provided by the SenterNovem IOP Precision Technology program of the Dutch Ministry of Economic Affairs. The machine has been realized at the TU/e GTD workshop.*

addition to this, to collect information on roughness and waviness, lines or small areas can be measured with much higher point spacing. Another useful addition would be if rough surfaces of pre-machined blanks can be measured on the machine. This gives a measurement tool that covers the whole production chain for measurement of form, waviness, and roughness.

1.2.2 MACHINE CONCEPT AND DESIGN

A brief description of the NANOMEFOS machine is given, with emphasis on metrology and aspects regarding the sensor. A detailed description is given in (Henselmans, 2009).

When comparing single-point, line and area measurement methods, single-point measurement seems best suited for the desired combination of universal applicability, product dimensions, accuracy and measurement time. Surface form is commonly used to define optical surfaces and is needed to determine the required material removal during production. The point-wise measuring of a continuous surface can be done by measuring absolute position, relative position, slope, slope difference or curvature. The latter three methods require integration of data to obtain position information needed to reconstruct the surface form. These three methods, as well as relative position measurement, do not allow the universal measurement of discontinuous surfaces and introduce scaling difficulties. Consequently, single point, absolute position measurement is preferred. Furthermore, absolute position measurement allows for direct measurement of alignment and markers and for relatively straightforward traceability via calibration artifacts.

For the setup of the measurement machine, orthogonal, cylindrical and polar setups have been compared for various stage layouts. It was found that a cylindrical setup as shown in Figure 1.3 is most suitable for the measurement task.

In this machine setup, the product (1) to be measured is mounted on a vertical air-bearing spindle (2), which rotates at constant velocity. A non-contact sensor (3) is positioned over the product using an r-stage (4) and z-stage (5), so that circular tracks can be measured. A ψ-axis (6) is incorporated to orient the sensor perpendicularly to the best-fit rotationally symmetric surface. This has the advantages that it greatly reduces the required acceptance angle of the sensor and lowers the sensitivity to tangential errors.

During the measurement of a track, the $rz\psi$-motion system is stationary, thus limiting dynamic errors. Consequently, deviation from rotational symmetry must be accommodated by the sensor, which therefore must have 5 mm measurement range. To allow for averaging, a track can be measured multiple times, after which

the sensor is repositioned and reoriented to measure the next track. If the departure from rotational symmetry of an off-axis surface is within the range of the sensor, it may be measured on-axis as if it is a freeform surface. If the departure from rotational symmetry is large, however, the off-axis component has to be measured off-axis. This, as well as non-circular components or discontinuities in the surface, such as holes, cause the measurement signal to be interrupted, which affects the choice of measurement method for the sensor.
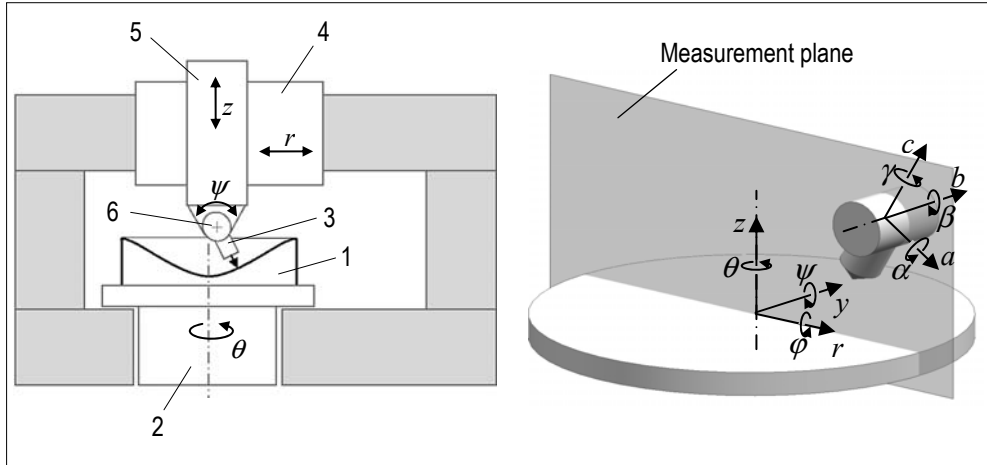
Figure 1.3: Concept of the measurement machine with its cylindrical setup (left) and definition of global coordinate system, relative to the machine base, and local coordinate system, relative to the sensor (right).

The plane of motion of the measurement spot is called the measurement plane. Two right-handed Cartesian coordinate systems have been defined, a global one, relative to the machine base, and a local one, moving with the sensor; both are shown in Figure 1.3, right. The global coordinate system has its origin at the intersection between the spindle's centerline and top surface; the z-axis coincides with the spindle centre line and the y-axis is orthogonal to the measurement plane. The local coordinate system has its origin at the intersection between the measurement plane and the ψ-axis centre line; the c-axis coincides with the sensor's measurement direction and the b-axis is orthogonal to the measurement plane.

If the position and orientation of product and sensor are referenced to a common observational frame of reference, the measurement problem has 13 Degrees Of Freedom (DOFs): the ryz-translations and $\varphi\psi\theta$-rotations of both freeform and sensor, and the distance between freeform surface and sensor in c-direction. Because of the lowered sensitivity to tangential errors, 7 of the 13 DOFs are less

critical. These less critical DOFs are constrained mechanically without applying real time metrology, while the most critical DOFs will be measured continually. In Figure 1.4, left, the 6 most critical DOFs are shown, these are z, r and ψ of the product, r and z of the ψ-axis centre line and the distance between sensor and surface in c-direction. Since these DOFs are all confined to the measurement plane, this reduces the three-dimensional measurement problem to a two-dimensional measurement problem.
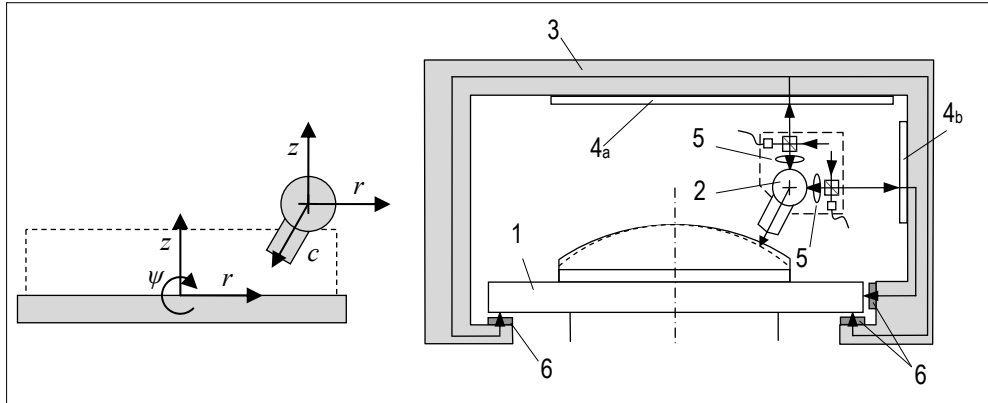
Figure 1.4: The 6 most critical degrees of freedom (left), and metrology frame concept (right).

In Figure 1.4, right, the metrology system is schematically depicted. The position of the spindle (1) and ψ-axis (2) are measured relative to a metrology frame (3), which is separated from the structural frame to increase accuracy. For the ψ-axis, the r- and z-position is determined using interferometers, which measure the distance between the ψ-axis and reference mirrors (4_a and 4_b) on the metrology frame. Thereto a mirror is applied on the ψ-axis and two cylindrical lenses (5) focus the beams on the ψ-axis centre line. The position of the spindle relative to the metrology frame is measured with capacitive sensors (6).

A photograph of the realized machine is shown in Figure 1.5. The base of the machine consists of a granite block assembly, which is suspended on four vibration isolators. To obtain an accurate plane of motion, instead of stacking the r- and z-stages, the z-stage is directly aligned to a vertical bearing face on the granite base by air-bearings. The air-bearings are force closed preloaded by opposing bearings. To prevent hysteresis in the structural frame that positions the sensor, separate preload frames are used so that the position frames do not significantly deform due to the preload. The ψ-axis consists of two radial air-bearings, located on both sides of the sensor and one axial air-bearing. The r- and z- stages and the ψ-axis and spindle are actuated with direct drive brushless motors driven by PWM amplifiers.

Figure 1.5: The NANOMEFOS machine.

1.2.3   MEASUREMENT UNCERTAINTY GOALS

For aspheric surfaces, the sensor's measurement direction is always perpendicular to the surface. Apart from zero first-order sensitivity to tangential errors this does also imply that only a small part of the measurement range and acceptance angle of the sensor are used. For freeform surfaces, on the contrary, the local slope is not always perpendicular to the sensor's measurement direction, leading to increased sensitivity to tangential errors. Furthermore, the measurement uncertainty for freeform surfaces increases due to the required range and acceptance angle of the sensor.

Because of the differences in measurement of aspheric surfaces and various gradations of freeform surfaces, a task specific measurement uncertainty is defined. In Figure 1.6, the budgeted measurement uncertainty as a function of local surface slope and product diameter is shown.

The variation of the overall measurement uncertainty for various surfaces is also reflected in the requirements for the sensor. The target for the sensor's contribution to the measurement uncertainty has been balanced with the rest of the measurement loop. This results in a required 2σ measurement uncertainty for the sensor of 10 nm for 0° surface tilt over a small measurement range, and increases to 35 nm for 5° local surface tilt over 5 mm measurement range (Henselmans, 2005).
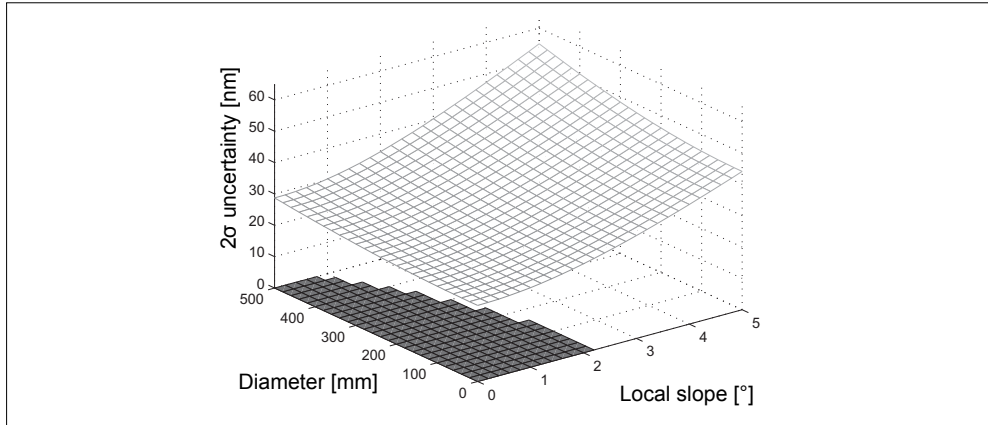
Figure 1.6: Budgeted 2σ measurement uncertainty as a function of local slope and product diameter.

1.3 REQUIREMENTS AND IMPACT ON SENSOR DESIGN

During the development of the NANOMEFOS machine, it was concluded that no existing distance sensor meets the requirements. Therefore, parallel to the development and realization of the machine, the development and realization of a distance measurement sensor has been undertaken. From the expected freeform characteristics and the design choices regarding the measurement method as a whole, the requirements of the sensor are formulated. The general implications of these requirements on design are also discussed.

Absolute distance measurement

As mentioned in the previous section, the NANOMEFOS machine relies on an absolute distance sensor to enable measurement of discontinuous surfaces.

Range, resolution and measurement uncertainty

Departure from rotational symmetry may be up to 5 mm peak to valley, which must be accommodated by the sensor, thus determining the range at 5 mm. The 2σ expanded measurement uncertainty requirement is 10 nm for surfaces perpendicular to the measurement direction, and 35 nm for surfaces with 5° local omnidirectional tilt. It is generally considered good practice to keep resolution an order of magnitude smaller than the intended measurement uncertainty. Hence, 1 nm resolution or better over the entire range will be aimed for.

Because of the required dynamic range of $5 \cdot 10^6$, it is unlikely that an analog measurement system can meet the requirements; therefore, an incremental measurement method such as interferometry or a linear scale is called for. Figure

1.7, left, depicts a sensor with a single-stage measurement system: an incremental primary measurement system (1) measures directly from the ψ-axis (2) to the freeform surface (3) over the entire range. An incremental non-contact measurement method that is suitable to measure directly to the surface with its significantly varying shape and tilt has not been found in literature.

It is expected that a dual-stage measurement system as schematically represented in Figure 1.7, right, will offer a solution. An analog primary measurement system (1), with the required resolution but with insufficient range, measures distance to the freeform surface (3) and is kept in range by a servo system (4). An incremental secondary measurement system (5) is used to measure the position of the primary system. Adding the distance measurements of the primary and secondary system gives an absolute measurement of distance between the freeform surface and the ψ-axis.

A consequence of this approach is that the allowable measurement uncertainty must be split up between the primary and secondary measurement system. Because of the limited foreknowledge of the subsystems, as an initial guess the primary and secondary system can be allotted the same value of measurement uncertainty. If the measurement errors of the systems are assumed uncorrelated, this works out to be 7 nm for each system.
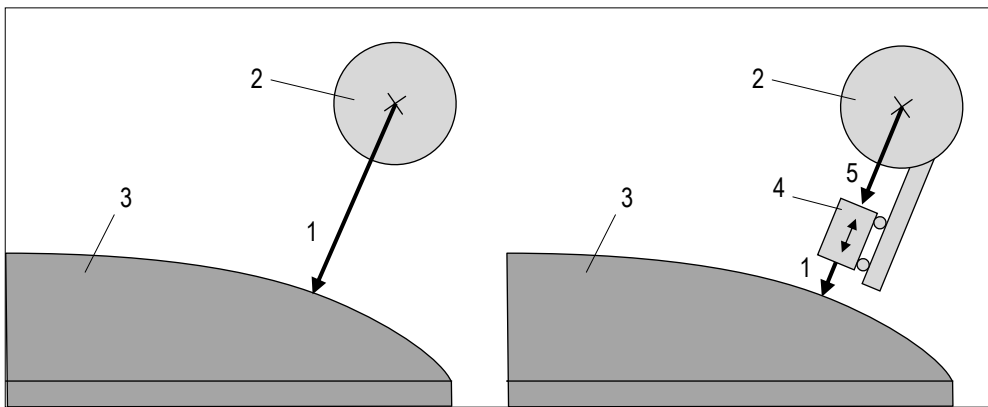
Figure 1.7: Schematic depiction of a single-stage measurement system (left) and a dual-stage measurement system (right).

Tracking bandwidth

The dual-stage approach gives rise to the derived requirement that the servo system must be able to follow the trajectory imposed by the freeform, with a tracking error smaller than half the range of the primary measurement system. Little is known about what kind of shapes to expect on future freeform optics, therefore it is not

possible to arrive at an exact requirement for the servo bandwidth. The trajectories necessary for surface measurement are expected to have most of the power content in the 1 Hz to 10 Hz range. Furthermore, the faster the servo system is, the better the overall system performance, since the measurement uncertainty of the primary system is expected to decrease for a smaller measurement range as well as for smaller tracking errors.

Guidance accuracy after calibration

If the dual-stage concept is used, it has to incorporate a linear guidance system. Because uncertainty in the lateral location of the measurement spot leads to uncertainty in surface form, the lateral guidance movement will be calibrated. The lateral accuracy of the guidance after calibration should be better than 50 nm to balance the measurement uncertainty caused by guidance run-out with that due to other sources of measurement uncertainty in the machine.

Non-contact operation and standoff

The measurement method must be non-contact because high scanning speeds and a low probability of surface damage are desired. For a single-stage measurement system, the standoff of the surface is not important. For the dual-stage measurement system, however, the minimum standoff of the surface is important, since the front part of the sensor can collide with the product. Machine control must have enough time to retract the sensor in case of an emergency or loss of the tracking signal, while also being robust to surface defects. This in combination with the tilt of the surfaces and the finite dimensions of the sensor, requires a standoff of 0.5 mm or more.

Surface characteristics

The surfaces to be measured can be metal, glass or ceramic; this rules out the use of sensors that require an electrically conducting target, such as eddy current and capacitive sensors.

The reflectivity of the surfaces varies between 3.5% and 99%. When an optical method is applied, it must therefore be able to function properly with differences in returning light intensity of a factor of almost 30. Characteristics that might vary with intensity include a sensor's response curve, measurement noise and measurement uncertainty. Except for local surface defects, the reflectivity for a single component is not expected to vary significantly over the surface.

The primary goal of the project is to measure high-end optics in the fine machining stage; most surfaces will thus have low surface roughness, however, local surface defects such as scratches, pits or dust, are likely to be encountered regularly. Moreover, when porous materials such as sintered Silicon Carbide are polished,

only a portion of the surface will have low surface roughness with interruptions due to pores and possibly particle boundaries. The measurement method chosen must either be robust for such surface defects, or offer a way to identify compromised data points.

Enabling the measurement of opaque surfaces of pre-machined blanks is considered a bonus, but not a core objective. Consequently, this will not be a consideration during selection of a measurement principle and design of the sensor. Nevertheless, since the required measurement uncertainty at these intermediate product steps is much lower than what is required during the fine machining stages, measuring rough surfaces might prove possible anyway.

Acceptance angle

The local surface tilt encountered during measurements depends on amplitude, number of waves per revolution, radial location and shape of the waveform superimposed onto the rotationally symmetric component of the freeform. For the most heavily freeform surfaces to be measured, a maximum surface tilt of ±5° is expected in both the α- and β-direction. Therefore, the sensor will need to have an omnidirectional acceptance angle of 5°.

Whatever primary measurement principle is chosen, to some extent, distance measurement will always depend on the tilt of the surface, if not inherently, then through limited component and alignment tolerances. Based on existing distance measurement principles, it is expected that the 35 nm measurement uncertainty for surface tilts up to 5° is going to be difficult to achieve without correction. For that reason, a system that incorporates calibration for surface tilt is pursued.

In principle, the local slope is approximately known from CAD data of the surface, and can thus be applied for correction. If, however, the CAD data is (locally) not accurate, for example due to misalignment or microstructure of the surface, this will lead to measurement errors. Hence, a system in the sensor that registers the surface tilt is favored.

As mentioned before, slope data can also be applied to reconstruct a surface, which, for phenomena with high spatial frequency, can have considerable advantages over distance measurement. Because of discontinuities in the surfaces, however, tilt measurement needs reference points from absolute distance measurements. In principle, distance measurement and slope measurement could therefore be used together, to combine the best of both worlds. Nevertheless, this solution is not pursued here because of the difficulty to incorporate a tilt measurement and distance measurement system into the sensor that are both accurate enough to achieve the requirements.

Instead, effort will be put into designing a distance measurement sensor with low measurement uncertainty, while the angle measurement system is only accurate enough to enable significant reduction of tilt dependency through calibration.

Sample rate, spot size and aliasing

To attain the envisaged short measurement times, scanning speeds of almost 1.6 m/s are needed when measuring a 0.5 m diameter optic. It is thought that a sample spacing of about 0.5 mm to 2 mm will suffice for form measurement, corresponding to a sample rate of about 3 kHz.

It is unclear which spectral content the surfaces to be measured will typically have, thus the contribution of aliasing to measurement uncertainty is hard to estimate, but it is thought that it can lead to significant measurement errors. Aliasing can be considerably suppressed if the measurement spot size is twice as large as the sample spacing. This can be achieved by choosing a large spot size or by decreasing the sample spacing. What spot size can be attained depends on the primary measurement principle; as discussed later, the spot size of the primary measurement system selected here is typically in the micrometer range.

Decreasing both radial and tangential sample spacing to micrometer level is rejected because it leads to impractical measurement times and amount of data. When the surface roughness is isotropic, most of the aliasing can be suppressed by decreasing sample spacing in tangential direction only, which can be done by increasing the sample rate. As will be discussed later, a small signal bandwidth of 150 kHz can be achieved by the sensor. In practice, the bandwidth is thus limited by the control and data logging system of the machine. If needed, the sample rate of data logging can be increased independently of the control system. Analog low-pass filtering before the ADC can also be used to suppress aliasing, thereby decreasing electrical measurement noise as well.

When measuring diamond turned optics, the assumption of isotropic surface roughness is not satisfied. In this case, attention must be given to the alignment of the measurement tracks relative to the grooves of the diamond tool.

Volume envelope

A large distance between ψ-axis centre line and measurement spot allows measurement into deep concave components; however, it tightens the requirements for the ψ-axis encoder and requires the rz-motion system and metrology system to have a larger stroke. A distance of 100 mm has been found to be a good compromise.

Part of the sensor can be placed in a pocket inside the ψ-axis, provided that it can be removed in one piece. Consequently, the sensor's width is limited by the ψ-axis

bearings. As discussed in Subsection 1.2.2, the *r*- and *z*-positions of the *ψ*-axis is measured with two interferometers that are reflected by a cylindrical mirror on the *ψ*-axis. Because the *ψ*-axis must be able to rotate between -45° for measuring concave surfaces, and +120° for interferometer nulling (discussed later), there must be a clear trace of 255° on the *ψ*-axis. The diameter of the *ψ*-axis is limited to 70 mm because this is the maximum allowable diameter of the setup used to calibrate the *ψ*-axis mirror roundness at NMi[3]. These factors together limit the volume envelope of the sensor, which is depicted in Figure 1.8.
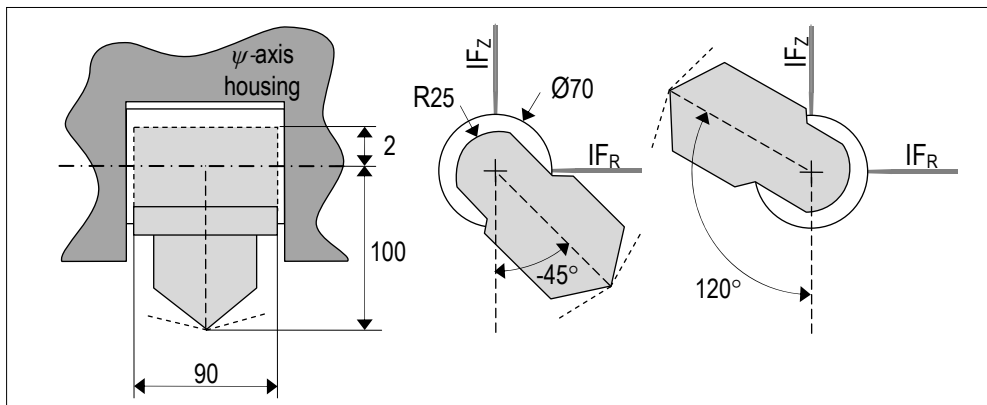
Figure 1.8: Approximate volume envelope for the sensor viewed along the r-axis with the sensor in vertical orientation (left) and viewed along the y-axis with the sensor oriented for measuring a steep concave surface (middle) and for nulling on the reference mirror (right).

Environment

The working environment will be a conditioned metrology laboratory that is clean and where the temperature is controlled to 20°C ±0.2°C. Most of the amplifiers are turned off during measurement of tracks because the axes are mechanically braked. When measuring radial scans of the product to enable drift correction or when spiral measurement is desirable, however, the PWM amplifiers will be turned on. This increases the risk of electromagnetic interference, especially with regard to the small current signals characteristic for intensity and position measurement of low power light.

Requirement overview

An overview of some of the requirements addressed here is given in Table 1.1.

---

[3] *VSL is the Dutch metrology institute; website: www.vsl.nl.*

Sensor property	Requirement
Measurement range	5 mm
Acceptance angle	5°
Resolution	1 nm
Measurement uncertainty for 0° surface tilt (2σ)	10 nm
Measurement uncertainty for up to 5° surface tilt (2σ)	35 nm
Sample rate	>3 kHz
Lateral guidance accuracy (in case of dual-stage system)	50 nm

Table 1.1: Overview of requirements for the sensor.

1.4 OBJECTIVE, METHODS AND OUTLINE

No commercial or experimental sensor combines the required characteristics. Therefore, the objective of this research is the development and realization of an absolute distance measurement sensor suitable for surface distance measurement in NANOMEFOS, for which the requirements have been discussed in the previous section.

The approach followed largely corresponds to the order in which this thesis is set up; hence, the method description and thesis outline are described together.

Chapter 2:   Optical surface measurement

It is believed that a dual-stage measurement system with an optical primary measurement method is best suited for the task. Various measurement methods known from literature are compared and evaluated with regard to their inherent properties and the possibilities for adaptation to the requirements for NANOMEFOS. To allow for correction of tilt dependent error through calibration, a third measurement system is added that measures through which part of the aperture the light returns.

Chapter 3:   Analytical differential confocal model & demonstrator

To enable predictions regarding performance and optimization, analytical models of the selected primary measurement principle are derived. In addition, to provide proof of principle and gain experience with the selected method, an experimental setup is built. To assess the models' usefulness as a design and optimization tool, the results from the models are compared to those experimentally obtained using the demonstrator.

Chapter 4: Differential confocal property analysis & optimization

To enable optimization of the primary measurement system, an understanding of the relations between the design parameters and the system's characteristics is desired. Therefore, various properties are identified that characterize the performance of the primary measurement system. After that, their dependence on the design parameters is studied with the aid of simulations using the previously derived analytical models. Based on this property study, the primary measurement principle is optimized for application in NANOMEFOS.

Chapter 5: Optical prototype design

First, the conceptual optical system design is made with regard to integration of the primary and secondary measurement principle, limiting measurement uncertainty and adaptation of the optical path to fit the volume envelope. Next, the general requirements for the optical components are formulated. Many of the components have to be adapted from commercially available parts or custom made to fit the volume envelope; hence, their specifications are drawn up.

Chapter 6: Optomechanical and mechatronic prototype design

The optomechanics are designed after their requirements and boundary conditions have been mainly determined by the optical system. Furthermore, a custom guidance mechanism and actuator for tracking the surface are designed; to achieve limiting measurement uncertainty, the emphasis is on accurate motion, high bandwidth and low dissipation.

Chapter 7: Realization, focusing unit performance and control

The realization of the system and testing of subsystems are discussed followed by the partly custom electronics and the control of the sensor.

Chapter 8: Experimental results and calibration

Performance tests of the sensor are presented with an emphasis on measurement errors and repeatability. Various calibrations and their expected effect on measurement uncertainty are discussed. Among these calibrations is a novel calibration method to measure the tilt dependency of distance sensors. To test the performance of the NANOMEFOS machine with the sensor installed, a tilted flat, which serves as a reference freeform with known surface form, is measured.

Chapter 9: Conclusions and recommendations

Based on the results and experiences, conclusions are drawn and recommendations for improvement and application are given.

2 OPTICAL SURFACE MEASUREMENT

Considerations regarding the conceptual design of the sensor are treated including a method to enable correction of measurement errors due to surface tilt. A dual-stage design with an interferometer as secondary measurement system is chosen. Various measurement methods known from literature are compared and evaluated with regard to their inherent properties and the possibilities for adaptation to the requirements for the primary measurement system. Differential confocal measurement is selected; some design considerations and an initial approach to the design of such a system are discussed.

2.1 SENSOR CONCEPT

Issues that mainly govern the conceptual design of the sensor are the required dynamic range and measurement uncertainty, the reduction of tilt dependent error, the tracking bandwidth and the system complexity.

2.1.1 SINGLE-STAGE DESIGN

A single-stage measurement system is potentially simpler than a dual-stage design and its uncertainty budget can be larger since it is not in series with another measurement system. Unfortunately, this solution is not feasible because of the lack of a suitable measurement method.

Interferometry is capable of achieving the required dynamic range of $5 \cdot 10^6$ in a non-contact manner. Traditional interferometers are relative measurement methods and can thus not be directly used to measure discontinuous surfaces. Various interferometry systems overcome the phase ambiguity of conventional interferometry by applying stabilized multiple wavelength sources, superheterodyne detection, frequency sweeping or combinations of these methods

(Hartmann, et al., 2008). Some of these techniques can achieve the required measurement range and resolution. In this application, however, the shape and tilt of the Surface Under Test (SUT) is undetermined, which causes loss of modulation-depth. This can be (partly) prevented by keeping the interferometer beam in focus at the SUT, which requires a focusing system to track the SUT. In this case, tracking errors lead to measurement errors that are difficult to compensate for. If a focusing system is needed anyway, the system becomes as complex as a dual-stage design with the added disadvantage that the tolerable tracking error is small. Therefore, it seems better to apply a primary measurement system that allows larger tracking errors in series with a traditional incremental measurement system as introduced in 1.3.

2.1.2 DUAL-STAGE DESIGN

A short-range primary measurement system is kept in range by a servo system, while an incremental secondary measurement system measures the tracking translation of the primary system. The distance measurements of the primary and secondary system are added to obtain an absolute distance measurement between the freeform surface and ψ-axis.

A significant advantage of a dual-stage design is that it allows for in situ calibration of the primary measurement system. This is advantageous in case the response of the primary measurement system is surface dependent or if it changes over longer periods of time.

Secondary measurement system

Two types of measurement systems are commercially available that can achieve the high dynamic range and low uncertainty required for the secondary measurement system: optical linear scales and interferometers. An interferometer can more easily be incorporated within the volume envelope while also satisfying the Abbe principle (Abbe, 1890). Interferometry adds no friction or damping to the system and requires little extra moving mass. Furthermore, the measurement head of an optical linear scale introduces a significant heat source close to the critical components of the system, while interferometry allows the heat sources to be placed away from the sensor. Therefore, interferometry is preferred as the secondary measurement method. A disadvantage of interferometry is its sensitivity to refractive index variation of air.

Primary measurement system

Because high measurement bandwidth is desirable, single point sensors are best suited. Most single-point distance measurement methods are not suitable as primary measurement method. As mentioned in Chapter 1, capacitive and eddy-

current sensors cannot be used because they need the surface to be electrically conducting. A range of scanning probe microscopy techniques exist that are non-contact and do not need the surface to be electrically conducting, examples are tapping and non-contact mode atomic force microscopy. The scanning speed that can be achieved with these methods however, is orders of magnitude below what is practical here. Ultrasonic sensors are not applicable because the highest resolution found for ultrasonic sensors is orders of magnitude lower than what is required. The category that remains is single-point optical distance measurement, on which the search for a primary measurement method is focused. A comparison between various measurement methods will be presented in Section 2.2. The suitable candidates for the primary measurement method require light to be focused on the SUT.

Surface tracking

In general, the measurement uncertainty decreases as the measurement range of the primary system decreases, therefore it is desirable to keep the tracking error small. To track the surface, various methods can be applied, some of which are depicted in Figure 2.1.
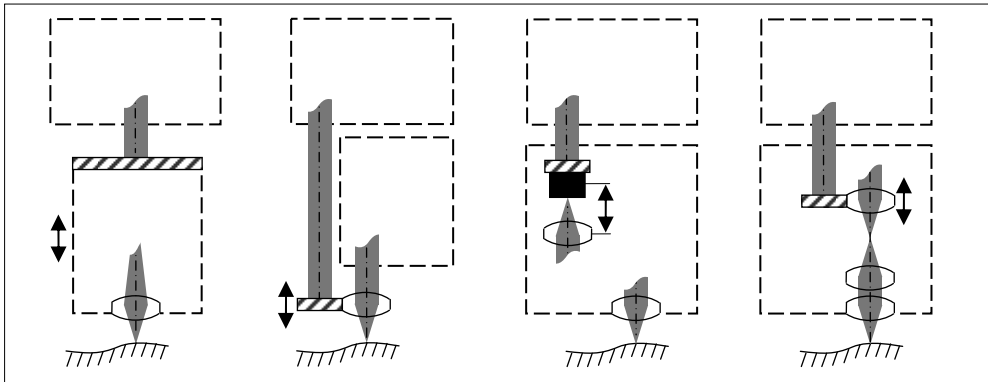
Figure 2.1: Four methods to track the surface: the entire primary measurement system is translated (far left), the objective lens is translated (center left) an optical component or detector of the primary measurement system is translated (center right) or a focusing subsystem such as a telescope lens pair is added (far right).

The most obvious is to translate the entire primary measurement system and measure its displacement with the secondary system, Figure 2.1, far left. This involves accelerating a relatively large mass and therefore is undesirable. Alternatively, for most measurement methods the object lens can be translated instead, as depicted in Figure 2.1, center left. Because of the low mass of the object lens, a high tracking bandwidth can be achieved, thus enabling small tracking

errors to be realized. This method requires a collimated light source if a near equal response of the measurement system is desired over the tracking range.

In some measurement systems, it is possible to translate one or several of the components instead, such as a detector or image lens, as in Figure 2.1, center right. Due to the low moving mass involved, high bandwidths can be obtained using this method also. In some cases, it is beneficial because of space considerations; however, it is only practical for small tracking distances since the divergence of the returning beam otherwise changes too much. Another solution is to incorporate dedicated extra components for focusing, in optical storage this is sometimes done with a telescope lens pair and axial translation of one of the lenses (McDaniel, et al., 1997), as shown in Figure 2.1, far right. This method again allows for high bandwidths but leads to a change of beam diameter for large translations. Tracking the surface by translating the objective lens seems the best solution.

Vibrating around best-focus

As an alternative to continuously tracking the surface close to best-focus, the primary system can be vibrated around best-focus and the signal of the secondary systems is sampled at the zero-crossing of the primary system. The advantage of this approach is that the primary measurement system can be used as a null-sensor, which can achieve superior resolution and measurement uncertainty compared to measurement systems with substantial range.

In case of a vibrating system it might be beneficial to split the translation of the long range tracking from the vibration movement; the objective can be translated for the long range tracking while vibrating around best-focus is accomplished by vibrating a detector or lens in the primary system.

It is thought that both vibrating and non-vibrating systems can be made to work. Nevertheless, a non-vibrating primary measurement system is preferred. The reason is that sampling the secondary system only at zero-crossings of a vibrating primary measurement system leads to lowered bandwidth and varying sample spacing, which reduces effectiveness of averaging multiple measurements of the same track. Furthermore, it increases system complexity and complicates control.

Illumination and light collection

For most of the primary measurement methods evaluated in Section 2.2, two types of illumination can be distinguished: one where the incident beam arriving at the objective lens is divergent and one where it is collimated. To enable tracking the surface by translating the objective lens, a collimated source is required. Figure 2.1 depicts the illumination of the SUT and collection of the reflected light for a SUT that is too far from the object lens, in best-focus and too close to the object lens.
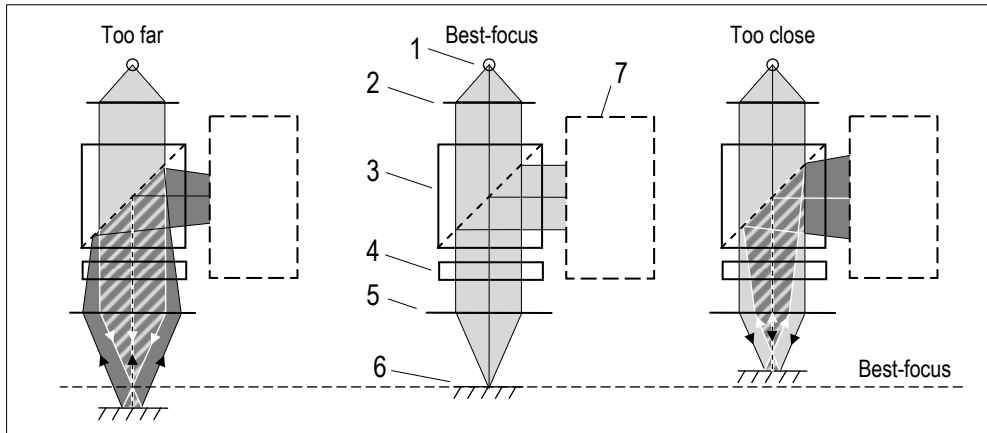
Figure 2.2: Illumination of the Surface Under Test (SUT) and the collection of the reflected light when the SUT is too far away from the object lens (left), in best-focus (middle) and too close to the object lens (right).

The light source (1) is often a laser with linearly polarized output, although some methods apply a white light source such as a lamp or LED. Unless a light source with collimated output is used, a gas laser for example, the system incorporates collimation optics and or beam shaping optics (2). The incident beam passes through a Polarizing Beam Splitter (PBS) (3). The polarization of the light source is aligned to the PBS such that nearly all light leaves the PBS in transmission. A Quarter Wave Plate (QWP) (4) is aligned with its fast axis oriented at 45° relative to the beam's polarization direction, so that the beam becomes circular polarized.

An object lens (5) focuses the beam onto the, in this case, specular SUT (6) which reflects the beam. The reflection causes the beam's circular polarization to reverse from left-handed to right-handed or vice versa. The returning light is collected by the objective lens and if the SUT is located at best-focus, the beam leaves the object lens collimated. It propagates through the QWP again and due to the beam's reversed circular polarization it leaves the QWP linearly polarized with its orientation orthogonal to the incident beam. The beam entering the PBS is reflected at the beam splitter plane because of the 90° rotated orientation of its polarization, separating it from the illumination beam.

When the SUT is too far away from the lens (left), the returning light leaves the object lens as a convergent beam; the opposite holds when the SUT is too close to the lens (right). The returning light that is separated from the incident light by the beam splitter, is processed by a system (7) which measures where the SUT is relative to best-focus.

2.1.3   APERTURE CORRECTION

The sensor has to achieve a measurement uncertainty of 35 nm for surfaces that are tilted up to 5°. If the SUT is tilted, the reflected cone of light tilts twice as much. This leads to an offset of the returning beam, as can be seen in Figure 2.3, left. The shift of the returning beam, r_b, can be expressed as:

$$\Delta r = f_1 \tan(2\alpha_t) \tag{2.1}$$

in which:

f_1 is the focal length of the object lens, and

α_t is the tilt of the SUT.
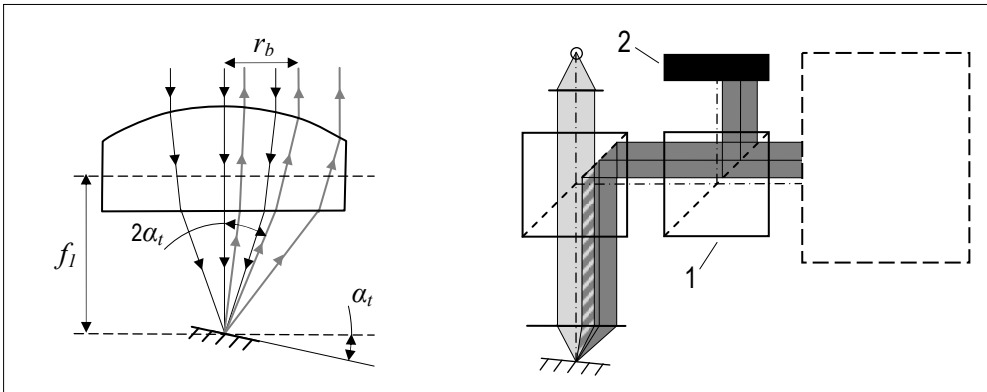
Figure 2.3: Beam offset due to tilt of the SUT (left) and addition of a beam splitter and Position Sensitive Detector (PSD) to coarsely measure surface tilt for calibration purposes (right).

The shift of the returning light leads to Tilt Dependent Error (TDE), which can be split up in three components, namely TDE due to:

- vignetting of the beam,

- inherent tilt sensitivity of the measurement principle, and

- aberrations in the system.

Measurement errors due to vignetting of the beam can be prevented relatively easily: by matching the beam diameter to the aperture such that the aperture can fully accommodate a returning beam with maximum offset.

TDE due to inherent tilt sensitivity of the measurement principle is an important selection criterion in Section 2.2, where various primary measurement methods are

compared. Furthermore, the inherent TDE is one of the dominant criteria in Chapter 4, in which the chosen primary measurement system is optimized.

The TDE due to aberrations in the system can be caused by aberrations in the design, manufacturing tolerances or alignment errors of the optical components. Even for measurement principles that have no inherent sensitivity to tilt, the aberrations alone are expected to introduce errors that are too large to meet the uncertainty requirement without correction. To enable correction through calibration, a third measurement system is added to measure which part of the aperture is used; it is depicted in Figure 2.3, right.

This aperture correction system consists of an extra beam splitter (1) and a two-dimensional Position Sensitive Detector (PSD) (2). The beam splitter reflects a small percentage of the beam onto the PSD, which produces four photocurrents from which the lateral position of the centroid of the beam as well as the incident power can be obtained.

A PSD is chosen for its combination of low rise-time and its integrating measurement principle. Quad-cells have even shorter rise-time but are more sensitive to intensity variations at the center of the beam, for example due to interference with a ghost reflection. Use of a CCD is rejected because of its long readout time.

The ratio of the spot diameter on the PSD to the maximum tilt induced offset depends on the Numerical Aperture (NA) of the beam. As will be treated in the following sections, a high beam NA is needed to achieve the required measurement uncertainty. Consequently, the maximum tilt dependent offset is roughly an order of magnitude smaller than the spot diameter on the PSD. Hence, the intensity variations over the spot give relatively large errors when compared to dedicated tilt measurement methods such as autocollimators and deflectometers, in which the spot on the PSD is much smaller than the tilt dependent offsets. Therefore, the aperture correction signal is not accurate enough to be directly used for reconstruction of the measured surface; however, it is expected to be accurate enough to reduce TDE by a factor of about 10 to 30 through calibration. A new method for calibration of TDE has been developed and will be discussed in Subsection 8.3.3.

The here presented aperture correction to enable correction of TDE through calibration is described in a patent application covering part of the sensor design (van Amstel, Cacace and Henselmans, 2007).

2.2 CHOICE OF PRIMARY MEASUREMENT METHOD

An important criterion for the choice of a primary measurement method is tilt dependence; this however, is not often quantified in literature. Furthermore, performance indicators such as range, measurement uncertainty and resolution, are hard to compare since they depend on design, realization, calibration and testing conditions. Therefore, it has been chosen instead to assess measurement principles mainly by comparison of characteristics inherent to their principle of operation. Their working principles are discussed first, followed by a comparison on which the choice of the primary measurement principle is based.

Because a system is pursued in which the light is collimated at the object lens, the following treatment of the focus methods consider collimated illumination, even if the method is usually realized with a divergent beam at the object lens. The only method discussed here for which this is not relevant, is triangulation.

2.2.1 TRIANGULATION

Triangulation sensors are widely used as distance sensors in automation and inspection. Most triangulation sensors apply an illumination beam that is oriented perpendicular to the SUT. This works well for diffuse surfaces but not for specular surfaces. An alternative layout that is suitable for measurement of diffuse as well as specular surfaces is discussed in (Tanwar, et al., 1984) and (Mitsui, 1986). In these layouts, the optical axes of the illumination and measurement system are oriented opposite to each other at an angle of 45° to the SUT, as is depicted in Figure 2.4.
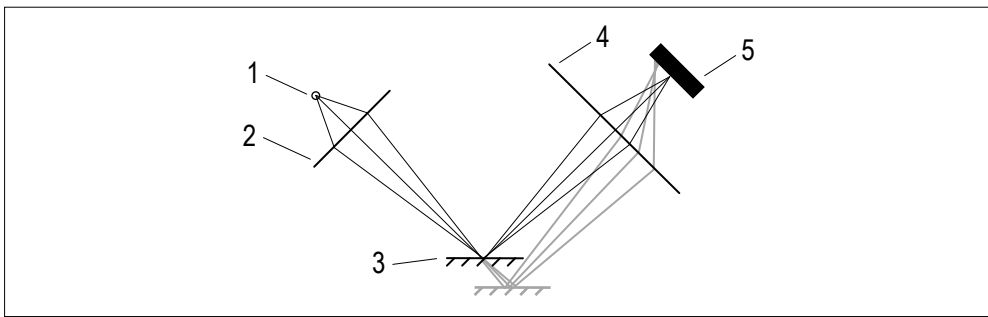
Figure 2.4: Layout of a triangulation method for specular surfaces.

The system comprises a light source (1) that is focused onto the SUT (3) by an illumination lens (2). Another lens (4) focuses the reflected light onto a PSD (5), which measures the lateral position of the centroid of the beam. When the SUT moves up or down, the spot moves over the PSD; the distance of the SUT is obtained by triangulation. Surface tilt does not principally effect the distance

measurement since the PSD is located at the image plane of the spot on the SUT. The lateral displacement of the measurement spot on the SUT is proportional to the measured distance and thus relatively easy to compensate for in software.

The sensitivity of the measurement principle is not fundamentally affected by the NA of the illumination beam, hence the spot can be chosen quite large. The size of the spot is however limited because of the adverse effect of large spot size on measurement uncertainty due to changing intensity inhomogeneities over the spot.

A disadvantage of this method is that it is unpractical to track the surface using only a part of the primary system; therefore, it has to be translated as a whole to track the surface. This has an adverse effect on tracking bandwidth and leads to relatively large dynamic errors in the machine. Another disadvantage is the large lateral dimensions of the sensor head, which is conflicting with measurement of concave optics. In (Gao, et al., 1997) a layout is presented that allows a slimmer sensor head by reflecting the incident and returning beams with mirrors.

Although a high spatial dynamic range can be achieved with PSDs, impractical optical magnification or small detectors are needed to achieve the required resolution. Furthermore, since out-of-focus light from within a large field can reach the detector, the measurement uncertainty of this type of sensors is often limited by stray light and ghost reflections.

2.2.2   WHITE LIGHT INTERFEROMETRY

White light interferometry can be applied on specular as well as rough surfaces; for reflective surfaces, a reproducibility of 0.5 nm has been reported (Deck, et al., 1994). The method applies a broadband measurement beam that is reflected by the SUT after which it is combined with a reference beam and interference is detected. To allow measurement of tilted surfaces the light is focused to a small spot.

Three types of focused white light interferometers are encountered: the Michelson, Linnik and Mirau type (Creath, 1989). The Mirau type is best suited because it can be used with high NA objectives while the reference mirror is incorporated in the objective lens, avoiding the need to translate a separate reference branch to maintain approximately equal optical path lengths. The layout of a Mirau-type white light interferometer is shown in Figure 2.5, left.

The light from a broadband source is focused by an objective (1) in which a splitter plate (2) is incorporated. This plate splits the beam into a reference and a measurement beam. The measurement beam propagates to the SUT located approximately at best-focus and the reference beam is focused on a small reference mirror (3). A lens (4) focuses the returning light on a detector (5) that measures the

intensity of the recombined beam. The short coherence length of the source causes interference to occur only if the optical path lengths of the reference and measurement beam are approximately equal; this enables absolute distance measurement. The interference signal (6) is depicted in Figure 2.5, right. Around best-focus the amplitude of the signal increases and has a narrow peak at best-focus. The top of this central peak can be used for distance measurement; due to the limited width of the central peak, a very small tracking error can be allowed.
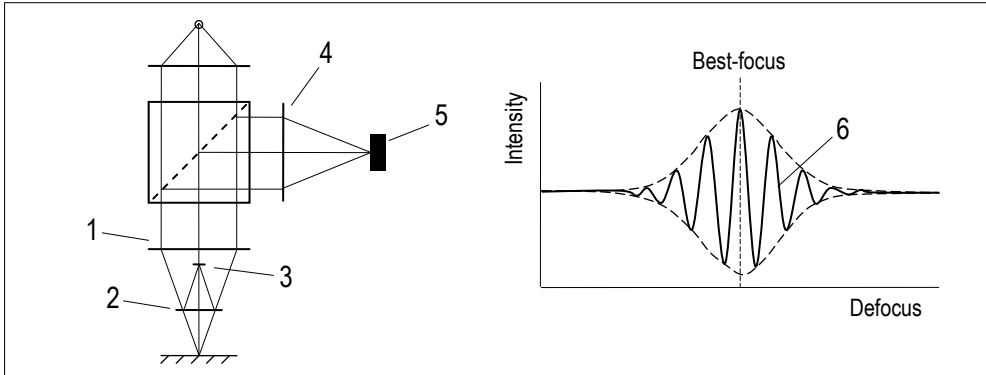
Figure 2.5: Schematic layout of a white light interferometer with a Mirau-type objective (left) and its response (right).

To circumvent the problem, the objective can be vibrating through best-focus as discussed in Subsection 2.1.2. This however requires extremely high sample rates and fast data processing, since reconstruction of an under-sampled signal such as applied in (Deck, et al., 1994) is not possible here, because of the varying scan speed in this application.

2.2.3   AUTO FOCUS METHODS

Auto focus detectors make up a category of sensors that measure distance by evaluating if the SUT is located at the focus of a light beam, or how far away from it. An overview of various auto focus methods is given in (Stout, et al., 1994).

The auto focus methods that have been evaluated for suitability as primary measurement system are the astigmatic, double Foucault, laser detector grating and double critical angle methods. The principles of operation of these four methods are applied in read-heads of optical storage systems (McDaniel, et al., 1997) and (Stan, 1998). In optical storage applications, apart from the Focus Error Signal (FES), which is used to stay in focus, a Tracking Error Signal (TES) and a data signal are needed to radially track the spiral and to read out the data respectively. Some of these methods have been adapted for distance measurement applications,

for which the TES and data signal are not needed, and in which, contrary to use in optical storage application, absolute distance uncertainty is important.

The working principles and specific properties of these methods will be discussed per method, the general properties are discussed at the end of this subsection.

Various confocal methods have been evaluated also. Strictly considered, these are auto focus methods as well; nevertheless, because of their specific characteristics they are treated separately in Subsection 2.2.4.

Astigmatic method

The astigmatic method, (Bouwhuis and Braat, 1978) has been widely used in optical storage applications; its use for surface measurement is reported in (Lou, et al., 1984) and (Mitsui, 1986). The method is depicted in Figure 2.6.
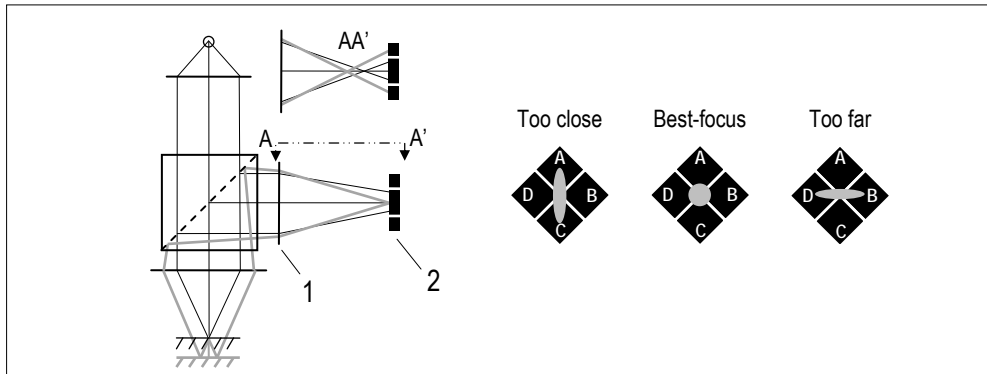
Figure 2.6: Layout of an astigmatic focus detector (left) and the spot on the quad cell for different SUT locations (right).

To detect the position of the SUT relative to focus, astigmatism is introduced into the returning beam by use of an astigmatic imaging system (1), which can consist of an astigmatic lens, a combination of two cylindrical lenses or a conventional lens combined with a cylindrical lens or a tilted plate. A photodiode quad cell (2) is located such that when the SUT is in best-focus, a round spot is formed at its center and the incident power on the four quadrants is equal. When the SUT has an offset from best-focus, an elliptical spot is formed, as shown in the middle of the detector images in Figure 2.6, right. Because the quad cell is oriented such that the edges of the quadrants are 45° rotated to the focal lines, the incident power on two opposing quadrants increases, while the incident power on the other two quadrants decreases. The orientation of the ellipse indicates whether the SUT is located too close to or too far from the lens.

The normalized Focus Error Signal (FES) can be calculated as:

$$FES = \frac{(A+C)-(B+D)}{A+B+C+D} \qquad (2.2)$$

Normalization renders the signal insensitive to intensity fluctuations, for example due to variation in laser power or the reflectivity of the SUT. The sensitivity of the FES to lateral spot displacement on the quad cell is zero in first order, because the change of incident power on one quadrant is compensated by the change of incident power on the opposing quadrant. Furthermore, the lateral spot displacement due to surface tilt is small since the quad cell is located close to focus. Consequently, the inherent tilt dependence of the method is relatively small. Because of the low sensitivity of spot position to surface tilt, the difference signal of opposing quadrants is not suitable to measure the tilt of the SUT for aperture correction.

Double Foucault method

A version of the double Foucault method that uses a double-wedge (Bouwhuis and Braat, 1978) is schematically depicted in Figure 2.7. Its use for surface measurement is reported in (Benschop, et al., 1991).
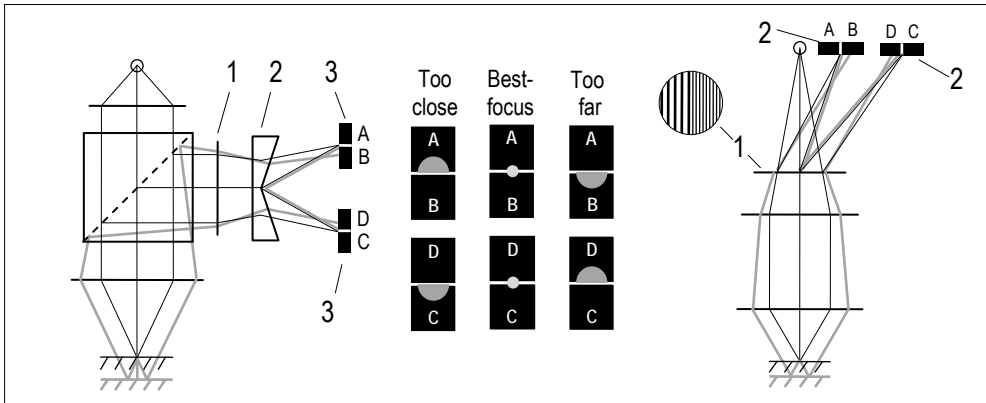
Figure 2.7: Layout of the double Foucault detector using a double wedge (left) and the layout of a laser diode grating unit (right); the spots on the detectors (middle) are similar.

The returning beam is focused by a lens (1) with a double wedge (2) placed behind it. The double wedge refracts the beam into two branches, which are focused onto the middle of two dual cells (AB and DC). The spots at the detectors are depicted in Figure 2.7, middle. When the SUT is located at best-focus, the incident power on both parts of each dual cell is equal. When the SUT is too close to the objective lens, the spots move away from the optical axis; when the SUT is too far away

from the objective lens, they move towards the optical axis. A FES can be obtained with (2.2). The letters indicating the detectors have been chosen such that the same formula can be used with consistent sign.

In optical storage systems, sometimes only one of the branches is used to determine the FES (Stan, 1998). For distance measurement, however, the dual design is beneficial since it reduces tilt sensitivity. Surface tilt causes a shift of the spots over the detectors. When the shift is orthogonal to the apex of the double wedge, the incident power in one branch increases while it decreases in the other. At best-focus, this effect cancels out because of the mirror-symmetry in the system. Outside best-focus, symmetry is broken and the effect of tilt is only partly compensated. In this direction, a coarse tilt measurement can be obtained by taking the difference of the power in both branches, using the relationship:

$$S_{tilt} = \frac{(A+D)-(B+C)}{A+B+C+D} \tag{2.3}$$

In principle, the method is insensitive to tilt that causes beam shift parallel to the apex of the double wedge.

Laser detector grating unit

Various types of Laser Detector Grating Units (LDGU) are described in (Ono, et al., 1986). LDGUs are distinguished from other layouts by application of a diffraction grating to separate the incident and returning beams and their integral packaging of laser, detector and diffraction grating into a single opto-electronic component. They employ various kinds of measurement methods and grating types. Figure 2.7, left, depicts a layout of a LDGU applying the same measurement principle as the double Foucault method discussed before. The use of this type of LDGU for surface measurement is described in (Visscher, 1992) and (Ehrmann, et al., 1998).

The signal processing and inherent measurement characteristics for this detector are the same as for the double Foucault method; the difference is that it is more compact. In optical storage applications, LDGUs are more stable due to the short thermal loop between the components; in this application, where environmental fluctuations are much smaller, this advantage is considered less important. Furthermore, the use of a separate light source is desirable because it increases the design freedom regarding the choice of light source.

Double critical angle method

The double critical angle method (Kohno, et al., 1988) is based on the critical angle method that is applied in optical storage (Stan, 1998). For explanatory purpose, the critical angle method will be considered first; it is schematically depicted in Figure

2.8, left. The method uses a critical angle prism (1) and a dual cell (2). The angle of the reflection surface of the prisms is chosen such that the angle of incidence of a collimated beam's wavefront is at the critical angle. Hence, the criterion for total internal reflection is met and all light is reflected to the dual cell.

When the SUT is too close to the objective lens, the returning beam is divergent, see detail A in Figure 2.8. On one side of the optical axis the angle of incidence of the wavefront increases, while on the other side of the optical axis it decreases. On the side where the angle decreases, the criterion for total internal reflection is not met and part of the light exits the prism through refraction, resulting in a decrease in incident power on photodiode A. When the SUT is too far away from the objective lens, the opposite happens as shown in detail B. The FES can be obtained by subtracting the signals:

$$FES = A - B \tag{2.4}$$

Using the Fresnel formulas (Hecht, 2002), reflectance as a function of angle of incidence can be calculated; for BK7, the relation is shown in Figure 2.8, right. Because the reflectance just below the critical angle is very sensitive to the angle, small displacements of the SUT result in relatively large signals.
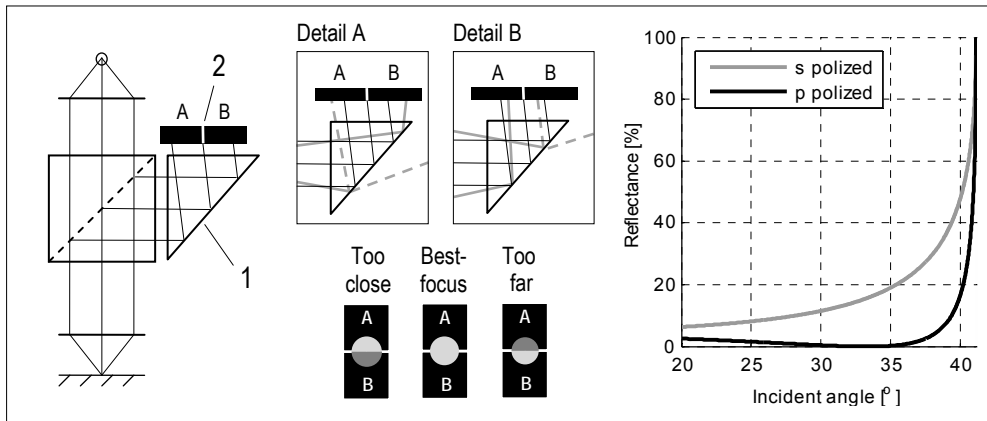
Figure 2.8: Layout of the critical angle method (left), details for divergent and convergent beams (middle top), spots on the detectors for various SUT positions (middle bottom) and reflectance as a function of angle of incidence for BK7 (right).

The schematic layout of the double critical angle method is depicted in Figure 2.9. This method features two improvements to the normal critical angle method. One improvement is the addition of a non-polarizing beam splitter (1) and a second measurement branch (2) with opposite orientation to the beam. When the signals are processed in accordance with equation (2.2), this renders the method insensitive to surface tilt, also when the SUT is out of focus. Just as with the double Foucault

method, tilt measurement in one direction can be obtained with equation (2.3). The spots on the detectors for various SUT positions are shown in Figure 2.9, right.

The other improvement is the use of two double critical angle prisms instead of two single critical angle prisms (3). This effectively squares the reflectance curve shown in Figure 2.8, making it the most sensitive auto-focus method found in literature.
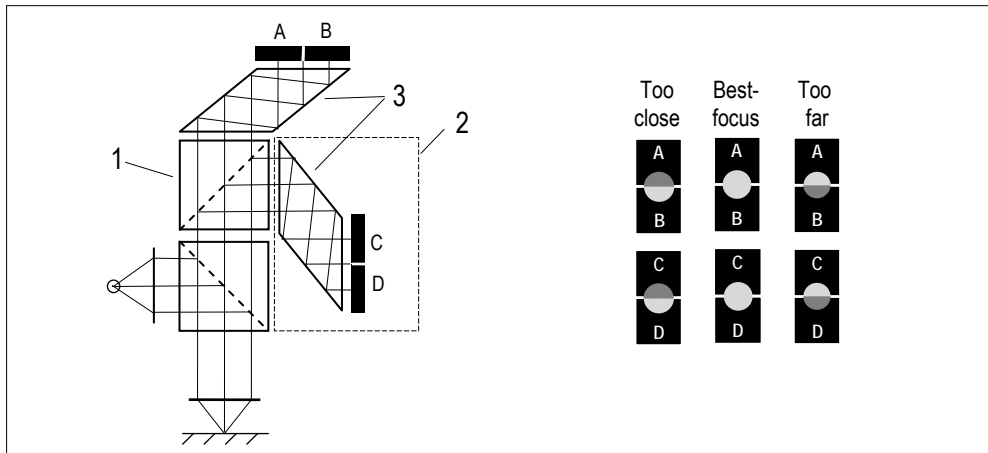
Figure 2.9: Layout of the double critical angle method (left) and the spots on the detectors for various SUT positions (right).

General characteristics of auto focus methods

Range and resolution of these systems vary with choice of parameters; when a resolution of 1 nm has to be realized the range and measurement spot are typically in the order of a few micrometers. The four evaluated auto focus methods are intensity robust because of the normalization applied in signal processing. Furthermore, they are all capable of measuring rough surfaces although the response will change significantly and uncertainty is expected to increase. This is not considered a problem since rough surfaces do not need to be measured as accurately as reflective surfaces.

Because it is relatively easy to attain a high intensity spot at the detectors, bandwidths of hundredths of kilohertz can be achieved with these methods. Tracking of the SUT by translating the objective lens is possible for all methods, allowing high bandwidth tracking as well.

All but the astigmatic method can be used for course measurement of the tilt in one direction. By realizing the measurement system a second time, rotated 90° around the optical axis, tilt measurement in two directions can be obtained. Moreover, the

resulting redundancy in distance measurement can help to decrease measurement uncertainty and improve resolution; however, it requires a more elaborate system. Alternatively, all auto focus methods allow relatively straightforward addition of a dedicated aperture correction system as discussed in Subsection 2.1.3.

Due to the large detector areas and lack of spatial filtering, out-of-focus light with a large range of angles can reach the detector, leading to measurement errors. The effects of ambient light can be suppressed by modulating the light source; this however has no effect on the stray light originating from the source itself or on ghost reflections. Consequently, the measurement uncertainty of this type of sensors is often limited by stray light and ghost reflections. The use of a coherent light source further increases the measurement uncertainty, due to amplification of coherent disturbances through interference.

2.2.4 CONFOCAL DISTANCE MEASUREMENT METHODS

Confocal distance measurement methods have evolved from confocal imaging, patented by Marvin Minsky in 1957 (Minsky, 1957). Confocal techniques are widely applied in optical microscopy because they strongly attenuate light originating from out-of-focus planes. This is highly advantageous when imaging the inside of thick transparent objects, such as biological tissue samples.

Confocal method

The use of confocal principles for surface profiling is reported in (Wilson, 1980) and (Hamilton, et al., 1982). A schematic layout of a confocal system is shown in Figure 2.10, left.
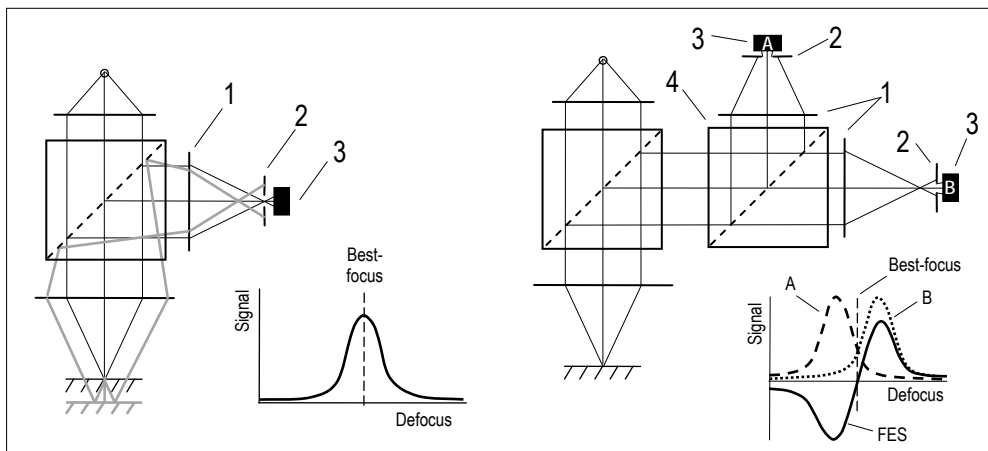
Figure 2.10: Layout of a confocal measurement method (left) and of a differential confocal measurement method (right).

A lens is placed in the returning light beam and an undersized pinhole is located in the focal plane. If the SUT is in best-focus, the image spot coincides with the pinhole so that most of the light reaches a photodiode placed behind the pinhole. When the SUT moves out of focus, the spot size at the pinhole increases so that the incident power on the photodiode decreases. The typical bell-shaped curve of the signal as a function of defocus is depicted in Figure 2.10 also.

A disadvantage of this method is that the amplitude of the bell-shaped curve changes when the intensity of the returning light changes, leading to measurement errors. Furthermore, because the signal is mirror symmetric around best-focus, it cannot be deduced from the signal on which side of best-focus the surface is located. The only part where the location can be unambiguously determined is best-focus, but here the derivative of the curve is zero, leading to zero sensitivity.

To simultaneously overcome the intensity dependence and the position ambiguity, the objective lens is sometimes vibrated through focus and the peak at best-focus is detected. This measurement method still suffers from the zero sensitivity at the top of the bell-shaped curve. Moreover, as discussed in Subsection 2.1.2, vibrating is not considered desirable.

Alternatively, as described in (Lee, et al., 1997), a flank of the bell-shaped curve can be tracked, which solves the problem of position ambiguity but not that of intensity dependence. Moreover, the flank of the curve is tilt dependent which also causes measurement uncertainty. Sometimes this method is called "differential confocal microscopy". The confocal method that is discussed next, uses the flanks of *two* bell-shaped curves, and is often referred to just as "differential confocal measurement", and occasionally as "*bipolar* differential confocal measurement". In the remainder of this thesis, the term "differential confocal" refers to confocal systems that use the flanks of *two* bell-shaped curves.

Differential confocal method

The differential confocal method is described in (Simon, 1970), (Fainman, et al., 1982), (Tan, et al., 2002), (Tan, et al., 2003) and (Zhao, et al., 2004). A layout of a differential confocal system is shown in Figure 2.10, right.

A non-polarizing beam splitter (4) splits the returning beam into two equal beams that are focused by two imaging lenses (1). Undersized pinholes (2) and photodiodes (3) are placed close to focus in both measurement branches; just behind focus in one branch and an equal distance in front of focus in the other branch. This leads to two bell-shaped curves: one on either side of focus. When the signal of one photodiode is subtracted from the other photodiode signal, a FES is obtained that has a zero-crossing at best-focus and a relatively steep and long near-

linear range around it; the bell-shaped curves and the difference signal are depicted in Figure 2.10.

When the difference signal is normalized by dividing it by the sum of the detector signals, the FES becomes insensitive to intensity variations; it is given by:

$$FES = \frac{B - A}{A + B} \tag{2.5}$$

With the SUT located at best-focus, the principle is insensitive to surface tilt since due to symmetry around focus, the lateral shift of the beam in one branch is compensated by the lateral shift in the other branch. If the SUT is too close to or too far away from the objective lens, symmetry is broken and the mechanism does only partially compensate the lateral shift of the beams. Because surface tilt, pinhole diameter and the offset from best-focus have a non-linear effect on the bell-shaped curves, and because of the effect of normalization, it is expected that there is at least one combination of pinhole diameter and offset that leads to minimal tilt dependent error.

Chromatic confocal method

Chromatic confocal sensors apply focus multiplexing by wavelength encoding through use of monotonic chromatic aberration (Molesini, et al., 1984). The popularity of confocal chromatic sensors has grown considerably over the last decade. A layout is shown in Figure 2.11.
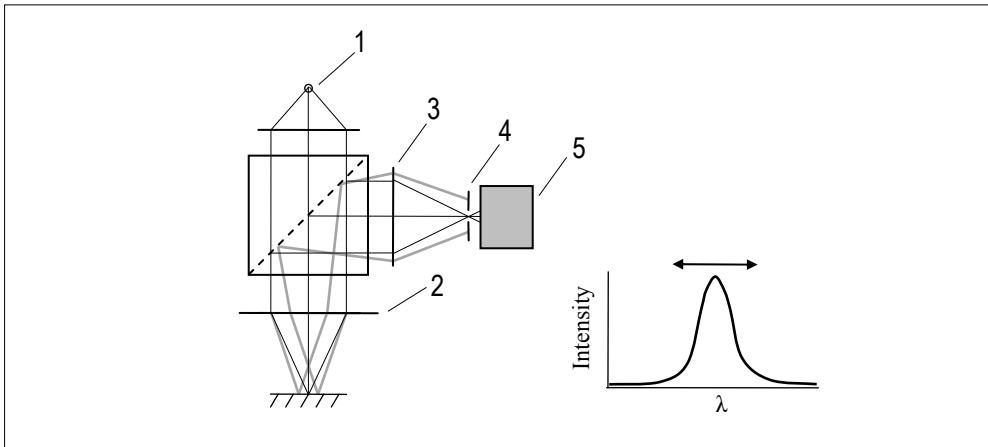
Figure 2.11: Schematic layout of a chromatic confocal measurement method.

A broadband light source (1) is focused by a lens system (2) that has strong monotonic chromatic aberration, hence the foci of light with various wavelengths

have different axial positions. As a result, light of a certain wavelength is always in focus (depicted in black in Figure 2.11). The returning light is focused by a lens (3) on a pinhole (4) through which most of the wavelength that is in focus at the SUT passes, while wavelengths that are focused far away from the SUT are blocked. A spectrometer is used to determine which wavelength is in focus. This results in an intensity peak that shifts to different wavelengths for different positions of the SUT. A single fiber, which automatically functions as a pinhole, is often used for both delivery and pickup of the light.

Particularly elegant of this method is that over the whole range the distance measurement is made by determining which wavelength is in focus, rendering the measurement principle insensitive to stray light, roughness and surface tilt. A practical disadvantage of this measurement method is the limited power of suitable broadband sources, especially when measuring materials with low reflectance. Consequently, the maximum bandwidth that can be achieved for a certain required resolution is relatively low when compared to other methods. The low power of the source also poses a problem for the integration of an aperture correction system, which, in spite of the inherent tilt insensitivity of the principle, is still needed to correct for tilt dependence due to aberrations.

The commercially available system with the highest resolution found has 2 nm resolution, 10 nm measurement uncertainty and 20 µm range (STIL A.S., 2008). This system has a length of 191 mm, making it too long to fit in the volume envelope; possibly, the optical train of such a sensor can be folded to solve this. Furthermore, the sensor's outer dimensions make it too blunt to measure surfaces that are tilted 5° and it has a mass of 380 g which has an adverse effect on tracking bandwidth and leads to relatively large dynamic errors in the measurement machine.

A chromatic confocal system in which most of the lens system remains stationary and only an objective lens at the front is translated (Dobson, et al., 1997) might be suitable to decrease the moving mass. Another possible solution to lower moving mass and simultaneously allow for a slimmer sensor head are recent miniaturization developments such as reported in (Ruprecht, 2008).

2.2.5   MEASUREMENT METHOD EVALUATION AND SELECTION

A qualitative evaluation of various inherent properties of the measurement methods is shown in Table 2.1. Triangulation, white-light interferometry and conventional confocal measurement are considered not suitable or practical; for the sake of completeness, they are included in Table 2.1 (dark gray background), nevertheless.

Sensor type	Tilt robust @ zero-crossing	Tilt robust over range	Inherent tilt measurement	Intensity robust	Stray light / ghost robust	Roughness robust	Bandwidth	Suitable for tracking	Incorporation of tilt measurement system
90° Triangulation	++	++	no	++	--	++	++	-	+
White-light interferometer	+	0	no	0	+	0	--	--	+
Astigmatic	+	+	no	++	-	0	++	++	+
Double Foucault	+	0	1D	++	-	0	++	++	+
LDGU	+	0	1D	++	-	0	++	++	+
Differential critical angle	++	++	1D	++	-	0	++	++	+
Confocal	++	-	no	--	++	0	++	-	+
Differential confocal	++	0	no	++	++	0	++	++	+
Chromatic confocal	++	++	no	++	++	++	0	0	-

Table 2.1: Qualitative appreciation of various inherent properties of the measurement methods.

Among other objections, the most important reason to abandon triangulation is that it cannot practically meet the resolution and measurement uncertainty requirements. White-light interferometry is rejected because it requires complicated sampling and data processing. Additionally, in normal tracking mode it allows only small tracking errors, while vibrating mode leads to a more complex system and uneven sample spacing. Conventional confocal measurement is not regarded suitable since it does not incorporate normalization, leaving the measurement susceptible to various sources of noise and measurement error. Furthermore, measuring the top of the bell-shaped curve leads to relatively large measurement uncertainty, due to the zero-sensitivity at the top of the curve, while measuring at the flank of the curve leads to large tilt dependent error.

This leaves six suitable methods: the four auto focus methods and two of the confocal methods. For all these systems, the measuring spot has to be in the order of a few micrometers to achieve the required measurement uncertainty and resolution. A weakness of single-point measurement with a small spot is its sensitivity to surface defects and aliasing. This weakness can however be partly

overcome by applying a higher sample rate than necessary in combination with data filtering.

Regarding many criteria, the remaining six measurement principles have about equal potential. Chromatic confocal measurement is less suitable for high bandwidth tracking, but this is compensated by a larger allowable measurement range. An important difference between the six measurement methods is, nevertheless, that for the four auto focus methods (light gray background in Table 2.1), stray light and ghost reflections can easily lead to measurement errors due to the large detector areas and lack of spatial filtering. This can pose a significant threat to the required 10 nm measurement uncertainty. An additional problem is that the effects of stray light and ghost reflections are often only observed when the system is already realized. For the two remaining confocal measurement methods, on the contrary, the chances are much smaller that significant amounts of unwanted light reach the detectors because of the spatial filtering by the pinholes. Therefore, the differential and chromatic confocal methods are considered the best candidates (white background in Table 2.1).

Of the chromatic confocal sensors found at the start of this research, the resolution and bandwidth were a bit lower than desired. Moreover, the low power light sources applied in chromatic confocal sensors make it more difficult to incorporate a measurement system for aperture correction, especially when measuring materials with low reflectance. Consequently, it has been decided to pursue a sensor that applies differential confocal measurement as the primary measurement method.

2.3 DIFFERENTIAL CONFOCAL DESIGN CONSIDERATIONS

Chapter 3 deals with the formulation of analytical models that describe the response of a differential confocal sensor. Parallel to the development of the models, a demonstrator has been built to gain experience with differential confocal measurement and to test the applicability of the models as a design tool. The demonstrator will be addressed in Chapter 3 as well.

Influence of system parameters on response

Without deriving a model, some approximate design rules can however be deduced by considering the principle of operation. A differential confocal sensor functions by detecting the amount of defocus of the SUT through measurement of the laser power that propagates through the pinholes. Hence, the sensor's response primarily depends on depth of focus, the diameter of the pinholes and their offset from best-focus. When dealing with optical distance sensors, spot size at the SUT is often

given much importance. The reason is that it determines the lateral resolution of the sensor. In this application, the lateral resolution is of less importance since the primary aim is to measure form of the surface, not micron-sized features.

The pinhole diameter and offset from best-focus can be freely chosen, while depth of focus depends on the wavelength and the convergence of the incident light. The convergence of the incident light can be characterized by the Numerical Aperture (NA), a dimensionless number that is defined as (Hecht, 2002):

$$NA = n_i \sin(\theta_{max}) \tag{2.6}$$

in which:

n_i is the index of refraction of the medium, in this case air ($n_i = 1$), and

θ_{max} is the half-angle of the cone of light.

Estimation of suitable design parameters

Because a measurement uncertainty of 10 nm and a resolution of 1 nm are required for the final design, it has been decided to initially aim at a demonstrator with high sensitivity, to see what can be achieved. To attain high sensitivity and little noise, the pinhole parameters must be chosen such that the incident power on the detectors is high and that the amount of light propagating through the pinholes changes strongly as a function of defocus. Intuitively, a pinhole diameter close to the spot size and a pinhole offset of about the depth of focus seem good choices. Furthermore, a short depth of focus leads to relatively high sensitivity. The higher the NA of the incident light is, the shorter the depth of focus will be. The NA of the incident light is however limited by the NA of the object lens that focuses it. The NA of a lens can be calculated with (2.6) as well; in this case, θ_{max} is the angle between the optical axis and the marginal ray (i.e. the outer ray that can propagate through the lens).

Numerical aperture and prevention of vignetting

Since the SUT can tilt up to α_t, the returning cone of light can tilt $2\alpha_t$. As discussed in Subsection 2.1.3, it is desirable to prevent vignetting of the returning beam. Therefore, a Gaussian beam is used with a diameter chosen such that the reflected beam can be fully collected by the object lens. Consequently, the maximum attainable half-angle of the incident light is reduced by $2\alpha_t$ for a given object lens.

Some theory of Gaussian beams is discussed at the beginning of Chapter 3; an important property for the discussion here is that only 86% of the power of a Gaussian beam is contained within the diameter as it is often defined. Hence, if the maximum permissible diameter of the beam is calculated using this definition of

diameter, a considerable part of the reflected beam still falls outside the aperture of the object lens for large tilts. For this reason, the diameter is multiplied by a factor of 1.5, thus giving an extended diameter that contains 99% of the beam's power.

In Figure 2.12, left, the lens is depicted with its maximum collectable cone of light, as well as the incident beam with its conventional defined diameter and its extended diameter. The diameter of the object lens, the extended beam and the incident beam are indicated by D_L, D_{ext} and D_B, respectively, their half-angles by θ_L, θ_{ext} and θ_B.
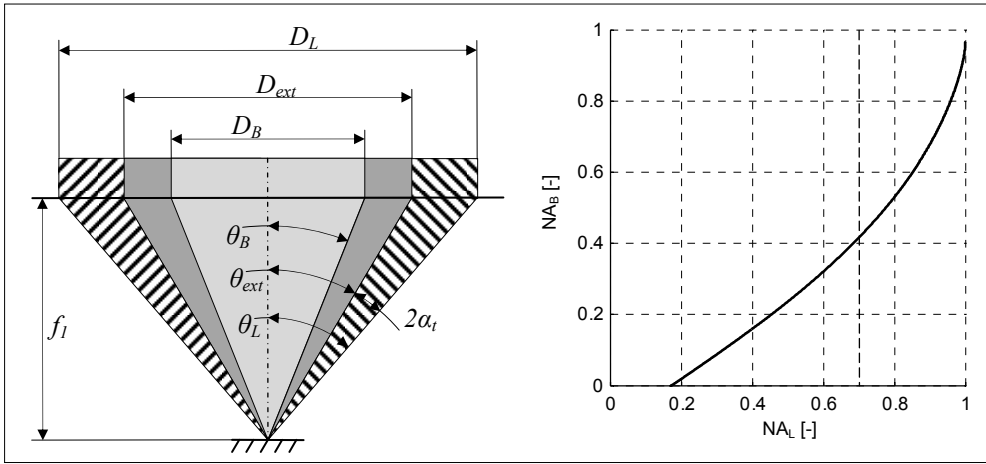
Figure 2.12: Maximum collectable cone of light of a lens, and the incident beam with customary defined diameter and extended diameter (left). A graph of the maximum allowable NA of the incident beam as a function of lens NA at 5° surface tilt (right).

Using goniometric relations, the NA of the beam, NA_B, can be expressed as:

$$NA_B = \sin(\arctan(2/3 \cdot \tan(\arcsin(NA_L) - 2\alpha_t))) \tag{2.7}$$

in which:

NA_L is the numerical aperture of the object lens, and

α_t is the tilt of the SUT.

In Figure 2.12, right, a graph of the maximum NA of the beam is shown as a function of the NA of the lens for a maximum allowable surface tilt of 5°. It can be seen that to achieve a high NA of the illuminating beam of light, and thus a short focal depth, an object lens with a relatively large NA is needed. Due to their volume and mass, a small aspheric singlet is thought practical for use as object lens. The NA of commercially available, suitable singlets is limited to about 0.7; this limit is indicated in the graph of Figure 2.12, right, with a dashed line.

Optical magnification

It is beneficial to apply image lenses with focal lengths considerably larger than that of the object lens. This way the spot size in image space is magnified by the ratio of the focal lengths and the depth of focus increases quadratic with this ratio. Hence, measurement errors resulting from pinhole displacements are attenuated and tolerances on initial alignment are eased.

It has been decided to pursue a dual-stage system in which a differential confocal sensor tracks the SUT by focusing the object lens, while an interferometer measures the translation of the object lens. A beam splitter and position sensitive detector will be added to allow aperture correction. Now that the conceptual design has been decided on, the foundation for a specific design will be laid in the next chapter, in which analytical models of the differential confocal system are formulated and a demonstrator setup is presented.

3 ANALYTICAL DIFFERENTIAL CONFOCAL MODEL & DEMONSTRATOR

An analytical dimensionless differential confocal model based on Gaussian beam theory will be presented. From this, a dimensional model for optoelectric behavior of a complete differential confocal sensor is derived. To gain experience with differential confocal measurement, a demonstrator has been built, which has resulted in insights and design rules for prototype development. Furthermore, it is shown that the models agree with the experimental results, building confidence in the use of the models as design tools.

When designing a sensor, its properties must be chosen such that they match the requirements of the envisaged application. Various properties characterize performance and robustness of distance sensors, some of which are:

- resolution,
- measurement range,
- measurement uncertainty,
- dynamic range, and
- tilt dependent error.

Since several of these properties are conflicting, it is useful if the relation between properties and design parameters is known, so that optimization becomes possible. For a differential confocal sensor, the fundamental design parameters that determine the aforementioned properties are:

- pinhole diameter,
- pinhole offset,

- laser wavelength,

- diameter of the collimated beam, and

- focal lengths of the object and image lenses.

Because of the size of the design space and the number of properties involved, it is valuable to understand the relationships at a fundamental level. Hence, analytical models, which relate Focus Error Signal (FES) to the design parameters and the translation of the Surface Under Test (SUT), are desired.

When evaluated in dimensional space, the properties of the sensor depend on all design parameters, which complicates optimization. To avoid this obstacle, the analytical expressions are derived in a dimensionless form. Since the *dimensionless* properties and *dimensionless* parameters are invariant to changes in the other, *dimensional,* parameters, this enables optimization to be performed in two steps. The first step of the optimization is to choose the *dimensionless* pinhole diameter and offset so that an optimal combination of *dimensionless* properties is achieved; the second step is to choose the remaining *dimensional* parameters so that the *dimensional* properties match the requirements of the application.

In some texts, converting dimensional physical quantities into dimensionless parameters, is called normalizing that quantity. Here, however, normalization of the FES and conversion into a dimensionless form, are two distinct operations. The term "normalization" is used only for dividing the difference between the pinhole signals through their sum, which varies as a function of the SUT's position. The term "dimensionless" is used only for expressions and parameters that have been made dimensionless by division through a dimensional physical quantity that is a sensor property, and thus independent of the position of the SUT.

In Chapter 4, a differential confocal property analysis and optimization based on the models presented here will be treated. As explained in that chapter, some properties are best treated in dimensional and others in dimensionless space. Furthermore, some properties depend on the normalized and others on the unnormalized FES. Hence, *dimensionless* as well as *dimensional* models of both the *normalized* and *unnormalized* signals will be presented here.

3.1   ANALYTICAL MODELING

Various analytical models for confocal systems (Sheppard, et al., 1978), (Corle, et al., 1986), (Wilson, et al., 1987), (Kino, et al., 1989) and (Valter, 1992), as well as differential confocal systems, (Corle, et al., 1987) and (Tan, et al., 2002) are found in literature. These analytical models describe the usual mode of operation for

differential confocal systems, where the lens is overfilled. Since the approach chosen here is based on a system with an underfilled lens and a Gaussian beam, these models are not valid. Therefore, new analytical models have been derived and are presented here.

A dimensionless analytical model describing a differential confocal system in its most basic form is treated first. The dimensional form that includes all the above-mentioned fundamental design parameters and accounts for the optoelectronics, is derived from this dimensionless model. The deduction of the basic model is included in the main text because of the insight this provides into the working principle and because of the importance of the models for the design process.

3.1.1 MODEL ASSUMPTIONS

The analytical models are based on the assumptions listed below.

- The system is aberration free. Although impossible to realize in real systems, one can get close by using well aligned, diffraction limited components.

- The SUT is perpendicular to the measurement direction. A numerical approach will be applied for evaluation of the response of a tilted SUT, as will be treated in Subsection 4.2.4.

- The SUT produces ideal specular reflection. Since the primary goal is to measure high quality polished optics this is considered a valid assumption.

- The SUT is flat. Some freeforms are expected to have local curvatures that will significantly influence distance measurement; this is not accounted for in the model.

- The laser output is a diffraction limited TEM_{00} Gaussian beam. For a laser with a M^2 close to 1, this is a realistic approximation (ISO standard 11146).

- Detectors exhibit linear responsivity. This closely resembles reality, as linearity is better than 1% for intensities and photocurrents sufficiently below the detector's saturation limits (OSI Optoelectronics, 2009).

When drawing conclusions from the models, these limitations must be taken into account.

3.1.2 GAUSSIAN BEAM THEORY

The laws of geometrical optics provide a good model for the behavior of light in many situations. When light is focused, however, geometrical optics predicts that the diameter of the intensity distribution, which gives the distribution of power per area, approaches zero. This would require infinite intensity to satisfy the laws of

energy conservation. Near focus, the laws of diffraction provide a much better estimate for the behavior of light, instead. At focus, light contracts to a distribution with finite dimensions and intensity from where it starts expanding again, as is schematically depicted by the solid line in Figure 3.1, left.
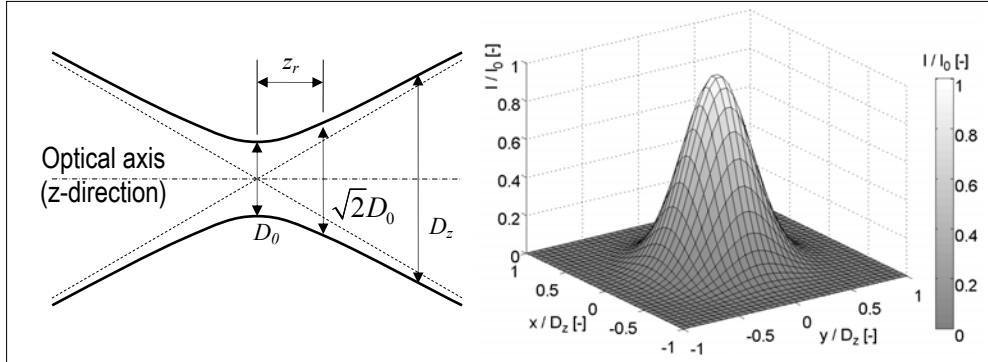
Figure 3.1: Gaussian intensity profile.

Most intensity distributions, such as a top-hat distribution, change shape as well as scale along the optical axis since the intensity distribution in focus is the Fourier-transform of the far field. Gaussian distributions exhibit the special property that they maintain their shape under Fourier-transformation. A laser operating in TEM_{00} mode will be used in the system, because it produces a Gaussian intensity distribution. This will give a gradual change of power transfer through the pinholes when the SUT moves through focus with only one inflection point per side of focus, ensuring near equal sensitivity over a large range. The lateral intensity distribution of a Gaussian beam is shown in Figure 3.1, right, and is given by (3.1) (O' Shea, 1985):

$$I_D = I_0 \, e^{-2\left(\frac{D^2}{D_z^2}\right)} \tag{3.1}$$

where:

I_D is the intensity at diameter D in the beam,

I_0 is the intensity at the center of the beam,

D is the diameter in the beam at which the intensity is I_D, and

D_z is the e^{-2}-diameter of the beam at position z.

The e^{-2}-diameter is defined as the diameter where the intensity is e^{-2}, or about 0.135, times the central intensity of the distribution; it scales with position along the beam's axis. This border is represented by the solid lines in Figure 3.1, left,

whereas the borders predicted by geometrical optics, which form the asymptotes for the e^{-2}-border, are represented by dashed lines. Of course, as diameter scales, the central intensity changes as well to satisfy energy conservation.

The smallest diameter of the beam is called the beam waist diameter D_0 and the region where it is located is called the beam waist. The beam waist region extends over a distance along the beam's axis defined by the z positions where the e^{-2} diameters are $\sqrt{2} \cdot D_0$ or less. The distance from this boundary to the beam waist is called the Rayleigh range and denoted by z_r. The divergence of the beam is described by the angle between the asymptotes, θ_b. If one of the three beam properties z_r, D_0 and θ_b is given, the others can be calculated provided that the wavelength of the light is known. A useful equation giving D_z at an arbitrary position along the optical axis is:

$$D_z = D_0 \sqrt{1 + \left(z / z_r \right)^2}$$

(3.2)

3.1.3 MODEL FOR DIMENSIONLESS FES

In the differential confocal system, the signal is generated by measuring the amount of power passing through the pinholes. From (3.1) it can be derived by integration that the power passing through a circle concentric with the beam's axis with diameter D, can be expressed as:

$$P_{enc} = P_b \left(1 - e^{-2\left(\frac{D^2}{D_z^2} \right)} \right)$$

(3.3)

where:

P_{enc} is the encircled power, and

P_b is the power of the laser beam at the pinholes.

To convert the encircled power to a dimensionless parameter, it can be divided by the total power of the beam at the pinhole; it is called the Fractionally Transferred Power (FTP) and is given by:

$$\overline{FTP} = \frac{P_{enc}}{P_b} = 1 - e^{-2\left(\frac{D^2}{D_z^2} \right)}$$

(3.4)

In the notation used in formulas, dimensionless parameters and variables are denoted with an overhead bar. FTP is dimensionless by definition and therefore the

notation is redundant, nevertheless, the bar is kept for the sake of clarity and consistency.

The situation can be treated as if both pinholes are in object-space, as can be seen in Figure 3.2.
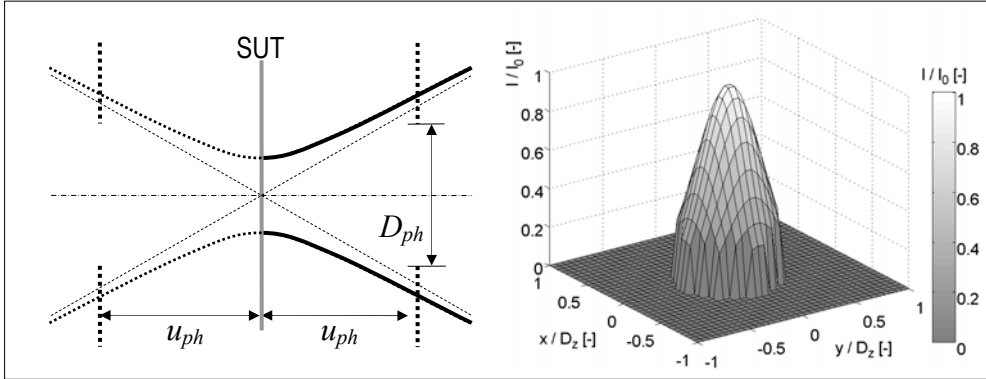
Figure 3.2: Pinholes as if within object space (left) and intensity distribution just behind a pinhole with $D_{ph}/D_z=0.8$ (right).

The circle through which energy is transferred is physically defined by the pinhole diameter, therefore $D=D_{ph}$. In (3.4) D_z is the parameter which varies with defocus as described in (3.2), where z is determined by the distance of the pinholes to the waist, thus $z=z_{ph}$. Substitution yields:

$$\overline{FTP} = 1-\mathrm{e}^{-2\left(\frac{D_{\mathrm{ph}}{}^2}{D_0{}^2\left(1+\left(z_{\mathrm{ph}}/z_{\mathrm{r}}\right)^2\right)}\right)} \tag{3.5}$$

To make the expression indifferent for the system's focal lengths, beam diameter, and laser wavelength, the other parameters are made dimensionless by dividing dimensions in radial direction by D_0 and dividing axial dimensions by z_r. Hence:

$$\overline{D}_{\mathrm{ph}} = \frac{D_{\mathrm{ph}}}{D_0} \tag{3.6}$$

$$\overline{z}_{\mathrm{ph}} = \frac{z_{\mathrm{ph}}}{z_{\mathrm{r}}} \tag{3.7}$$

$$\overline{D}_0 = 1$$

$$\overline{z}_r = 1$$

Substitution in (3.5) yields:

$$\overline{FTP} = 1 - e^{-2\left(\frac{\overline{D}_{\text{ph}}^2}{1+\left(\overline{z}_{\text{ph}}\right)^2}\right)} \tag{3.8}$$

Since the light is reflected at the SUT, an axial displacement of the SUT will result in twice as much axial displacement of the beam's waist at the pinholes. Therefore the distance between the pinhole and the beam's waist, $\overline{z}_{ph}$, for pinhole 1 can be written as:

$$\overline{z}_{\text{ph1}} = 2\overline{u}_{\text{sut}} + \overline{u}_{\text{ph}} \tag{3.9}$$

and $\overline{z}_{ph}$ for pinhole 2 can be written as:

$$\overline{z}_{\text{ph2}} = 2\overline{u}_{\text{sut}} - \overline{u}_{\text{ph}} \tag{3.10}$$

in which:

$\overline{u}_{\text{sut}}$ is the defocus of the SUT, divided by z_r, and

$\overline{u}_{\text{ph}}$ is the offset of the pinholes to the waist at best-focus divided by z_r.

Substituting these parameters into formula (3.8) and subtracting the FTP of pinhole 1 from the FTP of pinhole 2 will yield the expression for the dimensionless Focus Error Signal (FES):

$$\overline{FES} = \overline{FTP}_2 - \overline{FTP}_1 = e^{-2\left(\frac{\overline{D}_{\text{ph}}^2}{1+\left(2\overline{u}_{\text{sut}}+\overline{u}_{\text{ph}}\right)^2}\right)} - e^{-2\left(\frac{\overline{D}_{\text{ph}}^2}{1+\left(2\overline{u}_{\text{sut}}-\overline{u}_{\text{ph}}\right)^2}\right)} \tag{3.11}$$

To simplify the expression, the dimensionless exponential terms associated with FTP at the pinholes are denoted $\overline{e}_1$ and $\overline{e}_2$ and defined as:

$$\overline{e}_1 = e^{-2\left(\frac{\overline{D}_{\text{ph}}^2}{1+\left(2\overline{u}_{\text{sut}}-\overline{u}_{\text{ph}}\right)^2}\right)} \tag{3.12}$$

and

$$\overline{e}_2 = e^{-2\left(\frac{\overline{D}_{\text{ph}}^2}{1+\left(2\overline{u}_{\text{sut}}+\overline{u}_{\text{ph}}\right)^2}\right)} \tag{3.13}$$

Then the FES can be written as:

$$\overline{FES} = \overline{e}_1 - \overline{e}_2 \tag{3.14}$$

From equation (3.11) derivation of the *normalized* dimensionless FES is straightforward and yields:

$$\overline{\mathbb{FES}} = \frac{e^{-2\left(\frac{\overline{D}_{ph}^{2}}{1+\left(2\overline{u}_{sut}-\overline{u}_{ph}\right)^{2}}\right)} - e^{-2\left(\frac{\overline{D}_{ph}^{2}}{1+\left(2\overline{u}_{sut}+\overline{u}_{ph}\right)^{2}}\right)}}{2 - e^{-2\left(\frac{\overline{D}_{ph}^{2}}{1+\left(2\overline{u}_{sut}+\overline{u}_{ph}\right)^{2}}\right)} - e^{-2\left(\frac{\overline{D}_{ph}^{2}}{1+\left(2\overline{u}_{sut}-\overline{u}_{ph}\right)^{2}}\right)}} = \frac{\overline{e}_1 - \overline{e}_2}{2 - \overline{e}_2 - \overline{e}_1} \tag{3.15}$$

In formulas, the normalized FES and its properties are denoted with double-struck letters, in text and graphs however, this is refrained from for practical reasons; in these instances, normalization will be mentioned in the context instead. The models presented here allow finding optimal values with respect to sensitivity and range for various pinhole diameters and offsets. To transform the dimensionless values into physical values, $\overline{D}_{ph}$ and $\overline{z}_{ph}$ can be multiplied with D_0 and z_r in object space and scaled with system-magnification between the object- and image-space. Alternatively, they can be scaled with D_0 and z_r in image space.

With the models derived here, graphs of the FTP and the normalized as well as unnormalized dimensionless FES can be made, as shown in Figure 3.3 as a function of dimensionless defocus of the SUT. For this graph, both the dimensionless pinhole diameter and offset are set to 1; the equations used to calculate FTP, unnormalized and normalized FES are (3.8), (3.11) and (3.15) respectively. The shape of the pinhole signals and the effect of normalization on the FES can be seen. The *dimensionless* defocus can be multiplied by z_r in object space to find the *dimensional* defocus. For a short wavelength and a high NA, z_r is relatively small, which leads to high sensitivity and a short measurement range. For a long wavelength and a low NA, z_r is relatively large, which leads to low sensitivity and a large measurement range. For example, a wavelength of 400 nm and an NA of 0.65 gives a z_r of 0.25 µm, whereas a wavelength of 700 nm and an NA of 0.1 gives a z_r of 22 µm.

The three-dimensional graphs of Figure 3.4 show what influence modification of dimensionless pinhole parameters has on the FES. This can be useful to develop a qualitative understanding of changes to the FES as pinhole parameters are varied separately. To quantitatively explore these changes when both pinhole parameters are varied simultaneously, Figure 3.5 contains a matrix of nine combinations of $\overline{D}_{ph}$ and $\overline{z}_{ph}$.
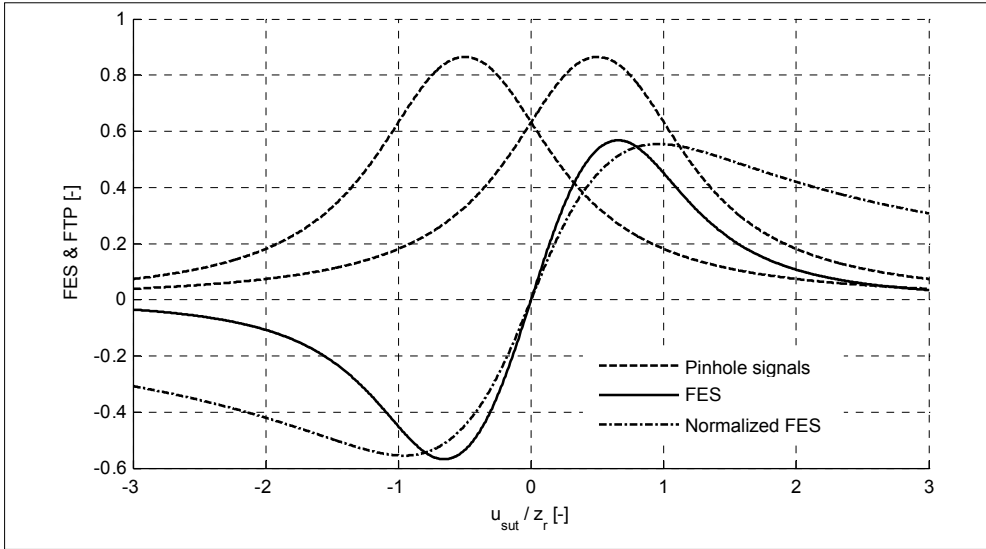
Figure 3.3: Unnormalized and normalized response for dimensionless pinhole diameter and offset equal to 1.
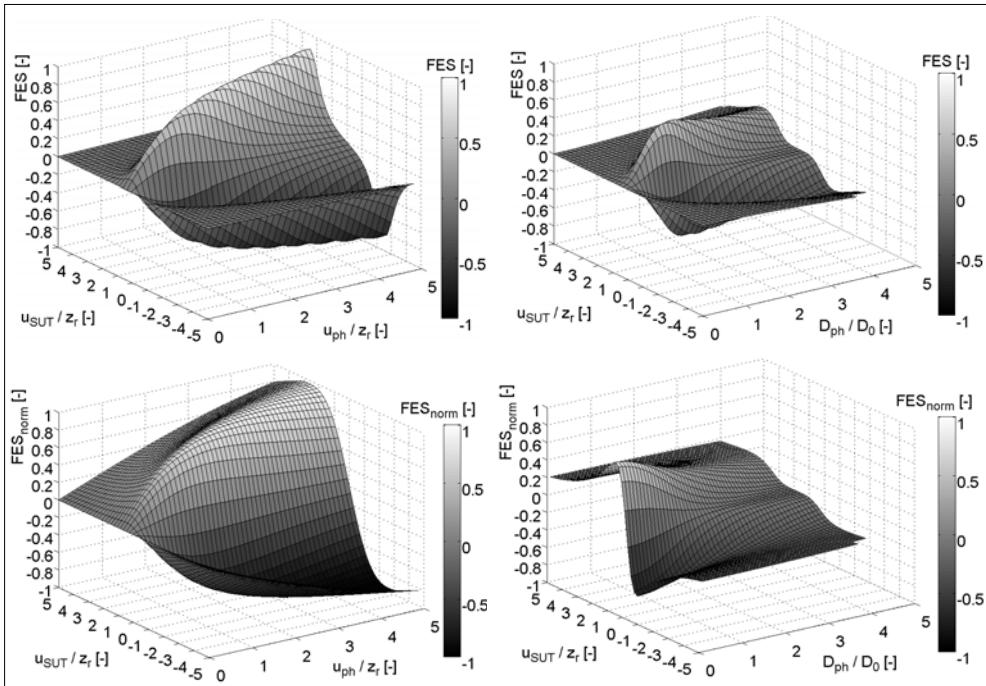
Figure 3.4: Influence of pinhole offset (left) and pinhole diameter (right), for both unnormalized FES (upper) and normalized FES (lower).

51
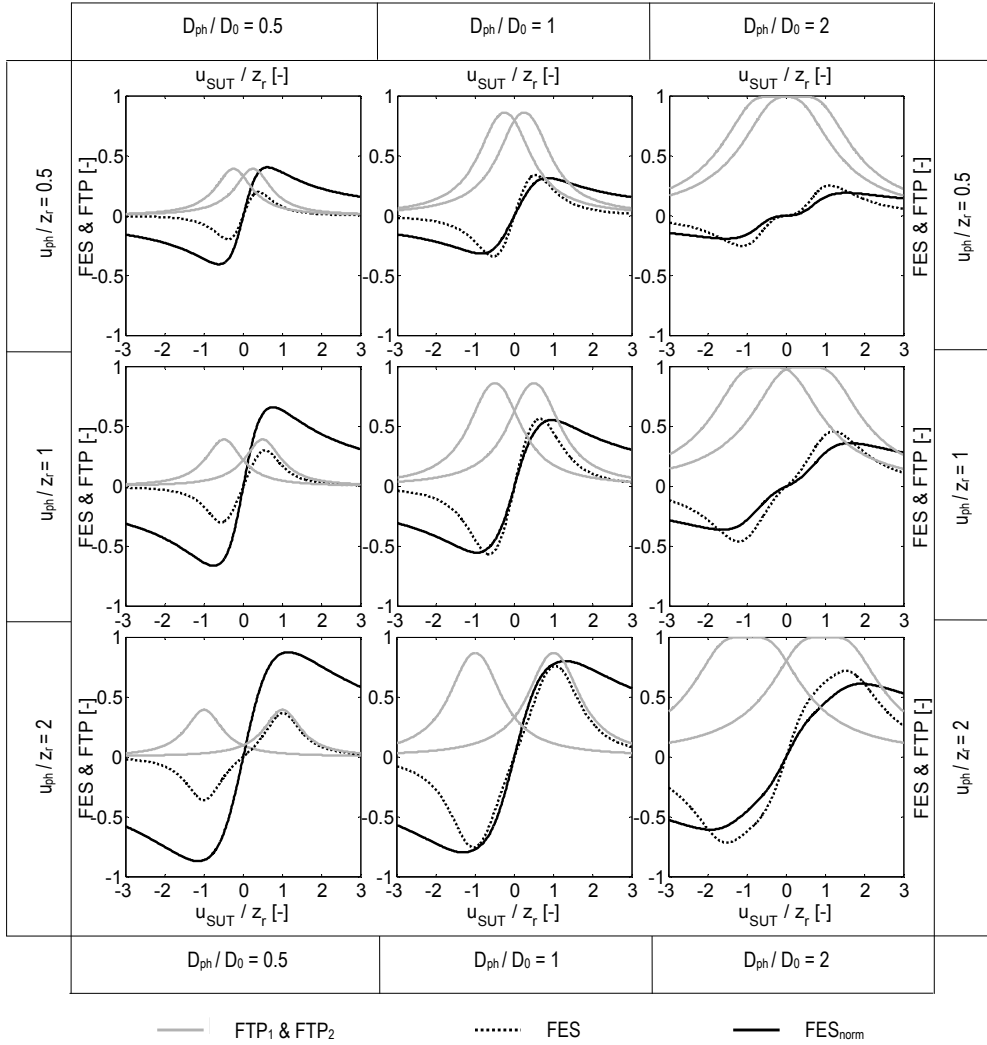
Figure 3.5: Responses for the nine combinations of dimensionless pinhole diameter and offset for dimensionless values of 0.5, 1 and 2.

For the smallest pinholes ($D_{ph}=0.5D_0$) in the left column, the bell-shaped curves are low since only a small portion of the power can pass through the pinholes. Accordingly, the FES curve is small as well. Nevertheless, for the two lower graphs of the left column the normalized FES reaches large values, because of division by a small signal. Thus, a large normalized FES does not necessary mean that the underlying measurement principle is robust.

For the large pinhole size (D_{ph}=2D_0) in the right column, the tops of the bell-shaped curves are flat, since there is a range of defocus where nearly all energy passes through the pinhole. This causes increased non-linearity for both the normalized and unnormalized FES. Non-linearity also increases for large pinhole offsets relative to the width of the bell-shaped curves, as in the lower left corner; here normalization improves the linearity. The top row reveals that both unnormalized and normalized FES curves stay small if the pinholes have a small offset from focus.

For the combinations shown here, the slope of the unnormalized dimensionless FES is steepest for a dimensionless diameter and offset of the pinholes equal to 1, i.e. a pinhole diameter equal to the diameter of the beam waist and a pinhole offset equal to the Rayleigh range. An analytical derivation presented in Appendix A, confirms that this combination of pinhole parameters indeed yields the highest sensitivity.

In Chapter 4, various properties will be defined to describe and compare the characteristics of different configurations.

3.1.4 MODEL FOR DIMENSIONAL FES AND OPTO-ELECTRONICS

The dimensionless FES, treated in the preceding subsection, does not provide insight in relationships between system parameters other than dimensionless pinhole parameters and does not reveal sources of noise and measurement errors. Therefore, it is beneficial to derive a dimensional model, which includes:

- focal-lengths of objective and image lenses, f_1 and f_2, respectively,

- beam diameter at the waist of the collimated laser beam, D_B,

- laser wavelength, λ,

- power of the laser, P_L,

- transmissivity of the optical train, η_{opt},

- reflectance of the Surface Under Test, R_{sut},

- responsivity of the photodiodes, R_{pd}, and

- gain in the signal processing electronics, G_{el}.

Because the last five terms represent linear subsystems that are in series, these terms are multiplication factors that together constitute a gain factor in series with the FES.

Furthermore, the *dimensionless* parameters are multiplied by physical quantities to attain dimensional parameters. These physical quantities are the Rayleigh range, z_r, and waist diameter, D_0 in object space. From Gaussian beam theory (O' Shea, 1985) it can be derived that D_0 and z_r can be expressed as:

$$D_0 = \frac{4\,f_1\,\lambda}{\pi\,D_B} \qquad \text{and} \tag{3.16}$$

$$z_r = \frac{4\,f_1^{\,2}\,\lambda}{\pi\,D_B^{\,2}} \tag{3.17}$$

To transform dimensions from the object space to the image space it is useful to express magnification of the system as a ratio between the focal-lengths:

$$M = \frac{f_2}{f_1}$$

As treated step by step in Appendix B, when the above-mentioned parameters and expressions are substituted in formula for the *dimensionless* FES (3.11), the following equation is obtained:

$$FES = G_{el}\,R_{sut}\,R_{pd}\,\eta_{opt}\,P_L\ldots$$

$$\left(e^{\left(-2\left[\frac{D_{ph}^{\,2}}{\left(\frac{4\,f_2\,\lambda}{\pi\,D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^{\,2}\,u_{ph}}{4\,f_2^{\,2}\,\lambda} + \frac{2\pi D_B^{\,2}\,u_{sut}}{4\,f_1^{\,2}\,\lambda}\right)^2\right)} \right]\right)} - e^{\left(-2\left[\frac{D_{ph}^{\,2}}{\left(\frac{4\,f_2\,\lambda}{\pi\,D_B}\right)^2 \left(1 + \left(\frac{-\pi D_B^{\,2}\,u_{ph}}{4\,f_2^{\,2}\,\lambda} + \frac{2\pi D_B^{\,2}\,u_{sut}}{4\,f_1^{\,2}\,\lambda}\right)^2\right)} \right]\right)} \right) \tag{3.18}$$

To shorten the notation, the exponential terms can be expressed as e_1 and e_2, thus:

$$e_1 = e^{\left(-2\left[\frac{D_{ph}^{\,2}}{\left(\frac{4\,f_2\,\lambda}{\pi\,D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^{\,2}\,u_{ph}}{4\,f_2^{\,2}\,\lambda} + \frac{2\,\pi\,D_B^{\,2}\,u_{sut}}{4\,f_1^{\,2}\,\lambda}\right)^2\right)} \right]\right)} \tag{3.19}$$

and

$$e_2 = e^{\left(-2\left[\frac{D_{ph}^{\,2}}{\left(\frac{4\,f_2\,\lambda}{\pi\,D_B}\right)^2 \left(1 + \left(\frac{-\pi\,D_B^{\,2}\,u_{ph}}{4\,f_2^{\,2}\,\lambda} + \frac{2\,\pi\,D_B^{\,2}\,u_{sut}}{4\,f_1^{\,2}\,\lambda}\right)^2\right)} \right]\right)} \tag{3.20}$$

giving:

$$FES = G_{el}\, R_{sut}\, R_{pd}\, \eta_{opt}\, P_L \left(e_1 - e_2 \right) \tag{3.21}$$

With $\mathbb{G}_{el}$, the gain in the normalizing signal processing electronics, the normalized FES becomes:

$$\mathbb{FES} = \mathbb{G}_{el} \left(\frac{ e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4f_2\,\lambda}{\pi D_B}\right)^2\left(1+\left(\frac{\pi D_B^2\, u_{ph}}{4f_2^2\,\lambda}+\frac{2\pi D_B^2\, u_{sut}}{4f_1^2\,\lambda}\right)^2\right)} \right)} - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4f_2\,\lambda}{\pi D_B}\right)^2\left(1+\left(\frac{-\pi D_B^2\, u_{ph}}{4f_2^2\,\lambda}+\frac{2\pi D_B^2\, u_{sut}}{4f_1^2\,\lambda}\right)^2\right)} \right)} }{2 - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4f_2\,\lambda}{\pi D_B}\right)^2\left(1+\left(\frac{\pi D_B^2\, u_{ph}}{4f_2^2\,\lambda}+\frac{2\pi D_B^2\, u_{sut}}{4f_1^2\,\lambda}\right)^2\right)} \right)} - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4f_2\,\lambda}{\pi D_B}\right)^2\left(1+\left(\frac{-\pi D_B^2\, u_{ph}}{4f_2^2\,\lambda}+\frac{2\pi D_B^2\, u_{sut}}{4f_1^2\,\lambda}\right)^2\right)} \right)} } \right) \tag{3.22}$$

or in the shortened form:

$$\mathbb{FES} = \mathbb{G}_{el} \left(\frac{e_1 - e_2}{2 - e_1 - e_2} \right) \tag{3.23}$$

The difference between equation (3.21) and (3.23) shows what the benefits of normalization are: the reflectivity of the SUT, R_{SUT}, the power of the laser, P_l, the transmission of the optical train, η_{opt}, and the responsivity of the photodiodes, R_{pd1}, disappear from the expression. This means in practice that variations in these parameters due to, for example, temperature change or noise will not affect the measurement. Hence noise and measurement uncertainty are reduced. Note that for photodiode responsivity and transmission of the optical train this only concerns the common part of their drift and noise. Actually they are not physically the same, however, if the photodiodes are placed closely together, electromagnetic interference and thermal drift in both photodiodes are expected to resemble each other for a large part. This does not count for some other sources of noise, such as photon shot noise.

3.2 DEMONSTRATOR SETUP

To provide insights and design rules for prototype development, a demonstrator has been realized. Furthermore, the demonstrator allows the analytical models to be compared to experimental data, so that confidence in their validity can be built. To limit research time and costs it has been decided to realize a basic system, such as schematically depicted in Figure 2.10, right. This does not include the additions for translating the object lens and for aperture correction.

3.2.1 DEMONSTRATOR DESIGN CONSIDERATIONS

The design considerations for the demonstrator are in many respects different from those for the prototype. In experimental science, insight grows as one goes along and therefore ease of modification is important. This is sometimes conflicting with other properties associated with good mechanical design, such as high stiffness, low mass, low hysteresis, thermal stability etc. These, however, are of lesser concern in a laboratory environment where temperature is well regulated, the setup is vibration isolated and orientation with respect to the gravity vector will remain unchanged.

With the demonstrator, the measurement uncertainty of 10 nm has not been aimed for. To evaluate uncertainty at this level, the displacement of the SUT relative to the objective lens has to be measured with a reference measurement system. Above-mentioned concessions mainly concern the optomechanical and optical design of the setup. The signal processing electronics are representative for the prototype.

Linos's Microbench building system (Linos photonics, Inc, 2008) is chosen because of the tradeoff between ease of modification, ease of alignment, robustness and stiffness matches the needs in this particular case. To improve stability, next to these standard parts, a central holder for some of the critical optics is machined that also serves as the central structural unit. The SUT will be substituted by a mirror, which is translated along the optical axis. A laptop with a DAQ system is used to enable automated driving of the system and logging of data.

3.2.2 INITIAL RESEARCH SETUPS

Before designing a demonstrator, two initial research setups have been built. The first setup has been put together using posts on an optical table. Although this system has many disadvantages from an optomechanical viewpoint, it has been used for the short setup time and its ample availability. A photograph of this setup is shown in Figure 3.6, left. This setup has been useful in developing a feeling for alignment requirements and for taking manual FES readings.

A second research setup has been built with the Microbench system and can be seen in Figure 3.6, right. This setup is equipped with a piezo actuated test mirror and data acquisition, which allowed for automated logging of FES curves and qualitative assessment. Mechanical stability, however, is not high enough to vary just one design parameter while maintaining an otherwise unchanged system.
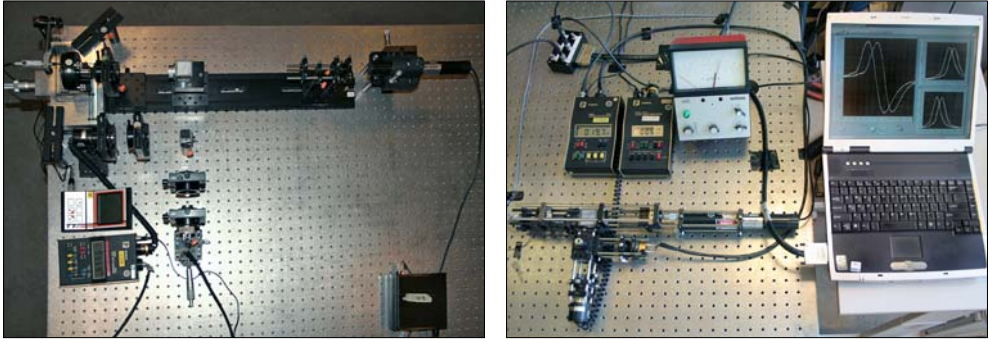
Figure 3.6: Initial research setups, using posts (left) and using the Microbench system (right).

3.2.3   LIGHT SOURCE

After gathering experience with the initial research setups, the demonstrator has been developed. Its optical components and beam path are shown in Figure 3.7. A 1 mW helium-neon-laser (1) is used as light source. The laser has linearly polarized output at 632.8 nm with a measured e^{-2}-diameter of 0.62 mm. A HeNe-laser has been chosen since these lasers typically produce Gaussian beams with good beam quality, and are readily available. The beam is expanded to 3.1 mm diameter by a five times Beam Expander (BE) consisting of two positive doublets (2 and 4). At the focus, a 50 μm non-reflecting pinhole (3) spatially filters the beam. With equation (3.16), it can be calculated that the waist diameter at the beam expander is 21 μm. Since the pinhole diameter is about 2.4 times the e^{-2}-boundary at the waist, the Gaussian intensity profile will not be disturbed significantly.

3.2.4   OPTICS

A 12.7 mm laser-line polarizing beam splitter cube (5) couples the beam in. An air-spaced quarter wave plate (6) for 633 nm is used. The objective lens is a 40x microscope objective (7) with an NA of 0.65 and a focal length of 4 mm. This combination of objective lens and beam diameter gives the maximum beam NA that can be accommodated by the objective, taking into account an extended beam diameter that is a factor 1.5 larger than the e^{-2}-diameter and a surface tilt of 5°, as is discussed in Section 2.3. The objective is designed for use with a cover glass of 0.17 mm thickness. A cover glass is however not incorporated in the setup, which will lead to some spherical aberration. Regarding focal length and NA, the objective is comparable to the aspherical singlets, which are envisaged as object lens for the prototype. Using equations (3.16) and (3.17), the waist diameter at the SUT can be calculated to be 1.0 μm with a Rayleigh range of 1.3 μm.
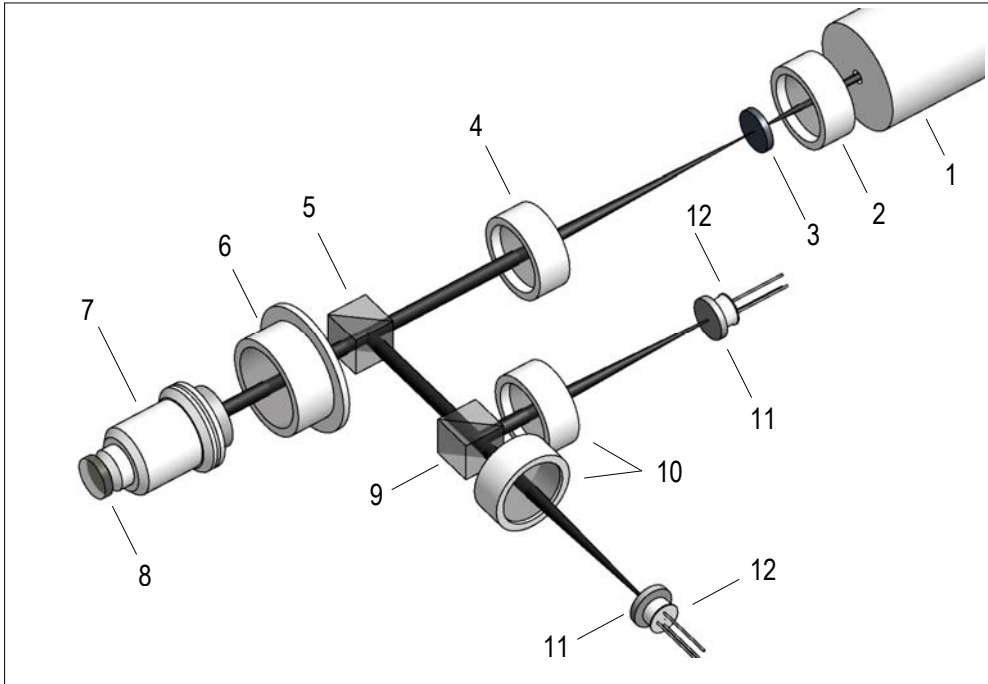
Figure 3.7: Optical components and beam path of the demonstrator.

As a substitute for the SUT, a 10 mm circular, metal-coated mirror (8) with protective dielectric coating is used. The mirror can be actuated along the optical axis by means of a piezo actuator (not depicted).

The returning beam is split into two beams with equal intensities by a laser-line non-polarizing beam splitter cube (9). Doublets with focal lengths of 60 mm (10) are applied as imaging lenses, giving a magnification factor of 15. High magnification is beneficial because Rayleigh range scales quadratic with magnification. Hence, measurement errors due to axial displacements of the pinholes relative to the focal points of the imaging lenses are attenuated by a factor of 225. Another benefit of high system magnification is that it loosens the tolerances on radial pinhole alignment.

From Figure 3.5, it seems that with respect to sensitivity and near-linear range, a pinhole diameter of about 1 to 1.5 times the waist diameter and a pinhole offset of about 2 times the Rayleigh range is a good initial choice. Furthermore, from the standpoint of energy utilization it can be calculated from (3.2) and (3.4) that at best-focus, for a dimensionless pinhole offset of 2 and diameter of 4/3, the FTP is about 0.5, which intuitively seems a good choice. The waist diameter can be estimated using (3.16) which gives a waist diameter of 15.6 µm and hence a

nominal pinhole diameter of 20.8 µm. Therefore, a 20 µm non-reflecting pinhole (11) is chosen, behind which the photodiodes (12) are placed.

3.2.5   OPTOMECHANICS

The Microbench system uses anodized aluminium mounting cells, which slide fit over four parallel 6 mm diameter hardened steel rods at the corners. Four functional branches are built in this way, which connect to a central optics holder, as shown in Figure 3.8. The demonstrator is not mounted rigidly to the optical table but instead is freely suspended by a foam layer on top of the vibration-isolated table. This prevents thermal snapping due to differences in thermal expansion coefficients and time constants, while at the same time allows easy accessibility of the bottom of the setup.
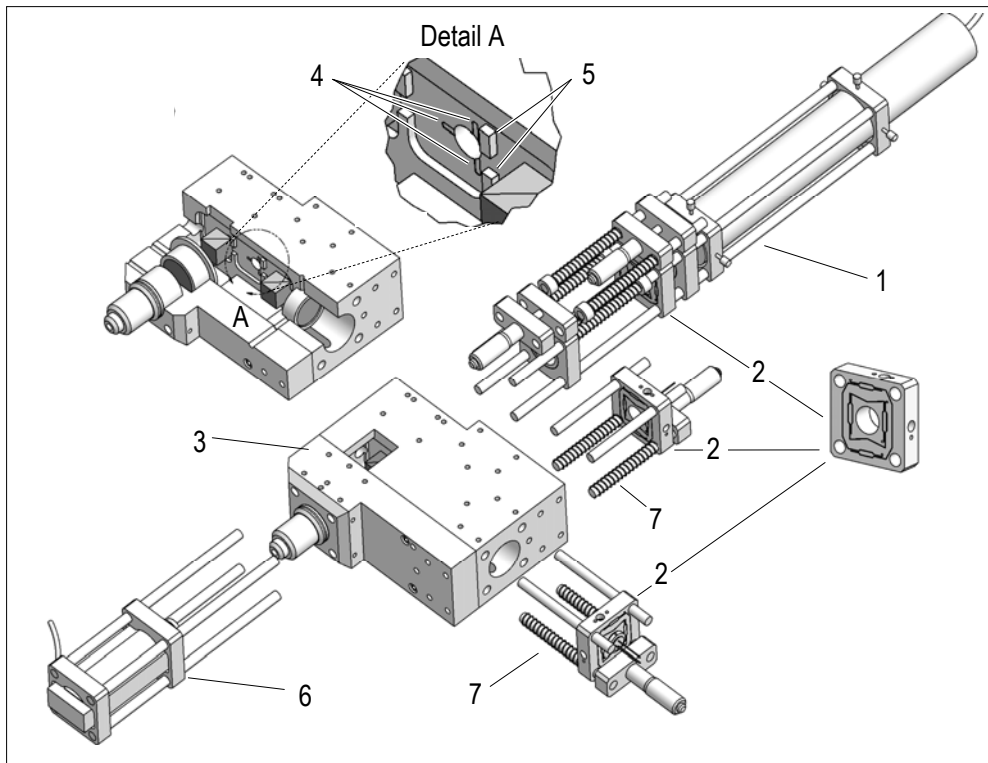

Figure 3.8: Exploded view of the setup showing its main components.

To minimize the number of adjustments in the system, positioning and orientation of the components is preferably achieved through dimensional tolerances. It remains necessary to align several of the components to enable compensation of

stacked tolerances in the rest of the system or because part precision does not suffice. These components and their adjustable degrees of freedom are:

- the 3 pinholes   3 translations (1 axial, 2 radial),

- the laser   3 rotations (1 axial, 2 tilt) and 2 translations (radial),

- the mirror   2 rotations (2 tilt) and 1 driven translation (axial), and

- the 2nd BE lens   1 translation (axial).

The only driven degree of freedom is the translation of the mirror along the optical axis.

Laser branch

In Figure 3.8, the laser branch (1) is shown with the laser mounted in two Microbench mounting plates using eight push screws. This is over-constrained while the laser housing experiences temperature differences of about 50 K between when the laser is turned off or on. Nevertheless, the pointing stability seems to be stable over weeks during which the laser goes through many on-off-cycles. A drawback of this way of mounting is that adjusting the tilt angles and radial translations requires several iterations because these degrees of freedom are not decoupled. Since these alignments do not need to be performed often, this is considered acceptable.

The primary doublet of the BE is locked in a fixed position. The pinhole for beam washing is aligned to focus with a manipulator for radial translations (2), which allows sub micrometer positioning, see Figure 3.8. It has two nested flexural guides consisting of leaf spring parallelograms, which are adjusted by two push-pull screw pairs. Axial positioning is achieved by sliding along the Microbench steel rods after which the position is locked by screws. These axial adjustments are made with a micrometer head, which is pretensioned by two compression springs. The same method is used for axial positioning of the secondary BE lens.

Central holder

The central optics holder (3) serves as mount for most of the optical components that have no adjustments: the beam splitters, the quarter wave plate, the objective lens and the two image lenses. The central holder is custom built for this setup and is machined by the GTD at Eindhoven University of Technology.

It consists of two parts, which are bolted together; their relative position is maintained by two dowels. The larger part has a pocket, as can be seen in Figure 3.8. During machining care is taken to achieve good alignment of the bores to the

bottom surface of the pocket. On this surface, the beam splitters are mounted and it is a reference surface during alignment procedures of other components.

The beam splitters are bonded with epoxy adhesive. To avoid having to control adhesive uniformity, the adhesive is not applied directly between the reference surface and the beam splitters. Instead, three 0.1 mm deep grooves (4) of 1.6 mm width are machined into the reference surface in the parts where the beam splitters are placed. When the adhesive shrinks during curing, the 0.1 mm thick adhesive layer pulls the beam splitters into contact with the reference surface. There are features (5) on the reference surface to position the beam splitters initially before alignment. In this way, during alignment only small movements are needed, preventing adhesive to flow between the reference flat and the beam splitters due to viscous transport.

The quarter wave plate, the microscope objective and the two doublets which form the images are mounted by placing them into slip fit holes where they are locked by means of two radial screws. For this purpose, the microscope objective is screwed into an adapter.

Mirror branch

In the mirror branch (6), two mounting plates accommodate a piezo-actuator. This actuator consists of a linear flexure guidance, driven by a piezo stack and is equipped with a capacitive feedback sensor. The actuator has a range of 38 µm, a resolution of 0.2 nm and 0.03 % linearity error over its range. The mirror is bonded to a stainless steel stalk and screwed to the actuator.

The actuator is operated in closed loop and controlled by a digital piezo controller. The piezo controller and actuator combination have been checked using an Agilent double-pass interferometer and proved to meet specifications to the level of precision attainable with the interferometer setup. Because the feedback loop is made by the controller, only trajectory points have to be sent from the laptop to the controller.

Detector branches

The pinholes in the detector branches (7) can be adjusted for three translations applying the same method as for the pinhole in the laser branch. With radial adjustment of the pinholes, the alignment errors in the rest of the system can be compensated. With the axial adjustments, it is possible to vary the pinhole offset or compensate for collimation errors. The photodiodes slip-fit inside a brass adapter tube, which is bonded to the pinholes using adhesive.

3.2.6   OPTOELECTRONICS AND DAQ

Photoconductive PIN diodes with an active area of 2.54 mm in diameter measure the amount of light that propagates through the pinholes. They are operated at 10 V reverse bias and their metal housings are grounded. The photodiodes are not placed directly behind the pinholes. If the distance between pinhole and detector goes to zero, like when a pinhole is applied as a mask on the surface of a photodiode, a part of the beam will be reflected back in itself if the beam is focused on the pinhole. This will illuminate the SUT with the reflected light, which leads to disturbed measurements. Hence the photodiodes are placed relatively far behind the pinholes. Due to the high divergence of the light behind the pinhole, this prevents almost all the light from reentering the system. Hence, to intercept the majority of the diffraction pattern behind the pinhole, the photodiode has to be large compared with the pinhole size of 20 μm. In general, larger photodiodes have higher noise levels. Photodiode size, however, does not influence shot noise, which is the dominant noise source for high bandwidth measurement at high photocurrents (Melles Griot, 2008).

The photodiodes are connected to an analog normalization circuit via an individually shielded four-core cable with low capacitance. Analog normalization is preferred over numeric normalization because of two reasons. The amplitude of the signals varies significantly for various design parameters and between a reflective and transmissive SUT, therefore a 16 bits DAQ might not have enough dynamic range to achieve the required resolution of 1 nm. In addition, for digital normalization to be effective for noise suppression at high frequencies, the data of the different channels has to be sampled at the same time, whereas low priced DAQ systems predominantly incorporate multiplexed ADCs.

The normalization circuit used is a PSD processing unit, modified to normalize photodiode signals. At 150 kHz and a total photocurrent of 500 μA, the circuit has a resolution of $15 \cdot 10^{-6}$ (C.S. Kooijman, ND). It has a normalization accuracy of 0.05 % for summed photocurrents of 1 μA to 500 μA. There are four output signals: the normalized differential photodiode signal (-10 V to 10 V), the summed photodiode signal (0 V to 10 V) and the two normalized photodiode signals (0 V to 10 V).

For data acquisition a 16 bits DAQ card in conjunction with a laptop and labVIEW is applied. The DAQ card features sixteen single ended or eight differential input channels with a multiplexed sample rate of 200 kS/s and two analog outputs. Including quantization, the total ADC noise is 1.5 LSB_{rms} for 10 V bipolar operation (National Instruments, 2007). The ADC is operated in differential input mode to minimize electromagnetic interference.

3.2.7   EXPERIMENTAL RESULTS

Two photographs of the demonstrator are shown in Figure 3.9. Due to the central mounting block, this setup is stable enough to adjust one degree of freedom while the alignment of the other components remains unchanged.
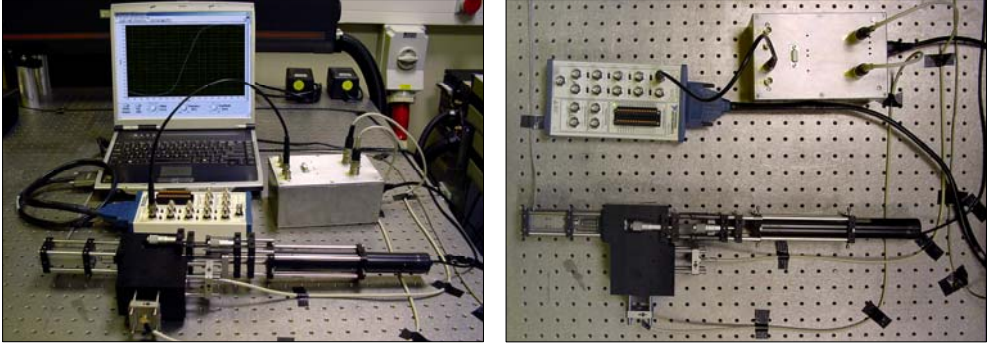

Figure 3.9: The demonstrator on a vibration isolation table.

The experimental results for various pinhole offsets, generated using the demonstrator, are the pinhole signals, FES, normalized FES curves and sensitivity. Furthermore, noise, short-term drift and quality of the FES for a system with 2 µm pinhole offset in object-space is investigated.

Measured pinhole signals, FES and normalized FES curves

Figure 3.10 shows experimentally obtained pinhole signals, FES and normalized FES curves for pinhole offset in object-space of 2 µm, 3 µm, 4 µm, 5 µm and 6 µm.

Sensitivity is one of the useful indicators of performance because it influences resolution and measurement uncertainty. To make this property easily applicable for other systems it is used in its dimensionless form and defined as:

$$\overline{S} = \frac{\delta \overline{FES}}{\delta \overline{u}_{\mathrm{sut}}} \tag{3.24}$$

where the FES is expressed in units of Fractionally Transferred Power (FTP).

The unnormalized dimensionless sensitivities for the zero-crossing for these configurations are 0.77, 0.67, 0.37, 0.12 and 0.06 respectively, while the normalized dimensionless sensitivities are 0.67, 0.87, 0.98, 0.56 and 0.48 respectively. This shows that there are system configurations for which the sensitivity of the underlying physical measurement principle has deteriorated significantly, although the normalized signal seems to exhibit sufficient sensitivity.

In dimensional space, the sensitivity of the normalized FES works out to be 9 V/µm, 12 V/µm, 14 V/µm, 8 V/µm and 7 V/µm respectively.
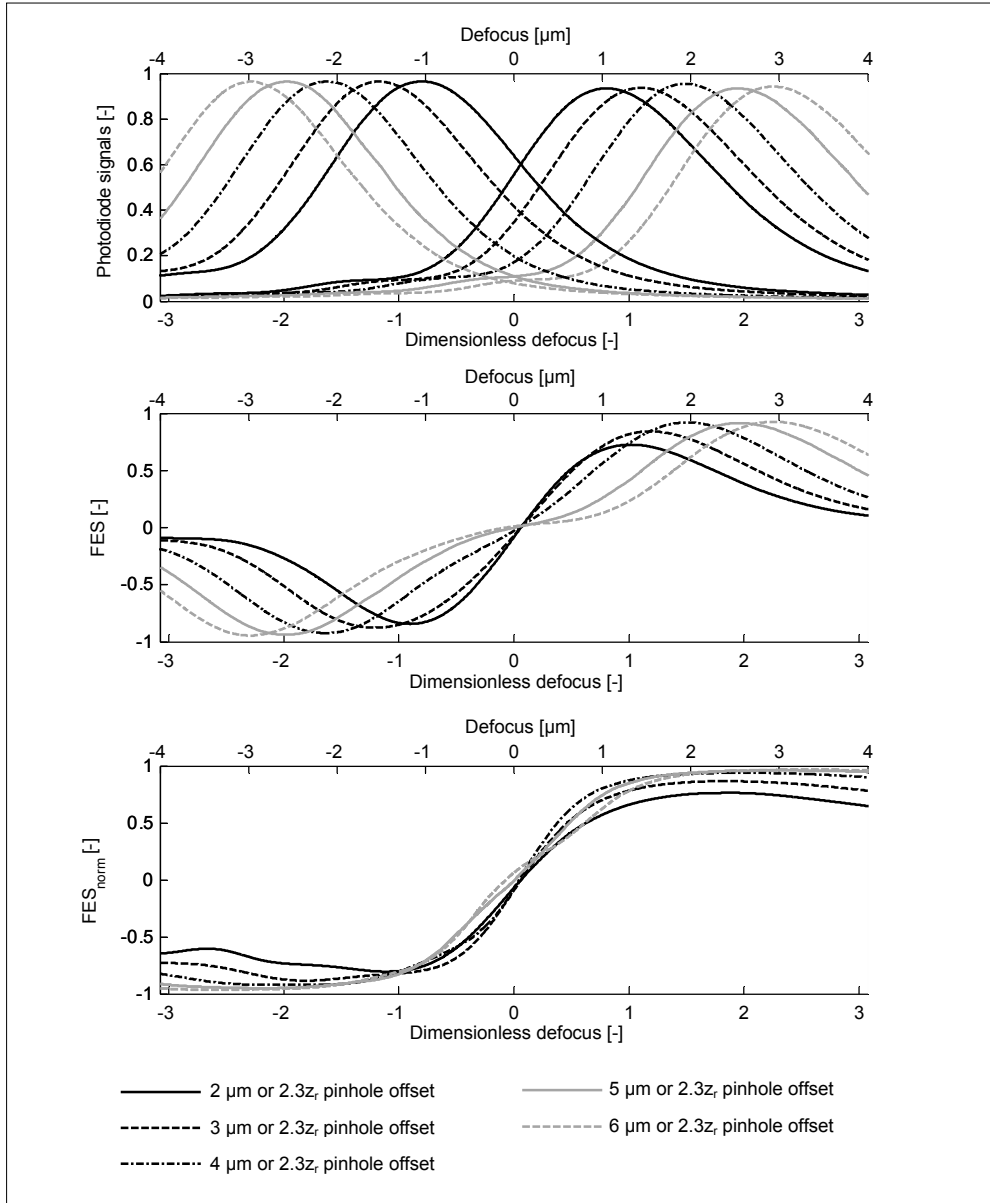

Figure 3.10: Experimental results, from top to bottom: pinhole signals, unnormalized and normalized FES for five different pinhole offsets.

Noise level

The electrical noise level of the normalized FES for 2 µm pinhole spacing has been analyzed at various positions close to zero-crossing. For this purpose, the mirror has been translated to various positions around best-focus and the FES has been sampled at 100 kHz. RMS noise levels are calculated over a 1 second window selected for low drift. The observed values vary between 0.7 and 0.9 mV. This can be converted to electrical nominal equivalent resolution by dividing through sensitivity, and works out to be 75 pm to 95 pm. Note that this is electrical nominal equivalent resolution; irregularities in the laser beam are expected to cause the real system resolution to degrade and vary over the measurement range.

The Thompson's multitaper method (Mathworks, 2007) is applied to estimate the power spectral density of the normalized FES and is shown in Figure 3.11. The calculations have been made with data taken at 100 kHz over a measurement time of 1 second, after subtraction of the signal mean. For the power spectral density estimate from 0 Hz to 50 kHz, the time-bandwidth product for the Thompson's multitaper algorithm has been set to 10 to yield low variance. For the 0 Hz to 1 kHz estimate a time-bandwidth product of 1.5 has been used to yield higher resolution at the expense of variance.
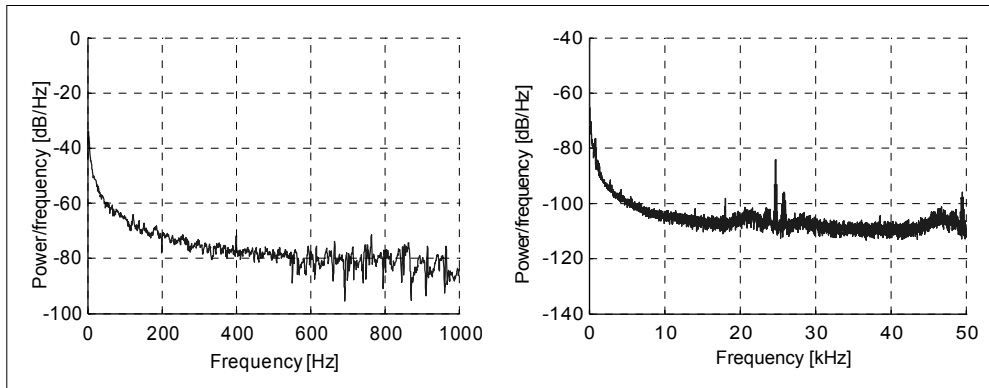
Figure 3.11: Power spectral density plot of the normalized FES around zero-crossing, the corresponding RMS of the signal is equivalent to 75 pm.

It is suspected that the peaks just under the Nyquist frequency are caused by the ADC, since they are present also, although at slightly different frequencies, when the ADC is short-circuited with a 50 Ω resistor. Apart from these peaks, the noise is approximately white from about 10 kHz and up. For a large portion, this white noise can be contributed to the electronics and ADC. Below 10 kHz, the 1/f noise becomes dominant. It is known that photodiodes are a source of 1/f noise but here it is probably the combined noise of various sources. There are no noticeable peaks at 50 Hz and 100 Hz, so the electromagnetic shielding and the common mode

rejection of the electronics and DAQ are sufficient. Furthermore, the absence of well-defined peaks at other frequencies in the sub kHz region indicates that mechanical vibration in the system does not contribute substantially to the measurement noise. This seems plausible, since the demonstrator is operated on a vibration isolation table, which has high attenuation at frequencies above 10 Hz. The NANOMEFOS machine will be supported by vibration isolators as well. Nevertheless, a higher level of vibration can be expected here because of vibration sources in the machine itself. Therefore, during opto-mechanical design of the prototype, the mass to stiffness ratio must be kept low.

Measurement drift

Although physically there is no difference between drift and the pseudorandom components considered as noise, the distinction is made here, based on practical considerations. The content of the fluctuations with frequencies higher than 0.5 Hz is considered to be noise, whereas the content with lower frequencies is considered to be drift. This 0.5 Hz limit is further discussed in Subsection 8.3.1.

The demonstrator has not been designed for high stability; hence, no long-term drift measurements are conducted. Short-term drift measurements, however, are believed to provide information, which might be useful. Therefore, a series of measurements has been taken, some with the system exposed directly to the ambient air and some with the system shielded. The measurements are taken over 5 second periods. Typical observed drift can be seen in Figure 3.12 on the left side; the worst-case observed drift is shown on the right side.
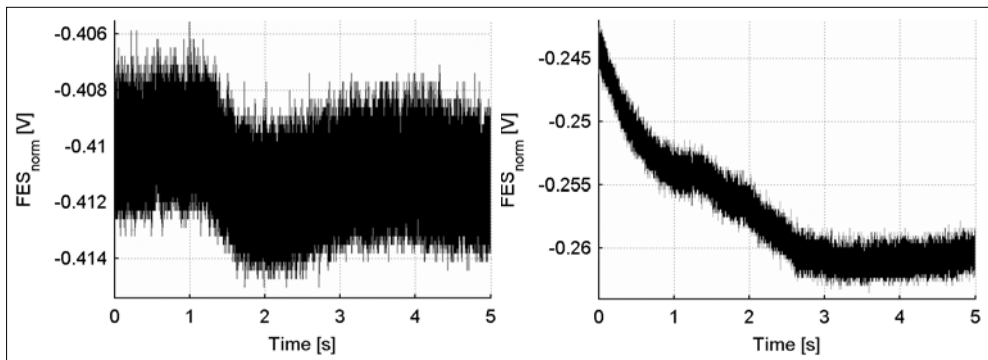

Figure 3.12: Typical (left) and worst-case (right) observed drift in demonstrator.

The Peak to Valley (PV) of the drift in typical measurements is equivalent to 0.3 nm. The PV for the worst case of drift observed is about 1.9 nm, however, it is not known how long the observed trend has been.

Influence of air currents and turbulence

The worst case of drift observed for the unshielded setup can be seen in Figure 3.13, left. The PV of the data in the graph is about 5 nm.

Since the SUT can rotate at considerable speed during freeform measurements, the layer of air through which the sensor measures can be turbulent. To test the effect of turbulence on the measurement in a qualitative way, airflow at the objective lens has been created. The result on the normalized FES is shown in Figure 3.13, right. This type of airflow is thought not to be representative for the airflow encountered in NANOMEFOS but it gives an indication of the effects that can be expected when measuring through a turbulent air layer. The PV of the data in the graph corresponds to about 3.5 nm.
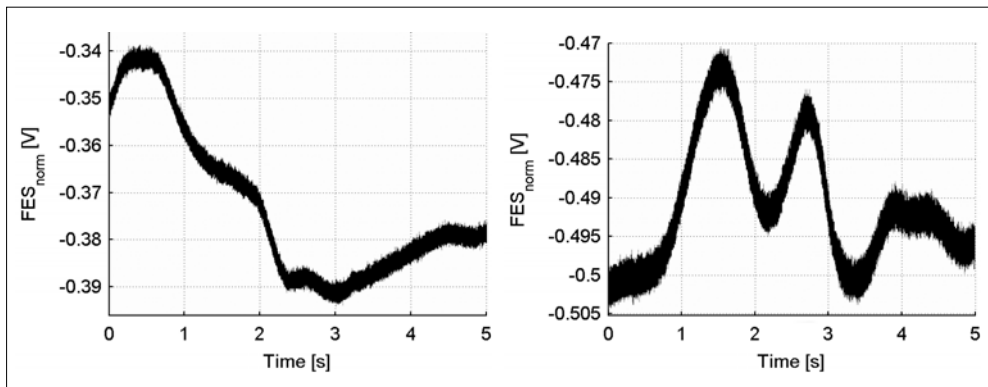
Figure 3.13: Worst-case signal drift for unshielded demonstrator (left) and signal fluctuations due to turbulent air currents (right).

Quality of the near-linear region

To see how well the near-linear part of the normalized FES reproduces, multiple measurements have been taken. During these measurements, the system had a dimensionless pinhole offset of 2.3 and a dimensionless pinhole diameter of 1.3. Comparison of these measurements around zero-crossing shows that there are deviations in the nanometer range that do not reproduce; consequently, they cannot be suppressed by calibration.

To obtain a quantitative estimate of this source of measurement uncertainty a least squares optimal polynomial of degree 6 has been fitted to the samples. Such a polynomial can be considered an approximation of a calibration curve taken over multiple measurements, with the exception that it adjusts its shape to low frequency noise since it is fitted to a single curve. Therefore, estimation performed based on such a polynomial does not apply to low frequency noise or drift. Fitting

is only applied to the near-linear range to acquire an acceptably tight fit to the bounded function using a polynomial, which is unbound, while keeping the degree reasonably low. The polynomial, the measurement data and the deviation from the fit are shown in Figure 3.14.
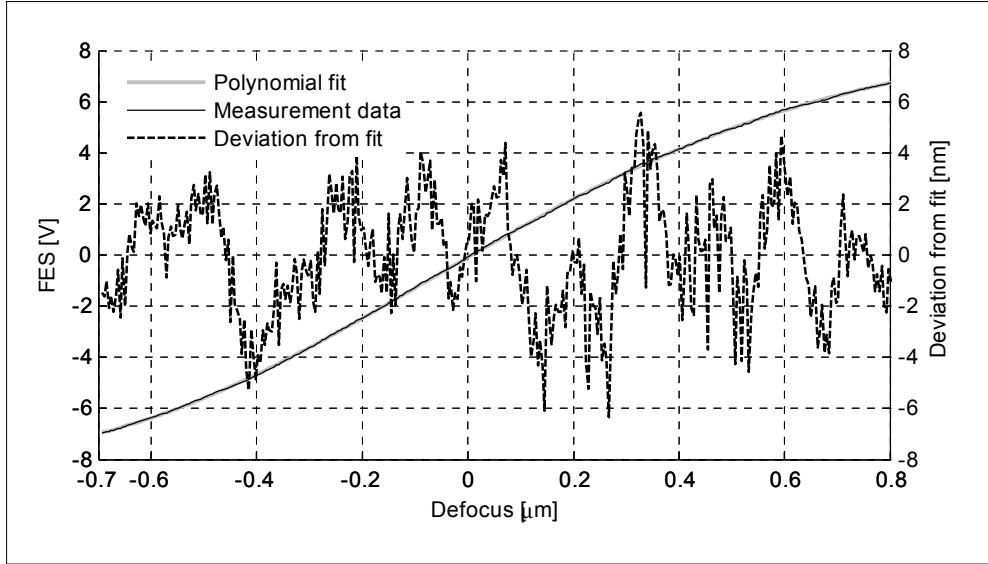
Figure 3.14: A polynomial fit to a normalized FES and the original measurement data in Volts, shown together with the deviation from the fit in nm.

The RMS of the difference between the polynomial and the measurement data has been calculated for three different measurement runs, to get an indication of measurement uncertainty. The RMS values found are 2.1, 2.1, and 2.6 nm.

3.2.8   CYCLIC DISTURBANCE

The results presented in the previous subsection are successful tests. A periodic disturbance, however, has often been observed during initial experiments. The disturbance is superimposed on the pinhole signals, FES and the normalized FES and periodic with respect to the axial position of the SUT. An extreme case is shown in Figure 3.15, left. Such a disturbance is disastrous for determining distance, since measurements become ambiguous because the disturbance is large enough to cause the slope of the FES to change sign. On top of that, the phenomenon does not reproduce over longer periods of time. Many experiments have been carried out to find the cause of the problem; only a few which lead to the solution and theories about the possible cause will be treated here.
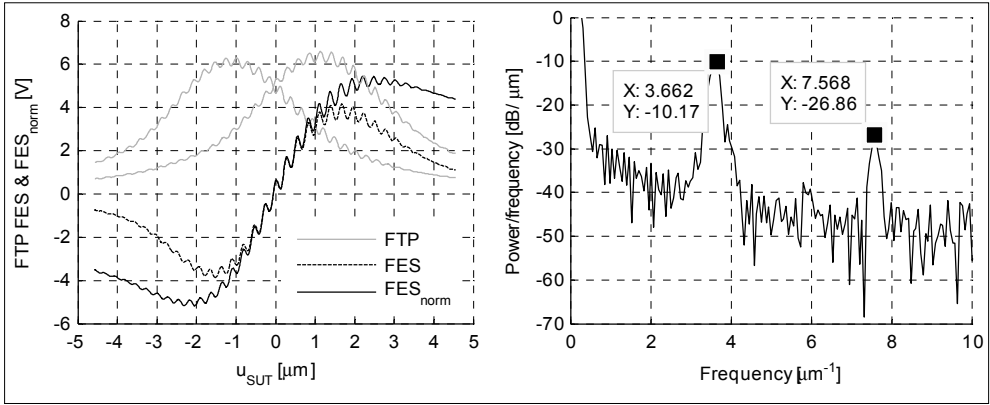
Figure 3.15: Periodic disturbance of the signals (left) and a power spectral density plot showing the periodicity (right).

From the shape of the disturbance, it is suspected that interference of the beam returning from the SUT with a ghost reflection might be the cause. In the unnormalized pinhole signals, the maximum modulation depth is just below 7 %. If both wavefronts have equal shape and uniform intensity over the entire aperture, a ghost reflection of 0.12 % is sufficient to cause this modulation depth. Two observations, however, do not seem to agree with this explanation: the period is not equal to half the wavelength and normalization does not suppress the intensity variations.

To find the periodicity of the signal, a power spectrum density plot with respect to spatial frequency in z direction has been made, as seen in Figure 3.15, right. For interference in air, a period of 316.4 nm is expected or a 158.2 nm period for double-pass interference. The two peaks, on the contrary, reveal a periodicity of 273 nm and 132 nm. Note that the periods of the two measured periodicities have a ratio close to 2:1, as expected for interference. By experimenting, both mechanics and electronics have been ruled out as likely causes, it is therefore assumed that the cause is optical interference.

If one considers a ray in the outer part of the beam, as depicted in Figure 3.16, it can be seen that its Optical Path Difference (*OPD*) is not equal to the displacement of the SUT (Δu_{sut}), as for a paraxial ray. Instead, it is equal to $\Delta u_{sut}/\cos(\alpha)$ which might cause the reduced period. Using similar triangles it can be shown that for a returning ray to have equal phase while the SUT is at a different axial position, the following criterion must be satisfied:

$$r_r = \sqrt{\left(\frac{i\,\lambda\,f_1}{2\,\Delta u_{sut}}\right)^2 - f_1^2} \qquad (3.25)$$

where:

r_r is the offset of the ray from the optical axis,

i is an integer, and

f_1 is the effective focal length of the objective lens.

Filling in the parameters of the demonstrator in (3.25) gives r_r=2.35 mm. Together with the beam diameter of 3.1 mm (3.1) yields that rays fulfilling (3.25) only have an intensity of 1 % of a paraxial ray. One expects the effective wavelength to be an average of the contributions over the entire diameter of the beam. Therefore, if all parts of the beam contribute equally to the interference phenomenon, it is impossible that this explains the entire observed periodic shift.
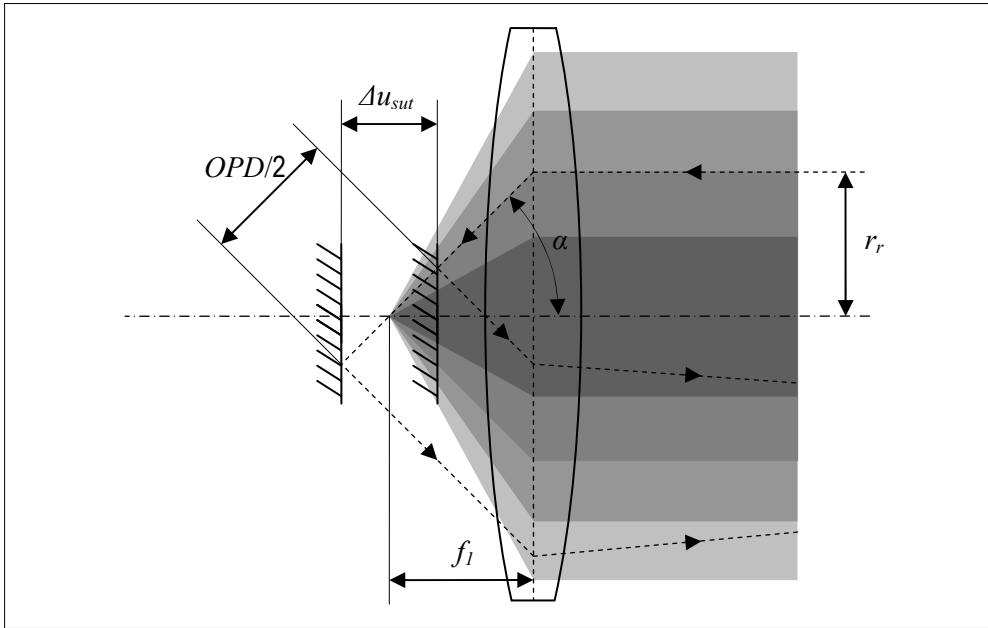
Figure 3.16: Path of an oblique ray for different positions of the SUT. It shows that the Optical Path Difference (OPD) is not equal to twice the axial displacement of the SUT, as is the case for paraxial rays.

Apart from the shortened periodicity, intensity variations should be suppressed by normalization if they occur in both detector branches. Inspection of the pinhole signals, however, reveals that their periodic disturbances are out of phase and that

their relative phase difference seems to change for various positions of the SUT. So either the pinholes experience effects of different processes, or the effects are non-uniform within the beams, and the pinholes filter out other parts of the intensity distribution as the SUT scans through focus. The latter is probably the case since the power of the beams has been measured with the pinholes removed, this showed in-phase and equal modulation in the detector branches.

To find which components contribute to the phenomenon, various experiments have been conducted to subsequently rule out components. First, the beam between the Polarizing Beam Splitter (PBS) and the Non-Polarizing Beam Splitter (NPBS) has been measured, which also contains the periodic intensity variation. Therefore, the NPBS and the detector branches are not the (only) cause of the phenomenon. Next, the laser has been slightly tilted so that the surfaces of the PBS and Quarter Wave Plate (QWP) are not perpendicular to the beam anymore, as shown in Figure 3.17, left. Because the pinholes in the detector branches are close to focus they block reflections from these surfaces, since these beams are at a different angle. In this configuration, the phenomenon is still observed, thus it is concluded that the resonance cavity which causes the problem must be constituted of the laser's exit window and the cat's eye consisting of the objective lens and SUT. Ideally, no light returns to the laser due to the QWP and the PBS. The extinct ration of the PBS is measured to be as high as 2500:1 for 633 nm. Furthermore, slight rotation of the QWP showed marginal effect on modulation depth.

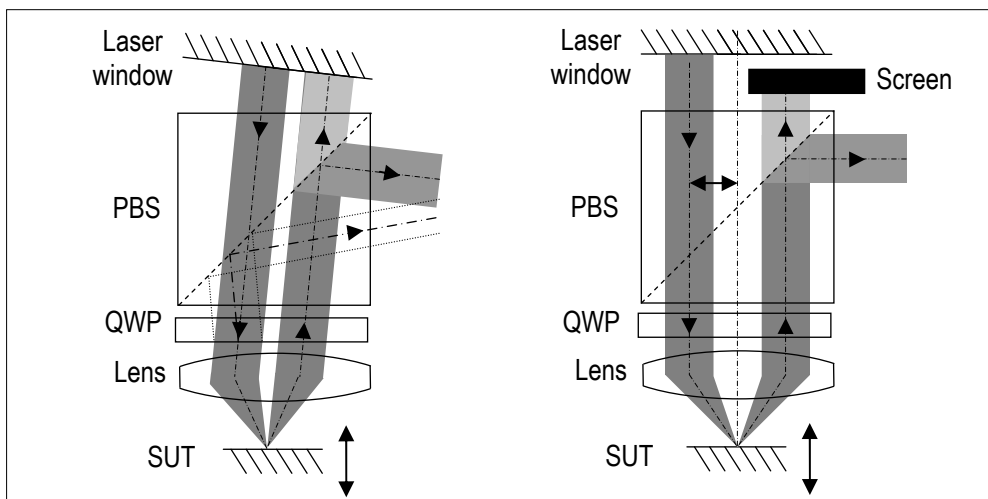

Figure 3.17: Cavity formed by the laser window and the cat's eye constituted of the objective lens and the SUT. A schematic representation of the beam paths when the laser is tilted (left). The offset is shown to distinguish between the incident and the returning beam. Proposed experiment where the laser is offset from the optical axis and the beam returning to the laser is blocked (right).

Since the laser window apparently is part of the cavity, back-coupling into the laser might be involved. Effects in the gain medium may account for the reduced period of the disturbance. HeNe lasers can operate at the 543 nm lasing band in the appropriate configuration. This gives a ½λ of 271.5 nm, remarkably close to the observed period. Single mode laser operation is obtained through proper resonator design, which aims at achieving a gain curve above the loss level at only one of the possible lasing wavelengths. Maybe, by adding an external resonator a system of two interacting cavities is created. If that is so, the periodic phenomenon might be due to the simultaneous measurement of a 633 nm and a modulating 543 nm beam. Alternatively, the 633 nm beam's power is modulated due to power drain from the gain medium by a modulated 543 nm resonance. Either way, the effect must still be non-uniform over the spots at the pinholes, because it will otherwise be canceled out by normalization. If a 543 nm resonance is indeed the cause, this also explains why the QWP and PBS do not prevent the beam from returning to the laser, since both are designed for 633 nm and do not function properly at 543 nm.

To obtain certainty whether the laser window and the cat's eye form the only cavity that plays a role in the phenomenon, it was proposed to conduct the experiment shown in Figure 3.17, right. The laser beam is given such an offset from the system's optical axis that the incident and returning beam do not overlap. This enables blocking the returning beam between the PBS and the laser by placement of a screen. Unfortunately, the clear aperture of the system is insufficient to accommodate two beams next to each other: therefore, this experiment has not been conducted.

Instead, a Faraday isolator has been inserted between the laser and PBS. This device allows light of a certain wavelength and polarization to propagate in one direction whereas in the other direction it is blocked, regardless of the polarization state. Experiments have shown that a Faraday isolator for 633 nm neutralizes the phenomenon. At that time, the 543 nm resonance had not been considered yet; otherwise, a long-pass cutoff filter may have been inserted as a test. If the 543 nm resonance indeed causes the disturbance, the effectiveness of placing the isolator can be explained if it increases the loss level sufficiently even if it does not block all of the returning 543 nm light.

Although an adequate solution is found to circumvent the problem, the cause of the periodicity shift of the interference remains elusive. This is work for further research.

3.3   COMPARISON OF MODEL TO EXPERIMENTAL DATA

Using the data obtained with the demonstrator, the validity of the analytical models can be tested against the experimental results. Therefore, the analytical and experimental results are plotted together in Figure 3.18. The signals from the demonstrator are scaled so that the maximal magnitude of the pinhole signals agrees with the maximal Fractionally Transferred Power (FTP) of the model. This is done since the power of the laser, the transmissivity of the optical train and the gain of the electronics are not exactly known. Note that small differences in the system's electrical gain and laser power have negligible influence on the quality of measurements and hence need not be included in the comparison anyway.

The measurement data shown is taken from a test in which nominal $\overline{D}_{ph}$ and $\overline{z}_{ph}$ are 2.25 and 1.3 respectively. The simulated signals are generated using the same dimensionless pinhole parameters and based on equations (3.19), (3.20) and (3.21). Other pinhole parameters show similar correspondence.
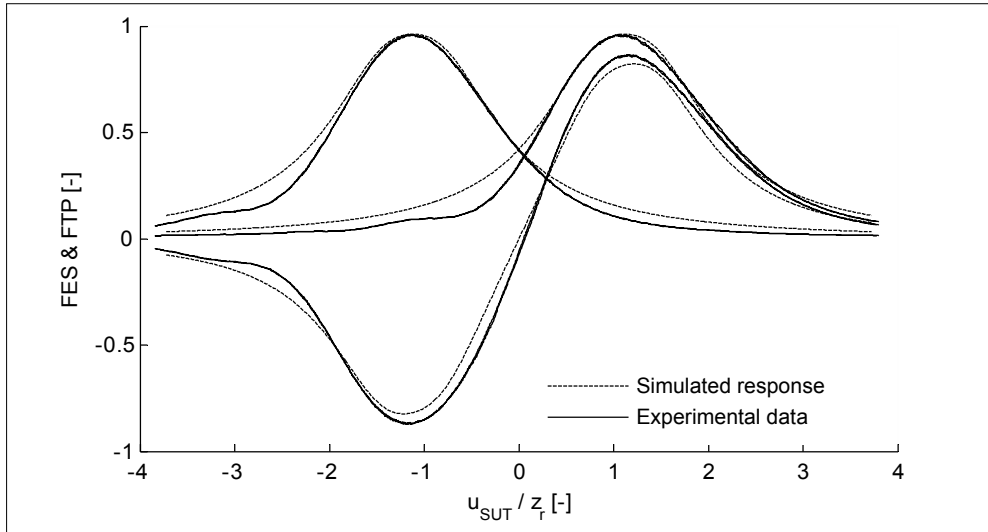
Figure 3.18: Correspondence between simulated and experimental curves for FTP and FES.

A small discrepancy between model and data can be observed. Relative to the top of the bell-shaped curves the extremities of the curves are slightly lower than predicted by the model. This effect is more pronounced on the left side of the bell-shaped curves. It is likely that this asymmetry is caused by aberrations in the system. For one, the lacking of a cover plate in front of the microscope objective is known to introduce spherical aberration to the beam. Furthermore, if the intensity

distribution of the laser beam is not exactly Gaussian, this might also introduce part of the observed deviations. Because the deviations from the test results are small and can be explained, the models are considered useful design tools.

3.4 CONCLUDING REMARKS

Models of the differential confocal system have been derived that relate the response of the system to its design parameters. Confidence in the analytical models has been developed since the data obtained through experiments and models agree.

The demonstrator has proved that with the differential confocal method, the resolution requirement can be amply exceeded within a measurement range sufficient for tracking surfaces with a heavily freeform shape. Furthermore, the demonstrator also has provided valuable experience with differential confocal measurement from which guidelines for design of the prototype can be deducted.

To prevent the occurrence of cyclic disturbances such as discussed in Subsection 3.2.8, a Faraday isolator will be placed between the laser and the rest of the optical system. Moreover, where practically possible, partially reflecting optical surfaces perpendicular to the laser beam should be avoided to avert the occurrence of other resonance cavities. This can be achieved in various ways: different optical components can be cemented together with index-matched adhesives and wedges can be applied to entrance or exit surfaces.

It is believed that combined with the improvements proposed in Section 2.1, the requirements for the non-contact sensor of NANOMEFOS can be met with the differential confocal measurement principle. Aiming at this goal, the analytical models presented in this chapter are considered valuable design tools, since they can be used to optimize the differential confocal system.

4 DIFFERENTIAL CONFOCAL PROPERTY ANALYSIS & OPTIMIZATION

Based on the analytical models presented in the previous chapter, the influence of the design parameters on the system properties is investigated. Using the results from this analysis, the differential confocal system is optimized for application in NANOMEFOS.

As mentioned in the previous chapter, the properties of the differential confocal system must be chosen such that they match the requirements for the NANOMEFOS project. Several of the properties that characterize performance and robustness of distance sensors were also mentioned, namely:

- resolution,
- measurement range,
- measurement uncertainty,
- dynamic range, and
- tilt dependent error.

Some of these properties are conflicting, such as measurement uncertainty and measurement range. Therefore, to arrive at an optimal sensor design, the parameters are chosen so that the tradeoff between properties is optimal for the application. To make this tradeoff, an analysis of the variation of these properties within the design space is desired.

Properties such as measurement range and inherent tilt dependent error can be numerically calculated using the models presented in the previous chapter. Measurement uncertainty, resolution and dynamic range, on the other hand, cannot

be calculated numerically because they do not only depend on the design parameters, but on many other component properties and environmental conditions as well, some of which are unknown. If these properties are not included in optimization, however, the design parameters found might be unfavorable regarding these properties. To avoid this, other system properties are defined that are proportional to the properties that cannot be quantitatively evaluated.

4.1 INVESTIGATED PROPERTIES

The properties that will be investigated are:
- sensitivity,
- near-linear range,
- near-linear optical signal range,
- tilt dependent error of the system,
- measurement noise, and
- electrical dynamic range.

The first four properties are purely optical in nature, whereas the latter two also account for properties of the optoelectronics. Which properties drive the design process depends on the application, and not all will be used for optimization of the prototype. Properties that are not critical in this application are treated nevertheless, because they can be of interest in applications with different requirements.

In this section, definitions of the investigated properties and their relation to the optimization process are treated; in sections 4.2 and 4.3, their variation within the design space is presented.

Sensitivity

Sensitivity is an often-specified characteristic of sensors; it describes how much the signal changes for a certain change of the measured quantity. For linear systems, the sensitivity is regularly used to convert the measurement signal to a measurement value; whether sensitivity is high or low is usually not of primary interest to the user. The resolution and measurement uncertainty, however, are of interest to the user, and usually both improve when sensitivity is increased.

To investigate the influence of dimensionless pinhole parameters on sensitivity, the dimensionless sensitivity is evaluated; it is defined as the slope of the dimensionless Focus Error Signal (FES) at zero-crossing. An arbitrary FES with the tangent line at zero-crossing is shown in Figure 4.1, left.
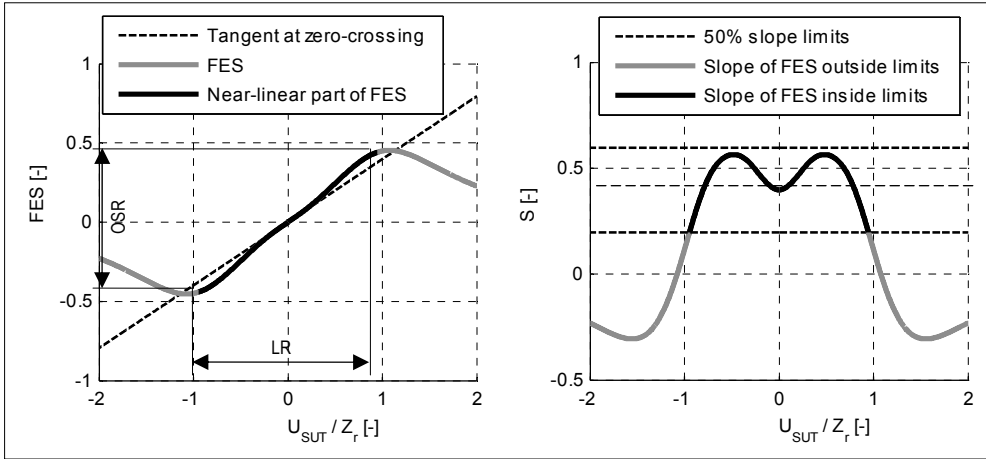
Figure 4.1: Arbitrary dimensionless FES with the tangent at zero-crossing, near-Linear Range (LR) and near-linear Optical Signal Range (OSR) (left) and its derivative with 50 % limits on slope deviation (right).

Given a certain *dimensionless* sensitivity, the *dimensional* sensitivity can be tuned by varying the Rayleigh range at the SUT. High sensitivity of the *unnormalized* FES leads to low measurement errors caused by an optical disturbance of a given power. High *normalized* sensitivity leads to high resolution at a given electrical noise level.

Dimensionless near-Linear Range

The dimensionless near-Linear Range (LR) is defined as the range of dimensionless defocus over which the local sensitivity of the FES stays within a certain percentage of the sensitivity at zero-crossing. This criterion is selected because of the influence of local sensitivity on resolution and sources of measurement uncertainty. The allowed deviation of local slope is chosen to be a 50% relative to the slope at zero-crossing. This is illustrated for an arbitrary FES in Figure 4.1 where the FES is shown on the left, and the local sensitivity on the right. The FES is said to be near-linear over the range in which the local sensitivity stays within 50% of the sensitivity at zero-crossing; the 50% slope limits are plotted as well.

Both unnormalized and normalized LR are of interest. A large *unnormalized* LR implies that for a large defocus, the *measurement error* caused by optical disturbances of a given power stays approximately equal. Whereas a large *normalized* LR implies that for a large defocus, the *resolution* stays approximately equal given a certain electrical noise level.

Because slope is used as criterion for linearity and not the absolute error between the FES and a tangent line at zero-crossing, the nonlinearity must be compensated for by calibration. Since such a calibration is necessary to achieve the required measurement uncertainty anyway, this involves no extra effort.

Depending on pinhole parameters, the sensitivity at zero-crossing can be either a local minimum (as in Figure 4.1) or maximum. The latter is desirable as this provides maximum resolution for measurement of slight freeform and rotation symmetric optics. At the outer ends of the LR, however, the resolution and measurement uncertainty deteriorates by about 50%. This is acceptable because that part of the FES is only used if the tracking error becomes relatively large, as is to be expected when measuring heavily freeform surfaces. For such freeforms the 2σ measurement uncertainty requirement is eased to 35 nm from the 10 nm for rotationally symmetric optics.

The LR of the unnormalized and normalized FES are not among the driving optimization criteria for this application. The near-linear optical signal range, however, is a driving optimization criterion and is directly related to the LR.

Dimensionless Optical Signal Range

The dimensionless Optical Signal Range (OSR) is a measure for the fraction of light that arrives at the pinholes and contributes to measurements within the near-linear range. As such, it is a measure of the useful fraction of the available physical quantity that is measured, i.e. the power of the light at the pinholes. This is important since for equal relative fluctuations in the intensity of the light, higher OSR leads to a higher ratio of measurement range to uncertainty. A large measurement range is important because it allows for large tracking errors, and hence improves robustness. Low measurement uncertainty is needed to achieve the 10 nm measurement uncertainty required for NANOMEFOS. A long Rayleigh range leads to a long measurement range but a high measurement uncertainty and vice versa. Thus, these two requirements are contradictory in this respect and a tradeoff has to be made. A high OSR leads to a high ratio of measurement range to measurement uncertainty, therefore OSR is a dominant design criterion. OSR is indicated in Figure 4.1, left.

Tilt Dependent Error

One of the most demanding requirements is the 35 nm uncertainty at tilts up to 5°. In Subsection 2.1.3 the addition of a Position Sensing Detector (PSD) is proposed to allow for aperture correction. In addition to calibration, it is desirable to keep the Tilt Dependent Error (TDE) of the uncorrected sensor to a minimum in the first place. The TDE of the differential confocal system can be thought of as consisting of two components caused by fundamentally different mechanisms. One is the

error resulting from aberrations in the optical system. This mechanism is expected to result in both shift of the zero-crossing and slope change of the FES. The aberrations can be split up in those inherent to the design of the optics in question, those caused by limited manufacturing precision of the optics and those due to misalignment. Suitable optics are available that according to specification are diffraction limited over the entire clear aperture. Hence, these errors are expected to be reduced to within the required 35 nm uncertainty zone using aperture correction.

The other mechanism causing tilt dependency of the system is inherent to the measurement method, irrespective of the quality of the optics used. This inherent error is induced by beam shift relative to the pinholes that occurs when the SUT is tilted. This particular mechanism causes slope change of the FES but no shift of the zero-crossing. The magnitude of this error is numerically calculated for different combinations of dimensionless pinhole parameters using a dimensional model. Because of the required 35 nm measurement uncertainty for 5° tilt, the TDE is one of the driving optimization criteria.

Measurement noise

The measurement noise of the system, together with the sensitivity of the normalized FES, gives an indication which resolution can be achieved with the system. Since most of the noise sources only appear in the dimensional model, noise is evaluated in dimensional space. Only the normalized FES is of interest when studying noise, as the real measurements are taken from the normalized signal. Noise in itself is not among the optimization criteria; instead, it is included in the calculation of electrical dynamic range.

Electrical dynamic range

Dynamic range is an important figure of merit for many measurement systems since it is defined as the ratio of the measurement range to the resolution. Because the real resolution depends on some effects that are hard to predict, here, only the Electrical Dynamic Range (EDR) of the system is studied. It is defined as the ratio of the LR to the resolution of the system, where resolution is estimated by dividing electrical noise by sensitivity.

Although EDR itself is a dimensionless property by definition, it is evaluated in dimensional space as it depends on measurement noise. For the same reason only the normalized FES is of interest when studying EDR. It is one of the more important optimization criteria for many other systems, however, for this system, it is not one of the driving criteria for optimization. Nevertheless, it must be high enough to achieve the required resolution in the area chosen based on the dominant criteria.

Overview of properties

The diagram in Figure 4.2 gives an overview of which properties relate to the *normalized*, and which relate to the *unnormalized* FES and whether they are calculated using a *dimensional* or *dimensionless* model.

	Unnormalized FES	Normalized FES
Dimensional		TDE Noise EDR
Dimensionless	Sensitivity LR OSR	Sensitivity LR

Figure 4.2: Diagram showing whether properties relate to the unnormalized or the normalized FES and if they are evaluated using the dimensional or dimensionless model. The properties are Tilt Dependent Error (TDE), measurement noise, Electrical Dynamic Range (EDR), sensitivity, near-Linear Range (LR) and Optical Signal Range (OSR).

Design parameters

The relationship between the aforementioned properties and the parameters in the exponential terms of equation (3.19) and (3.20) are investigated. These parameters are:

- dimensionless pinhole diameter,

- dimensionless pinhole offset,

- wavelength of the laser,

- diameter of the collimated beam, and

- focal length of the objective lens.

The diameter of the collimated beam and focal length of the objective lens (f_l) will be investigated simultaneously, in the form of Numerical Aperture (NA_b) of the beam. The relationship of f_l and D_B to the system's NA is given by:

$$NA_b = \sin\left(\tan^{-1}\left(\frac{D_B}{2 \cdot f_1}\right)\right)$$

(4.1)

4.2 OPTICAL PROPERTY ANALYSIS

4.2.1 SENSITIVITY

Dimensionless sensitivity of the normalized and unnormalized FES as a function of pinhole diameter and offset can be derived analytically from (3.11) and (3.15) by partial differentiation. Alternatively, it can be approximated numerically using the finite forward difference method. Here, the latter method has been chosen because of the simple implementation, while an analytical solution is used to check the results.

Figure 4.3 and Figure 4.4 show the dimensionless sensitivity at zero-crossing as a function of dimensionless pinhole diameter and offset for the unnormalized FES and for the normalized FES. It can be seen that for the unnormalized FES the sensitivity is at maximum when dimensionless pinhole diameter and offset are equal to 1. By manipulation and partial differentiation of (3.11), the dimensionless pinhole diameter and offset yielding maximum dimensionless sensitivity for the FES are found to be equal to 1 (see Appendix A), thus supporting the numerical results.
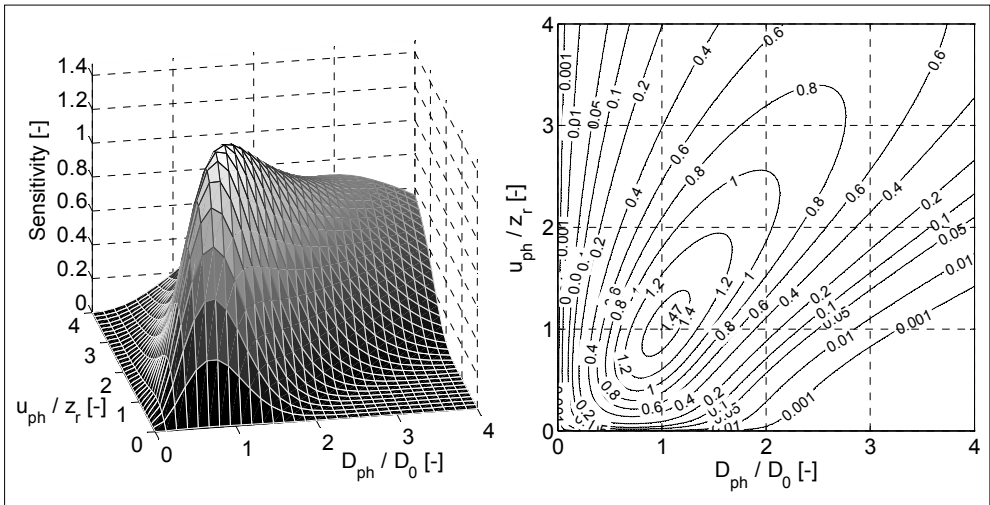
Figure 4.3: 3-D graph and contour plot of the sensitivity of the unnormalized dimensionless FES as a function of pinhole parameters.
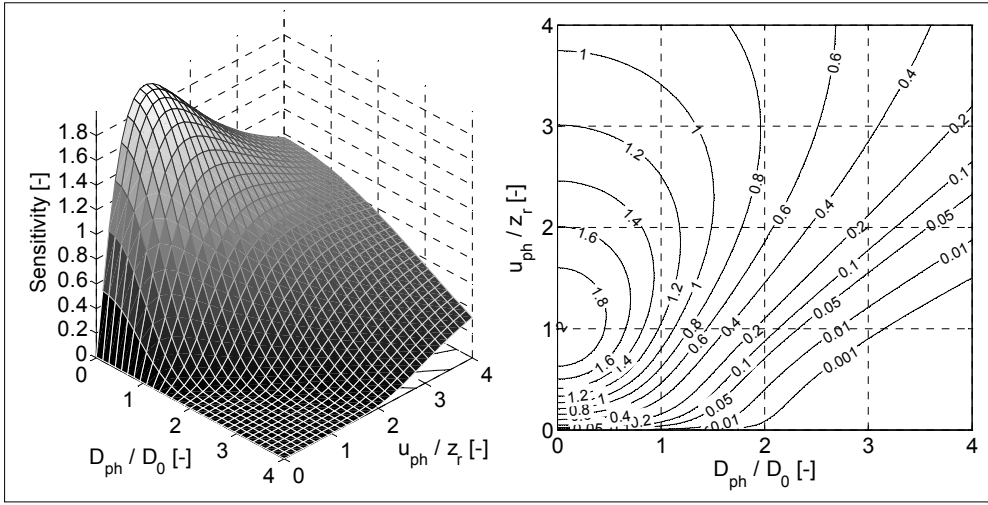
Figure 4.4: 3-D graph and contour plot of the sensitivity of the normalized dimensionless FES as a function of pinhole parameters.

Since sensitivity is evaluated in *dimensionless* space, these results are insensitive to changes of other system parameters. The *dimensional* sensitivity of the unnormalized and normalized FES are obtained using the following relationships:

$$S = \frac{G_{el}\, R_{sut}\, R_{pd}\, \eta_{opt}\, P_l}{z_r}\, \overline{S} \tag{4.2}$$

$$\mathbb{S} = \frac{\mathbb{G}_{el}}{z_r}\, \overline{\mathbb{S}} \tag{4.3}$$

where:

S	is the sensitivity for the FES,
$\mathbb{S}$	is the sensitivity for the normalized FES,
$\overline{S}$	is the dimensionless sensitivity for the FES, and
$\overline{\mathbb{S}}$	is the dimensionless sensitivity for the normalized FES.

Thus, for assessing the influence of NA and laser wavelength on sensitivity, only their influence on Rayleigh range matters. From Gaussian beam theory (O' Shea, 1985, pp 231) it can be derived that z_r depends on wavelength and NA_b as:

$$z_r = \frac{\lambda}{\pi\left(\sin^{-1}\left(NA_b\right)\right)^2} \tag{4.4}$$

Figure 4.5 shows a 3-D graph and contour plot of reciprocal Rayleigh range as a function of wavelength and NA. Only wavelengths within the visible spectrum are investigated since use of a visible light source eases alignment during assembly.
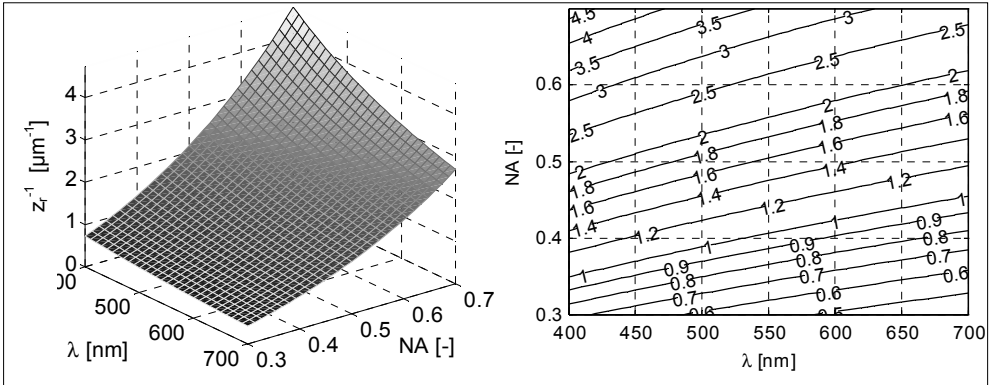
Figure 4.5: Dependency of reciprocal Rayleigh range [μm^{-1}] on beam properties.

From these figures, it can be seen that within a practical range of design parameters, NA has more influence on sensitivity than wavelength does.

4.2.2 NEAR-LINEAR RANGE

To determine the length of the near-Linear Range (LR) numerically, the local slope is approximated using the finite forward difference method for increasing defocus. Because the FES is an odd function, slope only needs to be evaluated for positive values of defocus. The LR for the unnormalized FES and for the normalized FES are depicted in Figure 4.6 and Figure 4.7.
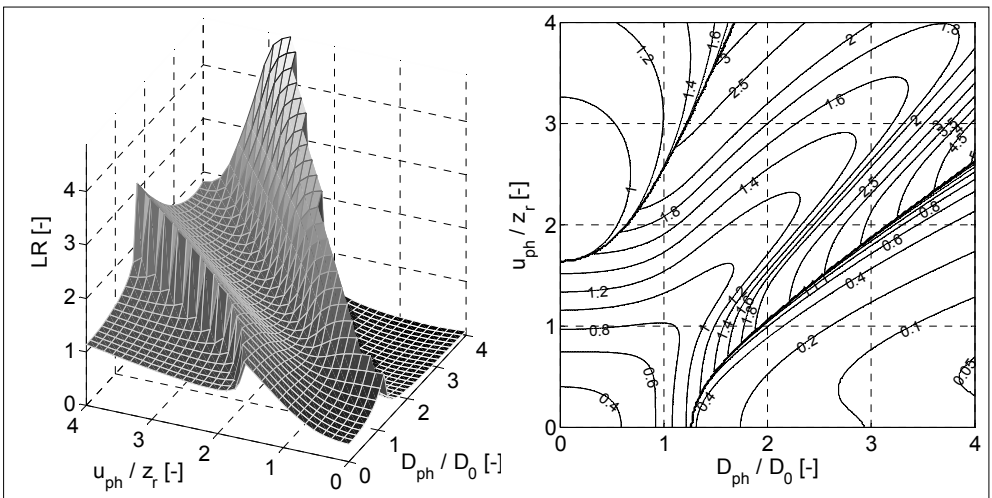
Figure 4.6: Plots of near-Linear Range (LR) for unnormalized dimensionless FES as a function of pinhole parameters.
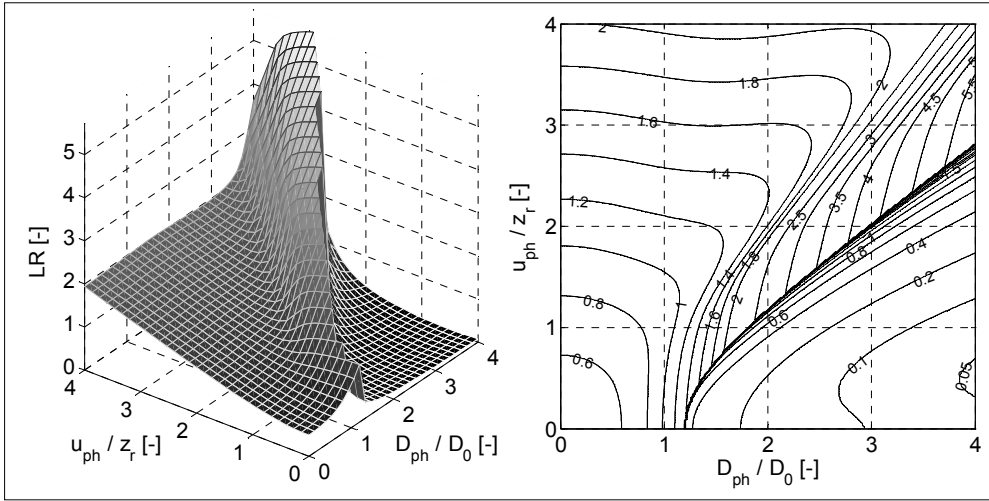
Figure 4.7: Plots of near-Linear Range (LR) of normalized dimensionless FES.

The LR, pinhole diameter and offset are evaluated in dimensionless space and are thus insensitive to changes in other design parameters. From dimensionless LR the dimensional LR can be obtained by multiplying with z_r. Therefore, it suffices to evaluate the effect of wavelength and NA on Rayleigh range. Using (4.4) the plots of Figure 4.8 are found showing the Rayleigh range versus wavelength and NA.
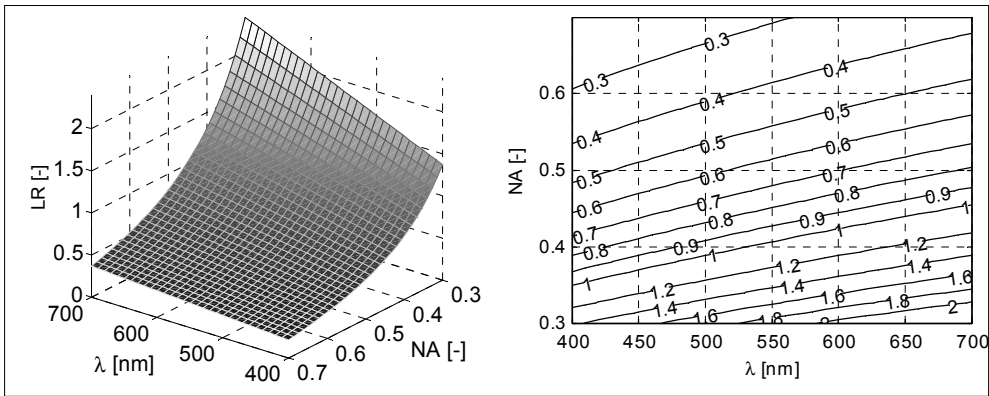
Figure 4.8: Dependency of Rayleigh range [μm] on beam properties.

4.2.3   DIMENSIONLESS OPTICAL SIGNAL RANGE

OSR is determined by evaluating the values of the FES at the extremes of the LR. It depends on dimensionless pinhole diameter and offset only. Figure 4.9 shows the OSR as function of dimensionless pinhole parameters.
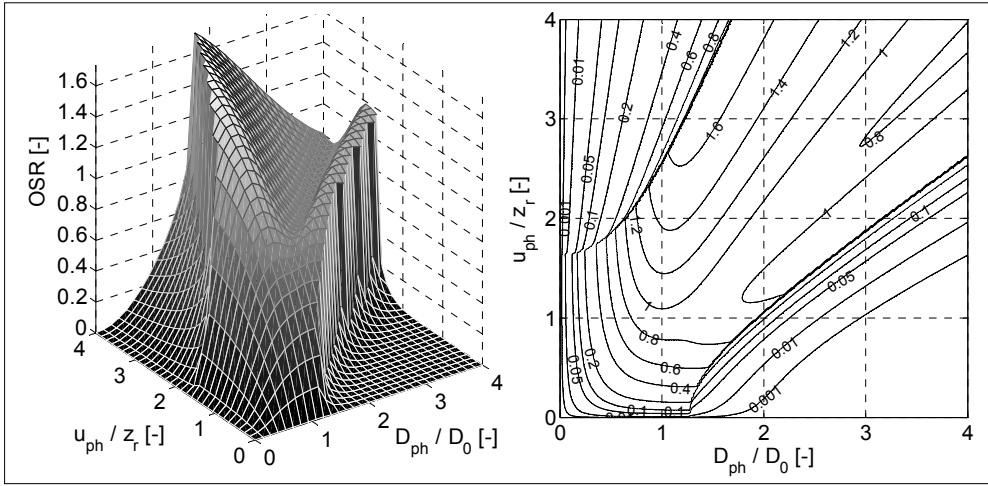
Figure 4.9: 3-D graph and contour plot of Optical Signal Range (OSR) as a function of dimensionless pinhole parameters.

4.2.4 TILT DEPENDENT ERROR

Tilt Dependent Error (TDE) is caused by a beam shift at the pinholes, which leads to a different Fractionally Transferred Power (FTP) for measurements at 5° than for measurements at 0°. To determine this shift, a ray-trace of the central ray of the returning beam is made, as schematically depicted in Figure 4.10.
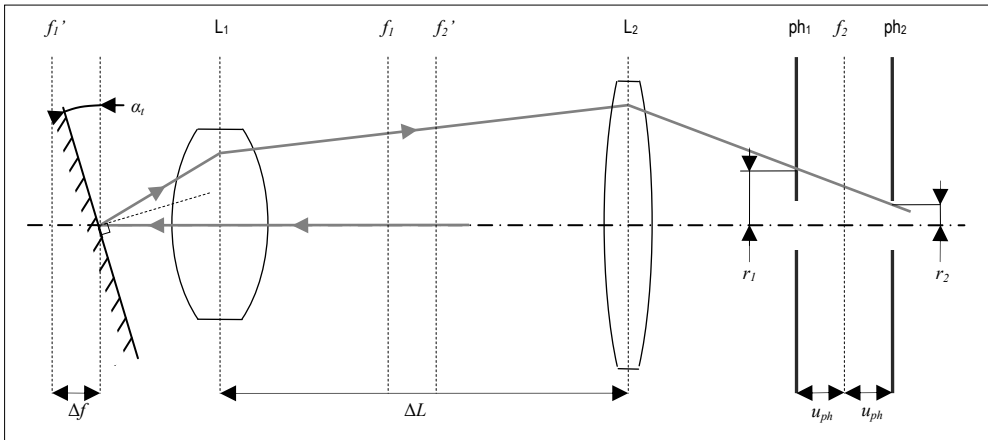
Figure 4.10: Ray-trace of the ray with maximum intensity for a tilted SUT.

Deduction using goniometric relationships yields the following equation for beam shift at the pinholes as a function of SUT tilt, defocus and system parameters:

$$r_1 = \left(\frac{f_2 \, \Delta f}{f_1} + \frac{\left(\frac{f_2 \, \Delta f}{f_1} - f_1 - \frac{(\Delta L - f_2) \, \Delta f}{f_1} \right) u_{ph}}{f_2} \right) \tan(2\alpha_t) \tag{4.5}$$

Thus pinhole parameters, beam properties, focal lengths of objective and image lenses and the distance between them, all influence tilt dependency. Therefore, tilt dependency can be minimized by proper design choices. The criterion used to evaluate TDE will be the uncorrected maximum measurement error within the LR caused by the inherent tilt dependency of the configuration.

The previous models and simulations are based on the expression for encircled power, (3.3), which is only valid for circles concentric with the Gaussian intensity profile; this is not the case when the SUT is at an angle. Hence, the amount of light passing through the pinholes is solved by numerical integration of the Gaussian beam profile.

To find the FTP for a tilted SUT, two square matrices are generated in which each entry represents the intensity of a Gaussian profile at that entry's position, so the matrix entries can be considered pixels located at the pinholes. Likewise, a mask is generated as a matrix containing a circle within which all entries are equal to one, whereas outside this circle all entries are zero's. To yield the FTP per pinhole, the mask is multiplied entry-wise with the matrix representing the Gaussian profile and subsequently all entries of the resulting matrix are summed. The resulting scalar is proportional to the energy transferred through the pinholes and is then divided by the sum of the entries of the matrix representing the unmasked Gaussian beam, yielding the FTP per pinhole.

To reach satisfactory precision a large number of pixels is required. Furthermore, to acquire the maximum TDE within the LR for each system configuration, these steps have to be repeated for a large number of increments within the LR for each set of pinhole parameters.

Because this leads to unpractical calculation times, an alternative method is desired. FTP values can be computed for varying beam shift and relative pinhole size, which are normalized by dividing them by the local e^{-2} boundary of the beam. In this way, a two-dimensional lookup table is generated containing FTP values for various combinations of relative beam shift and relative pinhole size. Consequently it is not necessary anymore to perform numeric integration for each combination of pinhole offset, pinhole diameter and increment of defocus. Instead, only the relative beam offset and relative pinhole size are calculated, after which FTP is approximated by cubic interpolation of the values in the lookup table.

TDE is evaluated in *dimensional* space, its shape as a function of *dimensionless* pinhole parameters, however, remains unchanged. A typical system is chosen for simulation and conclusions about pinhole parameters yielding minimal TDE are generally applicable. A three-dimensional graph and contour plot of TDE for the normalized dimensionless FES of a typical system are shown in Figure 4.11 as a function of dimensionless pinhole parameters. In the contour plot, a line along which the TDE is close to zero is plotted in gray; points along this line are favorable working points. For large dimensionless pinhole diameter combined with small dimensionless pinhole offset, the results are inaccurate due calculation with small numbers. Therefore, that part of the graph is omitted.
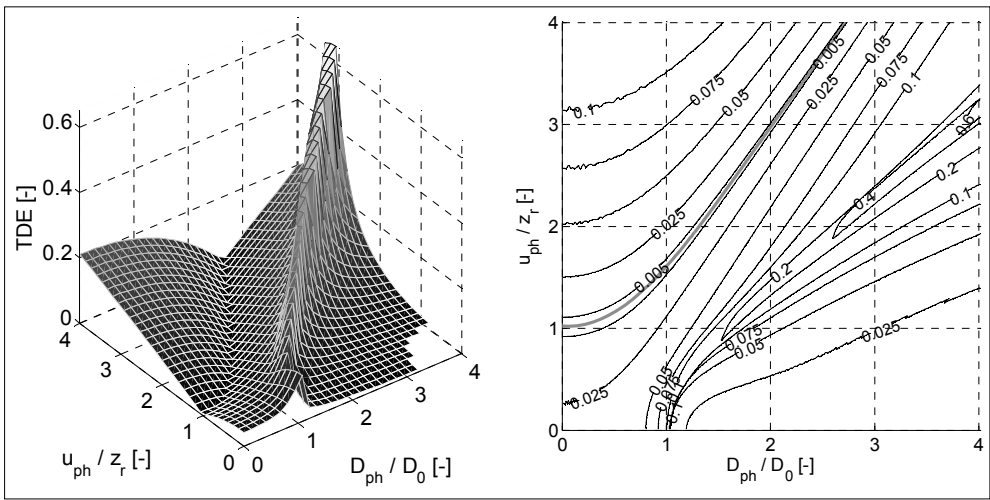
Figure 4.11: Three-dimensional graph (left) and contour plot (right) of Tilt Dependent Error (TDE) for varying dimensionless pinhole parameters with the values of TDE normalized by z_r. The line of minimal TDE is plotted in gray in the contour plot.

4.3   OPTOELECTRONIC PROPERTY ANALYSIS

In this section, an analysis of measurement noise and dynamic range for the differential confocal system is presented. Both are dependent on measurement bandwidth and the amount of light available. Fortunately, in this application the amount of light can be regulated, hence noise can be limited with little repercussions. It will be shown that even if the measurement bandwidth is chosen at 150 kHz, sub-nanometer noise and high dynamic range can be achieved.

Note that the values for noise and dynamic range calculated in this section are an electrical nominal equivalent and might not directly determine the system's measurement resolution. Distortions in the Gaussian intensity profile of the beam

can lead to local non-linearity in the LR of the FES, thereby degrading local resolution. Nevertheless, equivalent electrical dynamic range is important to consider, since the higher it is, the better the electronics can resolve small displacements of the SUT, even in parts with reduced sensitivity.

4.3.1 MEASUREMENT NOISE

As explained in Section 2.2, normalization is applied to the FES to suppress noise. Consequently, only the normalized FES is of interest here. Some sources may be identified by examining equation (B.8) from which can be concluded that the system is susceptible to photodiode noise and noise introduced by the electrical system. Other terms in (B.8), such as focal lengths, wavelength and beam width, have little influence, or are expected to show some drift over time, but negligible noise, due to the nature of the physical mechanisms they are related with.

In this subsection, the following sources are treated: photodiode noise, noise generated by the processing electronics, noise of the ADC, fluctuations in the intensity distribution of the laser, mechanical vibrations, ElectroMagnetic Interference (EMI) and electronic hum. Only the first three sources are fully random processes, the others are at least partly pseudo random or even periodic in nature. Nevertheless, they are difficult or impossible to distinguish from the measurement signal and all are therefore regarded as noise.

Fluctuation in the intensity distribution

The influence of total intensity fluctuations of the laser is suppressed by normalization. On the contrary, measurement noise might be introduced by changes of the intensity distribution within the laser beam. Predictive models about noise introduced by lasers in general are hard to make since noise levels differ considerably for different laser types and models and depend on the control electronics as well. Fortunately, the laser beam is symmetric around focus so to a large extend these fluctuations cancel out. The effectiveness of this mechanism, however, decreases with simultaneously increased defocus and surface tilt. Hence, a reasonable margin will be preserved to account for a possible contribution of this noise source and care must be taken to select a laser with a stable intensity distribution.

Mechanical vibrations

From examination of the power density spectrum of the demonstrator, it is found that mechanical vibrations of system components do not significant contribution to the noise level. Light and stiff design ensures that this is the case for the prototype as well. For that reason, the contribution of these sources is estimated negligible.

Photodiode noise

Three sources govern photodiode noise: 1/f noise, Johnson noise and shot noise (Melles Griot, 2008). The 1/f noise is only of significance for low frequency measurements since, as the name suggests, its power spectrum density decreases for increasing frequency. The other two sources produce broadband white noise and thus become dominant in high frequency systems.

Thermal noise (Schottky, 1918; Johnson, 1928; Nyquist, 1928) is caused by thermally induced statistical fluctuation in charge carriers, in the case of photodiodes this is dominated by thermal electron-hole pair generation. This, however, is negligible relative to shot noise for as long as the photocurrents at the detectors are relatively large, which will be so over the entire LR if the laser power is chosen high enough.

Therefore the only significant source of photodiode noise is the shot noise (Schottky 1918), an effect caused by quantum statistics. Because the photocurrents consist of large amounts of electrons with finite charge, the number of quanta contributing to a given measurement interval will be subject to random statistical variations. The RMS values of such fluctuations are described by:

$$I_{ns} = \sqrt{2\, q\, I_P\, f_{bw}} \tag{4.6}$$

where:

q is the electron charge,

I_P is the photocurrent, and

f_{bw} is the bandwidth of the measurement system.

Following IEEE convention (IRE Symbols Committee et al.1964), instantaneous values are indicated using lower case with upper case subscript (v_{FES}), incremental small signal quantities by all lower case (i_{ns}), quiescent large signals by all upper case (I_P) and RMS values by upper case with lower case subscript (I_{ns}).

The system's electronics have a small-signal bandwidth of 150 kHz, hence the relationship between shot noise and photocurrent becomes:

$$I_{ns} = 2.19 \cdot 10^{-7} \sqrt{I_P}$$

To determine the contribution of shot noise to total signal noise, normalization must be taken into account. In order to do so (B.8) can be rewritten as:

$$v_{\text{FES}} = \mathscr{G}_{\text{el}} \frac{\left(I_{\text{P1}} + i_{\text{ns1}}\right) - \left(I_{\text{P2}} + i_{\text{ns2}}\right)}{\left(I_{\text{P1}} + i_{\text{ns1}}\right) + \left(I_{\text{P2}} + i_{\text{ns2}}\right)} \tag{4.7}$$

In focus, I_{P1} and I_{P2} are equal and cancel out in the numerator, I_{ns1} and I_{ns2} are equal and the values of i_{ns1} and i_{ns2} in the denominator are negligible due to the large sum of I_{P1} and I_{P2}. Taking the fractional term, substituting I_{ns1} and I_{ns2} for i_{ns1} and i_{ns2}, quadratic addition and manipulation yields the RMS shot noise in equilibrium:

$$V_{\text{ns eq}} = G_{\text{el}} \sqrt{\frac{2 q f_{bw}}{I_{\text{P1}} + I_{\text{P2}}}} \tag{4.8}$$

This approximation is valid at zero-crossing when both photocurrents are equally large. The RMS shot noise remainder in the normalized signal relative to full signal range as a function of photocurrents is shown in Figure 4.12, left.
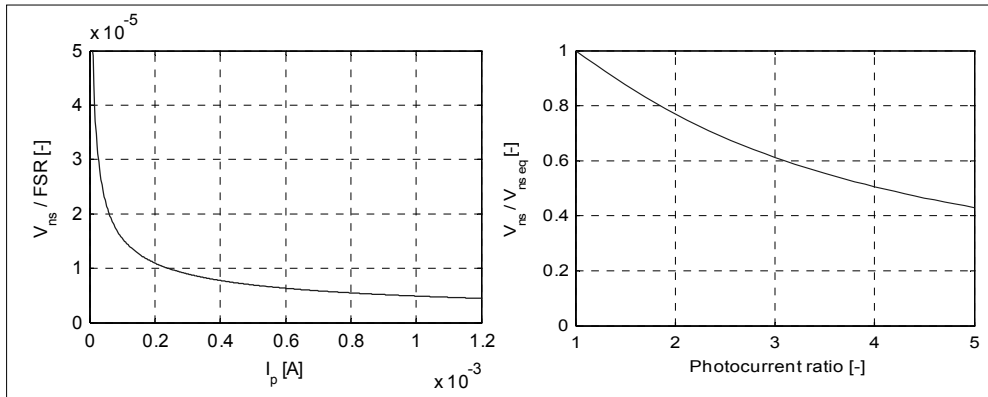
Figure 4.12: Influence of photocurrent on shot noise in normalized FES (left) and correction factor for shot noise at varying photocurrent ratios (right).

Figure 4.12, right, shows the relationship between normalized shot noise at zero-crossing and different parts of the LR, as found through a Monte Carlo simulation. For departure from focus within the LR, one photocurrent increases whereas the other photocurrent decreases. This difference in photocurrents can be expressed as the photocurrent ratio. First, the RMS noise value for the smaller photocurrent is read from Figure 3.10, left. Then, using the photocurrent ratio, the correction factor for the normalized shot noise can be determined from Figure 4.12, right.

The photocurrent at zero-crossing is dependent on the laser power, reflectivity of the SUT, responsivity of the photodiodes, transmissivity of the optical train and the dimensionless pinhole parameters. This relationship can be expressed as:

$$I_{\mathrm{P}} = \frac{1}{2} P_L \ R_{\mathrm{sut}} \ R_{\mathrm{pd}} \ \eta_{\mathrm{opt}} \left(1 - e^{-2\left(\frac{\overline{D}_{\mathrm{ph}}^{\,2}}{1 + (\overline{z}_{\mathrm{ph}})^2} \right)} \right) \tag{4.9}$$

where the expression between brackets represents the Fractionally Transferred Power (FTP). In spite of the benefit of high photocurrents, the laser power is limited to 10 mW for safety reasons. The performance critical situation with 3.5 % reflectivity of the SUT and a calculated transmissivity of 0.47, leads to a dependence on dimensionless pinhole parameters as shown in Figure 4.13, left.
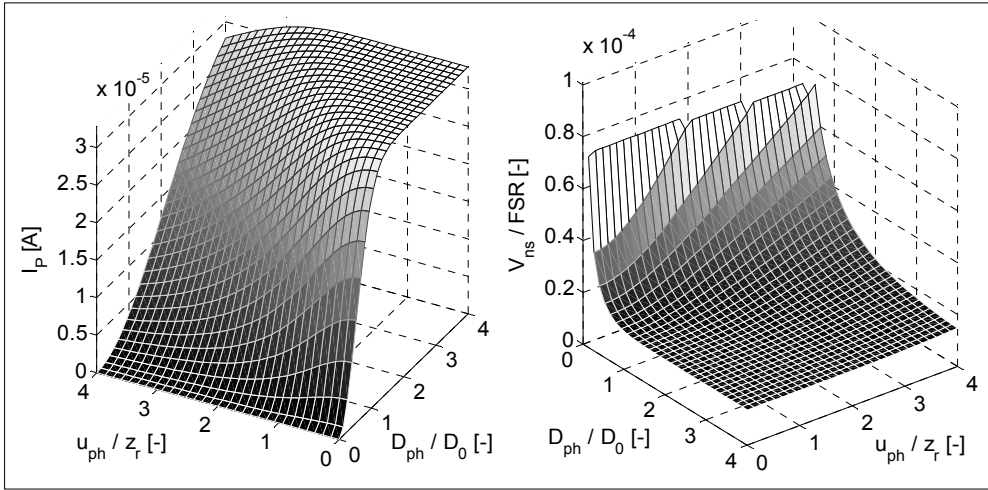
Figure 4.13: Single diode photocurrent in ampere (left) and normalized photodiode shot noise (right) expressed as part of Full Signal Range (FSR) for varying pinhole parameters.

The shot noise in the normalized FES as a function of dimensionless pinhole parameters is calculated from the photocurrents and (4.8), see Figure 4.13, right.

EMI and electric hum

During building of the demonstrator care is taken to minimize ElectroMagnetic Interference (EMI) and electric hum by twisting cables, reducing cable length between detector and electronics, shielding of detectors, cables and electronics, a stable power supply, avoiding ground loops and using differential inputs. As treated in Subsection 3.2.7 this proved successful since the demonstrator signals show negligible electric hum and EMI. The same electronics and measures are applied in the prototype as in the demonstrator. Thus, it is assumed these results can be repeated and no extra margin is preserved to account for possible contributions of EMI and electric hum to measurement noise.

Noise of normalization electronics

The noise level of the electronics used for the normalization and amplification are specified for various summed input currents and bandwidths. Because of the dependency on input current, like photodiode noise, this noise source is dependent on dimensionless pinhole parameters. In Figure 4.14 left, it is depicted relative to full signal range. Values between specified photocurrents are obtained by cubic interpolation.
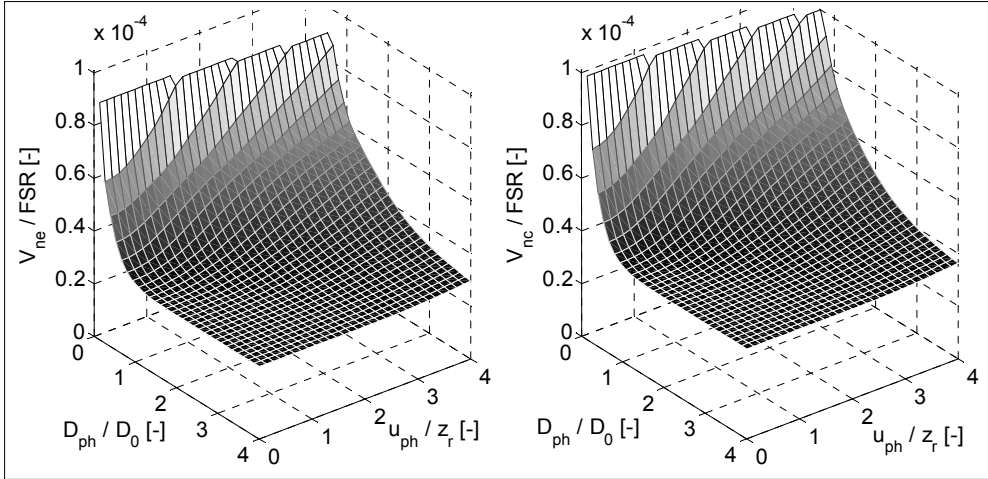
Figure 4.14: Noise introduced by the normalization electronics (left) and the combined electronic noise of the system expressed as part of Full Signal Range (FSR) for varying pinhole parameters (right).

ADC noise

Even for a good 16 bit DAQ, the ADC's noise level, V_{na}, constitutes a significant part of the combined noise level of the system. The ADC noise is considered constant at zero-crossing and independent of photocurrent. Here a 1 LSB RMS noise level including quantization noise at 16 bits is assumed, which works out to be $15 \cdot 10^{-6}$ relative to the full signal range.

Combined electrical noise

Since all electrical noise sources of the system predominantly produce uncorrelated white noise the RMS value for the combined system, V_{nc}, can be calculated by quadratic addition as:

$$V_{nc} = \sqrt{V_{ns}{}^2 + V_{ne}{}^2 + V_{na}{}^2}$$

(4.10)

The combined electrical noise of the system as a function of dimensionless pinhole parameters is shown in Figure 4.14, right.

4.3.2 ELECTRICAL DYNAMIC RANGE

The Electrical Dynamic Range (EDR) is defined here as the ratio of the LR to the resolution of the system, where resolution is estimated by dividing the combined electrical noise by the normalized sensitivity. Thus, EDR is independent of NA and wavelength since LR and resolution scale proportional to Rayleigh range, leaving their ratio unchanged. Thus, only the influence of the dimensionless pinhole parameters has to be investigated. As in Subsection 4.3.1, about noise, this is only done for the normalized FES since the EDR of the unnormalized FES is irrelevant. Figure 4.15 contains a three-dimensional graph and contour plot of EDR for unnormalized dimensionless FES as a function of dimensionless pinhole parameters.
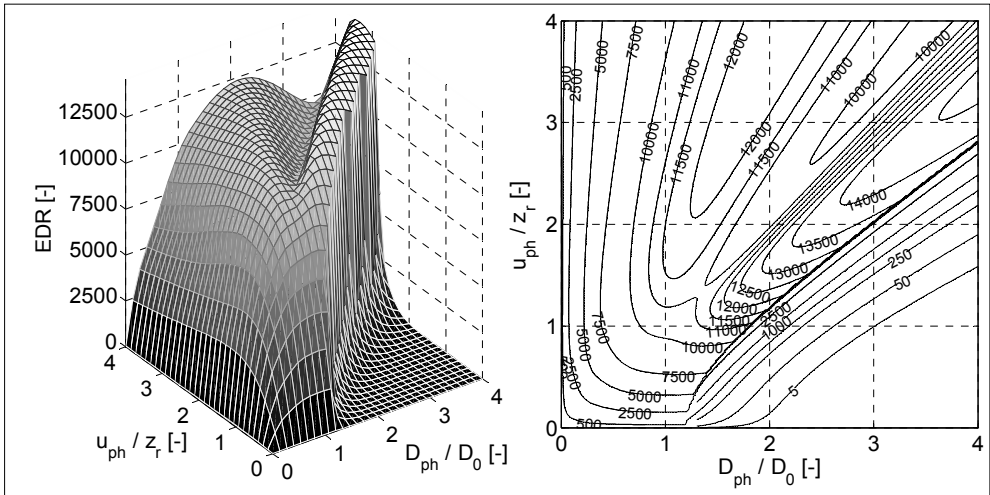
Figure 4.15 Electrical Dynamic Range (EDR) for varying dimensionless pinhole parameters.

The values presented here are calculated for a SUT reflectivity of 3.5 % since this is the performance limiting case. The same simulations have been performed with a reflectivity of the SUT of 95 %. In this case, for some pinhole parameters the laser power must be attenuated to prevent saturating the normalization electronics. These simulations show a dynamic range of up to three times higher, depending on pinhole parameters. Furthermore, to test the method, the same calculation is made using the parameters of the demonstrator, giving a resolution of 92 pm. This is

close to the values of 75 pm to 90 pm obtained by experiment and therefore provides some validation of the models of noise and EDR.

4.4 DIFFERENTIAL CONFOCAL SYSTEM OPTIMIZATION

The property analysis presented in the previous sections is used to optimize the differential confocal sensor. A tradeoff between different properties must be made since their optima do not coincide.

4.4.1 OPTIMIZATION APPROACH

First, a region of dimensionless pinhole offset and diameter is chosen based on the properties determined in the previous sections. In this application, the dominant criteria for determining this region are the Optical Signal Range (OSR) and Tilt Dependent Error (TDE). A high OSR is important because it leads to a high ratio of measurement range to uncertainty, two properties which are desirable, but mutually conflicting. A low TDE over the entire range is important to achieve the 35 nm uncertainty goal at 5° surface tilt, since aperture correction is expected to compensate only partly for this error. Furthermore, the EDR should be sufficiently high to allow the resolution specification to be met.

Once the optimal dimensionless pinhole parameters are known, so is the *dimensionless* near-Linear Range (LR) of the normalized FES. Therefore, the desired *dimensional* LR of the normalized FES can be achieved by choice of the Rayleigh range, z_r. The dominant requirement for this part of optimization will be the 10 nm measurement uncertainty at 0° tilt. When the desired value for z_r is known, the laser wavelength and NA can be varied to obtain such a z_r. Pinholes and imaging lenses can then be selected so that the dimensionless pinhole parameters are within the optimal region. Because aberrations and slight misalignment might occur in the real system, it is wise to choose a robust working point: close to the optimal region in the design space, the critical properties should not degrade rapidly. Choice of wavelength, NA, imaging lenses and pinhole diameters depend on which components can be acquired; for that reason, this selection will be treated in Chapter 5 on optical design.

4.4.2 OPTIMIZATION FOR NANOMEFOS

To determine the region of optimal dimensionless pinhole parameters, the contour plot of the OSR and the line of minimal TDE are plotted in Figure 4.16. Also plotted in this figure is the area in which the EDR is below 7000.
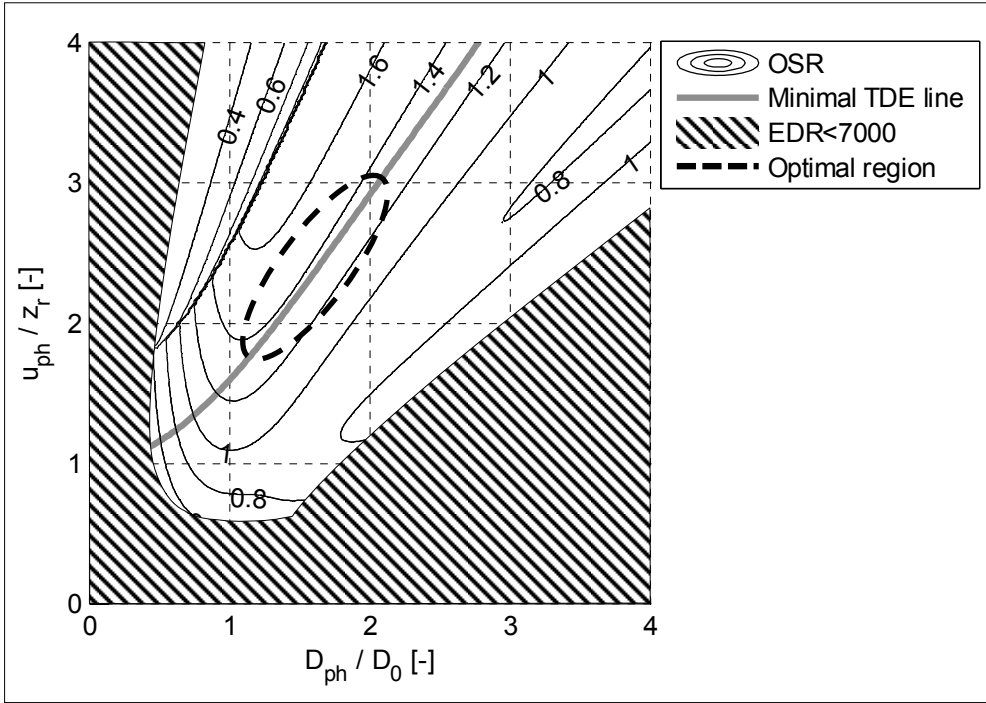
Figure 4.16: Combined plot of the optimization criteria plotted against dimensionless pinhole parameters. Shown are the contour plot of Optical Signal Range (OSR) in black, the line of minimal Tilt Dependent Error (TDE) in gray, the design space with Electrical Dynamic Range (EDR) below 7000, hatched, and the region with optimal pinhole parameters, encircled by a dashed line.

A large LR is desired to allow for tracking errors during measurements of heavily freeforms. The critical physical quantity being measured in the sensor is the incident power on both photodiodes. Measurements of incident power are subjected to various mechanisms, which introduce uncertainty in the measurements, such as: interference, contamination of the optics and degradation of coatings. As a rule of thumb, the ratio of long-term absolute measurement uncertainty to the measurement range which can be obtained by straightforward measurements of incident power is somewhere in the order of 1/10. This ratio can be increased to somewhere in the order of 1/100 if considerable measures are taken, such as calibration and compensation (van Amstel, 2008). In this case, however, the signals are normalized and there is the possibility for inline drift reconstruction by means of closure and for inline FES calibration before and after each measurement. Because these measures remove most of the dominant contributors to the measurement uncertainty, it is believed that for this application the ratio can be increased to 1/500. This assumption involves some risk, since it

remains to be seen that this ratio can be achieved in the prototype. To reduce this risk, the pinhole parameters can be chosen in a part of the optimal region, which gives high dimensionless LR and by tuning the Rayleigh range accordingly. Then, if in practice the uncertainty is too high, it can be reduced by inserting other pinholes that are also within the optimal design space but give lower dimensionless LR, without increasing the Rayleigh range. If equal uncertainty budgets are allocated to the interferometry and the differential confocal systems, the total measurement uncertainty of 10 nm works out to a required uncertainty of 7 nm. The dimensional LR should therefore be about 3.5 µm.

The minimum value for the EDR can be determined from the resolution requirement of 1 nm, the LR of 3.5 µm and a safety factor of 2, which is applied to create a margin for the noise sources not accounted for in EDR. This gives a lower boundary of 7000; the part of the design space with an EDR below this value is represented as the hatched area in Figure 4.16.

The design space is chosen around a part of the minimal TDE line and where the OSR is as high as possible. In Figure 4.16, this region is encircled in black. It can be seen that the OSR is nearly constant along the line of minimal TDE within this region. In addition, it is in the middle of the area for which the EDR is sufficiently large. In Figure 4.17, left, it is shown that the estimated EDR within this region is actually between 11000 and 12000.
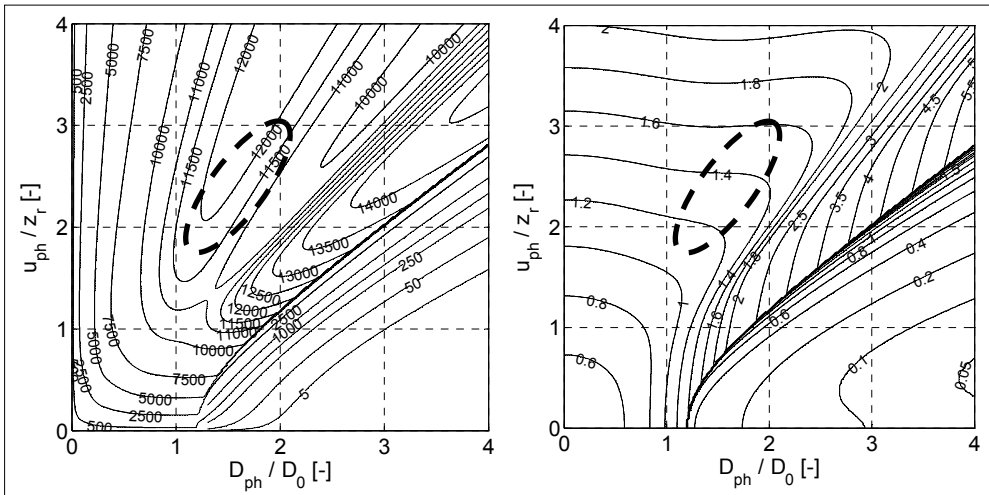
Figure 4.17: Electrical dynamic range (left) and near-linear range (right) of the normalized FES and the optimal region of the design space.

To determine what the dimensionless normalized LR is within the region of optimal design space, Figure 4.17, right, shows the contour plot of the normalized

LR and the optimal region. It can be seen that for the optimized system the *dimensionless* LR is between 1.1 and 1.6. Since a *dimensional* LR of about 3.5 µm is desired, the Rayleigh range should thus be between 2.2 µm and 3.2 µm. A Rayleigh range close to 2.4 µm is aimed for to allow lowering the measurement uncertainty at the expense of measurement range by changing pinhole parameters after completion of the prototype.

Which working point is chosen depends on application requirements and on the designer's preference and opinion. Although the aim here is to meet the requirements dictated by NANOMEFOS, it is believed that the presented process is illustrative for a generally applicable approach to the optimization of differential confocal sensors. For an application with different requirement the same type of procedure can be used, however, the optimization criteria might be chosen differently.

5 OPTICAL PROTOTYPE DESIGN

The optical layout and the considerations regarding optical design will be presented first. The choice of optical components is described next, and depends on the desired properties, partially found in the previous chapter, as well as on availability, manufacturability and cost. At the end of the chapter an overview of the optical system is given.

The optical system is of crucial importance to both the primary and secondary measurement methods and as such dictates the boundaries and requirements for the optomechanics and mechatronics of the sensor. Therefore, before the rest of the system is designed, the optical design needs to be known to a large extend. There are four main issues, which have considerable impact on the optical layout, these are:

- integration of an interferometer into the Differential Confocal System (DCS),

- folding of the optical path so that it fits inside the volume envelope,

- beam delivery and pickup to and from the NANOMEFOS machine, and

- the prevention of ghost reflections parallel to the optical axis.

By treatment of these four issues, the optical layout will take shape step by step. Next, some considerations on general selection criteria for components will be discussed followed by component specific requirements. In the last section of this chapter an overview and some properties of the resulting optical system will be given.

5.1 INTEGRATION OF AN INTERFEROMETER

In Subsection 2.1.2, it was proposed to track the Surface Under Test (SUT) by translating the objective lens of the DCS while measuring the displacement with an interferometer.

5.1.1 HOMODYNE VS. HETERODYNE INTERFEROMETRY

In commercially available interferometric distance measurement systems, there are two methods of detection: homodyne and heterodyne. The working principle of homodyne interferometers is quite similar to the first interferometer built (Michelson, 1887). A schematic representation of a modern homodyne interferometer is shown in Figure 5.1.
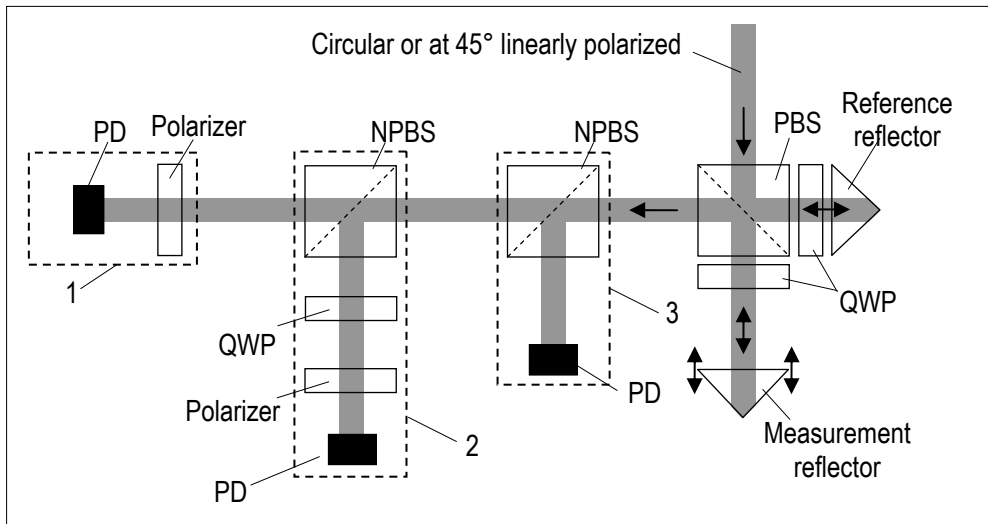
Figure 5.1: Schematic representation of a homodyne interferometer incorporating quadrature detection and intensity correction.

Typically, a monochromatic beam is used which is split by a Polarizing Beam Splitter (PBS) into a reference beam and a measurement beam with orthogonal polarization. After reflection by a reference and a measurement reflector respectively, the beams return to the beam splitter and are recombined. A photodiode (PD) and polarizer (1) are placed in the returning beam. The polarizer is oriented at 45° relative to the beam splitter and aligns the beam's polarization orientation so that interference occurs. Depending on the relative phase of the beams, the interference is (partially) constructive or destructive. When the measurement reflector moves, the optical path length of the measurement branch changes, thereby changing the phase difference between the beams. By analyzing

the signal from the photodiode, the cycles of constructive and destructive interference can be counted and interpolated, from which the distance the measurement reflector has moved can be calculated. The direction of the movement, however, cannot be determined using such a method. There are different ways of solving this problem; one solution is adding a Non Polarizing Beam Splitter (NPBS), a Quarter Wave Plate (QWP), a second polarizer and a second detector (2). The fast axis of the QWP is aligned with one of the polarization directions, and therefore shifts the phase difference of the beams by ½π, thus obtaining quadrature detection. This allows the direction of displacement to be determined as well as compensation of periodic errors (Heydemann, 1981). Furthermore, by adding another NPBS and photodiode (3), the intensity before interference can be measured to correct for errors arising from intensity variations.

A schematic representation of a heterodyne interferometer is shown in Figure 5.2.
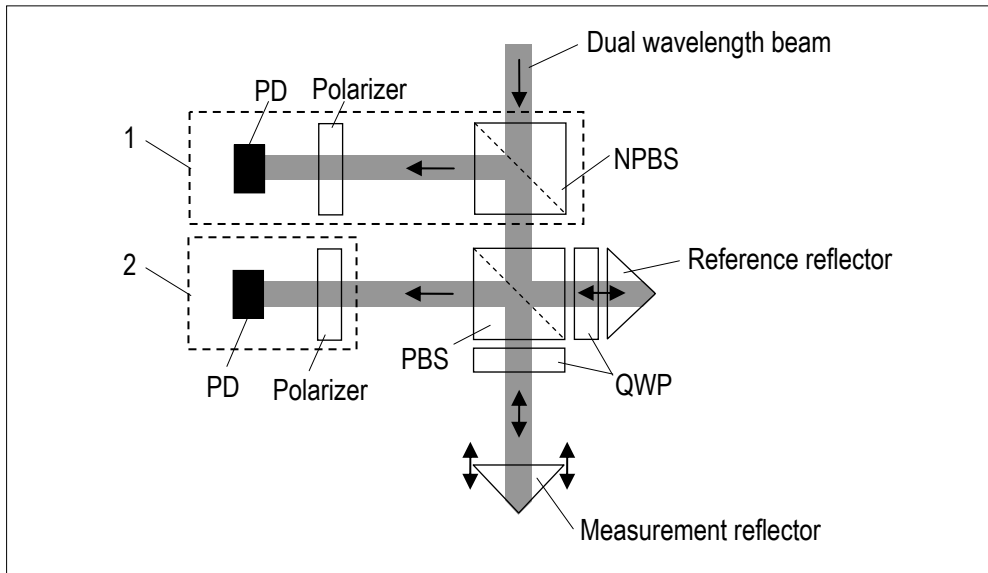
Figure 5.2: Schematic representation of a heterodyne interferometer.

In this type of interferometer, the reference and the measurement beams have slightly different wavelengths and orthogonal linear polarization. Such an output is most commonly produced with a Zeeman laser for which the frequency difference can be up to about 4 MHz. When these two components are made to interfere by a linear polarizer oriented at 45°, it results in intensity variations at the beat frequency. A reference frequency is generated in this way with the beam coming directly from the laser, for example by applying an NPBS, polarizer and photodiode (1). The remainder of the beam is split by a PBS into its two separate

frequency components. One component is reflected by a reference reflector constituting a reference arm of fixed optical path length, whereas the component at the other frequency is reflected by the measurement reflector. The beams are recombined at the PBS and a polarizer and photodiode (2) are used to cause interference and measure the beat frequency. This beat frequency is compared with the reference frequency taken directly from the laser. If the optical path difference between the reference and measurement arm of the interferometer is constant, the beat frequencies are equal. When the measurement reflector moves, however, the returning measurement beam experiences a frequency shift due to the Doppler effect, thereby causing the measurement beat frequency to change by the same amount. The shift in beat frequency is measured and from it, the rate of change in optical path length can be calculated. Through integration of the rate of change in optical path length, a displacement measurement is obtained.

Both heterodyne and homodyne interferometry systems can achieve the required uncertainty. Here a heterodyne system is chosen, mainly because the NANOMEFOS machine already incorporates a heterodyne interferometry system that is suitable for simultaneous measurement of three axes. Since only two axes are needed in the machine, the third can be applied to measure the displacement of the objective lens. The beam diameter of this system is 3 mm.

5.1.2   COAXIAL VS. DOUBLE PASS LAYOUT

To satisfy the Abbe criterion, the beams of both the Differential Confocal System (DCS) and the interferometer have to measure towards the same point. There are five fundamental ways to separate the two beams:

- position,
- wavelength,
- angle,
- polarization state, or
- temporal multiplexing.

Separation by angle is not practical for various reasons, one being the geometry of the measurement problem. Polarization state is already used to separate incident from returning beams and temporal multiplexing is not feasible because of the incremental nature of interferometry. Therefore, the options are reduced to separation by means of different wavelengths and spatial separation. Three configurations which enable measurement of the displacement of the objective lens and which satisfy the Abbe principle with respect to the focal point of the DCS are shown in Figure 5.3.
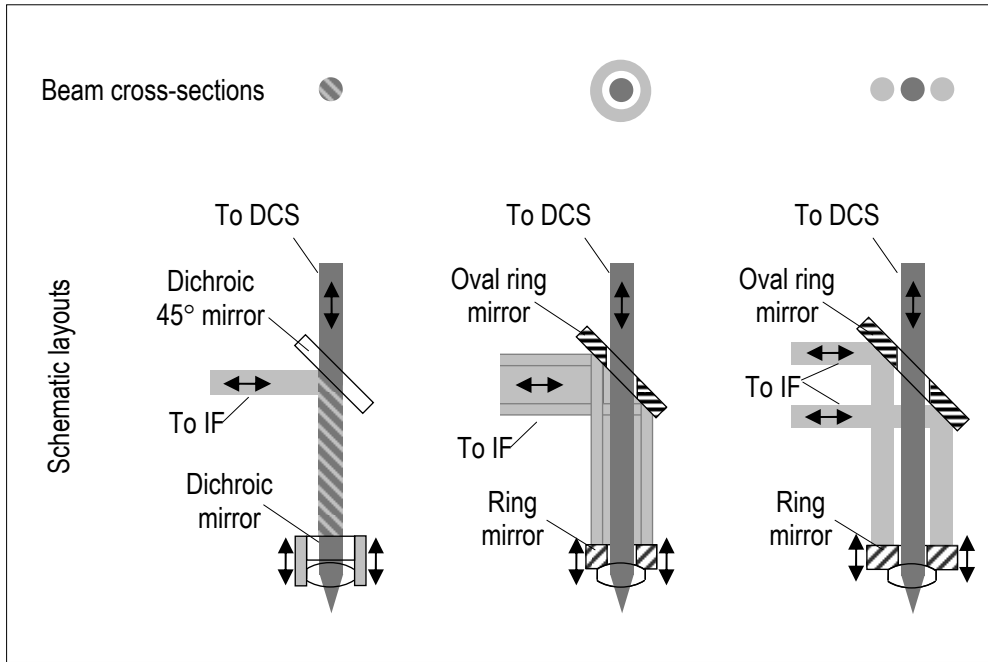
Figure 5.3: Three beam configurations to measure objective displacement in Abbe. Wavelength separation of coaxial Differential Confocal System (DCS) and InterFerometer (IF) beams by means of dichroic mirrors (left). Spatial separation by using a coaxial hollow IF beam around the DCS beam (middle). Spatial separation of DCS and IF beam by applying a double pass interferometer symmetrically with respect to the DCS beam (right).

The configuration of Figure 5.3, left, allows for the coupling in and separation of the two beams by use of different wavelengths and dichroic mirrors, i.e. mirrors which reflect light of specific wavelengths whereas other wavelengths are transmitted. A 45° dichroic mirror merges and splits the beams, a 90° dichroic mirror is stacked on top of the objective lens to reflect the interferometer beam, while the differential confocal beam is transmitted to the objective lens and the surface under test. The advantage of such a configuration is that the optical path from the optics to the sensor tip has a small cross-section, leaving ample space for an efficient actuator. Furthermore, the moving mass of this configuration is low, thus enabling actuation of the objective lens with low dissipation and high bandwidth. It is necessary that the wavelengths of the interferometer and DCS are chosen such that they allow proper separation by the dichroids. Since the IF system in NANOMEFOS applies a HeNe laser operating at 633 nm, choices for the wavelength of the DSC will be limited, but this does not have to be a problem. A more important disadvantage is that ghost reflections and leak-light might degrade

the performance of both the interferometry system and the DCS, since both systems operate at the same wavelength.

The configuration shown in Figure 5.3 (middle) applies a hollow interferometry beam that is coaxial with the DCS beam. It comprises an elliptical ring mirror to align the incident interferometry beam around the DCS beam and to separate the returning beams. In addition, a round ring-shaped mirror is attached to the objective lens to reflect the interferometer beam. The ring-shaped mirrors can either be plane mirrors with a bore through the centre or plates that are partially coated. Although the cross-section of the optical path is larger than for the two-wavelength configuration it is still relatively small. A big advantage compared to the two-wavelength option is that in this configuration no extra optical components need to be inserted into the optical path of the DCS. In addition, the chance that light from one system significantly influences the performance of the other system is greatly reduced. Unfortunately obtaining a hollow interferometer beam will probably cause considerable difficulties. No commercial stabilized laser suitable for high precision interferometry with this output is available. Obtaining the hollow beam by simply "cutting out" the center of a TEM_{00} beam, by using a glass plate with a reflecting dot coated on it, for example, will distort the wavefront due to diffraction. Alternatively the doughnut shaped TEM_{01*} mode has a clean wavefront but for its central intensity dip to be large enough to contain the DCS beam, its outer diameter needs to be large, offsetting the advantage of the small optical path cross-section.

Figure 5.3, right, shows a configuration consisting of a conventional double pass interferometer that applies an elliptical and a round ring-shaped mirror. This option has the advantage that it uses proven, commercially available technology. Furthermore, a double pass setup renders the setup insensitive to misalignment of the plane mirrors. A disadvantage is that the cross-section of the optical paths is quite wide in one direction. Together with the advantages of a rotationally symmetric actuator and guidance design, this leads to a rather large bore diameter through the actuator, thus limiting its efficiency.

The first two configurations need to be studied and tested before it is known whether they are applicable. Because of time constraints, financial reasons, the desire to limit risks in the realization phase, and the need for reliability regarding measurement uncertainty in the long term, it is decided to choose the double pass interferometer configuration.

5.1.3   HIGH STABILITY DOUBLE PASS CONFIGURATION

As will be treated in Section 5.5, the interferometer beam splitter dimensions are 20 x 20 x 20 mm and both the Polarizing Beam Splitter (PBS) and the corner cube

will be made of BK7. In a conventional double pass configuration as shown in Figure 5.4, left, the measurement beam travels a distance of five times the beam splitter edge length through BK7 and four times through the Quarter Wave Plate (QWP), which is 1 mm thick and made of crystalline quartz. The reference beam, however, only travels a distance of three times the beam splitter edge length through BK7. So, the measurement beam travels an extra 40 mm through BK7 and 4 mm through crystalline quartz compared to the reference beam. With linear thermal expansion coefficients of $\alpha_{BK7} = 7.1 \cdot 10^{-6}$ m/m/K and $\alpha_{quartz\perp} = 13.2 \cdot 10^{-6}$ m/m/K, refractive indices of $n_{BK7} = 1.517$ and $n_{quartz} = 1.547$ and thermo-optic coefficients of $dn_{BK7}/dT = 1.3 \cdot 10^{-6}$ /K and $dn_{quartz}/dT = -6.0 \cdot 10^{-6}$ /K this gives an Optical Path Difference (OPD) of 540 nm/K. Since the interferometer is operated in dual-pass a measurement error of 135 nm/K results. Therefore, it is instead chosen to apply a high stability double pass plane mirror configuration, which is shown in Figure 5.4, right.
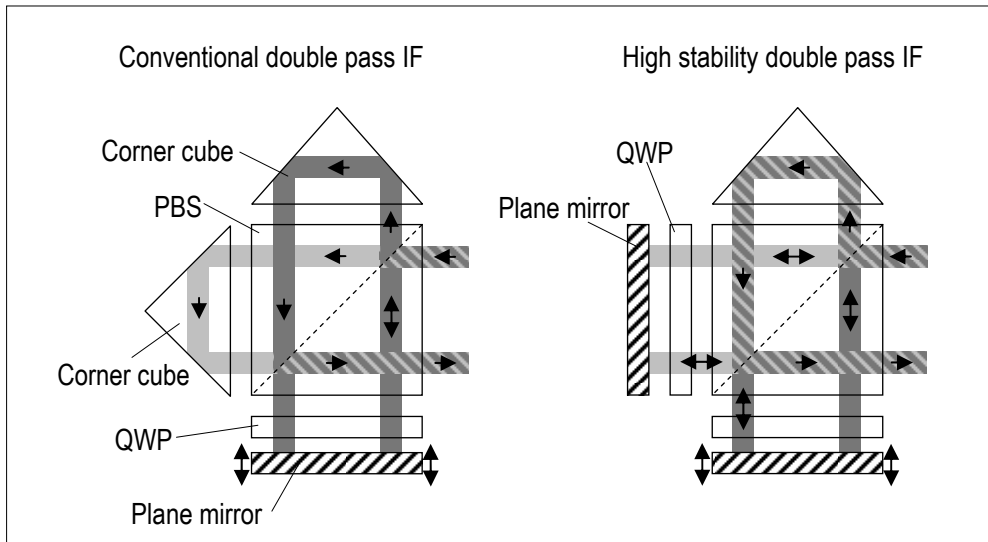
Figure 5.4: Measurement and reference beam paths for conventional plane mirror double pass configuration (left) and high stability plane mirror double pass (right) configuration.

In the high stability layout, the reference corner cube is substituted by a QWP and a plane mirror, thereby diverting the reference beam to the same corner cube as the measurement beam. Now both beams travel a distance of five times the beam splitter edge length through BK7 and four times the QWP thickness through crystalline quarts. Hence, if due to a uniform temperature change an OPD arises in one beam, it will also occur in the other beam, thereby canceling each other out.

5.1.4   COMPONENT INTEGRATION

The interferometer and the DCS have to be integrated into a compact, stable system. Therefore a layout was devised in which the same beam splitter and QWP are utilized by both the differential confocal and interferometry systems, as shown in Figure 5.5. In this configuration, the interferometer corner cube has its vertex removed so that it has an extra facet of optical quality, through which the DCS beam can enter or exit. A consequence of this layout is that either broadband or dual band optics need to be applied or that the DCS and the interferometry systems have to use a laser of nearly the same wavelength.
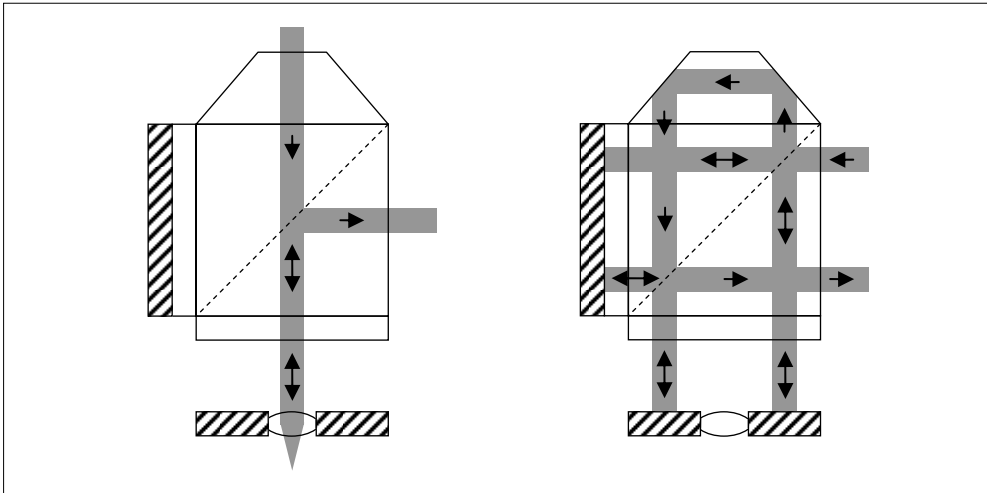
Figure 5.5: The Differential Confocal System (DCS) beam path (left) and the double pass interferometer beam path (right) in the optical monolith.

Furthermore, to obtain a compact, stable system the optical components will be bonded together using index-matched adhesive. This has the following advantages:

- the steps in refractive index are reduced from about 0.5 to about 0.03, giving reflectivities in the order of 0.01%, thereby limiting the risk of ghost reflections compared to AR coated surfaces,

- the shape specifications on the bonded surfaces used in transmission can be relaxed, because for transmission optics the OPD equals the geometrical error of the surface multiplied by the step in refractive index,

- direct bonding gives a compact optical layout,

- thermal conductivity between the optical elements is increased, which improves thermal uniformity and thereby the thermal stability of the interferometer,

- once bonded, the surfaces are not susceptible to contamination,

- less alignment is necessary during assembly because the tight angular tolerances on the bonded faces automatically align some of the degrees of freedom between the components, and

- short- and long-term mechanical stability is high.

A disadvantage is that index-matched adhesive is a UV curing adhesive that does not enable the components to be separated once it has cured. Hence, the complete optical monolith needs to be replaced if a part is misaligned, damaged, not within specification, contaminated or if there are air inclusions in the adhesive layer. Since most of the components are specials, this can potentially lead to considerable delays and high costs. It is thought, nevertheless, that the benefits outweigh the risks, especially bearing in mind that a cyclic disturbance, such as described in Subsection 3.2.8, might result from ghost reflections. The separate optical components and the optical monolith are shown in Figure 5.6.
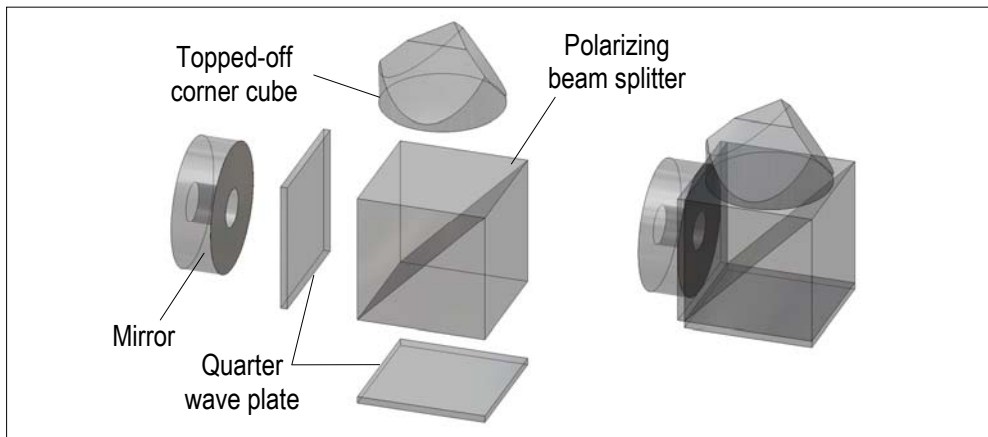
Figure 5.6: Exploded view of the optical monolith (left) and the cemented monolith (right).

The corner cube functions as a retro reflector at any rotation around the vertical, however, the orientation as shown is preferred, because in this way the interferometer beam is not intersected by mirror edges. The bore through the interferometer reference mirror has two functions: it is used during alignment (Section 6.1) and it allows light from the DCS laser that is not properly separated by the PBS to escape without degrading DCS performance.

This optical design, as well as some alternative layouts, is described in a patent application (van Amstel, Cacace and Henselmans, 2007).

5.1.5 MEASUREMENT REFLECTOR LOCATION

To enable measurement of concave optics, the distance between the ψ-axis centerline and the differential confocal measurement spot needs to be 100 mm nominal. Due to the volume envelope, the need for an efficient actuator and the delivery of the interferometer beam (Section 5.2) it is advantageous to place the IF/DCS beam splitter so that it is intersected by the ψ-axis centerline. Consequently, if the measurement reflector is directly attached to the objective lens, the deadpath length, i.e. the measurement beam's one-way optical path length in air, is about 75 mm. A change in the refractive index of air results in an apparent displacement of the measurement reflector: the deadpath error. Note that in the dual pass configuration the optical path length is actually four times the deadpath length, this however, cancels out with the increased sensitivity of one fringe per λ/4 displacement.

The refractive index of air changes with environmental conditions, the most important being: Temperature, Pressure and partial pressure of Water vapor indicated by T, P and W. There are various adaptations of the empirical Edlèn equation (Edlèn, 1966) that describe the influence of these parameters on the refractive index of air. From the adapted equation by (Birch and Downs, 1994) the linearized sensitivities to temperature, pressure and partial water vapor pressure can be derived to be $\delta n/\delta T = -9.30 \cdot 10^{-7}$ /K, $\delta n/\delta P = 2.68 \cdot 10^{-9}$ /Pa and $\delta n/\delta W = -3.63 \cdot 10^{-10}$ /Pa, for standard laboratory air, T = 20°C, P = 101325 Pa, W = 926.8 (Cosijns, 2004). For the deadpath length of 75 mm, this gives sensitivities of 69 nm/K, 0.2 nm/Pa and 0.03 nm/Pa to temperature, pressure and partial water vapor pressure, respectively.

The typical daily variation of temperature, pressure and water vapor content of laboratory air is ΔT = 0.1 °C, ΔP = 2000 Pa and ΔW = 100 Pa according to (Cosijns, 2004). These values give deadpath errors of 6.9 nm, 400 nm and 3 nm respectively. The NANOMEFOS machine therefore applies an environmental sensor to measure the change of these parameters and correct for it. The absolute 2σ measurement uncertainty of the sensor used for T, P and W is 0.2 K, 30 Pa and 24 Pa, respectively. If the relative measurement error over one measurement time would be the full absolute 2σ measurement uncertainty the deadpath error would result in an uncertainty of 16 nm. This is unlikely due to the short measurement times of about 15 minutes. Furthermore, the nature of the measurement problem allows drift to be reconstructed (Subsection 8.3.1). Nevertheless, there is a 2σ measurement uncertainty of 1 nm over the 5 mm stroke of the objective lens due to the uncertainty in the resulting scaling error.

Apart from fluctuations in the environmental conditions of the room, actuation of the objective lens is expected to introduce uncertainty in the environmental

conditions in the sensor. As will be shown in 6.3.2, dissipation of the actuator will be below 10 mW even for heavily freeform surfaces. The temperature of the deadpath air will thus increase little and gradually over the course of one measurement. Because of the slow gradual nature of such drift, it is expected that it can be mostly compensated for.

On the contrary, the influence of turbulence and fluctuations in pressure, caused by pumping due to translation of the objective lens, will be fast and hard to predict. This might be a significant source of measurement uncertainty. Applying a low vacuum or a helium atmosphere would greatly reduce the effect, but will introduce many practical problems, and is therefore rejected. Alternatively the deadpath error can be reduced by placing the reference mirror closer to the IF beam splitter.

Figure 5.7 shows both the concept where the reference reflector is directly attached to the objective lens (left) and a concept where a tube of low expansion material is used to connect the measurement reflector to the objective lens (right). This way the deadpath length can be reduced to about 5 mm, thereby reducing the deadpath error by a factor of 15.
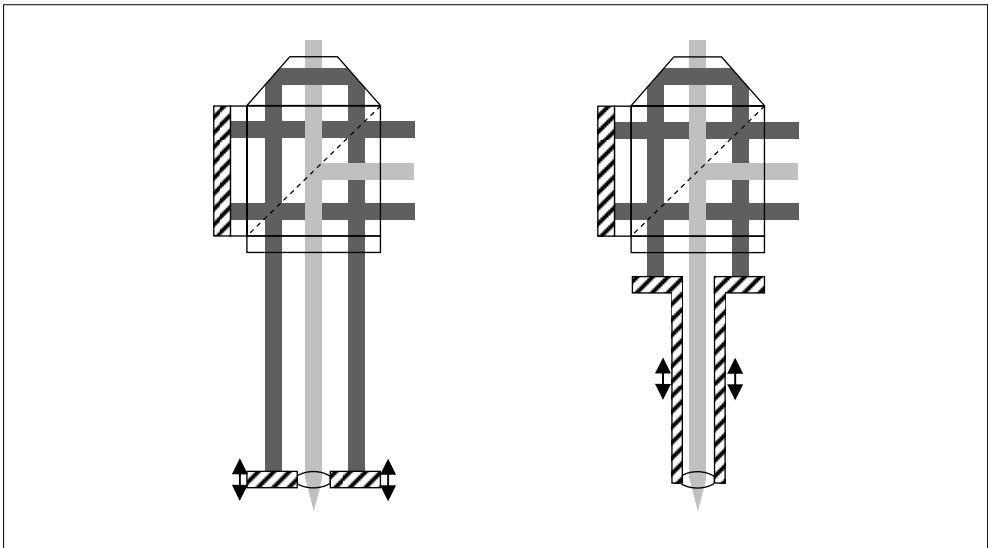
Figure 5.7: Location of the measurement reflector if it is directly attached to the objective lens (left) and when it is connected to the objective lens through a tube of low expansion material (right), to reduce deadpath length.

With a linear thermal expansion coefficient of $\alpha_{zerodur} = 0.05 \cdot 10^{-6}$ m/m/K for Zerodur of expansion grade 1 (Schott, 2008) a thermal sensitivity of 3.5 nm/K is obtained. A disadvantage is that it considerably increases the moving mass of the

system. This might be partly compensated by the opportunity to decrease the diameter of the bore through the actuator. Furthermore, depending on the mechanical design, some dynamic errors might be introduced.

The idea for this alternative solution arose when fabrication of the configuration depicted in Figure 5.7, left, was at an advanced stage. Because of time constraints and financial reasons, it was not a feasible option at that time and is consequently not further investigated. Either for a next generation sensor or an upgrade for the current sensor, it might be worthwhile to look into this configuration in more depth.

To limit the influence of turbulence and pressure fluctuations, some air vents and holes are added to the actuator design. The deadpath error due to turbulence and pressure fluctuations is nevertheless expected to be one of the major contributors to the overall measurement uncertainty.

5.2 BEAM DELIVERY AND PICKUP

Because of the tight volume envelope and because of thermal issues, the light sources for the Differential Confocal System (DCS) and the interferometer are not integrated in the sensor. It is most convenient to deliver the supply beams to the sensor by means of fibers. For both the DCS and interferometer, the polarization state of the supply beams is important to their performance. Even if a Polarization Maintaining (PM) fiber is used, the polarization state of the output varies significantly for varying stresses in the PM fiber. In the NANOMEFOS machine, this will be the case since movement of the stages and rotation of the ψ-axis will cause the PM fiber to deform. For the DCS beam, this is allowed because linear polarization can be restored by placing a polarizer behind the fiber collimator. The arising intensity variations in the beam are compensated by normalization of the detector signals. For the IF beam however, this is not possible since it contains two orthogonally linearly polarized beams.

Therefore, the IF beam must be supplied to the sensor by use of mirrors. It enters the ψ-rotor from the back, through a bore along its centerline. For the interferometer to work properly, the two orthogonal polarizations of the IF beam have to maintain aligned to the polarization axis of the beam splitter cube when the ψ-axis is rotated. For that reason, a Quarter Wave Plate (QWP) is fixed to the ψ-axis stator while a second QWP is bonded to the Polarizing Beam Splitter (PBS). If the stationary QWP is aligned so that its fast axis is at 45° relative to the beam's polarization directions, it transforms one component of the beam into left handed circular polarized light and the other component into right-handed circular polarized light. The second QWP is placed within the ψ-rotor with its fast axis

oriented at 45° relative to the PBS. This QWP transforms the right- and left-handed circular polarized components back into two orthogonal linearly polarized components under 45° relative to the QWP's fast axis. Now, when the ψ-rotor rotates the polarization directions of the beam stay aligned to the PBS.

Using Jones matrices (Jones, 1941) it can be shown that this method introduces a relative phase shift of 2ψ between the two linearly polarized components, as treated in (Henselmans, 2009). The relative phase shift causes a linearly increasing apparent displacement of the objective lens of 2λ per revolution, where λ is the wavelength of the laser. Applying the signal of the ψ-axis encoder, this error can be compensated for. Furthermore, some polarization mixing can be expected due to the limited tolerances on retardation and alignment of the QWPs. This causes a cyclic error, which can be suppressed by calibration.

After the IF beam exits the PBS and the measurement and reference beam are recombined, they can be made to interfere by a polarizer oriented under 45°. Since now the interferometer signal is carried in the frequency by which the intensity varies, the polarization state of the beam is not important anymore. Therefore, the beam can be delivered to the interferometer detector with a fiber. Considering the layout and assembly of the machine, the fiber pickup for the returning IF beam can best be placed at the front end of the ψ-rotor.

It seems favorable to allocate the space available between optical monolith and object lens to the actuator, the guidance and the mechanical interface between sensor and ψ-rotor. Combined with the limited volume available within the ψ-rotor this has been reason to consider placement of the DCS detectors on the ψ-axis stator, and transmit the DCS beam through a bore along the ψ-axis centerline at the front end. This has been rejected because rotation between parts of the optical system will most likely degrade reproducibility. Hence, the optical system and accompanying optomechanics must fit into the ψ-rotor.

5.3   FOLDING OF OPTICAL TRAIN

It is desirable that the sensor can be removed from the machine in one piece, so that the internal alignment of the sensor does not change during mounting and unmounting. Therefore, the optical system must fit into a pocket of about 41 mm x 84 mm x 37 mm within the ψ-rotor between the ψ-axis air bearings (Section 1.3). To make this possible, the optical path of the DCS is folded using mirrors and prisms. Figure 5.8 shows the outer perimeter cross-section of the ψ-rotor pocket (1), the optical components and the DCS (left) and IF (right) beam paths.
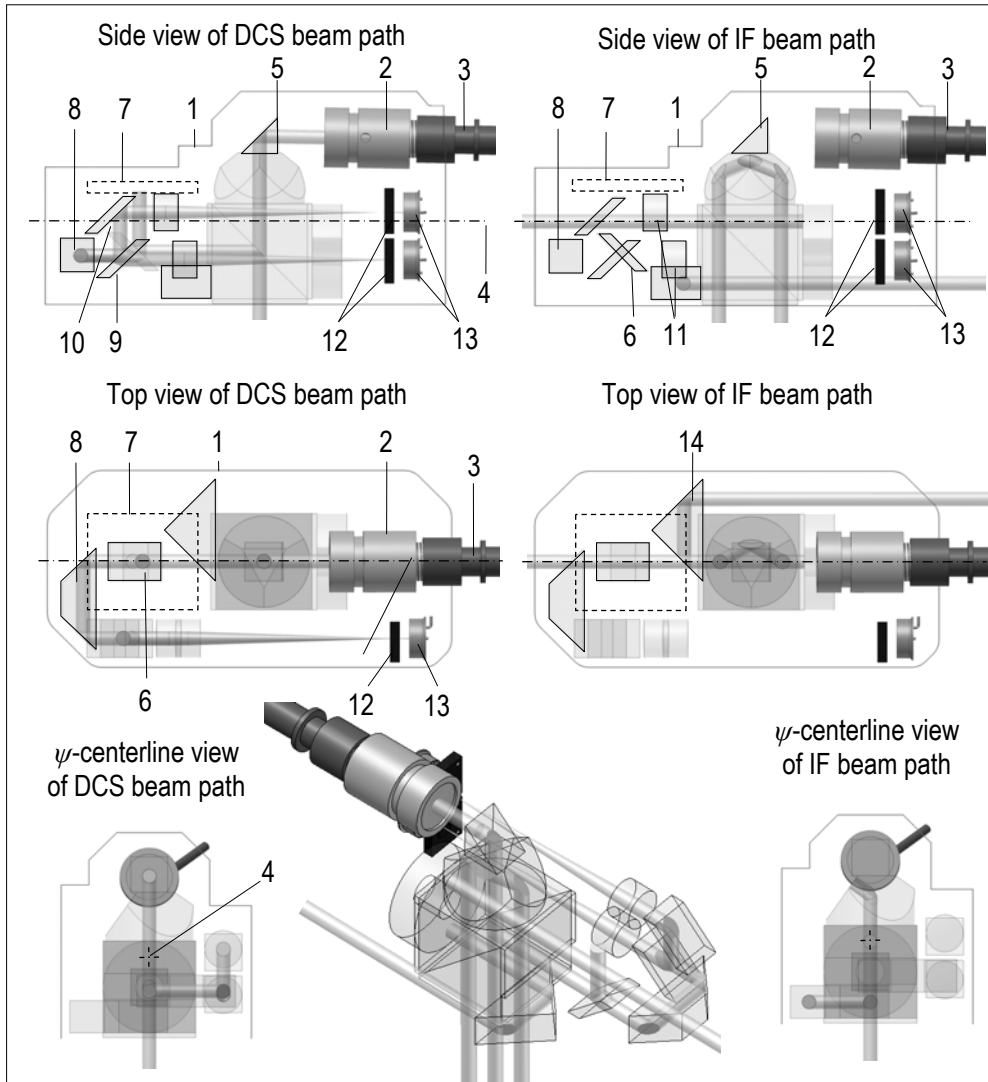
Figure 5.8: Various views of the folding of the Differential Confocal System (DCS) and InterFerometer (IF) beams; the outer perimeter cross-section of the ψ-rotor pocket is also shown.

The DCS supply beam is delivered through a PM fiber which connects to a collimator (2) using a fiber coupling (3). The collimator is oriented parallel to the ψ-axis centerline (4) to allow insertion of the collimator through a bore in the ψ-axis from its front end, after the sensor is mounted into the ψ-rotor. This is necessary because the fiber must be detached to slide the sensor into the pocket.

The collimated beam is reflected towards the PBS by a right-angle-prism (5) that is bonded onto the optical monolith. When the DCS beam returns from the surface under test it is coupled out by the PBS after which a small part is reflected by a beam splitter plate (6) to the Position Sensitive Detector (PSD) (7). The remaining beam is reflected backwards with an offset by a dove prism (8) so that the detector branches of the DCS can be placed next to the optical monolith.

A non-polarizing beam splitter (9) splits the beam in two. To stay within the volume envelope the reflected beam is reflected again by a folding mirror (10) so that it becomes parallel to the transmitted beam. This has the added benefit that the two detector branches experience similar temperature changes. A disadvantage is that the beams do not have similar wavefronts anymore since the folded branch experiences some aberration and polarization effects from the mirror, while the other one does not. The image lenses (11), pinholes (12) and photodiodes (13) in the detector branches are also shown. Because the IF beam is supplied from the back of the ψ-rotor while the interferometer fiber pick up is placed at the front, a right-angle-prism (14) is used to reflect the beam backwards with an offset so it can pass the optical monolith.

5.4 PREVENTION OF PARALLEL GHOST REFLECTIONS

It is desirable to limit the number of optical surfaces that cause ghost reflections parallel to the DCS beam, to minimize the risk that a cyclic disturbance as observed in the demonstrator (3.2.8) occurs. By bonding the optical components to each other, the number of such surfaces is already reduced from nine to three. To reduce disturbances caused by the remaining surfaces, the optical monolith can be slightly tilted by an angle α_{bs} relative to the incoming DCS and IF supply beams, as shown in Figure 5.9, left. A disadvantage is that it leaves less space for the DCS entrance facet at the virtual vertex of the corner cube. Furthermore, an offset Δc_{IF} is created between the IF reference and measurement beams, and the beams between the objective lens and the optical monolith are not perpendicular to the supply beams and the beams going to the detectors but instead make an angle of $90°-2\alpha_{bs}$. This method therefore introduces measurement errors and complicates the optomechanical design, and is rejected.

An alternative solution is to align the PBS well but to change the angle of the entrance facet of the DCS folding prism as well as the DCS supply beam, and to bond a wedge with angle β_w onto the surface where the DCS beam exits the optical monolith, see Figure 5.9 (middle and right). This eliminates two out of the three remaining surfaces that might produce parallel ghost reflections. A similar wedge can also be bonded to the dove prism, so that the beam's deviation is cancelled. In addition, the dove prism can be oriented under an angle $\beta_w/2$, so that its exit face

does not produce a parallel ghost reflection either. Despite the remaining surface that might cause parallel ghost reflection, this option has been chosen, since it seemed the best solution when the design was finalized.
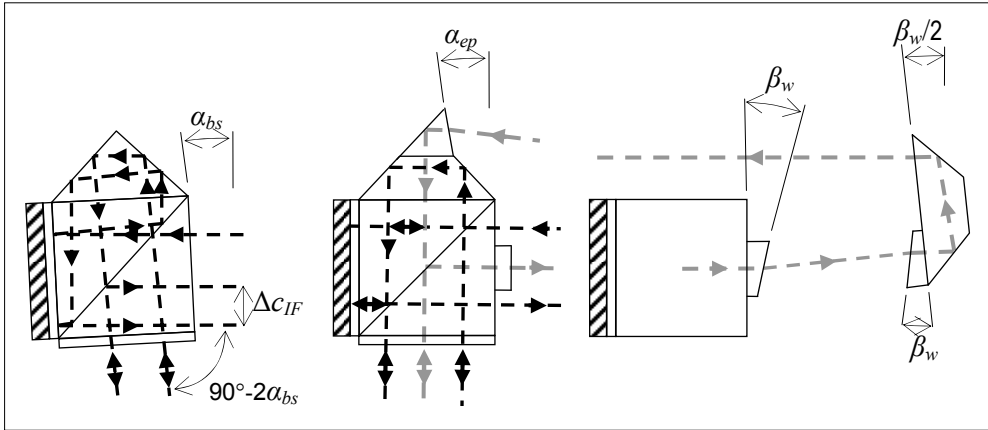
Figure 5.9: Effect of tilt of the optical monolith on beam alignment (left) and alternative solution to prevent parallel ghost reflections (middle and right).

Later the idea for a third layout arose which is shown in Figure 5.10, left. The PBS is rotated by an angle α_{bs}, the entrance facet of the DCS folding prism is given an angle of α_{ep} and two wedges of angle α_{bs} are bonded to the optical monolith, eliminating all optical surfaces causing ghost reflections parallel to the DCS beam.
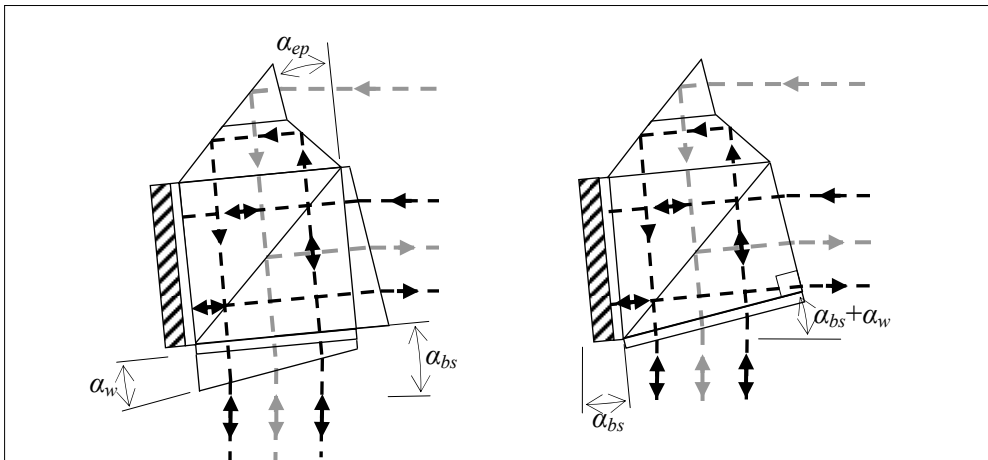
Figure 5.10: Elimination of parallel ghost reflections by tilt of the optical monolith in combination with 2 wedges (left) and a modified beam splitter geometry (right).

Using Snell's law it can be shown that the beam paths in the glass are aligned to the PBS, plane mirror and corner cube, while the beams between the objective lens and the optical monolith are perpendicular to the supply beams and the beams going to the detectors, when the angles are chosen such that:

$$\alpha_w = \sin^{-1}\left(\frac{1}{n_g}\sin\left(\alpha_{bs} + \alpha_w\right)\right) \tag{5.1}$$

where n_g is the refractive index of the glass. For small angles, (5.1) can be approximated by:

$$\alpha_w = \frac{1}{n_g - 1}\alpha_{bs} \tag{5.2}$$

which is accurate to within a couple of arc seconds for practical values of α and n_g.

Figure 5.10, right, shows an alternative layout where instead of adding two wedges, one of the 45-45-90 right-angle prisms of the beam splitter is replaced by a right-angle prism with different angles, thus reducing the number of bonded components. For this application, the layout with the wedges is preferred because it allows the use of a standard quarter wave plate at 0° incidence in between the beam splitter and the wedge. Also, it is more cost effective for a prototype or a small series, since both the beam splitter and the wedges can be purchased as standard components.

5.5 CHOICE OF COMPONENTS

The performance of the optical system depends on alignment and the characteristics of the components used. The cost of components, however, can increase rapidly when specifications are tightened. Decreasing the allowed surface shape error by a factor of two, for example, can sometimes more than quadruple costs. Therefore, it is important to identify for each component what the necessary requirements are.

5.5.1 GENERAL CONSIDERATIONS

Some of the selection criteria are applicable to most of the components and will be discussed first; the selection criteria for specific components are treated in Subsection 5.5.2.

Wavefront error

The wavefront error must be low for both the DCS and the interferometer, but for other reasons. For the DCS it is desirable that diffraction limited object and image spot sizes are obtained. This ensures a symmetrical Focus Error Signal (FES), a lower tilt dependency caused by aberrations, and optimal behavior for the design parameters found in Section 4.4, including minimal inherent tilt dependency.

For the interferometer, the variation in wavefront between the reference and measurement beam is more important than the absolute wavefront error. The reason is that relative phase variations between measurement and reference beam cause fringes to appear over their area of overlap. If the measurement and reference beam have a uniform intensity distribution and exactly one fringe is present, the modulation depth becomes zero since the power of the beam as a whole remains constant. In the real system, the measurement and reference beam have a Gaussian intensity distribution, so the middle of the interference pattern dominates the signal. Therefore, more than one fringe over the overlap area is needed to cause complete loss of modulation depth. To have a margin, the goal here is to keep the worst-case relative phase variation over the diameter of the beams to about half a wavelength or less.

For the DCS to achieve diffraction-limited performance, the light source and delivery must produce a diffraction-limited beam to start with, and subsequently, the optical components should not deteriorate the wavefront too much. The specification of wavefront errors introduced by optical components, however, is often incomplete or ambiguous. Wavefront and surface shape errors are usually given in submultiples of wavelength, but whether this concerns the design wavelength or a test wavelength is not always mentioned. Moreover, for beam splitters the distinction between reflected and transmitted beam is rarely made. Sometimes instead of the transmitted wavefront errors, only the surface shape errors of a component are specified.

For specification of custom components (ISO 10110-5:2005, 2005) has been adhered to, in which errors are stated as fringe submultiples at a 540 nm test wavelength. Nevertheless, a reference wavelength of 633 nm is used here when errors are expressed as submultiples of wavelength. This is done because it is the most commonly stated test wavelength by suppliers and, in this application, it is close to the design wavelength as well.

Wavefront errors are introduced by optical components through various mechanisms, namely:

- aberrations in the design,
- surface shape errors,

- stress- and therm-optic effects,

- mechanical and thermal deformation, and

- material inhomogeneity.

Here, components are selected of which the design is diffraction limited and the optical materials are fine annealed, leading to low nominal and material inhomogeneity errors respectively. Furthermore, the magnitude of errors introduced by the stress- and therm-optic effect, mechanical and thermal deformation and material inhomogeneity scale with size. In addition, the optomechanical and mechatronic design aims at limited thermal and mechanical loads. Consequently, surface shape errors are expected to be the dominant cause for most of the wavefront error introduced by components.

Surface shape error

Four situations can be identified with respect to surface shape errors, namely for:

- external reflective surfaces,

- internal reflective surfaces,

- external transmissive surfaces, and

- internal transmissive surfaces.

This classification is made because to convert surface shape error into wavefront error another multiplication factor must be applied for these situations. The multiplication factor is 2 for external reflection, $2n_g \approx 3$ for internal reflection, $n_{g1}-n_{g2} \approx 0$ for internal transmission and $n_g-1 \approx 0.5$ for external transmission. Here n_g is the index of reflection of the glass and the approximations are valid for the optical materials used. Therefore, for the same allowable wavefront error, very different surface shape tolerances can be applied for various surfaces.

Surface shape can be divided into a regular and irregular component, i.e. the power of the surface shape error and the remainder after subtraction of the power component, as specified by (ISO 10110-5:2005, 2005). The distinction is relevant since the power component only causes an axial focal shift. For the DCS beam this is less critical, because it can be compensated for by axial adjustment of the imaging lenses, and does not significantly affect performance.

Unfortunately, many optical suppliers do not specify surface shape according to ISO 10110-5 but only give the peak to valley deviation of the total surface shape error. Hence, for surfaces that are not custom made, the irregular error sometimes has to be asked from manufacturers or estimated from known specifications.

Because the DCS beam is only 1.9 mm in diameter, while most of the clear apertures are at least a couple of times larger, the wavefront error of the beam is expected to be significantly smaller than the wavefront error as calculated over the full aperture. Since supplier specifications are given over the clear aperture of the optics, these values are used in the treatment of surface shape errors.

Beam deviation

When the entrance and exit surface of a component have an angular misalignment or if the alignment of a reflective surface is slightly off, it leads to angular beam deviations. Deviation of the DCS beam causes a lateral shift of the image spots. This can be compensated for by radial alignment of the pinholes, which have ample radial stroke for the deviations that arise from component tolerances. For the IF beam, deviation causes reduction in overlap between reference and measurement beam. The interferometry system, however, can handle much larger reduction in overlap than caused by component tolerances. Furthermore, the guidance for the objective lens will be aligned to the incident DCS beam. The measurement errors that nevertheless arise from slight beam deviations will inherently be compensated by the calibration measurements that are necessary anyway. Consequently, beam deviation is not critical in this application.

Coatings

AntiReflection (AR), reflective and beam splitting coatings are applied on several surfaces. They are different for various substrate materials, for use at other angles of incidence, and for internal or external reflection. The reflective coatings are all protected metal coatings. Compared to dielectric coatings these have the advantage that their reflectivity depends less on polarization direction, are simpler to apply and can be removed after a failed coating run. On external reflective surfaces, optimized protected aluminium coatings are applied on a nickel bonding-layer.

For all folding prisms, the criterion for total internal reflection is met, since the smallest angle of incidence is larger than the critical angle, which is 43.4° for fused silica and 41.2° for BK7. Nonetheless, all these surfaces have been coated with protected aluminium so that reflectivity is not influenced by surface contamination. The downside of this solution is that reflectivity is decreased by about 10% to 15% (Van Erps, 2003), which is acceptable because enough light is available in both the DCS and IF systems. For these coatings, no nickel bonding layer is used, since it degrades reflectivity too much.

At a narrow wavelength band, reflectivities can be in the order of 0.25% for relatively simple dielectric AR coatings, to below 0.1% for more complex dielectric AR coatings. The risk that something will go wrong however, increases significantly with coating complexity. In case of a failed coating run, dielectric AR

coatings, unlike metal reflective coatings, cannot be removed by etching. Because most of the components are specials of which only one is made, relatively simple coatings are used where possible.

Furthermore, a polarizing beam splitting coating and a non-polarizing beam splitting coating at 45° incidence are needed for the central beam splitter cube and the DCS beam splitter plate, respectively. These are both dielectric coatings optimized for the wavelength used.

Costs

Optical components can often be conveniently produced in small to medium sized series, therefore standard catalog components are in general much less expensive than specials. Here, however, there are three main reasons why the use of standard components is limited, namely:

- there is little margin in the size of most of the optical components since they have to fit in the volume envelope while their clear aperture has to be bigger than the beams,

- many of the components have to be uncoated on at least one of their surfaces because they will be bonded together, and

- some of the components have unusual geometries.

To keep costs down, some of the specials are made by modifying catalog items. Most catalog items can only be ordered with either all surfaces AR coated or all surfaces uncoated. To keep coating costs low, most components are ordered uncoated so that the surfaces to be coated can be processed in the same coating batches.

5.5.2 COMPONENT SELECTION

First, the optical components will be treated that determine the fundamental design parameters within the optimal design space found in Subsection 4.4.2. These components are:

- the light source, consisting of the laser, isolator, fiber and collimator,

- the objective lens,

- the imaging lenses, and

- the pinholes.

The objective lens and light source will be chosen first. Their combined properties determine the Rayleigh range at the surface under test. Because of this interdependence, they will be treated simultaneously. For a given beam diameter and wavelength, the imaging lenses determine the Rayleigh range and beam waist

at the pinholes and are selected next. Then the pinholes are chosen such that the dimensionless pinhole parameters are within the desired design space.

The remaining components do not determine fundamental design parameters; nevertheless, they influence system performance through non-ideal behavior. These components are treated in the order in which they are placed in the DCS' optical path: the polarizer behind the collimator, the components of the optical monolith, the dove prism, the PSD splitter plate, the DCS splitter plate and DCS folding mirror.

Light source and objective lens

The light source is acquired as an assembled and aligned system consisting of a laser, Peltier elements, control electronics, a Faraday isolator, a fiber coupler, a polarization maintaining fiber and a collimator. This allows specifications for overall performance from suppliers to be obtained. For selection of the light source, the following criteria are most important:

- wavelength,

- beam diameter,

- tunable output power from 1 mW to 10 mW,

- beam quality, and

- polarization ratio.
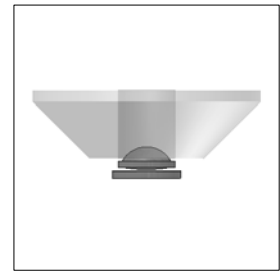
An aspherical singlet will be used as objective lens because of the small size combined with diffraction limited performance at large Numerical Aperture (NA). The dominant selection criteria are:

- the design wavelength,

- the Rayleigh range at an available beam diameter and wavelength,

- the NA of the clear aperture, and

- the introduced wavefront error.

Since both the DCS and interferometer share various components, it is beneficial if the wavelength of the DCS is close to that of the interferometry system. Therefore the wavelength of the DCS should be somewhere between 620 nm and 645 nm.

In Subsection 4.4.2, it has been determined that the Rayleigh range is preferably close to 2.4 µm and at least between 2.2 µm and 3.2 µm for optimal performance. At a wavelength of 633 nm, this is achieved if the beam has an NA of 0.28, for which the ratio of focal length to beam diameter must be about 1.7.

To prevent significant disturbance of the DCS beam, the clear aperture of the optical components must be at least 1.5 times the e^{-2} diameter of the beam. Hence, the lens must have a larger NA than the beam. In addition, 5° tilt of the surface under test will cause the returning beam to have an angle of 10° relative to the optical axis. This also has to be accommodated by the objective lens, resulting in a total needed NA of the objective lens of at least 0.52.

Considering the above-mentioned requirements the Panasonic EYLG50M073 lens and SRT-F640S10/APM/ISO/15m/OSYS fiber coupled diode laser from Micro laser systems have been selected.

The light source has the following specifications:

- a wavelength of 640 nm ± 5 nm,

- an output beam diameter of 1.9 mm,

- tunable output power from 0 mW to 10 mW,

- an M^2 below 1.1, and

- a polarization ratio better than 100:1.

The chosen objective lens has the following specifications:

- a design wavelength of 650 nm,

- a focal length of 3.3 mm,

- a numerical aperture of 0.6, and

- a wavefront aberration of $\lambda/18$.

The selected lens is designed for use with a glass plate in front of the lens, which is therefore added to prevent spherical aberration. Using equations (3.16) and (3.17) this combination of light source and objective lens can be calculated to give a waist diameter at the SUT of 1.42 µm with a Rayleigh range of 2.46µm.

Imaging lenses

Imaging lenses with a long focal length are desirable since this leads to high magnification, which is beneficial because Rayleigh range scales quadratic with magnification. This strongly reduces the system's sensitivity to distance variations between imaging lenses and pinholes. Furthermore, high magnification leads to loosened tolerances on radial pinhole alignment. Consequently, lenses with a focal length of 45 mm are chosen, leaving just enough space for the other optics and the pinholes and
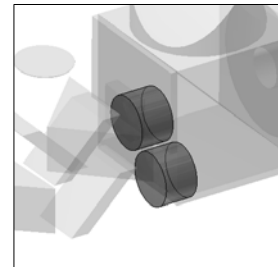
detectors. The resulting magnification is 13.6. Using equation (4.1) the NA of the imaging lenses is calculated to be 0.021. Because of the low numerical aperture, doublets can be applied without introducing significant spherical aberrations. Two standard doublets of 12 mm diameter are ground down to 8 mm diameter so that they fit in the available space. Using equations (3.16) and (3.17) the waist diameter and the Rayleigh range at the pinholes is calculated to be 19.3 μm and 0.457 mm respectively.

Pinholes

Non-reflecting pinholes are preferred to prevent light blocked by the pinholes from causing measurement errors. These non-reflecting pinholes are available with diameters of 20 μm, 30 μm and 50 μm, which give nominal dimensionless pinhole diameters of 1.04, 1.55 and 2.59. In Subsection 4.4.2 it has been determined that optimal system performance is obtained for values of dimensionless pinhole diameters ranging from 1.1 to 2.1, therefore pinholes of 30 μm are selected. For this dimensionless pinhole diameter, a dimensionless pinhole offset of 2.3 gives optimal performance.
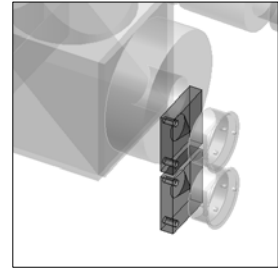
Polarizer plate

To obtain a stable polarization state of the DCS beam irrespective of stress in the polarization maintaining fiber, a Codixx Colorpol polarizer plate is placed between the collimator and the DCS folding prism. The polarizer plate is supplied with AR coatings and is placed at an angle of 1° to prevent parallel ghost reflections. The ratio of transmitted to blocked polarization is 83000:1 and the transmission of the aligned polarization direction is 90%. The polarizer plate introduces a wavefront error of $\lambda/5$.
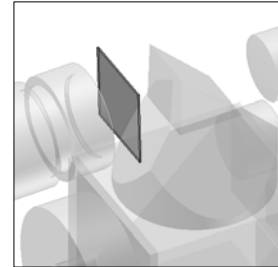
Quarter wave plates

For the Quarter Wave Plates (QWPs) crystalline quartz is chosen because of its stability and because durable AR coatings can be applied to it. Furthermore to achieve good temperature stability and ensure quarter wave retardance for both the DCS and IF beam wavelengths, compound zero order QWPs are used. The plates are cut to custom size from stock substrates. The QWP located in the supply beam of the interferometer is made larger (see picture on the right) so that it converts the DCS beam into circular polarized light.
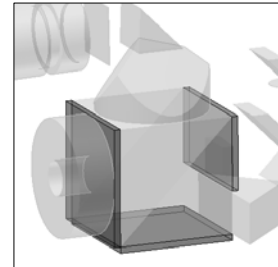
The QWPs have $\lambda/100$ retardation accuracy, and are designed for 633 nm: a standard wavelength. For the DCS beam, this causes an extra retardation error of $\lambda/360$, which is below the production tolerance of the wave plates. The specification for the irregular surface error is $\lambda/10$ for the external surfaces, leading to an irregular wavefront error of $\lambda/20$.

Differential confocal folding prism

The preferred material for the components of the optical monolith is BK7, since it is a commonly used, cost effective material, that provides a good match with crystalline quartz regarding refractive indices, as well as its linear thermal expansion coefficient.
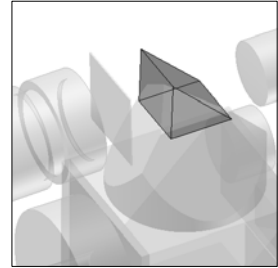
The DCS folding prism is made from a larger right-angle prism that has been cut to size and reworked. The internal reflective surface is the most critical and its irregular surface shape error is specified as $\lambda/23$ over the clear aperture. This causes an irregular wavefront error of about $\lambda/8$. The external transmissive surface is specified to have a surface shape error of about $\lambda/5$, giving an irregular wavefront error of $\lambda/9$.

Corner cube

The corner cube is a 19 mm diameter catalog component. The vertex is removed by grinding after which the surface is polished to optical quality. The transmitted wavefront error of the beam splitter cube is specified as $\lambda/4$. An advantage of the high stability interferometer layout (Subsection 5.1.3) is that both the reference and measurement interferometer beams are reflected at the same locations on the surfaces of the same corner cube. Hence, the wavefront errors added by the corner cube are identical for both branches, and thus less critical because modulation depth is not affected.
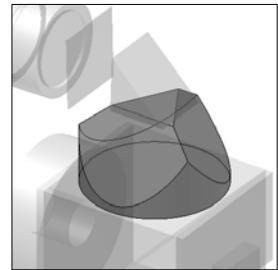
Polarizing beam splitter cube

A laser-line Polarizing Beam Splitter (PBS) cube of 20 mm edge length with an extinct ratio of 1000:1 is used. Since components will be bonded to the PBS at all external optical surfaces these should be uncoated. Therefore, the AR coatings of a catalog PBS cube are removed by polishing. The specification for the transmitted wavefront error is $\lambda/4$ over the clear aperture. In this application though, the external surface errors will
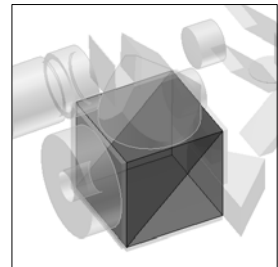
not contribute to the wavefront error because in the optical monolith they are internal surfaces. Thus, only the beam-splitting surface will add to the wavefront error, and only in reflection. Furthermore, the specified wavefront error contains the regular component and is specified over the full clear aperture. For the DCS beam, just the irregular wavefront error over a much smaller area is of importance and is estimated to be below $\lambda/10$. For the interferometer, however, the regular surface shape error of the splitter surface causes the wavefronts of the measurement and reference beams to make an angle. If the whole specified wavefront error of the beam splitter is assumed to be due to regular surface shape error of the splitter plane and the interferometer beam pitch of 13 mm is taken into account, a relative phase variation between measurement and reference beam of $\lambda/5$ is found.

Interferometer right-angle prism

The interferometer right-angle prism has been made from a 15 mm right-angle prism, which has been cut to size. The surface shape is less critical, because the interferometer reference and measurement beams are recombined when they are reflected by this component, and therefore suffer the same wavefront errors. The peak to valley surface shape can thus be specified at $\lambda/1$, which is relatively easy to manufacture.
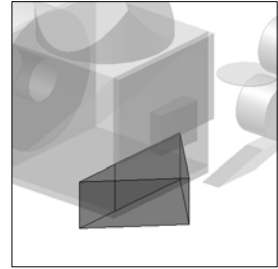
Interferometer reference mirror

The interferometer reference mirror is made from a standard circular mirror with a diameter of 20 mm and a thickness of 6 mm. A 5.2 mm hole is drilled through the center of the blank to enable alignment of the DCS beam and to allow leak light of the DCS beam escape the system. Since this component is bonded directly onto the optical monolith, the shape error of the reflective surface has been specified to be $\lambda/20$, which gives a wavefront error of about $\lambda/7$.
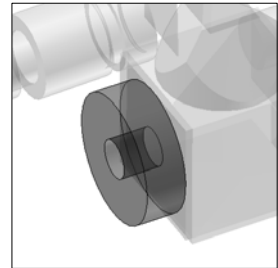
Interferometer measurement mirror

The interferometer measurement mirror is made from a window blank because a low parallelism error between front and back is desired. It is 20 mm in diameter and 6 mm thick. A hole is drilled through the center and the edge of the back surface has a 5 mm, 45° bevel applied to it to allow measurement into concave optics. The thermal expansion over the thickness of the mirror will lead to measurement errors since the objective lens is mounted on

124
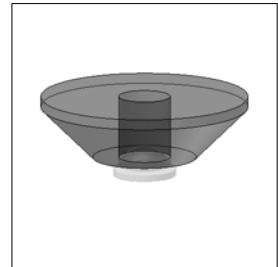
the back surface of the mirror. Hence a fused silica window is selected which has a low coefficient of thermal expansion of $5.5 \cdot 10^{-7}$ m/m/K. The resulting measurement error is 0.3 nm for a temperature change of 0.1 K. Because it is used as a first surface mirror in air, the surface shape error is specified as $\lambda/10$, which gives a wavefront error of $\lambda/5$.

Wedges

The wedges have a wedge-angle of 4° and are custom made because of their unusual dimensions of 7 mm by 7 mm and 7 mm by 6 mm. These dimensions are necessary to provide enough clear aperture for the DCS beam while fitting between the supply and returning interferometer beams. Since the wedges are used in transmission, they are specified to have an irregular surface shape error of $\lambda/5$, leading to a wavefront error of $\lambda/10$.
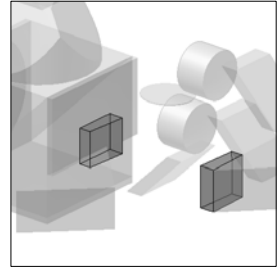
Dove prism

The dove prism is made by modification of the other part of the 15 mm right-angle that was cut up to make the interferometer right-angle prism. It has two internal reflecting surfaces that are specified to have an irregular surface shape error of $\lambda/23$, and one external transmissive surface that has an irregular surface shape error of $\lambda/5$. Assuming the irregular surface shape errors are uncorrelated, the dove prism introduces a wavefront error of $\lambda/5$.
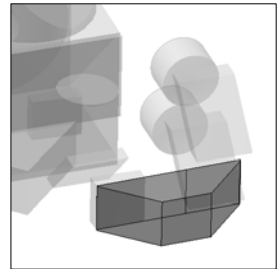
PSD splitter plate

The splitter plate for the Position Sensitive Detector (PSD) is custom made because no standard component was found which provides a clear aperture that is large enough while fitting in the available space. The plate is made of fused silica and measures 8 mm x 11.3 mm x 2 mm.
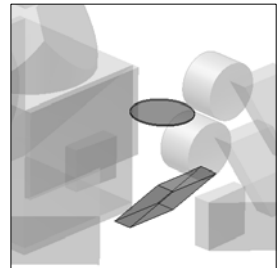
To sample the beam, the front surface is left uncoated so that at 45° incidence 8.2% of the s-polarized light and 0.7% of the p-polarized light is reflected to the PSD. There are no AR coatings of which the reflectance is close to 0 % for both the s and p polarizations at 45°. The AR coating on the backside of the plate must thus be optimized for the p or the s polarization or for overall performance. Interference of ghost reflections with the measurement spot on the PSD causes measurement errors, therefore it has been decided to optimize the AR coating for the s-

polarization and block the p-component with a polarizer in front of the PSD. The polarizer plate is a Codixx Colorpol that is slightly tilted. The wavefront aberration for the reflected beam is not critical and negligible for the transmitted beam.

Differential confocal splitter plate and folding mirror

To keep cost down, the DCS splitter plate and folding mirror have the same dimensions and surface shape specifications as the PSD splitter plate, so that they can be manufactured simultaneously.

A laser-line non-polarizing coating has been applied to the front side of DCS splitter plate, while the backside is AR coated for best overall performance. The wavefront error is $\lambda/12$ for the reflected beam and is negligible for the transmitted beam.
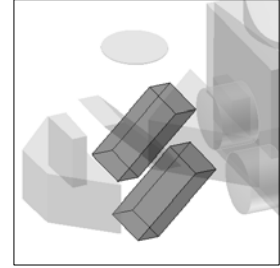
The 45° mirror that folds the reflected branch parallel to the transmitted beam is also identical to the PSD splitter plate but is coated with protected aluminium. The introduced wavefront error is $\lambda/12$.

5.6 OPTICAL SYSTEM OVERVIEW

Here some of the optical properties of the system are calculated and an overview of the optical train is given.

5.6.1 OVERALL SYSTEM PROPERTIES

Since the specifications of the components are known, the overall transmissivity of the system and the accumulated wavefront error can be estimated.

Accumulated wavefront error

If the contribution to the irregular wavefront error of each component is considered uncorrelated, the total accumulated wavefront error can be estimated by squared addition. This gives a total wavefront error of $\lambda/2.5$, if the wavefront errors over the clear apertures of the components are used. In general, a wavefront with an irregular error of $\lambda/4$ or better is considered diffraction limited. Because the DCS beam is only 1.9 mm in diameter, the wavefront error of the beam is less than the estimated $\lambda/2.5$ and is likely to be under the $\lambda/4$ criterion for diffraction limited performance.

Moreover, most of the wavefront distortion is introduced by the polarizer and the components used to fold the beam. At the time of component selection, the

polarizer was thought to introduce a wavefront error of about $\lambda/10$. Polarizers that introduce less wavefront errors are available, so in a next generation sensor or an upgrade for the current sensor, the wavefront error can be lowered further relatively easily. In addition, for an application where more space is available, the components used for folding the system can be omitted, which further lowers the accumulated wavefront error.

For the interferometer system, the worst-case relative phase variation between measurement and reference beam over the diameter of the beams can be calculated to be just over $\lambda/2$. This leaves margin for alignment errors.

Overall system transmittance

The transmittance and reflectivities of most of the standard components is specified; for the custom-built components it can be calculated from theory. The overall system transmittance can then be calculated by multiplication of transmittance and reflectance of the separate components, and yields an overall transmittance of about 46%, excluding the surface under test.

5.6.2   OPTICAL LAYOUT

An overview of the optical train including the detectors and the pick-up for the interferometer can be seen in Figure 5.11 and Figure 5.12 on the next pages. Figure 5.11 shows an exploded view while Figure 5.12 shows the components as they are placed in the prototype.

The optical design is complete. It sets the boundary conditions for the optomechanical and mechatronic design, which will be addressed in the next chapter.
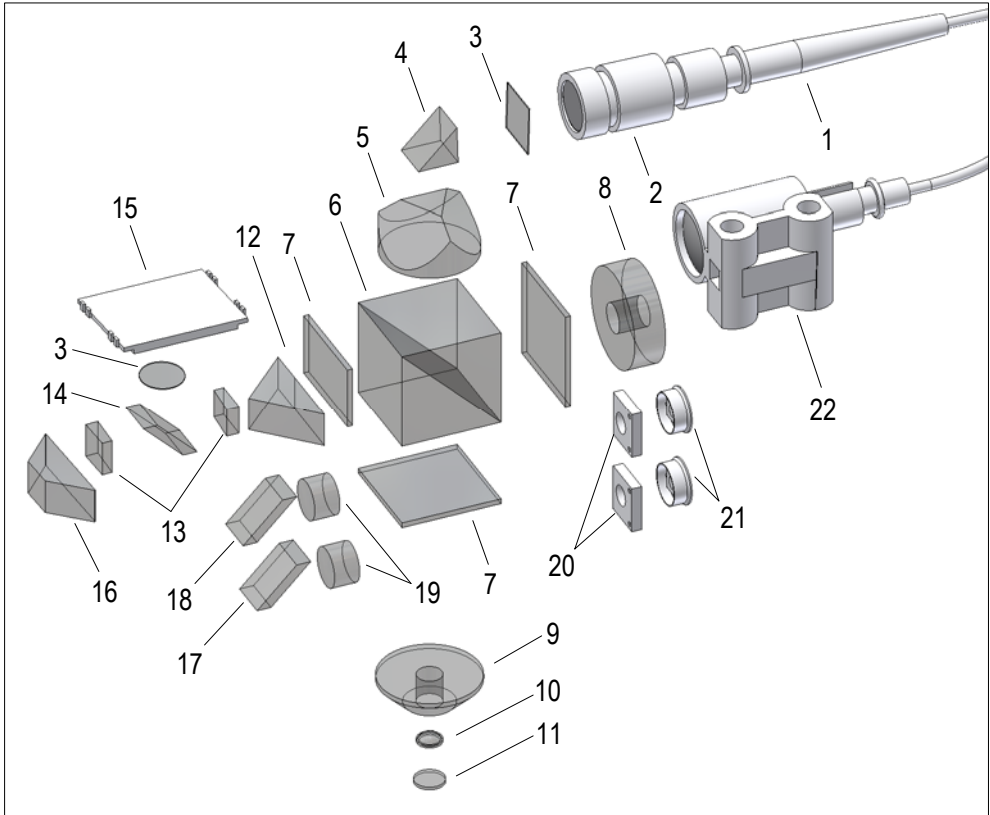
Figure 5.11: Exploded view of the optical train including the detectors and the interferometer pick up.

The components are:

1	fiber connector	12	IF right-angle prism
2	collimator	13	wedges
3	polarizer	14	PSD beam splitter
4	DCS folding prism	15	PSD
5	corner cube	16	dove prism
6	polarizing beam splitter cube	17	DCS splitter plate
7	quarter wave plates	18	DCS folding mirror
8	IF reference mirror	19	doublets
9	IF measurement mirror	20	pinholes
10	objective lens	21	photodiodes
11	cover glass	22	IF pick-up
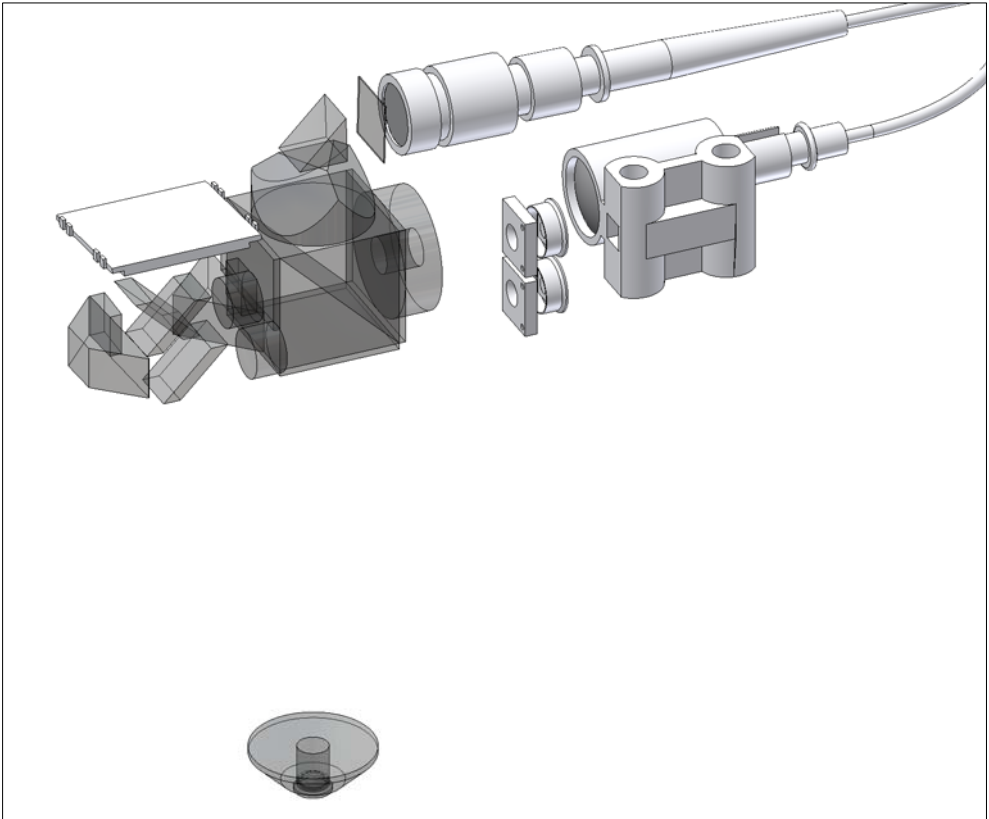
Figure 5.12: The optical train as applied in the prototype, the detectors and the pick-up for the interferometer included.

6 OPTOMECHANICAL AND MECHATRONIC PROTOTYPE DESIGN

The optomechanical and mechatronic systems are designed around the optical train. The general design considerations are discussed first, followed by a more specific treatment of the optomechanical subassemblies and the focusing unit.

6.1 DESIGN CONSIDERATIONS

Apart from wavefront errors introduced by optical components (Subsection 5.5.1), performance is also dependent on component alignment. Consequently, the primary goal of the optomechanics and mechatronics is to attain and maintain adequate positioning and orientation of the optics, both relative to each other and to the ψ-axis rotor. This must be achieved while stress in, thermal gradients over, and deformation of the optical components is kept within acceptable limits. Desirable characteristics of the opto-mechatronics that arise from abovementioned goal are:

- high dimensional stability,
- high thermal stability,
- small gravity induced deformation,
- high eigenfrequencies,
- low hysteresis,
- low parasitic movement due to adjustment and locking, and
- low dissipation for actuation over the stroke.

Dimensional stability depends mainly on the choice of materials and the design of interconnections. Most of the structure is made of aluminium because of its combination of low thermal sensitivity, i.e. its thermal expansion coefficient divided by its thermal conductivity, excellent machinability and low density. In addition, it can be blackened by anodization, which reduces the effects of stray light in the optical system.

To obtain high eigenfrequencies and low gravitationally induced deformation, a high stiffness to mass ratio is desired. Due to the low mass of the majority of the components in this application, this is not expected to be an issue for most components and subassemblies.

To ease machining of bonding features and facilitate handling and manipulation, the optical components are mainly bonded to mounting rings or holders, which are then mounted to the structure. If it is feasible to achieve sufficient tolerances during fabrication, alignment using component features is favored over manipulation. This keeps assembly efforts limited but involves tightening some of the components' dimensional requirements. When manipulation is necessary, external manipulators are applied and components are locked in place once they are aligned. The advantages are that it saves space and that long-term manipulator drift does not influence performance.

In general, the use of adhesives for locking is preferred over mechanical locking mechanisms, because it is more compact and requires less complicated mechanics. For critical Degrees Of Freedom (DOFs), the bond interface must be designed such that when the adhesive cures, shrinking does not lead to significant component drift. To achieve this, the interface is designed such that the shrinkage of the adhesive pulls two surfaces, which determine the position, into contact, or the adhesive is placed symmetrical around a component so that shrinkage in one direction is compensated by shrinkage in other directions.

Five situations are identified regarding alignment of components and subassemblies, namely where the DOFs to be constrained are:

- determined by component features after which the part is permanently fixed,

- initially manipulated after which the part is permanently fixed,

- initially manipulated after which alignment is permanently locked, while the aligned subassembly can be detached and reattached,

- re-adjustable and re-lockable, and

- actively controlled during operation.

Components that have to be permanently fixed are bonded with adhesive. For some bonds, facilities have been installed to enable breaking the bond if unforeseen troubles arise. This can be a pocket so that a lever can be inserted between the parts, or a threaded hole through one of the components allowing the components to be pushed apart with a screw.

For components that must be easy to replace or re-adjust, mechanical locking mechanisms will be used. The locking mechanisms are designed to prevent parasitic movement during clamping and incorporate flexures.

It must be possible to remove the sensor from the ψ-axis rotor and to re-install it while alignment is maintained. The same applies to the focusing unit, which contains the objective lens, cover glass and interferometer measurement mirror.

Kinematic couplings are well suited for connections where high repeatability is desired. In this application, the load on the interface is expected to vary significantly, because the relatively heavy focusing unit is rotated around the ψ-axis, and because dynamic reaction forces are generated by the focusing unit. The need for high stiffness and initial adjustability of multiple DOFs complicates the use of kinematic couplings within the available volume envelope. To circumvent the problem, semi-kinematic couplings are applied in which point contacts are substituted by small contact areas. Measures are taken to achieve good parallelism between the contact areas on both sides.

The only actively controlled DOF is translation along the optical axis of the interferometer mirror, objective lens and cover glass. This will be provided by the focusing unit, which incorporates a guidance mechanism and an actuator.

The guidance is a flexure-based design because of requirements such as: high precision, high bandwidth and environmental issues, as well as experience in this field. A voice coil with stationary magnets is used as actuator, due to the high bandwidth, low moving mass and high electromechanical efficiency it can provide.

6.2 OPTOMECHANICAL SUBASSEMBLIES

Before the various subassemblies are treated in depth, an exploded view of the final design in which the various subassemblies are identified is presented in Figure 6.1.
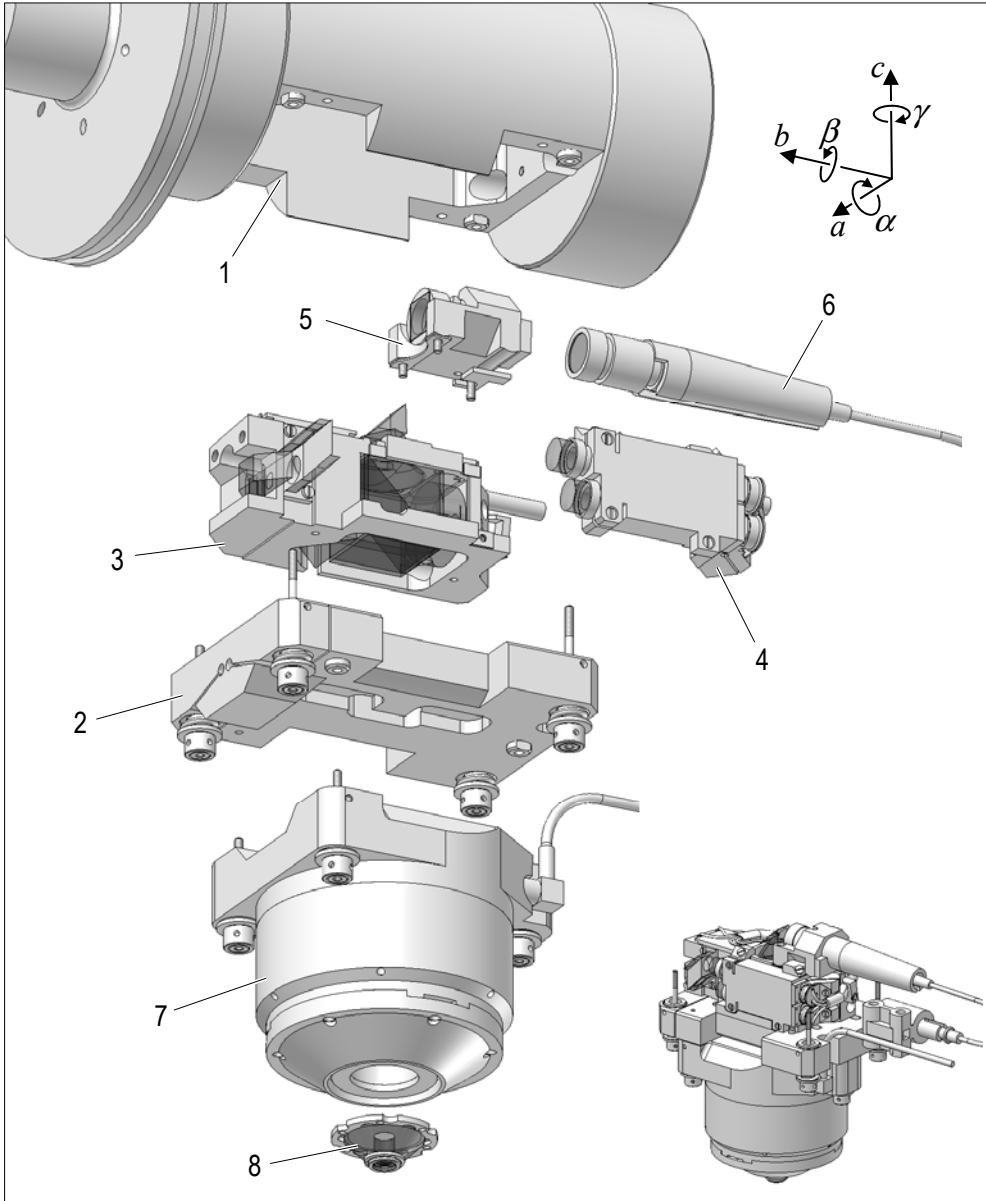
Figure 6.1: Exploded view showing the subassemblies of the optomechatronic system.

The ψ-axis rotor (1) has a pocket machined into it, which is closed off by the interface plate (2). Inside this pocket, the central optics unit (3) is bonded on top of the interface plate. The central optics unit contains most of the optics and forms the basis on which the differential confocal unit (4) and collimator unit (5) are

mounted. The collimator (6) is inserted in the collimator unit through a bore in the ψ-axis rotor front-end, after the sensor is installed. On the other side of the interface plate the focusing unit (7) is mounted, which holds the objective lens subassembly (8) consisting of the interferometer mirror, objective lens and cover glass.

6.2.1   CENTRAL OPTICS UNIT

The central optics unit, (3) in Figure 6.1, is composed of the central optics holder, a monolithic aluminium structure, which holds most of the optical components and subassemblies; it is shown in Figure 6.2.
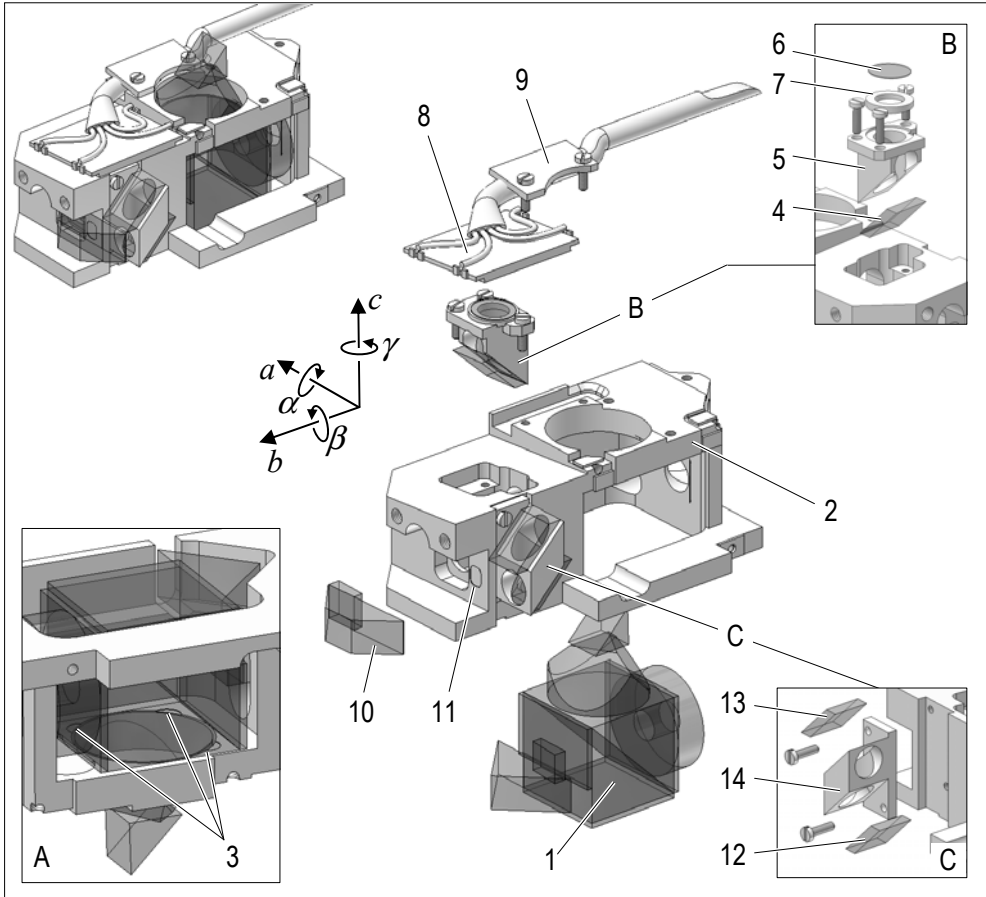
Figure 6.2: Overview of the central optics holder and the components it holds.

The optical monolith (1) is bonded to the central optics holder (2) with epoxy. To obtain alignment and a stable interconnection, the adhesive is applied in three 0.1 mm deep recesses (3) as depicted in Figure 6.2 in detail A. The 0.1 mm thick adhesive layer shrinks during curing, pulling the beam splitter into contact with the mounting surface. This prevents alignment errors due to non-uniformity in the adhesive layer or adhesive creep. Angular alignment in α- and β-direction is obtained by contact between the mounting surface and the optical surface. The *c*-position is not critical since it is compensated with other optics. Positioning in *a*- and *b*-direction is obtained by shims, which are later removed. Rotation in γ-direction is inspected using a traveling microscope and is manually adjusted.

An exploded view of the Position Sensitive Detector (PSD) splitter subassembly is shown in detail B of Figure 6.2. The PSD splitter plate (4) is bonded to an

aluminium holder (5) by two adhesive lines along the long edges. Alignment of this component is not critical since beam position on the PSD is inherently compensated for by calibration. The polarizer (6) at the top of the holder is bonded to a mounting ring (7) which fits into the countersunk bore through the PSD splitter holder. The top surface of the mounting ring is tilted -5° in β-direction to prevent parallel ghost reflections. Once rotational alignment is obtained, the ring is locked by three adhesive dots around the perimeter. The subassembly (B) is mounted in a pocket in the central holder by means of three screws. While the screws are tightened, a-, b- and γ-alignment is determined by three contact areas in the pocket, to which the holder is pushed with a lever.

The PSD (8) is bonded to the central optics holder using epoxy at various places around the perimeter. The PSD mounting surface of the central optics holder is tilted 5° in β- direction to prevent parallel ghost reflections from the silicon surface and protection window of the PSD. Alignment is not critical for this component because it will be inherently compensated for during calibration.

The PSD signal cable is bonded to the PSD, and the soldering bonds are covered with epoxy. This protects the soldering bonds, which are prone to progressing creep and eventually failure if they are under tension. As the epoxy has a high electrical resistance, it also prevents short-circuiting. The PSD cable is clamped in a slot in the central optics holder by a plate and two screws (9) to provide protection from tension in the cable during mounting and un-mounting of the sensor in the ψ-axis rotor.

The dove prism (10) is bonded directly to the central optics holder with epoxy applied in a 0.1 mm deep recess (11). This provides stable mounting in the same way as for the optical monolith (1 in Figure 6.2). The DOFs in a- and β-direction are manipulated to align the laser beam to the differential confocal unit in a- and c-directions. The relatively narrow contact area might lead to slight misalignment in α-, and γ-direction. These DOFs are however not critical because the system is inherently insensitive to γ-misalignment due to the 90° angle in γ-direction between the reflective surfaces, while α-misalignment is compensated by β-adjustment of the dove prism in combination with c-adjustment of the pinholes.

An exploded view of the differential confocal splitter subassembly is shown in detail C of Figure 6.2. The non-polarizing beam splitter plate (12) and differential confocal folding mirror (13) are bonded to an aluminium holder (14) by adhesive lines along their long edges. The distance between the reflective surfaces determines the pitch between the two differential confocal beams and must thus correspond to the pitch of the bores in the differential confocal unit. For this purpose, machining tolerances are precise enough since pinhole alignment in a- and c-directions will compensate beam misalignment. The system is insensitive to

misalignment of most DOFs of this subassembly due to the two parallel reflective surfaces oriented at 45° relative to the laser beam. The only DOF for which the system has first-order sensitivity is rotation around β, which, due to the wide holder base in c-direction, has good reproducibility. The base of the holder fits in a slot in the central optics holder. To determine α-rotation when the screws are tightened, a lever is used to push the holder into contact with the side of this slot.

6.2.2 DIFFERENTIAL CONFOCAL UNIT

The main structural component of the differential confocal unit (Figure 6.3) is the differential confocal holder (1): a monolithic aluminium part to which the imaging lenses, pinholes and photodiodes are mounted.

Because forces are exerted on the subassembly during pinhole alignment and locking, the connection between central optics holder and differential confocal holder must have sufficient stiffness to maintain alignment. For testing, it is convenient if removing and reinstalling the differential confocal unit requires little pinhole realignment. Reproducible mounting is however less important than sufficient stiffness, as the latter is critical to achieve proper pinhole alignment.

Since both the pinholes and the imaging lenses are held by the differential confocal unit, the system is sensitive to angular misalignment in α- and γ-directions while translations in a-, b- and c-direction and rotation in β-direction are less critical. No alignment is needed for any of the DOFs. Because of the aforementioned combination of requirements and the limited space available, a semi-kinematic mount is applied. The contact surfaces are defined by six plates (2) which are bonded to the differential confocal holder; the connection to the central optics holder is made using three screws (3). The plates are cut from a 0.5 mm thick, steel gauge block by wire EDM. The screws are positioned and oriented such that their axes intersect the collective center of the surfaces they load, and are parallel to the resultant of the normal vectors of the respective surfaces. The screw located at the three plates is tightened first, the screw at the two plates next and the screw at the single plate last. This subsequently constrains translations in a-, b- and c-direction, rotation in α- and γ- direction and rotation in β-direction, respectively.

The pinholes must be aligned relative to the foci in lateral direction, a- and c-, and in focus direction, b. To facilitate manipulation and locking, it has been decided to accommodate these alignments by focus adjustment of the lenses and lateral adjustment of the pinholes. An exploded view of the imaging lens mounting is shown in Figure 6.3, detail A. The imaging lenses (4) are bonded to lens holders (5) which have a position fit in bores in the differential confocal holder of which a 90° section is cut free by a radial and axial cut (6). The lens holders are clamped in the bore by a screw located close to the lens holder.
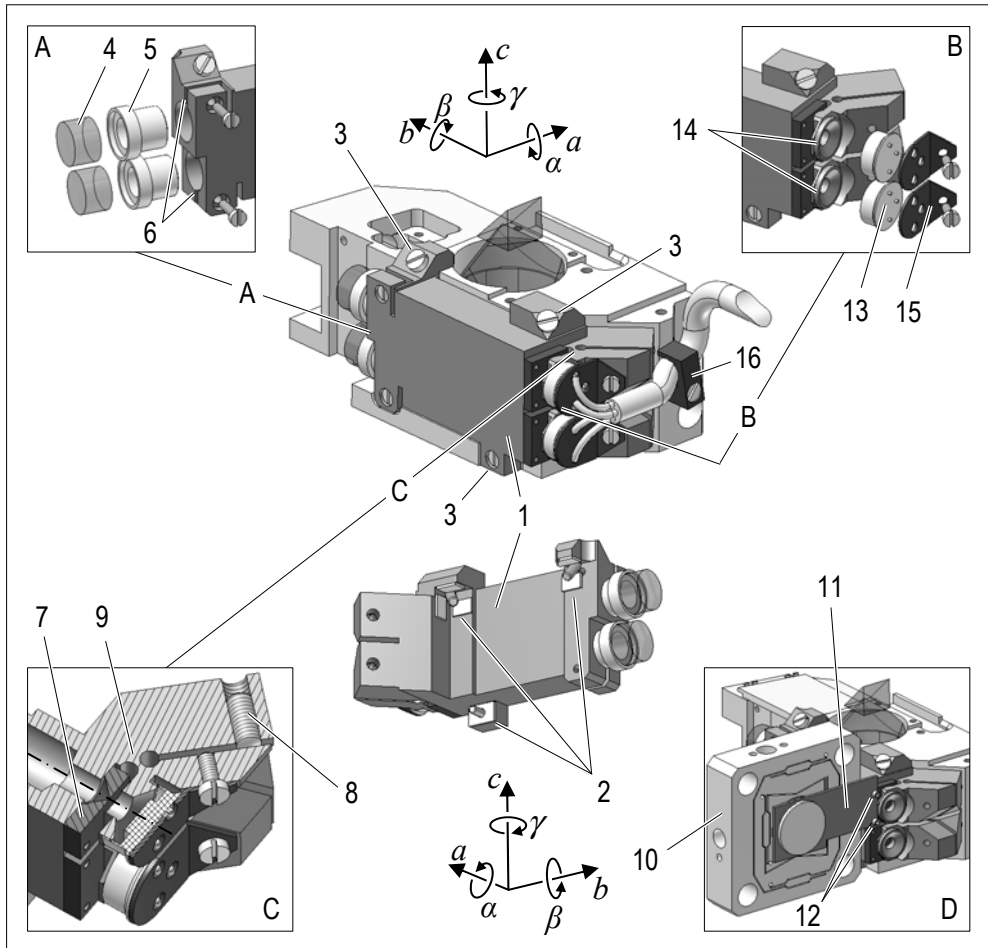
Figure 6.3: The differential confocal unit.

Detail B of Figure 6.3 shows an exploded view of the detector mounting. The pinholes (7) in detail C are contained in 2 mm thick aluminium plates, which are clamped between a flat and a cylindrical surface with a lever mechanism. The cylindrical contact surface has its axis parallel to the *c*-direction and has a bore through it, which is coaxial with the optical axis. When a push screw (8) is tightened, the lever rotates around an elastic hinge (9), which reduces transfer of parasitic forces and moments to the pinholes. The mechanism is made by milling followed by wire EDM in one direction. The screw can be reached through a hole in the central optics mount.

To align a pinhole, the lever mechanism is slightly tightened so that the pinhole slides over the flat surface of the differential confocal mount while an external

manipulator (10) is used for adjustment in *a*- and *c*-direction. This is the same manipulator as is incorporated in the demonstrator: it has two nested flexural guides consisting of plate spring parallelograms, which are adjusted by two push-pull screw pairs. A plate spring (11) connects the manipulator to the pinholes with two M1 screws (12). A plate spring parallel to the *a-c* plane is ideal for this purpose since it only constrains *a*, *c* and *β*, i.e. the DOFs that are not constrained by the contact surface of the differential confocal mount. Although the manipulator is designed for use around its center, the off-center usage here, worked quite good too.

The windows of the photodiodes (13) in detail B are removed on a lathe with a facing cut through their metal housing. The back surface of each photodiode is bonded to a metal plate with holes providing passage to the photodiode lead-pins. The photodiodes are each placed on a boss (14) on the lever ends, and the plates are fastened with screws. The photodiode signal cable is soldered to the photodiode pins and epoxy is applied over the solder bonds for protection. To prevent pulling on the differential confocal unit during mounting and un-mounting of the sensor, the cable is clamped onto the central optics holder using a clip (16).

6.2.3 COLLIMATOR UNIT

An exploded view of the collimator unit is shown in Figure 6.4. The main structural component is the collimator holder (1), containing the collimator and a polarizer.
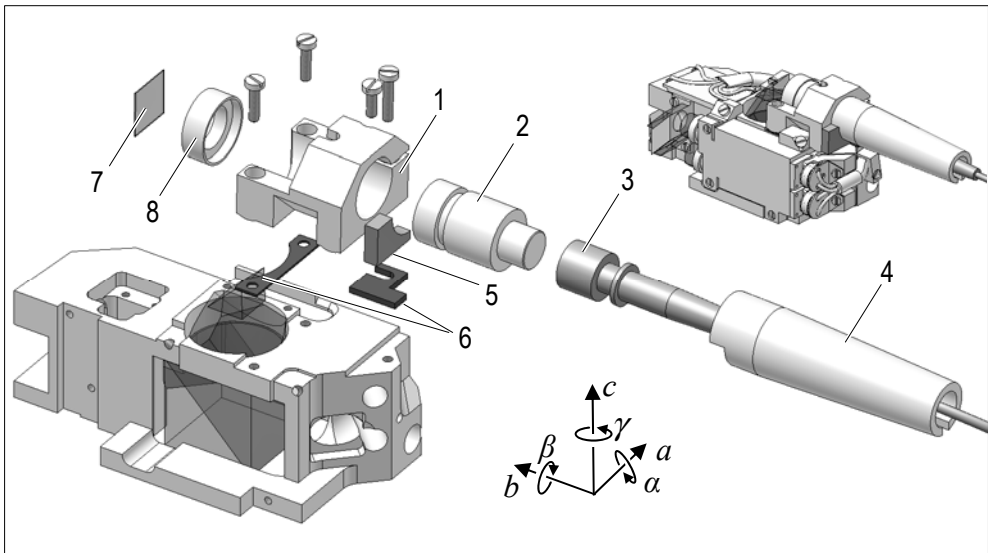
Figure 6.4: An exploded view of the collimator unit.

The collimator is supplied including a standard subcell (2) that incorporates collimation adjustment. It is inserted into a location-fit bore in the collimator holder; a radial cut and screw enable locking.

Connecting the fiber coupling (3) to the collimator, after installation of the sensor and collimator into the ψ-axis rotor, leads to unacceptable variation in beam pointing. Instead, the collimator and fiber are bonded together at the fiber coupler using epoxy and the collimator is removed from the assembly during installation. Once the sensor is installed in the ψ-axis rotor, the collimator is inserted into the collimator holder through a hole in the rotor. For this purpose, a placement tube (4) is bonded to the back of the collimator. It serves as a handle during insertion and as an axial stop since its diameter is larger than that of the collimator. A 90° section is removed from the front of the placement tube so that, once the polarization is aligned to the beam splitter cube, collimator rotation can be indexed by bonding a stop (5) to the collimator mount.

The remaining DOFs of the laser beam that must be aligned are a, c, α and γ. Three screws in oversized countersunk bores connect the collimator unit while allowing it to move in a-, b- and γ-direction. Translation in b-direction has no influence on system performance; a and γ are adjusted with micrometer heads which are temporarily fixed to the central optics mount.

The remaining DOFs, c and α, are aligned using shims (6), which is a time-consuming process. This is considered acceptable, because the adjustment has to be made only once. To allow α to be adjusted with shims, the bottom surface of the collimator holder has two cylindrical surfaces with their axes parallel to the a-direction. When all DOFs are adjusted, the three screws are tightened, locking the collimator unit in place.

The collimator polarizer (7) is mounted to a mounting ring (8) of which the front surface is tilted 5° in α-direction to prevent parallel ghost reflections. The mounting ring sits in the half bore at the front of the holder, and is thus easily removed and rotated, which is convenient during tests. The polarizer is locked by applying three adhesive dots at its edge.

6.2.4   INTERFACE PLATE

The interface plate (1) in Figure 6.5, is the main structural component connecting the central optics unit and focusing unit with each other and with the ψ-axis rotor. It is made of a 12.5 mm thick aluminium plate and contains no optics.
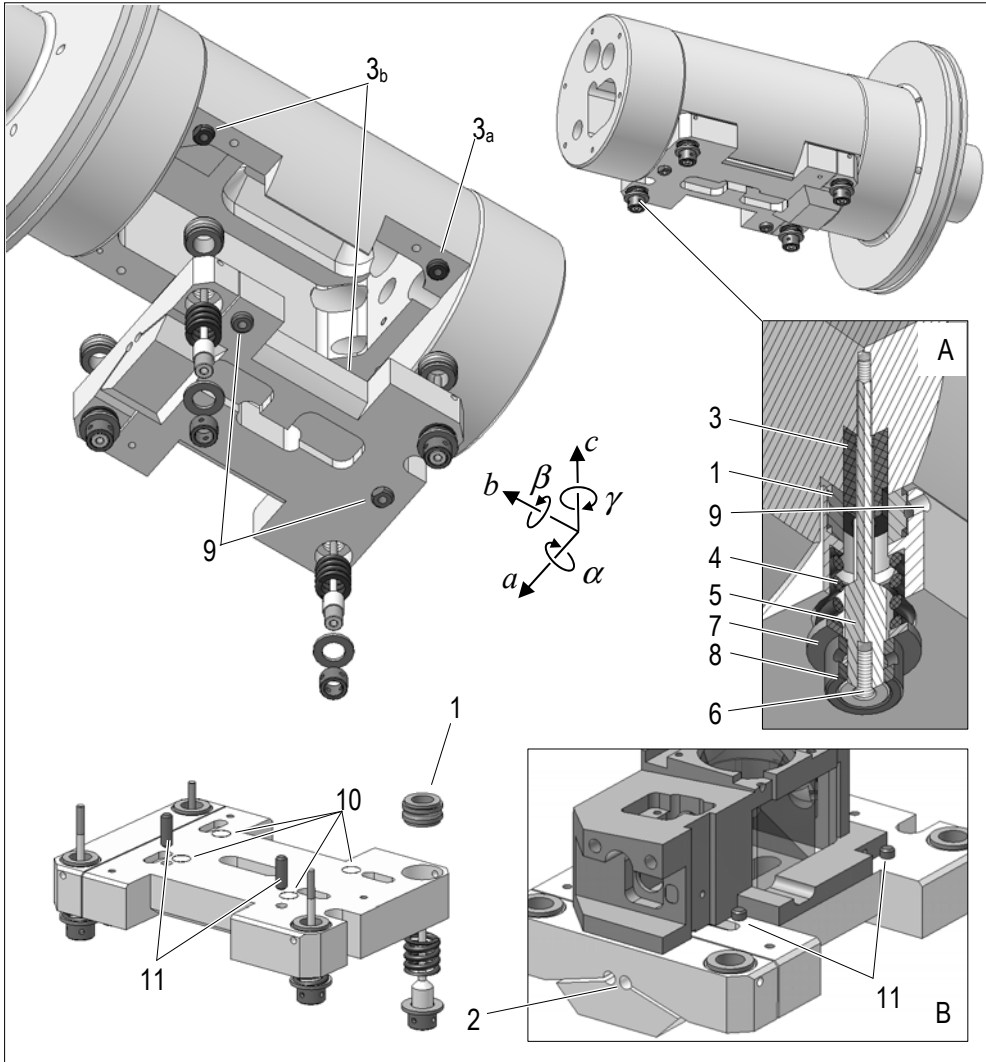
Figure 6.5: Interface plate with semi-kinematic mounting system.

For reasons mentioned in Section 6.1, a semi-kinematic coupling is used between the interface plate and ψ-axis rotor. It consists out of four hardened steel bushings (1), placed into holes in the interface plate where they rest on a flange. Their top surfaces are slightly higher than the surface of the plate and form the contact-areas that constrain the c-, α- and β-DOFs. By cutting a lever and a flexural hinge (2) in the plate, a whiffletree is obtained which prevents overconstraining. Three dowel pins are placed in the rotor: one with a circular circumference (3_a) and two that are flattened (3_b). The flattened dowel pins are oriented such that they do not constrain

the pitch of the bushings. Because of the positional fit of the dowel pins to the bushings, the circular dowel pin and one flattened dowel pin constrain the a-, b- and γ-DOFs. The second flattened dowel pin constrains the internal DOF of the rotor, which has decreased torsional stiffness due to the large pocket.

The outside of the ψ-axis rotor is the measurement mirror for the NANOMEFOS interferometers and its shape must consequently reproduce to nm level when the sensor is removed and re-inserted. Therefore, and to attain reproducible positioning of the sensor, care is taken to apply a reproducible preload force to the bushings while minimizing parasitic forces and moments introduced to the rotor. A cross-section of the configuration to achieve this is shown in Figure 6.5, detail A.

Generally, tension springs are convenient for preloading because they are self-aligning, experience almost no friction and have well-defined points of application. In this situation, however, they are difficult to apply. Hence, compression springs (4) are used here, providing a nominal preloading force of 212 N per bushing. Stepped stud bolts (5) with external M2 thread on one side and external M5 and internal M2 thread on the other side, are used as pull rods to compress the springs. The thin ends of the stud bolts stick through oversize bores in the dowel pins (3), and are screwed into the ψ-rotor where they are locked with epoxy. The usual threadlocking products are avoided since they are known to pollute optical surfaces by out-gassing.

The interface plate containing the bushings (1) is threaded over the stud bolts and dowel pins after which a custom tool is used to compress the springs. The tool screws into the internal M2 thread (6), and pulls the stud bolt while it pushes a washer (7) against the compression spring. A nut (8) is then screwed onto the M5 thread, thus retaining the washer and spring when the tool is removed. In this way, little torsion is introduced into the ψ-rotor.

Alignment of the sensor to the ψ-axis in α-direction is adjusted by lapping the bushings to thickness. The holes in which the bushings are placed are oversized, which allows alignment in a-, b- and γ-direction. Once the sensor is aligned to the ψ-axis, the bushings' positions are fixed by injection of epoxy through a channel (9) in the plate. This automatically sets the pitch of the bushings to that of the dowel pins, and enables locking of alignment while the sensor can be reproducibly removed and reinstalled.

The central optics holder is bonded to the top of the interface plate with adhesive applied in four 0.1 mm deep recesses (10). Pockets are milled alongside the basins to allow breaking the bond with a lever. To align the central optics holder to the interface plate in a-, b- and γ-direction two dowel pins (11) are press-fitted in the plate.

6.3 FOCUSING UNIT

The focusing unit has to translate the objective lens subassembly along the optical axis with a resolution of tens of nanometers over a stroke of 5 mm, while constraining its other DOFs. As tracking errors lead to increased measurement uncertainty, a high bandwidth is desirable. Therefore, the first parasitic eigenfrequency must be high, as well as the ratio of actuator force available for acceleration, to moving mass.

Since thermal disturbances also lead to an increase of measurement uncertainty, low dissipation is desired as well. To achieve this, the ratio of the force required to overcome stiffness, damping and inertia to the motor constant, must be low.

Because of the interdependence between actuator and guidance properties, and because they compete for available space, an integral design is desirable. Nevertheless, after an overview of the focusing unit, the two subsystems will be treated in separate subsections for clarity's sake. In Figure 6.6, the main constituents of the focusing unit are identified.
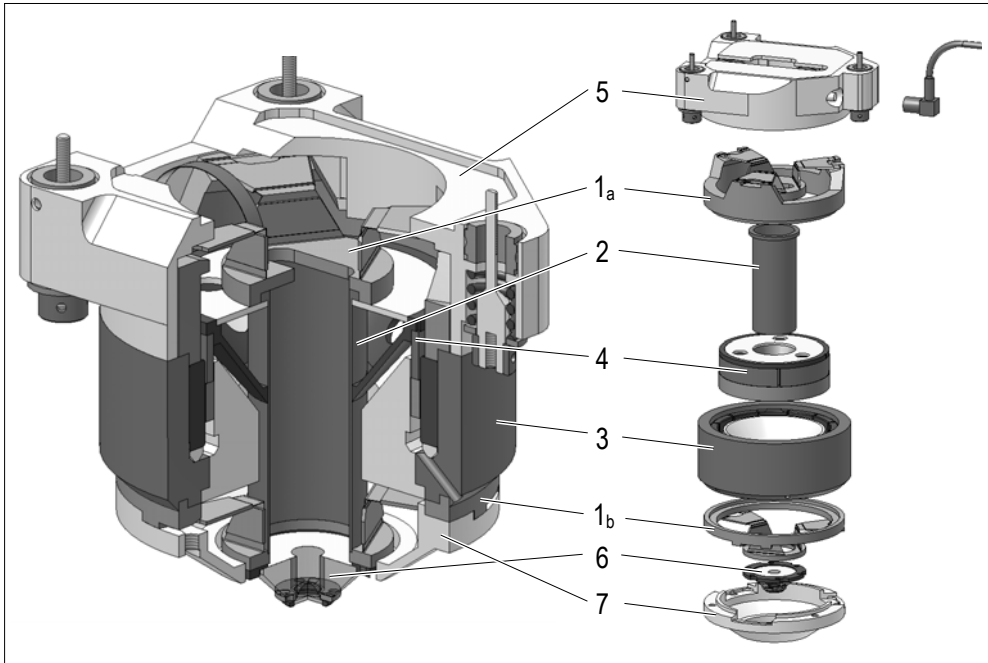
Figure 6.6: Overview of the focusing unit design.

The guidance mechanism is a flexure-based design consisting of two TiAlV monoliths (1_a and 1_b) which contain three folded plate springs each, and are placed

at the top and bottom of the available design space. The moving parts of the monoliths are linked by a concentric tube, which surrounds the optical paths. A voice coil actuator, consisting of a field assembly (3) and a coil assembly (4) is placed in between the two monoliths. A voice coil actuator is preferred for this kind of application because of various reasons: it is suitable for direct drive, can achieve high bandwidth response, is rotationally symmetric around the driving direction and, due to low hysteresis and absence of cogging, can achieve exceptionally high positional dynamic range. The outer part of the field assembly forms the structural link between the motionless parts of the guidance; the coil assembly drives the aforementioned concentric tube.

At the top, the stationary part of the guidance mechanism is attached to a holder (5) which forms a detachable semi-kinematic mount to the interface plate. At the bottom, the objective lens subassembly (6) is mounted to the moving part of the guidance, and a protective cover (7) is bonded to the motionless part of the guidance.

6.3.1 GUIDANCE MECHANISM[4]

The guidance must allow 5 mm stroke. The lateral movement will be calibrated and the lateral accuracy after calibration must be better than 50 nm.

Guidance method

Given the repetitive short-stroke nature of the movement, the required precision and bandwidth, and the environmental requirements, three guidance methods can be applied:

- air bearings,
- flexural bearings, and
- magnetic bearings.

Magnetic bearings are abandoned because of the considerable design and development effort involved. After an initial investigation, a flexural bearing method is favored, not so much for fundamental reasons, as for the available experience in this field, both in the area of design as in fabrication.

Concept design

It is expected that for most freeforms, the largest part of the power of the trajectories is contained below the first eigenfrequency of the guidance. Therefore,

---

[4] *The guidance design is largely based on the BSc thesis of S.K. Ravensbergen (Ravensbergen, 2006).*

and because flexural guidances display low damping, most of the driving force is required to overcome stiffness in driving direction. Consequently, the ratio of the guidance stiffness to the actuator's motor constant is an important design criterion.

Various flexure-based guidance mechanisms are discussed in (Rosielle and Reker, 2000); two concepts that have been evaluated for this application are schematically depicted in Figure 6.7. One concept is based on a double parallelogram in conjunction with a 1:2 lever to divide the stroke (van Seggelen, 2007), while the other consists of two sets of three folded plate springs (Broers, 1999).
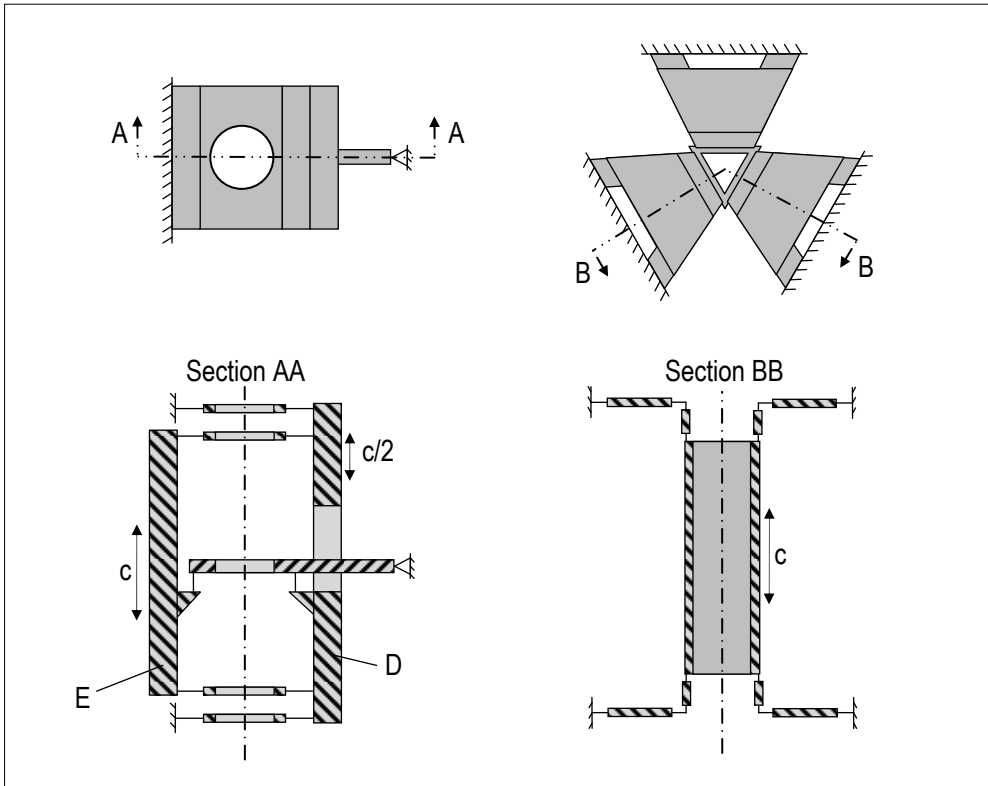
Figure 6.7: Two guidance concepts based on flexures; a double parallelogram with a 1:2 lever (left) and a six folded plate spring mechanism (right).

In the double parallelogram design, the lever mechanism constrains the ratio of movement of bodies D and E to be 1:2. Because of this, the deflection of the plate springs between body D and the fixed world is equal to that of the plate springs connecting body E to body D. Therefore, the shortening of the two plate spring sets over the stroke compensate each other.

In the folded plate spring design, each folded plate spring constrains a degree of freedom parallel to its fold line. The long parts of the folded plate springs facilitate axial movement while the short parts accommodate shortening of the long plate springs when deflected.

A large guidance base is desirable to obtain high precision motion; long flexures are desirable to accommodate the relatively large stroke. Combined with the need for a compact build, this causes the actuators to be best located between upper and lower plate spring sets for both concepts. For the parallelogram design, two flat Lorentz-type actuators can be placed at both sides of the optical axis between the two moving bodies. For the folded plate spring design, a voice coil located at the middle of the moving body seems best suited.

The rotation of the moving body around the optical axis is over-constrained for both concepts. For the double parallelogram design, this is permissible because it is manufactured monolithically. For the folded plate spring design, the upper and lower plate spring sets are monolithic and aligned by a cylindrical fit and an axial plane, allowing their rotation around the guidance axis to align.

Due to the geometry of the concepts, a longer flexure length is possible in the parallelogram design than in the folded plate spring design. Furthermore, in the parallelogram design, each plate spring deflects half the stroke, in contrast to the folded plate spring design, where each plate spring deflects over the full stroke. Therefore, low stiffness in driving direction is easier to obtain for the parallelogram design. Calculations show, however, that this difference is offset by the more favorable geometry for the actuator in the folded plate spring design, leading to about equal dissipation for the two concepts.

With a double parallelogram design, nanometer straightness after calibration can be achieved over millimeter stroke (van Seggelen, 2007). In this application though, high bandwidth response is desired. The simplicity of the mechanism and the rotational symmetry of the folded plate spring design leads to superior dynamic behavior. Hence, the folded plate spring concept is selected.

Material selection

In flexure design, an often-used criterion for material selection is the strain at onset of plastic deformation, $\sigma_{p0.2}/E$. This value determines the maximal stroke a flexure can provide without plastically deforming. Here, the mechanism is intended to go through a large number of cycles, so in this case the strain corresponding to fatigue stress at the required number of cycles, $\sigma_{fatigue}/E$, might provide a better criterion.

For 200 measurements per year, with estimated average characteristics (250 mm diameter optic with 4 waves per revolution, a track-spacing of 1 mm and 5

measurements per track), the number of cycles per year is found to be $5\cdot10^5$. Using a fatigue stress for 10^7 cycles, an often-specified value, allows for a lifecycle of 20 years. Most optics to be measured are expected to require less stroke than the full 5 mm, therefore, using this selection criterion provides a safety margin.

When dynamic performance is critical, the criterion $\sigma_{fatigue}/\sqrt{E\rho}$, is a better figure of merit than $\sigma_{fatigue}/E$, since the higher this ratio is, the higher the eigenfrequencies that can be obtained for a given stroke, or the larger the stroke that can be obtained for a given minimum eigenfrequency. In Table 6.1, $\sigma_{fatigue}$, E, ρ, $\sigma_{fatigue}/E$ and $\sigma_{fatigue}/\sqrt{E\rho}$ are shown for various metals. It can be seen that Ti-10V-2Fe-3Al has the highest $\sigma_{fatigue}/E$ and $\sigma_{fatigue}/\sqrt{E\rho}$. In this application however, Ti-6Al-4V, the most common titanium alloy, is used, because of its good properties combined with its wide availability and the available experience with EDM of this material.

Material	$\sigma_{fatigue}$ @ 10^7 cycles [MPA]	E [GPa]	ρ [kg/m^3]	$\sigma_{fatigue}/E$ [-]	$\sigma_{fatigue}/\sqrt{E\rho}$ [s/m]
Ti-10V-2Fe-3Al	910	107	$4.7\cdot10^3$	$8.5\cdot10^{-3}$	41
Ti-6Al-4V	700	114	$4.4\cdot10^3$	$6.1\cdot10^{-3}$	31
AerMet 310	1030	192	$8.0\cdot10^3$	$5.4\cdot10^{-3}$	26
Magnesium AZ92A-T6	190	45	$1.8\cdot10^3$	$4.2\cdot10^{-3}$	21
Titanium Grade 4	425	105	$4.5\cdot10^3$	$4.0\cdot10^{-3}$	20
Aluminium 2014-T451	670	73	$2.8\cdot10^3$	$3.6\cdot10^{-3}$	18
Copper UNS C62500	460	110	$7.2\cdot10^3$	$4.1\cdot10^{-3}$	16
Steel 4340	505	205	$7.9\cdot10^3$	$2.5\cdot10^{-3}$	13
Beryllium S-200F	261	303	$1.9\cdot10^3$	$0.8\cdot10^{-3}$	11
Stainless Steel 410	400	200	$7.7\cdot10^3$	$2.0\cdot10^{-3}$	10
Aluminium 6061-T451	97	69	$2.7\cdot10^3$	$1.4\cdot10^{-3}$	7
Beryllium Copper 20C	165	128	$8.3\cdot10^3$	$1.3\cdot10^{-3}$	5

Table 6.1: Material properties for some of the best suitable flexure materials.

Plate springs dimensions

The plate springs have been thickened in the middle, thus increasing their eigenfrequencies, buckling strength and stiffness in radial directions. This also results in a slight increase of stiffness in the driving direction (axial). A-shaped plate springs are used since this increases stiffness in radial direction. To maintain equal effective width of the thin sections at both ends of the plate springs, and thus equal bending, and to decrease the stiffness in driving direction, the middle part of the wide side is removed (see Figure 6.8 top left).

The long part of the plate springs is 17.1 mm long and the short part 8.1 mm, with thin parts of 2.7 mm long and 0.09 mm thick (see Figure 6.8, bottom left). This gives a combined stiffness of $3.6 \cdot 10^6$ N/m in radial direction and $1.3 \cdot 10^3$ N/m in driving direction. The stress for maximum displacement is 320 MPa. Hence, the stress reversal over the full stroke is 640 MPa, just below the allowable fatigue stress of 700 MPa. The total moving mass including coil assembly and objective lens assembly is 46 g.
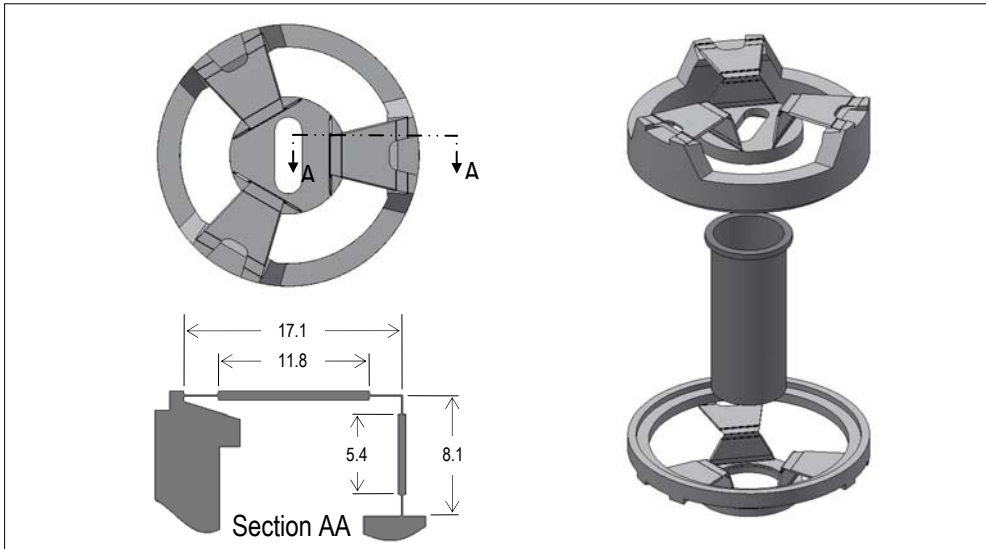
Figure 6.8: Plate spring dimensions (left) and exploded view of guidance mechanism (right).

6.3.2   VOICE COIL ACTUATOR

When the sensor is vertically oriented and the objective is in the highest position, the stiffness of the guidance and the weight of the moving parts require an actuation force of 3.6 N. To reserve enough force for acceleration and some margin for safety, the design goal is to obtain an actuator that can generate a nominal force of 5 N at least, combined with minimum dissipation and a high bandwidth.

Actuator topology

If the average magnetic flux density through the coil varies over the stroke, so will the force factor, complicating stable, high bandwidth control. To avoid this, the coil must experience approximately constant average magnetic flux density over the stroke. This is satisfied either by a long coil/short gap topology, where the coil extends beyond the magnetic air gap over the entire stroke, or by a short coil/long gap topology, where the coil stays within the uniform part of magnetic gap over the

entire stroke. In general, the short gap allows for higher magnetic flux density in the air gap, however a large part of the coil does not contribute to force generation, degrading efficiency. Furthermore, a long coil leads to a significant increase in moving mass. Therefore, the short coil/long gap topology has been chosen.

Lorentz force

Voice coil actuators consist of a coil, which is suspended in a radially oriented magnetic field, generated by a field assembly. When current flows through the coil, an axial reaction force between coil and field assembly results due to the Lorentz force. Because the magnetic flux is approximately perpendicular to the current, the Lorentz force law simplifies to:

$$F_{Lz} = BIl \qquad (6.1)$$

where:

B is the magnetic flux density,

I is the current through the conductor, and

l is the length of the conductor.

Alternatively, for a required force with a given actuator:

$$I = \frac{F_{Lz}}{Bl} \qquad (6.2)$$

Hence the force constant of the actuator is:

$$K_F = \frac{F_{Lz}}{I} = B \cdot l \qquad (6.3)$$

Actuator efficiency

The dissipated power in the coil can be expressed as:

$$P_{dis} = I^2 R_{coil} \qquad (6.4)$$

where:

P_{dis} is the power dissipated in the coil, and

R_{coil} is the resistance of the coil.

The resistance of the coil can be written as:

$$R_{coil} = \frac{\rho_w l_w}{A_w} = \frac{\rho_w V_{coil}}{A_w^2} \qquad (6.5)$$

where:

ρ_w is the resistivity of the wire material,

l_w is the length of the wire,

A_w is the cross-sectional area of the wire, and

V_{coil} is the conductor volume of the coil.

The dissipation can be found by substitution of (6.2) and (6.5) in (6.4), yielding:

$$P_{dis} = \left(\frac{F_{Lz}}{B\,l} \right)^2 \frac{\rho_w V_{coil}}{A_w^{\,2}} = F^2 \frac{\rho_w}{B^2 V_{coil}} \tag{6.6}$$

For a given trajectory and guidance stiffness, the dissipation must be as low as possible. The electromechanical efficiency of the actuator can be expressed by the motor constant, K_m, which can be obtained by manipulating (6.6):

$$K_m = \sqrt{\frac{F_{Lz}^{\,2}}{P}} = \sqrt{\frac{B^2 V}{\rho_c}} \tag{6.7}$$

K_m must thus be maximized for optimal performance.

Magnetic circuit

The magnetic flux density that can be obtained in the air gap is dependent on the type of magnetic circuit used. Flux-focus configurations can achieve air gap magnetic flux density up to 1.8 T (Wavre, et al., 1995). Here they cannot be applied, however, due to the 22 mm bore through the center of the actuator and the relatively short available length. Finite element simulations with Ansoft Maxwell® have been performed for various configurations. Within the available design volume, a conventional magnetic circuit as shown in Figure 6.9 proved favorable.
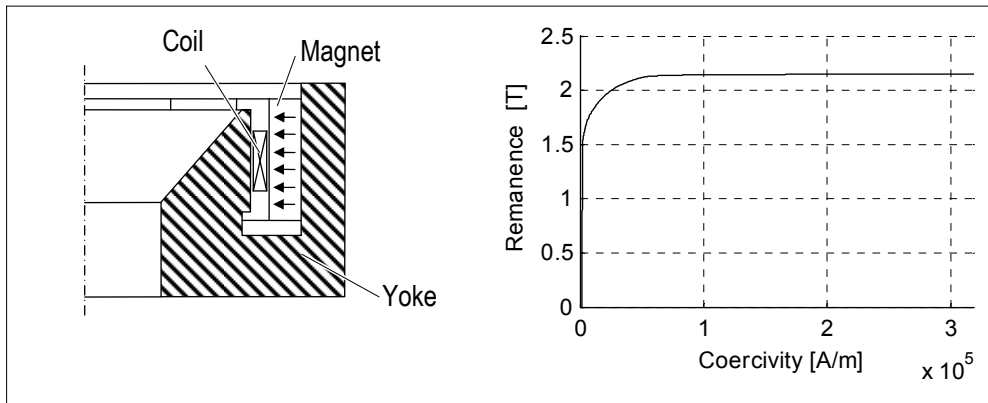
Figure 6.9: Schematic magnetic circuit layout (left) and saturation curve for Armco (right).

To achieve high magnetic flux density with relatively small magnets, Neodymium-Iron-Boron magnets with a maximum energy product of 398 kJ/m^3 are used. The remanence of the magnets is 1.44 T; the normal coercivity is $8.6 \cdot 10^5$ A/m. Because high strength ring-shaped Neodymium magnets with radial magnetization are hard to produce, twelve 30° segments are used. The yoke material is Armco, of which the saturation curve is shown in Figure 6.9. Other materials such as Permendur give slightly better performance, but could not be obtained in small quantities.

Coil volume and magnetic flux density are conflicting; nevertheless, due to non-linear behavior of the yoke material, $B^2 \cdot V_{coil}$ reaches a maximum for a finite air gap width. Simulations with Ansoft Maxwell® show that for limited leak flux, the highest motor constant is found for 2.1 mm air gap width and 44 mm middle diameter. The average magnetic flux density over the coil is about 0.8 T. The flux lines, flux density in the gap, yoke and leak flux are depicted in Figure 6.10.
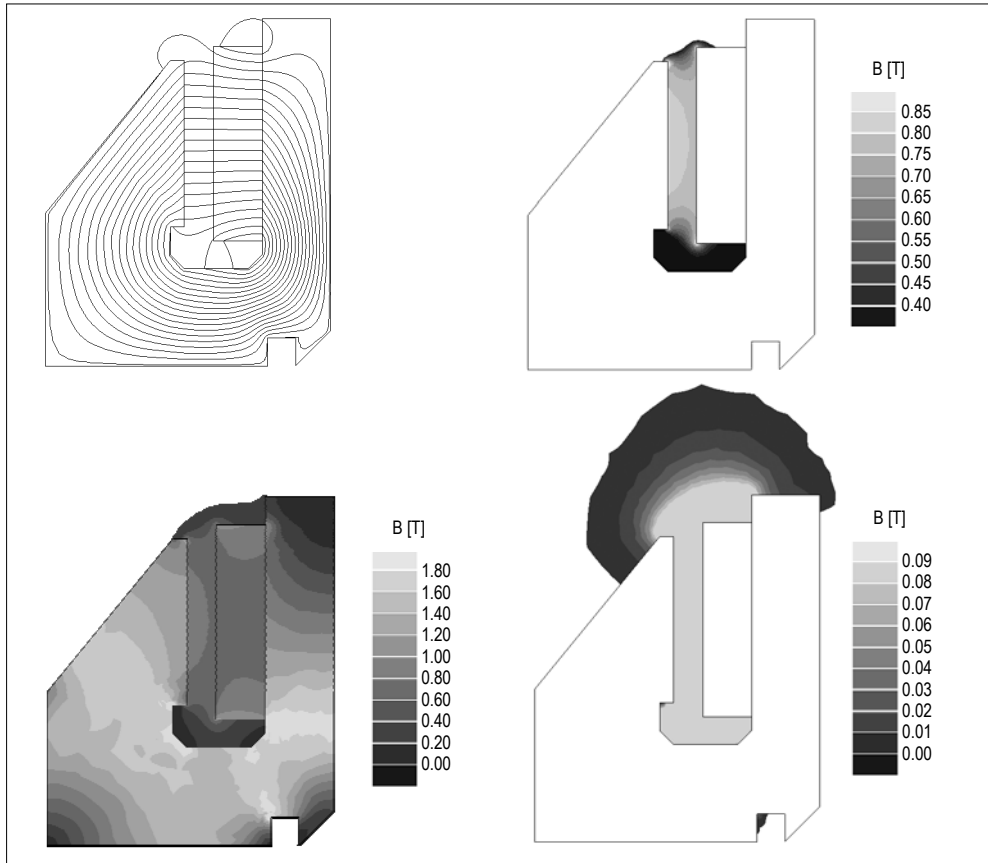
Figure 6.10: Magnetic flux lines in the field assembly (top left), magnetic flux density in the air gap (top right) and in the yoke (bottom left), and leak flux (bottom right).

Because the coil is straight guided and the tolerance train from coil to yoke mainly depends on fits that have been machined on a lathe, a small clearance between coil and yoke can be realized. The components that have large dimensional tolerances are the sintered magnet segments and the coil itself. Therefore, the segments are measured after fabrication, after which the yoke is finished. Since the ring of magnet segments has slightly larger diameter than specified, the air gap is increased to 2.2 mm. Keeping an inner and outer clearance of 0.1 mm and 0.2 mm, respectively, gives a coil cross-section of 1.9 mm radial and 6.3 mm axial with 44 mm middle diameter.

Coil

Both the force constant and motor constant increase for increasing conductor volume in the air gap. To utilize the maximum available air gap volume, a self-supporting coil will be used. A low viscosity epoxy is applied on each winding layer to make the coil self-supporting.

Furthermore, a high fill factor, i.e. the conductor volume divided by the available coil volume, is desirable. Fill factor is determined by winding type, coil geometry and wire and insulation thickness, and is highest for foil and rectangular wire winding. Because insulated foil or rectangular wire of proper dimensions could not be acquired within the available time and budget, standard square lay winding with round wire is applied. A 52 winding, four-layer coil made with 0.475 mm diameter wire closely matches the given coil dimensions. Taking into account the 0.025 mm thickness of the enamel insulation, a fill factor of 0.62 is found.

A winding test revealed that the actual outer diameter was smaller than calculated, except for a small bulge where the wires of one layer cross those of another layer. Therefore, 0.5 mm diameter wire is used instead of the 0.475 mm diameter wire and part of the bulge is removed on a lathe. This locally reduces the wire cross-section, thus increasing the local current density, which is permitted since the generated heat can be conducted to the surrounding material. Two layers of 13 windings and two layers of 12 windings fit in the volume, giving a total number of 50 windings and a fill factor of 0.66.

The electrical resistance of the coil is calculated to be 0.75 Ω. The value of self-inductance is obtained with the empirical Wheeler equation for multi-layer air core coils (Wheeler, 1928), and is 0.183 mH. Accounting for approximate doubling of inductance due to the yoke, the electrical time constant is estimated to be 0.5 ms. Note that a current amplifier with high bandwidth (Subsection 7.4.2) will be used, making this time constant irrelevant for the dynamics. With a continuous maximum permissible current density of 10 A/mm^2, the values found for the theoretical

maximum continuous current, voltage and force are I_{max} = 1.6 A, V_{Imax} = 1.2 V, F_{max} = 8.8 N. The force constant, and motor constant are, K_f = 5.5 N/A, K_m = 6.4 N/√W.

Coil assembly

In Figure 6.11 left, the coil assembly is shown. The coil (1) is bonded to the coil carrier (2) with high strength epoxy. To avoid damping by eddy currents, the coil carrier is made of Macor, a machinable ceramic that is a good electrical insulator. It also has low thermal conductivity, which, unlike in many other applications, is desirable here, because the thermal load of the coil is not critical and it is beneficial to isolate the optics from the heat generated by the coil. A groove (3) is machined into the carrier to accommodate the wires that connect the coil.

To attach the coil assembly to the tube (4) that is fixed to the guidance mechanism, an aluminium cone (5) with lid plate (6) is used. Holes are made in the cone and plate so that during translation, air can escape from one side of the coil assembly to the other. A face (7) on the aluminium cone functions as stroke limiter when it contacts a face (8) on the core of the field assembly. The coil assembly excluding payload, guidance tube and flexure monoliths, has a mass of 24 g, of which 10.5 g is contributed by the coil.
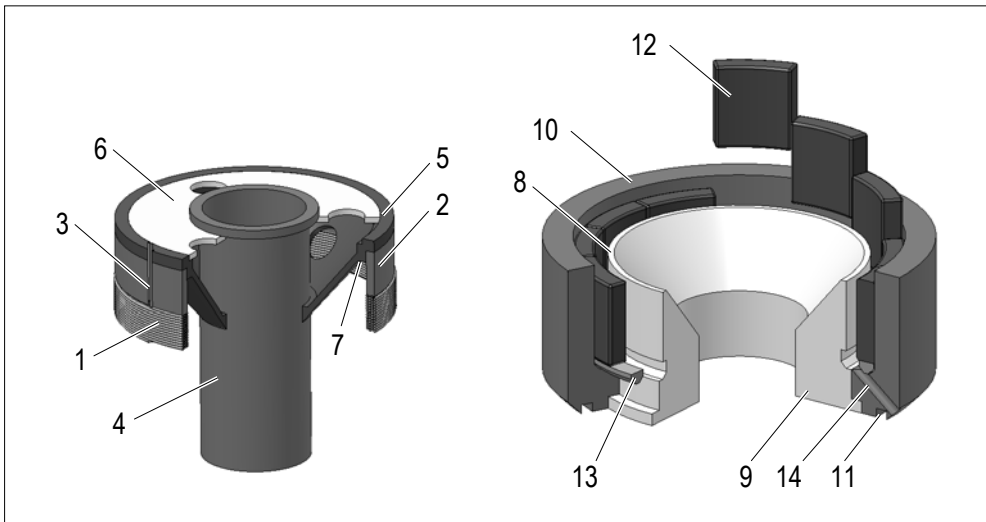
Figure 6.11: Overview of the components of the coil assembly (left) and the field assembly (right).

Field assembly

The Armco yoke consists of a turned core (9) and outer yoke (10) that have a tight fit to limit the barrier it poses for the magnetic field; they are bonded to each other

using adhesive. The inner surface at the top of the outer yoke and the groove (11) at the bottom of the outer yoke have positional fits to the top and bottom flexure monoliths.

The twelve magnet segments (12) are bonded in the outer yoke with epoxy. A nylon spacer (13) on the bottom surface of the outer yoke functions as an axial stop during placement of the magnet segments. In addition to the gap between the coil and the field assembly, six 2 mm diameter bores (14) through the outer yoke allow air to flow through when the coil moves up and down. The yoke as well as the magnets are nickel-plated to prevent corrosion. The field assembly has a mass of 400 g.

6.3.3   FOCUSSING UNIT PERFORMANCE

With a mass of 498 g the focusing unit is the heaviest part of the sensor; the moving mass, M_m, is 46 g. When the sensor is vertically oriented and the actuator is switched off, the nominal deflection due to gravity is 0.37 mm. To limit dissipation, the zero point of the guidance is chosen at this position.

With a force factor of 5.5 N/A and a peak-current of 2 A, at which the amplifier is limited, the theoretical acceleration of the objective lens is 240 m/s^2. The mechanical time constant can be expressed as (Yajima, et al., 2000):

$$\tau_m = \frac{M_m R_{coil}}{K_f^{\;2}} \qquad\qquad (6.8)$$

giving a value of 1.1 ms for the focusing unit including optics.

Eigenmodes

The theoretical first eigenfrequency is 28 Hz with the coil circuit open. Finite element simulations (Henselmans, 2009) show that the first disturbing eigenmode occurs at 1.4 kHz for rotation of the moving body around the optical axis, see Figure 6.12 left. Nominally, this mode does not affect the focus error measurement, aperture correction or interferometer measurement. Nevertheless, it might impose a limitation on the obtainable bandwidth of the control loop, although it will be hardly controllable/observable since it is not exited by the actuator force, nor measurable by the differential confocal system or interferometer.

The second disturbing eigenmode occurs at 1.7 kHz and is swinging of the objective lens assembly in lateral direction, to which both the focus error and interferometer measurements have low sensitivity. The first plate spring resonance occurs at 2 kHz.
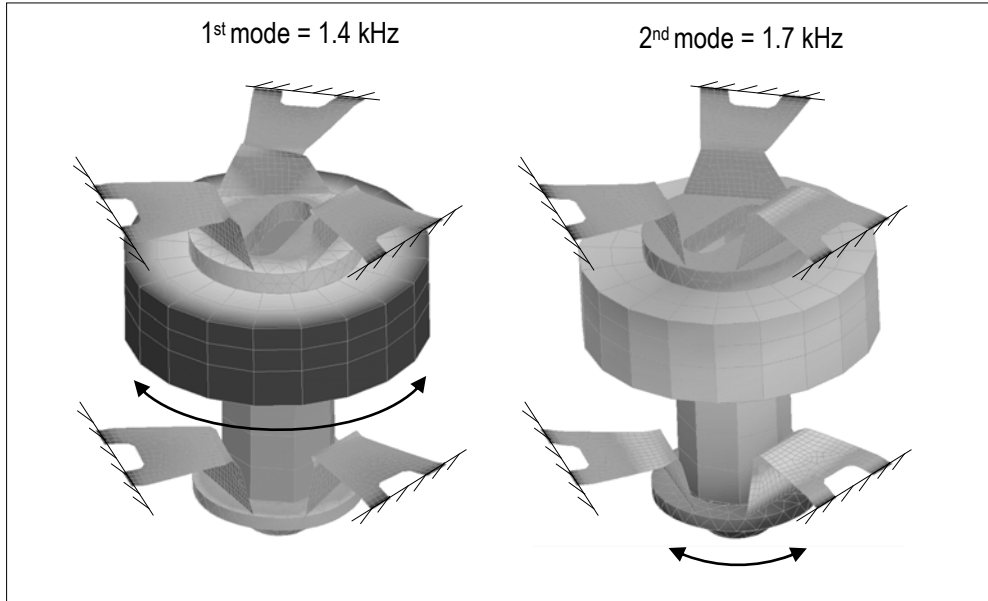
Figure 6.12: The first two parasitic eigenmodes of the system: rotation around the optical axis at 1.4 kHz (left) and swinging of the objective lens at 1.7 kHz (right).

Dissipation

To estimate how much heat is generated in the actuator during measurements, a sinusoidal trajectory with a peak-peak amplitude of A_{pp} and a frequency f is assumed. Because the zero point of the guidance can be chosen at the equilibrium position for each ψ-orientation, no force is required for weight compensation. When damping is not taken into account, the actuator force as a function of time can be expressed as:

$$F_{vc}(t) = M_m a(t) - F_c(t) = M_m \frac{A_{pp}}{2} 4\pi^2 f^2 \sin(2\pi ft) - c_{ax} \frac{A_{pp}}{2} \sin(2\pi ft) = \quad (6.9)$$

$$\left(M_m 4\pi^2 f^2 - c_{ax}\right) \frac{A_{pp}}{2} \sin(2\pi ft)$$

where:

M_m is the moving mass,

F_c is the spring force, and

c_{ax} is the axial stiffness of the guidance.

The root mean square value of the force can be expressed as:

$$F_{rms} = \frac{A_{pp}}{2\sqrt{2}}\left(M_m 4\pi^2 f^2 - c_{ax}\right) \tag{6.10}$$

Using the motor constant, K_m, the average power dissipation can be written as:

$$P_{mean} = \left(\frac{F_{rms}}{K_m}\right)^2 = \left(\frac{A_{pp}}{2\sqrt{2}K_m}\left(M_m 4\pi^2 f^2 - c_{ax}\right)\right)^2 \tag{6.11}$$

Based on (6.11), the dependence of dissipated power on trajectory frequency and stroke is calculated and shown in Figure 6.13.
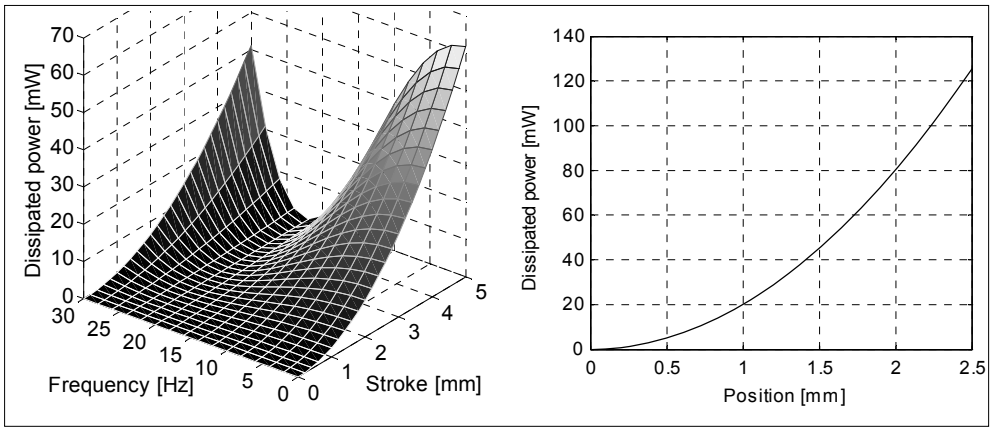
Figure 6.13: Nominal dissipated power for sinusoidal trajectories with various stroke and frequency combinations (left) and power dissipated when a c-position is maintained (right).

For a typical trajectory with four waves of 1 mm peak-peak amplitude over the circumference, power dissipation is just over 2 mW. For a peak-peak amplitude of 2.5 mm, this is about 14 mW and for a heavily freeform with 5 mm peak-peak waves, the dissipation will be just under 62 mW. Since damping has been neglected, the dissipated power is 0 mW at the eigenfrequency. The low power consumption around the eigenfrequency can however not be taken advantage of, because it is higher than most of the frequency content of trajectories for typical freeform measurements.

6.3.4 FOCUSING UNIT ADAPTER

The focusing unit is connected to the interface plate by an aluminium adapter (1), which is bonded to the upper guidance monolith, see Figure 6.14.
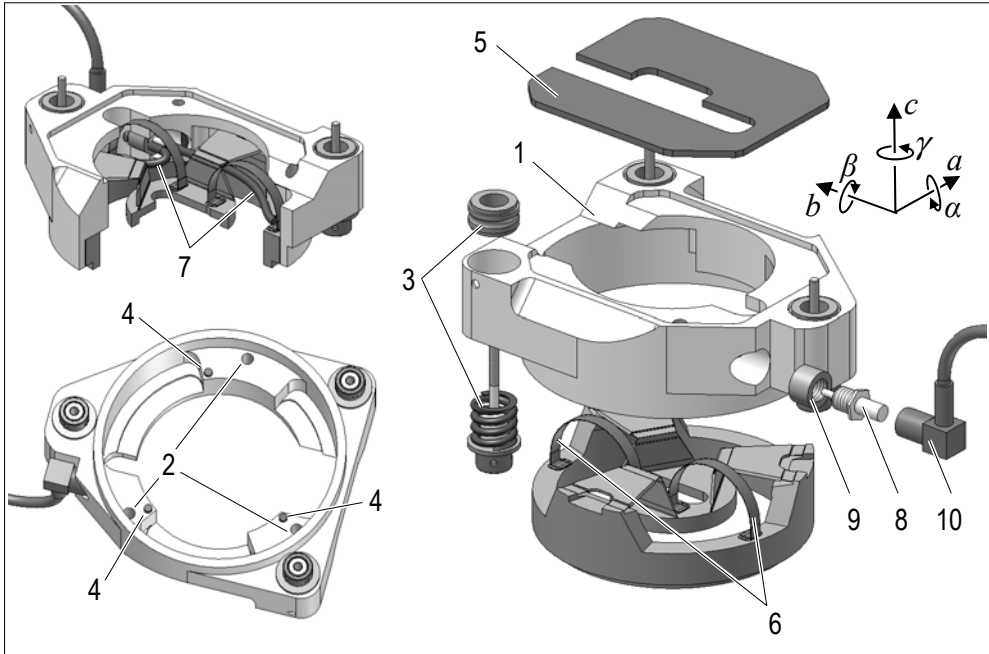
Figure 6.14: The focusing unit adapter providing a semi-kinematic mount, upper stroke limit and electrical connection.

The adapter is connected to the interface plate by a semi-kinematic mount consisting of three sets of bushings and springs (3), similar to the mount between interface plate and ψ-axis rotor. This allows alignment in α- and β-direction of the guidance to the laser beam and reproducible un-mounting and remounting.

Three 8 mm long pieces of 2 mm diameter rubber cable (4) are inserted into holes in the adapter and function as upper stroke limit. This provides a soft stop, needed for high-speed emergency retraction. A cover plate (5) is added to protect the plate springs and reduce the chance that dust enters the focusing unit.

The electrical connection between the fixed world and the moving coil is made via loosely braided flat litz wires without insulation (6). They are formed into semi-circular arches and placed opposite each other to limit disturbances acting on the moving body. The ends of the litz wires are soldered onto Kapton insulated solder pads and bonded to the upper guidance monolith with epoxy. Freestanding solid core wire (7) connects the litz wires to a bulkhead mounted male SMB connector (8). The connector is screwed into a nylon insulator (9) bonded in a bore in the adaptor. A coax cable with a female right angle plug (10) connects to the plug, allowing installation and removal of the focusing unit without threading wires.

6.3.5   OBJECTIVE LENS SUBASSEMBLY

Figure 6.15 shows the objective lens subassembly, which contains the interferometer mirror, objective lens and cover glass.
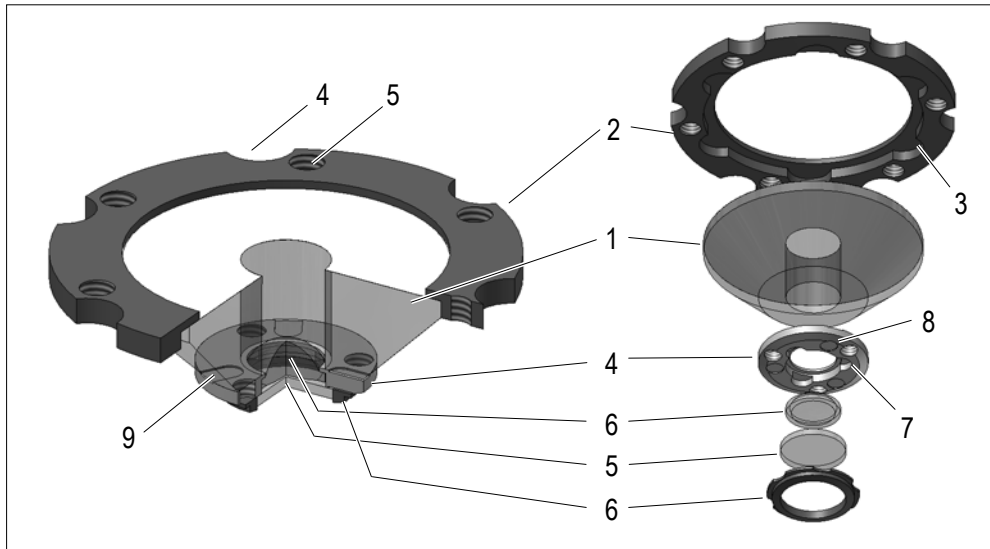
Figure 6.15: The objective lens subassembly.

The interferometer measurement mirror (1) sits on a flange in a TiAlV mounting ring (2) and is bonded to it using adhesive in six cavities (3) around the perimeter of the mirror.

The mounting ring in turn is bonded to the guidance with adhesive applied in six semi-circular cutouts (4). To facilitate breaking the bond if necessary, threaded holes (5) have been machined next to each cutout. The interferometer mirror is aligned in α- and β-direction with socket screws in three of these holes. To do so, the mounting ring is pushed onto the guidance using external springs while the screws are adjusted. Once the mirror is aligned, adhesive is applied in the cutouts and on the screws to lock them.

The aspheric singlet (6) is mounted in a TiAlV ring in much the same way as the mirror: it sits on a flange and is held in place by adhesive applied in three cavities (7) around its perimeter. The cover glass is bonded on top of the lens-mounting ring, with adhesive in three 0.1 mm deep recesses (8). The lens-mounting ring is bonded to the interferometer mirror with adhesive applied in three recesses (9) at its back. Manipulation in α- and β-direction is not necessary since the interferometer mirror is aligned in α- and β-direction and parallelism tolerances of the mounting surfaces are well within the angular alignment requirements of the

singlet. Lateral alignment below 0.01 mm can be obtained manually; the uncured adhesive is beneficial during this process due to its viscosity.

To protect the lens and guidance in case of a collision with a freeform, a Teflon glider (11) is mounted around the cover glass. The low friction coefficient of Teflon limits the lateral forces exerted on the guidance when being pushed back by the rotating freeform. This reduces the risk of damaging the optics and the guidance mechanism. Teflon is difficult to bond using adhesives and limited space is available to fix the glider mechanically. Therefore, the glider has a flange on the outside, over which a high strength epoxy is applied in the form of a closed ring. At six cutouts in the flange, the adhesive bonds to the lens-mounting ring, thus retaining the glider mechanically. With the placement of the glider, the free working distance becomes 1.45 mm.

6.3.6 PROTOTYPE OVERVIEW

The complete design is shown in Figure 7.8. The largest outer dimensions are 63 mm x 89 mm x 126 mm; the optics housing measures 39 mm x 44 mm x 81 mm. The mass of the assembly is just less than 1 kg, of which 500 g is contributed by the focusing unit, 270 g by the adapter and interface plate, and 200 g by the central optics holder and components.
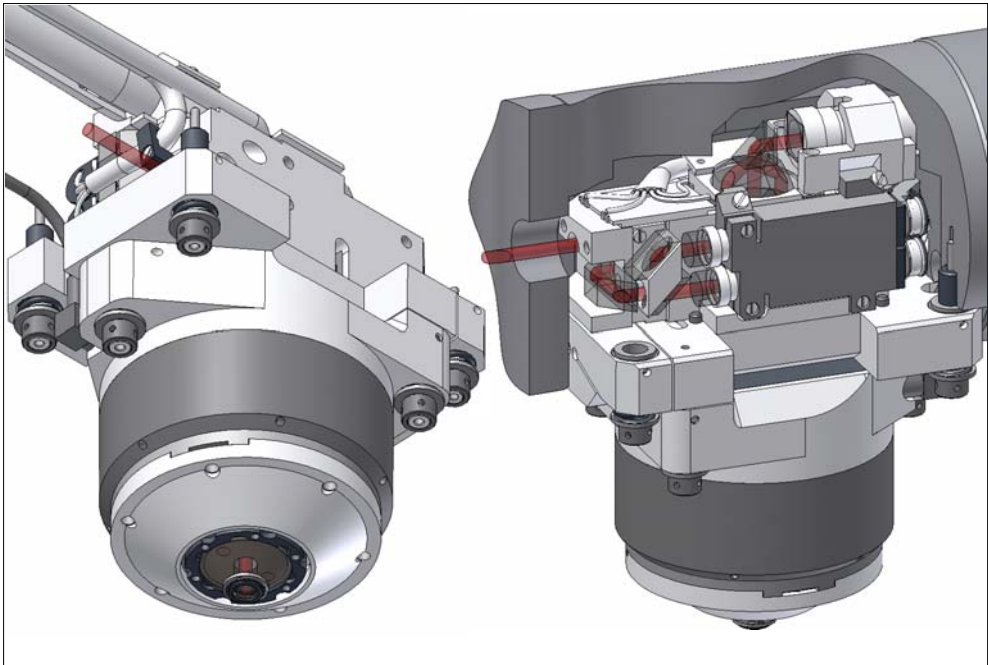
Figure 6.16: Overview of the complete prototype design.

With the optics and the mechanics designed, the components can be fabricated and the system assembled and tested. This phase, as well as the sensor electronics and control, are covered in the next chapter.

7 REALIZATION, FOCUSING UNIT PERFORMANCE AND CONTROL

Fabrication and assembly of the prototype is treated with an emphasis on experience gained during realization, and aspects that deviate from conventional fabrication methods. Next, the results of various performance tests of the focusing unit will be presented. At the end of the chapter, signal processing and control is discussed.

7.1 PROTOTYPE FABRICATION AND ASSEMBLY[5]

Realization of the prototype will be discussed and is illustrated with photographs taken during the various stages of fabrication and assembly. Optics, optomechanics and the focusing unit are treated in separate subsections and followed by photographs of the fully assembled sensor.

7.1.1 OPTICS

Figure 7.1 shows the optical components before bonding. Unfortunately, there was insufficient budget to test whether they meet specification through measurement of the optical surface shapes and the reflectivities of the coatings.

---

[5] *Modification of standard optical components and the fabrication of specials has largely been carried out by the optical workshop of TNO Science & Industry in Delft. Fabrication of mechanical components and assembly of the focusing unit has been realized by the GTD workshop of the Technische Universiteit Eindhoven, while assembly and alignment has been mainly carried out by the author with support and facilities of the GTD.*
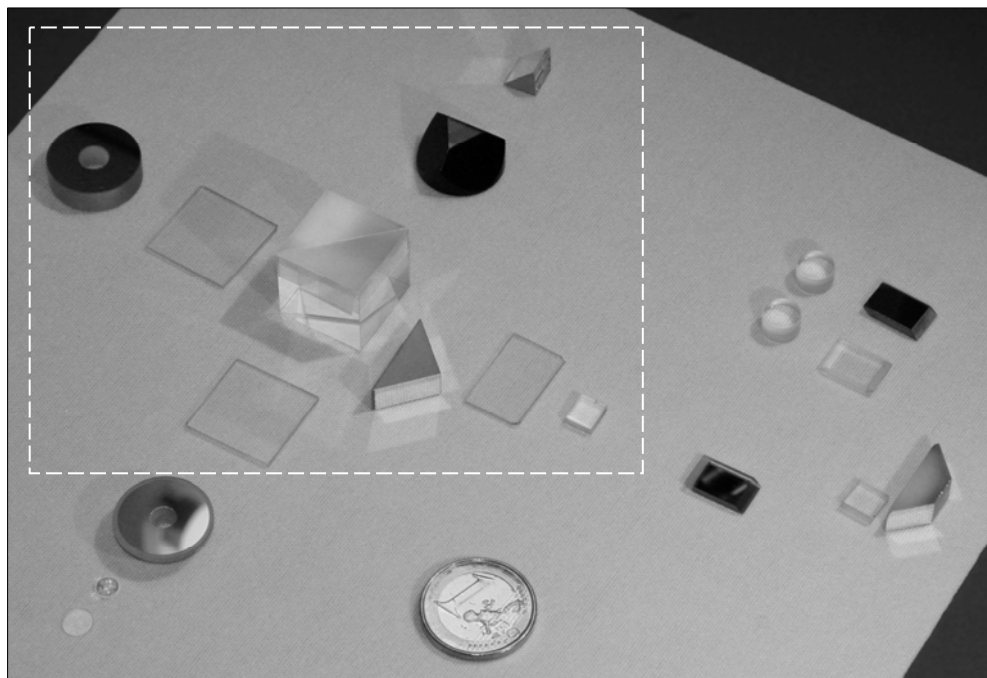
Figure 7.1: The optical components before assembly; the components of the optical monolith are shown in a dashed square with a euro coin as a reference of scale.

Figure 7.2 depicts a plastic jig used to position the components of the optical monolith during bonding; compared to metal jigs this reduces machining time and the likelihood of damaging the optics.
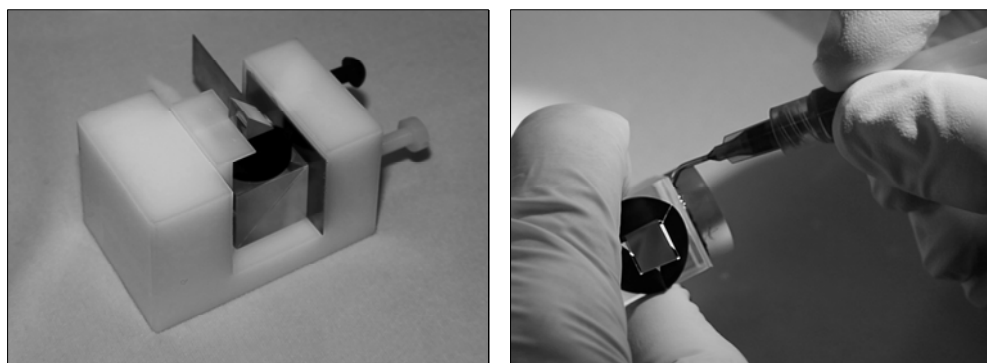
Figure 7.2: Positioning of the entrance prism on the transmissive facet of the topped-off cube corner by means of a jig (left) and sealing the clearance between interferometer reference mirror and quarter wave plate using adhesive (right).

The few degrees of freedom that need to be controlled more accurate than attainable by these jigs are inspected using a traveling microscope and aligned with shims and foil. Once alignment is achieved, the adhesive is cured with UV light.

Initially the interferometer reference mirror was meant to be bonded directly onto a Quarter Wave Plate (QWP) of the optical monolith. Therefore, unlike the two other QWPs that are coated on one side, this QWP must remain uncoated. By accident, however, all three QWPs were coated on one side.

Because the coating on the QWP is designed for a glass-air transition, it will not have low reflectance when bonded with index matched adhesive. To avoid the risk that bonding the mirror and the QWP might lead to an unforeseen effect, it was decided to leave an air gap of about 0.1 mm in between. Hence, adhesive is applied in two strips, left and right of the optical area. Two 110 µm diameter glass fibers are placed in the adhesive to align the mirror and achieve a stable air gap.

Since in this way the optical surfaces are not accessible for cleaning, the clearance around the edge is sealed with adhesive to prevent pollution from entering the air gap (Figure 7.2, right). Glass particles of about 2 µm diameter are mixed with the adhesive to prevent it from creeping into the air gap by capillary action.
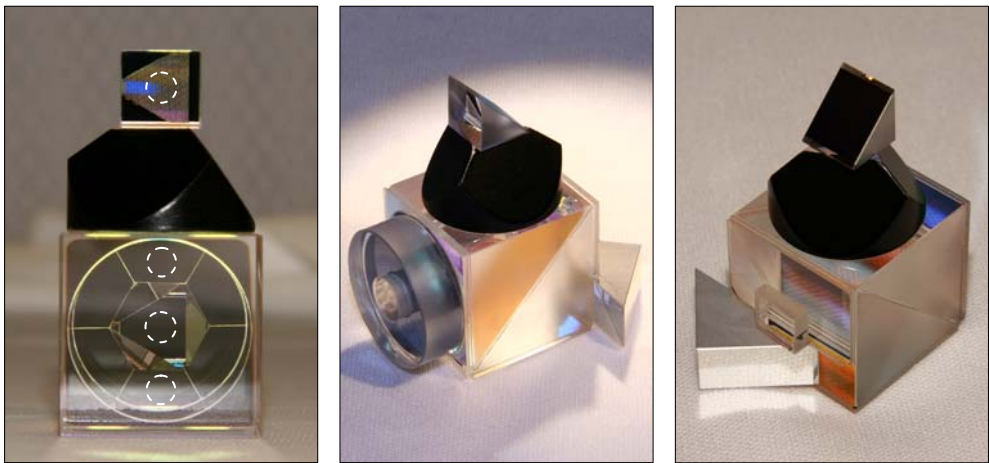
Figure 7.3: View of the partly assembled optical monolith from the side of the interferometer reference mirror (left) and the finished optical monolith (middle and right).

Figure 7.3, left, shows a view into the partly assembled optical monolith from the side of the interferometer reference mirror, with the beam splitter, corner cube and entrance prism in place. The edges of the topped-off corner cube and their reflections can clearly be distinguished. The dotted circles approximately indicate

the areas that are occupied by the differential confocal and interferometer laser beams. The finished optical monolith can be seen in Figure 7.3, middle and right.

7.1.2   OPTOMECHANICS

Figure 7.4, left, shows the aluminium optics holders that are contained in the ψ-axis pocket, before they have been anodized. Some screws (indicated by arrows) in the central optics holder that have no function in the finished sensor can be seen as well. They are used during assembly and testing to mount temporary alignment tools and sensors, and to connect the central optics holder with a stand.
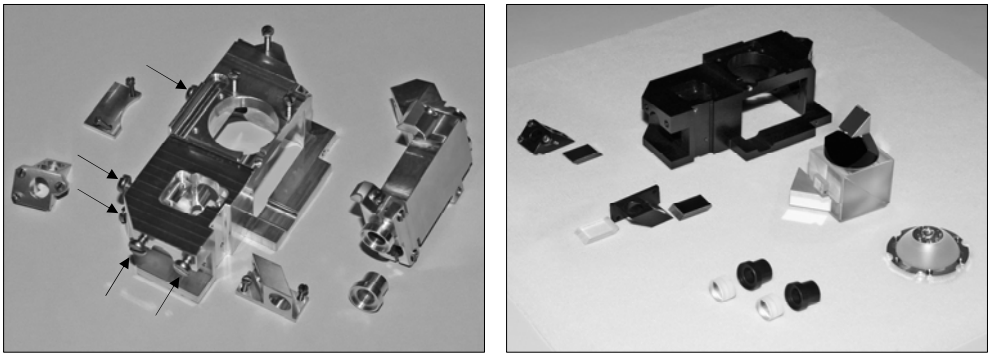
Figure 7.4: The central optics holder, differential confocal holder and some of the smaller optics holders before anodization (left), and the optics and matching holders before bonding (right).

In Figure 7.4, right, the optics paired with the anodized holders to which they will be bonded with epoxy. Because there are small threaded holes in the components, dimension-stable anodization is applied.

The optical monolith and the image lenses must be aligned to their holders, which is achieved through measurements of their edges with a traveling microscope. For the other components, alignment is either non-critical or obtained through manipulation of the mounts during final assembly.

Figure 7.5, left, shows the mounted optics. The dove prism is not attached here: it is one of the last components to be bonded since its positioning is used to laterally align the beam at the imaging lenses.
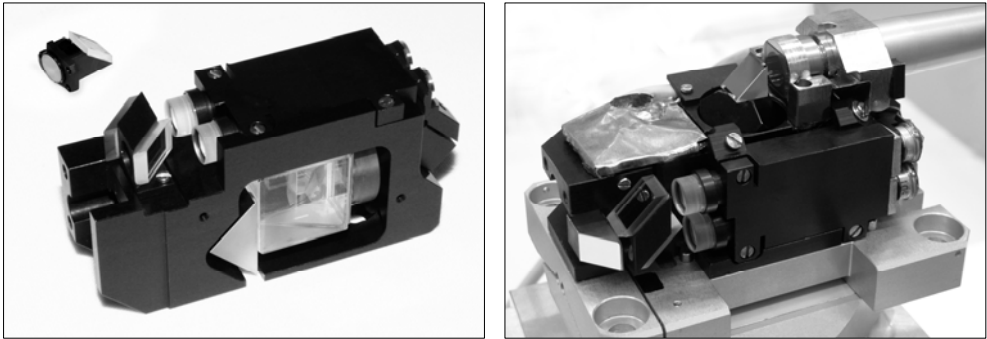
Figure 7.5: Central optics holder with mounted optics (left) and the central optics holder in the completed sensor (right).

In Figure 7.5, right, the central optics body is shown when fully assembled and bonded to the interface plate. After the system was aligned, separately tested and installed in the machine for further testing, the adhesive bond between central optics holder and interface plate failed. In some places, the adhesive had detached from the central optics holder while in other places from the interface plate, indicating bad adhesion to both surfaces. Because the recommended procedures for surface preparation had been followed, it was unclear what the cause could be. It was suspected, however, that the abrasion step might not have been thorough enough, since anodized surfaces are known to have high hardness.

The two parts have been re-bonded after being abraded for a second time, starting with a courser grid. In addition, extra epoxy with higher bond strength has been applied along the edges. Later it was found in (Anodizing - Wikipedia, 2009) that the type of passivation process applied after anodization has a large effect on bond strength. Passivation is achieved either by hydrating the oxide layer or by impregnation with a sealant. UK defense standards (DEF STAN 03-24/5, 2008) and (DEF STAN 03-25/5, 2008) require that anodized parts intended for adhesive bonding are not sealed. Communication with surface treatment companies revealed that even some non-sealed surfaces display low bond strengths because of additives that might be used to improve process control during hydration. In that case an adhesive primer offers a solution.

Anodizing of the prototype components was outsourced to two different companies, one of which has gone bankrupt since then. Upon inquiry with the company that is still in business, it was discovered that they indeed seal anodized layers with nickel fluoride, which is known to produce weak bonds with epoxy adhesives. In this case bond strength can be improved by using an etch primer.

The passivation process also influences dye-retention. This might be of importance in optical systems since black anodized parts are often used, while cleaning with solvents is common. Because of the interdependence of alloy composition, anodization process, passivation processes and dye-retention it seems best for a next generation sensor or for changes to the current sensor to specify anodization requirements in cooperation with the coating supplier.

7.1.3 FOCUSING UNIT

Figure 7.6, top left, shows the fabrication of the upper guidance monolith by wire EDM. The stress state of bulk material is usually non-uniform over its diameter due to forming operations and different cooling rates of the outside and inside of the material. To obtain identical plate springs, a round bar of TiAlV is used as starting material, because this has an approximately rotationally symmetric stress state.

At the stage of fabrication in the photograph, the plate spring at the bottom has been finished and the one at the top of the picture is only pre-machined by turning and milling. The outer contour of the plate spring on the left has just been machined, and work on the inner contour is about to start. It can be seen that a fair amount of material is left in the pre-machined state so that the material of the plate spring itself does not pick up stress during milling and turning. After six runs on the EDM machine, one for the inner and one for the outer contour of each folded plate spring, the component is finished. To achieve uniform plate spring thickness, the inner and outer contour of each plate spring is machined in consecutive cuts. The hexagonal shape of the holder that can be seen in the top left corner of the picture, is employed for indexing the rotation between the plate springs.

The cut of the inner contour of each plate spring is made from the moving part towards the stationary part, so that variations in plate spring thickness due to deformation caused by internal stress in the material is minimized.

Figure 7.6, bottom left, shows the coil after winding; the mandrel is greased so that epoxy does not bond to it. Once the epoxy has cured, the mandrel is taken apart in three sections: the part held by the chuck, a central disc around which the coil is wound, and an end disc. A round bar over which a bus can slide connects to the central disc to push the coil off the central disc. The same tool functions as a slide fit to the coil carrier to position the carrier concentric to the coil during bonding.

In Figure 7.6, right, the components of the guidance and actuator are displayed before assembly, along the centerline from top to bottom: upper flexure monolith, connection tube, cone cover plate, cone, coil carrier, coil, magnet spacer, outer yoke, core and lower guidance monolith. The magnet segments are located on the left.
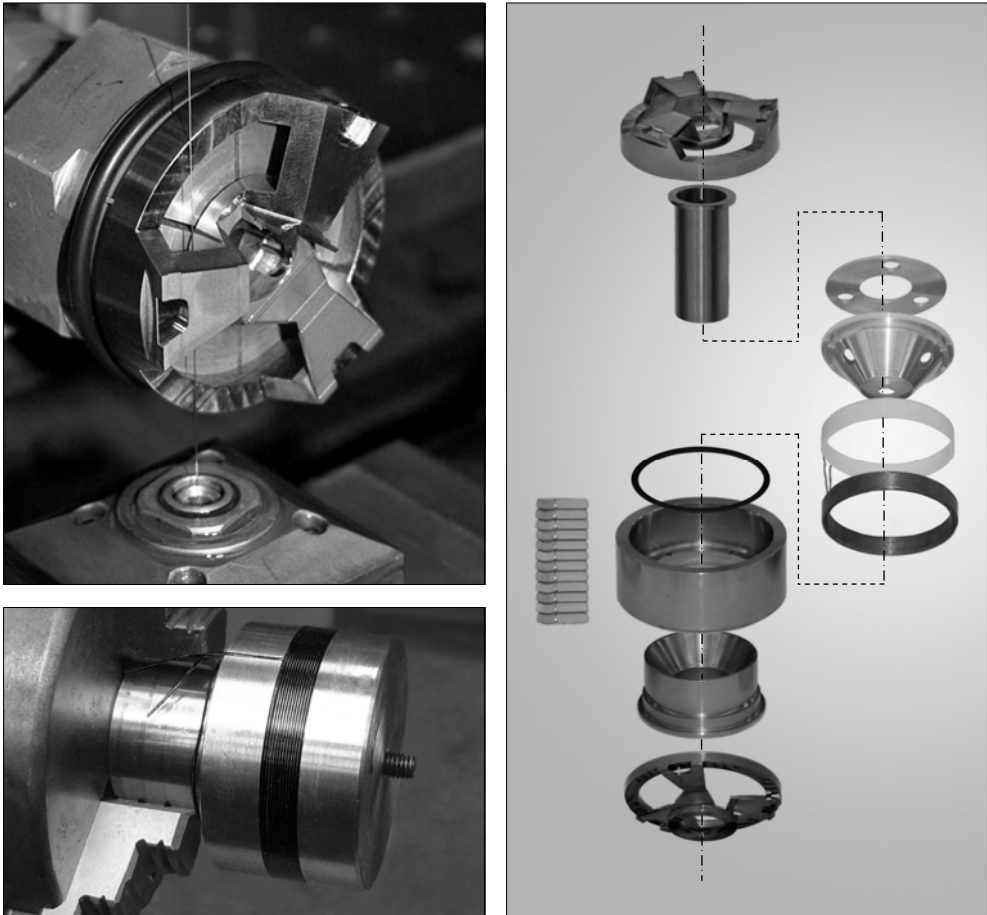
Figure 7.6: Fabrication of the upper plate spring monolith by wire EDM (top left), coil and coil winding mandrel in winding machine (bottom left) and photo montage of the guidance and actuator components before assembly (right).

The coil assembly is bonded with epoxy while the rest of the guidance and actuator has been assembled without adhesive. Only after verifying that it runs clear, it is reassembled while bonding the components. In Figure 7.7, left, the topside of the focusing actuator is shown with the folded plate springs and litz wire arches visible. The cover plate and the bushings of the kinematic mount have not been installed yet. In Figure 7.7, right, the focusing unit bottom can be seen with the optics in place. The coaxial SMB male connector in the adapter is just visible and the female connector of the connecting cable is shown.

Figure 7.7: Topside of the focusing unit with the litz wires and plate springs visible (left) and the bottom of the focusing unit with the optics already in place (right), the coaxial connector can also be seen.

When the sensor was fully assembled and aligned, the adhesive bond between the lower guidance monolith and the objective lens assembly failed, in spite of low loading. Communication with 3M, the manufacturer of the epoxy (Scotch-Weld DP100), did not clarify how the bond strength could be so low. Stronger bonds to titanium alloys can be obtained with DP470; use of DP100, however, should result in strength that is acceptable for this application.

After EDM of the monoliths, a thin layer could be observed in the form of colored gradients on the TiAlV. It is suspected that during EDM a thin anodization layer has formed that is too smooth for the adhesive to bond to. Therefore, before reattaching the components with DP470 the layer was removed.

7.1.4 PROTOTYPE OVERVIEW

The assembled prototype can be seen in Figure 7.8 with a 2 euro coin as a reference of scale. In Figure 7.9, the sensor is shown installed in the machine and locked onto a surface of a toric mirror.
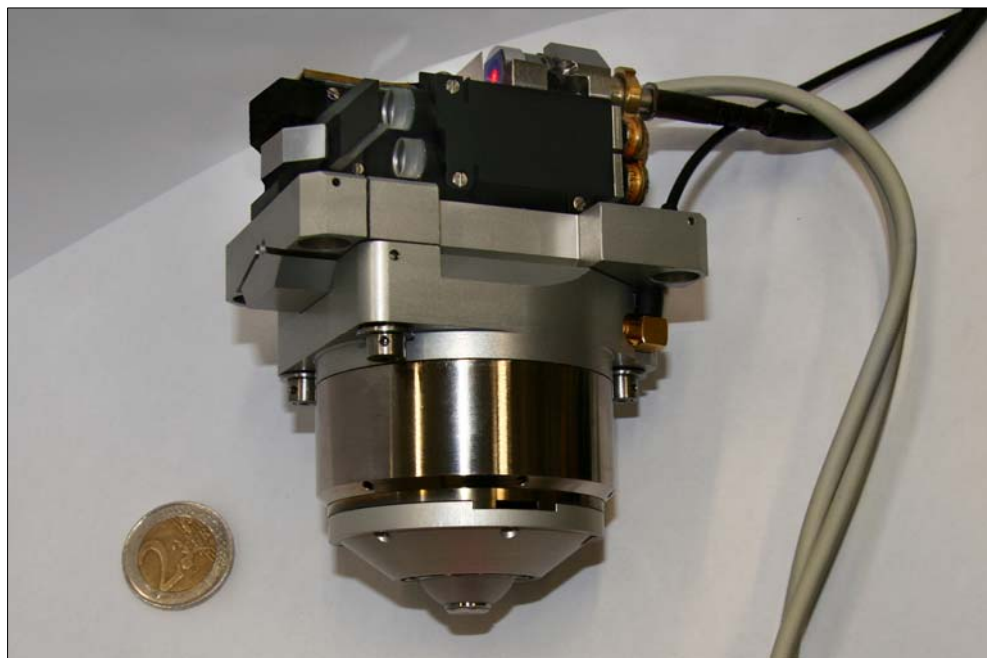
Figure 7.8: The completed prototype sensor before being installed in the machine.

Figure 7.9: The sensor installed in the machine.

7.2 FOCUSSING UNIT PERFORMANCE

Before, during and after assembly of the sensor, experiments have been carried out to investigate the performance of components and subsystems. This allows malfunctions to be detected and repaired in an early stage, it enables estimation of the properties of the complete system and can result in experience which might be useful for future designs. The results of the tests of the focusing unit are presented in this chapter as well as the response of the differential confocal system. A more specific investigation into the differential confocal system performance, such as: measurement noise, uncertainty and tilt dependency is given in Chapter 8.

7.2.1 VOICE COIL ACTUATOR

The electrical properties of the coil are determined to test whether it contains a short-circuit, since this would lead to reduced bandwidth and increased dissipation. The actuator's force factor and motor constant are of influence on the measurement uncertainty too, and therefore determined as well.

Coil

The electrical resistance of the coil including litz wires and connection wires has been measured to be 0.8 ± 0.1 Ω, which is close to the calculated value of 0.75 Ω. Hence, if there is a short-circuit in the coil, it consists only of few windings with a resistance smaller than the measurement uncertainty. To enable detection of a small number of short-circuited windings, the electrical resistance at 1000 Hz has been measured. A short-circuit in the coil creates a transformer in which one or several windings form a short-circuited secondary coil, while the remaining windings form the primary coil. If this is the case, the current in the short-circuited secondary coil will be amplified, leading to much higher resistance at high frequencies. The electrical resistance at 1000 Hz was measured to be 1.65 Ω, which seems about right; hence, the coil contains no short-circuited windings.

The induction of the coil was measured to be 164 µH, about 10% lower than the 183 µH that was estimated during coil design. When inserted in the air gap the coils inductance increased to 285 µH, giving an electrical time constant of 0.36 ms, 33% better than estimated.

Actuator characteristics

The force factor has been determined for various positions over the stroke since variations have influence on the stability of the control loop. This has been done by hanging weights from the assembled focusing unit and subsequently varying the current until the guidance returns to the position observed when unloaded. The values found at -2 mm, -1 mm, 0 mm, 1 mm and 2 mm are 5.6 N/A, 5.6 N/A 5 N/A

5.95 N/A and 5.85 N/A, respectively. The shape of the profile is unexpected and it does not seem likely that it is caused by measurement errors. Decreasing magnetic flux density at the outer parts of the air gap can give rise to force factor variations over the stroke. Furthermore, for high currents, the coil itself generates a magnetic field in the yoke, bringing about a variation of force factor as a function of current direction and magnitude. A combination of these effects, nevertheless, would not lead to a minimum of the force factor at the equilibrium position. What mechanism causes the variation has not been further studied. The force factor variation is within acceptable limits for controller design.

Using equations (6.3), (6.4) and (6.7) for force factor, dissipation and motor constant and substituting the measured resistance and the force constant, the motor constant can be calculated to vary between 5.6 N/√W and 6.6 N/√W over the stroke.

Temperature rise during measurement

As discussed in Subsection 5.1.5, a rise in air temperature of 0.1 K in the interferometer measurement path gives a measurement error of 6.9 nm. The nominal RMS current needed for a demanding sinusoidal trajectory is 0.3 A. Therefore, a DC current of 0.35 A was sent through the coil while measuring the air temperature in the field assembly with a thermocouple inserted through one of the air vents in the outer yoke. Ten minutes into the experiment, the temperature had stabilized at 0.1 K above the ambient temperature; after half an hour, the experiment was terminated.

During typical measurements, the temperature at the optical axis is expected to rise less because of mixing with ambient air, the long path between air gap and optical axis and the vicinity of heatsinks such as the yoke and the focus unit adapter. Furthermore, the rise of temperature is slow, so that it can partly be reconstructed from the drift.

7.2.2 GUIDANCE MECHANISM

The guidance straightness and the relation between force and axial displacement have been measured. From the force-displacement diagram the stiffness can be calculated, which is the dominant parameter in determining dissipation and of importance when designing a controller. The straightness of the guidance is of interest since it determines the lateral displacement of the objective lens relative to the laser beam and hence the lateral measurement spot displacement.

Because hysteresis in lateral direction contributes to uncertainty of measurement spot location, and hysteresis in driving direction influences the tracking error, these are analyzed as well.

Guidance straightness

To measure the angular error motion of the guidance, a mirror has been attached where the objective lens assembly will be placed, and an autocollimator is used to measure the tilt over the stroke. The graph in Figure 7.10, left, shows the results of this measurement. At 200 µrad tilt in α-direction and 70 µrad in β-direction, the angular error motion is much larger than expected. Given the stroke of 5 mm and a guidance basis of 44 mm, this translates to a misalignment of the guidance trajectories of the two monoliths of 3.5 mrad, which might be caused by contamination of the fits, fabrication errors or internal stress in the material. The repeatability is in the range of what would be expected: about 5 µrad.

Nominally, the distance measurement has no first order sensitivity to the angle of the focus lens assembly. In the realized sensor, however, misalignment of the interferometer beams to the differential confocal beam might lead to a significant Abbe error. This can be compensated for by interferometer calibration due to the high reproducibility.
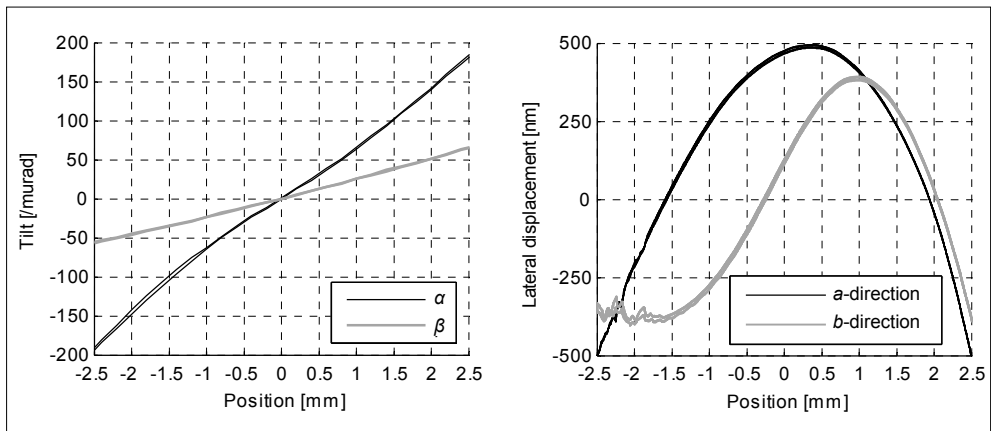
Figure 7.10: The angular error motion of the focusing unit guidance mechanism (left) and the corresponding lateral error motion at the measurement spot location (right) plotted against position.

The lateral error motion of the guidance over the stroke has been measured as well. To this end, a gauge block is mounted at the position of the focus spot and its lateral translation is measured with a capacitive probe. Since the guidance will be aligned to the differential confocal laser beam during assembly, only the non-linear component of the lateral displacement is of interest. The results in both *a-* and *b-*direction are plotted against position and shown in Figure 7.10, right.

Between -1.6 mm and -2.5 mm, some erratic displacements and a decrease of reproducibility can be seen, leading to the suspicion that in this region, the coil was scraping against the field assembly. Upon inspection with a stethoscope, a scratching sound was indeed heard. What caused the coil to run off-center is unclear. The problem was solved by slightly deforming the field assembly; however, all the straightness measurements are taken with the scraping coil.

The measured deviation from a straight line is about 1 μm in *a*-direction and 800 nm in *b*-direction, and as such, larger than expected. Since no reversal has been performed, the measurement contains both the error motion and the surface figure of the gauge block, averaged over the capacitive sensor's measurement area. Uncertainty in measurement spot position after calibration, however, is determined by the reproducibility, which can be measured accurately without a reversal. The reproducibility is mainly determined by the hysteresis, and is obtained by fitting a polynomial through the data and subtracting it from the measured data. The resulting graphs in *a*- and *b*-direction are shown in Figure 7.11.
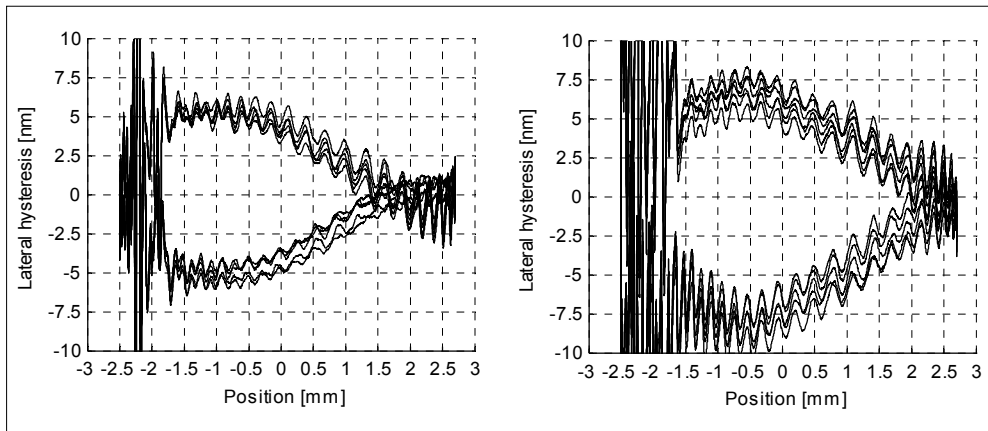
Figure 7.11: Hysteresis of the focusing unit guidance mechanism in a-direction (left) and b-direction (right) at the measurement spot location, plotted against actuation over the stroke.

The lateral accuracy of the guidance must be better than 50 nm after calibration. Since the maximum lateral hysteresis stays within 10 nm over the good part of the stroke, it is expected that this requirement can be met.

Force-displacement diagram

Figure 7.12, left, depicts the force-displacement diagram of the focusing unit. It is acquired by feeding a 0.1 Hz sinusoidal current to the actuator while measuring its axial position with the integrated interferometer.
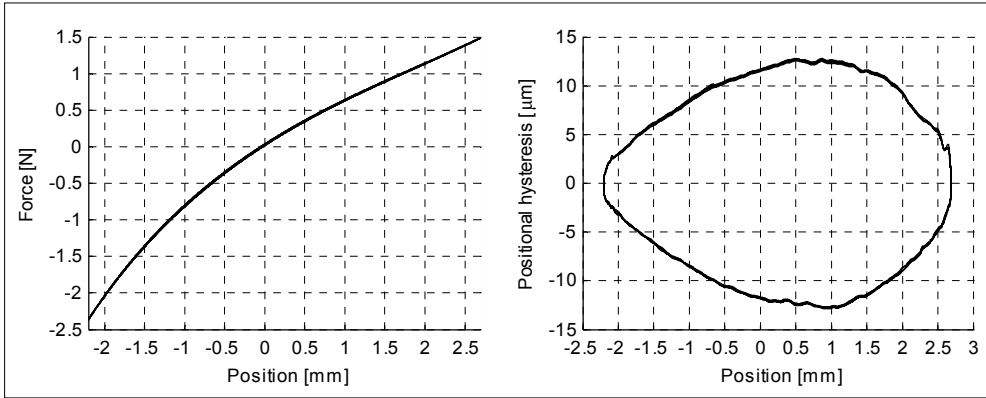
Figure 7.12: Force-displacement diagram of the focusing unit guidance mechanism (left) and the corresponding positional hysteresis in driving direction plotted against position (right) of the guidance mechanism.

The force is obtained by multiplying the current by the nominal force factor. Since the force factor varies slightly over the stroke, this gives rise to a maximum local force error of about 10%. Correcting for this error is not possible because the force factor has not been measured at enough positions to achieve an accurate fit. It can be seen that the local stiffness varies over the stroke: from 1700 N/m to 500 N/m; it is about 700 N/m at the equilibrium position, roughly a factor of two lower than calculated. The overall stiffness varies from 550 N/m for full retraction to 1100 N/m for full extension. A possible explanation is that internal stresses in the material of the guidance monoliths cause it to have negative stiffness. This, however, has not been investigated.

Hysteresis in driving direction

The hysteresis in driving direction is obtained by fitting a polynomial through the force-displacement data and subtracting it from the displacement signal taken over multiple cycles over the stroke. The resulting signals are then filtered to remove measurement noise and oscillations at the eigenfrequency in driving direction. Figure 7.12, right, shows the positional hysteresis plotted against position. The asymmetry is probably caused by the non-linear stiffness of the guidance. This is supported by the fact that the hysteresis in driving force, when plotted against the instantaneous force, yields a more symmetric graph, as can be seen in Figure 7.13, left.

It is suspected that the hysteresis is caused by friction in the litz wires or pumping of air. When the positional hysteresis is plotted against the velocity, a nearly straight relation is observed, with some virtual play around zero velocity, as shown in Figure 7.13, right. This indicates viscous damping and some static friction. Since

176

the hysteresis curve is quite smooth, it is expected that the control loop can sufficiently suppress hysteresis, even though it is multiple times larger than the measurement range of the differential confocal system.
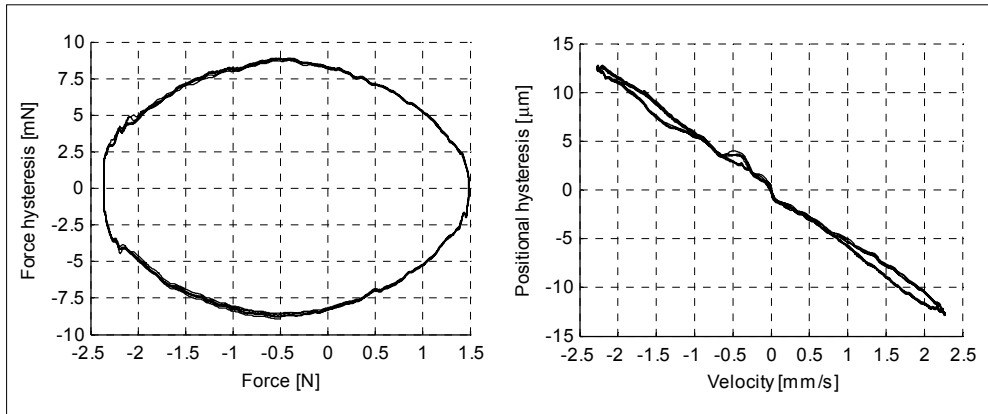
Figure 7.13: Hysteresis in driving force plotted against nominal driving force (left) and the positional hysteresis plotted against the velocity of the guidance.

The scraping of the coil against the field assembly was resolved and the focusing unit now meets the requirements. Nevertheless, some things remain unexplained, such as the cause of the large error motion and stiffness variation over the stroke, and how the coil could initially slide against the field assembly. Considering future guidance mechanism and actuator design, it would be interesting to investigate these things thoroughly; upon considering the project as a whole, however, it was decided to invest efforts in other areas.

7.3 DIFFERENTIAL CONFOCAL SIGNALS

There are two ways to measure the Focus Error Signal (FES) of the sensor: the focusing unit can be locked and the Surface Under Test (SUT) is scanned axially through focus, or a stationary SUT can be placed in front of the sensor and the focusing unit is actuated to scan through focus. Both methods are used, since they each have specific advantages, depending on which test is to be performed.

The results presented here are obtained while the sensor is mounted in the machine, with a stationary SUT placed on the machine's product table. The displacement of the focusing unit is measured with the integrated interferometer. Due to vibrations, this method is noisier than scanning the SUT, but it allows a much larger range to be covered.

Long range signals

Figure 7.14, left, shows a long-range plot of the FES for measurement of a bare Zerodur surface with a reflectivity of about 4.5 %, and a mirror with a protected aluminium coating, with a reflectivity of about 86 %.
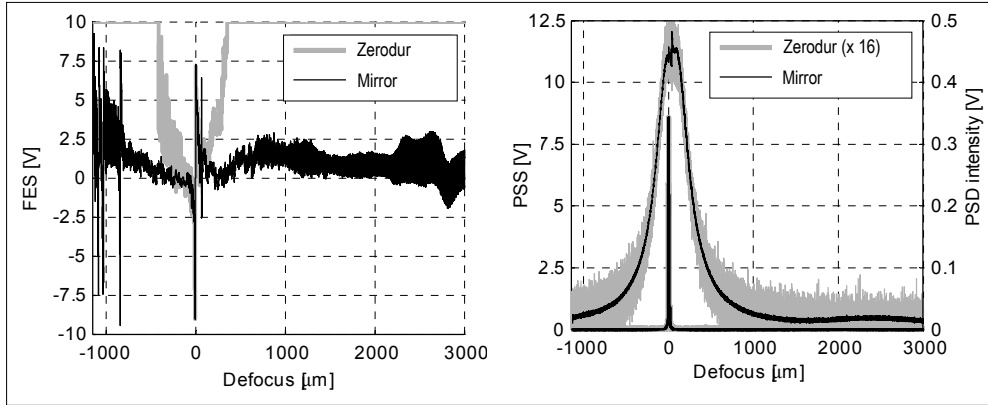
Figure 7.14: Long-range plot of the FES (left) and of the intensity signal of the PSD (1) and the photodiode sum signal (2) (right) for measurement on a Zerodur flat and a mirror.

The spike at 0 µm defocus is the FES around best-focus. At the maximal negative defocus depicted, the distance between the Teflon glider and the SUT is a few tenths of a millimeter. It can be seen that the signal obtained with the mirror stays low for large values of defocus, while for the Zerodur surface it increases to 10 V for large values of defocus. The reason is that for low photodiode currents, the normalization causes amplification of small unbalances in the photodiode signals.

For the aluminium-coated surface, on the far left of the FES plot, three spikes are visible. These are the FES curves of the front- and back-surface of the objective lens window and the front surface of the objective lens itself, reflected by the SUT. They can be distinguished from the real FES by their low intensity.

Figure 7.14, right, depicts the intensity signal of the Position Sensitive Detector (PSD) and the Photodiode Sum Signal (PSS), obtained with a Zerodur and an aluminium-coated SUT. For the uncoated SUT the signals are amplified by a factor of 16 to allow comparison of their shape with that of the mirror's signals.

With the PSD intensity signal the approach to focus can be detected; however, the top of the curve is not narrow enough to do so when close to best-focus. If the top of the curve is less wide so that its usable range overlaps with the usable range of the photodiode signals, this could be utilized as an input of a control algorithm for

focus finding, which will be discussed in Subsection 7.5.3. Thereto, a lens and a large pinhole can be placed in front of the PSD in a future design, so that the intensity curve is less wide but aperture correction remains unaffected.

Medium range signals

In Figure 7.15, left and right, the FES and PSS are shown from -60 µm to 60 µm, again for a Zerodur surface and a mirror. Between about -30 µm and 30 µm, the sign of the FES is used to detect whether the objective lens is too far away or too close to the SUT. To avoid ambiguity with positions outside this range or within the measurement range the PSS is evaluated, as treated in Subsection 7.5.3.
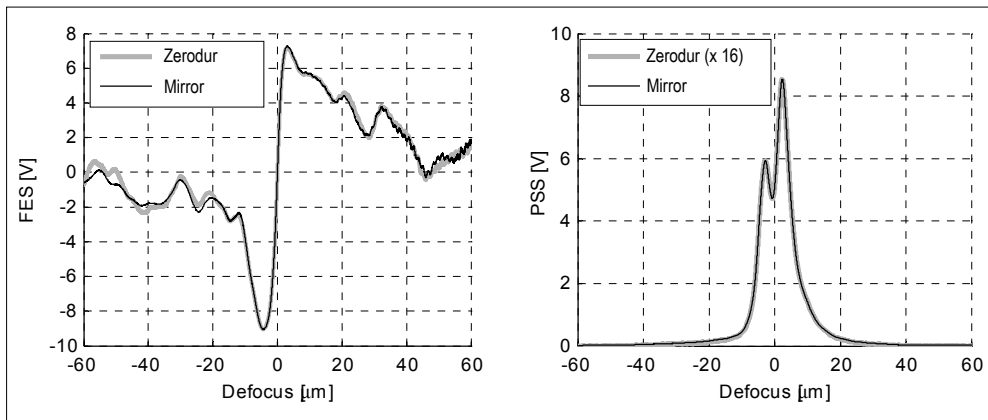
Figure 7.15: Medium-range plot of the FES for measurement on glass and a mirror (left) and of the PSS (right).

Short range signals

The FES close to best-focus is shown in Figure 7.16, left; the PSS and the separate photodiode signals are shown on the right. The part of the FES between -2.5 µm and 1.5 µm seems suitable for measurement. Although upon visual inspection the signals for the Zerodur and the protected aluminium mirror seem equal, there is a deviation of about 30 nm for the extremes of the measurement range. Therefore, a calibration scan must be performed for measurements on surfaces not encountered before.

An intensity difference between the pinhole signals is seen in the figure on the right side. The power of the two beams before the imaging lenses was measured to locate the cause. It was found that the coating of either the differential confocal splitter plate or the differential confocal folding mirror does not conform to requirements. A problem this causes is that the inherent tilt independence of the sensor is compromised, since a decrease in one pinhole signal due to tilt is not fully

compensated by a decrease of the other pinhole signal. Because of the long lead-time of optical components, it was decided not to replace the components. If the resulting tilt dependence is not sufficiently suppressed by aperture correction, the normalization electronics can be replaced by digital normalization, allowing the weaker signal to be amplified before normalization.
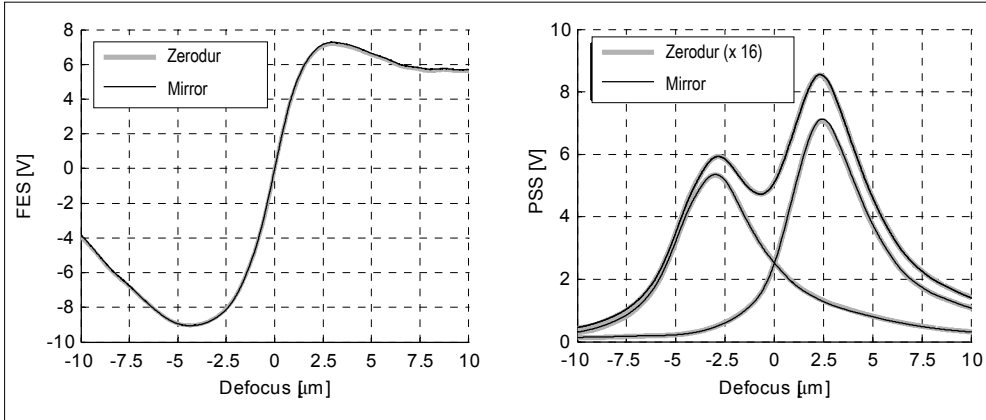
Figure 7.16: Short-range plot of the FES for measurement on a Zerodur flat and a mirror (left) and of the PSS and the separate photodiode signals (right).

An asymmetry with respect to the maximum of the pinhole signals can be observed: the right flanks of the curves are higher than the left flanks. This indicates aberrations in the system. Therefore, it is suspected that some of the optical components do not meet specification. This also leads to an increase of tilt dependence of the sensor. It can be concluded that the optical components must be tested before assembly for subsequent sensors. Tests that investigate the performance of the differential confocal system are presented in Chapter 8.

7.4 SIGNAL PROCESSING AND ELECTRONICS

The electronics that process the sensor signals and control the focusing unit and laser are integrated with the electronics for the rest of the NANOMEFOS machine. After an overview of the signal flow, some of the electronics specific to the sensor are discussed.

7.4.1 SIGNAL FLOW DIAGRAM

A flow diagram of the signals that are related to the sensor's measurements is shown in Figure 7.17.
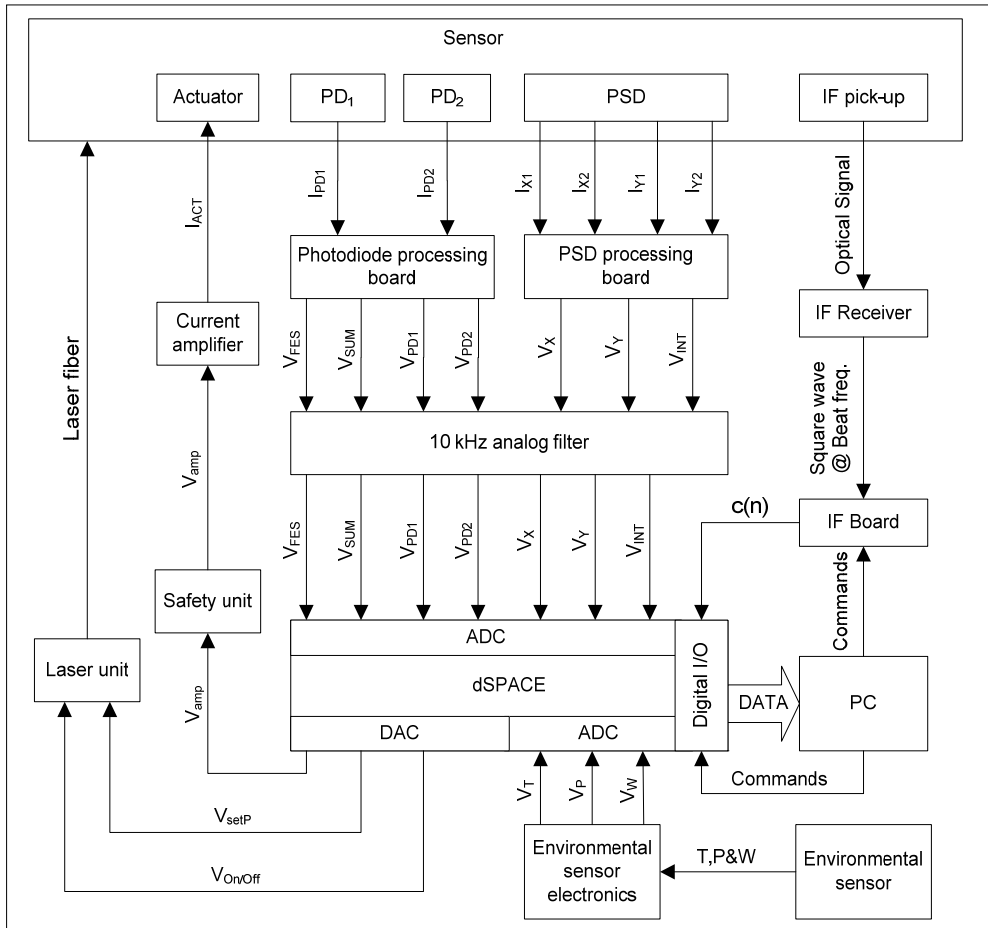
Figure 7.17: Flow diagram of the sensor's signals to and from the sensor.

At the center of the electrical system is a dSPACE real-time data acquisition and control unit, operating at a 10 kHz sample rate. Via a control desk interface on a PC, the user can send commands to the dSPACE unit, controlling the NANOMEFOS machine and sensor. During measurements, relevant signals are sent from dSPACE to the PC, where they are stored for offline processing.

The small current signals of the PhotoDiodes (PDs) and the Position Sensitive Detector (PSD) are processed by two electronics boards. Of the output signals, the Focus Error Signal (FES), photodiode signals and the PSD's x and y signals are normalized. The normalization signals (the summed photocurrents of the PDs and PSD) are also put out. The signals are analog-filtered before they are sampled at 10 kHz by the 16-bits differential input ADCs of the dSPACE unit. Apart from

distance measurement, the normalized FES is used for feedback control of the focusing unit and the PSS is used for focus finding, as will be treated in Section 7.5.

The InterFerometer (IF) pick-up collects the IF measurement and reference beam; the optical signal resulting from interference is sent through a fiber to an Agilent E1709A receiver. The receiver transforms the optical signal into an electrical square waveform at the beat frequency, which is converted into a digital distance signal by an Agilent N1231B interferometer counter board. The digital signal is fed to the dSPACE unit, where it is applied for feedback control.

A sensor close to the measurement volume measures the temperature and pressure of the air and the partial pressure of water vapor. The environmental sensor electronics convert the signals into three analog signals that are sent to the dSPACE unit.

The amplifier setpoints for the voice coil actuator are generated by dSPACE and range from -10 V to 10 V. The signal is routed through a safety unit, which overrides the setpoint so that the focusing unit fully retracts in case the dSPACE unit stops responding. The laser unit is controlled by dSPACE as well; the laser output is sent to the sensor via a fiber.

7.4.2   SENSOR ELECTRONICS

Most of the machine and sensor electronics are placed in a cabinet, which is shown in Figure 7.18. The units that are related to the sensor are indicated, these are: the dSPACE data-acquisition and control unit (1), the environmental sensor electronics (2), the dSPACE in- and output panels (3_a and 3_b), the analog signal filters (4), the focusing unit and ψ-axis amplifiers (5), the safety control unit (6), the laser unit (7) and the interferometer electronics (8).

Laser unit[6]

The laser unit is shown in Figure 7.18, right. It contains the fiber coupled diode laser (13) with integrated Peltier elements and a thermistor for temperature control, a laser driver (12) and a Peltier driver (10). Disturbances in the feed current of the laser and electronics can cause disturbances in the laser output, therefore the laser, laser controller, and Peltier driver are fed by separate analog power supplies (9).

---

[6]*Designed by M.H.C. Janssen of TU/e GTD (Laser and Peltier driver are standard components).*

Six BNC connectors (11) at the back allow access to reference voltages such as the thermistor voltage, and provide the connection to the dSPACE system, which sets the laser power level via a 0 V – 10 V signal, and switches the laser on and off. The status of the laser unit (off, enabled or on) is indicated by two lights in the front panel and communicated to dSPACE via a three-level signal. Stress relief and a protection cover (14) have been added to protect the fiber coupling.
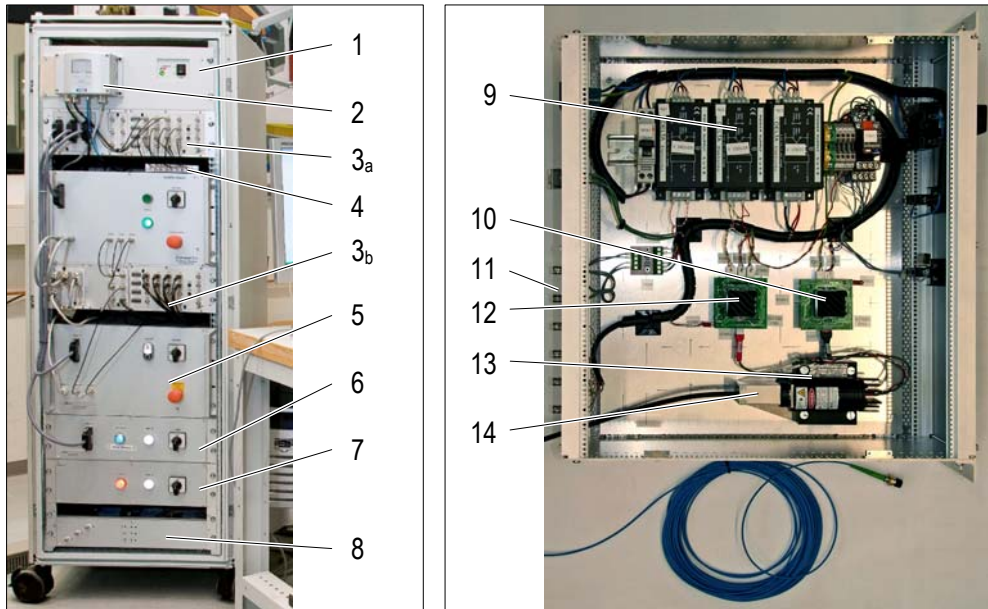
Figure 7.18: NANOMEFOS electronics cabinet (left) and laser electronics (right).

Actuator amplifier[7]

The amplifier for the voice coil actuator is a custom-built analog linear amplifier with current-feedback and a bandwidth of 50 kHz. An analog amplifier is preferred over a PWM amplifier, because of its better response around zero-current and because it is less likely to cause interference in the detector signals of the sensor.

The output current of the amplifier is limited to 2 A and a voltage of ±20 V. Accounting for the resistance of cables, connectors and coil, less than 3 V of the range is needed for steady state operation. The remaining voltage is available to improve the transient response of the focusing unit by countering the effects of back emf and inductance of the coil, which is the reason that high bandwidth current-feedback is desired.

---

[7] *Designed by M.H.C. Janssen of TU/e GTD.*

Shielding and cables

For measurements on glass, the currents generated by the photoconductive PIN diodes are in the order of 15 µA around zero-crossing and the currents generated by the PSD are even smaller. This makes the signals susceptible to electromagnetic interference and bandwidth loss due to cable-capacitance. To prevent this, a four-core, individually shielded, low-capacitance cable is used. Furthermore, the cable length between detectors and processing electronics is limited by placing the processing electronics in the machine's frame.

To shield the bonds between wire and lead pins, they are potted in epoxy that is covered with grounded foil, (1) in Figure 7.19, left. To shield the front of the detector surfaces, the insulating anodization layer has been removed from the bosses over which the PD housings are placed, thus grounding the aluminium structure. Grounded foil (2) is placed over the backside of the PSD as well. To prevent creating a ground loop, both the foil at the PDs and PSD are grounded via the PD cable. A copper wire (3) connects the foil at the PSD to the shielding of the PDs; the shielding of the PSD cable is terminated at the PSD side.
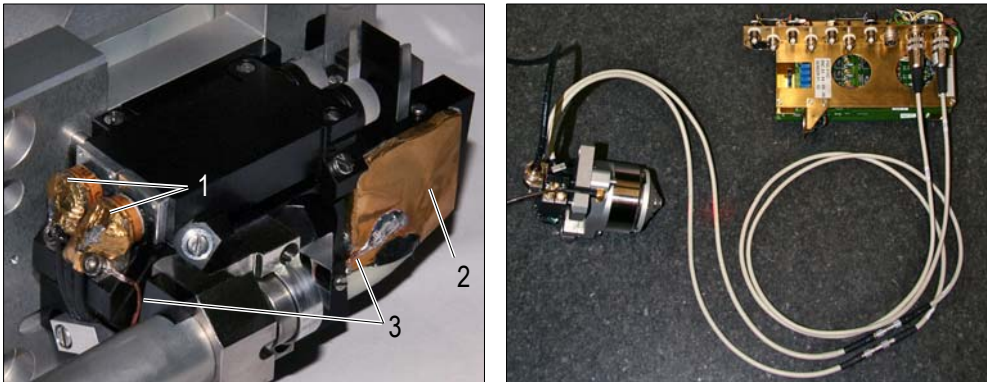
Figure 7.19: Shielding of the detectors (left) and the sensor connected to the PD and PSD processing boards (right).

In Figure 7.19, right, the sensor, signal cables and processing boards are shown. The connectors in the middle of the cables are added so that the cables do not have to be threaded through the machine during mounting and removing of the sensor.

Processing electronics

The PD processing board is the same as employed in the demonstrator setup. Apart from a few small modifications, the PSD and PD processing boards are similar; they are obtained from Philips Applied Technologies. The electronics have a small-signal bandwidth of 150 kHz. The circuit has a resolution of $15 \cdot 10^{-6}$ at 150 kHz and

a total photocurrent of 500 μA (C.S. Kooijman, ND). It has a normalization accuracy of 0.05 % for summed photocurrents of 1 μA to 500 μA.

Analog RC filters

Because the electronics of the sensor have a bandwidth of 150 kHz, while the dSPACE unit samples at 10 kHz, most of the bandwidth is not used. Through aliasing, part of the harmonic content above 5 kHz will shift to the 0 Hz to 5 kHz band, causing an increase of measurement noise. To suppress most of this effect, 10 kHz RC filters are inserted between the signal processing electronics and the dSPACE ADCs.

7.5   FOCUSSING UNIT CONTROL[8]

The controller diagram is presented, followed by a discussion about the controller performance. Finally, the control strategies for interferometer nulling, focus locking and tracking a surface will be treated.

7.5.1   CONTROLLER DIAGRAM

For control of the focusing unit, three sensor variables are relevant, one input: the actuator force F_{act}, and two outputs: the interferometer signal, IF, and the differential confocal signal, FES. The control system switches between three controllers for different situations, one open-loop controller and two closed-loop controllers. A schematic diagram of the system is shown in Figure 7.20.

When the IF signal is unreliable or needs to be reset, the system switches to the feedforward force controller, C_F. To move the objective lens to an approximate c-position, a corresponding force setpoint F_{set} is sent to the controller, which generates a second order profile between subsequent force setpoints.

To position the objective lens accurately along the c-axis, the system switches to the IF controller, $C_{IF,}$. A position setpoint, c_{set} , is sent to the trajectory generator, TG, which creates a third order trajectory. The resulting setpoints are sent to the IF controller, which applies feedback of the IF signal to suppress errors. The IF signal is proportional to the position of the objective lens relative to the rest of the sensor.

To track the Surface Under Test (SUT), the system switches to the FES controller, C_{FES} . It receives no other input than feedback of the FES, which is approximately

---

proportional to the distance of the SUT from best-focus. The trajectory imposed by the SUT can be viewed as a disturbance that must be suppressed.
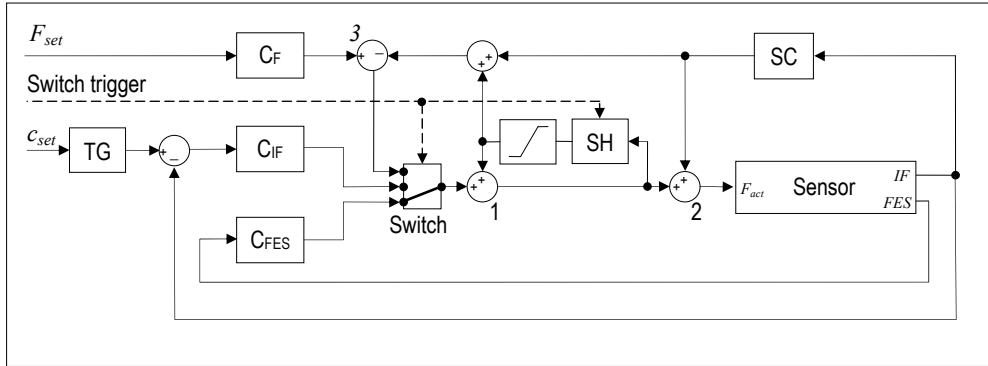
Figure 7.20: Schematic diagram of the focusing unit controller. It contains the sensor, Stiffness Compensation (SC) three controllers, C_F, C_{IF} and C_{FES} , and to allow smooth switching between the three controllers, a switch and a Sample and Hold block (SH).

The force needed to overcome guidance stiffness, represents a relatively large disturbance to both the FES and IF controller. To reduce the effort this poses to the integrating actions of these controllers, a parallel loop that contains a Stiffness Compensation block (SC) is incorporated. The force needed to overcome the guidance stiffness is calculated from the IF signal and added to the controller signal at 2. Because the stiffness varies over the stroke, a third degree polynomial is used for this calculation; the fitting error of the polynomial is smaller than the hysteresis. In this way, the system seen by the IF and FES controller is converted from a mass-spring system into an approximately pure mass system. Since the stiffness compensation should not be applied to the signal of the force controller, it is subtracted from this signal at 3.

To prevent steps in the control signal when switching between controllers, the controller states are reset during switching. Furthermore, a Sample and Hold (SH) loop is placed in series with the switch, and added to the control signal at 1. This signal is subtracted from the signal of the force controller at 3 as well. A saturation block can be included to prevent integrator windup for large values of SH.

7.5.2   PERFORMANCE

The frequency response of the focusing unit to the interferometer and differential confocal outputs are presented. The performance of the present controllers is discussed next as well as options for improvement. To conclude, a typical tracking error while measuring a flat that is tilted 1.6 mm, at 1 rev/s is presented.

Open-loop frequency response

To characterize the focusing unit's open-loop dynamic behavior, its frequency response to both the interferometer and differential confocal output has been measured. Thereto, the multisine method is used: a method in which a plant is fed a signal only containing sines that fit an integer number of times in the signal length, and with frequencies up to the Nyquist frequency. The frequency response can be obtained by measuring the content of these bins in the output signal.

The bode plot of the response to the IF output is shown in Figure 7.21. Because the stiffness and force factor vary over the stroke, the frequency response is measured at four positions along the guidance axis: -2.1 mm (1), -1.3 mm (2), 0 mm (3) and 1.9 mm (4), corresponding to force offsets of -2.2 N, -1.1 N, 0 N and 1.1 N.
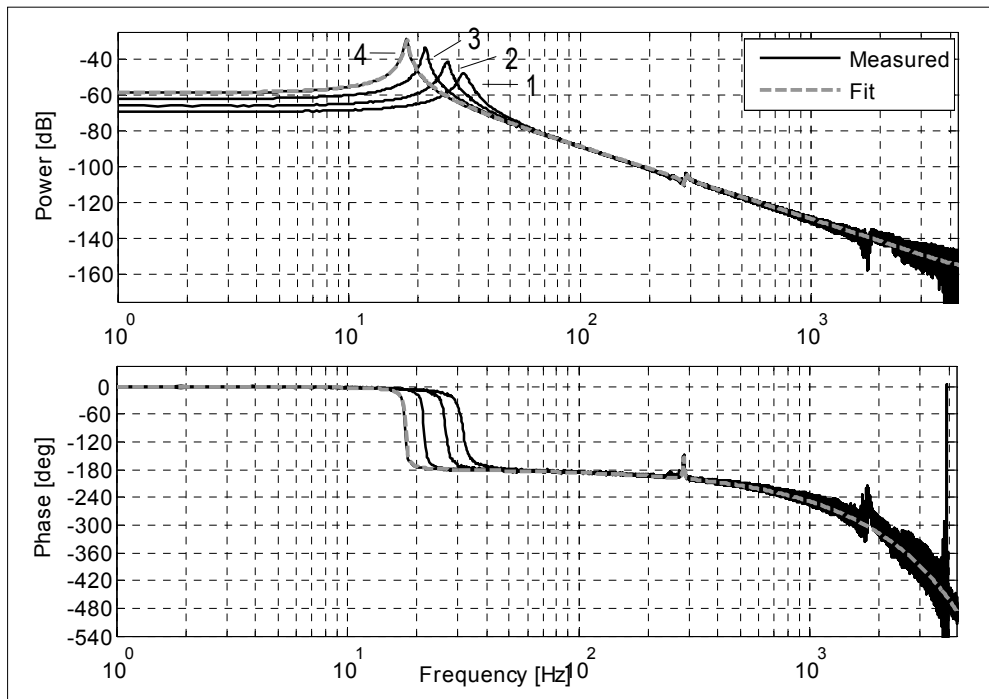
Figure 7.21: Bode plot of the response of the focusing unit to the interferometer output, with power in m/N.

The focusing unit response closely approximates that of a pure mass-spring-damper system with low damping ratio: about 0.018 for the fitted system. Due to the non-constant stiffness, the eigenfrequency in driving direction varies between 18 Hz and 31 Hz. The first parasitic resonance is an antiresonance-resonance pair at 285 Hz, which is further discussed after the response to the FES output has been

presented. The next resonances are resonances of the focusing unit's guidance mechanism at about 1700 and 1800 Hz, higher than the 1400 Hz and 1700 Hz predicted by finite element analyses. This discrepancy is thought to arise because in the model the fill-factor of the coil is not taken into account, leading to an overestimated mass of the coil. Since all resonances have little effect on the amplitude and phase lag, the control bandwidth of the IF controller is expected to be limited by the phase lag and noise of the interferometer system.

To measure the frequency response to the FES output, the objective lens is positioned at best-focus by the interferometer controller and perturbations are added to F_{act}. The influence of the IF controller is calculated out of the results to obtain the open-loop response. Normally, a control loop with low stiffness would be applied; here however, a stiff controller is needed to stay within range of the FES. For low frequencies, the data is unreliable because of high disturbance suppression of the controller. Where the response is predominantly determined by the mass, the results are accurate. Measurement at one c-position suffices since at high frequencies the response is independent of stiffness in driving direction. The results for both the IF and FES output are plotted in the Bode plot in Figure 7.22. The analog filter was not used for these measurements.
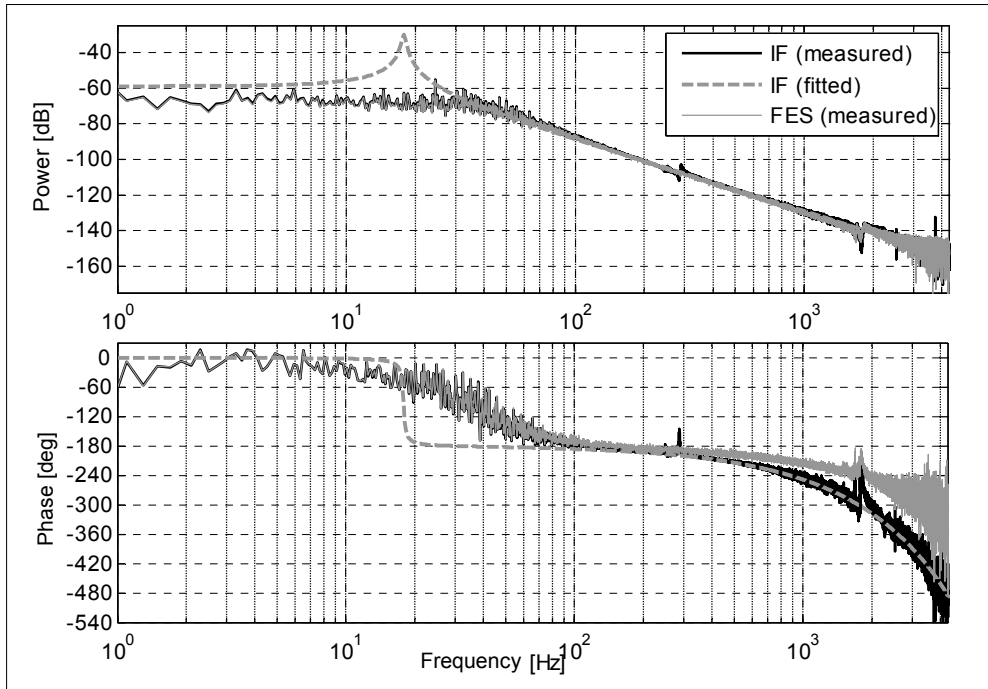
Figure 7.22: Bode plot of the frequency response to the FES output and the IF output as well as the response of the model fitted to the IF; power is in m/N.

The antiresonance-resonance at 285 Hz that is present in the IF response is not observed in the FES response. It is suspected that the ψ-axis unit, to which the field assembly and interferometer optics are rigidly connected, is rotating on its mount in φ-direction. Figure 7.23 depicts a free body diagram in which the situation is modeled as two translating masses: the mass of the ψ-axis unit, M_1, and the moving mass of the focusing unit, M_2.
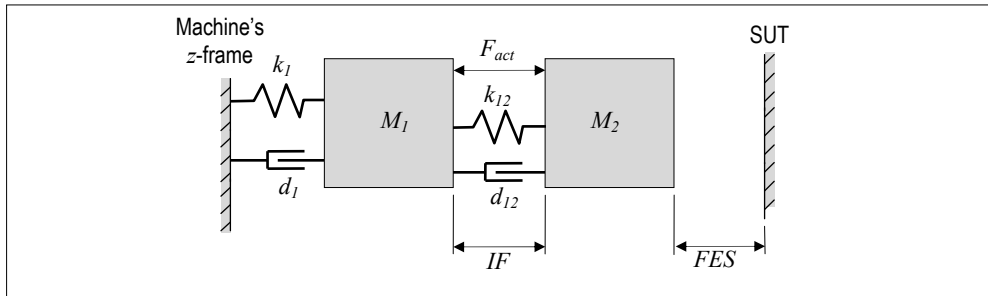
Figure 7.23: Free body diagram of the double mass-spring-damper model comprised of the sensor and ψ-axis mass (M_1) and the focusing unit's moving mass (M_2).

The reason that the antiresonance-resonance is observed in the IF output but not in the FES output, is because the input is a force between M_1 and M_2, between which the interferometer measures the distance, while the differential confocal system measures from M_2 to the fixed SUT. A model which closely approximates the IF response has been fitted and is depicted in both the Bode plots, the model's response at 285 Hz can be seen in Figure 7.21. The parameters of the fitted model are $M_1 = 5.0$ kg, $M_2 = 0.046$ kg, $k_1 = 1.6 \cdot 10^7$ N/m, $k_{12} = 485$ N/m, $d_1 = 153$ N/m/s and $d_{12} = 0.17$ N/m/s, which seem realistic values. Since displacement of the ψ-axis is measured by the machine's r- and z-interferometers, the measurement error caused by vibration of the ψ-axis is compensated for.

For the FES, the phase lag is hardly affected by the resonances at 1700 and 1800 Hz. The phase lag in the frequency response to the FES output is considerably smaller than that of the IF output. Almost all the phase lag in the FES signal beyond 180° can be contributed to the 10 kHz sample rate of the dSPACE unit, confirming that the sensor electronics cause hardly any phase lag at these frequencies.

Controller performance

For the time being, both closed-loop controllers use a PID-type controller consisting of a lag/lead (integrating action) and a lead/lag (differential action) compensator and a second order roll-off filter. The present controllers have not been optimized yet and have been designed with large margins. The IF controller

has a bandwidth of 100 Hz and the FES controller has a bandwidth of 250 Hz. The response of the interferometer controller is not critical for measurements of continuous surfaces, however good IF controller performance is needed for measurement of discontinuous surfaces.

Because measurement errors increase for increasing tracking errors, the sensitivity to disturbances of the FES plant has significant effect on measurement uncertainty. The profile imposed by the SUT, of which the main part of the frequency content is expected to be in the range from 1 Hz to 10 Hz, is the dominant disturbance. Therefore, the aim is to improve disturbance suppression in this frequency band.

Various methods can be applied to this end. Shifting the lead/lag to higher frequencies increases the bandwidth. This extra bandwidth can stabilize a system with a higher loop gain, thus improving low frequency disturbance suppression. From the frequency response shown in Figure 7.22 it can be expected that the bandwidth can be increased to 800 Hz. Furthermore, a double integrator can be applied in the range where the largest disturbances are expected. For measuring heavily freeform surfaces, repetitive control can be considered as well.

Tracking error

In Figure 7.24, left, the FES is shown for a stationary SUT with the FES controller locked onto focus. The periodic deviation has a frequency of 50 Hz; it is caused by hum in the amplifier signal. In open-loop, the PV amplitude of this disturbance is about 150 nm. With the loop closed, the amplitude is suppressed by a factor of 10, to about 15 nm. Since this disturbance is measured by both the interferometer and the differential confocal system, it does not cause significant measurement errors.
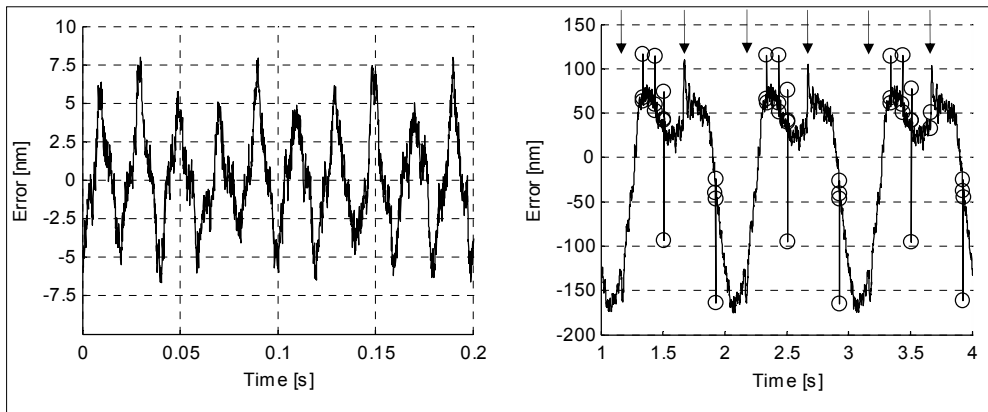
Figure 7.24: Tracking error for a stationary SUT (left) and for a SUT that is tilted 1.6 mm and rotating at 1 rev/m (right).

In Figure 7.24, right, the differential confocal signal is shown during tracking of a SUT that is tilted 1.6 mm while it is rotating at 1 rev/s. The maximum tracking error is about 170 nm. It is expected that with an improved FES controller this tracking error can be reduced substantially. The tracking error is asymmetric about the zero-crossing. Since effects of varying stiffness over the stroke should be suppressed by the stiffness compensation, it is not yet known what causes this asymmetry.

Some spikes appear in the FES (encircled in Figure 7.24, right); these spikes are caused by dust and surface defects. Because these artifacts disturb specular reflection, they are identified using the photodiode sum signal and excluded from the measurement data. When the FES controller's bandwidth is increased, it might be fast enough to react to some of these artifacts; in that case, the controller can evaluate the photodiode sum signal to disregard the artifacts. The spikes indicated by the arrows reproduce regardless of the position of the sensor and therefore cannot be surface artifacts. Since they occur where the velocity is zero, they are probably caused by virtual play, as discussed in Subsection 7.2.2.

7.5.3   CONTROL STRATEGY

Software running on the dSPACE unit decides when to switch between controllers and which setpoints are fed to the controllers. The strategy for this real-time decision processes will be discussed here for three different tasks: interferometer nulling, finding and locking onto focus and tracking a surface.

Interferometer nulling

Since interferometry is a relative measurement method, the interferometer's measurements have an arbitrary offset after the signal has been interrupted. Hence a mirror is incorporated into the ψ-axis housing, which serves as a reference; see Figure 7.25.

First, force control mode is used to retract the focusing unit, and the interferometer is nulled with the end-stop of the guidance serving as a course reference. The approximate stop position is subtracted from the interferometer signal, which can now be applied for course positioning. Next, the actuator force is reduced to zero and the IF controller is switched on. A position setpoint that fully extends the focusing unit is given, while the FES and photodiode sum signal are monitored to detect if the nulling mirror is in range of the differential confocal system. When this happens, the system switches to the FES controller, thus locking onto the FES signal. Because this determines the position of best-focus relative to the ψ-axis centre line, the interferometer signal can now be accurately referenced.

Figure 7.25: The sensor while referencing the interferometer to the reference mirror.

The equilibrium position of the guidance for zero actuator force changes according to its orientation relative to the gravity vector. To eliminate actuator dissipation due to weight compensation, the average objective lens position over a track can be chosen as a function of ψ-angle.

Controller states for locking and tracking

To achieve locking onto focus and tracking of a SUT, the algorithm that switches between controllers and generates setpoints, must determine in what region the SUT is relative to best-focus. By evaluating the magnitude of the FES and the Photodiode Sum Signal (PSS), the algorithm distinguishes between four such regions, corresponding to four controller states. These states are:

- "*blind*": it cannot be deduced whether the SUT is behind or in front of best-focus,

- "*far*": the SUT is too far away from the objective lens, but close to best-focus,

- "*ok*": the SUT is in the region where the FES is near-linear and the FES controller can be used, and

- "*near*": the SUT is too near to the objective lens, but close to best-focus.

Typical PSS and FES curves are depicted in Figure 7.26, left and right, respectively. The four states can be identified by evaluating whether the FES and PSS are below, between, or above thresholds that are indicated as well. Furthermore, the ranges of the states are also depicted in Figure 7.26, right.
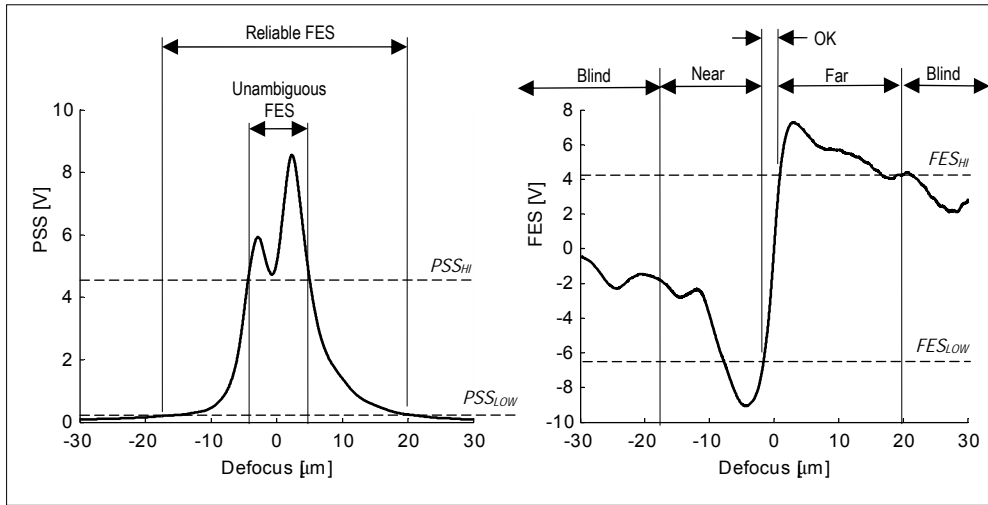
Figure 7.26: Typical Photodiode Sum Signal (PSS) and FES curves with various thresholds used to switch controller states.

An overview of the states and the corresponding conditions is given in Table 7.1. The *blind*-state corresponds to the range where the photodiode signals are so small that the FES is unreliable. For higher *PSS,* the FES is reliable but ambiguous in the sense that one value of the FES might correspond to multiple positions of the SUT. In this region the sign of the FES does indicate, however, whether the objective lens is too far from or too close to the SUT, corresponding to the states *far* and *near*. When the PSS is between PSS_{LOW} and PSS_{HI} and the FES is between FES_{LOW} and FES_{HI}, the FES is reliable and unambiguous; this is the *OK*-state where the FES controller can be used.

State	Condition	Action when tracking
Blind	$PSS < PSS_{LOW}$	Retract
Far	$(PSS_{LOW} < PSS < PSS_{HI}$ & $FES > 0)$ | … $(PSS > PSS_{HI}$ & $FES > FES_{HI})$	Apply fixed negative force
OK	$PSS > PSS_{HI}$ & $FES_{LOW} < FES < FES_{HI}$	Switch to FES control
Near	$(PSS_{LOW} < PSS < PSS_{HI}$ & $FES < 0)$ | … $(PSS < PSS_{LOW}$ & $FES < FES_{LOW})$	Apply fixed positive force

Table 7.1: Controller states with corresponding conditions regarding Photodiode Sum Signal (PSS) and normalized Focus Error Signal (FES), as well as the actions to be taken during tracking.

Hysteresis in the thresholds makes the decision process noise-robust. Because the reflectivity of the SUT varies for different materials, the thresholds used to evaluate the PSS should be adapted for different surface reflectivities; these values may be stored in a database of various surface types.

Focus finding and locking

When a user gives the command to start tracking a surface, the focusing unit fully retracts after which it starts to extend gradually in IF control. As the SUT comes in range, the system enters the *far*-state and subsequently the *OK*-state. When the *OK*-state is reached, the system switches to the FES controller, thus locking onto best-focus.

Tracking a surface

While tracking a surface, the focusing unit might not be able to follow the surface, so that the signals get outside the *OK*-range. In case the controller ends up in the *far*- or *near*-state, a large negative or positive force is applied to return to the *OK*-state. If the blind state is detected, the focusing unit is fully retracted to avoid a collision between the sensor and the SUT.

Alternatively, dust or a surface defect might disturb the signals, causing the controller to jump from the *OK*-state to a false *far*-, *near*- or *blind*-state, which would trigger an unwanted response and may lead to loss of focus. To avoid this, the PSS and FES signals are run through a recursive moving average filter before they are fed to the switching algorithm.

In this chapter, realization of the sensor has been discussed as well as testing of some of the subsystems. Furthermore, the sensor electronics and control of the sensor have been addressed. With the sensor finished, more elaborate tests have been performed from which the performance is evaluated; this is covered in the next chapter, together with calibration of the sensor. Full surface measurements made with the sensor installed in the machine are presented from which the effectiveness of aperture correction is evaluated.

8   EXPERIMENTAL RESULTS AND CALIBRATION

Tests of the performance of the sensor's measurement systems are presented. Next, calibration is discussed including a calibration method for tilt dependent measurement error of distance sensors. Based on the test results and taking into account the proposed calibrations, the attainable measurement uncertainty is estimated. The chapter is concluded with NANOMEFOS measurement results of a tilted optical flat.

8.1   DISCUSSION ON MEASUREMENT UNCERTAINTY

Although the measurement uncertainty is estimated in Section 8.4, it is useful to discuss some aspects of it here. The reason is that in the first section of this chapter, measurement results are presented on which the uncertainty estimation is based. To facilitate comparison between the estimated uncertainty contributions and the experimental results from which they are derived, contributions to uncertainty are summarized in short tables throughout the first section. In Section 8.4, these tables are combined and an estimate of uncertainty is calculated.

Because the requirements regarding measurement uncertainty are task-specific, uncertainty is estimated for situations that are representative for three different measurement tasks. These are the measurement of Rotationally Symmetric (RS) surfaces, Medium FreeForm (MFF) surfaces and Heavily FreeForm (HFF) surfaces. For these measurement tasks, different parts of the range of the acceptance angle, object lens stroke and the differential confocal measurement range are used, which has consequences for the achievable measurement uncertainty. Typical values of these characteristics are listed in Table 8.1. A fourth case is added that is not representative for typical freeform measurements but that

illustrates the performance of the sensor if it is employed over its Symmetric Full Range (SFR). Even though the measurement range of the differential confocal system is from -2.5 µm to 1.5 µm, the tracking error is chosen to be ± 1.5 µm, because it is assumed symmetric around the zero-crossing.

Measurement characteristic Surface type	Object lens stroke	Tracking error	Tilt angle
Rotationally Symmetric (RS)	< 0.1 mm	< 100 nm	≈ 0°
Medium FreeForm (MFF)	< 2.0 mm	< 200 nm	< 2°
Heavily FreeForm (HFF)	< 5.0 mm	< 500 nm	< 5°
Symmetric Full Range (SFR)	< 5.0 mm	< 1.5 µm	< 5°

Table 8.1: Three measurement tasks with characteristics that are representative for these tasks and which influence measurement uncertainty. The fourth option is included to illustrate the performance of the sensor when employing its symmetric full range.

Some of the error sources that might be significant have not yet been measured, while some other error sources have not been measured over the whole range or at sufficiently small intervals. Consequently, a complete formal evaluation of measurement uncertainty such as described in (Taylor, et al., 1994) is not possible at this time.

Given the importance of the sensor's measurement uncertainty, it has been decided to pursue a rough estimate instead. The uncertainty requirement is stated over a 2σ interval, this value is however not obtained from measurements directly. To arrive at values from which the uncertainty can be estimated nevertheless, two situations are distinguished. If a source of uncertainty is random or pseudo random, such as noise, its RMS value is calculated. For a large number of samples this is a good estimate of the standard deviation of the source and the 2σ contribution to uncertainty is estimated by multiplying the measured RMS value by 2.

In the case where the component is systematic in nature, such as measurement error due to tilt, the single-sided maximum of that error is used, which is easily obtained from experimental results. This method tends to overestimate the contribution to 2σ uncertainty of a source. Hence, it will result in a conservative estimation of the uncertainty and thus leave margin for sources that have not been taken into account. For practical reasons, sources are considered uncorrelated and are therefore summed by quadratic addition. It is likely that in reality, some error sources are strongly correlated since they are dependent on the shape of the SUT, such as errors due to surface tilt and due to the position of the object lens. Because the correlation is not known, in these cases, squared addition can lead to either over- or under-estimation.

The error values stated in the tables will be uncorrected values until measurement uncertainty is estimated in Section 8.4, unless stated otherwise. To provide an overview of various sources and to allow comparison of their contribution to uncertainty, some of the sources of which the contribution is small enough to neglect, are included in the tables nevertheless.

8.2 MEASUREMENT TESTS

Various aspects of the sensor have been tested before and after integration in the machine. Because of the aforementioned failure in the adhesive bonds, the alignment of the optical system has changed significantly between tests. Not all tests have been repeated after re-assembling the detached components, mainly since the time lost with re-alignment already put strain on the schedule, while the NANOMEFOS machine relied on the availability of a suitable sensor. Therefore some of the results presented here, are obtained with a differently aligned system than currently installed in the machine.

8.2.1 DIFFERENTIAL CONFOCAL SYSTEM

There are two ways to measure the Focus Error Signal (FES) of the sensor: the focusing unit can be locked and the Surface Under Test (SUT) is scanned axially through focus, or a stationary SUT can be placed in front of the sensor and the focusing unit is actuated to scan through focus. Both methods are used, because they each have specific advantages, depending on which test is to be performed.

Measurement noise and resolution

The nominal resolution of the sensor can be estimated by measuring the noise level in the measurement signal and dividing it by the local sensitivity. The noise level of the FES has been measured at zero-crossing and four locations along the FES curve. For this purpose, the sensor is mounted in a stand and supported on foam for vibration isolation. The guidance is mechanically locked and the SUT is manually adjusted to the desired position. The signals have been recorded for measurements on a mirror and on Zerodur at 200 kHz over 2 seconds, with a 16-bit DAQ.

As the sensor electronics have a bandwidth of 150 kHz while the machine samples at a maximum rate of 10 kHz, most of the bandwidth is not used. Through aliasing, part of the harmonic content between 5 kHz and 150 kHz will however shift to the 0 to 5 kHz band. To suppress this effect, a 10 kHz RC filter is inserted between the processing electronics and the DAQ. Therefore, noise measurements up to 100 kHz have also been performed with a 10 kHz RC filter in place. Since the Nyquist frequency is 10 times higher than the cutoff frequency of the RC filter, the power at the Nyquist frequency is attenuated 20 dB, or a factor 100.

The RMS noise values found for the aluminium coated SUT are 0.17 nm when filtered and 0.27 nm at full bandwidth; for the Zerodur sample these values are 0.16 nm and 0.41 nm, respectively. The RMS values along the filtered FES are shown in Figure 8.1 for an aluminium-coated (left) and a Zerodur (right) SUT.
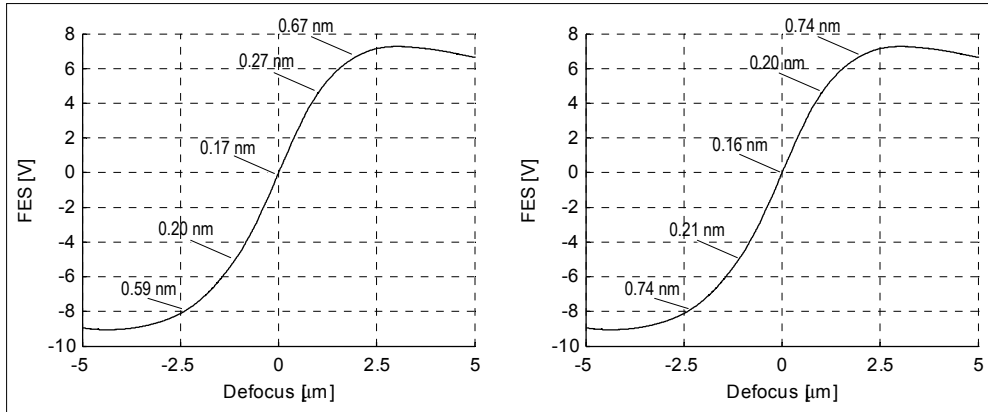
Figure 8.1: RMS noise levels, at different parts of the FES curves for measurement on a mirror (left) and uncoated Zerodur (right) with a 10 kHz RC filter applied.

As expected, the noise level increases for decreasing sensitivity, increasing measurement bandwidth, and decreasing reflectivity. For the filtered signal, though, the noise levels for high and low reflectivity are almost identical. It is thought that for the filtered signal, the noise is dominated by low frequency noise sources that are not intensity dependent. For the unfiltered signals, however, the intensity dependent shot noise gains in significance due to its white nature and the large frequency range over which it adds up. Power spectral density plots of the signals indeed reveal increased power levels at low frequencies.

To quantify how much the different parts of the measurement chain contribute to the noise level, two reference experiments have been performed. To determine the noise level of the DAQ, the ADC channel has been sampled while it is short-circuited with a 50 Ω resistor. To find the noise generated by the signal processing electronics, 1.2 MΩ resistors were connected to the inputs instead of photodiodes. At a 10 V bias voltage as applied to the photodiodes, this yields a summed current of 17 µA, similar to that when measuring Zerodur at zero-crossing. The DAQ has been found to have an RMS noise level equivalent to 0.10 nm and the RMS noise level of the electronics and DAQ combined is equivalent to 0.12 nm. The largest part of the noise can thus be attributed to the optical system and detectors.

The power spectral density plots from 0 to 100 kHz are depicted in Figure 8.2 for a mirrored SUT, left, and for a Zerodur SUT, right. The power spectral density

estimates for the ADC (1), the signal processing unit and ADC combined (2), the sensor with 10 kHz RC filter (3) and unfiltered sensor with 150 kHz bandwidth (4) are shown. To estimate the power spectral density the Thompson's multitaper (Mathworks, 2007) method is applied with a time-bandwidth product of 6, as a tradeoff between variance and memory use; the data is referenced to an equivalent of 1 nm at zero-crossing.
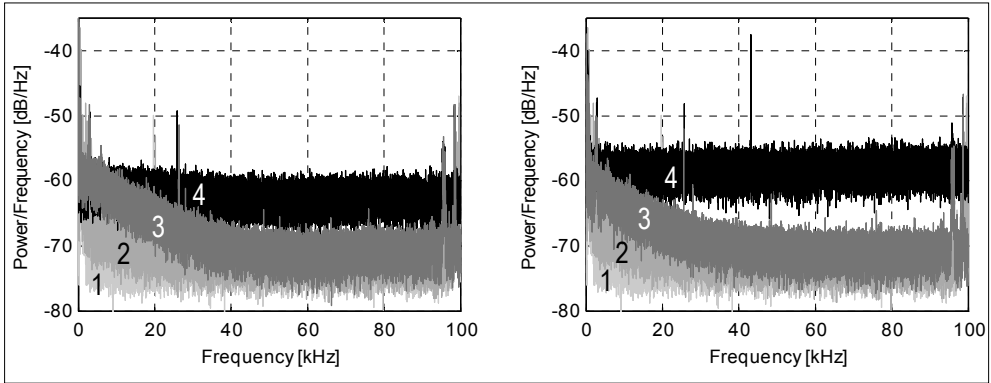
Figure 8.2: Power spectral density plot with a reference of 1 nm, of the ADC (1), the signal processing unit and ADC (2), the complete sensor with 10 kHz RC filter (3) and the 150 kHz bandwidth unfiltered sensor (4), for a mirror (left) and Zerodur (right).

It can be seen that the RC filter pulls the noise level down to almost the base level of the ADC. Some of the peaks, such as the ones at 20 kHz are present in the ADC signal but disappear when the electronics are connected, probably because the input impedance influences oscillations in the ADC. The peak at 45 kHz in the unfiltered signal for a Zerodur SUT is not observed in the filtered signal. Whether this frequency is introduced by aliasing or is a temporary effect is unknown; it was observed multiple times. As the noise levels satisfy the requirements, no actions have been taken to reduce them further or to identify sources of resonance peaks.

Since the noise varies over the range, its contribution to the measurement uncertainty will be different depending on the tracking error. The 1σ measurement uncertainty is equal to the RMS noise values; because 2σ uncertainty will be used for uncertainty estimation, this value is stated in Table 8.2.

Measurement task **Error / uncertainty source**	**RS**	**MFF**	**HFF**	**SFR**
Uncertainty due to noise (2σ)	0.3 nm	0.4 nm	0.5 nm	1.5nm

Table 8.2: Contribution to the measurement uncertainty of the noise in the differential confocal system.

Repeatability of FES curve

For these tests, the guidance mechanism is locked and the SUT is mounted on a flexure parallelogram driven by a piezo stack. The high voltage amplifier driving the piezo actuator is connected to a function generator, while the displacement of the SUT is measured with a nm-resolution capacitive probe. For measuring the repeatability of the FES, this method is favored over measurements in the machine with a stationary SUT, because it is less noisy. Due to the thermal sensitivity of the loop from the objective lens to the SUT, this method is not suitable for measurement of long-term zero-crossing drift. Local sensitivity changes in the FES, however, can be measured accurately in this way.

To evaluate FES repeatability, the response of the system at a certain moment must be characterized in a universally applicable form. Therefore, multiple scans over the near-linear part of the FES are recorded and a 13th order polynomial is fitted. The fitting residual is shown in Figure 8.3, left.
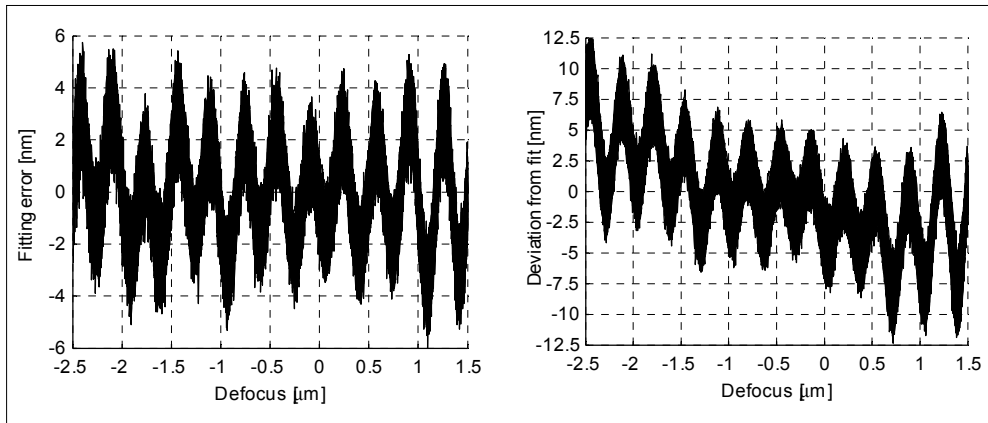
Figure 8.3: Fitting residual of a polynomial fitted through 28 FES curves (left), and the difference between the polynomial and 55 FES curves recorded 5 days later at a 0.7° C higher temperature and just after start up of the laser and electronics (right).

The RMS value of the signal is 1.9 nm, which is within acceptable limits. The ripple as observed in the demonstrator is visible here as well, but far less pronounced. Remarkably, the periodicity of the ripple is closer to what is expected for interference of light with a wavelength of 640 nm: the measured periodicity is between 306 nm and 310 nm. The periodicity of the ripple in the demonstrator was 273 nm.

To test the repeatability of the sensor, 55 FES curves were recorded 5 days later, just after startup of the laser and electronics and at a temperature 0.7° C higher than during measurement of the calibration data. The deviation of these curves from the

calibration polynomial can be seen in Figure 8.3, right; the RMS value of the deviation is 4.1 nm. The deviation due to long-term drift can be prevented by inline calibration before and after surface measurement, this is discussed in Section 8.3.

The sensor as it is currently installed in the machine has not been tested for long-term stability, since this can be calibrated inline. The influence of the cyclic error, however, is relevant and measured by scanning the sensor through focus on the nulling-mirror 12 times. The resulting residual is shown in Figure 8.4; to facilitate visual comparison the vertical axis is scaled identical to the one in Figure 8.3, left.
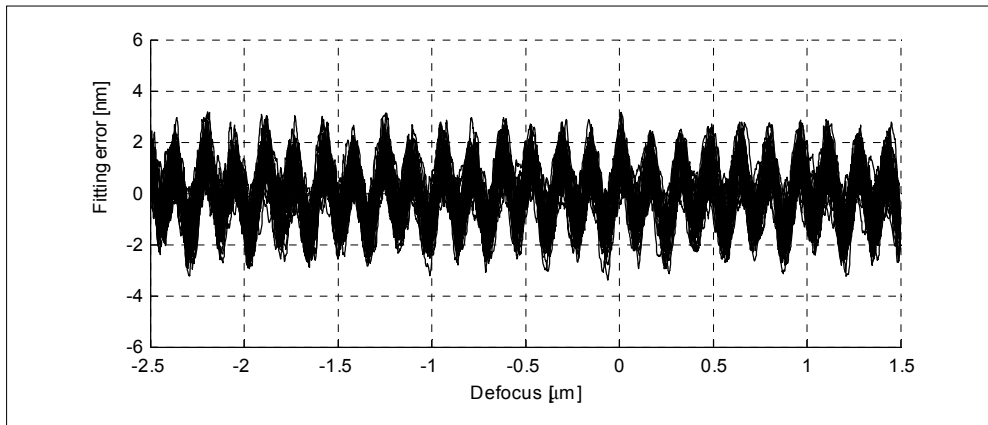
Figure 8.4: Residual of polynomial fitted through 12 FES curves of the current sensor.

It can be seen that the periodic disturbance is less pronounced and that the period is shorter than in Figure 8.3, left. Probably these differences are caused by the changes in alignment. At 157 nm, the period is about half that observed in earlier setups. This value is close to the 2nd harmonic which has a periodicity of 160 nm, corresponding to dual pass interference.

The RMS of the deviation *including* noise is 1.3 nm. For estimating the contribution to the measurement uncertainty, the maximum single sided error at the centre of the noise band is taken: this is 1.1 nm. Because the maximum error does not vary over the range, it has the same value for each of the tracking errors of interest, as is represented in Table 8.3.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Error due to cyclic error (max)	1.1 nm			

Table 8.3: Contribution to the measurement uncertainty of the cyclic error.

Dependence on objective lens position

Collimation error, aberrations, misalignment and lateral error motion of the guidance mechanism can lead to dependence of the FES on the c-position of the object lens. Both for experimental determination of this dependence as well as for calibration purposes, it is convenient to distinguish between zero-crossing shift and change of the shape of the FES due to translation of the object lens. The zero-crossing shift of the FES for different c-positions has been measured with the sensor installed in the machine. A mirror is mounted to the objective lens assembly and adjusted to best-focus with an elastic mechanism. After the adjustment mechanism of the mirror has stabilized, the focusing unit is translated through its axial range using the IF controller, while recording the FES and the interferometer signal. The experiment has been performed multiple times and reproduced within the noise band; the resulting curve is shown in Figure 8.5, left.

The curve of the zero-crossing shift due to the stroke of the object lens is smooth and reproduces well. Because the position of the object lens is well known from the interferometer signal, it is thought that with the 5th order polynomial that has been fitted through the data, the zero-crossing shift due to object lens stroke can be corrected to at least 0.5 nm over 2 mm stroke and 1 nm over the full stroke.
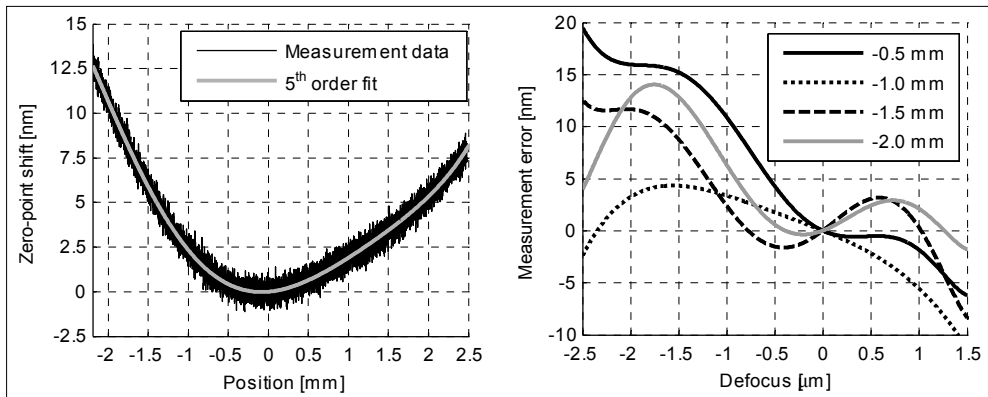
Figure 8.5: Zero-crossing shift of the FES as a function of c-position of the objective lens (left) and measurement errors of the FES as a function of defocus for various c-positions of the objective lens (right).

The change of shape of the FES curve for various c-positions of the objective lens has been measured when the sensor was not yet installed in the machine and with a system alignment different from the current sensor. The guidance mechanism is locked at five positions and the SUT is scanned through the measurement range with a piezo and elastic guidance. The recorded FES curve at the middle of the range of the guidance mechanism is used as reference to which the other curves are

compared. To this end, a 13th order polynomial is fitted through the reference curve. This polynomial serves to relate the voltages of the other curves to values of defocus. To obtain the measurement error, the corresponding values of defocus of the reference curve are subtracted from these values. A plot containing the resulting curves of measurement errors is shown in Figure 8.5, right.

For practical reasons, the measurements have only been performed for negative values of *c* (extension of the focusing unit). It can be seen that the deviations of the various curves differ significantly for the *c*-positions investigated. If these measurements are to be applied for calibration purposes, the measurements must not only be repeated with the current alignment and over the whole range, but also taken at smaller *c*-intervals to allow interpolation between calibration curves. Such measurements can quite easily be performed now that the sensor is installed in the machine so that the IF control can be used to scan the object lens through focus at various *c*-positions. This error is small however for small tracking errors, and consequently has relatively little influence on the overall measurement uncertainty. Therefore, this error is not considered worth the effort involved with calibration.

Again, the maximum single sided error at the centre of the noise band is of interest. Inherently, the zero-crossing shift is independent of tracking error; the maximum occurring FES error on the contrary, is strongly dependent on the tracking error. The estimates taken from the experimental data are stated in Table 8.4 together with the estimated value of uncertainty of the zero-crossing shift after correction.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Zero-crossing shift due to lens stroke (max)	-	2.5 nm	13 nm	13 nm
Zero-crossing uncertainty after correction (2σ)	-	0.5 nm	1 nm	1 nm
FES error due to lens stroke (max)	-	1.5 nm	4 nm	20 nm

Table 8.4: Uncorrected measurement errors due to the stroke of the object lens.

Tilt dependence

To investigate the influence of surface tilt, the FES has been recorded with the SUT perpendicular to the measurement direction and for a surface tilt of 4° in four orientations: in positive and negative α and β-directions. The curves of the SUT are shown in Figure 8.6, left.

To determine the Tilt Dependent Error (TDE) the same approach as for the axial position dependence has been applied: the curve for the perpendicular SUT is used as reference and the TDE of the other curves is calculated with a 13th order polynomial. The TDE curves can be seen in Figure 8.6, right.
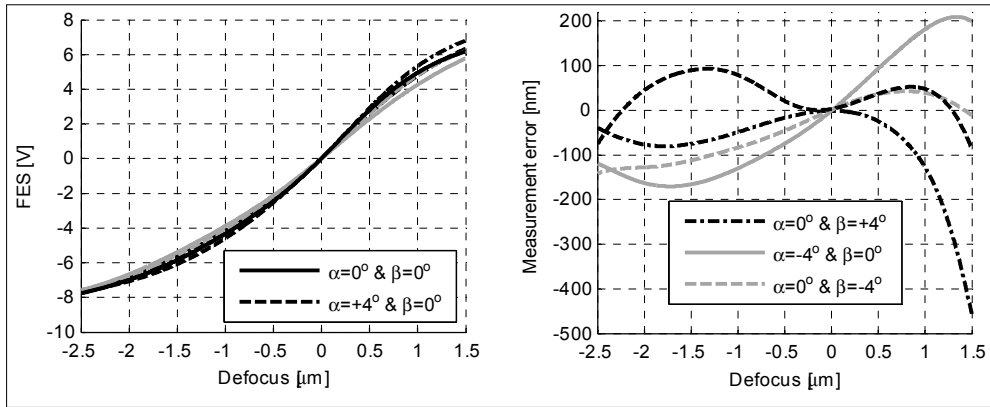
Figure 8.6: FES curves for various surface tilts (left) and the resulting measurement errors (right) for the initial alignment of the sensor.

To determine how much the TDE has changed after reassembly following the adhesive failures, the measurements have been repeated with the sensor installed in the machine. The results of these measurements are shown in Figure 8.7.
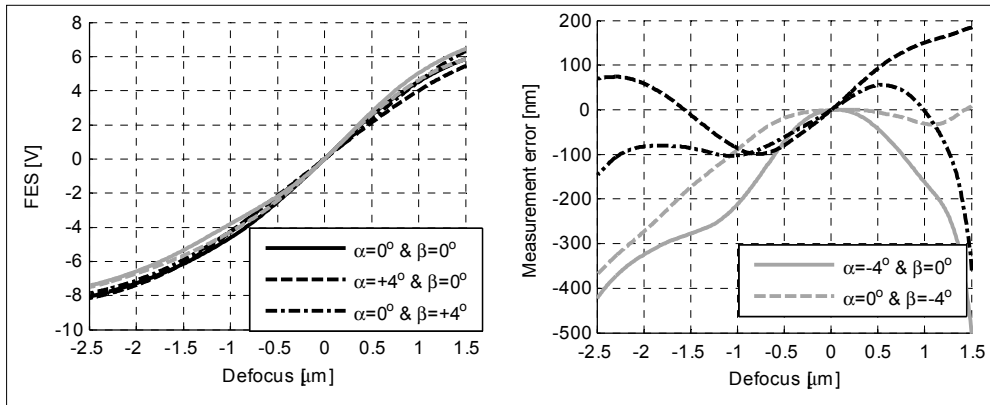
Figure 8.7: FES curves for various surface tilts (left) and the resulting measurement errors (right) for the sensor as installed in the machine.

To enable aperture correction, zero-crossing shift due to surface tilt must also be taken into account and TDE curves for defocus must be taken at orientations and amounts of tilt that are close enough to allow interpolation between TDE curves. Because of the size of the calibration space, such measurements are best automated, for which the NANOMEFOS machine can be used. The measurement of zero-crossing shift due to surface tilt requires a more elaborate procedure that is discussed separately in Subsection 8.3.3. This procedure also allows measurement of the TDE over the range.

The maximum single sided FES error is taken from Figure 8.7, right, and depends strongly on the tracking error; the values are stated in Table 8.5.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
FES error due to 4° surface tilt (max)	-	37 nm	92 nm	500 nm

Table 8.5: Uncorrected FES errors due to 4° surface tilt.

8.2.2 APERTURE CORRECTION SIGNALS

The nominal tilt sensitivity of the aperture correction system is 0.23 V/°, hence the full signal range from -5° to 5° surface tilt is 2.3 V. The RMS noise level for both the x and y signal of the Position Sensitive Detector (PSD) has been measured with a Zerodur SUT at best-focus and is about 3.7 mV for the x signal and 4.4 mV for the y signal. This is equivalent to about 0.02°, or 2‰ of the full signal range. Between subsequent measurements however, some zero-crossing drift was observed, which can be suppressed by means of inline calibration on the nulling target before and after each surface measurement. The remaining 3σ uncertainty in the tilt measurement due to drift is estimated to be 0.1°. The following measurements will have the zero-crossing offset removed.

The *x* and *y* signals of the PSD, V_x and V_y, are shown in Figure 8.8, left, for measurement on a Zerodur flat that is tilted about 0.8° and rotated around the θ-axis at 2π rad/s.
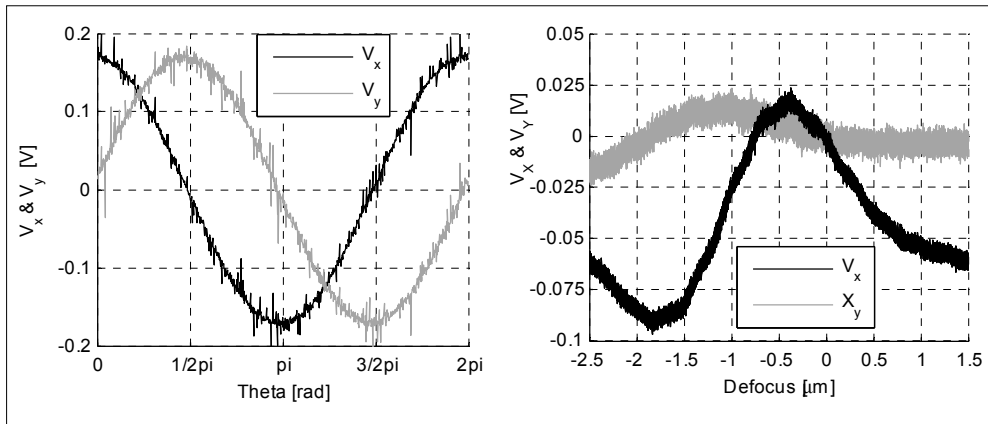
Figure 8.8: PSD x and y signals for a Zerodur SUT that has been tilted 0.8° and that is rotated at 2π rad/s (left) and PSD x and y signals for a perpendicular Zerodur SUT that is translated through focus (right).

As expected, a sine and cosine signal are generated. The spikes seen in the signals are much larger than the measurement noise and they reproduce for repeated passes over the same location on the surface, hence they are thought to be caused by surface artifacts, for which correction indeed is desirable.

The change of the PSD x and y signals as a function of defocus has been measured using a SUT that is scanned through focus by a piezo and elastic guidance mechanism; the results are shown in Figure 8.8, right. It can be seen that especially the x signal varies significantly over the range of the differential confocal system. The mechanism that causes this variation is not known, although it is suspected that the cover window of the PSD plays a role. For the type of PSD used it was not possible to remove the cover glass without significant risk of damaging the PSD, for future designs it is therefore desirable to order a PSD without a cover glass.

Dependence of the PSD x and y signals on the object lens position has also been determined. To this end, the same method has been applied as for the determination of the dependence of the FES zero-crossing on the object lens position. Figure 8.9 shows the x and y signals of the PSD plotted against the object lens position with a 4th order fit through the data. It can be seen that the maximum zero-crossing shift is 2 mV for the x signal and 7 mV for the y signal.
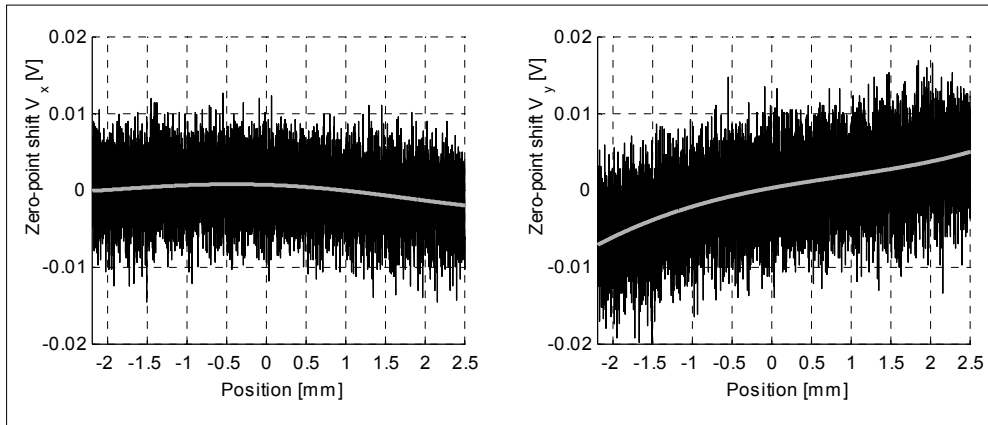
Figure 8.9: Zero-crossing shift of the PSD signals for a perpendicular Zerodur SUT in best-focus as a function of the c-position of the object lens, for the x signal (left) and the y signal (right).

An overview of the different contributions to tilt measurement error is given in Table 8.6. The 2σ uncertainty due to noise has been obtained again by multiplying the RMS by 2.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Uncertainty due to noise (2σ)	0.04°			
Uncertainty due to drift (2σ)	0.10°			
Error due to object lens stroke (max)	-	0.01°	0.03°	
Error due to defocus (max)	0.04°	0.07°	0.16°	0.39°
Tilt measurement uncertainty (2σ)	**0.11°**	**0.13°**	**0.19°**	**0.41°**

Table 8.6: Various contributions to the tilt measurement error.

8.2.3   DUAL-STAGE MEASUREMENT

Operation of the interferometer and the differential confocal measurement system together as a dual-stage distance measurement system has been tested by measuring distance to a stationary SUT. To limit the influence of the full measurement loop of the machine and effects due to machine dynamics, these measurements have been carried out on the nulling mirror. In this case, the displacements measured by the differential confocal system and the interferometer are mainly vibrations of the object lens caused by hum in the amplifier signal. Therefore, the displacements measured by both systems, ideally are identical in magnitude but opposite in sign. This is used to test how well the dual-stage system works by adding the signals and observing the remaining displacement.

In Figure 8.10, left, the differential confocal measurement (1) and the interferometer measurement (2) are shown together with their sum (3).
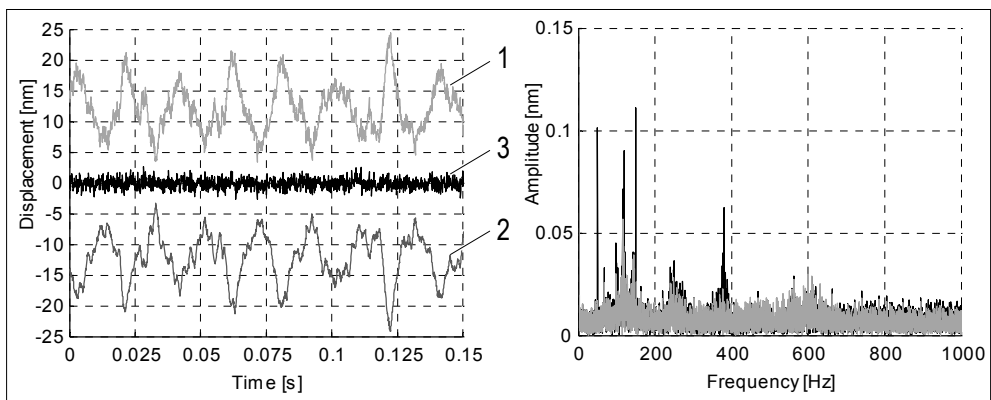
Figure 8.10: Differential confocal measurement (1), the interferometer measurement (2) and their sum (3), for measurement on a stationary SUT (left) and the harmonic content of the sum signal (right) without (black) and with (gray) compensation of data-age.

RMS of the FES and interferometer signals are 3.9 nm; the RMS noise of the signal after addition of the differential confocal and interferometer signal is 0.82 nm.

In Figure 8.11, left and right, the harmonic content of both the interferometer signal (gray) and the sum of the interferometer signal and differential confocal signal (black) is shown. The peaks at 50 Hz (1) and 150 Hz (2) are caused by hum in the amplifier signal (as discussed in the subsection on the control performance, Subsection 7.5.2); the other peaks correspond mainly to eigenfrequencies in the machine. It is concluded that the dual-stage method works satisfactory since these peaks are suppressed to almost the base level of the noise.

If there is a data-age difference between the interferometer data and the differential confocal data, this will reduce the accuracy of the dual-stage method. Therefore, the harmonic content of the sum signal was plotted for various temporal shifts between the signals. In Figure 8.10, right, the result is shown without temporal shift (gray) and with a $1 \cdot 10^{-4}$ s delay of the differential confocal data (black). Tests that are more accurate have shown that the interferometer lags $8 \cdot 10^{-5}$ s behind the differential confocal system. For high scanning speeds at steep angles, this leads to significant measurement errors and consequently it must be corrected.
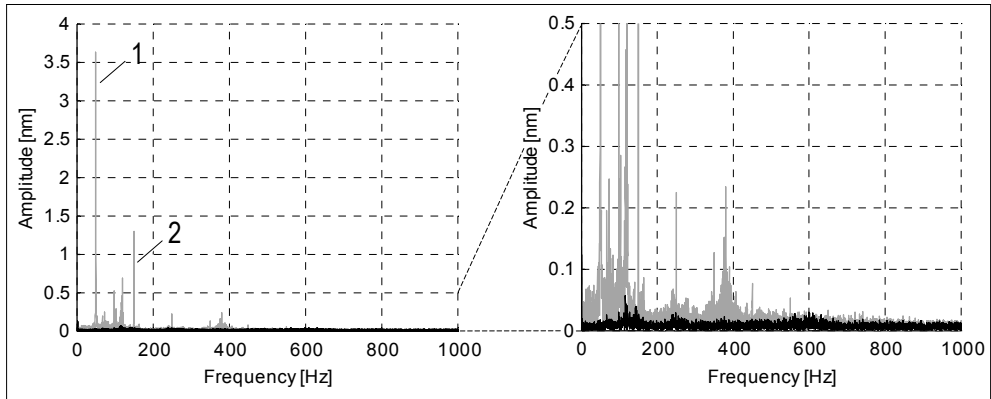
Figure 8.11: The harmonic content of the interferometer signal (gray) and of the sum of the interferometer signal and differential confocal signal (black).

In Table 8.7, the 2σ value that will be used for the uncertainty estimation is stated.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Uncertainty due to dual-stage noise (2σ)	1.6 nm			

Table 8.7: Contribution to the measurement uncertainty of the noise in the summed differential confocal and interferometer signal.

These values also include the uncertainty due to the noise of the differential confocal system at zero-crossing. Because this value increases over the differential confocal measurement range, it will be incorporated as well.

8.2.4   INTERFEROMETER SIGNAL

Errors in the polarization state of the interferometer laser source, non-ideal optical components such as the polarizing beam splitter, quarter wave plates and receiver, and component misalignment lead to periodic errors in the interferometer signals (Cosijns, 2004). These errors mainly occur at the 1st and 2nd harmonic frequencies, of which only the first is likely to be significant in this application. To determine an upper limit for these errors, the sensor's interferometer can be compared with the machine's z-axis interferometer. To do so, the focus of the sensor is locked on a stationary SUT that is placed on the machine's product table, and the z-axis of the machine is moved up and down. By adding the differential confocal signal and the two interferometer signals, a surface height measurement is obtained that should remain unchanged, but instead contains various error sources among which the periodic error. Since both the c and the z interferometer will have some amount of periodic error, it partly cancels out if the phase of the error is not equal. Therefore, to prevent underestimation of the periodic error, the measurement must be conducted multiple times with the SUT at various heights that are not an integer multitude of half the wavelength apart.

The surface height measurement obtained with a translational speed of 10 µm/s is shown in Figure 8.12, left, over 3 µm range.
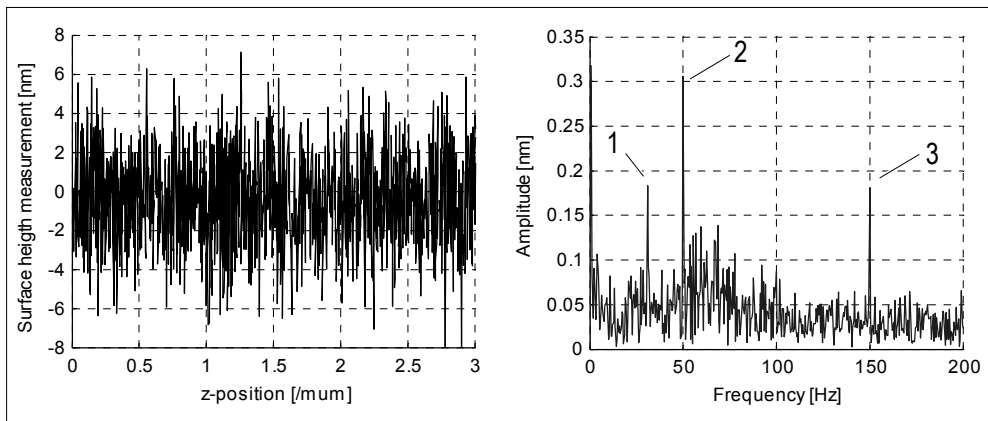
Figure 8.12: Surface height measurement taken with translating z-stage (left) and harmonic content of that signal (right) to identify periodic errors in the interferometer signal.

The periodic error cannot be observed by visual inspection due to noise sources in the measurement chain, therefore a plot of the harmonic content of the signal is shown in Figure 8.12, right. Since the interferometer uses light with a wavelength of 633 nm and the translational speed is 10 µm/s, the first harmonic corresponds to 31.6 Hz. The peak at this frequency (1) represents a periodic error with an amplitude of 0.18 nm, hence the contribution to the measurement uncertainty is considered negligible. The peaks at 50 Hz (2) and 150 Hz (3) are visible here as well. The influence of these vibrations is not suppressed in this experiment because not all the channels necessary for correction of data-age were recorded.

8.3 CORRECTION AND CALIBRATION

As expected, the measurements presented in the previous section show that the required measurement uncertainty cannot be met without corrections and calibrations. With respect to this, measurement machines have a significant advantage over manufacturing machines since corrections and calibrations may be applied to a measurement afterwards, whereas they have to be performed real time during production of a component. This does not only allow for more complex calculations with equal processing capacity, it also increases possibilities concerning averaging and permits information from the end of the procedure to be taken into account in the correction of data that is generated at the beginning of the procedure.

8.3.1 DRIFT AND NOISE

Except for the timescale of the phenomena, physically there is no difference between drift and some of the pseudorandom components regarded as noise; the frequency at which the division is made is somewhat arbitrary. Here, the distinction is made based on practical considerations resulting in discrimination between long-term drift, short-term drift and noise.

Drift over periods that are longer than typical measurement times (about 15 minutes for Ø500 mm optics) are considered long-term drift, and can be compensated to a large degree by inline calibration before and after surface measurement. Drift over timescales between 15 minutes and 2 seconds is considered short-term drift; faster fluctuations are considered noise.

Tracks will be measured multiple times, which enables partial reconstruction of short-term drift, so that it can be corrected for. Ideally, measurements of the same part of the SUT give the same height measurement every time. In reality, subsequent measurements of the same part of the surface are however expected to vary due to a combination of drift, noise and positioning error. Hence, single

samples cannot be used to determine drift between separate measurements of the same track. Instead, multiple samples can be averaged to suppress the influence of noise. Since the noise in surface height measurement has an RMS of about 0.9 nm for the whole metrology loop (Henselmans, 2009), averaging samples taken over a part of a track will sufficiently reduce the influence of noise. Averaging over only a part of the track will however lead to significant errors due to lateral position variation of the measurement spot. For most surface forms, averaging over a whole track suppresses this error as well.

By applying a symmetric moving average filter, the drift can thus be estimated at the sample rate most of the time. An exception to this are measurements in the first and last half revolutions of a particular track and during the time it takes to move from one circular track to the next, since the relation between samples of drift taken on the previous and new track is lost. This missing information can be estimated to some extent by extrapolation of the drift in the old and new track, but this will introduce additional uncertainty. Because the spindle will rotate at 1 rev/s, the sample spacing when moving from track to track will therefore be a little larger than 1 s, approximately corresponding to a sample rate of 1 Hz and a Nyquist frequency of 0.5 Hz.

Currently, such drift correction is not yet applied in the data processing of the machine. Instead, another method (Henselmans, 2009) is used that requires less complicated data processing. Before and after each full surface measurement, various radial scans are made, which contain negligible drift because they take little time to perform. Height measurements obtained during these radial scans are compared to corresponding measurements of averaged circular tracks and used to correct for drift during the full surface measurement. The disadvantage of this method is that it leads to surface artifacts in the form of height differences of several nanometers between subsequent tracks. This mainly leads to a less accurate result in the mid-spatial range, which is quite well visible in plots of surface measurements. It nevertheless reduces the RMS error of full surface measurements significantly.

Possibly the two methods can be combined to benefit from the advantages of both methods. The scan method can for example be used to determine the drift between the first and the last track, while the averaging method is used to reconstruct the development of the drift in-between. Alternatively, extra averages of the inner most and outer most tracks can be taken before and after each measurement and combined with the averaging method to suppress drift.

For an estimate of the contribution of drift to measurement uncertainty, the use of the averaging method is assumed. The standard deviation of the averaged value relates to the single sample standard deviation as:

$$\sigma_{ave} = \frac{\sigma_{hm}}{\sqrt{K}}$$
(8.1)

where:

σ_{ave} is the standard deviation of the average,

σ_{hm} is the standard deviation of height measurement, and

K is the number of samples.

Assuming that the deviation of the averages is uncorrelated, the uncertainty of the drift over a total measurement can be estimated by:

$$u_{drift} = \sqrt{L \sigma_{ave}^2} = \sqrt{L \frac{\sigma_{hm}^2}{K}}$$
(8.2)

where:

u_{drift} is the uncertainty due to drift after correction, and

L is the number of track measurements.

With the 0.9 nm RMS noise in the height measurement and a measurement consisting of 900 track measurements sampled at 1 kHz, this gives a 1σ uncertainty of 0.85 nm in the drift reconstruction over the full surface measurement. In Table 8.8, the 2σ uncertainty due to drift after correction is stated.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Uncertainty due to drift (corrected) (2σ)	1.7 nm			

Table 8.8: Contribution to the measurement uncertainty of drift after compensation.

8.3.2 DIFFERENTIAL CONFOCAL SYSTEM CALIBRATION

To suppress long-term drift and to account for differences in the FES for different SUT materials, the sensor can be calibrated before and after each measurement. Because of the dual-stage design with the interferometer as a secondary measurement system, this is relatively straightforward.

Drift in the zero-crossing of the FES is corrected automatically since it is the reference during interferometer nulling, as was discussed in Subsection 7.5.3. Next, the sensor can be aimed at the SUT and the object lens can be translated to calibrate the FES using the interferometer signal. The measurement error due to the cyclic disturbance cannot be corrected this way.

The zero-crossing shift due to the stroke of the object lens is corrected by calibration, while the change of the FES curve for different positions along the stroke of the object lens will not be corrected, as was discussed in Subsection 8.2.1. Aperture correction is also part of the FES calibration; nevertheless, it is treated separately because it is a quite specific subject.

8.3.3 APERTURE CORRECTION

The changes in the shape of the FES for different tilts can be measured relatively easily by the method described in Subsection 8.2.1. This method does however not allow the zero-crossing shift of the sensor to be measured, since the position of the SUT changes when its tilt is changed.

The only method to evaluate tilt dependence of distance sensors that was found in literature is by measurement of reference artifacts, for example a precision sphere. As high accuracy is needed, the sphere must be calibrated and the roundness error has to be taken into account. Besides a high accuracy artifact, this method requires a precision motion and measurement system to position the sphere and sensor relative to each other and measure their relative position. The r-axis of the machine can be used to translate the sensor in one direction, and the sphere can be placed eccentric on the θ-spindle to enable scanning in the y-direction. In this manner, the sensor can be calibrated for angles in all directions. Because the reference mirrors of the machine have not been calibrated yet, applying a small sphere is beneficial since this limits the range over which the z-mirror is used. For a Ø20 mm sphere, the sag at 5° tilt is 38 µm, so the sensor must track the surface to stay in focus. A disadvantage of this method is that the influence of surface curvature on the distance measurement is not known and cannot be separated from the influence of tilt.

An alternative method for calibration of tilt dependency in optical distance sensors that does not rely on a reference artifact has been developed here. The essence of the method is that it allows simultaneous measurement of distance to the SUT as well as the lateral position of the measurement spot on the surface by using a PSD as SUT. The layout is schematically depicted in Figure 8.13, left.

A PSD (1), of which the cover plate is removed, is placed at an angle ξ_{psd} on the θ-spindle (2). The ψ-axis (3) is rotated ψ_s, so that the measurement direction is perpendicular to the PSD (top picture). When the spindle is rotated, the angle of the PSD relative to the sensor changes and the centre of the returning beam describes a circular trajectory. When the spindle is rotated 180°, the surface tilt relative to the measurement direction is twice ξ_{psd} (bottom picture). By repeating this procedure for various tilts of the PSD, the centre of the reflected laser beam describes a series

of circles of different diameter that are tangent at the center of the aperture, which corresponds to zero surface tilt.
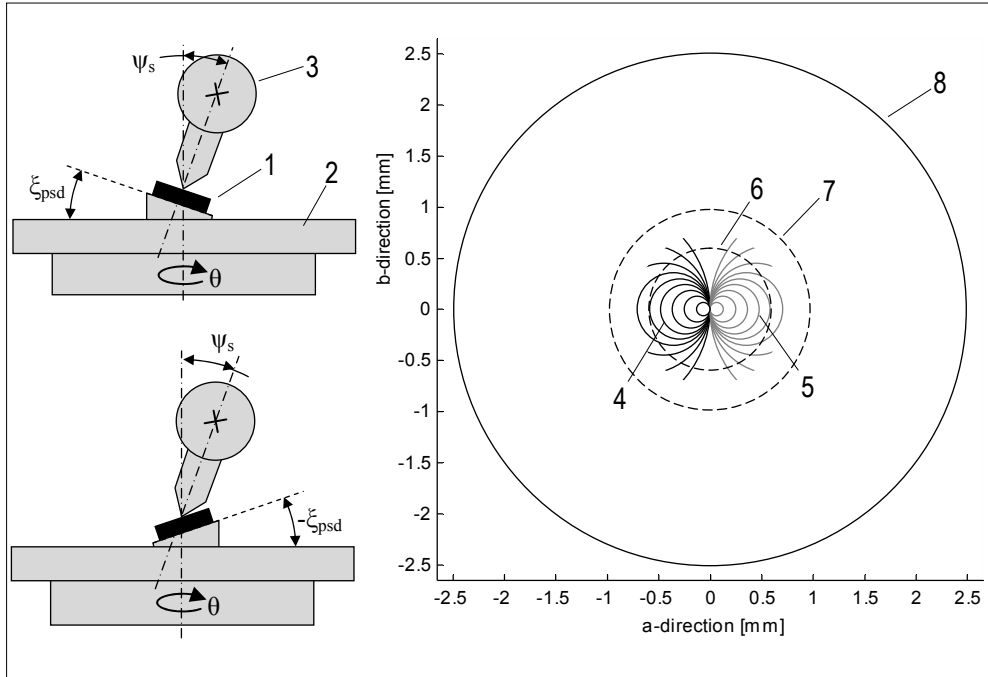
Figure 8.13: Schematic representation of the layout used to calibrate the tilt dependency (left) and the theoretical trajectories of the centre of the beam at the object lens aperture (right).

The pattern as it is at the object lens and the aperture correction PSD, is shown in Figure 8.13, right. For positive ψ_s, the circles extend in negative *a*-direction (4, in black), for negative ψ_s, the circles extend in positive *a*-direction (5, in gray). Because all measurements have a common point at zero surface tilt, they can be related to each other even though the position of the PSD changes when its angle is changed.

The area that is used for tilts up to 5° is indicated with a dashed circle (6), as well as the e^{-2}-diameter of the beam (7) and the aperture of the object lens (8). To allow interpolation between different data points, the spacing between circles is chosen about a factor 20 smaller than the diameter of the laser beam.

If the measurement spot is located exactly on the θ-axis, the height of the SUT only varies due to error motion of the air-bearing spindle, which has been measured to be below 4.5 nm (Henselmans, 2009). Therefore, the remaining change in height

measurement is tilt dependent error of the sensor. If the measurement spot is not located on the θ-axis however, it will result in an additional change of height at the measurement spot. Such misalignment can be detected with the PSD on the spindle, since in that case the spot will describe a circular trajectory on the PSD; this is used to centre the spot on the θ-axis.

The lateral position of the light is measured in the silicon layer while the light that is collected by the differential confocal system is reflected at the surface. This leads to an offset between the real position of the measurement spot and the position measured by the PSD. When the SUT is rotated, this positioning error causes a height error, which can be expressed as:

$$e_{pt} = d_{pt} \frac{\sin\left(\dfrac{\alpha_{tmax}}{2}\right)}{n_{si}} \tan\left(\frac{\alpha_{tmax}}{2}\right) \approx \frac{d_{pt}}{n_{si}} \left(\sin\left(\frac{\alpha_{tmax}}{2}\right)\right)^2 \tag{8.3}$$

where:

e_{pt} is the measurement error due to penetration of the light in the PSD,

d_{pt} is the penetration depth,

α_{tmax} is the maximum tilt to be corrected, and

n_{Si} is the index of refraction.

The penetration depth and refractive index of silicon depend on the kind of dopants and the concentration, hence values stated in literature vary. Two values found for 640 nm light are 3.6 µm (Cuser, 2004) and 2 µm (Kohn, 2009). For the calculation, a penetration depth of 5 µm and an index of refraction of 3.9 are used here. This gives a distance error of 2.4 nm, which is considered acceptable. For applications where the required acceptance angle is larger, the error might become significant since it is approximately proportional to the square of the tilt.

A sine plate with an integrated 2 mm PSD and spacers corresponding to the desired tilts were made; the setup during calibration is shown in Figure 8.14. The structural components are made of Invar so that measurements can be started immediately after handling the sine plate and spacers to change the tilt angle. By working in a dark laboratory and digital filtering to suppress noise and hum (optimized for 50 Hz and 100 Hz), a lateral resolution of about 20 nm was achieved. Before each measurement, the spot was aligned to within 50 nm, which causes a maximum error of 2.2 nm in surface height.
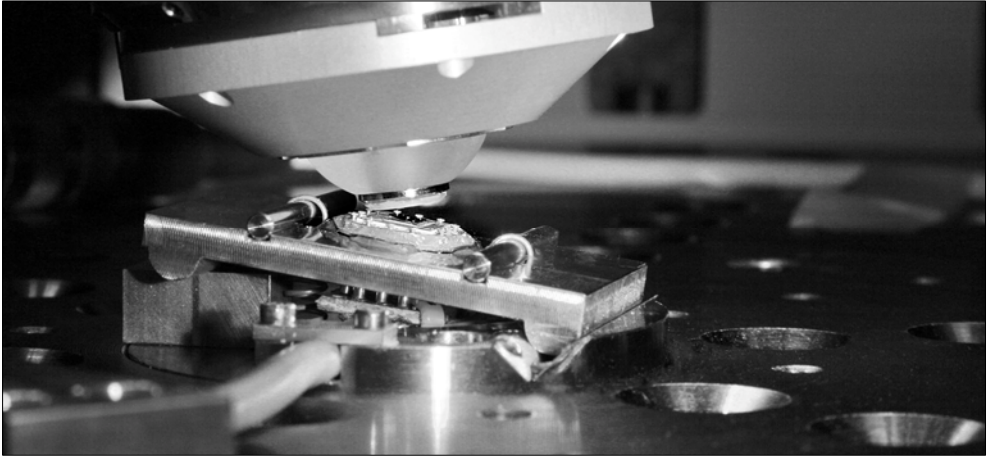
Figure 8.14: Photograph of the sine plate with an integrated 2 mm PSD during calibration of tilt dependent error.

Using this method, each trajectory has been subsequently measured 3 times at 1 kHz, resulting in a data set of about 1 million points. The down-sampled data is plotted in Figure 8.15, left. The zero-crossing shift varies between -312 nm and 499 nm for maximum to minimum tilt in β-direction, which is more than expected. There is about 100 nm variation due to tilt in α-direction. It is suspected that the degraded alignment after adhesive failure is partly to blame for the large tilt dependent error in β-direction. Figure 8.15, right, depicts a least squares 10th order polynomial plane that has been fitted through down-sampled data with an RMS fitting residual of 13 nm. The fitting residual is shown in Figure 8.16.
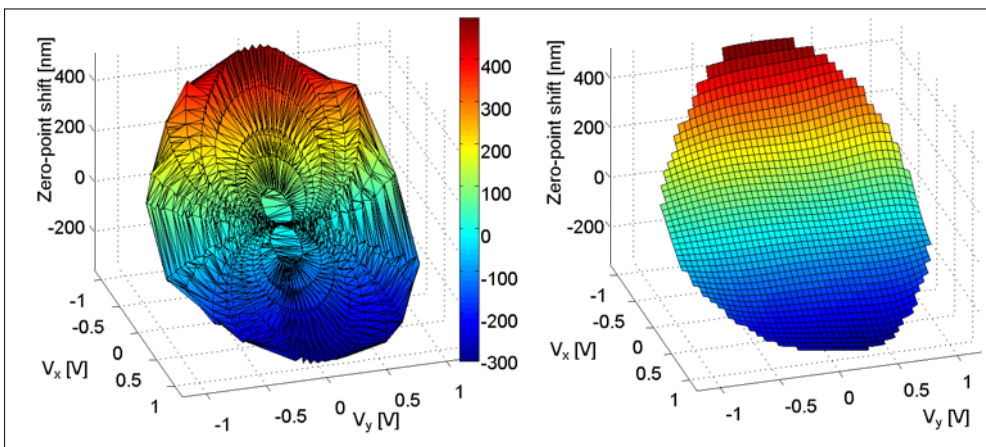
Figure 8.15: Measured zero-crossing shift for tilt angles up to 5°, plotted against the aperture correction PSD signals (left) and a fitted 10th order polynomial plane (right).
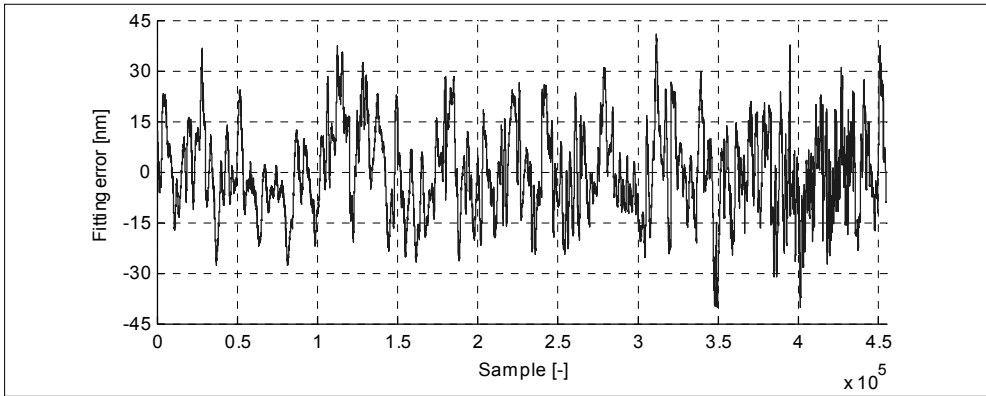
Figure 8.16: The fitting residual of the 10th order polynomial plane.

To avoid calculating the 66 terms of the two-variable polynomial for every measurement point, interpolation of an equidistant lookup table generated with the polynomial can be used for correction.

During the calibration procedures, it is relatively easy to measure the FES errors due to tilt too, by translating the object lens through focus at various orientations. To do this within acceptable time with sufficiently small measurement increments, the process must be automated. Because at the time of the measurements, there was no budget left for software development, this has not been carried out.

In Table 8.9, the error and uncertainty contributions to uncertainty in the calibration data for tilt dependent error are stated. The 2σ uncertainty due to measurement variability is calculated from the RMS fitting error and (8.1), using $K=3$. The total uncertainty in the aperture correction data is estimated by quadratic addition and stated as well.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Error due to spindle error motion (max)	-	4.5 nm		
Error due to penetration of PSD (max)	-	0.4 nm	2.4 nm	
Error due to centering tolerance (max)	-	0.9 nm	2.2 nm	
Uncertainty due to fitting error data (2σ)	-	15 nm		
Uncertainty in aperture correction data (2σ)	**-**	**15.7 nm**	**16 nm**	**16 nm**

Table 8.9: Error and uncertainty contributions to uncertainty in calibration data for aperture correction.

8.3.4   INTERFEROMETER CALIBRATION AND CORRECTION

As discussed in Subsection 5.1.5, the change of refractive index of air with temperature, pressure and partial pressure of water vapor causes significant errors. Therefore, an environmental sensor is placed in the machine to correct for such changes. The most of the remaining error after correction has typical timescales longer than 2 s and can be further suppressed by reconstruction of drift (Subsection 8.3.1). This also counts for thermal drift of the interferometer components.

Nevertheless, there is a remaining 2σ measurement uncertainty of 1 nm over the 5 mm stroke of the objective lens due to the uncertainty in scaling caused by refractive index changes. This error cannot be corrected by reconstruction of drift. In addition, because the environmental sensor is not placed in the focusing unit, local effects will not be corrected for using refractive index correction. If these local effects result in short-term drift such as heating of air due to dissipation in the focusing unit's actuator, these effects again can be suppressed by reconstruction of drift. This does not apply to quickly changing effects such as pressure fluctuations inside the focusing unit due to pumping, for which no method of correction is proposed. It is important to quantify the magnitude of this error for different measurement tasks so that it can be included in the uncertainty estimation. This can be done by comparison to a reference interferometer measuring the position of a reflector mounted to the object lens assembly.

Interferometer errors such as the cosine error, Abbe error, lateral displacements combined with mirror shape and periodic errors are all small due to the alignment and quality of the optical components. Furthermore, for this application these errors typically reproduce well, so they can be calibrated simultaneously by comparison to a reference interferometer that is traceable to the SI unit of length. It is estimated that they can be corrected to a 2σ uncertainty of about 2 nm. This calibration can be performed with the same setup as used to investigate the magnitude of the errors due to pumping.

Table 8.10 gives an overview of the sources of uncertainty in the interferometer signal.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Uncertainty due to scaling error (2σ)	-	0.4 nm	1 nm	1 nm
Uncertainty in calibration (2σ)	2 nm			
Uncertainty due to fast pressure change (2σ)	-	3 nm	10 nm	10 nm
Uncertainty in interferometer signal (2σ)	**2 nm**	**3.6 nm**	**10.2 nm**	**10.2 nm**

Table 8.10: Contributions to uncertainty in the interferometer signal.

8.4 ESTIMATION OF MEASUREMENT UNCERTAINTY

The experiments covered are thought to give a representative indication of the magnitude of the errors and uncertainties that are introduced into the measurements. With the previously presented data, the measurement uncertainty of the sensor will be estimated for the four previously defined measurement tasks. First, the uncertainty in the tilt measurement will be used to derive to what level of uncertainty the tilt dependent error can be reduced. When this is known, it is used to calculate the uncertainty in the differential confocal measurement. Finally, the uncertainty in the differential confocal signal, the interferometer signal, the dual-stage measurement and the drift reconstruction will be added to arrive at the uncertainty of the sensor as a whole.

Measurement uncertainty due to tilt

The sensor's measurement uncertainty due to tilt has been split into two components: shift of the zero-crossing and the change of the FES curve. To compensate for the zero-crossing shift, the tilt measurement of the aperture correction system is used to obtain a correction value from the calibration data. Hence, uncertainty in tilt measurement introduces uncertainty in the correction values that depends on the local slope of the calibration data. Therefore, the maximum slope within the tilt range of each measurement task has been looked up in the fitted correction plane.

By multiplying the uncertainty of the tilt measurement with the local slope of the fitted correction plane, the uncertainty in the correction data due to tilt measurement uncertainty is found. The tilt uncertainty, maximum local slope in the calibration data and the resulting uncertainty in the correction values are stated in Table 8.11.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Slope of zero-crossing calibration data (max)	-	95 nm/°	143 nm/°	
Tilt measurement uncertainty (2σ)	-	0.13°	0.19°	0.41°
Zero-crossing shift uncertainty due to tilt (2σ)	**-**	**12.3 nm**	**27 nm**	**58.6 nm**

Table 8.11: Uncertainty in correction values for zero-crossing shift due to local surface tilt.

For the correction of the change of the FES curve due to surface tilt, the same approach is used. Because, the change of the FES curves has however not been fully calibrated yet, the maximum slope of the correction data is not known. To arrive at an estimate, it is assumed that the change of the FES is evenly distributed

over the tilt, so that the maximum local slope in the data can be found by dividing the value for 4° surface tilt by four degrees. The estimated maximum local slope in the calibration data, the tilt uncertainty, and the resulting uncertainty in the correction values of the FES is stated in Table 8.12.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Slope of FES calibration data (max)	-	37 nm/ 4°	92 nm/ 4°	500 nm/ 4°
Tilt measurement uncertainty (2σ)	-	0.13°	0.19°	0.41°
FES uncertainty due to surface tilt (2σ)	**-**	**1.2 nm**	**4.4 nm**	**51 nm**

Table 8.12: Uncertainty in correction values for change in the FES shape due to surface tilt.

In Table 8.13, the contributions to the total uncertainty due to tilt are listed together with the total that is obtained by quadric addition.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Zero-crossing shift uncertainty due to tilt (2σ)	-	12.3 nm	27 nm	58.6 nm
FES uncertainty due to surface tilt (2σ)	-	1.2 nm	4.4 nm	51 nm
Uncertainty in aperture correction data (2σ)	-	15.7 nm	16 nm	16 nm
Uncertainty due to tilt, after correction (2σ)	**-**	**20 nm**	**32 nm**	**79 nm**

Table 8.13: Total uncertainty arising in the sensor's distance measurements due to surface tilt.

Total uncertainty for the differential confocal system

The total measurement uncertainty in the differential confocal measurements is obtained by quadratic addition of the separate components that were stated in Section 8.2 and the uncertainty due to surface tilt. Besides these contributions the maximum observed error due to turbulence in the demonstrator is included. This data is taken from the demonstrator because such an experiment has not been conducted with the final prototype. These separate contributions and the resulting total are listed in Table 8.14.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Uncertainty due to noise (2σ)	0.3 nm	0.4 nm	0.5 nm	1.5nm
Error due to cyclic error (max)	1.1 nm	1.1 nm	1.1 nm	1.1 nm
Zero-crossing uncertainty due to lens stroke (2σ)	-	0.5 nm	1 nm	1 nm
FES error due to lens stroke (max)	-	1.5 nm	4 nm	20 nm
Uncertainty due air turbulence (max)	3.5 nm	3.5 nm	3.5 nm	3.5 nm
Uncertainty due to tilt, after correction (2σ)	-	20 nm	32 nm	79 nm
Uncertainty differential confocal system (2σ)	**3.7 nm**	**20 nm**	**32 nm**	**82 nm**

Table 8.14: Total measurement uncertainty of the differential confocal system.

The measurement uncertainty of the differential confocal system for the measurement of rotationally symmetric surfaces seems low. It should be noted however, that this estimation has as starting point that long-term drift is compensated by inline nulling and calibration and that short-term drift is suppressed through multiple measurements of the same tracks. In some other applications, these methods to improve performance might not be feasible.

Total uncertainty estimate for the sensor

By quadratic addition of the uncertainty in the differential confocal signal, the interferometer signal, the dual-stage measurement and the drift reconstruction, the total uncertainty in measurements obtained with the sensor is estimated. The relevant values as well as the final estimate are shown in Table 8.15.

Measurement task Error / uncertainty source	RS	MFF	HFF	SFR
Uncertainty differential confocal system (2σ)	3.7 nm	20.1 nm	32 nm	82 nm
Uncertainty in interferometer signal (2σ)	2 nm	3.6 nm	10.2 nm	
Uncertainty due to dual-stage noise (2σ)	1.6 nm			
Uncertainty due to drift (corrected) (2σ)	1.7 nm			
Uncertainty of complete sensor (2σ)	**4.2 nm**	**21 nm**	**34 nm**	**83 nm**

Table 8.15: Estimate of the total measurement uncertainty of the sensor.

The task specific 2σ measurement uncertainty of the sensor for measurement of a Rotationally Symmetric (RS), Medium FreeForm (MFF), and Heavily FreeForm (HFF) surface are estimated to be 3.3 nm, 21 nm and 34 nm, respectively. The

illustrative case when using the signals of the sensor over their Symmetric Full Range (SFR), at 83 nm, shows a considerably larger measurement uncertainty. The reason is that an increase of defocus leads to a significant increase of measurement error of both the differential confocal as well as the aperture correction system. Hence, keeping the tracking error small is key in achieving a low measurement uncertainty. This emphasizes the importance of optimizing the FES controller.

The influence of surface curvature has not been taken into account. The effect of pressure changes due to fast axial movement of the focusing unit has been included but is not based on observations. Furthermore, the effect of combinations of factors such as simultaneous defocus, surface tilt and object lens translation, might lead to larger measurement errors than estimated from their separate contributions. This can all lead to an underestimation of uncertainty. Conversely, when adding systematic errors to estimate uncertainty, the worst-case error that has been observed has been used every time; this leads to an overestimation of the measurement uncertainty.

8.5 SURFACE MEASUREMENT[9]

Testing the combined performance of the machine and sensor can be done by measurement of an optical flat that is placed on the product table under a small angle. Such a tilted flat measurement is particularly suitable for validation since it has similar characteristics as freeform measurement: it requires a relatively large vertical measurement range and has a changing local surface tilt at the measurement spot. At the same time its surface shape, unlike that of freeform surfaces, can be readily measured with low measurement uncertainty by conventional techniques. This makes it a suitable artifact with which the effectiveness of aperture correction can be tested. A photograph of the measurement of a tilted flat is shown in Figure 8.17; in the insert, a picture of the NANOMEFOS machine is shown.

---

[9] *The experiment of which the results are discussed in this section, as well as the data processing have been performed by R. Henselmans. More about the NANOMEFOS machine and other validation measurements regarding the machine can be found in (Henselmans, 2009).*

Figure 8.17: Measurement of a tilted flat (main) and the NANOMEFOS machine at the GTD workshop (insert).

A Zerodur flat that is 0.5° tilted has been measured at 1 rev/s, a sample rate of 1 kHz and a track spacing of 1 mm. To suppress drift, the method using radial scans before and after the full surface measurement has been applied. The results have been evaluated with and without aperture correction to investigate its effectiveness. In Figure 8.18, the graphs of the results are shown, without aperture correction on the left, and with aperture correction on the right. The scales and the color-coding in these graphs are equal to facilitate comparison.
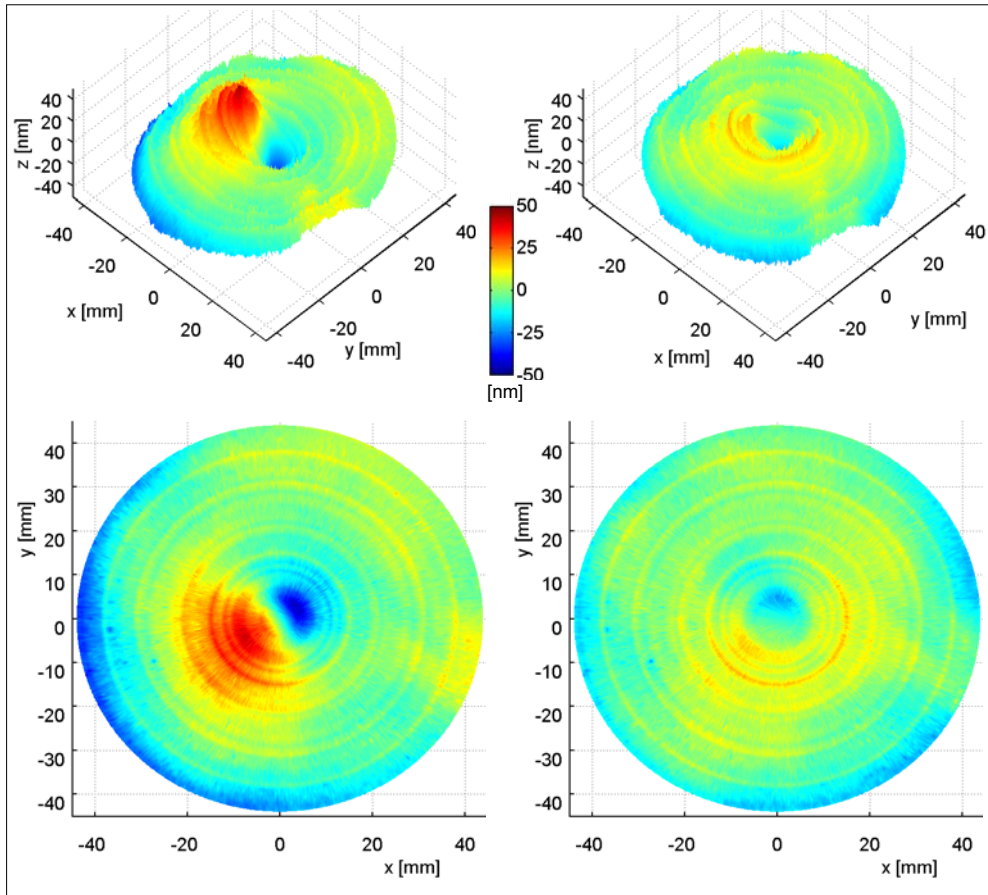
Figure 8.18: Surface height map after removal of tilt and piston of an optical flat tilted at 0.5° as measured with the NANOMEFOS machine without aperture correction (left) and with aperture correction (right).

The surface measurement without aperture correction shows significantly more deviation from flat than the measurement with aperture correction. The uncorrected measurement has a distinctive form with a PV of 97 nm and an RMS of 14.4 nm. For the corrected measurement, the deviation from flat is greatly reduced; it has a PV and RMS of 52 nm and 7.8 nm, respectively.

The result of the measurement with aperture correction is shown in Figure 8.19 alongside a measurement performed at NMi VSL using a Fizeau interferometer. Because the measurement data is processed with different software, the color-coding in these graphs is not the same.
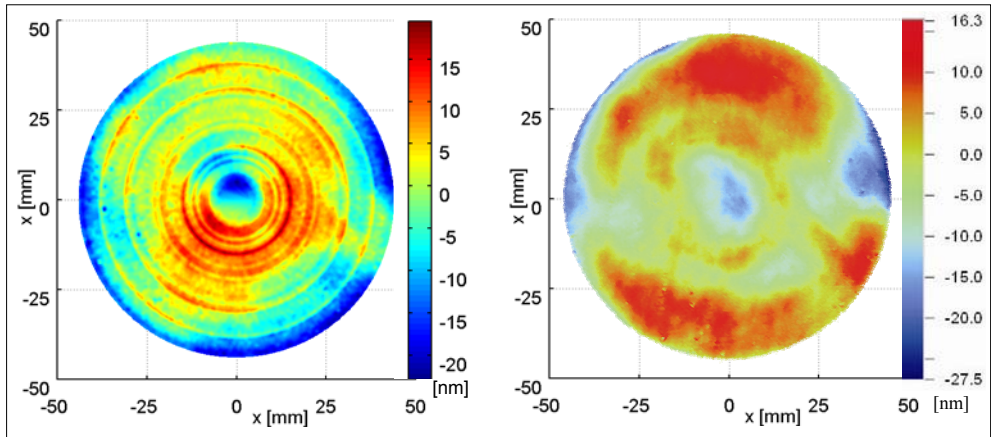
Figure 8.19: Surface height map after removal of tilt and piston of an optical flat tilted at 0.5°, as measured with the NANOMEFOS machine (left) and as measured at NMi VSL using a Fizeau interferometer (right).

The PV of the interferometric measurement is 44 nm ± 20 nm (2σ) and the RMS is 7.1 nm. Although the machine itself has not been calibrated and the sensor has been only partially calibrated, good agreement is already observed with the traceable interferometric measurement. Nevertheless, in the measurement obtained with the NANOMEFOS machine, some measurement artifacts can be identified. The concentric rings are caused by the currently applied method of drift compensation; this might be partially prevented by making changes to the drift correction method. In addition, the measurement seems to include some positive curvature, which can be caused by the horizontal reference mirror of the machine since it has not been calibrated yet. The dip in the middle of the surface is introduced in the data processing; this problem has now been resolved.

Despite these measurement artifacts, the uncertainty is already better than the 2σ measurement uncertainty of 30 nm that is required for measurement of such surfaces. Furthermore, the concept of reducing tilt dependent error through application of aperture correction has been demonstrated to work well.

9   CONCLUSIONS AND RECOMMENDATIONS

In this final chapter, conclusions about the work described in this thesis are presented, and possible relevance to domains outside of this project is briefly discussed. In addition, some areas are suggested which are thought to deserve further attention.

9.1   CONCLUSIONS

A tilt robust, single point distance sensor with millimeter range and nanometer uncertainty has been developed, realized and tested. The need for such a sensor arose when it was found that no commercial or experimental sensor combines the characteristics required for the NANOMEFOS project. In this project a universal measurement machine for freeform optical surfaces has been developed, which until now has been the missing link in the production chain for single-piece high-end freeform optics. Some of the requirements for the sensor are:

- 5 mm measurement range,

- a 2σ measurement uncertainty of 10 nm for surfaces perpendicular to the measurement direction and 35 nm for surfaces with tilts up to 5°,

- 1 nm resolution,

- capability to measure surfaces with reflectivities ranging from 3.5% to 99%,

- non-contact measurement,

- absolute measurement, and

- a minimum sample rate of 3 kHz, preferably higher.

To achieve these requirements, a conceptual design has been made consisting of a dual-stage measurement system: a primary measurement system tracks the surface

under test by translating its object lens, while the secondary measurement system measures this translation. After an assessment of various measurement principles through comparison of characteristics inherent to their principle of operation, differential confocal measurement has been selected as the primary measurement system. Dual-pass heterodyne interferometry is used as secondary measurement method. To allow for correction of tilt dependent error through calibration, a third measurement system has been added. It consists of a beam splitter and position sensitive detector, which measures through which part of the aperture the light returns. A patent application has been filed that covers this method of tilt dependent error correction.

Analytical models of the differential confocal measurement principle have been derived to enable predictions regarding performance and optimization. The models show satisfactory agreement with the experimental results generated using the demonstrator, thus building confidence that the models can be applied as design and optimization tools.

Various properties that characterize the performance of a differential confocal measurement system have been identified. Their dependence on the design parameters has been studied on the basis of the models. The results of this study are applied to optimize the sensor for use in NANOMEFOS.

With the measurement methods chosen and the optimal parameters known, the optical system has been designed. The interferometer and the differential confocal systems are integrated in a compact design, which is also covered in the aforementioned patent application. To obtain a system that fits the allotted volume envelope, many components are custom made and the optical path of the differential confocal system is folded using prisms and mirrors.

The optomechanical and mechatronic system has been designed around the optical system. A custom focusing unit has been designed that comprises a guidance mechanism and actuator to enable tracking of the surface. To achieve a low measurement uncertainty, it aims at accurate motion, high bandwidth and low dissipation. The lateral position of the guidance reproduces within 20 nm and from the frequency response, it is expected that a control bandwidth of 800 Hz can be realized. Power dissipation depends on the form of the freeform surface and is a few mW for most expected trajectories.

Partly custom electronics are used for signal processing and to drive the laser and the focusing unit. Control strategies for interferometer nulling, focus locking and surface tracking have been developed, implemented and tested.

After the system was realized, tests have been performed to evaluate the system's performance. Various calibrations and their expected effect on measurement uncertainty are discussed. One of these calibrations is based on a new method to measure the tilt dependency of distance sensors. Due to misalignment resulting from adhesive failures and suspected aberrations in some of the components, the uncorrected tilt dependency is not as low as expected. Aperture correction however, enables realizing the required measurement uncertainty nevertheless.

To test the performance of the NANOMEFOS machine with the sensor installed, measurements are performed on a tilted flat, which serves as a reference freeform with known surface form. These tilted flat measurements demonstrate the reduction of measurement error obtained by aperture correction.

The approach followed during this research has resulted in the successful realization of a novel sensor for NANOMEFOS; it is now incorporated in the machine and installed at TNO. So far, it has been used for internal research and third party measurements; a project in which it will be used as a measurement tool for production of freeform optics has started. As such, it forms a crucial link in the production chain for single-piece high-end freeform optics and it is hoped that it can contribute to realization of lighter, compacter, high-end optical instruments with superior performance to conventional high-end optical systems.

The sensor realized has the following specifications:

- 5 mm measurement range of the secondary system,

- -2.5 µm to 1.5 µm tracking range,

- sub-nanometer resolution,

- an (estimated) 2σ measurement uncertainty of:

    - 4.2 nm for measurement of rotationally symmetric surfaces (no surface tilt, a measurement range <0.1 mm and a maximum tracking error of 100 nm),

    - 21 nm for measurement of medium freeform surfaces (tilt up to 2°, a measurement range of 2 mm and a maximum tracking error of 200 nm),

    - 34 nm for measurement of heavily freeform surfaces (tilt up to 5°, a measurement range of 5 mm and a maximum tracking error of 0.5 µm),

- the capability to measure surfaces with reflectivities ranging from 3.5% to 99%, and

- a small-signal bandwidth of 150 kHz.

Besides providing the sensor required for NANOMEFOS, additional contributions of the work presented in this thesis are listed below.

229

- The analytical models developed and the optimization approach applied constitute a design method that is applicable to other differential confocal systems as well. The expression of various design parameters and properties in dimensionless form has the advantage that the pinhole diameter and pinhole offset can be chosen separately from the choice of numerical aperture and laser wavelength.

- Combinations of design parameters that result in near-zero sensitivity of the differential confocal measurement principle to surface tilt is discovered and can be used in other differential confocal systems as well.

- Aperture correction by means of a beam splitter and a position sensitive detector as proposed here, in combination with calibration of tilt dependent error can be applied in various optical distance sensor types. In the current market for distance sensors, the focus is more on resolution than on measurement uncertainty and tilt dependence. It is expected however, that due to the continuing trend towards higher accuracy products, the importance of measurement uncertainty and tilt dependency will grow. In this respect, addition of aperture correction would constitute an improvement to other optical sensor types as well. Because of this potential, various manufacturers of distance sensors have been approached.

- A novel calibration method for tilt dependent error that is universally applicable to optical distance sensors has been developed.

- A layout has been designed that, due to wedges bonded to the beam splitter, eliminates ghost reflections parallel to the optical axis, while allowing the use of standard quarter wave plates. This layout can decrease measurement uncertainty in various other measurement systems as well.

The characteristics of the sensor make it suitable for various applications; one could think of surface profiling and 2.5D measurement platforms. In scanning applications, a clear advantage of this sensor compared to other single point optical sensor types is its ability to combine low measurement uncertainty with a high measurement bandwidth, which is up to 150 kHz for this design. Due to the developed design methodology, optimization of properties such as resolution, measurement uncertainty, acceptance angle and measurement range for other applications is relatively straightforward. Furthermore, if the design volume is larger than in this application, folding of the optical path is not necessary and standard optical components can be used. This leads to a significant reduction in costs and complexity, while the decrease in the number of reflective surfaces improves the wavefront quality.

9.2 RECOMMENDATIONS

Although the sensor has shown satisfactory performance in test measurements of a tilted flat, some changes to the current system can further improve performance while being relatively straightforward to apply.

Measurement errors increase significantly for increasing tracking errors. Nevertheless, the present controller for surface tracking is not yet optimized for the freeform surfaces anticipated, which mainly require high disturbance suppression in a frequency band of about 1 Hz to 10 Hz. It is expected that the bandwidth of the controller can be increased from the 250 Hz it is now, to about 800 Hz. This allows application of a higher loop gain, thus improving low frequency disturbance suppression. Additionally, a double integrator can be applied in the relevant frequency band. To measure heavily freeform surfaces, repetitive control can be considered as well.

The two measurement branches of the differential confocal system do not have equal power, probably because the non-polarizing beam splitter does not meet specification. Consequently, inherent tilt dependency is increased since the incident power change at one detector due to surface tilt is not entirely compensated by that at the other detector. It would be advisable to restore the balance by replacing the beam splitter, or by changing the gain of the signals. Changing the gain can be done either analog or digitally; both require the processing electronics to be changed.

Besides these modifications, some matters of the current sensor deserve further research. It would be interesting to test the calibration method for tilt dependency as applied here against other methods than the tilted flat measurement performed here, for example against measurements of a reference sphere.

The focus error signal will not be the same for measurement of rough surfaces as for measurement of specular surfaces. Furthermore, tracking will require another approach since for rough surfaces, the power of the returning light is much lower due to scattering at the surface. Vibrating the objective lens might be beneficial in this case. To develop a new tracking strategy, research is needed in which the signals for surfaces of different materials and roughness are investigated.

The measurement uncertainty of the sensor cannot yet be determined because the magnitude of errors in the interferometer measurements due to changes in the refractive index of air are not known. In this respect, the measurement errors due to pressure variations in the focusing unit that are caused by pumping are especially of interest because they are expected to vary too quickly to compensate. A method to quantify these errors is to let the focusing unit describe various trajectories while

measurements taken by the sensor's interferometer are compared to those with a reference system. For this reason, a mirror can be mounted to the object lens assembly, and an external interferometer can be used to record its displacements.

During tests with the demonstrator, a cyclic disturbance in the focus error signal was observed. A solution was found that is also applied in the sensor design for NANOMEFOS. Nevertheless, the cause of this cyclic error remains elusive; especially the periodicity that is higher than what would be expected from interference raises questions about what mechanism causes the cyclic disturbance. Because a solution to the problem has been found, developing a sound theory about the underlying mechanism is not urgent. It would be interesting however, to investigate it further, since this possibly leads to a reduction of costs.

BIBLIOGRAPHY

Abbe E. `Meßapparate für physiker´ [Journal] // *Zeitschrift für instrumentenkunde*. - Vol. 10, pp. 446-448, 1880.

Agilent technologies. `*Optics and laser heads for laser-interferometer positioning systems, product overview´* [Datasheet] retrieved from : http://cp.literature.agilent.com/litweb/pdf/5964-6190E.pdf, 2009.

Amstel W.D. van [Private communications], 2008.

Amstel W.D. van, Cacace L.A. and Henselmans R. `*Optical sensor with tilt error correction´* [European patent application] // European application number EP-A1-07108395.0. - European application date : May 16, 2007.

Anodizing [Web site] *Wikipedia* http://en.wikipedia.org/wiki/Anodizing - June 24, 2009.

Benschop J. and van Rosmalen G. `Confocal compact scanning optical microscope based on compact disc technology´ [Journal] // *Applied optics*. - Vol. 30, No. 10, pp. 1179-1184, 1991.

Boers, S.B. `Ontwerp van een draaibank met elastische geleidingen voor fabricage van niet-rotatiesymmetrische lenzen´ [PhD Thesis]: Technische Universiteit Eindhoven, 1999.

Bouwhuis G. and Braat J.J.M. `Video disk player optics´ [Journal] // *Applied optics*. - Vol. 17, No. 13, pp. 1993-2000, 1978.

Corle T.R., Chou C.H. and Kino G.S. `Depth response of confocal optical microscopes´ [Journal] // *Optics letters*. - Vol. 11, No. 12, pp. 770-772, 1986.

Corle T.R., Fanton J.T. and Kino G.S. `Distance measurements by differential confocal optical ranging´ [Journal] // *Applied optics*. - Vol. 26, No. 12, pp. 2416-2420, 1987.

Cosijns S.J.A.G. `*Displacement laser interferometry with sub-nanometer uncertainty´* [PhD Thesis] - ISBN 90-386-2656-8 : Technische Universiteit Eindhoven, 2004.

Creath K. `Calibration of numerical aperture effects in interferometric microscope objectives´ [Journal] // *Applied optics*. - Vol. 28, No. 15, pp. 3333-3338, 1989.

Cuser L. `*Optical properties of silicon´* [PDF] – Fredericksburg, Virginia Semiconductor, Inc., 2004.

De Nicola S., Carbonara, G., Finizio, A. and Pierattini, G. `Measurement of the temperature dependence of quartz refractive indices – A technique´ [Journal] // *Applied physics B: Lasers and optics.* - Vol. 58, No. 2, pp. 133-135, 1994.

Deck L. and Groot P. de `High-speed noncontact profiler based on scanning white-light interferometry´ [Journal] // *Applied optics*. - Vol. 33, No. 31, pp. 7334-7338, 1994.

DEF STAN 03-24/5 [UK Defense standard] - Glasgow : UK Ministery of defence, 2008.

DEF STAN 03-25/5 [UK Defense standard] - Glasgow : UK Ministery of defence, 2008.

Dobson S.L., Sun P. and Fainman Y. `Diffractive lenses for chromatic confocal imaging´ [Journal] // *Applied optics*. - Vol. 36, No. 20, pp. 4744-4748, 1997.

Downs K.P. and Birch M.J. `Correction to the updated Edlèn equation for the refractive index of air´ [Journal] // *Metrologia*. - Vol. 31, No. 4, pp. 315–316, 1994.

Edlèn B. `The refractive index of air´ [Journal] // *Metrologia*. - Vol. 2, No. 2, pp. 71–80, 1966.

Ehrmann K., Ho A. and Schindhelm K. `A 3D optical profilometer using a compact disc reading head´ [Journal] // *Measurement science and technology.* - Vol. 9, pp. 1259-1265, 1998.

Erps J. van, Desmet, L., Vervaeke, M., Stijns, E. and Thienpont, H.`Investigation on metal reflection coatings of free-space optical interconnect components with integrated fan-out DOEs´ [Conference proceedings] // *Symposium IEEE/LEOS Benelux chapter.* pp. 277-280, Enschede, 2003.

Fainman Y., Lenz E. and Shamir J. `Optical profilometer: a new method for high sensitivity and wide dynamic range´ [Journal] // *Applied optics*. - Vol. 21, No. 17, pp. 3200-3208, 1982.

Gao W. and Kiyono S. `Development of an optical probe for profile measurement of mirror surfaces´ [Journal] // *Optical engineering*. - Vol. 36, No. 12, pp. 3360-3366, 1997.

Hamilton D.K. and Wilson T. `Three-dimensional surface measurement using the confocal scanning microscope´ [Journal] // *Applied physics*. - Vol. 27, No. 4, pp. 211-213, 1982.

Hartmann L., Meiners-Hagen K. and Abou-Zeid A. `An absolute distance interferometer with two external cavity diode lasers´ [Journal] // *Measurement science and technology*. - Vol. 19, pp. 1-6, 2008.

Hecht E. `*Optics - Fourth edition´* [Book] - ISBN 0-321-18878-0 : Addison Wesley, 2002.

Henselmans R. `*NANOMEFOS Probe Requirements´* [Report] - Technische Universiteit Eindhoven, 2005.

Henselmans R. `*Non-contact measurement machine for freeform optics´* [PhD Thesis] - ISBN 978-90-386-1607-0 : Technische Universiteit Eindhoven, 2009.

Heydemann P.L.M. `Determination and correction of quadrature fringe measurement errors in interferometers´ [Journal] // *Applied optics*. - Vol. 20, No. 19, pp. 3382–3384, 1981.

IRE Symbols committee et al. `IEEE standard letter symbols for semiconductor devices´ [IEEE standard] // *IEEE transactions on electron devices*. - Vol. 11, No. 8, pp. 392-397, 1964.

ISO 10110-5:2007 `*Optics and photonics - Preparation of drawings for optical elements and systems - Part 5: Surface form tolerances´* [ISO standard] - 2007.

ISO 11146:2005 `*Lasers and laser-related equipment – Test methods for laser beam widths, divergence angles and beam propagation ratios´* [ISO standard] - 2005.

Johnson J.B. `Thermal agitation of electricity in conductors – The experiment´ [Journal] // *Physical review*. - Vol. 32, No. 1, pp. 97-109, 1928.

Jones C. and Hurwitz H. Jr. `New calculus for the treatment of optical systems´ [Journal] // *Journal of the optical society of America*. - Vol. 31, pp. 488–499, 1941.

Kino G.S. and Corle T.R. `Confocal scanning optical microscopy´ [Journal] // *Physics today*. - Vol. 42 No. 9, pp. 55-62, 1989.

Kohn C. `*The theory of SPV for measuring diffusion length´* [PDF] - TN201 *Rev. 1.1* : Semilab, 2009.

Kohno T., Ozawa, N., Miyamoto, K. and Musha, T. `High precision optical surface sensor´ [Journal] // *Applied optics*. - Vol. 27, No. 1, pp. 103-108, 1988.

Kooijman C.S. `*PSD processing electronics´* [Report] – Philips : Centre for Manufacturing Technology, No Date.

Lee C.H. and Wang J. `Noninterferometric differential confocal microscopy with 2-nm depth resolution´ [Journal] // *Optics communications*. - No. 135, pp. 233-237, 1997.

Linos photonics, Inc [Product catalog], 2008.

Lou D.Y., Martinez A. and Stanton D. `Surface profile measurement with a dual-beam optical system´ [Journal] // *Applied optics*. - Vol. 23, No. 5, pp. 746-751, 1984.

Mathworks `*Spectral analysis - Statistical signal processing - Signal processing toolbox´* [Matlab R2007b help document], 2007.

McDaniel T.W. and Victora R.H. `*Handbook of magneto-optical data recording: materials, subsystems, techniques.´* [Book] - ISBN 978-0-8155-1391-9 : William Andrew Publishing/Noyes, 1997.

Melles Griot [Product catalog], 2008.

Michelson A.A. and Morley E.W. `On the relative motion of the earth and the luminiferous ether´ [Journal] // *American journal of science*. - 3rd series, Art. 34, pp. 333–345, 1887.

Minsky M. `*Microscopy apparatus´* [United States patent] // United States patent number 3,013,467. - United States application date: November 7, 1957.

Mitsui K. `In-process sensors for surface roughness and their applications´ [Journal] // *Precision engineering*. - Vol. 8, No. 4, pp. 212-220, 1986.

Molesini G., Pedrini, G., Poggi, P. and Quercioli, F.`Focus-wavelength encoded optical profilometer´ [Journal] // *Optics communications*. - Vol. 49, No. 4, pp. 229-233, 1984.

National Instruments *NI DAQCard-6036E* [Datasheet], 2007.

Nyquist H. `Thermal agitation of electric charge in conductors – The theory´ [Journal] // *Physical Review*. - Vol. 32, No. 1, pp. 110-113, 1928.

Ono Y., Yamanaka Y. and Kimura Y. `*Optical head comprising a diffraction grating for directing two or more diffracted beams to optical detectors´* [United States patent] // United States patent number 4,945,529. - United States application date: December 10, 1986.

O'Shea D.C. `*Elements of modern optical design´* [Book] - ISBN 0-471-07796-8 : Wiley-Interscience, 1985.

OSI Optoelectronics `*Photodiode characteristics and applications´* [Application note], 2009.

Ravensbergen S.K. `*Probe guidance and focus drive for the NANOMEFOS optical sensor´* [Report] - Technische Universiteit Eindhoven, 2006.

Rosielle P.C.J.N. and Reker E.A.G. `*Constructieprincipes 1 - Bedoeld voor het nauwkeurig bewegen en positioneren´* [Lecture notes in Dutch] - Technische Universiteit Eindhoven, Lecture notes 4007, 2000.

Ruprecht A. `*Konfokale Sensorik zur Hochgeschwindigkeits-Topografiemessung technischer Objekte´* [PhD thesis] - ISBN 978-3-923560-58-5 : Universität Stuttgart, 2008.

Savio E., De Chiffre L. and Schmitt R. `Metrology of freeform shaped parts´ [Journal] // *Annals of the CIRP*. - Vol. 56/2, pp. 810–835, 2007.

Schott AG `*Technical information TIE-19: Temperature coefficient of the refractive index´* [Digital article] - retrieved from: http://www.schott.com/... advanced_optics/english/our_products/materials/optical_glass.html, 2008.

Schott AG `*Zerodure´* [Datasheet] - retrieved from: http://www.schott.com/... advanced_optics/english/our_products/zerodur/zerodur.html, 2008.

Schottky W. `Über spontane Stromschwankungen in verschiedenen Elektrizitätsleitern´ [Journal] // *Annalen der Physik*. - Vol. 57, pp. 541–567, 1918.

Seggelen J.K. van `*NanoCMM - A 3D coordinate measuring machine with low moving mass for measuring small products in array with nanometer uncertainty´* [PhD Thesis] - ISBN 978-90-386-2629-1 : Technische Universiteit Eindhoven, 2007.

Sheppard C.J.R. and Wilson T. `Depth of field in the scanning microscope´ [Journal] // *Optics letters*. - Vol. 3, No. 3, pp. 115-117, 1978.

Simon J. `New Noncontact Devices for Measuring Small Microdisplacements´ [Journal] // *Applied optics*. - Vol. 9, No. 10, pp. 2337-2340, 1970.

Stan S.G. `*The CD-ROM drive: a brief system description.´* [Book] - ISBN 978-0792381679 : Springer, 1998.

STIL A.S. `*Monobloc optical pens - Micrometric measurement range.´* [Datasheet], 2008.

Stout K.J. and Blunt L. `*Three-dimensional surface topography´* [Book] - ISBN 978-1-85718-026-8 : Jessica Kingsley Publishers Ltd, 1994.

Tan J. and Wang F. `Theoretical analysis and property study of optical focus detection based on differential confocal microscopy´ [Journal] // *Measurement science and technology*. - Vol. 13, pp. 1289-1293, 2002.

Tan J. and Zhang J. `Differential confocal optical system using gradient-index lenses´ [Journal] // *Optical Engineering*. - Vol. 42, No. 10, pp. 2868-2871, 2003.

Tanwar L.S. and Kunzmann H. `An electro-optical sensor for microdisplacement measurement and control´ [Journal] // *Journal of physics E: Scientific instruments*. - Vol. 17, pp. 864-866, 1984.

Taylor B.N. and Kuyatt C.E. `*NIST technical note 1297 - Guidelines for evaluating and expressing the uncertainty of NIST measurement results´* [Report], 1994.

Valter D. `Dependence of two- and three-dimensional optical transfer functions on pinhole radius in a coherent confocal microscope´ [Journal] // *Journal of the optical society of America, A - Optics image science and vision.* - Vol. 9, No. 5, pp. 725-731, 1992.

Visscher M. `*The measurement of the film thickness and the roughness deformation of lubricated elastomers´* [PhD Thesis] - URN:NBN:NL:UI:25-381654, Technische Universiteit Eindhoven, 1992.

Wavre N. and Thouvenin X. `Voice-coil actuators in space´ [Conference] // *Proceedings of the Sixth European space mechanisms & tribology symposium* : pp. 227-232. – Zurich : European Space Agency, 1995.

Wheeler H.A. `Simple Inductance Formulas for Radio Coils´ [Conference] // *Proceedings of the I.R.E.* - Vol. 16, No. 10, pp. 1398-1400, 1928.

Wilson T. and Carlini A.R. `Size of the detector in confocal imaging systems´ [Journal] // *Optics letters*. - Vol. 12, No. 4, pp. 227-229, 1987.

Wilson T. `Imaging properties and applications of scanning optical microscopes´ [Journal] // *Applied physics*. - Vol. 22, pp. 119-128, 1980.

Yajima H., Wakiwaka, H., Minegishi, K., Fujiwara, N. and Tamura, K. `Design of linear DC motor for high-speed positioning´ [Journal] // *Sensors and actuators*. - No. 81, pp. 281–284, 2000.

Zhao W., Tan J. and Qiu L. `Bipolar absolute differential confocal approach to higher spatial resolution´ [Journal] // *Optics express*. - Vol. 12, No. 24, pp. 5013-5021, 2004.

APPENDIX A. DERIVATION OF PINHOLE PARAMETERS FOR MAXIMUM SENSITIVITY

Axial pinhole offset for maximum sensitivity

The position of the pinholes along the optical axis that results in the highest sensitivity of the system will be there where the infinitesimal change of power transmission through the pinholes is at a maximum for an infinitesimal axial displacement of the waist. The cross-sectional intensity distribution of a laser beam has the same shape everywhere along its axis, i.e. only the width and height of the Gaussian distribution scale. Therefore, the optimal position of the pinholes can be found by evaluating the relative change of the cross-sectional area of the beam, as will be shown here. Note that the optimal axial position of the pinholes can be calculated while the pinhole diameter is unknown. The optimal pinhole diameter will be calculated further on, using the axial position along the beam that is found here.

For a laser operating in TEM_{00} mode the width of the beam is given by (O'Shea, 1985):

$$D_z^2 = D_0^2 + \theta_f^2 z^2 \tag{A.1}$$

Where:

z is the axial distance from the beam's waist,

D_z is the diameter of the beam at position z defined by the e^{-2} point,

D_0 is the diameter of the beam at its waist defined by the e^{-2} point, and

θ_f is the divergence (Note that the subscript f is used to indicate full-angle, contrary to θ which is used for calculation of NA where θ denotes the half-angle).

The cross-sectional area of the beam at position z is given by:

$$A = \frac{1}{4}\pi D_z^2 = \frac{1}{4}\pi(D_0^2 + \theta_f^2 z^2) \qquad (A.2)$$

Taking the derivative of A with respect to z:

$$\frac{d}{dz}A(z) = \frac{1}{2}\pi\theta_f^2 z \qquad (A.3)$$

Dividing the derivative of A with respect to z by A itself gives the relative change of A for an infinitesimal axial displacement:

$$\frac{\frac{d}{dz}A(z)}{A} = \frac{2\theta_f^2 z}{D_0^2 + \theta_f^2 z^2} \qquad (A.4)$$

The optimal axial position of the pinholes is given by the value of z for which $\dfrac{\frac{d}{dz}A(z)}{A}$ reaches its maximum, thus for:

$$\frac{d}{dz}\frac{\frac{d}{dz}A(z)}{A} = \frac{2\theta_f^2 D_0^2 - 2\theta_f^4 z^2}{(D_0^2 + \theta_f^2 z^2)^2} = 0 \qquad (A.5)$$

$$2\theta_f^2 D_0^2 - 2\theta_f^4 z^2 = 0 \qquad (A.6)$$

$$z^2 = \frac{D_0^2}{\theta_f^2} \qquad (A.7)$$

$$z = \frac{D_0}{\theta_f} = z_r \qquad (A.8)$$

Where:

z_r is the Rayleigh range.

Hence, the optimal pinhole offset with respect to sensitivity of the unnormalized FES is equal to the Rayleigh range.

Pinhole diameter for maximum sensitivity

The diameter of the pinholes resulting in the highest sensitivity of the system will be the diameter that leads to the highest infinitesimal change of power transfer through the pinholes for an infinitesimal axial displacement relative to the beam waist. Moving of the beam waist relative to the pinholes leads to widening or tightening of the beam diameter at the position of the pinholes.

The radial intensity distribution of a Gaussian beam is given by:

$$I_D = I_0 \, e^{-2\frac{D^2}{D_z^2}} \qquad (A.9)$$

Where:

I_D is the intensity at diameter D in the beam, and

D is the diameter in the beam at which the intensity is I_D.

From this expression, it can be derived that the energy passing through a circle (in this case the pinhole) which is concentric with the beam's axis, the so-called encircled power, can be expressed as:

$$P_{enc} = P_b(1 - e^{-2\frac{D^2}{D_z^2}}) = P_b(1 - e^{-2\frac{D_{ph}^2}{D_z^2}}) \qquad (A.10)$$

Where:

P_b is the power of the laser beam, and

D_{ph} is the diameter of the pinholes.

The ratio of the encircled power to the total power of the beam is called the Fractionally Transferred Power (FTP) and is given by:

$$FTP = 1 - e^{-2\frac{D_{ph}^2}{D_z^2}} \qquad (A.11)$$

D_z can be expressed as (O'Shea, 1985):

$$D_z = D_0\sqrt{1 + (z/z_r)^2} \qquad (A.12)$$

Now, if instead of z one uses dimensionless $\bar{z}$ by normalizing with z_r, as discussed in Chapter 3 and substituting 1 for z_r, the expression for D_z simplifies to:

$$D_z = D_0 \sqrt{1 + \bar{z}^2}$$ (A.13)

The FTP can be written as:

$$FTP = 1 - e^{-2 \frac{D_{ph}^2}{D_z^2 + \bar{z}^2 D_z^2}}$$ (A.14)

The sensitivity of the system depends on the absolute value of the derivative of the FTP with respect to $\bar{z}$.

$$\frac{d}{d\bar{z}} FTP(\bar{z}) = \frac{4 D_{ph}^2 D_0^2}{\left(D_0^2 + \bar{z}^2 D_0^2\right)^2} \, e^{-2 \frac{D_{ph}^2}{D_0^2 + \bar{z}^2 D_0^2}}$$ (A.15)

Substituting $\bar{z}$ is equal to 1, i.e. z is equal to z_r, as calculated to be the optimal pinhole distance in the previous subsection, this simplifies to:

$$\frac{d}{d\bar{z}} FTP(\bar{z} = 1) = \frac{D_{ph}^2}{D_0^2} \, e^{-\frac{D_{ph}^2}{D_0^2}}$$ (A.16)

For optimal performance with given beam parameters, the pinhole size needs to be chosen in such a way that the FTP changes as quickly as possible for an infinitesimal change in waist location relative to the pinhole. In other words, the optimal D_{ph} is equal to the value of D_{ph} at which the absolute value of the derivative of the FTP with respect to $\bar{z}$ has its maximum, thus for:

$$\frac{d}{dD_{ph}} \frac{D_{ph}^2}{D_0^2} \, e^{-\frac{D_{ph}^2}{D_0^2}} = \frac{2 D_{ph}}{D_0^2} \, e^{-\frac{D_{ph}^2}{D_0^2}} - 2 \frac{D_{ph}^3}{D_0^4} \, e^{-\frac{D_{ph}^2}{D_0^2}} = 0$$ (A.17)

For the optimal pinhole diameter this gives:

$$D_{ph} = D_0$$ (A.18)

Hence, the unnormalized differential confocal system has the highest sensitivity when the pinholes have a diameter equal to the waist diameter and an offset from best-focus that is equal to the Rayleigh range.

APPENDIX B. DIMENSIONAL MODEL DERIVATION FOR A DIFFERENTIAL CONFOCAL SYSTEM

From the dimensionless FES, (3.11), the FES for a complete differential confocal system can be derived. Because many of the additional terms represent linear subsystems that are in series, their terms are multiplication factors in the model. Furthermore, the *dimensionless* parameters are multiplied by the physical quantities with which they were made dimensionless. The FES then becomes:

$$
FES = G_{\mathrm{el}} \; R_{\mathrm{sut}} \; P_L \left(R_{\mathrm{pd1}} \; \eta_{\mathrm{opt1}} \; \mathrm{e}^{-2\left(\frac{D_{ph}^2}{M^2 \, D_0^2 \left(1 + \left(\frac{u_{\mathrm{ph}}}{M^2 \, z_{\mathrm{r}}} + \frac{2 \, u_{\mathrm{sut}}}{z_{\mathrm{r}}} \right)^2 \right)} \right)} - \dots \right.
$$

$$
\left. R_{\mathrm{pd1}} \; \eta_{\mathrm{opt1}} - R_{\mathrm{pd2}} \; \eta_{\mathrm{opt2}} \; \mathrm{e}^{-2\left(\frac{D_{ph}^2}{M^2 \, D_0^2 \left(1 + \left(\frac{-u_{\mathrm{ph}}}{M^2 \, z_{\mathrm{r}}} + \frac{2 \, u_{\mathrm{sut}}}{z_{\mathrm{r}}} \right)^2 \right)} \right)} + R_{\mathrm{pd2}} \; \eta_{\mathrm{opt2}} \right) \tag{B.1}
$$

in which:

G_{el} is the gain in the signal processing electronics,

R_{pd} is the responsivity of photodiodes at the wavelength used,

η_{opt} is the transmissivity of the optical train,

R_{sut} is the reflectance of the SUT,

P_L is the output power of the laser,

f_1 and f_2 are the focal-lengths of objective and image lenses respectively, and

$$M = \frac{f_2}{f_1} \qquad \text{is the magnification of the system.}$$

From Gaussian beam theory (O' Shea, 1985) it can be derived that Rayleigh range and waist diameter can be expressed as:

$$D_0 = \frac{4 f_1 \lambda}{\pi D_B} \tag{B.2}$$

$$z_{\mathrm{r}} = \frac{4 f_1^2 \lambda}{\pi D_B^2} \tag{B.3}$$

with:

λ the wavelength of the laser source, and

D_B the waist of the laser beam.

Substitution yields:

$$FES = G_{\mathrm{el}} \, R_{\mathrm{sut}} \, P_L \left(R_{\mathrm{pd1}} \, \eta_{\mathrm{opt1}} \, \mathrm{e}^{\left(-2 \left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \, \lambda}{\pi D_B} \right)^2 \left(1 + \left(\frac{\pi D_B^2 \, u_{\mathrm{ph}}}{4 f_2^2 \, \lambda} + \frac{2 \pi D_B^2 \, u_{\mathrm{sut}}}{4 f_1^2 \, \lambda} \right)^2 \right)} \right) \right)} - \ldots \right.$$
$$\left. R_{\mathrm{pd1}} \, \eta_{\mathrm{opt1}} - R_{\mathrm{pd2}} \, \eta_{\mathrm{opt2}} \, \mathrm{e}^{\left(-2 \left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \, \lambda}{\pi D_B} \right)^2 \left(1 + \left(\frac{-\pi D_B^2 \, u_{\mathrm{ph}}}{4 f_2^2 \, \lambda} + \frac{2 \pi D_B^2 \, u_{\mathrm{sut}}}{4 f_1^2 \, \lambda} \right)^2 \right)} \right) \right)} + R_{\mathrm{pd2}} \, \eta_{\mathrm{opt2}} \right) \tag{B.4}$$

To shorten the notation, the exponential terms can be expressed as e_1 and e_2, thus:

$$e_1 = \mathrm{e}^{\left(-2 \left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \, \lambda}{\pi D_B} \right)^2 \left(1 + \left(\frac{\pi D_B^2 \, u_{\mathrm{ph}}}{4 f_2^2 \, \lambda} + \frac{2 \pi D_B^2 \, u_{\mathrm{sut}}}{4 f_1^2 \, \lambda} \right)^2 \right)} \right) \right)} \tag{B.5}$$

and

$$e_2 = \mathrm{e}^{\left(-2 \left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \, \lambda}{\pi D_B} \right)^2 \left(1 + \left(\frac{-\pi D_B^2 \, u_{\mathrm{ph}}}{4 f_2^2 \, \lambda} + \frac{2 \pi D_B^2 \, u_{\mathrm{sut}}}{4 f_1^2 \, \lambda} \right)^2 \right)} \right) \right)} \tag{B.6}$$

The FES can now be written as:

$$FES = G_{\text{el}}\, R_{\text{sut}}\, P_L \left(R_{\text{pd1}}\, \eta_{\text{opt1}}\, e_1 - R_{\text{pd1}}\, \eta_{\text{opt1}} - R_{\text{pd2}}\, \eta_{\text{opt2}}\, e_2 + R_{\text{pd2}}\, \eta_{\text{opt2}} \right) \tag{B.7}$$

The normalized FES can be written as:

$$\mathbb{FES} = \mathbb{G}_{\text{el}} \frac{
\left(R_{\text{pd2}}\, \eta_{\text{opt2}} - R_{\text{pd2}}\, \eta_{\text{opt2}}\, e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2\, \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{-\pi D_B^2\, u_{\text{ph}}}{4 f_2^2\, \lambda} + \frac{2\pi D_B^2\, u_{\text{sut}}}{4 f_1^2\, \lambda}\right)^2\right)}\right)} \right) - \cdots
}{
\left(R_{\text{pd2}}\, \eta_{\text{opt2}} - R_{\text{pd2}}\, \eta_{\text{opt2}}\, e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2\, \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{-\pi D_B^2\, u_{\text{ph}}}{4 f_2^2\, \lambda} + \frac{2\pi D_B^2\, u_{\text{sut}}}{4 f_1^2\, \lambda}\right)^2\right)}\right)} \right) + \cdots
}$$
$$\cdots \frac{
\left(R_{\text{pd1}}\, \eta_{\text{opt1}} - R_{\text{pd1}}\, \eta_{\text{opt1}}\, e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2\, \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^2\, u_{\text{ph}}}{4 f_2^2\, \lambda} + \frac{2\pi D_B^2\, u_{\text{sut}}}{4 f_1^2\, \lambda}\right)^2\right)}\right)} \right)
}{
\cdots \left(R_{\text{pd1}}\, \eta_{\text{opt1}} - R_{\text{pd1}}\, \eta_{\text{opt1}}\, e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2\, \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^2\, u_{\text{ph}}}{4 f_2^2\, \lambda} + \frac{2\pi D_B^2\, u_{\text{sut}}}{4 f_1^2\, \lambda}\right)^2\right)}\right)} \right)
} \tag{B.8}$$

where:

$\mathbb{G}_{el}$ is the gain in the normalizing signal processing electronics.

or, in the simplified form, as:

$$\mathbb{FES} = \mathbb{G}_{\text{el}} \left(\frac{\left(R_{\text{pd2}}\, \eta_{\text{opt2}} - R_{\text{pd2}}\, \eta_{\text{opt2}}\, e_2 \right) - \left(R_{\text{pd1}}\, \eta_{\text{opt1}} - R_{\text{pd1}}\, \eta_{\text{opt1}}\, e_1 \right)}{\left(R_{\text{pd2}}\, \eta_{\text{opt2}} - R_{\text{pd2}}\, \eta_{\text{opt2}}\, e_2 \right) + \left(R_{\text{pd1}}\, \eta_{\text{opt1}} - R_{\text{pd1}}\, \eta_{\text{opt1}}\, e_1 \right)} \right) \tag{B.9}$$

When enough light is used, the photodiode noise becomes negligible, as does the photon shot noise for both detector branches. When high quality photodiodes from the same batch are used one can assume that R_{pd1} and R_{pd2} are more or less equal to each other. Also when a high quality laser line non-polarizing beam splitter is used η_{opt1} and η_{opt2} can be considered equal. With these assumptions, the expression for the FES now becomes:

$$FES = G_{\text{el}}\, R_{\text{sut}}\, R_{\text{pd}}\, \eta_{\text{opt}}\, P_L \cdots$$
$$\left(e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2\, \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^2\, u_{\text{ph}}}{4 f_2^2\, \lambda} + \frac{2\pi D_B^2\, u_{\text{sut}}}{4 f_1^2\, \lambda}\right)^2\right)}\right)} - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2\, \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{-\pi D_B^2\, u_{\text{ph}}}{4 f_2^2\, \lambda} + \frac{2\pi D_B^2\, u_{\text{sut}}}{4 f_1^2\, \lambda}\right)^2\right)}\right)} \right) \tag{B.10}$$

or:

$$FES = G_{el}\, R_{sut}\, R_{pd}\, \eta_{opt}\, P_L \left(e_1 - e_2 \right) \tag{B.11}$$

And the normalized FES becomes:

$$\mathbb{FES} = \mathbb{G}_{el} \left(\frac{e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^2 u_{ph}}{4 f_2^2 \lambda} + \frac{2\pi D_B^2 u_{sut}}{4 f_1^2 \lambda}\right)^2\right)}\right)} - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{-\pi D_B^2 u_{ph}}{4 f_2^2 \lambda} + \frac{2\pi D_B^2 u_{sut}}{4 f_1^2 \lambda}\right)^2\right)}\right)}}{2 - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{\pi D_B^2 u_{ph}}{4 f_2^2 \lambda} + \frac{2\pi D_B^2 u_{sut}}{4 f_1^2 \lambda}\right)^2\right)}\right)} - e^{-2\left(\frac{D_{ph}^2}{\left(\frac{4 f_2 \lambda}{\pi D_B}\right)^2 \left(1 + \left(\frac{-\pi D_B^2 u_{ph}}{4 f_2^2 \lambda} + \frac{2\pi D_B^2 u_{sut}}{4 f_1^2 \lambda}\right)^2\right)}\right)}} \right) \tag{B.12}$$

or:

$$\mathbb{FES} = \mathbb{G}_{el} \left(\frac{e_1 - e_2}{2 - e_1 - e_2} \right) \tag{B.13}$$

DANKWOORD / ACKNOWLEDGEMENTS

Mijn afstudeer-, freelance- en promotiewerk bij elkaar opgeteld, heb ik nu bijna vier en een half jaar aan dit project gewerkt. Morgen gaat mijn boekje naar de drukker en hoewel een project als dit nooit helemaal klaar is, voelt dat als de afronding van een levensfase. Ook al werk je hard en doe je wat je kunt, geluk blijft een belangrijke factor. Ik denk dat ik ontzettend veel geluk heb gehad met de betrokken, interessante, enthousiaste en getalenteerde mensen om me heen, zowel privé als professioneel, en met de kans om zo'n veelzijdig project aan te pakken. Naast het leuke, spannende werk heb ik het fascinerend gevonden om te zien hoe mensen met zulke verschillende persoonlijkheden en specialismen verenigd zijn door hun passie voor techniek.

Nick, ik kijk er altijd met bewondering naar hoe je al je ervaring en kennis ogenschijnlijk zo gemakkelijk ombuigt naar andere situaties. Buitenom je technische begeleiding heb ik je morele bijstand ervaren als een grote steun. Nadat ik in het begin soms moeite had met je metaforen, ben ik deze gaandeweg juist erg gaan waarderen. Soms tijden na een gesprek, in een heel andere setting, word ik herinnerd aan één van die metaforen en begrijp ik ineens wat je bedoelde. Dank je.

Maarten, bedankt voor de kans die je me hebt geboden om via deze weg te promoveren en voor de fijne, effectieve en betrokken begeleiding. Ik heb de openheid die je oproept en de drive die je hebt om te helpen als zeer prettig ervaren. Als ik na een IWO alles nog eens door mijn hoofd laat spelen verbaast het me vaak hoe je aanvoelt wat ik op een bepaald moment moet horen om een voor mij optimaal resultaat te behalen. Zoals jij het zou kunnen zeggen… gaaf!

Rens, het was een genot om met je te werken en stap voor stap je machine te zien ontstaan. Ik heb onze lange en soms intensieve samenwerking als erg motiverend ervaren en heb veel van je geleerd. Daarbij ben je altijd bijzonder behulpzaam geweest, zowel toen ik nog aan de sensor werkte als toen ik aan mijn thesis bezig

was. Bedankt voor de mooie tijd in een leuk project. Ik hoop dat onze wegen elkaar nog vaak zullen kruisen.

Wim, jouw inzet en goede adviezen en onze gepassioneerde technische discussies hebben enorm geholpen bij het tot stand komen van zowel de sensor zoals die nu is, als bij dit proefschrift. Daarvoor bedankt! Bovendien vind ik het fijn dat we in de loop van de jaren naast collega's ook vrienden zijn geworden.

Geerten, jou wil ik bedanken voor alle regeltechnische uitleg, de stimulerende gesprekken en je droge humor. Het was mooi om te zien hoe je met je rustige stijl en gestructureerde aanpak zulke grote hoeveelheden werk wist te verzetten.

Graag wil ik iedereen van de Gemeenschappelijke Technische Dienst bedanken voor de altijd vriendelijke, goede hulp. Of het nou om vervaardiging of het lenen van gereedschap ging, vaak kreeg ik niet alleen waar ik naar vroeg, maar ook een waardevolle tip of een nuttig truckje. Peter, je hebt bijna alle metalen onderdelen van de sensor gemaakt en wist altijd weer te vertellen hoe we de nauwkeurigheid van de machines het beste konden gebruiken. Michiel, hoewel je meer aan de machine hebt gewerkt dan aan de sensor was je altijd bereid om te helpen. Rinus, bedankt voor het aanstekelijke enthousiasme waarmee je de elektronica voor je rekening hebt genomen en de begrijpelijke, diepgaande uitleg die je er altijd graag bij geeft. Erwin, het was inspirerend om te zien hoe bij jou praktijk en theorie zozeer met elkaar verweven zijn.

Van het VSL: Rob, bedankt voor je input op metrologisch vlak, Gerard door je enthousiasme en grondige aanpak vond ik het leuk om samen met je te meten.

Simon, bedankt voor de waardevolle bijdrage die je aan de sensor hebt geleverd tijdens je stage. Ik hoop dat je er net zo trots op bent als ik.

Jacob Jan, bedankt voor de mooie kans die je me hebt gegeven om na mijn afstuderen voor TNO als freelancer verder te werken aan dit leuke project.

De mensen bij TNO wil ik bedanken voor de hulp en middelen die ik kreeg bij mijn experimenten. In het bijzonder wil ik Mariët, Andries, Harry, Johan, Kees, Pieter, Reinder en Teun bedanken. Voor het aanpassen en vervaardigen van de optische componenten: Rob, Bob, Jacob, Maarten en Casper, bedankt!

Graag wil ik de leden van de begeleidingscommissie en Lou, Eddy en Casper van het IOP Precisie Technologie bedanken voor hun input en steun.

Hierbij wil ik Prof. Rob Munnig Schmidt (TU Delft), Prof. Gerrit Kroesen (TU/e) en Prof. Paul Urbach (TU Delft) bedanken voor hun inspanningen om dit dik uitgevallen boek te lezen en mij van nuttige feedback te voorzien.

De mensen uit het Constructies & Mechanismen lab en de W collega's waarmee ik te maken heb gehad gedurende de afgelopen jaren wil ik hier graag bedanken. In het bijzonder wil ik de dames van het DCT secretariaat bedanken: Petra en Lia.

Ian en Ellen, naast de leuke feesten en goede gesprekken wil ik jullie hier bedanken voor de gastvrijheid die ik genoot toen ik veel experimenteerde in Delft, en bij jullie in Rotterdam mijn derde thuis had. Wanneer komen jullie nou eens een paar dagen per week in Eindhoven werken? ;-)

Bastiaan, we zijn vrienden voor het grootste gedeelte van ons leven en we hebben samen veel avonturen beleefd. Daarnaast ben je voor mij ook een belangrijk klankbord voor grote beslissingen, bedankt.

1s, Duck, Merijn en Robbert, de vrijdag avond is voor mij behalve lekker muziek maken een gezellig ritueel met vrienden geworden waarin we de week afsluiten en het weekend inluiden. Rock on!

Verder wil ik de volgende mensen bedanken voor hun vriendschap, de waardevolle momenten en voor het begrip tijdens de afgelopen periode waarin ik niet genoeg tijd heb ingeruimd voor alle dingen die in het leven belangrijk zijn: / Furthermore I would like to thank the following people for their friendship, valuable moments and understanding during the past period in which I didn't reserve enough time for all the important things in life: Anil&Akin, Arjen, Freek, Geeske, Gerdien, Giaco, Immie, Menno, Milko, Michiel&Mojca, Raimond&Yvonne, Rian, Ronald, Seçil, Tina, Tjeerd, Vincent en Wouter.

Mijn ouders: gelukkig zijn er te veel dingen om op te noemen waar ik jullie voor wil bedanken. Daarom wil ik me hier beperken tot jullie bijdrage aan dit proefschrift en de weg die daar heen heeft geleid. Cobi, jou wil ik bedanken voor alle hulp op taalkundig vlak en de eindeloze inzet en het geduld waarmee je dat hebt gedaan. Michele, jou wil ik bedanken voor het samen klussen, het sleutelen aan de auto in mijn kindertijd en de weken aan boord met die fascinerende uren in de machinekamer. Aan jou heb ik mijn liefde voor techniek te danken, waardoor ik werk ben gaan doen waar ik plezier aan beleef. Waar ik vandaan kom verandert niet meer; samen hebben jullie ervoor gezorgd dat dat goed voelt. Bedankt!

Petra, bedankt voor het plezier, de passie en de liefde die we samen beleven, en voor je steun tijdens het afgelopen jaar, dit terwijl je het zelf ook heel druk hebt gehad. Met jou hoop ik later samen op een bankje te zitten en terug te kijken op twee levens waarin we, zonder van elkaar afhankelijk te zijn geweest, meer geluk en voldoening zullen hebben gekend dan zonder elkaar. Stijn, jou wil ik bedanken voor het voor jouw leeftijd uitzonderlijke begrip dat je toont en voor alle lol die we samen hebben. Ik hoop dat je je passie in het leven snel vindt!

CURRICULUM VITAE

Lennino Cacace was born on October 22, 1980 in Middelburg, the Netherlands. After finishing his secondary education at the Christelijke Scholengemeenschap Walcheren in 1999, he studied Mechanical Engineering at the Technische Universiteit Eindhoven. In June 2006, he graduated (Cum Laude) at the Constructions & Mechanisms group (part of the Control Systems Technology section) with the master thesis "Development of a nanometer accuracy non-contact probe for NANOMEFOS". With this thesis, he won the Corus young talent graduation award for mechanical engineering and materials science, granted by the Royal Holland Society of Sciences and Humanities.

In March 2006, anticipating his graduation, he founded the company AC Optomechanix, through which he continued to work on the development of a sensor for NANOMEFOS as a freelancer. He focused on writing a Ph.D. thesis on this subject following the completion of the project in the autumn of 2008. After receiving a Ph.D. degree, he intends to keep on working for high-tech companies through AC Optomechanix.