Linear motion systems; a modular approach for improved straightness performance

Gert-Jan Nijsse

Linear motion systems; a modular approach for improved straightness performance

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.F. Wakker,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen
op dinsdag 12 juni 2001 om 10:30 uur

door

Gerard Johannes Pieter NIJSSE

Werktuigkundig ingenieur
geboren te Goes

Dit proefschrift is goedgekeurd door de promotor:
Prof.dr.ir. H.F. van Beek

Samenstelling Promotiecommissie:

Rector Magnificus,	voorzitter
Prof.dr.ir. H.F. van Beek,	Technische Universiteit Delft, promotor
Prof.dr.ir. J. van Eijk,	Technische Universiteit Delft
Prof.dr.ir. P.H.J. Schellekens,	Technische Universiteit Eindhoven
Prof.dr.ir. J. van Amerongen,	Technische Universiteit Twente
Prof.dr.ir. J.C. Compter,	Technische Universiteit Eindhoven
Dr.ir. A.J.J. van der Weiden,	Technische Universiteit Delft
Dr.ir. F. Auer,	ASML

Published and distributed by: DUP Science

DUP Science is an imprint of
Delft University Press
P.O. Box 98
2600 MG Delft
The Netherlands
Telephone: +31 15 27 85 678
Telefax: +31 15 27 85 706
E-mail: DUP@Library.TUDelft.NL

ISBN 90-407-2187-4

Keywords: Straight motion, Feedback control, Zero stiffness actuators

Printed in The Netherlands

Contents

List of symbols and abbreviations

Symbol	Unit	Description
a	m / s^2	Acceleration
A	m^2	Area
B	T, Vs / m^2	Magnetic induction, flux density
B_r	T, Vs / m^2	Remanent induction
c	Ns / m	Damping coefficient
C	F	Capacity
d	m	Distance, displacement, diameter
DoF	$-$	Degree of freedom
e	$-$	Relative error
E	N / m^2	Modulus of elasticity
E	J, Nm	Energy
f	Hz	Frequency
f_{bw}	Hz	Bandwidth
F	N	Force
F_{wp}	N	Bearing force in the working point
FEA	$-$	Finite Element Analysis
h	m	Height, thickness
H	A / m	Magnetic field strength
H_c	A / m	Coercitive field strength
H_{ci}	A / m	Intrinsic coercitive field strength
I	A	Current
I	m^4	Moment of inertia
J	A / m^2	Current density
k	N / m	Stiffness
l	m	Distance, displacement
L	m	Length
ℓ	m	Length
m	kg	Mass

M	kg	Mass
$\mathcal{O}$	–	Order of magnitude
p	N / m^2	Pressure
price	f	Price in Dutch guilders; f1.00 ≈ $ 0.40
Q	C	Electrical charge
r	m	Radius
R	m	(Reduced) radius
RH	%	Relative Humidity
s	m	Pitch, distance
t	s	Time
T	K, °C	Temperature
T_c	K, °C	Curie temperature
V	V	Voltage
W	Nm	Energy, work
x	m	Distance, displacement
y	m	(Lateral) distance, displacement
z	m	(Vertical) distance, displacement
z_{wp}	m	Vertical coordinate of the working point
α	rad	Angle
α	K^{-1}	Coefficient of thermal expansion
δ	m	Deviation (in position)
ϵ	–	Strain
ε_0	F / m	Dielectric permittivity of vacuum; $\varepsilon_0 \approx 8.8542 \cdot 10^{-12}$ F / m
ε_r	–	Relative dielectric permittivity
η	kg / ms	Dynamic viscosity
μ_0	H / m	Magnetic permeability of vacuum; $\mu_0 = 4\pi \cdot 10^{-7}$ H / m
μ_r	–	Relative magnetic permeability
ν	–	Poisson constant
σ	N / m^2	Stress
τ	N / m^2	Shear stress

Chapter 1

Introduction to straight motion

1.1 Background, history and trends

In the course of the history of precision engineering, there has been a growing need for accurately defined shapes and geometries, in order to measure and to manufacture in an accurate and repeatable way. The most basic geometries are a straight line (1-dimensional) and a flat plane (2-dimensional). Much effort has been spent to approximate these mathematically ideal shapes in a mechanical way.

The approximation of these ideal shapes is used for several purposes. Firstly, to realize a physical *form standard* in order to compare or to measure the shape of objects, e.g. manufactured products. Secondly, to realize a *motion profile* accurately describing the ideal line or plane in order to move a tool or instrument relative to an object, e.g. in manufacturing processes or alignment systems. This *form standard* or *motion profile* has to define or exclude certain degrees of freedom.

A solid object in a three-dimensional space has six degrees of freedom: three translations along the axes of an orthogonal coordinate system, defining its position, and three rotations around each of these axes, defining its orientation. When the coordinate system is chosen having one axis in the direction of the desired motion, straight motion can be described as the motion in only one of the six degrees of freedom, while all other five coordinates are equal to zero. This is depicted in figure 1.1.

Therefore, the way to realize a straight motion is to keep the undesired deviations in five degrees of freedom as small as possible. This formulation immediately implies that in practice straight motion only is an approximation of the mathematical description, having a certain inaccuracy. Many different solutions have been de-

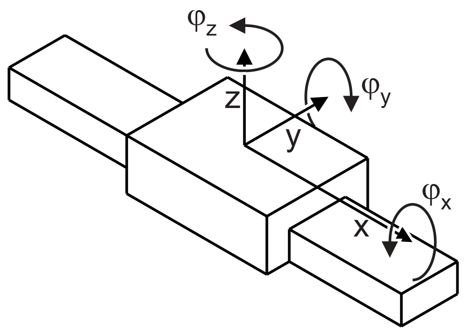


Figure 1.1: A solid object in a three-dimensional space has six degrees of freedom: three translations along the axes of an orthogonal coordinate system and three rotations around these axes. Straight motion allows one coordinate to vary while deviations in all other five degrees of freedom have to stay zero.

vised and applied to achieve the highest possible straightness. The state of the art in straight motion will be discussed in more detail further in this thesis.

In the course of history a development is visible in the design of straight motion systems, aiming at several different goals, depending on the application in which the straight motion system has to be used. These goals include:

High accuracy, high velocity: In most system designs there is a tradeoff between accuracy and velocity. To increase the velocity as well as the accuracy, generally more design effort and money will be needed. Typical values for the aimed accuracy range from the micrometer level to the nanometer level. The velocity can vary from the μm / s to m / s range. Instead of velocity also bandwidth can be read.

Integration vs. modularity: Dependent on the application a designer will either try to integrate functions or split them up in functional units. Both approaches can be advantageous.

Small size: Often also the volume is minimized in order to reduce the mass and the necessary mounting space and to achieve the highest possible velocity and bandwidth.

Low power consumption: From economical point of view, but also for reasons of thermal behavior, the power consumption can be reduced by e.g. smaller size (see above) and optimized drivers or controllers.

Low cost: From economical point of view as well, efforts are made to decrease production cost. This only makes sense for larger product series; for prototyping purposes in a research environment this is of minor interest, as long as the research budget is sufficient.

Because this thesis describes work in the field of precision engineering, the focus will be mainly on *accuracy* optimization. Besides that, also *modularity* and *cost* play an important role. Many of the described systems can be classified as 'mechatronic' systems. This term covers the overlap area between (precision) mechanics, electronics and (sometimes) physics or optics.

When concentrating on accuracy aspects in the development history of *mechanical* straight motion systems, it seems that the achieved performance approaches the micrometer range asymptotically. With regard to mechanical straight motion systems without feedback control, only in a few special cases and under strict conditions, accuracies in the submicrometer range or even lower are reported, for example in [65]. It appears that in practice it is very difficult to guarantee accuracy below $\mathcal{O}(1\,\mu\text{m})$[1] because of manufacturing errors, external loads, wear, temperature dependent deflections, etc.

[1]The symbol $\mathcal{O}$ will be used throughout this thesis to indicate an order of magnitude.

1.2 Definitions and classifications related to straightness and flatness

Throughout this thesis the coordinate convention as displayed in figure 1.2, taken from figure 1.1, will be used. The direction of motion is addressed as the x-direction. Gravity acts in the negative z-direction. Using this convention the angles around the three main axes—roll, pitch and yaw—correspond to φ_x, φ_y and φ_z, respectively.

Unless otherwise specified, the direction of motion of all described systems and tests will be the x-direction, whereas the vertical direction always is addressed as z.

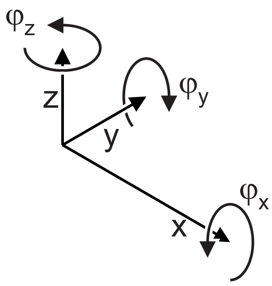


Figure 1.2: The coordinate system convention that will be used in this thesis. The gravity acts in the negative z-direction, the direction of linear motion is the x-coordinate, unless otherwise specified.

To get an impression of common orders of magnitude, the straightness tolerance of steel straightedges, classified according to the DIN 874 norm, is shown in table 1.1. The straightness deviations are defined relative to a best fit ideal straight line.

Class	**Max. deviation [μm]**
00	$1 + \ell/150$
0	$2 + \ell/100$
1	$4 + \ell/60$
2	$8 + \ell/40$

Table 1.1: DIN 874 Classification of steel straightedges, in which ℓ is the length of the straightedge in [mm].

When regarding straight motion systems, besides the *translation* also the *orientation* must be taken into account, so that actually five parameters are necessary to describe the straightness errors during motion. In table 1.2 the naming convention is given, as it is recommended by VDI 2617.

Because the linear motion systems in this thesis have only one motion axis (motion in the x-direction), the first subscript 'x' in e.g. 'Txy' can be omitted without causing

Travel direction	x-axis VDI 2617	this thesis
Translational errors		
straightness	Txy	y
straightness	Txz	z
Rotational errors		
roll	Rxx	φ_x
pitch	Rxy	φ_y
yaw	Rxz	φ_z

Table 1.2: Example of naming convention of the guide errors for a linear motion system, according to the directions given in VDI 2617. Because the systems considered in this thesis have only one motion axis, a simplified notation has been used, which is unambiguous.

ambiguity. Therefor, a simpler notation has been adopted.

Throughout this thesis, when addressing 'straightness errors' in five degrees of freedom, the translational errors as well as the rotational errors are meant.

1.3 Definitions related to accuracy and performance

In the following list the most important parameters of positioning systems are summarized. They are valid for a positioning system having a desired position as an input (for example, the setpoint of a servo system) and an physical position as an output. In figure 1.3 some of the most important accuracy related parameters are explained graphically [63].

1. Accuracy: the maximum difference between the commanded position and the actual position. The accuracy can be given for a point or for a certain interval.
2. Resolution: The minimum step or increment that can be made, so that it just can be discerned from the position noise or position uncertainty. However, when time averaging is applied, changes in the average position smaller than the position noise level can be defined and / or detected.
3. Repeatability: the maximum difference between different actual positions for the same commanded position.
4. Non-linearity: the maximum deviation from the ideal linear relation between input (commanded position) and output (actual position). Different definitions exist for absolute non-linearity (concerning the overall range), incremental non-linearity (concerning a smaller working interval) and relative non-linearity (expressed as a percentage of the regarded range).
5. Bandwidth: (in positioning systems) the frequency at which the frequency response function of the open–loop system crosses 0 dB for the first time; (in sensor systems) the frequency range up to where the output / input amplitude ratio of a sinusoidal signal is larger than -3 dB. (Other definitions exist.)

Note that the above definitions are related to position generating devices, but they can easily be translated to analogous definitions for position sensing devices.

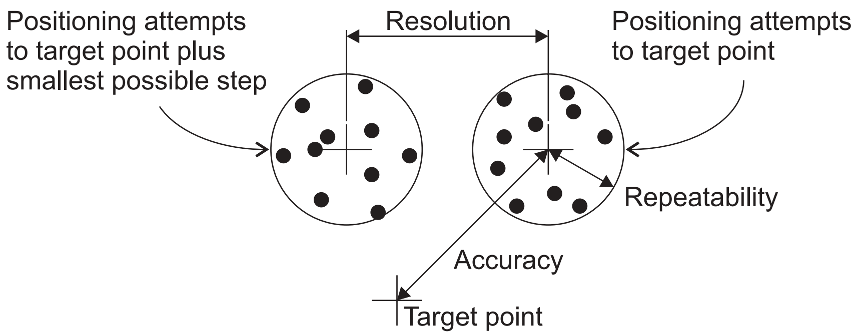


Figure 1.3: Illustration of basic terms relevant to positioning: accuracy, repeatability and resolution [63]. Analogous definitions can be derived for position sensors.

1.4 Goal of the research and contents of this thesis

The goals of this research are:

- to give a summary of the state of the art in linear and straight motion and to determine the bottle-necks in performance
- to investigate the improvements in straightness performance that can be achieved by means of feedback controlled straight motion systems
- to describe a new way to cope with environmental disturbances c.q. vibrations, in order to increase the motion accuracy

The contents of this thesis mainly follow the division mentioned above. Chapter 2 gives an overview of conventional and experimental straightness measurement systems and techniques with brief comments on performance, advantages and drawbacks and applicability in straight motion systems. In chapter 3 the state of the art in straight motion is considered. From each working principle or technology the most representative examples will be discussed, with comments on performance, advantages and drawbacks. For a well-designed system eventually there will be a limit to the performance, dictated by one or some parameters. For the described systems this limit will be investigated.

The next two chapters are the core of this thesis. In these chapters, attempts will be made to push forward the limits mentioned in the previous chapter, by combining known technologies and inventing new technologies, respectively. The emphasis will lie on straightness accuracy, position accuracy and low cost. Chapter 4 describes new experiments and developments in feedback controlled straight motion. It will appear that with the help of mechatronic components—sensors, actuators, control techniques—a substantial improvement can be achieved.

However, when aiming for the sub-micrometer and nanometer range, new limits are encountered because the environmental disturbances get increasing influence. It becomes necessary to isolate the positioning system from the (vibrations of the) world.

Therefor, in this thesis a new approach is presented, using the *zero stiffness principle*, integrated in the actuator system. The theory and new practice of zero stiffness bearings—to integrate vibration isolation in the actuator system of a feedback controlled straight motion system—are presented in chapter 5. Finally, concluding remarks and recommendations are made in chapter 6.

Chapter 2

Measurement of straightness

Before describing existing systems for the *generation* of straight motion, first some attention has to be paid to the methods of *measuring* straightness, in order to understand in particular the active straight motion systems that use a straightness sensor in their control system.

In most cases, measuring *straightness deviations* is equivalent to measuring *position deviations* in more degrees of freedom relative to a suitable straightness reference. From these position deviations the straightness deviation can be calculated; it will be expressed in units of distance [m] and angle [rad]. Therefore, this chapter will start with the description of position sensors[1], after which the most important types of straightness sensors will be discussed. Some of these position and straightness sensors are well-known and commercially available, in which case also a price indication is given, if possible. Some sensors are new because they are designed within the scope of this thesis, like the combined capacitive position / velocity sensor (section 2.1.2), the capacitive straightness sensor (section 2.2.1) and the laser straightness sensors (sections 2.2.2.1 and 2.2.2.2).

[1] Throughout this thesis, no distinction will be made between *position* and *displacement* sensors. Actually, absolute and relative position can be distinguished, while displacement could be the change in position over time. No special attention will be paid to these differences either.

2.1 General displacement sensors

This section will describe very briefly the most important displacement sensors which meet the following specifications:

- sub-micrometer resolution
- contactless operation
- electrical (analog or digital) output signal
- sufficient bandwidth (preferably > 1 kHz) for application in feedback controlled systems

In figure 2.1 the range–resolution ratio of various sensors has been plotted against the bandwidth. This is only one of the possible characterizations of position sensors, but it gives a reasonable general impression of the performance.

In the next, a summary will be given of the different sensor types and their most important properties. Much of the data is taken from an extensive literature and market research report on ‘nanosensors’ [80]. This report describes a variety of commercially available sensors, which have a resolution in the range between 1 nm and 1 μm.

Position sensors can be divided into four categories: optical, inductive, capacitive and miscellaneous. For each of the categories a short list will be given here, whereas in table 2.1 an overview of typical values for the most important specifications is given. Of course, such a short list will not be comprehensive. Its only intention is to give an overview of typical properties for commercially available sensors.

First, consider the *optical sensors*, which form a large part of the available sensors capable of measuring positions in the nanometer range.

Triangulation sensors consist of a light source (diode laser or LED), a focusing system and a light spot detection device. They can work at a considerable nominal distance (several millimeters). The overall dimension of a typical triangulation sensor unit is $50 \times 50 \times 20$ mm^3, which is rather large relative to other sensors.

Photonic sensors or fiber optic sensors mostly have two working areas: one region very close to the reflecting surface, where the light intensity strongly increases with distance, and one region at a larger distance, where the light intensity decreases due to the increasing distance. When a light conducting fiber is applied, the sensor tip only consists of a tiny fiber bundle, while the conditioning electronics can be located elsewhere. There exist IC sensor units having all optical and electrical components in a volume as small as $6 \times 2 \times 1.5$ mm^3.

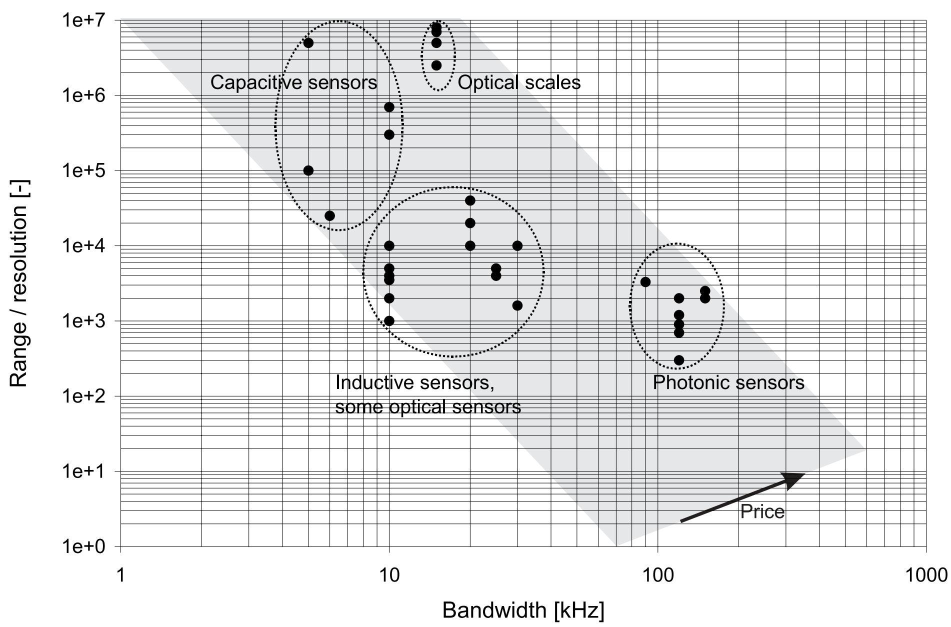


Figure 2.1: Overview of sensors that are commercially available, and capable of measuring in the nanometer range [80]. The range–resolution ratio of various sensors has been plotted against the bandwidth. The grey zone represents the area of tradeoff between relative accuracy and bandwidth. In general, the price increases when moving in the direction indicated by the arrow, but for different types of sensors considerable deviations occur. Note that the range / resolution ratio is only one possible way of characterizing position sensors.

Laser interferometers are widely applied in precision measurement technology. A complete measuring system for one degree of freedom occupies several dm^3 space. Various optical configurations exist, capable of measuring lateral displacements (using a Wollaston prism and two mirrors) or angles. Besides the rather bulky 'common' laser interferometer systems, there are compact units utilizing diode lasers, being only $\varnothing 17$ mm $\times$ 73 mm in size.

Laser autofocus sensors strongly resemble the optical system known from the compact disc player. Their working range is relatively small, several micrometers.

Optical scales are widely applied as well. There exist some types having nanometer resolution. Also 2-DoF and 3-DoF versions (with limited angular range) are known.

Second, among the *inductive sensors* several types are available that have a resolution in the sub-micrometer range. However, the nanometer level is reached only by some customized sensor types that are not in the manufacturer's catalogi, because they are only available in large quantities for the OEM[2] market.

Third, the *capacitive sensors* form a large group. In the next section, more detailed attention will be paid to this category, because of some new developments that are especially interesting for the field of multi-DoF (straightness) measurement.

Finally, the category '*miscellaneous sensors*' contains e.g. some hypothetical sensors based on the imaging elements of STMs and AFMs[3]. Their favorable properties are: very high resolution (sub-nm), small dimensions (mm range) and high bandwidth ($10 - 100$ kHz) [12]. The reasons that they are not yet widely used as universal position sensors are a.o.: the complex manufacturing process (due to which mass production is difficult), their fragility and their sensitivity to contaminations. In addition, since they are used mainly to map surfaces, the measuring range is very small, in the micrometer range.

2.1.1 Capacitive displacement sensors

Several types of capacitive sensors are commercially available, most of them consisting of one sensor electrode and an accompanying target electrode, using various types of electronic circuitry (see section 2.1.1.2). The resolution can be very high, down to the sub-nm level, at the cost of bandwidth. The measurement range is determined by the area of the electrodes and the sensitivity of the electronics, which both are constant. The resolution and bandwidth can be mutually exchanged.

In the recent past, much research has been done concerning the application of other capacitive sensor techniques in position sensing [24, 81, 28, 8]. Therefor, only a

[2]OEM: Original Equipment Manufacturer

[3]STM: Scanning Tunneling Microscope; AFM: Atomic Force Microscope.

Principle	**Range [mm]**	**Resolution [nm]**	**Bandwidth [kHz]**	**Size $[mm^3]$**	**Price [kƒ]**	**Note**
Triangulation	+/-0.25	10…50	10…20	50x50x20	10…20	1
Photonic, front slope	0.0025…0.15	2.5…25	0.01	Ø5x70	0.01…25	
		15…150	150			
Photonic, back slope	0.75…1.5	15…250	0.01			
		2000	150			
Laser interferometer	10000	1…50	v_{max}/Amplitude	100x100x400	10…50	2
Diode laser interferometer	1000	10	v_{max}/Amplitude	Ø17x73	10…20	3
	500	1.6	v_{max}/Amplitude			
Laser autofocus sensor	0.0044	5	100	Ø20x64	5	
	0.013	20				
Optical scale	70	1	v_{max}/Amplitude	20x20x20	10	2
	68x68	10			25	
	100x100x50mrad	5			5	
Inductive	1	100	20	Ø7x35	3	
Inductive, custom / OEM	<1	<1	n/a	n/a	>10	
Capacitive (synchr. det.)	<<1	Range/1e3	20	1000…10000	5	
		Range/1e6	2			
Capacitive (MMO)		Range/1e4	0.005	500	n/a	4
Capacitive (static)	+/- 0.5	200	0.01	1000	n/a	4
'xv-sensor' (dynamic)	+/- 0.005		1.5			4

Table 2.1: Overview of a selected set of specifications for a group of sensors, which are capable of measuring positions with a resolution in the $1 \dots 100$ *nm range. Notes: 1) Working distance* 0.25 *mm; 2) Typically* $v_{max} = 1$ *m/s, dependent on electronics; 3)* $v_{max} = 0.2$ *m/s; 4) Experimental sensor.*

short summary of the most relevant theory and practice will be given here, which is sufficient to compare capacitive sensors with other sensors and to understand the new developments, e.g. in section 2.1.2.

2.1.1.1 Capacitive position measuring principle

The basic equation providing the relation between geometry and capacitance of a parallel two-plate capacitor is the well-known formula

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d} \tag{2.1}$$

which is valid under the assumption that the electric field is homogeneous. In this formula C is the capacitance in Farad [F], ε_0 is the permittivity in vacuum, being approximately $8.8542 \cdot 10^{-12}$ F/m, ε_r is the relative permittivity of the medium between the capacitor plates, A is the projected overlapping area of the capacitor plates in [m^2] and d is their relative distance in [m]. As a rule of thumb, for a bi-plane capacitor in air with two 1 cm^2 electrodes with 1 mm gap the capacitance will be about 1 pF (actually 0.88542 pF).

The assumption of a homogeneous field only holds when A/d is sufficiently large and/or proper shielding is applied. A rule of thumb is that the emitter electrode must overlap the guarding around the receiving electrode with five times the nominal electrode distance d_0, the so-called $5d$-rule [81]. In that case the error made due to fringe effects will be smaller than 1 ppm.

From equation 2.1 follows that there are principally three methods to convert a displacement to a capacitance, as is shown in figure 2.2:

1. Couple the displacement u to a change in ε_r, e.g. move a piece of different-ε_r material between the capacitor plates. Some typical values for ε_r are shown below.

Material	Value of ε_r
Grounded shielding vane	0
Vacuum, air	1
Most electrical insulators	$1 \ldots 10$
Water	80
Piezo ceramics	2000
Electrical conductor	∞

2. Make the displacement u cause a variation in A.

3. Let the displacement u be equal to δd.

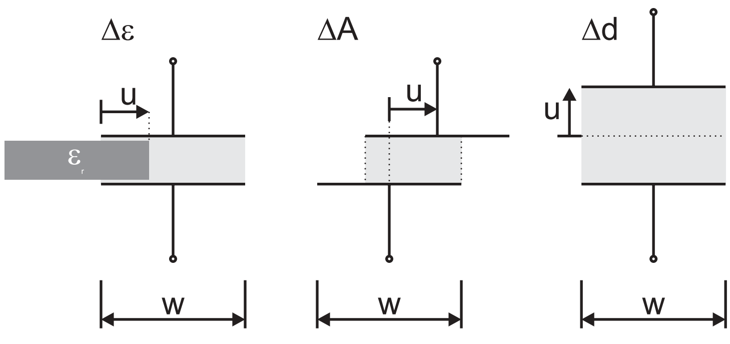


Figure 2.2: According to equation 2.1, there are three principal methods to obtain position information using a general two-plate capacitor.

These three methods each have their own characteristics. The influence of the measuring configuration on the working range, linearity and sensitivity is summarized in table 2.2. Besides these main variables some other parameters have a certain influence on the capacity. These influences are listed in table 2.3 [8].

2.1.1.2 Electronic circuits for capacitance detection

Charge–discharge measurement

One of the most fundamental methods of measuring capacitances is charging a

Variable	Range	Linearity	Sensitivity
ε_r	w	depends on the geometry of the medium material	$\frac{\partial C}{\partial \varepsilon_r} = \frac{\varepsilon_0 A}{d}$ is constant
A	w	depends on the shape of the capacitor plates	$\frac{\partial C}{\partial A} = \frac{\varepsilon_0 \varepsilon_r}{d}$ is constant
d	$\ll w$	hyperbolic instead of linear	$\frac{\partial C}{\partial d} = -\frac{\varepsilon_0 \varepsilon_r A}{d^2}$ depends on d

Table 2.2: The influence of the chosen method (see figure 2.2) on the range, linearity and sensitivity of capacitive position sensors. The variable w is the electrode width.

Parameter	Influence	Remark
Humidity	$\frac{\partial \varepsilon_r}{\partial RH} = 2 \cdot 10^{-6}$ [%RH^{-1}]	Condensed water gives big problems.
Temperature	$\frac{\partial C}{\partial T} = 2 \cdot 10^{-6}$ [F/°C]	Due to $\frac{\partial \varepsilon_r}{\partial T}$, for dry air, pressure 1013.25 mbar.
	$\frac{\partial C}{\partial T} = 2.5 \cdot 10^{-6}$ [F/°C]	For typically moist air: $RH =$ 50%, $P_w = 11.75$ mbar.
Pressure	$\frac{\partial C}{\partial P} = 53 \cdot 10^{-8}$ [F/mbar]	At 0°C.

Table 2.3: Influence of environmental parameters on the capacitance of a parallel two-plate capacitor [8].

capacitor up to a certain voltage, after which it is discharged, e.g. by connecting one electrode to the electrical ground. The amount of charge flowing to the ground is measured by a charge detecting device. From the basic equation, relating the capacitance C to the charge Q and the voltage V,

$$Q = C \cdot V \tag{2.2}$$

it follows that the capacity C is

$$C = \frac{Q}{V} = \frac{1}{V} \int I(t)dt \tag{2.3}$$

which can be measured using an integrating current–voltage converter, which is in fact a charge detector.

However, much attention has to be paid to the quality of the necessary switches within the circuitry, because they will inject disturbing charges into the measuring circuit. Without special precautions (shielding), the resulting high-frequency disturbances even may interfere with other electronic equipment.

Synchronous detection

A better, although more complex, method of capacitance measurement is synchronous detection. In this method, the unknown capacitor C_x and a reference capacitor C_{ref} are subject to two oscillating voltages V_{in} and $-V_{in}$, respectively, as is depicted in figure 2.3. When $C_x = C_{ref}$ charge will flow from C_x to C_{ref} and vice versa, and the input of the synchronous detector will be neutral. At the moment that a capacitance difference $C_x - C_{ref}$ occurs, the detector measures the amplitude of the signal component having the same frequency as the excitation signal V_{in}.

An advantage of the synchronous detection method is that it is differential. When C_{ref} is chosen equal to $C_{x,nom}$ the gain can be optimized, taken into account the expected variations in C_x, in order to achieve maximum sensitivity.

Another advantage is that only the signal having the same frequency and phase as the excitation signal will be detected. This means that this circuit is not sensitive to e.g. 50 Hz interference, provided that the excitation frequency is sufficiently high.

Typical specifications of a synchronous detector are given in table 2.4.

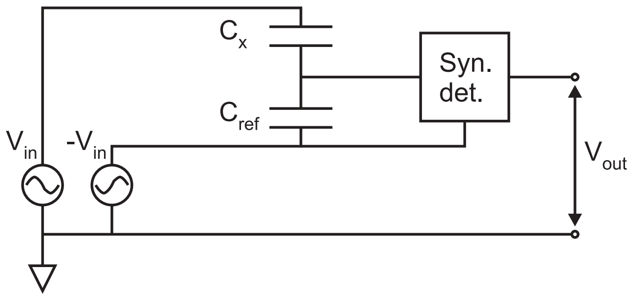


Figure 2.3: Scheme of a differential synchronous detector for capacitance measurement. The detector measures the amplitude of the signal component having the same frequency as the excitation signal V_{in}. The amplitude of the output signal V_{out} will be proportional to the capacitance difference between C_x and C_{ref}.

Parameter	**Value**	**Unit**	**Remark**
Range/resolution	10^3	–	20 kHz bandwidth
	10^6	–	20 Hz bandwidth
Price	5	kf	

Table 2.4: Typical specifications of capacitance measurement systems using synchronous detection.

Martin oscillator

In [70] and [36] a relatively new concept in capacitance detection is presented. A voltage-controlled switched-capacitor relaxation oscillator described first by Martin [42] has been used in a slightly modified form by Toth and Li [13, 70, 36], see figure 2.4. This capacitance measurement circuit is called the Modified Martin Oscillator, or MMO for short.

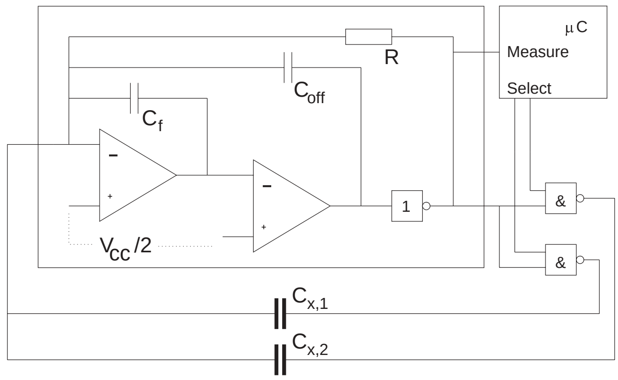


Figure 2.4: Electrical scheme of the Modified Martin Oscillator. The microcontroller (μC, upper right) switches between the unknown capacitors $C_{x,1}$ and $C_{x,2}$ (or any desired number of unknown capacitors). The output of the circuit, which becomes the 'Measure' input into the microcontroller, is a block signal with a period time proportional to the capacitance value. The microcontroller generates a digital number representing the measured capacitance value of $C_{x,i}$.

A microcontroller is used to switch between the unknown capacitors $C_{x,1}$ and $C_{x,2}$ (or any desired number of unknown capacitors). The output of the circuit, which becomes the 'Measure' input into the microcontroller, is a block signal with a period time proportional to the capacitance value, according to

$$T = 4R(C_{off} + C_{x,i}) \tag{2.4}$$

in which C_{off} is the offset capacitance that can not be made equal to zero.

To eliminate the unknown parameters R and C_{off}, three sequential measurements are done (switching on $C_{x,1}$, $C_{x,2}$ and none, respectively) which results in three measured periods $T_{x,1}$, $T_{x,2}$ and T_{off}. Now the *capacitance ratio* between $C_{x,1}$ and $C_{x,2}$ can be determined from

$$\frac{C_{x,1}}{C_{x,2}} = \frac{T_{x,1} - T_{off}}{T_{x,2} - T_{off}} \tag{2.5}$$

If desired, one of the capacitors $C_{x,i}$ can be replaced by a stable reference capacitor C_{ref}. Thus, the *absolute capacitance* of an unknown capacitor can be calculated.

After doing these calculations, the microcontroller generates a digital number representing the measured capacitance value of $C_{x,i}$ or the capacitance ratio between two C_x's, dependent on the programmed algorithm.

Typical specifications of a MMO-based capacitance measurement system are listed in table 2.5.

Parameter	Value	Unit	Remark
Range	$0\ldots25$	pF	
Resolution	100	aF	Range/Resolution $\approx 10^4$
Bandwidth	5	Hz	$t_{meas} = 100$ ms
Non-linearity	100	ppm	2 pF range

Table 2.5: Typical specifications of the Modified Martin Oscillator based capacitance measurement system [8].

A property of the MMO is that the signal on the reference electrode (generally the straightness reference bar) switches from minimum to maximum level—typically ± 5 V or ± 10 V— triggered by the oscillator, where the frequency of the resulting block signal is proportional to the capacitance value that is detected at that moment. This is the reason that only one measurement electrode can be used at the same time. For application in a straightness sensor using one single reference bar and at least five measurement electrodes, this means that multiplexing has to be applied.

2.1.1.3 Parameters influencing the quality of capacitive position detection

When measuring positions by means of capacitive sensors, several parameters which influence the quality of the measurement have to be taken into consideration. These parameters concern the electrodes, the wiring, the environment and the electronic circuitry. In the basic formulas for capacitive sensors (e.g. equations 2.1 and 2.3) these parameters are assumed to be ideal, which can be a too coarse simplification. In figure 2.5 a graphical representation is given of the three groups of error sources and attention points, as a result of literature research and experimental work [15]. The first two groups (electrodes and wiring / environment) can be divided into mechanical and electrical parameters. The parameters will be briefly discussed below, following the numbering as shown in figure 2.5. In [15] a more detailed practical overview is given.

1. Mounting: Alignment errors depend on the sensing mode (Δx, Δz or $\Delta\varepsilon$). The electrode distance error Δz has the largest influence.

2. Motion profile: The parasitic error motions of the moving object (e.g. $\Delta\varphi_x$ while measuring Δz) can cause errors which can be computed if the error motions are known.

3. Dynamical effects: When applied in environments with strong vibrations, the stiffness of the electrode structure and its mounting (e.g. a glue layer) are of importance. In practice, tests showed problems above 1 kHz when using double sided adhesive tape to attach a several cm^2 electrode made of a piece of epoxy-based printed circuit board (PCB).

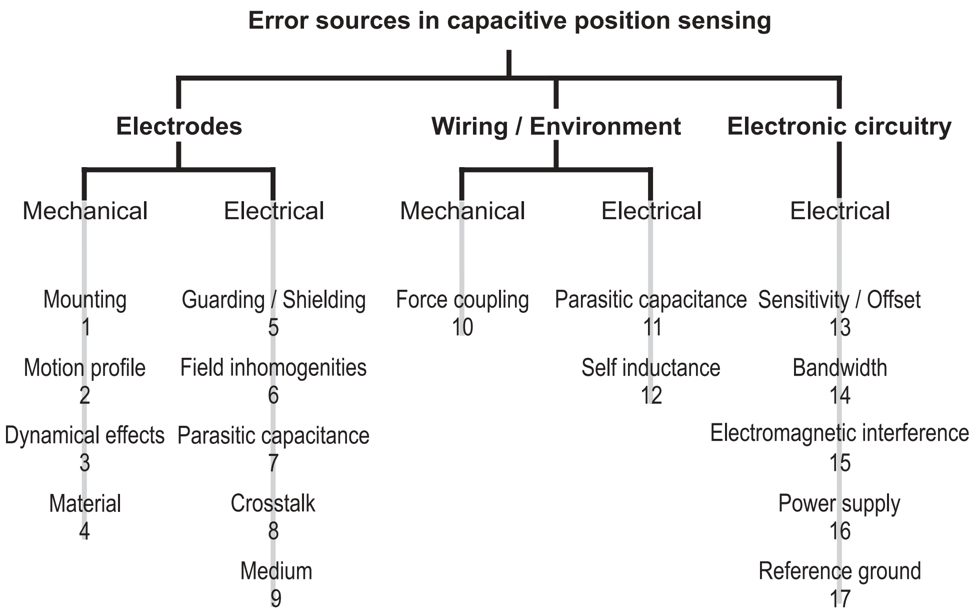


Figure 2.5: Inventory of possible problems and attention points when measuring position using capacitive sensors. Three main error categories can be distinguished: the errors concerning electrodes, wiring and electronics.

4. Material: The influence of manufacturing tolerances can be calculated, typical line width for etched PCB is 200 μm; surface roughness has a negligible influence unless the roughness is near the order of magnitude of the electrode distance; global form errors result in a relatively small error due to averaging.

5. Grounding: The ground circuit is very important, especially when high-frequency electronic circuitry is applied; thick and stranded wires are preferred, ground loops have to be avoided. Shielding is necessary to avoid cross-talk between sensing electrodes and to minimize the influence of environmental electromagnetic disturbances.

6. Field inhomogenities: The basic formulas for the capacitance values assume a uniform field. Therefore, care must be taken to keep the field as uniform as possible by means of a proper guarding around the measuring electrode.

7. Parasitic capacitance: Most measurement circuitry is designed in such a way that it is more or less insensitive to the parasitic capacitance of the connecting cables (see also point 11). However, within the electrode structure also parasitic capacitances can occur, e.g. by unearthed metal objects near the electrodes.

8. Crosstalk: In the case that measuring electrodes lie close to each other, it is

possible that surrounding electrodes disturb the measurement signal of an active measurement electrode. This can be solved by means of a proper guarding around the electrodes, which reduces the capacitance between two measurement electrodes with up to a factor 100.

9. Medium: The capacitance depends on the relative dielectric permittivity ε_r, which is in turn dependent on the air pressure, relative humidity and temperature.

10. Force coupling: Especially in multi-electrode sensors, the wiring can cause substantial forces on the sensor unit. Solutions can be thin wires laid in a goose-neck shaped loop or wireless transmission of measurement signals.

11. Parasitic capacitance: Most measurement circuitry is designed in such a way that it is more or less insensitive to the parasitic capacitance of the connecting cables. However, care must be taken that this parasitic capacitance is not more than 10^2 to 10^3 as large as the nominal capacitance.

12. Self inductance: At high frequencies, the self inductance of the wires normally does not influence the measurement of the capacitance, but it can cause a non-zero potential in the ground circuit, which decreases the measurement accuracy.

13. Sensitivity / offset: The electronic circuitry has to be calibrated because the component tolerances allow for a substantial variation in performance.

14. Bandwidth: Dependent on the type of electronic circuitry applied (see section 2.1.1.2) there is a limit to the frequency of the detectable capacitance variations.

15. Electromagnetic interference: At high (switching) frequencies, care must be taken to ensure that parts of the signal emitting surface do not act as antennas, disturbing other sensitive electronic circuits at large distance.

16. Power supply: Especially in the case of charge / discharge electronics it is important that the power supply has a very low-impedance grounding to reduce the sensitivity to high-frequency noise generated by the charge / discharge switches.

17. Reference ground: It is recommended that the measurement signals have their own reference ground, separated from the power supply ground, to be less subject to high-frequency disturbances.

The research summarized above resulted in five special points of attention during design and assembly of capacitive sensors:

1. Use a guarding system according to the ‘5*d*’-rule. All electrically conducting elements near the electrodes have to be grounded.

2. Keep the cable lengths as short as possible. Prevent vibration of wires.
3. Make the lay-out as symmetrical as possible, to eliminate the sensitivity for parasitic motions.
4. Use a high quality floating power supply with a low-impedance grounding circuit.
5. Avoid the use of electronic circuitry generating (high-frequency, high energy) pulses. Especially in charge / discharge type electronics, care has to be taken with regard to shielding.

2.1.2 Capacitive velocity sensors combined with displacement sensors

2.1.2.1 Introduction

A very straightforward method to obtain motion information from capacitive sensors is detecting a current flowing from one capacitor plate while there is a constant voltage on the other plate. The output current then is proportional to the relative velocity of the plates. When charge rather than current is measured, position will be output instead of velocity, as follows from equation 2.8. This will be explained in more detail in section 2.1.2.2.

The main disadvantage of this measurement method is that its sensitivity decreases with decreasing frequency, so that it is insensitive to static displacements. This is the main reason that this current-detection method hardly is applied in position sensing technology. Its advantages are, as mentioned already, its simplicity and also its high bandwidth and high resolution, which are only limited by the performance of the current detection device.

A promising new idea in the field of sensor technology is the combined use of such a simple velocity sensor and a displacement sensor to obtain accurate displacement information with a high bandwidth. This approach appears to be very fruitful when applied to capacitive sensors because it cancels out the drawbacks of both methods, while optimally combining the advantages [55]. This combination will be described in further detail in this section.

The capacitive sensor based on the Modified Martin Oscillator, as described in section 2.1.1.2, has a good accuracy but a low bandwidth (order of magnitude $1 - 10$ Hz). Actually, the term ‘output rate’ instead of ‘bandwidth’ is more correct in this case since it is a digital system, giving a measurement result at certain intervals.

On the other hand, a high bandwidth is possible with a very simple analog capacitive velocity sensor (order of magnitude 5 Hz − 100 kHz or 0.1 Hz − 1.5 kHz, dependent on the gain). Because of the underlying measurement principle—charge detection—the frequency response of this type of sensor will have a high-pass characteristic. The output of such a velocity sensor may be integrated to give position information.

The combination of these two methods, having a low-pass and high-pass characteristic, results in

- *either* a low-cost position sensor with improved bandwidth (e.g. DC to 1 kHz instead of DC to 10 Hz)
- *or* a low-cost velocity sensor with drift compensation (e.g. DC to 100 kHz instead of 5 Hz to 100 kHz)
- *or* a dual sensor with both a position and a velocity output (e.g. DC to 1 kHz)

The latter option has an additional advantage when using state-space controllers which use the velocity as an input signal. In many cases this kind of control systems simply take the derivative of an available position signal. However, when this position signal contains (quantization) noise, the derived velocity signal will contain substantially more noise than the real velocity. The magnitude of this kind of disturbances for a quantized position signal can be calculated by

$$\delta v = \delta x \cdot f_s \tag{2.6}$$

where δv is the resulting quantization step in the velocity, δx is the quantization step in the position signal and f_s is the sample frequency. This is explained in more detail in section 4.4 on page 117. In general, when the controller requires a velocity signal, it appears to be advantageous to measure velocity directly.

2.1.2.2 Capacitive velocity measurement principle

The measurement principle of the capacitive velocity sensor will be explained here further. The basic formula expressing the relation between the capacitance of a two-plate capacitor and its geometrical properties was given in equation 2.1:

$$C = \frac{\varepsilon_0 \varepsilon_r \cdot A}{d(t)}$$

in the case that ε and A are constant and d is a time-varying distance, under the assumption that the electric field is homogeneous. Together with

$$Q = C \cdot V \tag{2.7}$$

and

$$I = \frac{\partial Q}{\partial t} \tag{2.8}$$

this gives for the current I from the second plate (supposing a constant voltage difference V between the capacitor plates):

$$I = \frac{\partial}{\partial t}\frac{\varepsilon_0\varepsilon_r AV}{d(t)} = -\frac{\varepsilon_0\varepsilon_r AV}{d^2(t)} \cdot v(t) \tag{2.9}$$

in which $v(t) = \frac{\partial}{\partial t}d(t)$ is the motion velocity. Equation 2.9 implies that the output current of a velocity sensor based on this principle will also be dependent on the unknown squared distance $d^2(t)$. When the ratio between the amplitude u of the motion and the nominal distance d_0 is sufficiently small, this influence will be small, as will be shown hereafter.

First, assume a sinusoidal time-dependent displacement of the measurement electrode

$$x(t) = d_0 + u\sin(\omega t) \tag{2.10}$$

with u the displacement amplitude, so that the velocity will be

$$v(t) = u\omega\cos(\omega t) \tag{2.11}$$

which gives for the current I according to equation 2.9

$$I(t) = \frac{\varepsilon_0\varepsilon_r AVu\omega\cos(\omega t)}{(d_0 + u\sin(\omega t))^2} \tag{2.12}$$

Second, define a hypothetical current I' without the influence of the position dependance:

$$I'(t) = \frac{\varepsilon_0\varepsilon_r AVu\omega\cos(\omega t)}{d_0^2} \tag{2.13}$$

so that the relative difference between the actual and the ideal current becomes

$$\frac{I(t)}{I'(t)} = \frac{d_0^2}{(d_0 + u\sin(\omega t))^2} \tag{2.14}$$

Third, define the quotient

$$f = \frac{u}{d_0} \tag{2.15}$$

to express the motion amplitude as a fraction of the nominal working distance. Substituting this in equation 2.14 gives

$$\frac{I(t)}{I'(t)} = \frac{d_0^2}{(d_0 + \sin(\omega t) f d_0)^2} \tag{2.16}$$

or, expanded and simplified

$$\frac{I(t)}{I'(t)} = \frac{1}{1 + 2f\sin(\omega t) + f^2\sin^2(\omega t)} \tag{2.17}$$

from which it is clear that the worst case occurs when

$$\omega t = \frac{\pi}{2} + 2k\pi, \quad k \in \mathbb{N} \tag{2.18}$$

resulting in

$$\frac{I(t)}{I'(t)} = \frac{1}{1 + 2f + f^2} \tag{2.19}$$

In practice, the values for f will be small enough to neglect the quadratic term with respect to $(1 + 2f)$, so that the simplified expression for f becomes

$$\frac{I(t)}{I'(t)} = \frac{1}{1 + 2f} \tag{2.20}$$

The relative error e is equal to

$$e = 1 - \frac{I(t)}{I'(t)} = \frac{2f}{1 + 2f} \tag{2.21}$$

which simplifies even further, for small values of f, to

$$e \approx 2f \tag{2.22}$$

This means that e.g. for a ratio $u/d_0 = 0.001$ the maximum relative error is about 0.2%.

Note that this error might be avoided by choosing measurement method 1 or 2 ($\Delta\varepsilon_r$ or ΔA, see the list on page 14) instead of the method described above, which uses Δd as an input variable. The latter, however, in general leads to a higher sensitivity, especially when the nominal electrode distance d_0 is small compared to the electrode width w.

2.1.2.3 The combination of position and velocity sensing

The combination of the position sensor and the velocity sensor—or xv-sensor[4] for short—enables us to use the high bandwidth of the velocity sensor (integrated towards positions) while the static accuracy is provided by the slow position sensing part. In daily life, this is more or less analogous to two-way loudspeakers, where the frequency range in which a certain sound quality is required, has been split up so that each speaker unit provides a certain part of it.

An illustration of this principle as applied in capacitive sensors is given in figure 2.6, where some simulated signals are shown. First, the real position (top, left)

[4]The name xv-sensor is chosen merely because of the two different underlying measurement principles, which are based on x and v. In practice, when the v-measuring part applies a charge amplifier instead of a current–to–voltage converter, v is integrated towards an additional x. The name 'xv-sensor' does not intend to suggest that also a velocity signal necessarily is present.

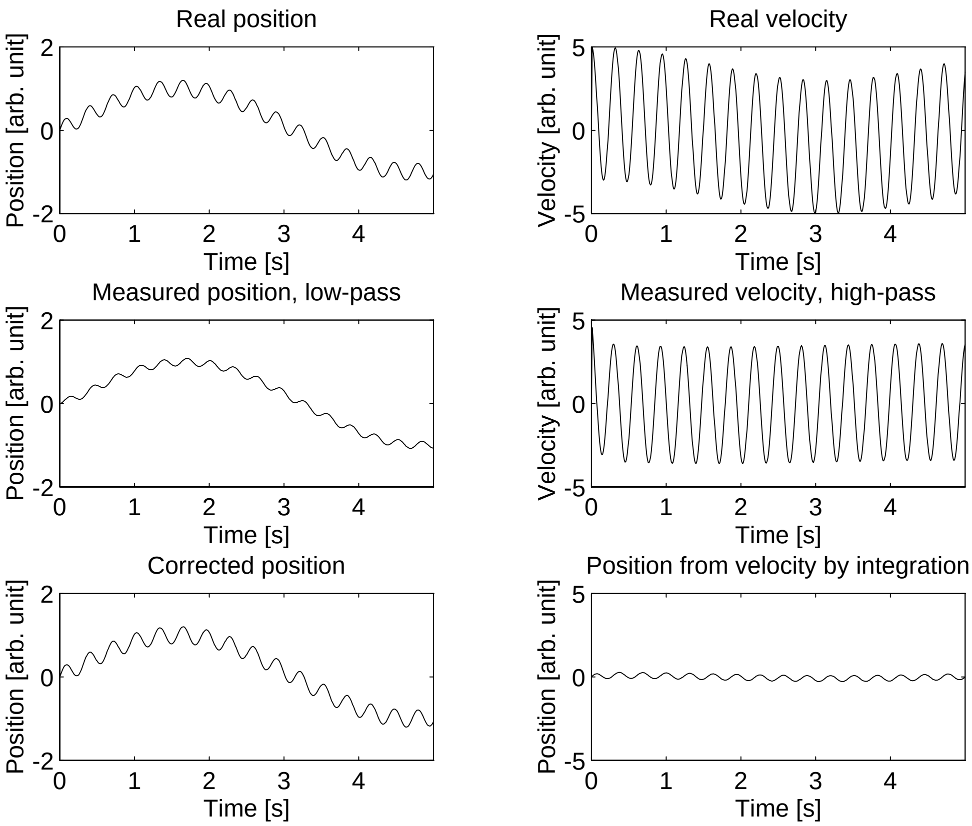


Figure 2.6: Simulated example of the combination of a position sensor having a low-pass characteristic with a velocity sensor having a high-pass characteristic. The real position signal has 1 and 20 rad/s components, the low- and high-pass characteristics have been modeled as first order Butterworth filters with 10 rad/s cross-over frequency. The lower left picture is the sum of the measured (low-pass) position and the integrated (high-pass) velocity signal.

A question that arises is why this approach is advantageous, since a direct high-bandwidth position measurement is more straightforward and also will extract the position information from 'the moving electrons' in a more pure way, without possible cross-over irregularities.

The answer to this question is that indeed theoretically the moving charge contains all information about relative movement of the electrodes, given the initial position as a boundary condition. However, in practice it is impossible to measure charge or current without loss or noise. This results in a decreasing accuracy at decreasing frequency, because the measured current is proportional to the velocity, and thus it becomes impossible to detect static positions. In the past, many methods have been developed to overcome this problem, resulting in a variety of capacitance detecting circuits as described in section 2.1.1.2, making use of various methods like modulation, oscillation, etc., to extract static position information from the two electrodes. The development process of each of these circuits has had its own design goals, for example maximum accuracy or maximum bandwidth. There is always a compromise between bandwidth, accuracy and cost.

The underlying idea for the combined system described in this section is that the combination of two low-cost units (for the measuring of capacitance) has advantages above using only one expensive method.

Several problems have to be solved to make this combination. First the way to add the two signals in the frequency domain has to be investigated. The *analog* method first integrates the velocity signal resulting in a position signal with a high-pass character, or directly uses the output of a charge amplifier. By using a frequency-dependent weighting function this can be added to the low-pass position information. When, as in figure 2.6, both 'sensors' have unity gain and their Bode plots are complementary, this weighting function can be a unity gain as well. In practice, as this will generally not be the case, some attention has to be paid to the cross-over region weighting function.

A *digital* method might be more flexible. In a digital environment filters, delays and arithmetic operations are easily adjustable. One special case is worth mentioning: a digital position sensor that gives one measurement every time interval t_{meas}, combined with an analog velocity sensor. In that case the integrated velocity can be used as position information, causing a certain drift because of the high-pass behavior of the velocity sensor. At certain intervals now the signal of the slow position sensor can be used to eliminate that drift.

2.1.2.4 Experimental results of the combined position-velocity sensor

Experiments with a combined analog position-velocity sensor, also called *xv*-sensor for short, show that excellent results can be achieved with relatively simple means. For these experiments, two electrode plates have been designed. The first plate contains the 'receiving' electrodes, which have a diameter of 15 mm and are surrounded by a grounded shielding surface. The second plate contains the 'sending' electrodes, which have to be larger than the receiving electrodes. The plates are placed parallel to each other, at a nominal working distance of 250 μm. One pair of sending / receiving electrodes is intended to be connected to a Modified Martin Oscillator based measurement circuit. The other pair is connected to a constant voltage source (50 V) and a current-to-voltage converter.

Furthermore, a test set-up has been built in order to move one of the two parallel electrode plates in the direction perpendicular to the electrode surface, with a mechanical bandwidth of about 1 kHz.

A Bode plot of the combined sensor is displayed in figure 2.7, the measured specifications of the velocity sensing part and the total combined sensor are given in table 2.6 and 2.7, respectively. As a displacement sensor an eddy current sensor (see section 2.1) has been used, low-pass filtered at 1 Hz. The reason that this (analog) sensor has been used instead of the (digital) MMO based capacitive sensor is that there were practical problems with regard to the implementation of this sensor readout in the dSpace test environment. The *xv*-sensor's signals have been compared with the signal from a photonic sensor (see also section 2.1), which serves as a reference sensor.

Parameter	**Value**	**Unit**
Frequency range	1 – 1500	Hz
Resolution	50	μm/s
Sensitivity	0.12	mV/(μm/s)

Table 2.6: Measured specifications of the xv-sensor prototype's velocity sensor part having one electrode with a constant voltage (50 V) and another electrode connected to a current-to-voltage converter.

Parameter	**Value**	**Unit**	**Remark**
Range	$\pm$500	μm	Static
	$\pm$5	μm	Dynamic
Noise level	0.2	μm	
Accuracy	< 0.5	μm	Rel. to photonic sensor
Frequency range	DC . . . 1500	Hz	

Table 2.7: Measured specifications of the combined xv-sensor prototype.

In figure 2.8 a comparison is shown of the integrated velocity output signal of the capacitive velocity sensor (giving a position signal having a slow drift) and the position output signal of the photonic sensor. The input is a 100 Hz mechanical vibration with a small amplitude of about 2 μm. The phase difference is caused by the low-pass filtering of the photonic sensor's output. Without this filtering the reference sensor's signal contained even more noise.

The maximum absolute difference between the two signals is about 0.2 μm. This

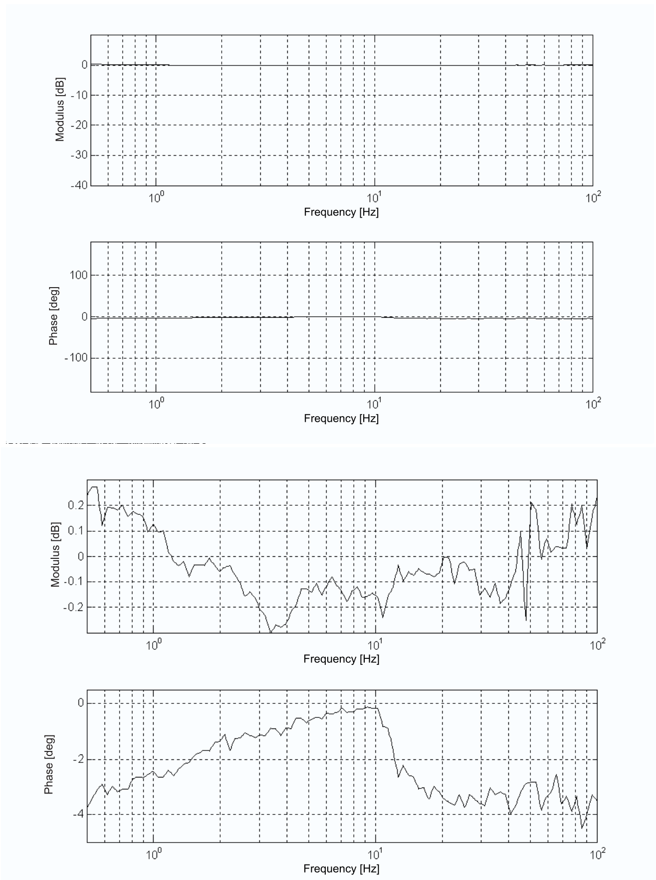


Figure 2.7: Bode diagrams (above: normal scale; below: expanded scale) of the combined capacitive position–velocity sensor. Displayed is the transfer function from the reference sensor (a fiber optic sensor, see section 2.1) to the combined capacitive position–velocity sensor. Ideally this should give a flat response of 0 *dB and* 0° *over the whole frequency range. The maximum deviation observed is as low as* 0.3 *dB over the whole frequency range of interest.*

comparison is done using the moving average of the photonic sensor output in order to eliminate the noise component.

It is clear that the capacitive sensor's signal is preferred, in particular for control purposes, where the high-frequency noise can be a severe limitation for the system performance.

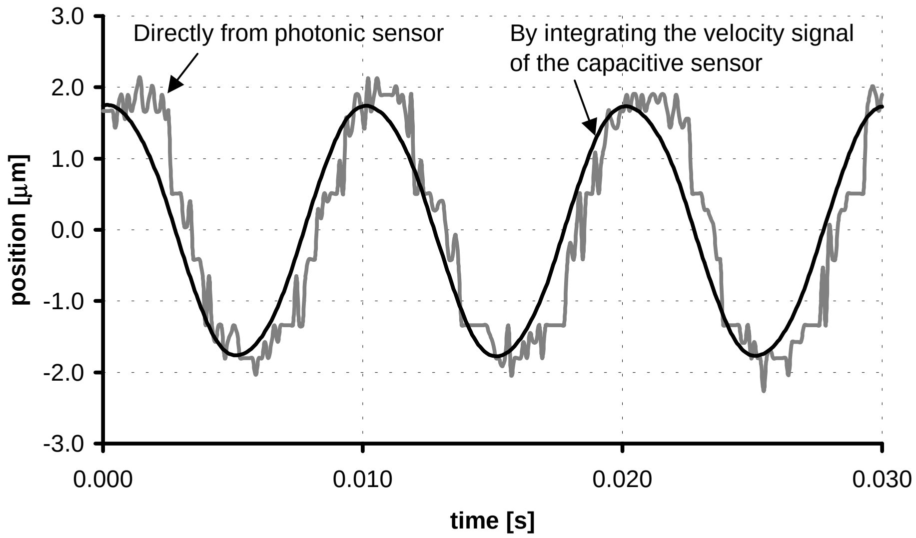


Figure 2.8: Graph showing the integrated velocity signal of the capacitive velocity sensor and the position signal of the photonic sensor for a small 100 *Hz vibration input. The maximum absolute difference between the two signals is* 0.2 μ*m, after low-pass filtering the photonic reference sensor output to remove most of the noise. Note that there is a phase difference, caused by this necessary filtering of the photonic reference sensor.*

In figure 2.9 the position signal of the combined *xv*-sensor and a photonic sensor are compared in the low frequency range. The input signal for both sensors is a small position fluctuation containing two different frequencies, of which the highest one just lies in the critical overlap zone of the *x*- and the *v*-part. The maximum difference appears to be smaller than 0.5 μm. Note the noise on the photonic sensor's signal. This noise probably is caused by the influence of ambient light and irregular reflection in combination with the finite number of glass fibers. The reason for this assumption is that in the near slope mode at a very short distance to the reference object, where the sensitivity is higher at equal output levels, the noise is considerably decreased.

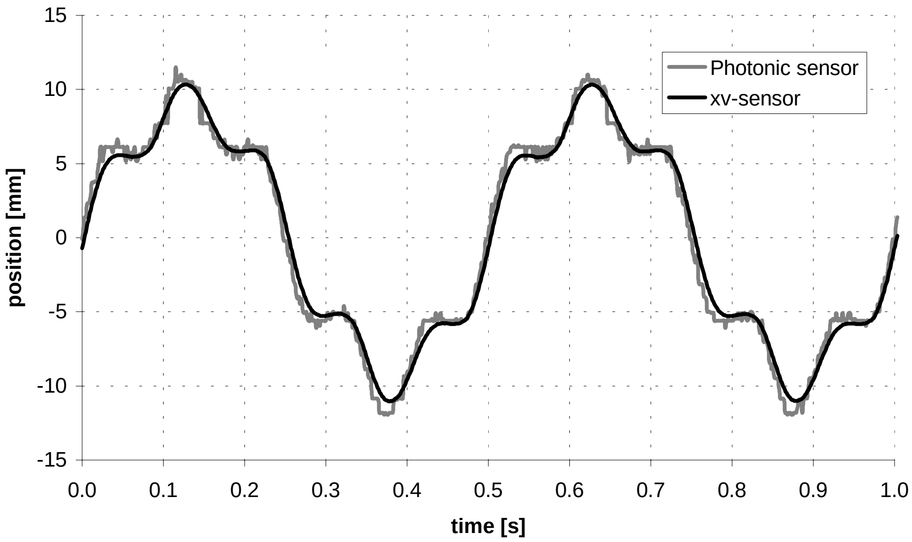


Figure 2.9: Comparison of the position signal of the xv-sensor and of a photonic sensor for a small signal containing two different frequencies, of which the highest one just lies in the critical overlap zone of the x- and the v-part. The maximum difference is smaller than $0.5\ \mu m$. *Note the noise on the photonic sensor's signal, having an amplitude of* $\pm 0.5\ \mu m$.

2.1.2.5 Applicability of the xv-sensor in straightness measurement

In practice, the x- and the v-part of a xv-sensor will have their own detector electrodes. It is preferable to use a single straightness reference as a sending electrode for both types of detector electrodes, to prevent complicated alignment procedures. However, the v-sensor requires a constant potential on the sending electrode whereas the x-sensor based on the Modified Martin Oscillator (MMO) applies a block signal on the sending electrode, having a frequency proportional to the capacitance value. This frequency typically is in the order of 50 kHz.

Several possibilities are proposed to realize a 5-DoF straightness sensor using five position (x-) sensing elements and five velocity (v-) sensing elements, while avoiding unwanted interference between the required reference signals:

1. Use of two parallel straightness reference bars, electrically isolated from each other.

2. Use of the same straightness reference bar both for x- and v-measurement; Transformation of the switching levels of the MMO from $-10\ldots10$ V to

$0 \ldots 20$ V; the high switching frequency, typically 50 kHz, can be filtered out for the low-bandwidth x-sensor giving a DC potential of 10 V.

3. Replace the capacitive x-measurement principle by another method, e.g. optical, which is independent on the electrical potential of the reference electrode's surface.

When using method number 1, much attention has to be paid to the alignment of the sensor elements and the two straightness reference bars. Method number 2 is the most experimental of the three. Cross-talk and interference may be limiting factors for successful application. Method number 3 is trivial and will not cause problems. The three methods mentioned above are not investigated further experimentally.

2.1.3 Concluding remarks concerning displacement sensors

This section will be finished giving a short overview of highlights concerning position sensors. Table 2.8 shows several possible design goals, together with a suggestion for a 'best' sensor choice.

Most important parameter	**Preferred sensor**	**Section**
Maximum resolution	Capacitive	2.1.1
Minimum cost	IC photonic	2.1
Minimum size	IC photonic	2.1
Maximum bandwidth	Photonic, inductive	2.1

Table 2.8: Overview of different design goals leading to a preferred sensor type.

The most interesting trend in sensor technology is the increasing application of IC fabrication techniques. This enables a substantial miniaturizing, together with the additional advantage that all signal conditioning circuitry can be integrated in the sensor package as well. Thus, the sensitivity for electromagnetic interference is greatly reduced.

2.2 Straightness sensors

In this section, attention will be paid to sensor technologies suitable for straightness error measurement, i.e. measuring position deviations from a straightness reference in five degrees of freedom.

Up to now, no commercial 5-DoF straightness sensors are available, except for one system by Applied Precision Inc., based on laser interferometry. Its measurement resolution is 0.2 ppm for the travel direction (x), sub-micrometer for the lateral directions (y, z) and sub- μrad for the three angles ($\varphi_{x,y,z}$).

Obviously, any type of sensor capable of measuring a distance relative to a reference surface can be used in such a set-up that five measured distances give the information about the desired five degrees of freedom, possibly after a coordinate transformation. In that way, any of the sensors described in section 2.1 could be applied. Here only one capacitive sensor and three optical straightness sensors will be described, because of their small volume and / or special configurations.

2.2.1 Capacitive straightness sensors

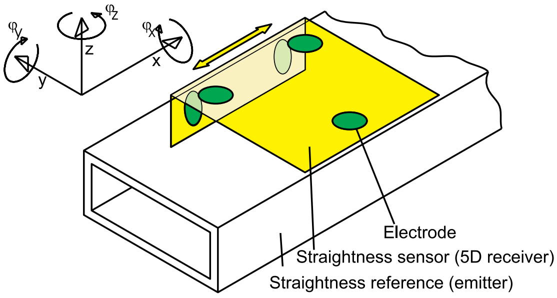
In the last years, extensive research has been done on the applicability of capacitance measurement in the field of positioning [24, 28]. On the one hand, new and / or improved electronic circuitry has been developed [7, 23, 31, 41, 69, 70, 36]. On the other hand, different configurations have been designed to measure straightness with one compact sensor unit [81, 8, 66, 50].

In this section, a capacitive position sensor is described that can measure all 5 degrees of freedom needed for straightness measurement simultaneously relative to a straightness reference, using a very simple sensor structure [50].

In [66] the concept of a capacitive straightness sensor has been presented. Based on that knowledge an L-shaped sensor unit has been designed. Figure 2.10 shows a schematical representation. Each of the five circular electrodes is used to detect the local distance between the sensor surface relative to the straightness reference. From these measurements, the actual straightness errors have been calculated using a coordinate transformation, which depends on the geometry of the sensor configuration.

The main advantage of this kind of capacitive sensors is their design flexibility. It is relatively simple to realize an electrode structure tailored to the requirements.

The electrode patterns were etched in double sided printed circuit board and, together with the analog part of the electronics (see text below), mounted in a small box of approximately $80 \times 60 \times 40$ mm^3. This has the advantage that only switching signals and the pulse width modulated output signal have to be transmitted by the long connecting cable, which results in a lower noise level than in the case that analog signals have to be transported through a long wire. The sensor unit has



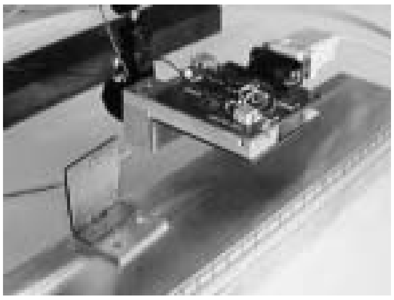


Figure 2.10: Left: Lay-out of the electrodes in the capacitive straightness sensor. Each of the electrodes measures the distance to the reference surface, an aluminum straightedge. A coordinate transform of the output signal vector gives the actual coordinates. Right: Two prototypes of capacitive straightness sensors; the largest one has all necessary measurement electronics integrated in the sensor unit (for this photo, the cover has been removed).

been mounted on a linear stage with about 300 mm travel and straightness in the order of 1 to 10 μm.

As a straightness reference, a hollow aluminum bar with rectangular cross section has been used. Its absolute straightness is approximately 30 μm; its roughness is in the micrometer range and does not have a measurable effect on the sensor accuracy because of the spatial averaging effect. Short-term form stability is estimated to be in the sub-micrometer range, due to the good thermal characteristics of aluminum. The nominal distance between the sensor electrodes and the straightness reference is 0.5 mm.

For detecting the actual capacitances, various electronic principles are known, as summarized in section 2.1.1.2. In this first prototype the Modified Martin Oscillator has been applied because of its ease of use, low cost and good accuracy. A dedicated chip set was developed earlier [70, 36], consisting of two ASICs[5]: the so-called UTI (Universal Transducer Interface) and a dedicated multiplexer chip called MUX. The output of this system is digital; the capacitance value acts on an oscillator signal in such a way that a block signal is generated of which the period is proportional to the capacitance value. This block signal is fed into a microcontroller that detects this period and sends the numerical capacitance values to a computer.

The electronic circuit can be split into two parts: the analog part with the oscillator and switches can be integrated in the sensor unit and properly shielded, while the microcontroller and communication part can be located elsewhere, because that part is much less sensitive to electromagnetic disturbances.

The low-cost capacitance detection circuit, as described above, has a low measurement frequency, due to the rather slow microcontroller and RS-232 serial communi-

[5]ASIC: Application Specific Integrated Circuit

cation, but it is fast enough for calibration purposes. For feedback purposes a faster measurement system is under development, with an aimed bandwidth in the order of 1 kHz.

The highest resolution is achieved around a measurement time $t_{meas} = 100$ ms. For shorter times the $1/f$ noise increases, for longer times the drift increases. Because of the low measurement speed of the UTI this 5-DoF straightness sensor has been used in quasi-static applications only. Its measured specifications are listed in table 2.9.

Parameter	**Value**	**Unit**	**Remark**
Nominal gap	$0.25 \ldots 4$	mm	$\approx 3 \ldots 0.2$ pF
Resolution	0.25	μm	y,z, at 0.5 mm gap
	7.5	mrad	φ_x,φ_y,φ_z, at 0.5 mm gap
Noise and drift	$<$ resolution		during 20 minutes
Non-linearity	0.28	%	Full range
Measurement frequency	1	Hz	5-DoF
	5	Hz	1-DoF

Table 2.9: Measured specifications of a capacitive straightness sensor which detects five degrees of freedom simultaneously.

The following additional remarks have to be made:

1. The large measurement range (> 1 mm for each of the electrodes) does not require very precise alignment of the sensor unit, relative to the straightness reference and the stage. However, the absolute non-linearity error will increase when the nominal working distance increases, so that the nominal gap preferably is as small as possible. In practice, this will result in an air gap in the same order as the desired working range.

2. It is possible to increase the measurement frequency at the cost of lower resolution or less degrees of freedom, when the circuitry, as described above, is applied. To avoid this compromise, currently new versions of the measurement circuitry are under development at the TU Delft's Electronics department. The goal is to realize a measurement frequency in the kHz range, the same resolution and the same number of degrees of freedom.

3. The measured straightness (actually: position) values are always relative to a straightness reference. As long as form stability of the straightness reference is guaranteed to be below a certain value ϵ, absolute repeatability is possible down to the order of magnitude of ϵ, *or* the resolution of the sensor, whichever is larger.

2.2.2 Optical straightness sensors

2.2.2.1 Three-beam laser straightness sensor

The generation of a laser beam seems to be a simple way to realize a straightness reference, because it describes a 'straight line' that can be directed along any desired trajectory. Along the length of this trajectory the position of a moving object relative to this beam can be detected using, for example, photodiodes. However, some uncertainties affecting the straightness, described further in section 2.3.4, have to be taken into account.

In [45] and [30], a straightness sensor has been described based on the use of three parallel laser beams. A schematic representation of the measurement principle and geometry is given in figure 2.11.

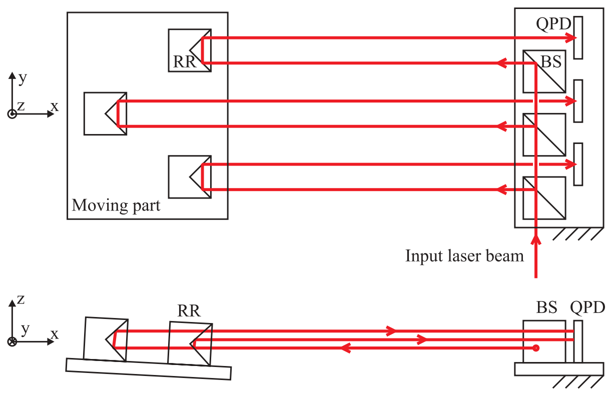


Figure 2.11: A straightness sensor based on three parallel laser beams. The three parallel beams are derived from one incoming beam by means of three beam splitters (BS). Then the three beams are reflected by three retroreflectors (RR) located on the moving target, so that they reach the three quadrant cell photodiodes (QPD). Due to a dedicated triangular placement of the three retroreflectors, straightness deviations of the moving part can be obtained in five degrees of freedom, all except the x-direction [30]. As an example, in the lower figure the laser beams have been sketched for a rotation around the y axis.

For the mounting of the optical components a special method has been developed (patented by EPFL, Lausanne), which enables flexible aligning and attaching by means of laser spot welding. In this method, the optical component is attached to a tripod, which is in turn placed on the mounting surface. Then the component is aligned by elastically deforming the tripod. When the position is adjusted (which is detected by an optical measurement device), its position is 'frozen' by laser spot welding the three feet of the tripod to the mounting surface.

The main drawback of this method is the parallellity of the three beams after they have been redirected by the beam splitters, which directly influences the accuracy of the sensor. Moreover, the internal stresses increase the sensitivity of the tripod's position to thermal disturbances.

Misalignment of the retroreflectors or the quadrant cell photo diodes only causes offsets or second-order effects.

The (small) divergention of the laser beam has no influence because the quadrant cell photo diodes determine the y- and z-position of the point where the beam has its mean intensity.

Using this three-beam laser straightness sensor, a resolution for deviations in the y- and z-direction in the order of 0.1 μm has been achieved in a motion range of 600 mm. The resolution in the angular directions

2.2.2.2 Single-beam laser straightness sensor

An alternative method presented here avoids the alignment of three beams by splitting the beams at the moving target rather than at the origin of the beam. This way, the parallellity is limited only to the shape tolerances rather than the alignment tolerances of the optical components. High accuracy can be achieved, especially by machining the appropriate parts together in one machining cycle. Moreover, the necessary optical components are considerably less expensive, because instead of three retroreflectors only one prism is needed, as is shown in the table below.

Number of required optical components		
Item	3–beam	1–beam
Laser	1	1
Beam splitter	3	3
Prism	0	1
Quadrant cell Photo Diode	3	3
Retroreflector	3	0

In figure 2.12 a possible geometrical solution has been sketched. The three returning beams together with the quadrant cell photodiodes give six signals for five degrees of freedom position information, so that one signal is redundant and can be used for verification purposes.

A possible disadvantage of this configuration is that the angle measurement depends on the distance x between the moving part and the fixed part. For a sensor of this type in an application requiring a certain angular range, a correction will be necessary for the distance x. For zero point sensors it only means that the angular sensitivity depends on x, which may be corrected as well.

A sensor of this type has been used with good results in a six degrees of freedom piezo manipulator, as is described in section 4.3.

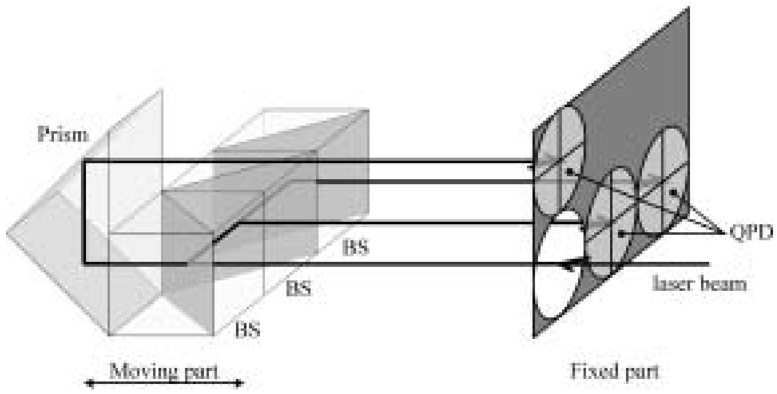


Figure 2.12: Schematic representation of a 5-DoF straightness sensor using only one laser beam without the need for careful alignment.

2.2.2.3 Photonic straightness sensor

The integrated photonic sensor as described in section 2.1 is very suitable for use in a multi-DoF sensor, e.g. a sensor analog to the capacitive 5-DoF sensor that was described in figure 2.10. Instead of the capacitive electrodes five photonic ICs can be applied.

No special precautions are necessary with regard to the electrical shielding and cabling, because all electronic circuitry is integrated in a shielded housing, together with the sensing elements. For the same reason the connection to the controller is very compact and straightforward: the only inputs needed are a 5 V power supply and a ground signal, the outputs are five measurement signals in the range of 0 . . . 5 V that can be fed directly into the controller. The total component cost will be in the order of 10 . . . 100 guilders.

2.2.3 Concluding remarks concerning straightness sensors

1. Up to now, no *commercial* 5-DoF straightness sensors are available, except for one system by Applied Precision Inc., based on laser interferometry.

2. Some *experimental* 5-DoF straightness sensors exist, but they are still in the prototype phase.

3. Most of the experimental 5-DoF straightness sensors consist of combinations of 1-DoF sensors.

4. Capacitive and optical sensors offer a large flexibility in realizing multi-DoF position sensors.

2.3 Straightness references

Throughout this thesis, the term 'straightness reference' will be used for any object that can be used to relate straightness measurements to. In practice, straightness references will be objects having suitable straightness tolerances. Of course, to obtain absolute straightness information, it is necessary to know the shape of the straightness reference with sufficient accuracy. That means, it must be traceable to an (international) standard or to a perfect geometrical shape.

Assuming a perfect straightness *sensor*, different situations can occur, dependent on the available information about the straightness reference. This is indicated in table 2.10.

Straightness reference quality	Type of information obtained
Absolute straightness	Absolute accuracy
Form stable, known shape	Relative accuracy, abs. repeatability
Form stable, unknown shape	Absolute repeatability
Not form stable, known initial shape	Short-term relative accuracy
Not form stable, unknown shape	Short-term repeatability

Table 2.10: Overview of the information that a perfect sensor can obtain, given a straightness reference having the properties as listed in the left column.

Straightness references can be divided in two categories: mechanical (including straightedges, taut wires and liquid surfaces) and optical. The fundamental difference is that mechanical references consist of particles (atoms) that are arranged in a more or less known way, while optical references make use of 'free' particles (photons) that are travelling in a more or less well-defined way. In the sections below, each of these straightness references will be described in more detail.

2.3.1 Mechanical straightedges

The simplest way to create a mechanical straightness reference is to take a rectangular bar and to support that in a stable and well-defined way. In principle, any bar can be used, as long as it has at least one 'flat' surface and one 'straight' edge. In practice, it is preferred to have a prismatic bar with a certain guaranteed straightness tolerance. Commercially available rectangular bars having a specified straightness tolerance are called straightedges. They can be bought in several accuracy classes, for example according to the DIN norm table on page 4, or be custom-made in any precision mechanical or optical workshop. Straightedges are widely used in the 'classical' measurement practice, and they can also serve as a target surface for various types of distance sensors.

To achieve the highest possible accuracy one might at first instance consider a kinematic support with knife edges and / or point contacts. However, sometimes a kinematically overdetermined support or clamp may be more suitable. For example, in the case of varying thermal or mechanical loading of the straightness reference bar more contact points are desired to improve the contact stiffness, because the stiffness of a kinematical ‘point contact’ is relatively low. Or in the case that high stability is required, a support consisting of a row of knife edges can be applied, which are individually adjustable to create a stable support that results in the best possible flatness of the supported bar. In most cases, however, a kinematic support will be used.

In the next paragraphs we will consider the most important aspects with regard to straightness references, in order to achieve straightness down to the nanometer level.

2.3.1.1 Location of support points

First, consider the location of the support points of a mechanical straightness reference bar. They determine the shape as deflected due to the gravitational force. Four classical situations are of importance [4]:

1. Airy point support: $x \approx 0.2113 \cdot l$; provides parallel end surfaces, used for long end gauges and line scales with the pattern in the top plane
2. Bessel point support: $x \approx 0.2203 \cdot l$; results in minimum shortening in the horizontal direction, used for line scales with the pattern in the neutral plane
3. Minimum bending: $x \approx 0.2232 \cdot l$; overall bending is minimal, equal deflection for middle point and end points
4. zero deflection in the middle point: $x \approx 0.2386 \cdot l$; bending between the support points is minimal

in which x is the distance from the end of the bar to the support point and l is the total length.

The point supports mentioned above are only valid for prismatic bars. They are calculated based on the simple beam theory, which can result in considerable deviations from the real shape.

By using the *plane stress theory* instead of the simple beam theory, more accurate predictions can be made of the deflections of the bar and the optimum location of the support points, because also the external dimensions of the bar are taken into account [59]. In that case the displacements of the *top surface* of the bar are calculated, instead of the displacements of the neutral line. A prismatic bar, having support locations calculated according to the plane stress theory, can have about 35% smaller peak–to–peak gravitational deflection than the same bar, having support locations calculated according to the simple beam theory. This is illustrated hereafter.

Consider a rectangular, prismatic bar, supported by two knife edges [59]. Its dimensions are: length $l = 685.5$ mm, width $t = 67.5$ mm and height $h = 137.7$ mm. The support lines are each on a distance $b \cdot l$ from the ends.

In figure 2.13 a comparison is shown between the deflection curves according to the *simple beam theory* (dashed lines) and according to the the model using also the *plane stress theory* (solid lines). Firstly, choosing the case of minimum bending, the value $b = 0.2232$ has been applied (upper figure). The expected beam shape for this value of b has been plotted for each of the two theories. While, based on the simple beam theory, a peak–to–peak straightness error of about 38 nm is expected, it appears to be about 48 nm, according to the more accurate plane stress theory. Secondly, for the same case, the value $b = 0.2285$ has been applied (lower figure). This value is the optimum value for minimum bending, according to the plane stress theory. It appears that the peak–to–peak straightness error is about 30 nm. This means that, by applying the augmented theory, an improvement of the peak–to–peak straightness error of $\frac{48-30}{48} \times 100\% = 37.5\%$ can be achieved. The support points are shifted towards the center of the bar by $(0.2285 - 0.2232) \cdot 685.5\,\text{mm} = 3.63\,\text{mm}$ or 0.53%.

It should be noted here that the above values concern cross sections in the length direction of the bar only. The gravitational sag $z(x)$ is not dependent on the width t of the bar, in these equations. When three-dimensional predictions are required, analytical calculations will become so laborious that finite element analysis will be the most preferable alternative.

Attempts also have been reported [17] to reduce the gravitational deflection by immersing the straightness reference bar in a high-density liquid, $C_2H_2Br_4$. However, the hydrostatic pressure appeared to induce deformations in the same order as the deformations due to gravity. Moreover, there were disturbing effects because of fluid dynamics.

2.3.1.2 Hertz contact stresses at the supporting points

Because the contact situation in the support points of the straightness reference determines the stability and repeatability of the straightness reference's position and orientation, it is important to know about elastic or plastic deformations in the contact zones.

According to the well known Hertz theory, as summarized in e.g. [32], for the contact of a ball (index 1) on a flat surface (index 2) the radius r of the contact area is given as

$$r = \sqrt[3]{3F\frac{R}{E}} \tag{2.23}$$

with

$$\frac{1}{E} = \frac{1-\nu_1^2}{2E_1} + \frac{1-\nu_2^2}{2E_2} \tag{2.24}$$

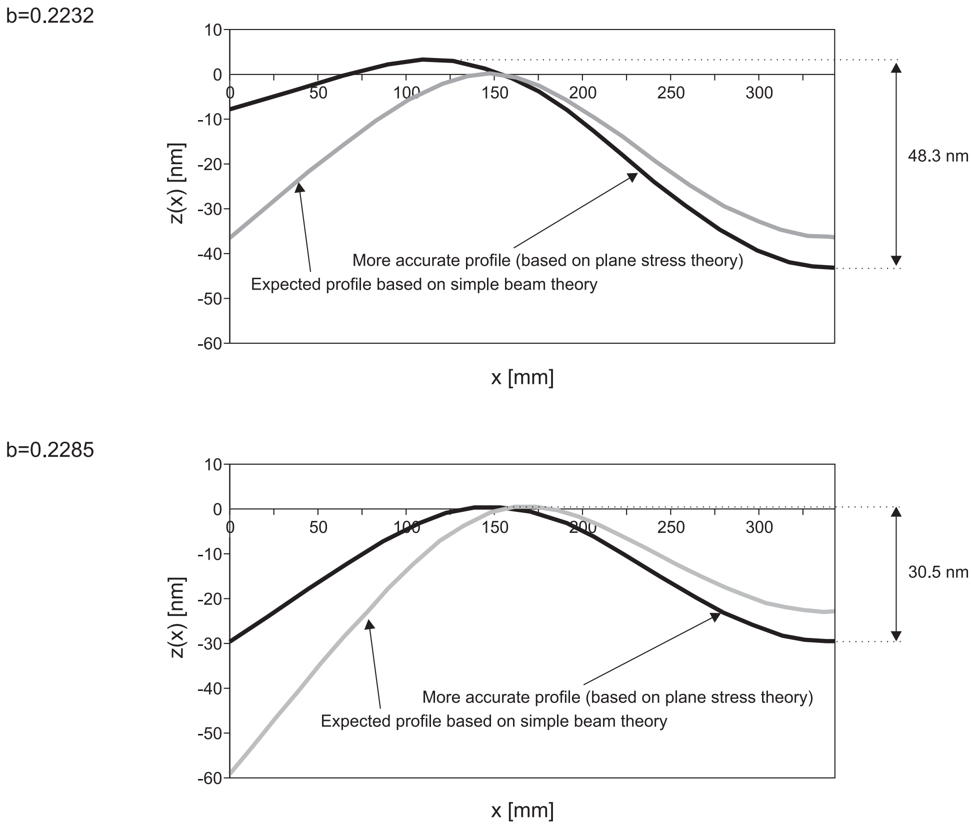


Figure 2.13: A comparison between the deflection curves according to the simple beam theory (dashed lines) and the model using also the shear theory and the plane stress theory (solid lines). A bar with length $l = 685.8$ mm, height $h = 137.7$ mm and width $t = 67.5$ mm is supported on two knife edges located at a distance $b \cdot l$ from the ends. Because the sag profile is symmetrical, only the left half of the bar has been displayed. The upper figure shows the calculated deflection curves using $b = 0.2232$, which should be optimal according to the simple beam theory; the lower figure shows the calculated deflection curves using $b = 0.2285$, which should be optimal according to the plane stress theory. There is a considerable difference between the predicted deflections (dashed and solid lines), and also between the calculated optimum support locations (upper and lower figure)[59].

and

$$\frac{1}{R} = \frac{1}{r_{11}} + \frac{1}{r_{12}} + \frac{1}{r_{21}} + \frac{1}{r_{22}} \tag{2.25}$$

In the case of a ball with radius R_1 on a flat surface $r_{11} = r_{12} = R_1$ and $r_{21} = r_{22} = \infty$ reducing equation 2.25 to

$$\frac{1}{R} = \frac{2}{R_1} \Leftrightarrow R = \frac{1}{2}R_1 \tag{2.26}$$

Now the average contact pressure is given by

$$p_{avg} = \frac{F}{\pi r^2} \tag{2.27}$$

which gives, by solving for F and substituting 2.23,

$$F = 9\pi^3 p_{avg}^3 \frac{R^2}{E^2} \tag{2.28}$$

It appears that the maximum allowable average contact pressure in order to prevent plastic deformation is [32]

$$p_{avg,critical} \approx 1.1 \cdot R_r \tag{2.29}$$

in which R_r is the elastic limit.

Finally, substituting equation 2.29 in 2.28 gives

$$F_{critical} = 9\pi^3 \frac{R^2}{E^2} p_{avg,critical}^3 \approx 12\pi^3 \frac{R^2}{E^2} R_r \tag{2.30}$$

representing an upper limit for the applied contact force as a function of the reduced contact radius, the reduced modulus of elasticity and the elastic limit.

Numerical example

If a steel bar with $E_1 = 2.1 \cdot 10^{11}$ N / m^2 is supported by ceramic balls with $E_2 = 3 \cdot 10^{11}$ N / m^2 and $R_2 = 3$ mm, then $E = 2.7 \cdot 10^{11}$ (according to equation 2.24) and $R = 1.5 \cdot 10^{-3}$ (according to equation 2.25), assuming the Poisson constant $\nu = 0.3$. Further assume that the elastic limit for steel is $4 \cdot 10^8$ N / m^2, which is lower than the elastic limit for ceramic materials.

Applying equation 2.30, this results in a maximum allowable contact force $F_{critical} = 0.72$ N while the elastic deformation radius $r = 23$ μm. This means that already for small forces the steel bar is permanently damaged on microscopic scale. This may lead to poor repeatability after removal and relocation of the kinematic support. A graph of the maximum allowable load as a function of the ball radius R for this case is given in figure 2.14.

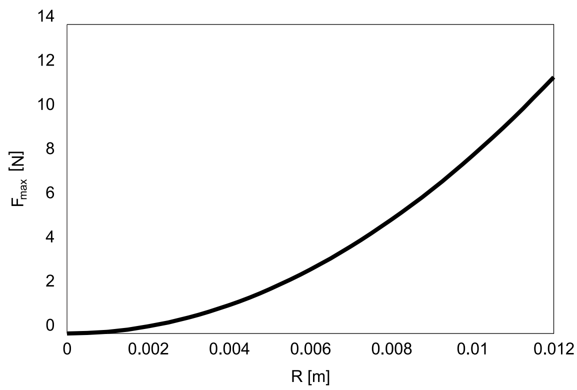


Figure 2.14: The maximum allowable force in a kinematic support point of a ceramic ball on a flat steel surface, as a function of the ball radius R, using the values of the numerical example mentioned on page 42.

In [62] a generalized approach is presented to analytically calculate the position errors as a result of global forces on a kinematically supported bar. Because it is not possible to derive these errors directly from the applied forces, another approach has been made. First, the global forces and torques are translated to forces and torques at the contact points by means of a matrix M_{cg} (c: contact points, g: global) containing geometrical information. Then, from these contact forces the contact deflections are calculated, using the theory of Hertz. Subsequently, the global error motions are calculated from the contact deflections by another transformation matrix $M_{gc} = M_{cg}^T$.

In practice, for carefully placed objects on kinematically correct support points, a repeatability in the order of 10 . . . 100 nm can be achieved [72].

2.3.1.3 Other stresses in straightness reference bars

Not only the supporting points introduce stresses. Also thermal gradients caused by local heat sources result in stress. Besides that, the initial stresses due to machining processes remain in the material unless special heat treatments are applied.

The modulus of elasticity E varies with temperature, an effect which is often disregarded. Hooke's law $\sigma = E \cdot \epsilon$ implies that in that case the strain $\epsilon = \sigma / E(T)$ will depend on temperature as well. It follows that

$$\epsilon = \frac{\sigma}{E} \Rightarrow \frac{\Delta\epsilon}{\epsilon} = -\frac{\Delta E}{E} \tag{2.31}$$

The magnitude of this effect depends on the material. For steel alloys the relative temperature dependency of E ($\Delta E / E$) around 20°C varies from −65 to

-700 ppm / °C. For aluminum and its alloys this rate is in the order of magnitude of -500 ppm / °C [11].

In combination with the thermal and / or mechanical stresses in the bar, the variable $E(T)$ results in temperature dependent deformations. In the case of substantially loaded elements or elements containing large stress concentrations, this $\Delta E(T)$ effect may not be negligible.

2.3.2 Taut wire straightness reference

For lengths up to several meters a taut wire is a possible straightness reference. Its advantages are low cost, good availability of different kinds of wire of constant (mass-produced) quality, predictable shape of the hanging wire and applicability for several measurement methods. The main disadvantages are the sensitivity to vibrations and the possibility of kinks in the wire.

Because this type of straightness reference is not applied further in this research, no attention will be given to all its aspects. Only the vibration behavior will be mentioned shortly here.

2.3.2.1 Vibrations in a taut wire

The *natural frequency* f_0 of a taut wire having length l, mass per meter length μ and tension force F is given by Mersenne's law:

$$f_0 = \frac{1}{2l}\sqrt{\frac{F}{\mu}} \tag{2.32}$$

The *amplitude* of the vibrations obviously depends on the magnitude of the disturbance. Furthermore the parameter EA/l—the axial stiffness of the wire—determines the amplitude [34].

The natural frequency of the taut wire preferably has to be chosen higher than the bandwidth of the applied sensor system. In that case the disturbing influence will be small. However, this requirement limits the geometrical parameters as can be concluded from equation 2.32. For a high natural frequency and given wire length the specific mass must be as low as possible while the force has to be high. Therefore, the optimal material for such a taut wire straightness reference has a high ratio σ_y/ρ, where σ_y is the yield stress of the material and ρ is the specific mass.

Because of the disturbing vibrations occurring after mechanical contact, a contactless measurement method has to be used, for example capacitive, optical or inductive.

Numerical example

Consider a steel wire having a yield stress of $\sigma_y = 1 \cdot 10^9$ N / m^2. $E = 2 \cdot 10^{11}$ N / m^2 and $d = 0.1$ mm. Then $A \approx 8 \cdot 10^{-9}$ m^2, $l = 1$ m, $m = 6.3 \cdot 10^{-5}$ kg, $g = 10$ m / s^2, $F_{t,max} = \sigma_y \cdot A = 8$ N.

This maximum tension force results in a first natural frequency

$$f_0 = \frac{1}{2l}\sqrt{\frac{F}{\mu}} = \frac{1}{2 \cdot 1}\sqrt{\frac{8}{6.3 \cdot 10^{-5}}} \approx 178 \text{ rad/s} \approx 28 \text{ Hz} \tag{2.33}$$

which makes it not suitable for sub-micrometer straightness measurement with high bandwidth.

Note that f_0 does not depend on the diameter d, because both the maximum force F and the mass per meter μ are proportional to d^2.

2.3.3 Liquid surface as a straightness reference

Some publications mention application of liquid surfaces as straightness reference for ranges in the order of 1 m. Actually, only the coordinates z, φ_x and φ_y can be measured with respect to a liquid surface.

Several points ask attention:

- Evaporation; this causes a (very slow) drift in the measured signal.
- Thermal expansion; for liquids, in general this is one order of magnitude larger than e.g. steel and aluminum.
- Vibrations; this limits the applicability to systems with a proper vibration isolation.
- Meniscus effect; this affects the flatness of the surface, especially near the walls of the container. The deviation can be reduced by choosing the right combination of liquid and wall coating and by measuring away from the walls.
- Earth surface curvature; this causes a deviation $\delta = R(1 - \frac{1}{2}\sqrt{4 - \frac{x^2}{R^2}})$, where x is the distance and R is the Earth radius, both in [m]; for example, $x = 1$ m gives $\delta \approx 20$ nm.

Yamaguchi [78] compared different solutions and found out that a 7:1 mixture of glycerol and water has a good combination of meniscus flatness and surface vibration damping.

However, except for the ease of alignment and the absence of deformations at support points, liquid surfaces do not have any advantages above mechanical straightedges.

2.3.4 Optical straightness references

In many measurement applications, a laser beam serves as a straightness reference, e.g. in quadrant cell photodiode (QPD) based sensors or laser interferometers (see also section 2.1). Because of the apparent absence of mechanical disturbances it might seem that a laser beam (or rather, the line connecting the weighted mid-points of the cross-sectional intensity distribution along the beam) describes an ideal straight line. In a perfect vacuum environment this is largely true, but under normal laboratory conditions there are several disturbing influences. Most of these disturbances are related to the refractive index n that is a function of temperature, pressure, humidity and gas concentrations. Table 2.11 gives an overview of environmental influences and their order of magnitude.

Effect	**Condition**	**Magnitude**	**Condition**
Temperature	Global	-1 ppm/°C	
Temperature	Gradients	depends	Causes bending
Pressure	Dry air	0.28 ppm/mbar	around 1 bar
Relative Humidity RH		-0.06 ppm	at RH incr. from 48% to 56%
CO_2 concentration		0.1 ppm	[CO_2] from 0.03% to 0.06%
Sound	60 dB(A)	0.56 ppb	dB rel. to 20 μPa

Table 2.11: Variations in the refractive index n due to variations in environmental parameters. The most important factors appear to be the temperature (gradients) and the air pressure.

Besides the environmental disturbances mentioned in table 2.11, there are some additional uncertainties:

1. A laser beam is no mathematical line but has a varying intensity over its cross section. Therefore, the intensity distribution has to be taken into account. It may be difficult to detect the true reference line, mostly the line connecting the points of mean intensity. Moreover, it may be difficult to guarantee the stability of the intensity distribution.

2. The position of a laser beam is not well defined mechanically with respect to the mechanical environment. Hence it can be necessary to fix the position of the beam by, for example, a pinhole / lens combination or an active control system [72].

3. Optical detection equipment easily is influenced by ambient light unless the detector is properly shielded, the light source is modulated or infrared or ultraviolet light is used (provided that the ambient light contains no IR or UV radiation).

4. To obtain straightness information (in five degrees of freedom) using light, one needs three parallel beams, as described in section 2.2.2.1, making the

mechanical alignment a potential error source. There are possibilities to avoid or reduce this, as described in section 2.2.2.2.

Uchikoshi [72] reports an open-loop pointing stability of a laser beam in the order of 10 nm / 100 mm, under several precautions (vibration isolation, temperature stabilized laser, tubing around the beam). It appears that air flow has a large influence on the position stability: a beam cover reduces the deflections (measured at 200 mm distance) from 500 nm to 20 nm. The residual errors are mainly due to thermal deformations inside the laser tube. - With the use of an experimental position stabilizer an accuracy of $2 \cdot 10^{-8}$ rad has been achieved, which is equivalent to 2 nm / 100 mm. This stabilizer uses two quadrant cell photo diodes (QPD) and piezo actuators driving mirrors to deflect the beam in such a way that its position and orientation is constant in four degrees of freedom (y, z, φ_y and φ_z) relative to the position and orientation of the QPDs.

2.3.5 Concluding remarks concerning straightness references

1. The most uncompromised straightness reference is a laser beam in vacuum, although care has to be taken with regard to the pointing stability and the intensity distribution in the beam's cross section. In ambient air, the straightness of a laser beam will be in the 100 nm range (at a stroke in the order of 100 mm), due to refractive index variations and the mechanical pointing stability of the laser tube. By carefully controlling the environmental parameters, a straightness deviation in the order of 10 nm / m can be achieved [74]. Measurement systems that allow the application of a laser beam as a straightness reference are e.g. quadrant cell photodiode (QPD) based sensors or laser interferometers.

2. The most practical straightness reference is a solid straightedge. The straightness errors of a straightedge due to manufacturing inaccuracies can be in the order of $0.01 \ldots 1\ \mu$m (at a stroke in the order of 100 mm). An advantage with respect to a laser beam straightness reference is that the straightness errors of a solid straightedge can be calibrated. Care must be taken with regard to form stability, deformations due to temperature and temperature gradients. A wide variety of sensors exists that can measure distance with respect to a straightedge. In some cases (e.g. using certain types of optical sensors) additional requirements have to be set to the surface quality.

2.4 Measurement principles and techniques

In this section some methods will be summarized that are used to improve measurement quality (accuracy, sensitivity, linearity etc.) or to make a certain measurement method more suitable for a specific task.

2.4.1 Differential measurement

Measuring differentially is a very common method to reduce a sensor's sensitivity for environmental disturbances. It utilizes the situation that an input parameter—c.q. the displacement to be measured—acts on two sensors simultaneously, with equal magnitude but opposite sign. Since environmental disturbances acting on the sensors mostly will have equal sign on both sensors, they will be canceled out. Stochastic disturbances, however, will be summed up in such a way that the resulting standard deviation $\sigma_{tot} = \sqrt{\sigma_1^2 + \sigma_2^2}$.

Another well-known effect of measuring differentially is that sensors with a large but equal non-linearity (e.g. capacitive sensors with a hyperbolical relation between position and capacitance) become more linear around the symmetry point. Besides that, the sensitivity will be doubled.

2.4.2 Zero point measurement

In tracking control systems and many other applications (e.g. an electromagnetically compensated balance), the only purpose of a position sensor is measuring a certain distance within a range that is in the same order as the sensor resolution. In that case the requirements for such a sensor become quite different from the standard parameters.

While the requirements to the *resolution* and *stability* are not affected, the *absolute accuracy* and the *linearity* become much less critical and may be reduced by more than an order of magnitude. In order to achieve the highest possible resolution, the *sensitivity* has to be as high as possible, in most cases by increasing the gain. However, in practice there is a limit to the gain, dictated by the maximum tracking error and the sensor's output range.

2.4.3 Two and three point straightness measurement

2.4.3.1 Working principle

A well-known method to obtain information about the straightness of an object in one degree of freedom is the so-called three point method, depicted schematically in figure 2.15 [4]. The vertical distance of each third point to the straight line through the two preceding points is a measure for the straightness. When for a series of measurement points at a pitch s these distances are recorded, a set of data will be obtained containing information about the straightness—actually, the flatness along one line—of the object under test.

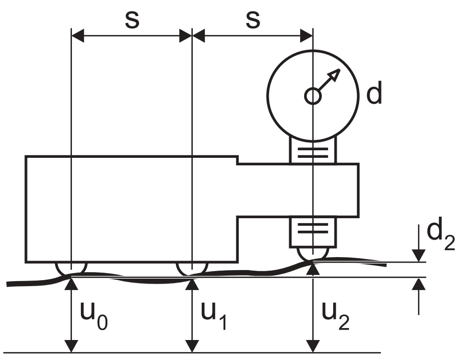


Figure 2.15: A measure for the straightness of an object can be obtained by measuring the deviation of each third point relative to the imaginary straight line through the two preceding points [4].

To do such measurements in a mechanical way, instruments are commercially available. They are provided with two fixed support points having a pitch s. Besides that, they have a third support point, again at distance s, that is movable in the vertical direction, i.e. perpendicular to the line through the first two support points. After calibration (zeroing) on a flat reference surface, the straightness information can be collected by reading the vertical deviations of the third point, while after each measurement the whole instrument is moved a distance s.

From simple geometrical relations follows (see also figure 2.15):

$$d_{i+2} = -u_i + 2u_{i+1} - u_{i+2} \tag{2.34}$$

where u are the actual values of the object's height, d is the measurement value and s is the distance between the measurement points.

In general, one is not interested in values for d but in the straightness information u. There is no directly visible relation between u and d, as is shown in figure 2.16, although there is a unambiguous geometrical relation. To obtain the values for u from the measured d, equation 2.34 is written as

$$u_{i+2} = -u_i + 2u_{i+1} - d_{i+2} \tag{2.35}$$

or

$$u_i = -u_{i-2} + 2u_{i-1} - d_i \tag{2.36}$$

When the unknown values u_{i-1} and u_{i-2} are assumed to be equal to zero, an approximation of u is obtained, as is depicted in figure 2.16.

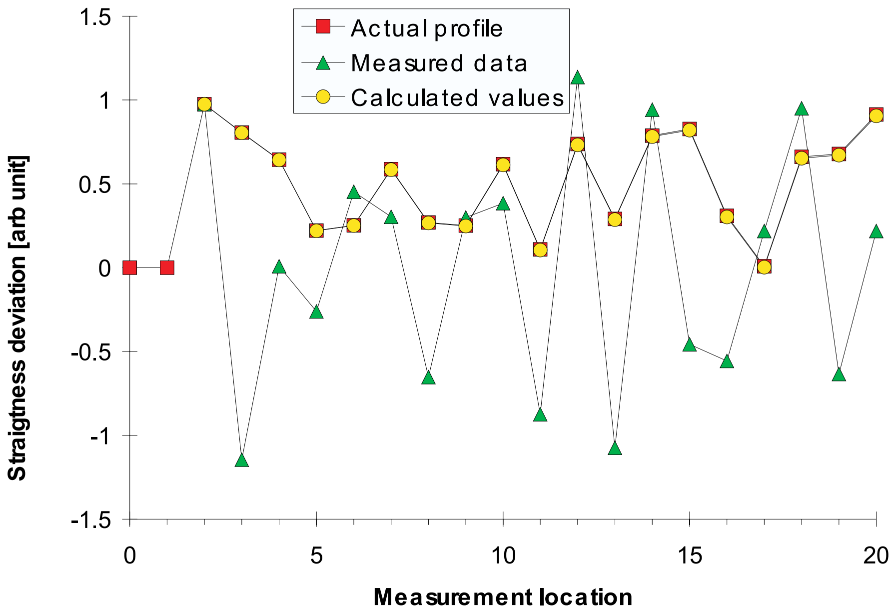


Figure 2.16: A set of simulated random values for the actual surface height values u, together with the measured surface height values d. There is no directly visible relation between the u and d. Also shown is the estimation (calculated values) based on the measured values d. Except for the position and orientation of the curve, the estimation exactly matches, assuming that there is no measurement error.

Due to these two unknown points, two parameters are undetermined in this approximation: an offset and a slope. However, these parameters only define the position and orientation of the object under test relative to the world or any other reference object. They do not have any influence on the straightness error values.

As a matter of fact, the straightness deviations are determined with respect to an imaginary straight line, of which the position and orientation in space are unknown. In general, this also will not be required.

2.4.3.2 Capacitive three point method

The mechanical straightness measurement method as sketched in the previous section can be applied using capacitive sensor elements as well. Instead of the mechanical contact points, capacitive distance measurements are used in that situation.

A problem that arises here, is that the two base points have no zero distance relative to the object under test. Therefore, the accuracy of the straightness measurement (of the third point, relative to the first two) is not defined by the repeatability of the mechanical support points, but depends on the accuracy of the distance measurements.

When the horizontal displacement s is known sufficiently accurate, as well as the distances u_i en u_{i+1}, then the same principle holds, with some minor adaptations following from geometrical considerations.

The measurement at the first sensor element must be equal to the previous measurement from the second sensor element:

$$u_i = u'_{i+1} \tag{2.37}$$

and

$$u_{i+1} = u'_{i+2} = -u'_i + 2u'_{i+1} - d'_{i+2} \tag{2.38}$$

where the accent addresses the corresponding values at the *preceding* measurement cycle.

2.4.3.3 Two point method

When using sensor elements, measuring the relative distance to a straightness reference surface, instead of mechanical contact points, two instead of three measurement locations can give sufficient information about straightness too. This is explained in [68] where this method has been used to measure the straightness of a workpiece in a lathe. Two inductive sensors at relative distance s along the x-axis are mounted on the tool holder, measuring the relative distance in the z-direction between the tool holder and the workpiece, see figure 2.17.

When d_i is the detected relative distance at a measurement location i, u is the tool motion profile and v is the workpiece surface profile (all measured in one degree of freedom, e.g. the z-coordinate), then

$$\begin{aligned} d_{0,B} - d_{0,A} &= v_1 \\ d_{1,A} - d_{0,A} &= v_1 - u_1 \\ d_{1,B} - d_{0,B} &= (v_2 - v_1) - u_1 \\ d_{2,A} - d_{0,A} &= v_2 - u_2 \\ d_{2,B} - d_{0,B} &= (v_3 - v_1) - u_2 \\ (etc.) & \end{aligned} \tag{2.39}$$

so that in general can be formulated for the distance values detected at measurement point i:

$$\begin{aligned} d_{i-1,B} - d_{0,B} &= (v_i - v_1) - u_{i-1} \\ d_{i,A} - d_{0,A} &= v_i - u_i \end{aligned} \tag{2.40}$$

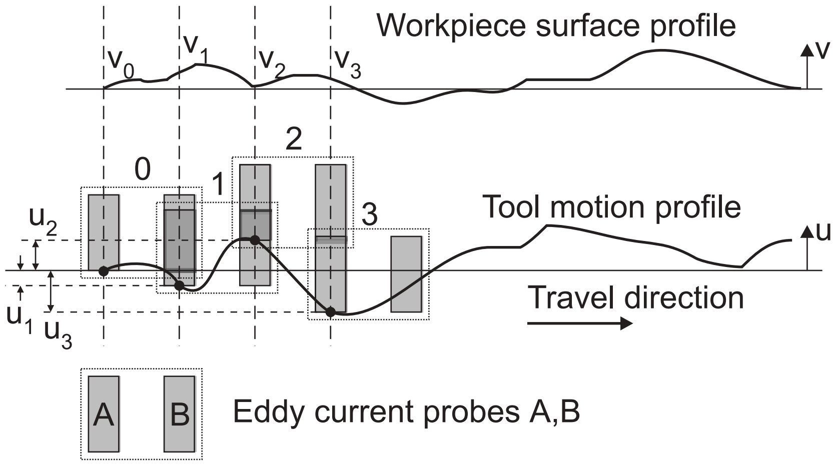


Figure 2.17: Illustration of the two point method. It uses two sensors to measure the relative distance between the tool holder and the machined workpiece in a lathe. The two equations obtained, each with two variables u and v, lead to determination of the absolute tool motion as well as the workpiece profile.

From this follows that the straightness u_i and v_i is

$$\begin{aligned} u_i &= u_{i-1} + d_{i-1,B} - d_{i,A} \\ v_i &= u_i + d_{i,A} - d_{0,A} \end{aligned} \tag{2.41}$$

Attractive in this method is that there is no absolute reference needed, because all measurements are relative to an imaginary ideally straight line. The disadvantage is that measurement errors add up, so that after a series of measurement points a considerable absolute straightness error can occur. Moreover, the straightness profile will be obtained only in discrete points at a certain interval, equal to the spacing of the sensor elements. Furthermore, in the determination of u_i the angle relative to the travel direction is assumed to be constant regardless of the tool motion. Angle deviations will lead to measurement errors proportional to these angle variations and to the smoothness of the workpiece profile.

2.4.3.4 Continuous three point methods

Another technique, that can be seen as an extension to the three point measurement method described above, is the *continuous* three point measurement method.

Unlike the two and three point methods described above, which only can be applied for evaluation or calibration purposes, it appears to be possible to obtain real-time continuous straightness information using the three point Error Separation

Technique. Using this technique resulted in substantially lower straightness errors of a cylindrical workpiece in a diamond turning machine [75, 35, 71]. This technique still is not widely applied. No further research has been done on this topic within the scope of this thesis.

2.4.4 Straightedge reversal techniques

2.4.4.1 Straightedge reversal

According to Estler [17] the maximum straightness accuracy will be achieved when using a form stable mechanical straightedge in combination with a laser interferometer as a measurement device. By means of averaging an effective measurement resolution of 2.5 nm can be achieved. Together with a correction for the refractive index of air this results in an absolute straightness accuracy in the order of 25 nm.

All straightness measurements are done with respect to a straightness reference. As long as this reference's shape does not change, this results in absolute repeatability. Often this is sufficient, but in some cases one wants more information about the absolute accuracy, i.e. one wants to calibrate the straightness reference. The problem often is, especially when a high-quality form standard is applied, that it is difficult to find a straightness reference having an even better straightness, preferably one order of magnitude.

A solution for this problem is to use the straightedge's own geometry by means of reversal techniques, utilizing the principle of *error separation*. An overview of several such techniques is given by Evans *et al.* [20].

The most basic method is sketched in figure 2.18. Assuming a repeatable movement of the guidance, the straightness error S of the used straightness reference can be calculated by reversing the straightness reference while doing two measurements I_1 and I_2, thus separating the absolute straightness of the guidance motion $M(x)$ and the straightness deviation $S(x)$ of the straightedge.

The indicator outputs I at position 1 and 2 are

$$I_1(x) = M(x) + S(x) \tag{2.42}$$

$$I_2(x) = -M(x) + S(x) \tag{2.43}$$

from which follows that

$$M(x) = \frac{I_1(x) - I_2(x)}{2} \tag{2.44}$$

$$S(x) = \frac{I_1(x) + I_2(x)}{2} \tag{2.45}$$

This would be an ideal way to obtain absolute straightness information, if the relocation accuracy of the straightedge support were zero. However, this repeatability

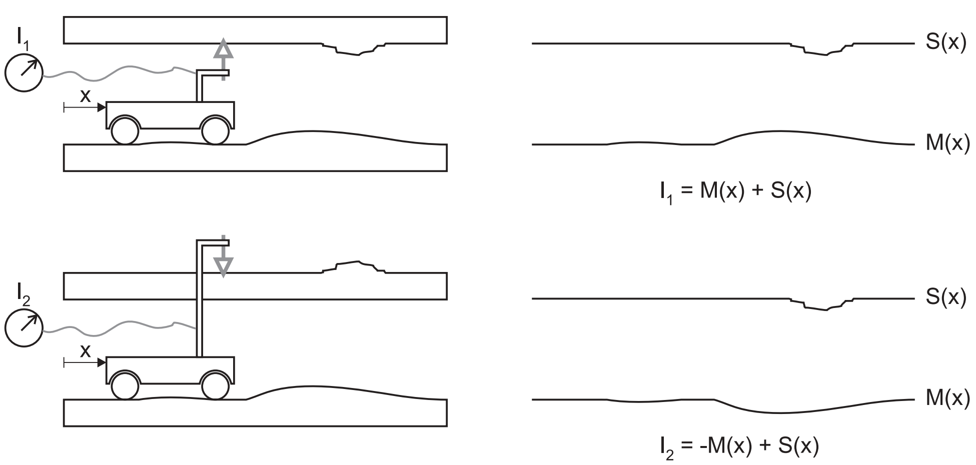


Figure 2.18: Straightedge reversal technique. By doing two measurements I_1 and I_2, the influence of the (repeatable) guidance error motion $M(x)$ can be separated from the absolute straightness deviation $S(x)$ of the straightedge.

error Δx in practice is not zero, decreasing the efficiency of this reversal technique. Now the equations 2.42 and 2.43 become

$$I_1(x) = M(x) + S(x) \tag{2.46}$$

$$I_2(x) = -M(x) + S(x + \Delta x) \tag{2.47}$$

so that the absolute guidance motion M contains an error δ_x

$$M(x) - \delta_x = \frac{I_1(x) - I_2(x)}{2} \tag{2.48}$$

where

$$\delta_x = \frac{S(x) - S(x - \Delta x)}{2} = \frac{-\Delta x}{2}\frac{d}{dx}S(x) \tag{2.49}$$

By introducing the *smoothness* as the first position derivative of $S(x)$, the maximum allowable measurement error δ_x together with the maximum smoothness $\frac{d}{dx}S(x)$ determine the maximum allowable relocation error Δx.

Another method is called the 'three straightedge test', described in [20]. As the name suggests, it uses three different straightedges in three measurement cycli. The advantage is that no repeatable guidance movement is necessary because the guidance motion M is eliminated from the measurements by measuring differentially. The difference with the straightedge reversal technique is that here in each cycle two straightedges are compared with each other instead of measuring each

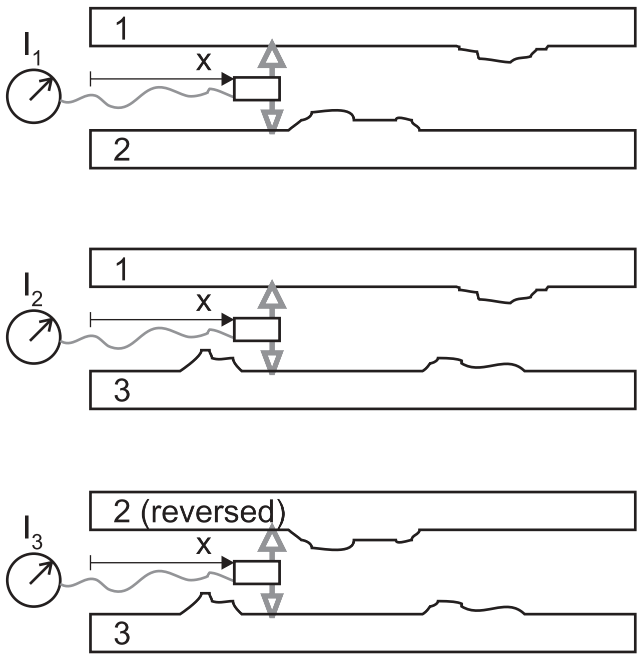


Figure 2.19: Three straightedge technique. In three measurement cycli three straightedge surfaces $S_i(x)$ are compared by means of the indicator I_i so that they all become known, irrespective of the guidance motion error $M(x)$, which does not need to be known or even repeatable [20].

straightedge relative to the guidance motion. An illustration of this method is given in figure 2.19.

The disadvantage of the three straightedge test is that the sensitivity to relocation errors becomes larger than using the method depicted in figure 2.18. The random error due to these relocation changes is $\sqrt{3/2} \approx 1.22$ times as large, compared to methods using only two straightedges.

To reduce these random errors and the relocation errors, it is advantageous to use two straightedges instead of three and measure differentially. Such a two-straightedge method, which is independent on the repeatability of the mechanical guidance, is described in [27]. An illustration is given in figure 2.20.

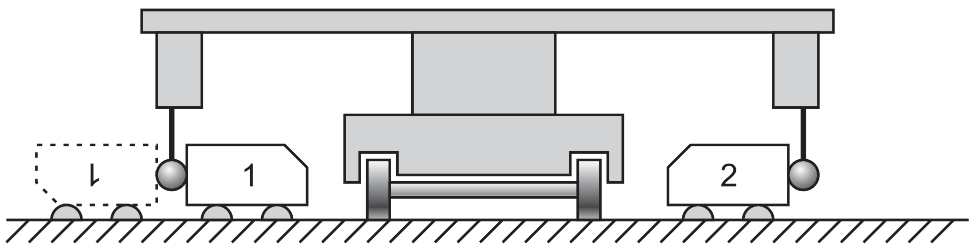


Figure 2.20: Two-straightedge test according to Hoffrogge [27]. To reduce random errors and to minimize relocation errors it is advantageous to use the lowest possible number of straightedges and measurement cycli. Here, in two measurement cycli, the straightedges 1 *and* 2 *are compared, while number* 1 *is inverted in the second run. This gives the absolute straightness of the two straightedges, irrespective of the repeatability of the guidance (which is an advantage with respect to the method sketched in figure 2.18). The direction of movement is perpendicular to the plane of drawing.*

According to Uchikoshi [72] it is very difficult in practice to obtain a relocation accuracy better than 10 nm due to surface imperfections. This means that for very accurate straightedges (25 nm, as mentioned above) the accuracy improvement that is to be expected using reversal techniques will be relatively small, because the uncertainty in the support of the straightedges is in the same order as the straightness errors.

For less accurate (order of magnitude μm) but dimensionally stable straightedges these reversal techniques offer a possibility for relatively simple calibration. After two or three measurement cycli the straightness profile of the used straightedge is known, after which a lookup table can be made to achieve absolute straightness accuracy down to the level of calibration accuracy (short-term) and straightedge stability (long-term).

2.4.4.2 Laser interferometer optics reversal

In section 2.1 the principle of lateral displacement measurement utilizing a Wollaston prism and two mirrors has been mentioned. This straightness measurement is in first order insensitive to displacements and rotations of the moving prism, i.e. when moving in x-direction while measuring z, the coordinates y, φ_x, φ_y and φ_z do not influence the measured value of z.

However, the straightness measurement is sensitive to flatness deviations of the reflector surfaces. In [29] it is shown that also this disturbing influence can be eliminated by means of a similar reversal technique. While the Wollaston prism moves along the x-axis and straightness deviations of the z-axis are measured, the straightness reflector is rotated 180° around the x-axis in the second measurement cycle. In this way, nothing changes, except for the flatness deviations of the reflector which change sign, so that they can be eliminated from the equations.

2.4.4.3 Concluding remarks concerning reversal techniques

1. Reversal techniques are useful for the calibration of straightness references. The more accurate the straightness of the reference is, the less improvement can be expected due to the limited relocation accuracy after reversing.

2. For the use in real-time feedback systems reversal techniques are not suitable, because a crucial aspect of the working principle is that one of the measured parameters physically changes sign.

2.4.5 Metrology frames

Anticipating on chapter 4 and 5, where systems using position error feedback will be described, the principle of the metrology frame will be mentioned here briefly.

When measuring the position or straightness deviation of the moving object relative to a straightness reference, care has to be taken to ensure that the reference's straightness is not influenced by environmental conditions. In particular, mechanical forces resulting from a change of environmental conditions should not deform the straightedge.

This is achieved by making a physical separation between the mechanical parts that are subject to internal or external forces and the mechanical parts that play a role in position measurement. Internal forces can be e.g. the varying forces due to the displacing mass of the carriage or the dynamical forces as a consequence of bearing imperfections.

Thus, two paths appear in the motion system: a *force path*, which conducts the occurring forces to the world, and a *measurement path*, which provides for a stable position reference. The mechanical parts forming the measurement path are called the *metrology frame*.

The physical separation obviously can only be accomplished to a certain degree, because there will always be a mechanical interconnection between the two paths. In practice, mostly the world (i.e. a baseplate that is assumed to be undeformable) is the connecting link.

Application of a metrology frame is only useful when the required inaccuracy is smaller than the errors due to deformations of the mechanical guidance. Otherwise the sensor system can measure relative to a part of the guidance, which is less complicated and less expensive.

2.4.6 Concluding remarks concerning measurement principles

1. When the deterministic errors are larger than the stochastic errors, it is advantageous to measure differentially to eliminate common-mode errors such as temperature effects. However, stochastic errors will increase. Moreover, the required number of sensor elements is twice as high and also some more signal conditioning circuitry is needed.

2. It is always advantageous to apply a zero point measuring system whenever possible. Because the range / resolution ratio of most sensors is constant, decreasing the measurement range generally leads to a higher resolution.

3. Two and three point methods may be useful as a kind of inherent calibration. Only the continuous methods are suitable for real-time measuring in a feedback loop. The disadvantage is that a double or triple number of sensors is needed, together with some computing power. In addition, there always is a certain accumulation of measurement errors, that may result in drift.

4. Metrology frames are necessary in applications where considerable forces occur, internal or external, to ensure that these forces do not influence the closed-loop positioning accuracy.

Chapter 3

Bearing systems for straight motion systems

This chapter gives an overview of existing technologies for high precision straight motion. Because the most crucial element of a straight motion system is the bearing part, which separates the moving part from the fixed part, straight motion systems can be characterized by the type of bearing that is applied.

The first three sections of this chapter describe bearing systems having mechanical contact, the next sections describe contactless bearing systems.

3.1 Sliding bearings

One of the most basic bearing types is the sliding bearing. As the name suggests, the moving part is guided along a baseplate while being in sliding contact. This sliding contact is dry friction, unless lubrication is applied. When the lubrication is such that there is no contact any more between the moving part and the baseplate, then the bearing is called hydrostatic or hydrodynamic. These bearing types are described shortly in section 3.4. A special case occurs when the lubrication medium is gaseous instead of liquid; then the bearing is called aerostatic or aerodynamic. These bearing types are described shortly in section 3.5.

Sliding bearings are applied in a wide variety of machines and instruments, with a corresponding wide range of specifications.

The most important properties of mechanical straight motion systems utilizing sliding bearings are listed below [63].

- Material combinations

1. Cast iron on cast iron: inherent lubrication (graphite in iron), micrometer accuracy possible if scraped, wear due to equal hardness of surfaces
2. Cast iron on steel: better wear resistance than cast iron on cast iron, high stiffness
3. Bronze on steel: lubricant in pores released by heating, high speed applications, low precision
4. Polymers on virtually anything: minimal stick-slip, contamination sensitive (embedded contamination particles in polymer cause wear)
5. Virtually anything on ceramics: more expensive, e.g. ruby bearings; machineable very flat, low internal stress build-up due to the brittle nature of ceramics

- Velocity and acceleration: due to relatively high coefficients of friction most applications are those where low or medium velocities and accelerations are required, order of magnitude 0.25 m / s and 0.1 g, respectively.
- Motion range: unlimited by means of coupling of guidance elements.
- Loads: allowable loads can be high in the case of plain sliding bearings. For kinematically designed bearings this is considerably less. Contact pressures are less than 1 MPa in general, sometimes up to 10 Mpa.
- Accuracy: depends on the guidance type, order of magnitude micrometers. By applying special materials and design choices (e.g. a light preload) submicrometer positioning accuracy is possible in the travel direction.
- Repeatability: $0.1 \ldots 1\,\mu$m after running-in thoroughly. For heavily preloaded systems: 2 μm.
- Resolution: limited by stick-slip. For PTFE the difference between static and dynamic coefficient of friction (responsible for the stick-slip phenomena) is less than 10%, which is the lowest known value.
- Preload: for precision guidances 5 . . . 10% of the allowable load. For higher required stiffness a higher preload is needed.
- Stiffness: can be very high, more than 10^8 N / m. Kinematical guidances have lower stiffness due to point contacts.
- Vibration and shock resistance: very good, but considerably lower for kinematical guidances.
- Damping: low perpendicular to travel direction, low to medium in the travel direction
- Friction: coefficients of friction are in the range of 0.03 (steel / PTFE) to 0.2 (steel / steel).

- Sensitivity to environmental influences: transmissibility for external vibrations depends on stiffness, mass and damping.

In order to reduce the influence of the bearing on the repeatability errors, the best approach is to follow the kinematic design rules. However, this implies the use of point contacts rather than surface contacts resulting in a lower stiffness.

In order to achieve maximum smoothness, it is preferred to have as large as possible bearing pads, in contrast to the kinematical point contacts mentioned above. Thus, in general, a compromise has to be found between maximum repeatability and maximum smoothness.

3.1.1 State of the art in sliding bearings

Probably the highest possible straightness accuracy for a sliding bearing has been reported by Lindsey *et al.* [38]. They describe a linear slideway used for profile and texture measurement in a range of 50 mm. Because the instrument is built nearly entirely from materials having near-zero thermal expansion, the susceptibility to thermal disturbances is very low. Under laboratory conditions ($\Delta T < 0.1$ K) a static noise level of 20 pm and repeatability of 1.5 nm RMS over a traverse of 40 mm (or 0.25 nm over 5 mm) have been achieved. These values all are measured in the (vertical) z-direction.

The slideway consists of Zerodur parts with kinematically located bearing elements in between. These bearing elements are described in a patent application by the same author [37]. The underlying principle is that for very thin (several micrometers) polymer layers the difference between the static and dynamic coefficients of friction decreases considerably, resulting in a much better repeatability and accuracy than is achieved when using bearing pads of bulk material. The bearing pads of this linear motion system are spherical surfaces (made of steel with a porous bronze layer) having a radius in the order of magnitude $50 \dots 500$ mm and being coated with $2 \dots 3$ μm PTFE. The bearing pads are in contact with a Zerodur slideway. During the running in process also on this slideway a thin layer of PTFE is formed. After stabilization of this process a very repeatable behavior has been observed.

However, these excellent results only have been achieved in carefully maintained laboratory conditions. This means:

- A quiet laboratory;
- Thermal stability of the room: $\Delta T < 0.1$ K;
- Acoustic shielding: 80 mm thick wood and polystyrene foam;

- Replacement of thermally sensitive aluminum parts of the Talystep[1] measurement stylus by fused silica[2];
- No external load on the carriage, except for the static load of a specimen of which the surface roughness and shape has to be measured.

When the carriage has to bear a certain payload or external forces are exerted on it, the stiffness of the sliding bearing pads becomes important. This stiffness can be approximated by Hertz' theory as has been explained in section 2.3.1.2.

In general, this system remains an open-loop system, which means that it is very difficult to guarantee the performance under unknown or strongly varying external influences.

[1]Talystep is a trade mark of Rank Taylor Hobson Ltd.

[2]Coefficients of thermal expansion are approximately: Aluminum alloys 24 ppm/K; Invar36 2.2 ppm/K; fused silica 0.5 ppm/K; Zerodur 0.02...0.1 ppm/K.

3.2 Ball bearings

The standard ball bearings are well-known already for a long time. They are commercially available in a wide range of geometries and specifications. To a lesser extent also linear ball bearings (LBBs) are available.

The straightness accuracy of linear ball bearings depends on the quality of the guidance and the roundness of the balls. For high-quality LBBs the overall straightness is in the order of 1 μm/cm, expressed as a function of the travel distance in the motion direction.

Through the time many improvements have been made concerning materials, ball roundness and diameter, maintenance and lifetime. However, for normal applications the total accuracy never will reach far in the submicrometer range. This is caused by the inevitable material imperfections, contaminations, wear, small plastic deformations etc. which result in position disturbances in the submicrometer or even micrometer range. At larger velocity the rotating frequency of the bearing balls is clearly visible in the measured position disturbances.

3.2.1 State of the art in linear ball bearings

One interesting development is described by Futami [21]. He divides the motion range of a linear guidance based on a ball bearing in three regions:

1. Displacements smaller than 100 nm; the balls can be regarded as linear springs.
2. Displacements between 100 nm and 100 μm; the balls act like non-linear springs.
3. Displacements larger than 100 μm; the rolling contact results in a hysteresis loop.

Implementing this model in the control system of a linear positioning system resulted in a position resolution in the nanometer range. Only the direction of motion has been regarded in this experiment; no attention has been paid to the other degrees of freedom. A similar approach might be possible there, although only the first and the second displacement regions are of interest there.

The most important advantages of this approach are the high accuracy for low velocities and the high (virtual) stiffness, both in the travel direction. The disadvantages are the complex control scheme, the low damping and the high requirements to the control bandwidth when higher velocities are desired.

3.3 Elastic bearings

3.3.1 Elastic elements

In a straight motion system using elastic elements, five degrees of freedom have to be fixed by means of elastic elements in a statically determined way. Two examples are

1. The use of five wire springs
2. The use of two plate springs, of which one has a flexure hinge in order to prevent an overdetermined mechanism

as shown also in figure 3.1. Of course, many other combinations of wire springs and plate springs are possible [16, 60].

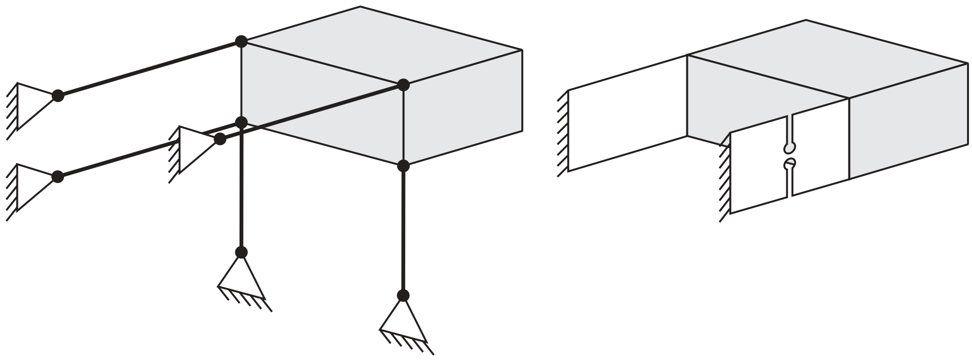


Figure 3.1: Two examples of fixing five degrees of freedom of an object in a statically determined way. In the left picture, the thick black lines represent wire springs, which are stiff only in their axial direction. Other combinations of wire springs and plate springs are possible.

When an object is supported by elastic elements, always a compromise has to be found concerning the stiffnesses in the degrees of freedom that are of interest. For example, a wire spring in the x-direction fixes movements in the x-direction; more precisely, it has an axial stiffness k_x of magnitude

$$k_x = \frac{EA}{l} \tag{3.1}$$

while the lateral stiffnesses $k_{y,z}$ are

$$k_y = k_z = \frac{12EI}{l^3} \tag{3.2}$$

so that for a wire spring with a circular cross section, radius r, the ratio between axial and lateral stiffness becomes

$$\frac{k_x}{k_{y,z}} = \frac{Al^2}{12I} = \frac{\pi r^2 l^2}{3\pi r^4} = \frac{l^2}{3r^2} \tag{3.3}$$

Therefore, it is preferred to have r as small as possible and l as large as possible. However, also the buckling properties are dependent on the dimensions l and r. For the same wire spring the critical load F_{cr} for buckling is

$$F_{cr} = \frac{4\pi^2 EI}{l^2} = \frac{\pi^3 Er^4}{l^2} \tag{3.4}$$

It is clear that it is preferable to have a large r relative to l, while r has much more influence than l. Dependent on the expected axial and radial loads, optimum dimensions in a specific situation can be found for wire springs using equations 3.3 and 3.4.

To reduce the buckling sensitivity, solutions like the one depicted in figure 3.2(b) are known. Compared to a normal wire spring (a) the axial stiffness of spring (b) is three times higher, the lateral stiffnesses are only 20% higher and the buckling strength is nine times higher, assuming that the thick part of the spring covers 90% of the length.

3.3.2 Parasitic displacements in elastic guidances

When displacing one end laterally, without rotation at both ends, a wire spring will introduce a parasitic displacement in its axial direction. This parasitic displacement δx amounts [16]

$$\delta x = 0.6\frac{z^2}{l} \tag{3.5}$$

describing a paraboloid. A wire spring with stiffened middle part, as sketched in figure 3.2(b), behaves more like a bar between ball hinges resulting in a parasitic displacement

$$\delta x = 0.5\frac{z^2}{l} \tag{3.6}$$

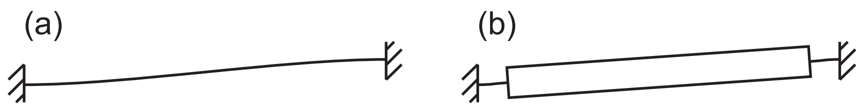


Figure 3.2: (a) Normal wire spring; (b) Wire spring element with increased buckling stiffness, higher axial stiffness and slightly higher lateral stiffness.

These parasitic displacements can be avoided by introducing an intermediate body $\mathcal{H}$ in series with the object $\mathcal{O}$ to be displaced, see figure 3.3. When this intermediate body's displacement is half of the displacement of the object and its parasitic displacement has the opposite direction, the net parasitic displacement can be eliminated, except for manufacturing errors. An even better design is the configuration analog to figure 3.3, but now having two intermediate bodies $\mathcal{H}$ placed symmetrically besides the object $\mathcal{O}$.

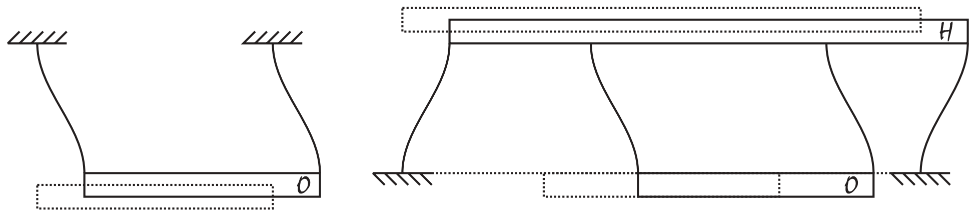


Figure 3.3: By introducing an intermediate body $\mathcal{H}$ in series with the object $\mathcal{O}$ to be displaced, the net parasitic displacement can be eliminated to an extent depending on part manufacturing errors.

3.3.3 Hysteresis in elastic guidances

In elastic guidances, the hysteresis phenomena influences

- the damping; energy will be dissipated during vibration cycli.
- the force–displacement characteristic; this will become non-linear and ambiguous, because the force–displacement curve will become a loop instead of a line.

The hysteresis phenomena is influenced by

- stress; around or above the fatigue stress limit hysteresis rapidly increases. In practice, this limit will be avoided in well-designed spring systems.

The hysteresis phenomena can be divided in two categories:

- internal material hysteresis; this effect is small compared to
- hysteresis in mounting parts, e.g. clamped plate spring ends.

Typical values for the hysteresis in well-designed plate spring systems are between 0.1 and 1% [16].

3.3.4 Straightness approximating mechanisms

Various mechanisms are known that approximate a straight motion. They can be attractive when relatively short travel ranges are required, especially when elastic hinges are applied [60].

The maximum straightness accuracy that can be achieved depends on the accuracy of the length of the bars and the geometrical accuracy of the hinges. For a *four-bar mechanism* having a basic bar length ℓ, a straightness accuracy better than $1.8 \cdot 10^{-3} \cdot \ell$ over a travel length of $1.15 \cdot \ell$ has been reported [60]. In general, this is not sufficient to work with sub-micrometer accuracy, because then the maximum ℓ has to be as low as 0.56 mm. To achieve better straightness, a smaller part of the maximum travel lengths will have to be chosen, or the basic bar length has to be chosen larger relative to the required travel length. On the other hand, the straightness errors of such an elastic mechanism will be well predictable. Sub-micrometer repeatability is not impossible, if the hysteresis is small and the external force variations are small as well.

3.3.5 State of the art in elastic guidances

In the literature some interesting straight motion systems have been reported, based on elastic deformations.

By consequently using the techniques mentioned above (statically determined mechanisms, parasitics compensation) and by carefully assembling and adjusting the spring lengths and distances, straight motion systems have been realized in practice, having the following properties [60]:

- Motion range ± 7.5 mm
- Largest outer dimension approximately 0.5 m
- Lateral accuracy 10^{-4} m / m
- Angular accuracy 10^{-6} rad

Another straight motion system based on elastic deformations, designed for a totally different purpose, is described in [58]. It consists of a central part that is guided along a straight line by means of wire springs. Another preloaded elastic element is connected to the moving part in such a way, that the total elastic energy of the entire system essentially remains constant. By doing this, the stiffness in the travel direction can be made close to zero, while it remains stiff in the lateral directions. This elastic straight motion system has been designed for the application in a seismic vibration sensor, where the low stiffness in one direction was required to achieve a low eigenfrequency, whereas the high stiffness in the other directions was necessary to minimize the parasitic sensitivity for lateral vibrations.

The specifications of this low-stiffness linear motion system are:

- Motion range ± 0.5 mm
- Largest outer dimension 50 mm
- Lowest eigenfrequency, axial 0.25 Hz
- Lowest eigenfrequency, lateral 650 Hz

Realization of a stiffness very close to zero in a certain degree of freedom results in a 'free floating mass' behavior in that degree of freedom. This is required in the case of the vibration sensor mentioned above, but it can also be interesting to aim for zero stiffness in actuators, although that looks contradictorily at first sight. More attention will be paid to zero stiffness systems in chapter 5.

3.4 Hydrostatic and hydrodynamic bearings

The *hydrostatic* bearing is well-known and is used in many different applications. No attention will be paid here to the details of their design, only their properties that can make them interesting for high-accuracy applications will be mentioned briefly.

The most important advantages of hydrostatic bearings are

- Simple geometry, low-cost manufacturing;
- High stiffness (preloaded systems);
- Good damping;
- No wear, because there is no direct contact;
- High position resolution possible; due to spatial averaging by the relatively large bearing surface, the position mainly depends on pressure variations and much less on roughness or flatness errors of the surface;

whereas its disadvantages are

- Continuous circulation of lubricant, necessary to maintain a bearing film;
- The thickness of the lubricant film is load-dependent.

As for the *hydrodynamic* bearings, they are mainly applied in rotating bearings that operate continuously, at a more or less constant angular velocity. This is because the pressure that results in a bearing force, is proportional to the relative velocity of the bearing parts. In a linear bearing this relative velocity becomes zero at the end point where the direction of movement changes sign. Thus there occurs a point without bearing force, which could cause damage to the bearing surfaces. Moreover, a variable bearing force is not preferred in most positioning applications.

In section 4.4 the research on an extreme example of hydrodynamic bearings is presented, using high-viscosity grease as a lubricating medium. This results in a very high stiffness and damping and enables a position resolution in the nanometer range due to a high degree of spatial averaging.

3.4.1 A water hydrostatic bearing

Slocum *et al.* [64] present a hydrostatic bearing constructed of ceramic parts and using water as a bearing fluid. By using a self-compensating design, very high stiffness has been achieved, in the order of $2 \cdot 10^9$ N / m. The straightness errors are smaller than the straightness and roughness errors of the guiding rail, due to the spatial averaging effect of the fluid; the vertical position error is less than 0.5 μm over a 100 mm range.

3.5 Aerostatic and aerodynamic bearings

Although the correct general name is 'gas bearings', often these bearings are called 'air bearings' because normal or dried air is most often used for lubrication. The standard compressed air facilities for pneumatic tool purposes may not be suitable for precision air bearings because mostly this compressed air contains too much oil particles, forming an oil film on the bearing surfaces which leads to increasing friction forces.

3.5.1 Aerostatic bearings

Gas bearings are well-known construction elements for high-precision, low friction applications. They exist in several shapes; rotating and plain. By the supply pressure and the geometry of restrictions and air chambers the maximum load and the stiffness are determined.

By applying a two-sided bearing the stiffness can be considerably increased, because a preload is possible. Under certain conditions it is also possible to achieve zero or infinite (static) stiffness by means of internal compensation techniques. The bandwidth of this stiffness is mainly determined by the internal dynamical properties of the pneumatic system.

The main advantages of gas bearings are:

- No mechanical contact, therefore no wear
- Very low friction, low damping
- Clean operation, except for possible oil particles in pressurized air
- High accuracy due to spatial averaging

while the main disadvantages are:

- High quality surface machining required
- Gas consumption
- Not suitable for vacuum applications

For very high accuracy another problem arises: due to the pressure ripple on the gas supply always a certain position noise will be present. Values for the position uncertainty are typically 50 nm.

3.5.2 Aerodynamic bearings

3.5.2.1 Spiral groove bearings

The most well-known example of aerodynamic bearings is the spiral groove bearing, analogous to the hydrodynamic spiral groove bearing. This type of bearing virtually only is applied in the rotating mode, because its performance is based on the generation of pressure by a velocity difference. Also here the disadvantage is, that at zero velocity there is no bearing force and thus there will be mechanical contact.

Linear bearings always have a return point where the direction of movement changes sign and where the velocity is zero, which makes them less suitable for the application of aerodynamic bearings.

3.5.2.2 Squeeze film bearings

A special type of aerodynamic bearing is the squeeze film bearing. Instead of a supply pressure or a velocity dependent pressure build-up in normal aerostatic or aerodynamic bearings, here the relative distance between the bearing surfaces is varied sinusoidally with a high frequency in order to obtain a bearing pressure. In [53] this is described as follows: "In a squeeze film bearing, high frequency transverse oscillations of one of the bearing surfaces provides a pumping action. The oscillatory squeeze motion results in a time averaged pressurization effect primarily due to the compressibility of the gas film, and the degree of pressurization increases monotonically with the amplitude of the oscillation relative to the average gap."

Figure 3.4 gives an illustration of the working principle. The variation in the air gap height, sinusoidal with time (curve 1), can be translated to variation of pressure with time (curve 3), if a behavior of the air layer according to Boyle's law (curve 2) is assumed. The time average of the pressure (curve 3) is higher than the ambient pressure P_0, resulting in a net bearing force.

Due to the inertia of the suspended object and the high oscillating frequency, the position of the suspended object will be virtually insensitive to the bearing vibrations.

In [1] an experimental set-up is described for a rotating bearing utilizing squeeze film air bearings. This set-up utilizes piezoelectric elements for the axial and radial vibrations that are necessary for the bearing function. In order to minimize energy consumption, the piezo elements have been excitated in their natural frequency. Measured results of these experiments are displayed in table 3.1.

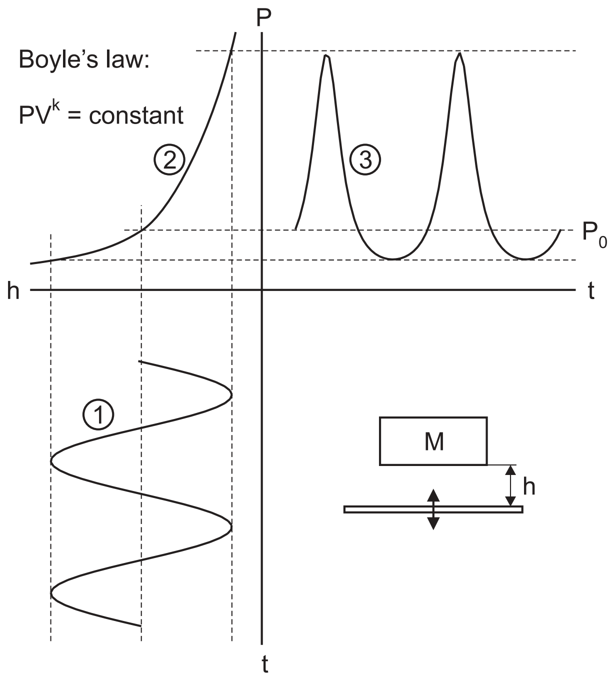


Figure 3.4: Working principle of a squeeze film bearing. The variation in the air gap height, sinusoidal with time (curve 1), can be translated to variation of pressure with time (curve 3) when a behavior of the air layer according to Boyle's law (curve 2) is assumed. The time average of the pressure (curve 3) is higher than the ambient pressure P_0, resulting in a net bearing force [1].

Parameter	Value		Unit
	Axial	Radial	
Vibration frequency	27	24.8	kHz
Vibration amplitude	4.4	3.2	μm
Nominal air gap	10	10	μm
Power	0.4	2.9	W
Stiffness	NA	10^5	N/m
Load capacity	NA	2.5	N
Motion error	NA	10	nm

Table 3.1: Measured results from experiments with a rotating bearing utilizing squeeze film bearings in the axial and radial direction [1].

3.5.3 State of the art in aerostatic bearings

The state of the art in aerostatic bearings is represented in table 3.2 where the typical specifications of a commercially available high-precision linear motion system are listed. Under laboratory conditions slightly better values are possible.

Parameter	Value	Unit	Remark
Travel	$150 \ldots 1000$	mm	various standard lengths
Resolution	0.02	μm	(after 200x interpolation)
Load capacity	25	kg	(uniformly distributed on slide)
Max. velocity	up to 500	mm/s	application dependent
Typ. acceleration	$1.0 \ldots 5.0$	m/s^2	load and duty cycle dependent
Pitch	± 0.5 $< \pm 5$	arcsec arcsec	per 25 mm travel full travel
Yaw	± 0.5 $< \pm 5$	arcsec arcsec	per 25 mm travel full travel
Orthogonality	5	arcsec	stacked XY configuration
Flatness	± 0.5 $< \pm 3$	μm μm	per 50 mm travel full travel
Straightness	± 1 $< \pm 4$	μm μm	per 50 mm travel full travel
Pos. accuracy	± 0.5 $< \pm 2.5$	μm μm	per 25 mm travel per 300 mm travel
Pos. repeatability	± 0.25	μm	per 25 mm travel, 3σ

Table 3.2: Typical technical specifications of a commercially available (Anorad) linear motion system using air bearings.

From table 3.2 it can be seen that even the high-end straight motion systems will not operate in the sub-micrometer range. This is due to the fact that the positioning accuracy is directly related to the manufacturing accuracy achievable.

3.6 Electromagnetic bearings

The technology of magnetic bearings is relatively new. Only in the last years a substantial development towards practical useability and commercial availability has been reported. The main reason for this is that only in the last decennia the control engineering theory and practical electronic circuitry have been developed sufficiently far to enable the closed-loop control of magnetic circuits. The control of electromagnetic bearings requires high accuracy and high bandwidth, for the sensor system as well as for the actuator part. The name 'active magnetic bearing' is often used for this kind of systems, to distinguish them from 'passive' bearing systems

Up to now, most effort has been put in the development of *rotating* magnetic bearings, because the industrial demand for rotating motion is larger than for translating motion. Note that this is analogous to the developments in the field of ball bearing technology and electromotors, where the development of rotating systems has always been earlier than the development of translating systems.

In this thesis no attention will be paid to active rotating magnetic bearings. Only linear and planar types of magnetic bearings will be briefly described. Auer [3] and Molenaar [44] discuss linear magnetic bearings in detail.

3.6.1 General active magnetic bearings

Electromagnetic *propulsion* has been known for a long time. Based on the techniques used in rotating electromotors, linear motors have been devised utilizing for example magnetic reluctance or inductance, or the stepping motor principle. However, in most of these electromagnetic linear motors the *bearing* system is a linear ball bearing or a linear aerostatic bearing. Electromagnetic actuator technology enables the integration of these two functions (propulsion and bearing) in one element, as will be described in section 3.6.2.

With increasing needs for high accuracy, zero wear (no particles allowed in clean rooms) and contactless operation (e.g. in vacuum applications), the development of accurate contactless electromagnetic bearings for linear motion systems becomes more and more of interest.

The most simple linear electromagnetic bearing consists of a C- or E-shaped yoke with a coil wound around it, exerting a force on a ferromagnetic plate that is free to move perpendicularly to the bearing force. To improve the linearity of the current–force relation, a better configuration appears when using two yokes with coils, one on each side of the moving plate. Moreover, to improve the bandwidth of the controlled bearing, a certain magnetic preload force has to be applied by means of a bias current through both of the coils, or by means of permanent magnets [3].

The disadvantage of these C- or E-shaped electromagnetic actuators is that the motion range in the bearing direction is very small, in the order of 1 mm, while in the directions perpendicular to the bearing direction there is no force generation. Therefore, applying only this type of actuator elements, it will not be possible to realize a six-DoF controlled bearing system having a long motion range in one degree of freedom.

3.6.2 The SPU: a short-stroke planar active magnetic bearing

One fundamental improvement to the idea of six-DoF controlled electromagnetic bearing system is the use of the magnetic *bearing* flux in the air gap for the *propulsion* force as well. By placing current-carrying conductors in the air gap of the *bearing* magnets, a Lorentz force can be generated for the *propulsion* perpendicular to the bearing force. It appears that this Lorentz force virtually does not disturb the bearing force.

The electromagnetic building block providing such a suspension and propulsion force by using the same magnetic field, has been called Suspension and Propulsion Unit or SPU [3].

The advantages of this SPU are:

- No mechanical contact, so the bearing is suitable for clean room applications.
- High resolution; there is no inherent limitation for the positioning resolution. The applied sensors in the control system determine the eventual positional accuracy [73].

The disadvantages are:

- The necessary preload results in a high power consumption and corresponding heat dissipation.
- The Lorentz coils in the air gap need some space so that a relatively large air gap is necessary. This can be solved by applying magnetically conductive coils, e.g. iron wire.
- The Lorentz force is an order of magnitude smaller than the reluctance force. As a result, the acceleration that can be achieved in the travel direction will be relatively small, compared to common linear motors.
- Due to the C- or E-shaped yokes the stroke in the travel direction is limited by the yoke dimension, because the flux has different sign at the pole surfaces. This may be solved by means of commutation, to achieve a translation range larger than the yoke dimension.

The geometry of the yokes and moving parts requires a moving part having a dimension at least equal to the sum of the required stroke and the electromagnet yoke dimension. For linear bearings this will not be a major problem. However, for planar bearings this design becomes quite space consuming, since in the case of three support points also three platens are necessary, each having an area larger than the product of the total range in x and y.

The problem of heat dissipation can effectively been solved by the application of permanent magnets. These permanent magnets provide for a static flux resulting in an 'offset force' while the electromagnets only have to supply a control force [61]. The bearing element presented there is called Non-Coplanar Suspension Unit (NCSU). The term 'non-coplanar' means that the permanent magnetic part virtually does not interfere with the electromagnetic part. However, the disadvantage of the short stroke (perpendicular to the bearing force) remains.

3.6.3 The PAMB: a planar active magnetic bearing

The next step in planar electromagnetic bearing design is the elimination of large yoke surfaces by a novel geometry utilizing crossed flux guidance bars [44]. In fact, this geometry consists of three NCSUs folded out three-dimensionally and combined in an inventive way. This prototype suspension and propulsion system has been called the Planar Active Magnetic Bearing (PAMB).

This new design enables a large planar stroke for a relatively light moving part, together with the low energy consumption due to the application of permanent magnets. The propulsion forces can also be generated by means of coils around the moving yoke (Lorentz force), but another possibility is to wind the coils around the *stator* parts, so that the moving part only consists of iron, carrying neither permanent magnets nor coils [43].

A comparison between the three bearing types is given in table 3.3.

Parameter	**SPU**	**NCSU**	**PAMB**
Motion range	10×10 mm	10×10 mm	150×150 mm
Power dissipation (coils)	80 W	$\mathcal{O}$(W)	$\mathcal{O}$(mW)
Accuracy	$\mathcal{O}(\mu\text{m})$	$\mathcal{O}(\mu\text{m})$	$\mathcal{O}(\mu\text{m})$

Table 3.3: Comparison between the three electromagnetic bearing types as they have been developed successively at the TU Delft. Values are typical values or orders of magnitude. The reason for the low power dissipation in the coils of the NCSU *and the* PAMB *is the use of permanent magnets for generating an offset force.*

The remaining disadvantages in the PAMB design are the size (approximately three times the stroke) and the Lorentz-type of propulsion force, which is an order of magnitude smaller than is possible using reluctance-based forces. The propulsion force can possibly be improved by means of techniques derived from linear reluctance motors, but this still is subject of research.

Chapter 4

Various experiments on straight motion systems

4.1 Introduction

In conventional mechanical linear bearing systems, the most important sources of straightness errors can be distinguished as follows:

1. Form errors of the guidance due to machining inaccuracy (initial error) or wear (time-varying error)
2. Irregularities of the bearing (unroundness of bearing balls, hysteresis, air fluctuations, etc.)
3. Deflections due to external forces and inertial forces of the load
4. Thermal deflections
5. Displacements due to external vibrations

One of the conclusions that can be drawn from chapter 3—summarizing the state of the art in straight motion systems—is that purely mechanical systems encounter rapidly increasing problems when the required accuracy approaches the sub-micrometer region.

Therefor, in this chapter experiments are described on the possibilities of feedback control techniques. The goal is to reduce the straightness errors of an open-loop linear motion system with at least a factor ten, so that the eventual accuracy will be in the sub-micrometer range.

A general block scheme of open-loop straight motion systems is shown in figure 4.1. On each of the mentioned blocks disturbing external influences may act. An illustration of the possible error sources for positioning systems, with a rough indication of their magnitude, is shown in figure 4.2.

By adding a feedback loop to the straight motion system, the influence of most of the error sources listed above can be reduced [49]. In that case the problems in the *mechanical domain* are converted to *the electronic and the control engineering domain*, where there are more possibilities to solve them. This is expressed in figure 4.3, referring to figure 4.1. Note that in these general block schemes the names outside the boxes represent a specific experimental setup as described in section 4.2. However, leaving this out of consideration here, these block schemes apply for virtually all types of positioning systems.

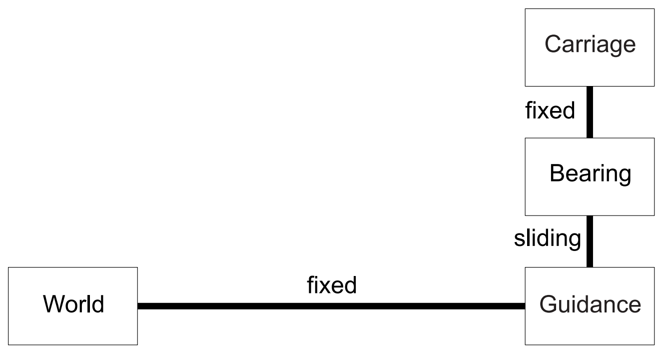


Figure 4.1: Block scheme of general (open-loop) straight motion systems. The thick black lines represent the interactions between the elements. In this example a sliding bearing is assumed, but any other bearing type is possible as well.

A practical implementation of a straight motion system as proposed in the block scheme of figure 4.3 is shown in figure 4.4. To illustrate the principle, only two of the five degrees of freedom are depicted in this schematic view.

A system that measures the actual straightness errors and generates a correcting force or displacement, so that the positioned table follows a straightness reference, will have the following effect on the various position errors that were listed above:

1. Form errors of the guidance due to machining and wear: these errors are quasi-statically repeatable and can be eliminated by a correcting system based on a look-up table and / or real-time control.

2. Position errors due to bearing irregularities (e.g. unroundness of bearing balls, air bearing pressure fluctuations): these errors are dynamical and often dependent on the travel velocity. They can be reduced depending on the dynamical properties of the mechanical structure and the performance of the control system.

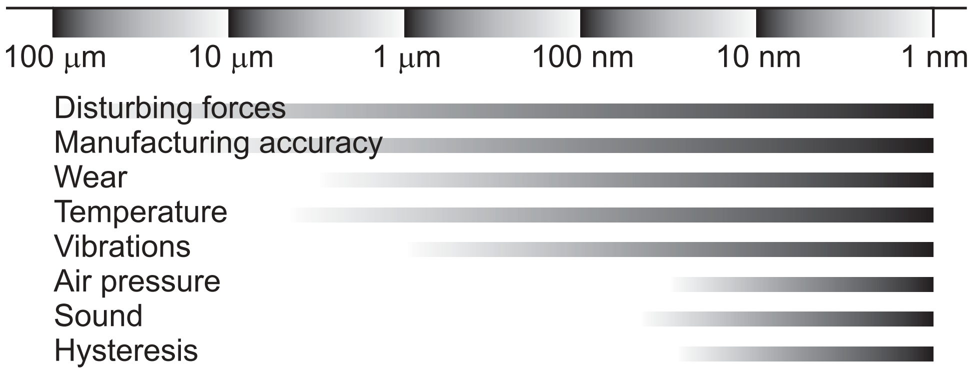


Figure 4.2: Typical values for possible error sources in open-loop positioning systems. Feedback controlled systems are able to compensate for the most important errors and can achieve a final accuracy in the sub-micrometer range.

3. Position errors due to external forces: these errors can be eliminated, within a certain bandwidth.

4. Position errors due to thermal drift: these errors are quasi-static; they can be eliminated by real-time control.

5. Position errors due to external vibrations: these errors can be reduced (as described in this chapter) or prevented (as described in chapter 5).

The measurements are preferably carried out with respect to an *unloaded* mechanical straightness reference (e.g. a straightedge). As can be derived from figure 4.4 the final straightness accuracy and stability of the table is limited by the quality of the straightness reference. However, as long as *form stability* rather than absolute straightness of the reference can be guaranteed, it is possible to achieve absolute (or strongly improved) accuracy by calibrating the straightness reference with an external instrument and making a lookup-table. Without this calibration, only absolute repeatability can be achieved. Often this absolute repeatability also will be sufficient, because there are not many applications in which absolute accuracy really is required.

Summarizing, dependent on the needs in a certain application, three options are available for the choice of a straightness reference:

No special precautions → Only short-term repeatability

Guaranteed form stability → Long-term absolute repeatability

Guaranteed form stability with additional calibration → Long-term absolute accuracy

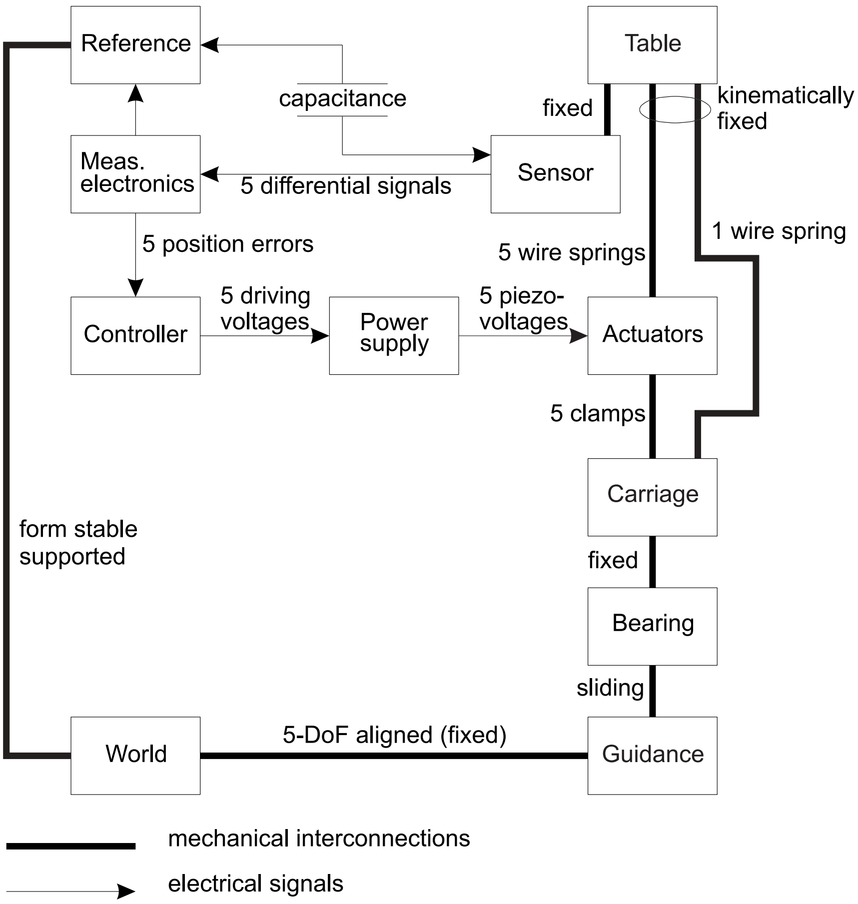


Figure 4.3: Block scheme of a feedback controlled straight motion system. Thick solid lines represent mechanical connections, thin lines mean electrical signals or interactions. An additional table has been placed on top of the open-loop system (figure 4.1) by means of an actuator system. In this general block scheme the descriptions outside the boxes represent a specific experimental setup that will be described in section 4.2.

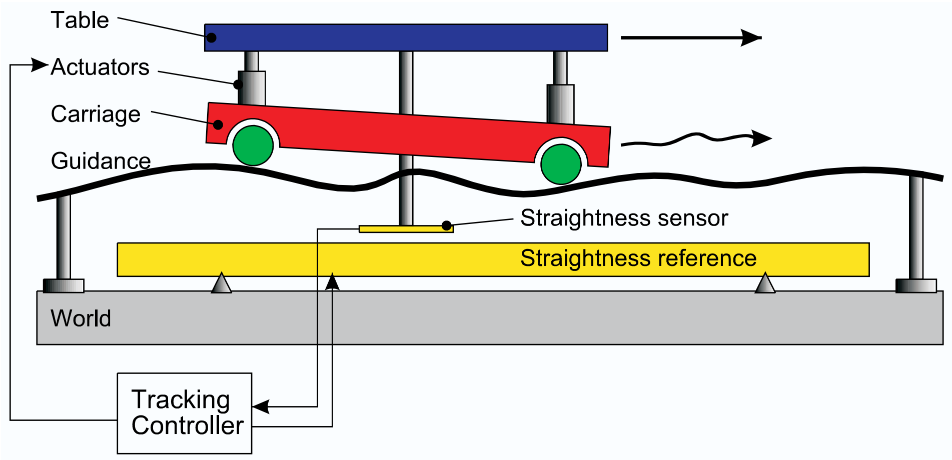


Figure 4.4: Schematic representation of a practical error feedback implementation according to the block scheme of figure 4.3. For clarity, only two actuators are drawn, although for straightness error feedback in five degrees of freedom, five actuators will be required.

In table 2.10 on page 38 a similar, slightly more detailed list is given.

The only weak point in the system presented in figure 4.3 is the connection between the table and the sensor, as this is the only part that is not controllable in the closed loop. This is visible clearly in figure 4.4, where the distance between the table and the sensor is quite exaggerated. In general there will be defined a certain point or object on the positioning table (e.g. a tool tip or an optical point of focus) that has to move along the desired straight line c.q. along the path prescribed by the straightness reference. Therefore, care has to be taken concerning the dynamical and thermal behavior of the mechanical parts between the actual sensing location and this reference point on the positioned table.

4.1.1 A consideration with regard to energy consumption

In the case that only geometrical disturbances of the guidance are present, i.e. no external forces have to be compensated for, the following important statement applies:

> *The amount of energy, necessary to make the suspended object follow a perfect straight trajectory, will be close to zero, when an ideal error feedback is realized. In that ideal case, the actuators only have to adjust their own length—that is, accelerate part of their own mass—to prevent*

the suspended object from accelerating in undesired directions, while maintaining a bearing force.

Even in the case that external forces are present, the energy consumption will be virtually equal to zero because there is no net displacement in the direction of the disturbing force—in the ideal case. But in practice this is not completely true, since an actuator having a finite stiffness k_a and bearing a load F_l will have an initial deflection

$$\delta_{init} = \frac{F_l}{k_a} \tag{4.1}$$

When an external force F_e is applied, an additional deflection

$$\delta_e = \frac{F_e}{k_a} \tag{4.2}$$

occurs which has to be eliminated by the actuator, resulting in an amount of work

$$W = F_e \cdot \delta_e = \frac{F_e^2}{k_a} \tag{4.3}$$

which means that the additional work W is equal to zero when the actuator stiffness is infinite.

In the next sections, experimental set-ups will be described that have been built (a) to demonstrate the feasibility of a position error feedback control for positioning systems and (b) to aim at a positioning accuracy in the sub-micrometer range.

4.2 A straight motion system with five degrees of freedom position error feedback

In this section, an experimental straight motion system will be presented. The development of this system consisted of two phases:

1. Design and realization of a straight motion system, applying the newly developed capacitive straightness sensor (see section 2.2.1), in order to show the practical applicability of this sensor and to demonstrate the feasibility of the straightness error feedback concept. Because of the low measurement frequency of the capacitive sensor prototype, only quasi-static operation will be possible. Most of this work has been carried out within the scope of a Master's thesis [49].

2. Anticipating on the development of similar capacitive straightness sensors having a larger bandwidth, dynamical operation of the straight motion system has been investigated using photonic sensors, in order to identify the system's parts that limit the dynamical positioning performance.

4.2.1 Structure of the straight motion system

The experimental straight motion system has been built according to the structure of the system as shown in figure 4.3 [49]. A short description of the prototype's elements according to the block scheme elements in figure 4.3 is given in table 4.1, a drawing of the complete system is shown in figure 4.5.

Table	Aluminum table, $160 \times 160 \times 20$ mm^3, mass $M \approx 2$ kg
Coupling	Wire springs between actuators and table
Actuators	Stacked piezo, Spindler&Hoyer PA1000-25, stroke 25 μm
Carriage	Aluminum frame
Bearing	PTFE on steel, sliding bearing
Guidance	Two steel bars, straightness errors < 30 μm
Sensor	Capacitive 5-DoF sensor [66]
Electronics	Modified Martin Oscillator [70]
Reference	Steel straightedge
Controller	Discrete, proportional and integrating action

Table 4.1: Listing of structural elements of the experimental feedback controlled straight motion system.

An aluminum *table* is supported by five stacked piezoelectrical *actuators* to control five degrees of freedom with respect to the movement of the *carriage*, being an aluminum frame. The actuators are connected to the table by means of wire springs providing high stiffness ($5 \cdot 10^6$ N / m) in actuating direction and a sufficiently weak coupling ($1 \cdot 10^4$ N / m) in all other directions, because the piezo elements do not allow lateral forces. By means of a sliding *bearing* (PTFE on steel) the carriage is moved along a *guidance* consisting of two steel rods, resulting in open-loop straightness errors in the order of magnitude of 10 μm.

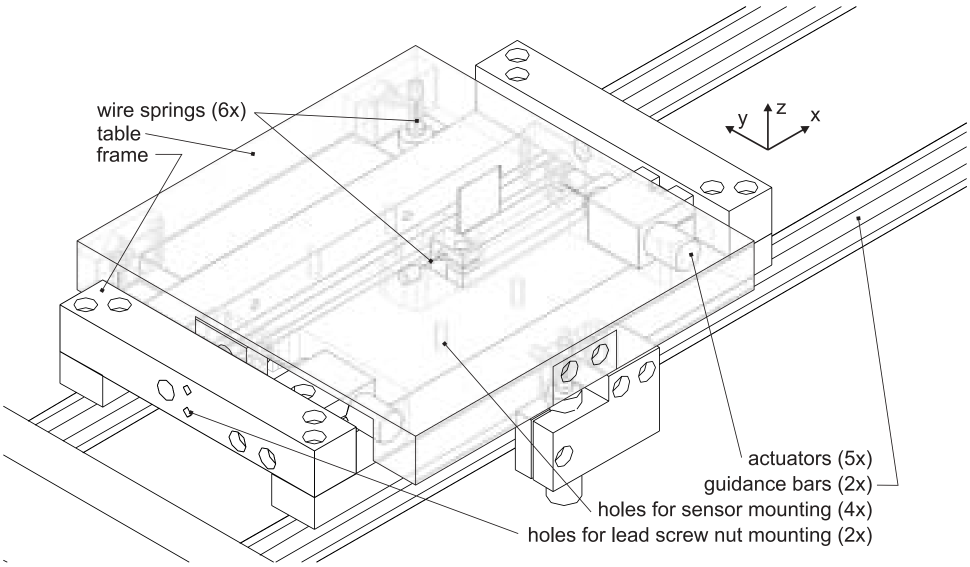


Figure 4.5: 3D drawing of the experimental 5-DoF error feedback straight motion system, showing large similarity to the schematical representation in figure 4.4. Partly visible are the five cylindrical piezo actuators, enabling adjustment of the table's position relative to the frame in five degrees of freedom. Also visible are some of the six wire springs, connecting the table to the actuators and the frame by means of clamping blocks. For clarity, the table has been made semitransparent. The two steel rods serve as a guidance for the five PTFE sliding pads, which are not well visible in this drawing. Omitted in this drawing are the 5-DoF straightness sensor (to be mounted under the table), the straightness reference (which should be located between the two guidance rods) and the lead screw that drives the frame in the x-direction.

As a straightness sensor, an experimental 5-DoF capacitive sensor has been used, as described in [66] and summarized in section 2.2.1. The only difference between this sensor and the sensor shown in figure 2.10 is the layout of the electrodes and the corresponding algorithm to calculate the positions from the measured capacitances. The sensor applied here has the advantage of a simple flat geometry, which is easily adaptable to other configurations. Moreover, it can be used on any electrically conducting bar serving as a straightness reference, as long as there is at least one plane and one edge available. Also it has the advantage that the sensing electrode area provides for a spatial averaging, so that surface roughness and other short-wavelength form errors of the straightness reference virtually do not affect the quality of the measurement. The largest disadvantage is the low measurement frequency, which is limited mainly by the digital part of the electronic circuitry.

Figure 4.6 shows the electrode layout of the applied sensor, which is manufactured of double sided printed circuit board (PCB). Note that two different measurement principles have been applied, as described also in figure 2.2. Three of the electrodes 1 through 3b measure the perpendicular distances between the electrode and the reference bar, from which the z, φ_x and φ_y can be found. The electrodes 4 and 5 measure the overlapping area at the edge, from which y and φ_z can be derived. The electrodes 3a and 3b are used to correct for the z-sensitivity of the electrodes 4 and 5. The use of two different measurement principles causes slightly different specifications in the y- and z-related coordinates, as will appear in the next section. The reason for choosing such a layout is that the whole electrode structure is in one plane, which is very convenient from PCB manufacturing point of view.

The choice for this sensor, manufactured from PCB material, implies a.o. the following uncertainties:

1. Temperature changes will cause linear expansion and possibly warping of the PCB plate. This error will be in the order of 1 μm ($\Delta\ell = \alpha\ell\Delta T = 50\,[\mathrm{ppm/K}] \cdot 1.6[\mathrm{mm}] \cdot 5[\mathrm{K}] = 0.4\,\mu\mathrm{m}$).

2. PCB material deforms under the influence of relative humidity changes of the ambient air. A typical value for the maximum water absorption in PCB material is 0.09%.

3. Geometric errors due to manufacturing inaccuracy (which will be in the order of 10 μm for standard PCB manufacturing methods); the effect of this kind of errors has been discussed in [8].

In this experiment, the feasibility of the method is more important than the long-term stability of the components. The disturbing influences mentioned above might be minimized by choosing appropriate materials and by realizing a geometrical accuracy that is less than about ten times the desired measurement accuracy [8].

To detect the capacitance variations in this sensor, a prototype of a Modified Martin Oscillator-based measuring circuit has been used (see section 2.1.1.2). This results in a range / resolution ratio of 10^4 at a measurement frequency in the order of 1 Hz for 5 degrees of freedom.

After assembling the complete structure, it has to be properly aligned relative to the straightness reference to ensure that the straightness deviations all lie within the working range of the piezo actuators (being 25 μm). Since it is very difficult in practice to obtain this alignment directly from the manufacturing accuracy, an additional alignment facility has been added to the guidance. This alignment facility consists of five alignment screws that are placed in such a way that the guidance's position and orientation can be adjusted relative to the straightness reference (which is kinematically attached to the world) in five degrees of freedom.

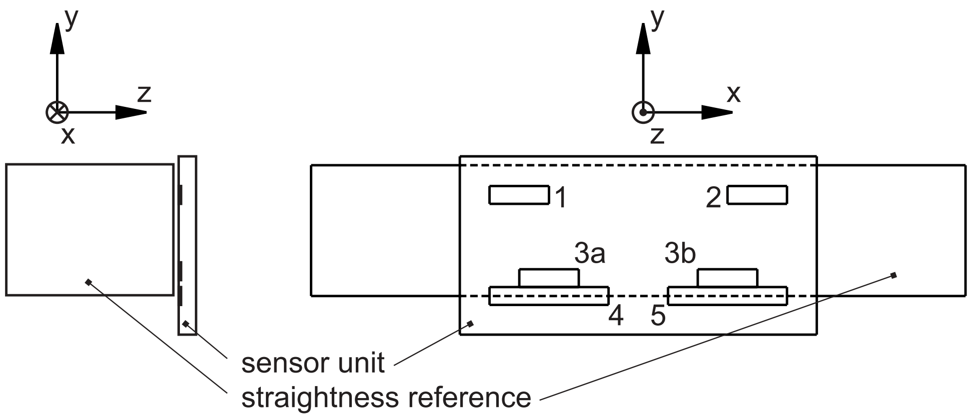


Figure 4.6: Electrode layout of the straightness sensor used in this experimental set-up. The left part of the picture shows the front view, the right part shows the top view. The straightness reference is kinematically supported on the world, the sensor unit moves with the table along the x-axis. There is an air gap of 0.5 mm between the sensor surface and the reference surface. The sensor electrodes 1, 2, 3a and 3b detect the height z and the angles φ_x and φ_y, the electrodes 4 and 5 detect the lateral motion y and the angle φ_z. Note the different measurement principles, varying distance versus varying overlapping area, as mentioned also in figure 2.2.

This alignment procedure is as follows:

1. Measure all five degrees of freedom by means of the straightness sensor at two positions x_1 and x_2 along the guidance, giving two vectors $\vec{x}_1$ and $\vec{x}_2$;

2. Calculate the necessary displacements for each of the five alignment screws $s_1 \ldots s_5$ from
$$\vec{s} = \mathbf{M}_1\vec{x}_1 + \mathbf{M}_2\vec{x}_2 \tag{4.4}$$
in which $\mathbf{M}_i$ are matrices containing geometrical information about the location of the alignment screws and the actuators;

3. To be able to monitor the adjustment of each screw, determine the effect that each alignment screw has on the coordinates, while the table remains on the last position x_2, using the equation
$$\vec{x}_s = \mathbf{M}_s\vec{s} \tag{4.5}$$

4. Adjust each of the alignment screws s_i according to equation 4.4, watching the influence on the table coordinates as predicted by equation 4.5. A proper algorithm can give the instructions to turn each of the adjustment screws until a specified coordinate reaches a certain given value.

By applying this procedure, each of the alignment screws has to be used only once. The angles φ_y and φ_z cannot be aligned by means of the screws on the guidance. Therefore, they have to be adjusted by aligning the sensor relative to the table, or the table (including the sensor) relative to the actuators. Because there are no special adjustment devices implemented for this alignment, this preferably should be done before the final alignment of all other degrees of freedom, as it is described above.

In the controller for quasi-static closed-loop operation, a proportional gain has been applied to the position error, after which this control signal has been added to the control signal value of the previous time step, thus creating a simple integrating action. The value of the gain has been determined in such a way, that the controller response on an arbitrary step disturbance cancels that disturbance. This has given acceptable results for the experiments described in the next section. Due to the long measurement time (2 seconds), not much attention has been paid to fine-tuning of the controller.

4.2.2 Quasi-static measurement results

In table 4.2 and figure 4.7 an overview is given of the different kinds of measurements that are possible on the straight motion system under test. In these measurements, the straightness sensor, the controller and an additional calibration sensor are the major items. It will be clear that the additional calibration device can not be used instead of the straightness sensor, because it will be able to measure in only one or two degrees of freedom, e.g. in the case of a laser interferometer or an autocollimator.

	Cntr	**Sens**	**Cal**	**Resulting information**
A	off	on	off	Guidance errors rel. to straightness reference
B	off	off	on	Guidance errors rel. to calibration device
C	off	on	on	Straightness reference errors
D	off	off	off	No information
E	on	on	off	Tracking error rel. to straightness reference
F	on	off	on	Tracking error rel. to calibration device
G	on	on	on	Straightness reference errors
H	on	off	off	No information

Table 4.2: Overview of the resulting information when testing a feedback controlled straight motion system by switching on or off the three mentioned components: 'Cntr' is the tracking controller, 'Sens' is the 5D straightness sensor used by the controller, 'Cal' means an external absolute position measurement system for calibration purposes.

Figure 4.8 shows for all five degrees of freedom the open loop and the closed loop straightness measurements of type A and E (see table 4.2 and figure 4.7) over a

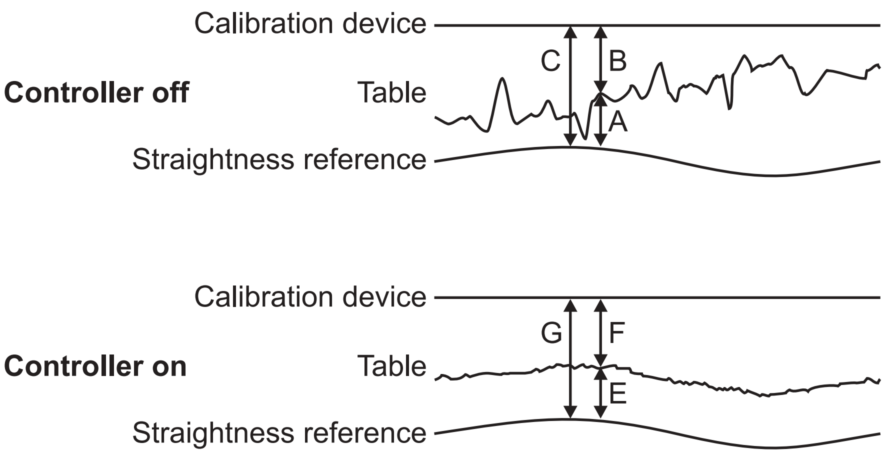


Figure 4.7: Illustration of the different kinds of measurements that are possible on the straight motion system, using the 5D straightness sensor relative to a straightness reference. The letters A *through* G *refer to table 4.2. An external measurement system may be applied as a calibration device (*B, F*) to obtain absolute error information in one or more degrees of freedom, in order to evaluate the 'absolute' quality of the straightness reference.*

travel range of 40 mm. The travel velocity was about 10 mm / min or $1.5 \cdot 10^{-4}$ m / s. For y and z the setpoint values are subtracted from the measured values to display zero in the working point. It is clear that the quasi-static disturbances and misalignments are reduced down to the sensor resolution.

The residual higher-frequency (relative to the measurement frequency) errors remain in the same order of magnitude, although the contribution of frequencies in the order of magnitude of the measurement frequency has been increased. These high-frequency position errors are caused by the simple control action in combination with the noise of the position sensor. Their magnitude might be reduced by setting the proportional gain to a lower value, at the cost of an increasing time constant.

The remaining errors in z are smaller than in y direction. This is due to the sensor configuration, as described in [66], which results in a better sensitivity for z than for y. Therefore, the resolution in φ_x and φ_y (calculated both from z-measurements) is better than the resolution in φ_z (calculated from y-measurements). Further, the resolution in φ_y is better than in φ_x because the angular sensitivity is directly related to the mutual electrode distance, which is larger for φ_y.

It appears that drift and other quasi-static disturbances can be fully eliminated, resulting in a static infinite stiffness (as long as the system is kept within its control range). This is shown in figure 4.9 where a static load of 20 N has been applied for some time. The remaining position uncertainty is in the same order of magnitude

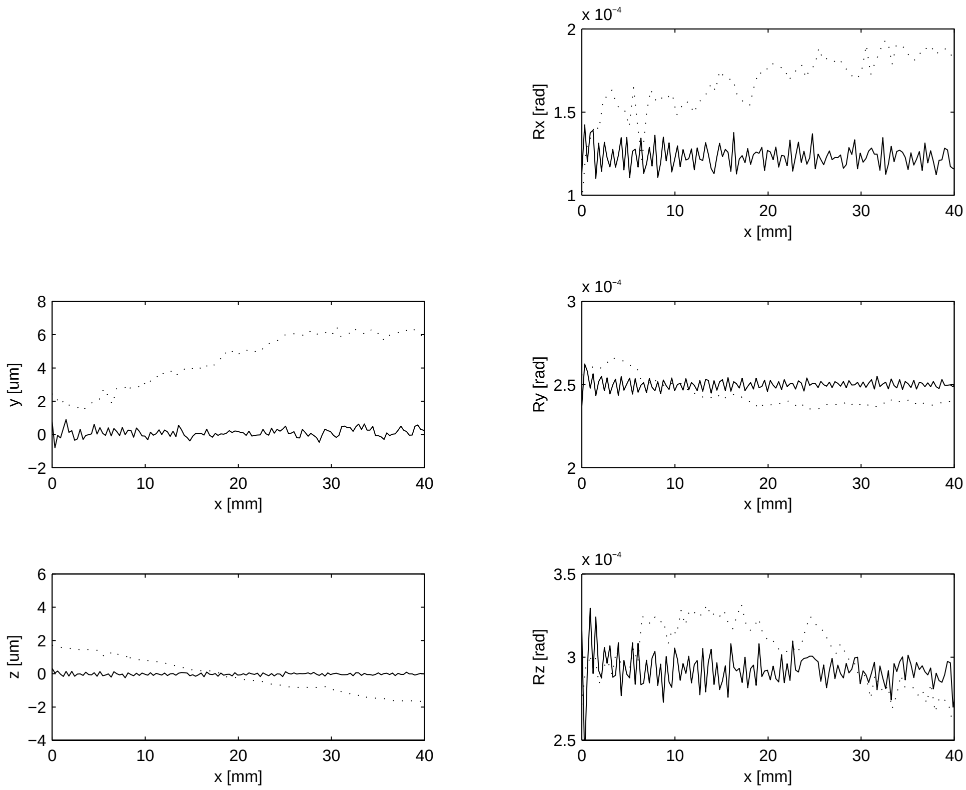


Figure 4.8: Straightness deviations in five degrees of freedom, measured by the capacitive 5D-sensor, without (dotted line) and with (solid line) active control. The travel range is 40 *mm, the travel velocity is about* 10 *mm/min or* $1.5 \cdot 10^{-4}$ *m/s. It is clear that the quasi-static disturbances and misalignments are reduced down to the sensor resolution. The residual higher-frequency (relative to the measurement frequency) errors remain in the same order of magnitude.*

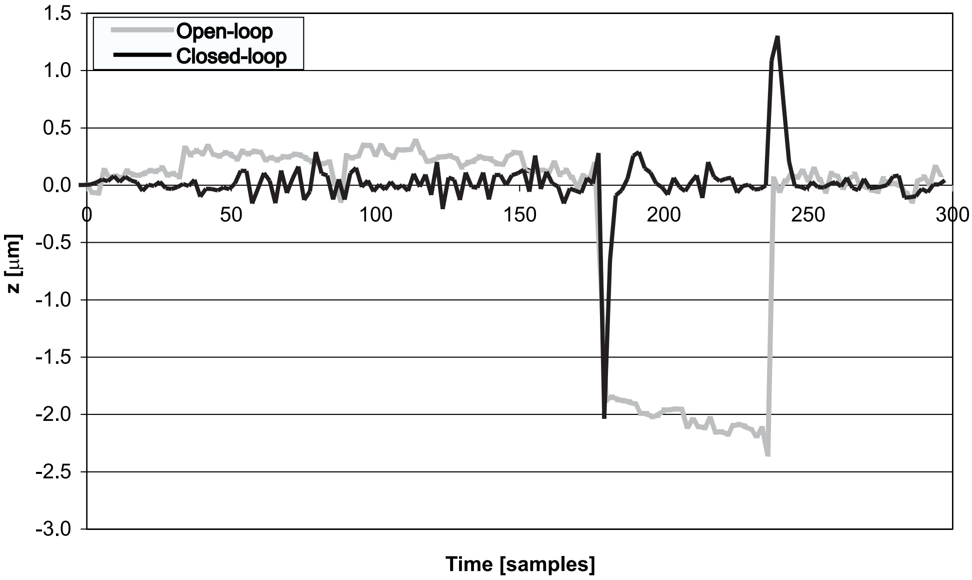


Figure 4.9: Response to a 20 *N static disturbance, open loop (dotted line) and closed loop (solid line). The peaks in the closed-loop signal are caused by the low measurement frequency, about* 1 *Hz. After less than three samples the position has been brought back to its original value. In this experiment, the sample time* $t_{sample} \approx 1$ *s.*

as the sensor resolution for all five degrees of freedom.

In table 4.3 a summary is given of the measured results with this quasi-static set-up.

From this first experimental set-up the following conclusions can be drawn:

1. Sub-micrometer straightness tracking and repeatability has been achieved with mechanically simple parts and a 5 degrees-of-freedom position control. The final accuracy is mainly determined by the resolution and stability of the sensor and by the straightness and form stability of the straightness reference.

2. Capacitive sensors are very suitable for straightness measurements. Due to the inherent spatial averaging by the electrode area, smoothness is guaranteed. Roughness and small surface defects of the electrodes and straightness reference do not influence the performance.

3. With low-bandwidth active control as described in this section, the position noise level remains in the same order of magnitude but the drift is eliminated. By using faster measurement electronics and a more sophisticated controller (probably a PID controller will be sufficient), also the higher-frequency position errors can be reduced, up to a certain bandwidth.

Parameter	Value	Unit	Remark
Straightness deviations y	± 0.5	μm	$\approx$ sensor resolution
Straightness deviations z	± 0.3	μm	$\approx$ sensor resolution
Straightness deviations φ_x	± 15	μrad	$\approx$ sensor resolution
Straightness deviations φ_y	± 5	μrad	$\approx$ sensor resolution
Straightness deviations φ_z	± 20	μrad	$\approx$ sensor resolution
Repeatability	< 1	μm	$\approx$ sensor resolution
Long-term stability	3	μm	Open-loop, 24 hours, $\Delta T \approx 5°$C
Stiffness	$\rightarrow \infty$	N/m	Statical

Table 4.3: Summary of measurement results for the quasi-statically operated straight motion system using straightness error feedback. The short-term deviations all are in the same order as the sensor resolution, from which can be concluded that the sensor resolution is the performance limiting factor. The long-term stability gives an impression of the magnitude of (mainly) the thermal expansion.

4. Near-infinite static stiffness has been realized; dynamical performance will be mainly limited by sensor bandwidth and eigenfrequencies of the mechanical parts.

4.2.3 Dynamic measurement results

After showing that the idea of position error feedback is feasible in practice, modifications have been applied to improve the system's dynamic performance. These measures include:

1. Replacement of the PTFE sliding bearing by a ball bearing to enable higher travel velocity and to introduce certain disturbing displacements, sufficient for performance evaluation purposes. An overview of the set-up is given in figure 4.10.

2. Replacement of the wire springs which couple the table to the piezo actuators by preloaded ball-on-flat contact points to increase the stiffness. This turned out to be not very effective, because relatively high preload forces are required for high stiffness, which may reduce the stroke of the piezo actuators.

3. Replacement of the experimental capacitive straightness sensor by fast (more than 20 kHz bandwidth) photonic sensors, which measure the table position relative to a plane mirror reference surface. Due to limited space only in the y- and φ_z-direction sensors have been applied, which is sufficient for proving the working principle.

4. Replacement of the simple quasi-static controller by a PID-controller.

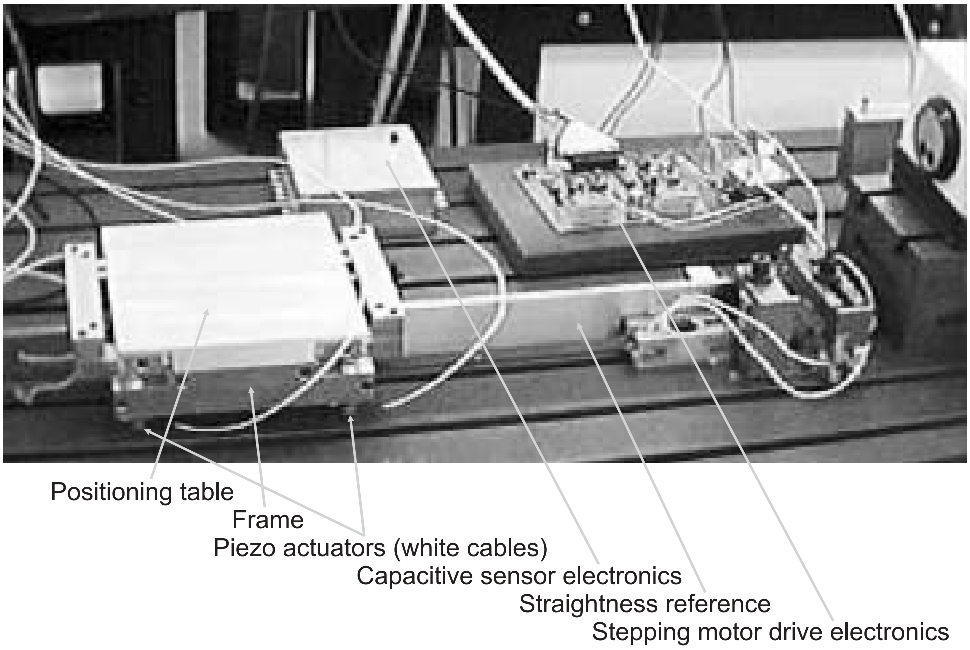


Figure 4.10: Photograph of the experimental set-up, after replacing the sliding bearings by ball bearings. The stepping motor drives a lead screw, with which the frame is moved in the x-direction. The straightness sensor system is not visible here, because it is mounted under the table.

4.2.3.1 Piezo actuator hysteresis

In the experimental straight motion system, stacked piezo actuators have been applied. One of the most important disadvantages of these piezo actuators, besides the short stroke, is the relatively large amount of hysteresis they introduce. In stacked piezo actuators, consisting of many thin layers of piezo electric material, this effect is relatively large and can amount to 15% of the total stroke[1]. In figure 4.11 measurements are shown of the table displacement as a function of the amplifier input voltage.

Note that this input voltage is not the same as the actual voltage over the piezo actuator. The gain factor of the amplifier is equal to 16, so that the total output range ($-10\ldots150$ V) is covered by a $0\ldots10$ V input signal. Furthermore, there are two dynamic effects that have to be considered.

Firstly, the amplifier's maximum current has to be sufficiently large to supply the

[1]This percentage is valid using conventional piezo driving amplifiers, which have a voltage controlled output. When charge rather than voltage is controlled, the hysteresis may decrease to less than 1%. This relatively new method is not considered further in this thesis.

piezo actuator with the desired amount of charge—that is, to bring it to the desired potential—within the required time. The impedance of the stacked piezo actuators used is mainly capacitive, being approximately 2.5μF. Considering the maximum current the amplifier used can deliver, 20 mA, the time needed to bring the piezo actuator's potential from -10 V to 150 V can be estimated by

$$C = \frac{Q}{V} = \frac{\int I dt}{V} \tag{4.6}$$

which gives, assuming that I is constant:

$$C = \frac{It}{V} \Rightarrow t = \frac{CV}{I} = \frac{2.5 \cdot 10^{-6} \cdot 160}{20 \cdot 10^{-3}} = 20 \text{ ms} \tag{4.7}$$

resulting in an upper limit for the frequency, assuming a full-stroke, triangular-shaped output signal:

$$f_{max} = \frac{1}{2 \cdot 20 \text{ ms}} = 25 \text{ Hz} \tag{4.8}$$

When a sinusoidal output signal is assumed, this frequency is even lower. The slew rate is calculated to be $\frac{160 \text{ V}}{20 \text{ ms}} = 8000$ V / s. With a maximum amplitude of 80 V, the signal becomes

$$V(t) = 80 \cdot \sin(\omega t) \tag{4.9}$$

so that

$$\frac{dV}{dt} = 80 \cdot \omega \cdot \cos(\omega t) \tag{4.10}$$

which gives the maximum value for ω:

$$80 \cdot \omega = 8000 \Rightarrow \omega = 100 \text{ rad / s} \approx 16 \text{ Hz} \tag{4.11}$$

This is the maximum frequency at which the sinusoidal full stroke signal is followed 'exactly', assuming that the amplifier does not introduce phase delay.

When not the current limitation of the amplifier, but the time constant of the piezo actuator would be determining the bandwidth, it would mean that for the same bandwidth ω

$$\omega = \frac{1}{R \cdot C} \Rightarrow R = \frac{1}{\omega \cdot C} = \frac{1}{100 \cdot 2.5 \cdot 10^{-6}} = 4 \text{ k}\Omega \tag{4.12}$$

which is a very unlikely value for the resistance R, so that it can be concluded that the bandwidth of the system will be mainly limited by the maximum current of the amplifier.

Secondly, dynamic loads on the piezo elements may cause (high) disturbing voltages due to the inverse piezoelectric effect. These voltages should not influence the supplied current for the piezo actuator. In other words, the output impedance of the amplifier has to be sufficiently high. On the other hand, this voltage might be used as a measure for the force on the actuator.

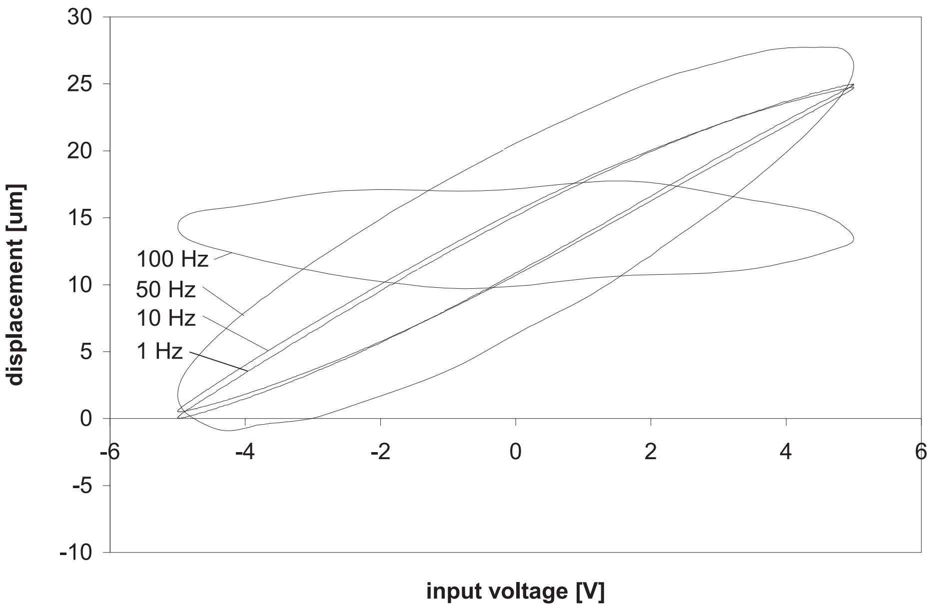


Figure 4.11: Effect of the hysteresis behavior of the stacked piezo actuators on the table displacement. This graph shows the table position y when two parallel actuators drive the 2 kg table in the y-direction. Input is the piezo amplifier input signal (this is not the actual voltage at the actuator) , varied sinusoidally with different frequencies. At $f = 1$ Hz, the curve approximates the statical hysteresis curve. At higher frequencies, the influence of the piezo amplifier, the support point elasticity and the table mass increases.

There are two main methods to solve the hysteresis problem. The first is to add an internal feedback loop in each piezo actuator, e.g. by applying an internal strain gauge or a capacitive position sensor, in such a way that the output displacement becomes linearly proportional with the input voltage. Actuators equipped with such sensors and accompanying amplifiers are commercially available. The second method is to model the hysteresis behavior and to include this model in the present controller [19]. Neither one of these methods has been applied in this experimental set-up.

The error made due to hysteresis in the *open-loop* system can not be translated directly into positioning uncertainty of the *closed-loop* system. Its effect can be regarded as a certain disturbing force that depends on the history of the actuator's movement. This history contains a.o. the direction of movement and the total displacement in that direction. The maximum magnitude of the disturbing force depends on the difference between the gains at the ends of the (full stroke) hysteresis curve. Dependent on the disturbance sensitivity of the closed-loop system, the hysteresis can e.g. increase oscillations at the end of a stepwise setpoint change of the position.

A measurement of the open-loop transfer function from the piezo amplifier input voltage to the table position is given in figure 4.12.

4.2.3.2 Dynamic modeling of the mechanical system

After determining the mechanical behavior by measuring the actuator transfer function H_{act}, we can define requirements for the controller, dependent on the sensor properties. The block scheme of the complete feedback system is given in figure 4.13.

Because the closed-loop system will act as a tracking system—the table has to follow the straightness reference profile—the block scheme of the system has been drawn in such a way that the disturbance signal is located at the left (instead of the setpoint signal z_{nom} which will be zero or constant) and the table position is at the right. This simplifies the calculation of the transfer functions, as will appear in the next.

For an ideal tracking system, the actuator output z_{act} has to be equal to the disturbance z_{dist}.

When the open-loop transfer function from figure 4.13 is given by

$$H_{OL} = H_{sens}H_{cntr}H_{act} \tag{4.13}$$

then the closed-loop transfer function will be

$$H_{CL} = \frac{z_{act}}{z_{dist}} = \frac{H_{OL}}{1 + H_{OL}} \tag{4.14}$$

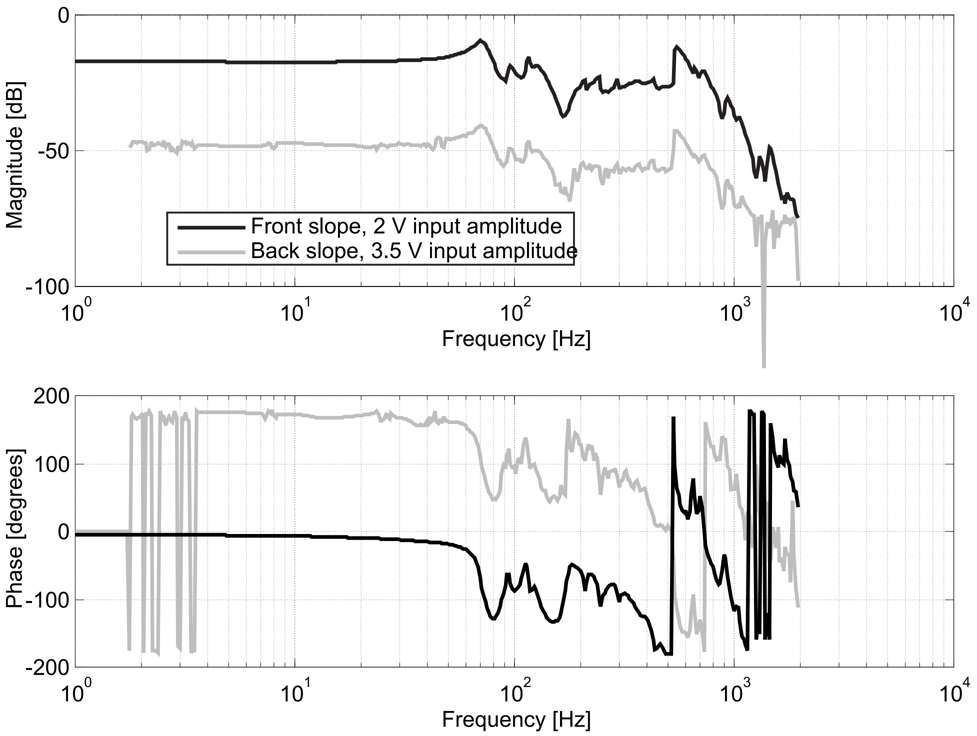


Figure 4.12: Open-loop transfer functions $H_{sens}H_{act}$ of the table driven in the y-direction by two parallel piezo actuators. Input is the piezo amplifier input voltage, output is the photonic sensor output corresponding to the table position in the y-direction. Two different measurements are displayed here: one using the back slope of the photonic sensors and an input amplitude of 3.5 *V, and one using the front slope of the photonic sensors and an input amplitude of* 2.0 *V. There is a scale factor between the measured signals, because of the different sensitivity values of the sensor modes. Also there is an apparent phase difference of* 180 *degrees between the two measurements, caused by the different sign of the sensitivity in the front slope and the back slope, respectively. The magnitude at the frequencies* 1*,* 10*,* 50 *and* 100 *Hz is related to the vertical width of the hysteresis curves in figure 4.11.*

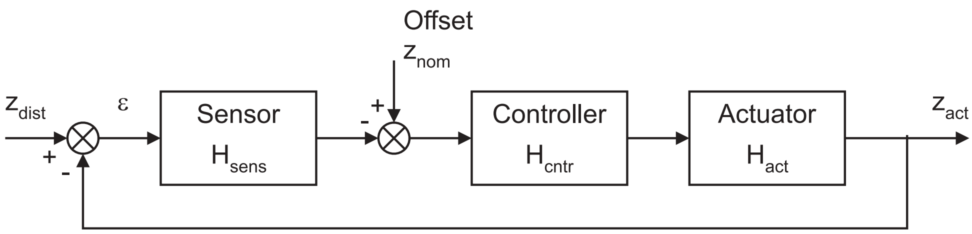


Figure 4.13: Block scheme of the feedback controlled straight motion system. The 'actuator' block also contains the dynamics of the table and its mechanical connection to the actuators.

from which follows that for an ideal tracking system where $z_{act} = z_{dist}$:

$$H_{CL} = 1 \Rightarrow H_{OL} \to \infty \tag{4.15}$$

This is not achievable in practice. But, fortunately, this also is not necessary, for in practice there is a certain maximum allowable deviation ε so that $z_{act} = z_{dist} - \varepsilon$. Now the closed-loop transfer function becomes

$$H_{CL} = \frac{z_{act}}{z_{dist}} = \frac{z_{dist} - \varepsilon}{z_{dist}} = 1 - \frac{\varepsilon}{z_{dist}} \tag{4.16}$$

so that the required open-loop transfer function becomes

$$\frac{H_{OL}}{1 + H_{OL}} = 1 - \frac{\varepsilon}{z_{dist}} \Rightarrow H_{OL} = \frac{z_{dist}}{\varepsilon} - 1 \tag{4.17}$$

However, the maximum allowable deviation ε is a too conservative requirement, when realizing that the present error, $z_{dist}(f)$, depends on the frequency. In general, $z_{dist}(f)$ tends to decrease with increasing frequencies. Thus, for higher frequencies the maximum allowable *relative tracking error* is higher. When at high frequencies the disturbance amplitude becomes lower than the specified maximum position error, the allowable tracking error even will be more than 100%.

From the expected error spectrum and from the desired accuracy, $< 1\ \mu$m in this experimental set-up, the required open loop transfer function—i.e. a lower limit of its value for each frequency—can be calculated using equation 4.17.

Subsequently, the required controller transfer function can be determined using equation 4.13. This gives

$$H_{cntr} = \frac{H_{OL}}{H_{sens} H_{act}} \tag{4.18}$$

$H_{sens}H_{act}$ follows from the measurements according to figure 4.12. The sensor transfer function H_{sens}, because of the high bandwidth of the photonic sensors, may be assumed to be flat in the frequency range of interest, so that H_{sens} simply is the sensitivity of the sensor. With this information, a lower limit for H_{cntr} can be determined.

However, in practice, it is difficult to determine the frequency-dependent disturbances $z_{dist}(f)$. Moreover, there are some additional requirements to H_{OL} concerning stability. Therefore, in this experimental situation, control parameters have been determined by applying the Ziegler-Nichols method, giving control parameters that provide stability and reasonable performance. For the y-direction, this has resulted in a PID controller having the following parameters (valid for sensor operation in the front slope):

$$\begin{aligned} P &= 2.55 \\ I &= 816 \\ D &= 1.275 \cdot 10^{-3} \end{aligned}$$

so that the controller transfer function becomes

$$H_{cntr} = 2.55 + \frac{816}{s} + 1.275 \cdot 10^{-3} \cdot s \tag{4.19}$$

and the cross-over frequencies for the integrating and differentiating action become

$$\begin{aligned} \omega_i &= \tfrac{I}{P} = 320 \text{ rad/s } = 51 \text{ Hz} \\ &\text{and} \\ \omega_d &= \tfrac{P}{D} = 2000 \text{ rad/s } = 318 \text{ Hz} \end{aligned} \tag{4.20}$$

Using this controller, the measurements in the next section have been done.

For a system having six degrees of freedom, of which five have to be controlled actively, the controller in general will be necessarily multivariable, because of the mutual coupling between the input and output parameters. However, when the input parameters (in this case the coordinates y, z, φ_x, φ_y and φ_z) can be expressed in an orthogonal coordinate system, so that they become independent on each other, it is possible to build a decoupled controller, consisting of only diagonal terms. This is done in the following way.

The five sensor outputs represent the five coordinates, but they are coupled as a result of the capacitive sensor's geometry. Therefore, a transformation matrix has to be build that translates the sensor outputs to orthogonal coordinates. These coordinate signals can be input in five parallel SISO (Single Input, Single Output) controllers. The five outputs of these controllers are multiplied by another transformation matrix representing the geometric placement of the actuators. Thus a control system is generated, consisting of five independent controllers, which is much more easy to handle.

4.2.3.3 Measured dynamic performance

Using the modified straight motion test set-up, the following measurements have been done to examine the dynamic behaviour:

1. Open-loop frequency response of the electromechanical system consisting of piezo amplifier, piezo actuator and table
2. Open-loop and closed-loop time response of the y-direction during motion in the x-direction
3. Step response of the closed-loop system by applying a step to the setpoint in the y-direction

These experiments will be described hereafter.

Open-loop frequency response

In figure 4.12 a measurement was shown of the frequency response of the electromechanical system consisting of piezo amplifier, piezo actuator and table. This frequency response defines the transfer function $H_{act} \cdot H_{sens}$ (see also figure 4.13).

In this experiment the table has been driven by two parallel piezo actuators in the y-direction. The input is the piezo amplifier input voltage, the output is the photonic sensor output voltage corresponding to the table position in the y-direction. The amplitude has to be interpreted in relation with the hysteresis behavior as depicted in figure 4.11, i.e. with an uncertainty margin being related to the hysteresis band.

Open-loop and closed-loop time response

In figure 4.14 a typical time response in the y-direction is shown, for both the open-loop and the closed-loop situation. In the x-direction a cyclic motion ($f \approx 1.8$ Hz, $\delta x \approx 10$ mm) has been applied. Note that the time base is not synchronized because the measurements have been done sequentially and the data has been merged afterwards.

In this figure it is visible that a reduction of the error motions is achieved roughly down to the order of the sensor resolution, being approximately 0.5 μm for the photonic sensor's far slope (see section 2.1). Note the remaining vibration of about 8 Hz, which is already present in the open-loop signal; probably this is caused by the unroundness of the ball bearing elements.

The remaining errors are well reproducible, since their main cause is the unroundness of the ball bearings. This means that their magnitude has a direct relation with the position in the x-direction so that this error might be reduced even further by means of learning control techniques.

Closed-loop step response

In the determination of the closed-loop step response the front slope of the photonic sensors has been used (see section 2.1) to achieve the best possible position resolution, being about 10 nm. In this situation, where only a step on the setpoint is

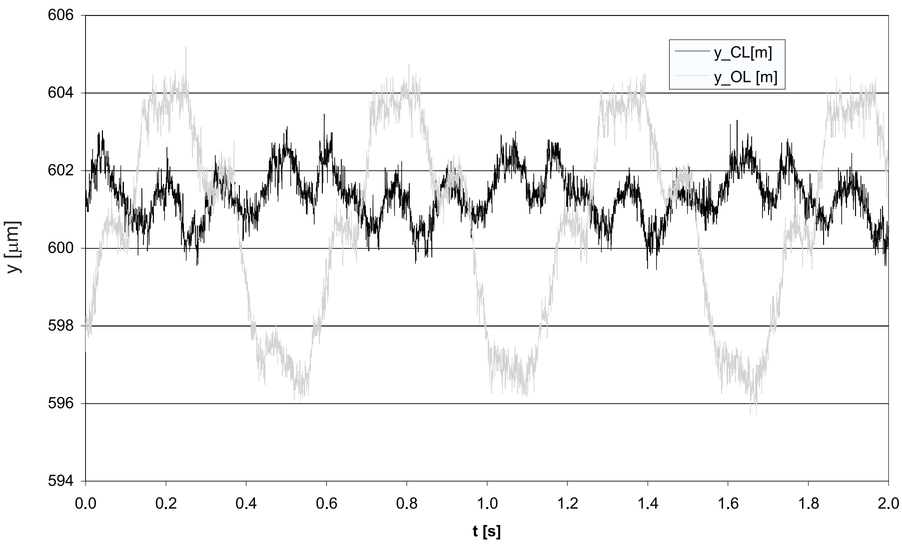


Figure 4.14: Typical time response in the y-direction during cyclic motion in the x-direction ($f \approx 1.8$ Hz, $\delta x \approx 10$ mm), for both the open-loop and the closed-loop situation; the time base is not synchronized. A reduction of the error motions is achieved down to the order of the sensor resolution, being approximately $0.5\ \mu m$ for the photonic sensor's back slope. Note the remaining vibration of about 8 Hz. Probably this is caused by the (unmodeled) piezo hysteresis behavior in combination with the unroundness of the ball bearing elements. The remaining errors are well reproducible so that they possibly can be reduced even further by means of adaptive control techniques.

generated, there is no risk of damaging the sensor tips that are at a distance of only $20 \ldots 40\ \mu$m.

In figure 4.15 the time response is shown for a 10 μm step in the y-direction. The delay or dead zone that is visible in this graph is due to the limited sample time, $t_s = 0.01$ s, of the data collecting program.

From this figure it is clear that the time constant of the system can be estimated to be in the order of 10 ms, corresponding to a bandwidth in the order of 10 to 100 Hz bandwidth. This is in accordance with the measured open-loop bandwidth in figure 4.12.

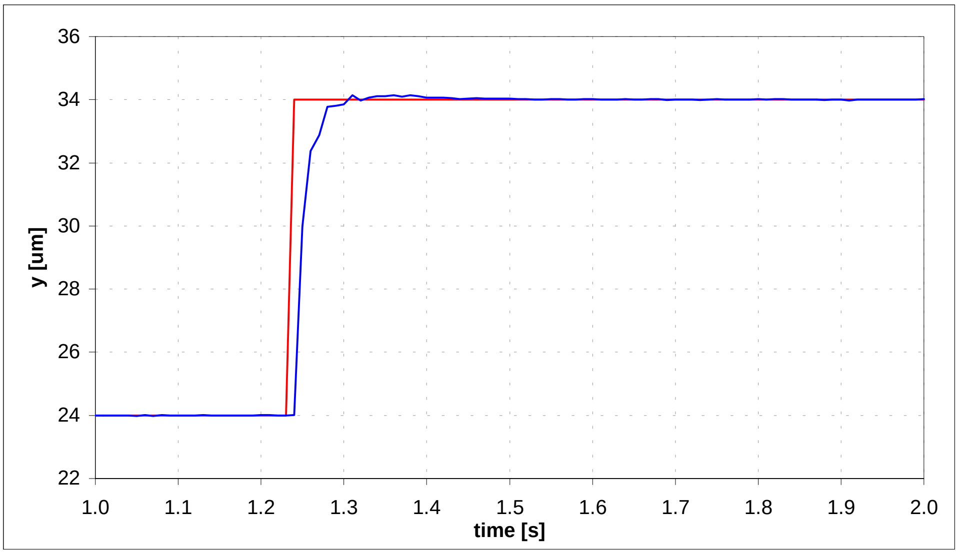


Figure 4.15: Typical step response of the closed-loop system. The delay or dead zone is due to the sample time, $t_s = 0.01$ s. The time constant of the system is in the order of 10 ms. For this measurement the near slope of the photonic sensors has been used, resulting in a position noise level as low as 10 nm, together with a large sensitivity.

4.3 A six degrees of freedom piezo manipulator with long stroke

Based on the same idea of straightness error compensation as described in the previous sections of this chapter, a six DoF piezo manipulator has been designed, built and tested [25]. The goal of this project was to realize a motion system meeting the following requirements:

Linear motion i.e. a long stroke in at least one direction

Inertial Sliding Motion principle, also known as 'piezo stepping'

Low velocity in the order of 1 mm / s

Sub-micrometer position accuracy relative to a straightness reference

Low cost by using no 'exotic' materials, no very accurately machined parts, one single motion principle and shared power supplies

4.3.1 Principle of motion

As a motion principle, the piezo inertial sliding effect has been adopted, abbreviated as ISM (Inertial Sliding Motion). The theoretical and physical background of this phenomenon has been described thoroughly in [77]. Summarized very shortly, motion is generated in the following way (see also figure 4.16):

1. An object to be translated is placed on top of—usually three—bearing elements;
2. These bearing elements each consist of a contact element, usually spherical, connected to a shear mode[2] piezo crystal;
3. By applying a ramp-shaped voltage on the piezo, it will deform proportionally to the input and take the object with it due to (static) friction;
4. At the maximum voltage, the voltage drops to the minimum value with a sufficiently high slew rate, thus causing the piezo to retract rapidly while the object virtually does not move due to its inertia;
5. Step 3 and 4 are repeated, resulting in a linear motion of which the velocity depends on the frequency and amplitude of the input signal.

[2] Depending on the polarization direction and the electrode geometry a piezo element can deform *expanding*, *shearing* or *bending* as a result of a voltage change.

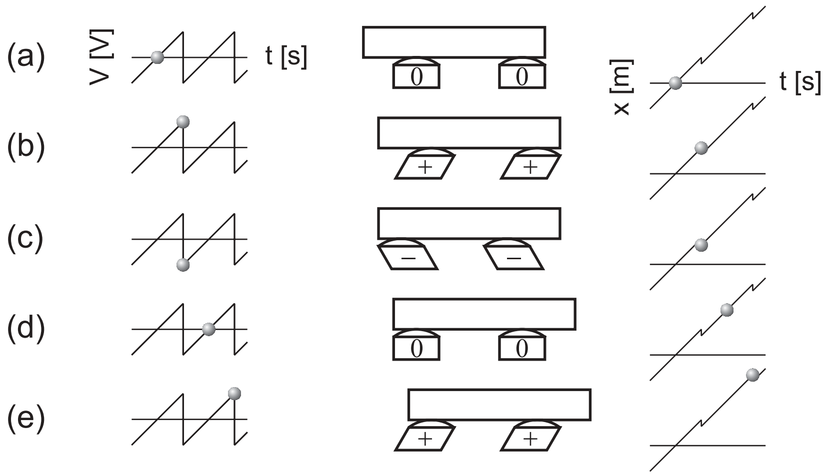


Figure 4.16: Principle of piezo inertial sliding motion (ISM). For clarity the piezo deformations and the position irregularities are exaggerated. (a,b) During the increasing ramp signal the table is moved; (c) During the fast retraction the table does not react much due to inertia; (d,e) Linear motion with virtually constant velocity can be realized.

The most critical moment in this method of driving a table is the moment that the piezo driving voltage changes sign. Under unfavorable circumstances the table velocity can decrease or even change sign for a short period. The main parameters that have influence on the velocity's smoothness, discussed also in [77], are listed below.

1. Material combination of the bearing surfaces
   (a) the difference between *static and dynamic coefficient of friction* of the contact surfaces preferably is as small as possible, because that facilitates the 'release phase' (point (b) in figure 4.16) at the return point of the piezo element.
   (b) the *hardness* of the bearing surface has to be high to reduce plastic deformations that may hinder smooth motion.
   (c) the *stiffness* of the contact region has to be high to improve the relative motion during the 'slip phase' (phase (b→c) in figure 4.16).
   (d) the *wear resistance* of the contact surfaces is related to the hardness. Smooth motion will be hampered by surface damage.
2. Voltage characteristics of the piezo power supply
   (a) the *slew rate* has to be sufficiently high in order to make the 'slip phase' as short as possible.

(b) the *bandwidth* has to be in the kHz range, in order to provide for a 'sharp sawtooth' to make the 'release phase' as short as possible.

Concerning the material choice, it appeared that very good results can be achieved using a combination of a ruby (Al_2O_3) sphere on a TiCN coated steel surface [77]. Because this is rather expensive, in this project only steel has been used on different uncoated surfaces like aluminum and glass, which gave satisfying results, although the 'design margin' appears to be smaller.

The experimental set-up built according to the requirements mentioned above is shown in figure 4.17 and 4.18. Note that the φ_z coordinate is operated by both the bottom and the top stage. Therefore, an additional sensor element will be necessary.

The arrangement of this stacked three-layer set-up gives the most straightforward access to all six degrees of freedom, while the motion range is considerable, as is listed in table 4.4.

DoF	Range
x	limited by baseplate dimension
y	limited by baseplate dimension
z	10 mm
φ_x	45°
φ_y	45°
φ_z	unlimited

Table 4.4: Motion range for the six degrees of freedom of the piezo manipulator.

When recalling the block scheme in figure 4.3, it appears that the design concept of this piezo manipulator is identical with the concept of the experimental linear motion system with straightness error feedback, presented in section 4.2. The most remarkable practical differences are the absence of a carriage or frame, and the fact that the bearing system is an integrated part of the actuators.

4.3.2 Principle of piezo power sharing

As can be concluded from the description of figure 4.17, 15 piezo elements ($[3 \times 2] + [3] + [3 \times 2]$) will be needed to address all degrees of freedom. The required power supplies each have to be capable of generating ±200 V with a very high slew rate, which makes them rather complicated and accordingly expensive. Although the 15 elements can be reduced to 12 groups by combining the elements that do not need to be driven independently, this still is a large number.

Therefore, alternatives were searched for. The most attractive idea was to use two power amplifiers, each capable of driving all 15 piezo elements simultaneously. Both amplifiers have a constant output of a sawtooth-shaped drive signal, but they are opposite in sign in order to enable bidirectional motion. Now each of the piezo

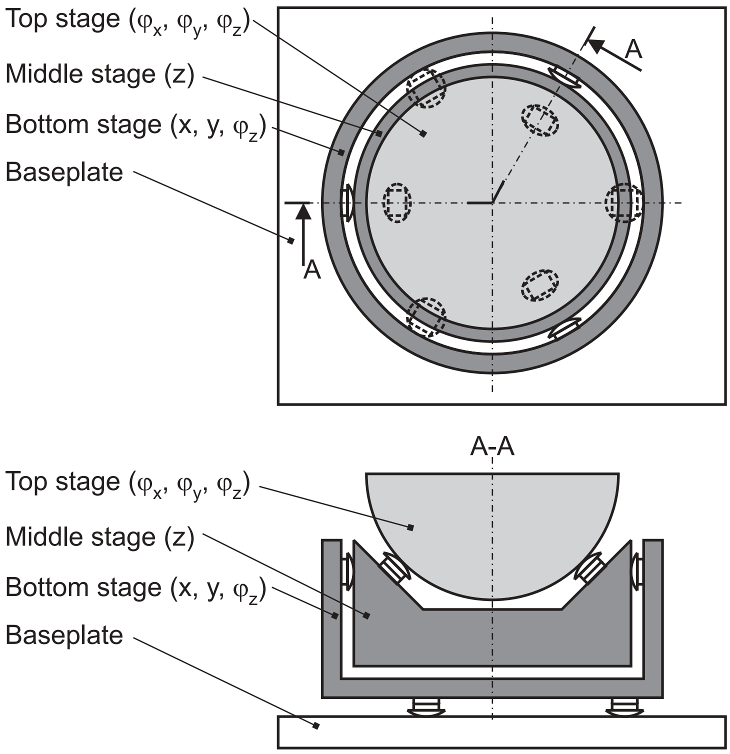


Figure 4.17: Schematical top and side view of the 6-DoF piezo manipulator. The bottom stage is a platform driven by three 'xy' piezo pairs, enabling motion in x-, y- and φ_z-direction. This platform supports the middle stage by clamping it between three 'z' piezo elements that move this middle stage up and down. The top stage is a half sphere supported by another three 'xy' piezo pairs oriented tangentially around the sphere, providing for the rotations φ_x, φ_y and φ_z. For the motion in the horizontal plane and the rotations, the preload is provided by gravity only. The middle stage, which moves in the vertical direction, has an adjustable preload spring (not shown here, see also figure 4.18).

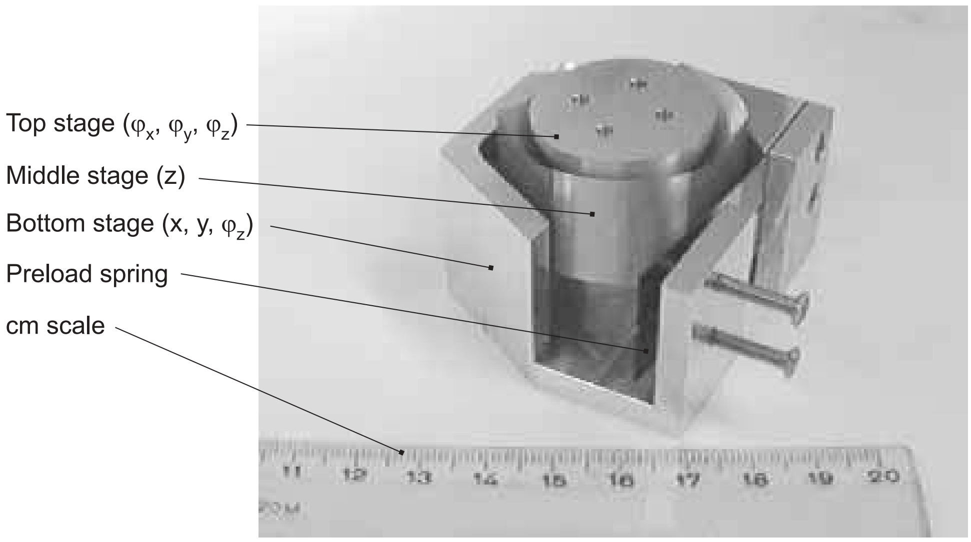


Figure 4.18: Photograph of the complete 6-DoF piezo manipulator, without sensor (to be mounted on the top stage).

elements gets its own pair of switches to make a connection with one of the amplifier outputs or the ground. To simplify the switching scheme and to reduce the requirements on the high-voltage switches, the switching point has been defined at the zero crossing of the sawtooth slope. Thus, the minimum position increment is one step, which limits the resolution. Nevertheless, when the period of the driving signal is chosen sufficiently short, sub-micrometer position resolution is achievable.

4.3.3 Measurement results

For real-time measurement of the straightness deviations of the moving top table, a prototype of the single beam laser straightness sensor has been used (see section 2.2.2.2). For the controller, a very simple algorithm has been applied: the measured and desired position value are fed in a comparator and dependent on the output the drive signal is made positive or negative, with maximum amplitude. This causes the table's position to 'oscillate' with an amplitude of about one step, being approximately 75 nm, but this is not disturbing in practice.

Because of the poor (optical) shielding of the first prototype of the laser sensor, the position noise was about $\pm 2\ \mu$m (peak to peak), which is too high to meet the requirements. However, it was possible to achieve a closed-loop positioning resolution in the order of the sensor resolution, the latter obviously being the performance limiting factor. With improved sensor resolution the position resolution will be improved accordingly, down to the level of the piezo step size.

The velocity achievable strongly depends on the bearing surface's material combination. A spherical steel bearing pad on a flat spring steel plate appeared to give well repeatable results. However, the steel-on-steel contact resulted in surface damage, so that the long-term repeatability was poor. Therefore, a glass surface is recommended. In figure 4.19 a typical open-loop position graph is shown. On the secondary vertical axis, the deviation relative to a fitted first order polynomial is shown.

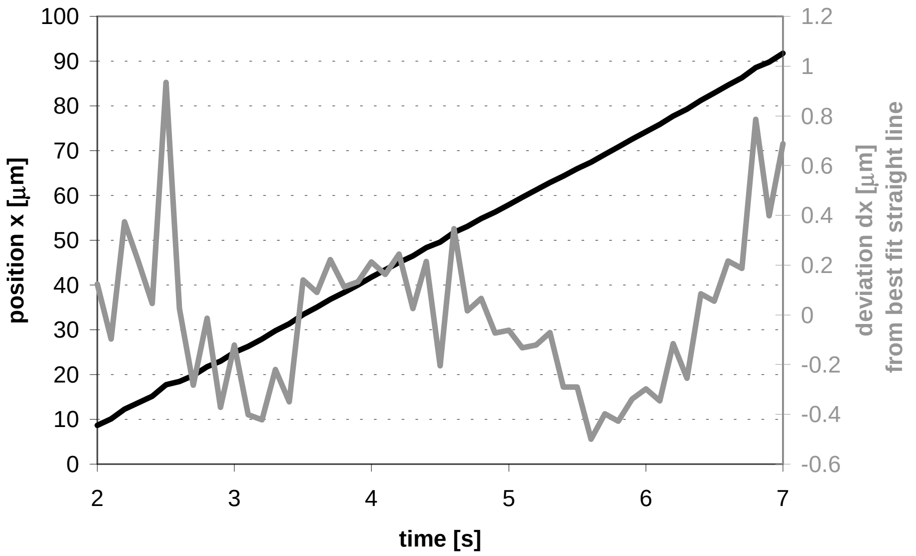


Figure 4.19: A typical measurement of the x-coordinate when the manipulator is driven in the x-direction, without feedback control of the position. The amplitude of the drive signal is 160 V*, the frequency is* 200 Hz*. On the secondary vertical axis (grey) the deviation relative to a fitted first order polynomial is shown. The maximum deviation from the ideal straight line is less than* 1 μm *over a travel of nearly* 100 μm*. This means that the step size repeatability is better than* 1%.

An example of the result of the control system is given in the figures 4.20 and 4.21. Figure 4.20 shows the open-loop measurement of the lateral position error in the y-direction while moving with a low, constant velocity (approximately 10 μm / s) in the x-direction. Clearly visible is the noise from the sensor, but also a non-constant, parasitic velocity in the y-direction is present. Figure 4.21 shows the same measurement, but now the feedback control of y and φ_z is active. The noise remains in the same order of magnitude (since it is caused by the sensor), but the deviations that are not originating from the sensor (e.g. the errors due to the unequality in the step size of the different piezo actuators, or due to variations in the coefficient of friction) are effectively compensated for. Some relevant parameters are summarized in table 4.5.

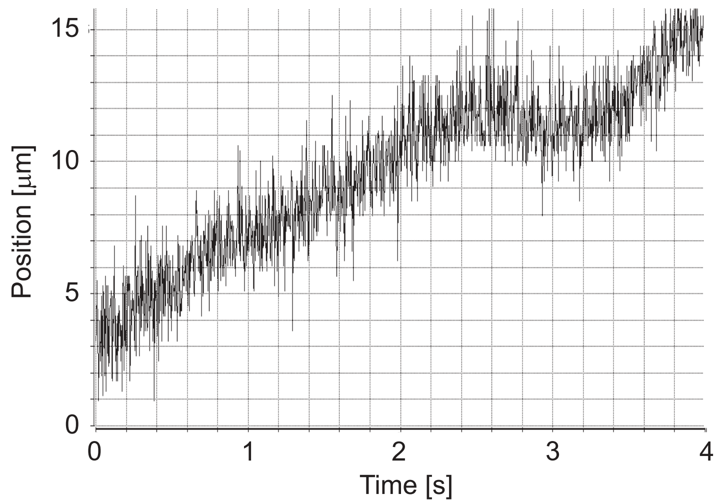


Figure 4.20: Open-loop measurement of the y coordinate while moving with constant velocity (approximately 10 $\mu m/s$) in the x-direction. Clearly visible is the noise from the sensor, but also a non-constant velocity in the y-direction is present.

Parameter	Value	Unit	Remark
Piezo dimensions	$5 \times 3 \times 1$	mm^3	Shear mode
Driving signal amplitude	200	V	Sawtooth-shaped
Driving signal frequency	200	Hz	
Bearing pad material	Steel		on glass surface
Manipulator mass	0.3	kg	approximately
Maximum step size x, y	250	nm	
Maximum velocity x, y	50	μm/s	Nominal velocity 20 μm/s

Table 4.5: Some relevant parameters of the six-DoF piezo manipulator.

The following conclusions have been drawn from these experiments:

1. A modular manipulator has been realized having unlimited stroke in x, y and φ_z, 10 mm in z and 45° in φ_x and φ_y. It can be applied for linear and planar positioning purposes, e.g. in microscopy.

2. Simple bearing surfaces (e.g. steel on steel, steel on glass) are applicable only for a limited range of preloads.

3. Because of the non-optimal material combination the maximum velocity is limited to 50 μm/s.

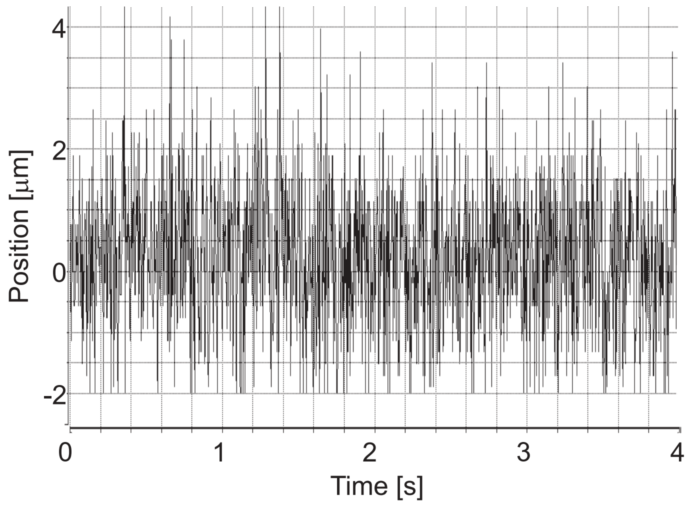


Figure 4.21: Closed-loop measurement of the y coordinate while moving with constant velocity (approximately 10 $\mu m/s$) in the x-direction. The noise remains in the same order of magnitude as in the case of the open-loop situation (since the noise is caused by the sensor) but the deviations that are not originating from the sensor are effectively compensated for.

4. Resolution of the feedback controlled system is in the order of the sensor noise; it must be noted that for the first prototype of the laser sensor the resolution was about ± 2 μm, which does not meet the requirements. After completion of these experiments the resolution was improved with at least one order of magnitude by means of better electrical and optical shielding.

5. When multiple piezo elements have to be driven, it is advantageous to share the power supply by means of high-voltage switches.

4.4 A linear motion system for low velocities

4.4.1 Introduction and concept

In some applications, for example high precision diamond turning or optical disc master writing, there is a need for linear motion systems with very high straightness accuracy at very low velocity, mostly less than 1 mm / s. A list of requirements for a representative application is given in table 4.6.

Parameter	**Value**	**Unit**
Stroke x	150	mm
Travel velocity v_x	$10\dots30$	μm/s
Straightness deviation y	1	nm/mm
Straightness deviation z	250	nm
Payload	25	kg
Disturbance force F_{dist} (up to 75 Hz)	$0.1\dots0.5$	N
Stiffness $k_{y,z}$	$2\cdot10^8$	N/m
Rigid body eigenfrequency	300	Hz
Internal eigenfrequency	500	Hz
Environment	(U)HV	

Table 4.6: Requirements for a high-precision linear motion system, serving as a guideline for the design of the high-viscosity sliding bearing discussed in this section.

From chapter 3 it is clear that no existing bearing type or positioning system directly meets these specifications. Therefore a combined system has been devised, consisting of a *coarse stage* and a *fine stage*. The goal is to make an optimum combination, so that for example the long stroke / low accuracy of one system, together with the short stroke / high accuracy of another system, results in a complete system having long stroke *and* high accuracy.

One possible implementation is the use of a six-DoF short-stroke piezo manipulator—for example, a system like the one presented in section 4.2—on top of a coarse stage having high dynamic stiffness. Besides piezo actuators also electromagnetic actuators might be useful, as will be noted further in this section.

The *coarse stage*, on which this section focuses, has to meet a number of requirements concerning accuracy and bandwidth to ensure that the top stage is capable to eliminate the residual errors. These are:

1. High dynamic stiffness
2. Very smooth motion, no high frequency disturbances
3. Maximum straightness deviation 1 μm

4. Travel velocity $0 \ldots 1$ mm/s

To realize such a coarse stage, a very straightforward approach has been followed. A simple steel sliding bearing has been used, while applying *high viscosity grease* as a lubricant [67]. The most important advantages of grease in relation to the *motion* of the stage are:

1. The viscous behavior of the grease lubrication provides a simple linear relation between the driving force F_x and the travel velocity v_x.

2. High frequency deviations from the ideal momentary position and high frequency straightness errors, that are normally caused by surface roughness and vibrations of the bearing, are eliminated to a great extent by the grease layer, due to the spatial averaging effect and the damping, respectively.

The most important advantages of grease in relation to the *bearing* of the stage are:

1. A low cost lubricant and bearing with a high damping ratio is achieved.

2. The bearing provides for high stiffness at high frequencies and near-zero stiffness at very low frequencies.

3. Easy and low cost assembly of the bearing.

4. No special requirements are needed for the bearing surfaces.

The *overall advantages* of the bearing are:

1. Compact geometry

2. Low cost

3. Flexible design

The *overall disadvantages* of the bearing are:

1. The thickness of the grease layer, although being not very critical, must be controlled within certain limits. A too thin grease layer will affect the smoothness of motion due to roughness of the bearing surfaces and possible stick-slip.

2. Grease is a non-Newtonian fluid. The apparent viscosity depends on the shear rate and shear amplitude.

3. The viscosity of the grease depends on the temperature.

The rheological behavior of greases in a lubricated bearing is a complex problem. Grease is a non-Newtonian mixture of substances, having a viscosity being not constant but depending on e.g. temperature and deformation velocity. When subject to oscillatory shearing of a certain amplitude, the prismatic grease bearing exhibits a non-linear relation between input force and output displacement. At low shear rates (the shear rate is defined as the relative velocity between the bearing surfaces divided by the gap height) the apparent viscosity is high, at high shear rates it is lower. Besides that, the viscosity also depends on the shear amplitude.

A first-order approximation of the behavior of the grease bearing can be obtained by regarding the system as a mass-damper system, having

$$m\ddot{x} + c\dot{x} = F_x \tag{4.21}$$

as the equation of motion in the x-direction. The damping constant c depends on the viscosity and on the geometry of the bearing, according to

$$c = \frac{A}{h}\eta \tag{4.22}$$

where A is the bearing surface, h the grease film thickness and η the dynamic viscosity.

The three important questions that have to be answered are

1. What is the ratio between the driving force F_x and the travel velocity $\dot{x}$ and what is its maximum fluctuation?
2. What is the smoothness of the travel velocity $\dot{x}$?
3. What are the magnitude and smoothness of the lateral straightness errors?

4.4.2 Experiments and results

4.4.2.1 The relation between force and velocity

A first test set-up has been built to investigate the open-loop behavior of the coarse stage concerning the three main questions mentioned above. A principle sketch of this set-up for the measurement in the x-direction is shown in figure 4.22. A steel slider (coarse stage) has been placed against two steel bars, separated by a grease film of approximately 20 μm thickness. The applied grease is called ‘Kilopoise 0001’ having a dynamic viscosity $\eta = 0.25\,\text{kPoise} = 25\ \text{Ns}/\text{m}^2$. A nearly constant force has been applied by means of a hanging mass and a pulley. In figure 4.23 the measured results are displayed. Note that no precautions have been taken to keep the thickness of the grease layer constant. In these experiments only onetime movements have been made.

Care has to be taken with regard to the point where the driving (pulling) force acts, to ensure that the resulting drive force is parallel to the guidance plates. This point has to be on the working line of the resulting friction force, which can be derived from the geometry of the grease layers.

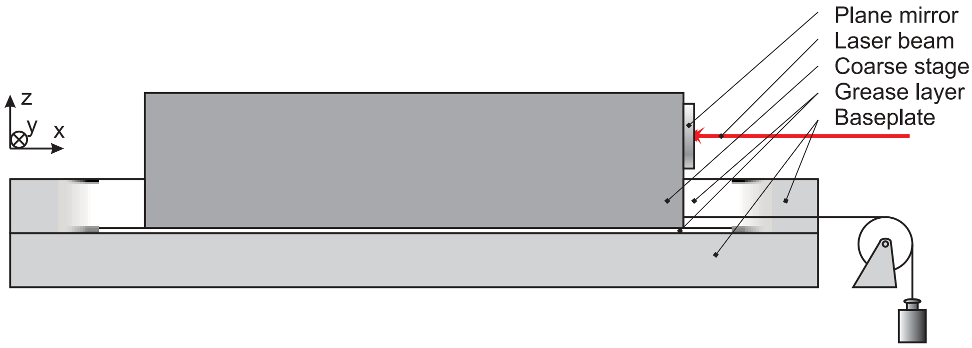


Figure 4.22: Schematic set-up for measurement of the position and velocity of the coarse stage in the x-direction. A steel slider (the coarse stage) has been placed on an L-shaped baseplate, separated by a grease film of approximately 20 μm thickness. A fairly constant force has been applied by means of a hanging mass and a pulley. See also figure 4.26 for a front view of nearly the same set-up; there another mirror is used and the pulley and mass are not shown.

From figure 4.23 it can be concluded that the relation between the applied force and the average travel velocity, F_x/v_x is fairly linear. When a least-squares first order polynomial fit is made, the relation becomes:

$$v_x \approx 47.6 \cdot F_x \quad (F_x \text{ in N}, v_x \text{ in } \mu\text{m/s}) \tag{4.23}$$

The difference between the measured value and the linearization of equation 4.23 is in the order of 10%; for the five last points the relative errors are -12%, -4%, -0.4%, -0.1% and 2%, respectively. This might suggest a rather large non-linearity at lower velocities. However, when the point (0,0) is excluded from the measurements, the fit of the first order polynomial is considerably better. In this case the linearized relation between force and velocity becomes:

$$v_x \approx 51.1 \cdot F_x - 15.3 \quad (F \text{ in N}, v_x \text{ in } \mu\text{m/s}) \tag{4.24}$$

resulting in relative errors of only -0.3%, -0.2%, 0.6%, -0.9% and 0.3%, respectively. The value of F_x where the graph crosses zero is

$$F_{x,v_x=0} = 0.30 \quad [\text{N}] \tag{4.25}$$

This leads to the conclusion that there is a certain elastic behavior, so that the first portion of the pulling force is 'stored' in a displacement rather than a velocity. This

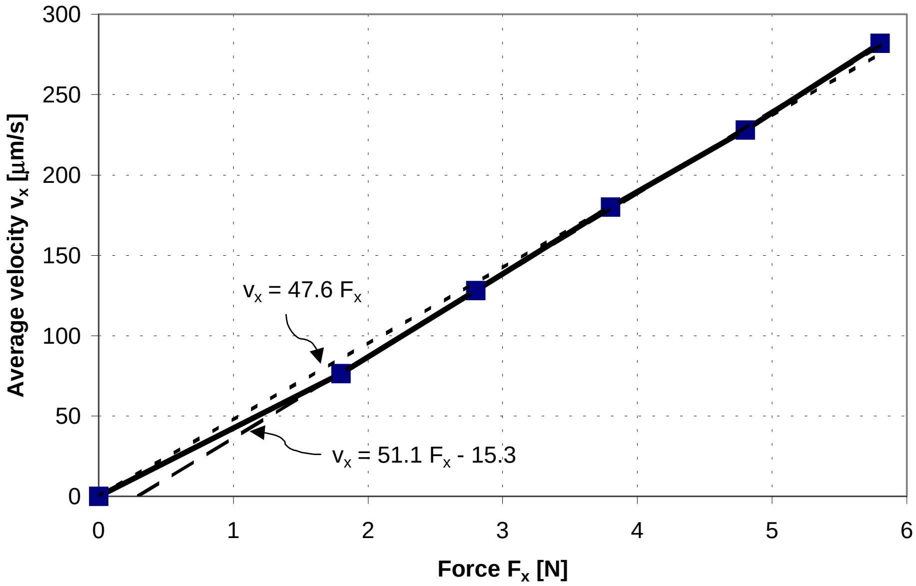


Figure 4.23: Measured average velocity v_x as a function of the applied force F_x. The relation between velocity and force is linear within 10%*, including the point* $(0,0)$*, and within* 1%*, excluding the point* $(0,0)$ *from fitting.*

phenomenon is also known under the name viscoelasticity[3]. This conclusion is further confirmed by a later experiment where the force F_x has been changed stepwise from 2.8 N to zero. The measurement results of this experiment are shown in figure 4.24.

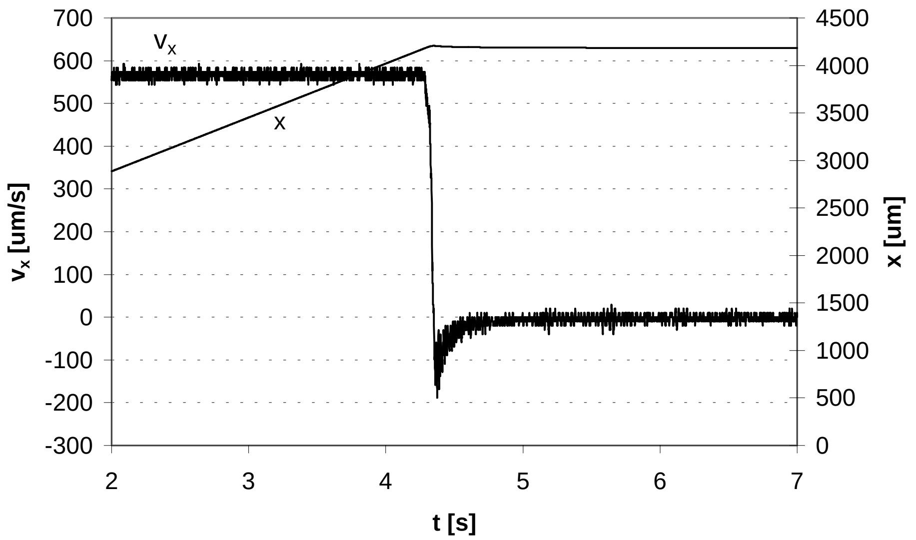


Figure 4.24: The effect of a step input on the travel velocity has been examined by instantly removing the pulling force. The noise in the velocity is caused by the quantization of the sampled x values, from which v has been calculated by numerical differentiation. The 'overshoot' after removal of the pulling force is caused by the viscoelastic behavior of the grease film.

In figure 4.24 is shown that the coarse stage's velocity becomes negative for a short period after removing the pulling force. This is in accordance with viscoelastic theory. After removing the shear force component, there is a certain elastic recovery of the deformation.

In reality, the curve that describes the relation between force and velocity probably will not cross zero at the calculated point, but will bend towards the point $(0, 0)$.

[3]The viscoelastic behavior can be modeled as the series connection of a damper (the viscous part) and a spring (the elastic part). During motion, part of the input energy is stored in the spring, the rest of the energy is dissipated in the damper.

4.4.2.2 The smoothness of the travel velocity

The travel velocity v_x is expected to be constant (assuming the driving force is constant) and smooth (because of the high damping). However, figure 4.24 shows a considerable amount of noise in the velocity. The observed noise on the velocity signal is not inherent to the physical process, but rather to the quantization noise of the position signal, as can be seen in figure 4.25, due to which the first derivative, calculated numerically, becomes much more noisy than the real velocity is. The mean value of the velocity is very constant: the maximum velocity deviation over the period $2 \ldots 4$ s is as low as $\pm 2\ \mu\text{m/s}$. This value has been derived from the maximum slope of the position deviation from a fitted first order polynomial in this region (not displayed).

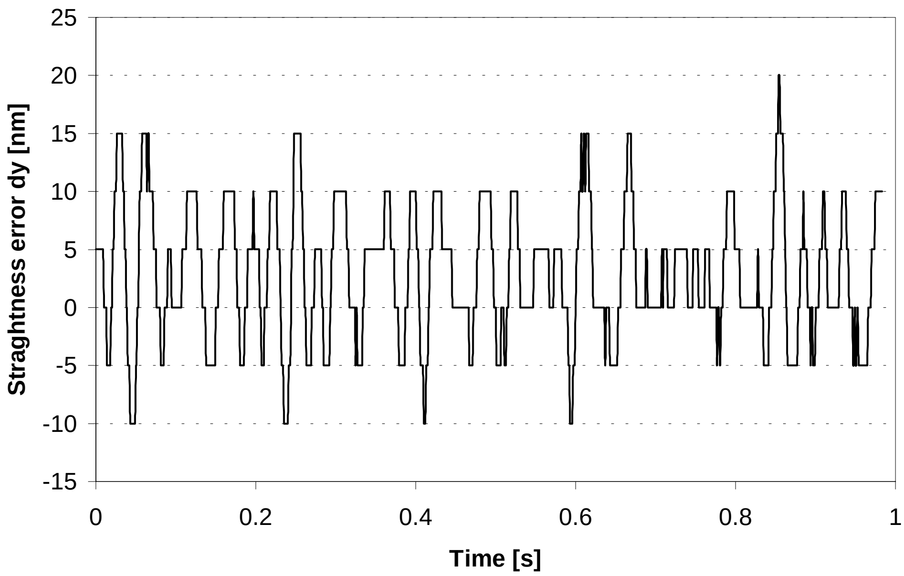


Figure 4.25: Graph showing the quantization effect observed when using a plane mirror laser interferometer in combination with a PC card for the digital data processing. The digital data acquisition causes a truncation of the measurement signal.

The laser interferometer used in these experiments has a standard resolution of 10 nm. When a plane mirror interferometer unit is used, the measured distance has to be divided by 2, resulting in an effective resolution of 5 nm which is clearly recognizable in figure 4.25. The signal from the laser interferometer is fed into a PC by means of a special data acquisition card, resulting in a sampled signal with selectable sample frequency, up to 100 kHz.

The coarse quantization step in the velocity, δv, caused by differentiating such a discrete-time quantized signal, depends on the quantization step in the position, δx, and the sample frequency f_s according to

$$\delta v = \delta x \cdot f_s \tag{4.26}$$

assuming a situation that the measurement signal oscillates around the actual value by one increment. For example, with a position resolution of 10 nm and a sample frequency of 1 kHz, the resolution for the numerically calculated velocity becomes 10 μm / s, being in the same order of magnitude of the actual velocity, 10 to 100 μm / s.

The noise in the calculated velocity can be reduced by low-pass filtering of the measurement values, but in the case of a closed-loop controlled system this may reduce the system's bandwidth.

4.4.2.3 The magnitude and smoothness of the straightness errors

To investigate the straightness errors in the y-direction, a set-up has been built as shown schematically in figure 4.26. A Zerodur straightness reference mirror has been placed on top of the coarse stage. The distance between the moving mirror and the laser interferometer has been measured, giving the lateral displacement of the mirror relative to the interferometer's beam splitter. To obtain the y-coordinate of the coarse stage, this measured displacement of course has to be corrected for the misalignment of the mirror relative to the coarse stage. The flatness of the mirror is specified to be better than $\lambda/8$ or approximately 80 nm, which limits the absolute accuracy.

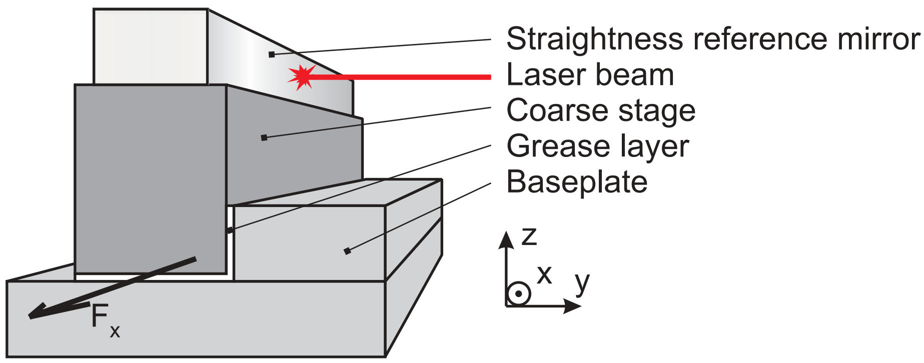


Figure 4.26: Scheme of the straightness measurement set-up for the y-direction. A Zerodur plane mirror has been placed on top of the coarse stage in order to measure the displacements in the y-direction while moving in the x-direction. For the measurement a plane mirror laser interferometer has been used.

Figure 4.27 gives an example of the short-term (10 s) behavior of the coarse stage in the y-direction during motion. It appears that the deviations are very small, in the order of the resolution of the laser interferometer. The smallest possible output step is 5 nm in the case of a plane mirror interferometer, but the resolution is larger, around ±10 nm due to (mainly) atmospheric influences.

Figure 4.28 shows a measurement during two minutes while moving in the x-direction with a velocity $v_x = 75\ \mu$m/s. The measured data has been compared to the best fitting first order polynomial, in order to eliminate alignment errors of the mirror. The difference has been plotted. Also here the straightness errors are only slightly larger than the resolution of the laser interferometer. Therefore only an upper limit can be estimated for the straightness errors.

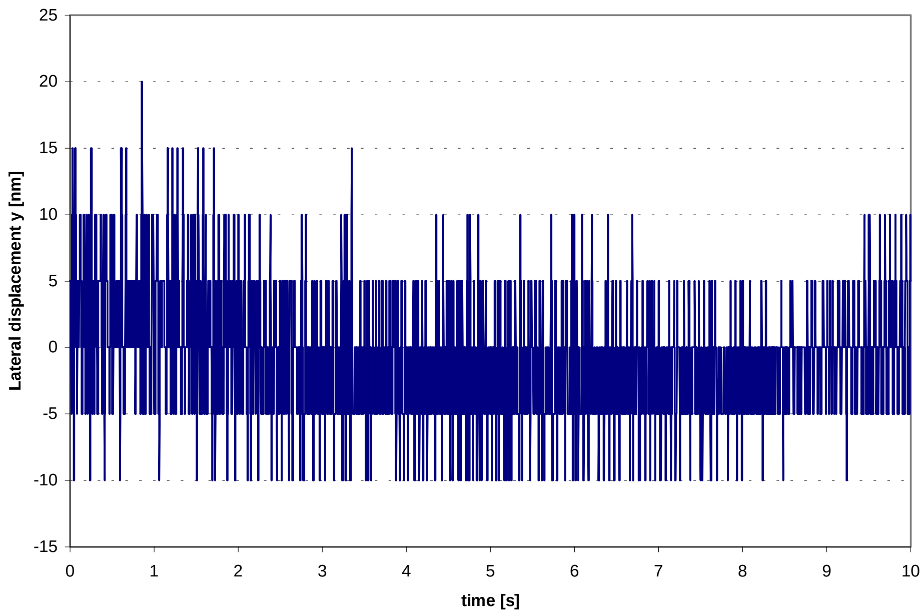


Figure 4.27: Short-term straightness errors in the lateral y-direction during motion in the x-direction. The velocity $v_x \approx 10\ \mu m/s$. The quantization error due to the limited resolution of the laser interferometer is clearly visible. The actual straightness curve is expected to be much smoother due to the averaging effect of the grease film, but because the straightness errors have the same or smaller order of magnitude as the sensor resolution, they are not distinguishable.

4.4.2.4 Summary of measurement results

In table 4.7 the most important measurement data have been summarized. It is clear that excellent straightness and smoothness can be achieved using a high viscosity lubricant.

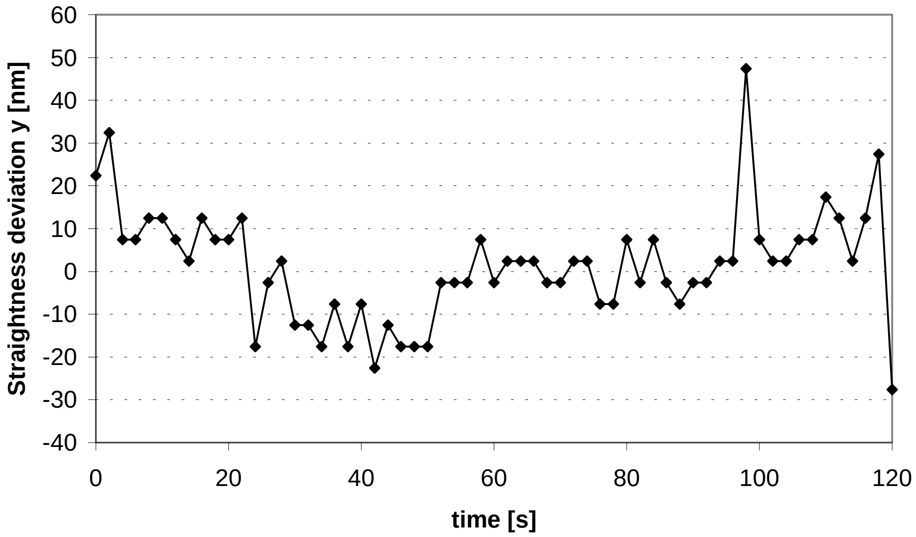


Figure 4.28: Measured straightness deviations in the lateral y-direction during 9 mm motion in the x-direction. The velocity $v_x = \frac{9\ mm}{120\ s} = 75\ \mu m/s$. *Displayed is the deviation from a linear polynomial fit through the data in order to remove the misalignment error. The measured straightness errors appear to be just slightly larger than the resolution of the laser interferometer, but they could even be caused by the flatness error of the mirror, which is specified to be less than* $\lambda/8 \approx 80\ nm$.

It has to be noted that these results may not be very repeatable because the grease film thickness is only estimated and no precautions were taken to keep the film thickness constant. However, the measurements give a fair impression of the order of magnitude of the main parameters that can be expected when designing a grease bearing based on these experiments.

4.4.3 Closed-loop experiments and results

Although the open-loop results look promising, the deviations are still too large to meet the requirements listed in table 4.6. Moreover, they are not well enough predictable. Therefore a closed-loop system will be necessary. There are two basic possibilities to realize this:

1. Use a fine stage on top of the coarse stage, for example a 5- or 6-DoF piezo actuated table. This option has been mentioned already in the start of this section. The specifications of the coarse stage meet the requirements listed on page 110.

Parameter	Value	Unit
Bearing surface	4500	mm^2
Grease film thickness	≈ 20	μm
Propulsion force F_x	≈ 1	N
Travel velocity v_x	40	$\mu m/s$
Linearity F_x/v_x	1	%
Straightness error, δy	± 20	nm
Relative error $\delta y/x$	≈ 10	nm/mm

Table 4.7: Summary of the measurement results obtained with the high viscosity grease bearing in an open-loop situation.

2. Leave the concept of coarse / fine stage and apply actuator forces directly on the coarse stage to achieve the required accuracy.

The first possibility results in a modular, well-predictable system, which is expected to meet the requirements listed in table 4.6. This option will be left out of consideration here.

Instead of that, a test set-up has been built according to figure 4.29, in order to evaluate the closed-loop bearing properties and its dynamical behavior in one degree of freedom. The test set-up consists of a steel guidance that is mounted to the world and a rectangular slider fitting around it in such a way that there remains a gap of 50 μm at all four sides which is filled with grease of high viscosity. This is done by greasing the guidance plate with a layer slightly thicker than 50 μm, and providing the slider with slightly beveled edges (approximately 1 : 10000). The dynamic viscosity of the grease is $\eta = 250$ Poise $= 25$ kg / m·s at 25°C. The upper and the lower part of the slider contain an electromagnet capable of generating a pulling force of about 200 N, so that a bi-directional correction force in the z-direction is available.

On top of the slider a Zerodur plane mirror is placed in order to measure straightness deviations in y resp. z by means of a laser interferometer.

On the front side of the slider a small frame has been mounted in such a way that a mass on a rope via a pulley exerts a constant force through the center of gravity of the slider.

Firstly, an experiment has been done to estimate the time constant of the actuator. In figure 4.30 a series of measured (open-loop) step responses has been displayed. The current through one the coils was switched on and off (corresponding with 1 A and 0 A, respectively), while the height z was measured with a laser interferometer. It is clearly visible that there is a certain elastic response after switching the current on, having a small time constant. After a longer time a relatively constant velocity $v_z \approx 60$ nm / s is achieved due to the viscous shearing effect and the squeeze film effect. The same time constants dominate the behavior after switching off the current again. This behavior agrees with the viscoelastic theory, as already mentioned in section 4.4.2.1. The very low velocity, and corresponding low bandwidth, is caused

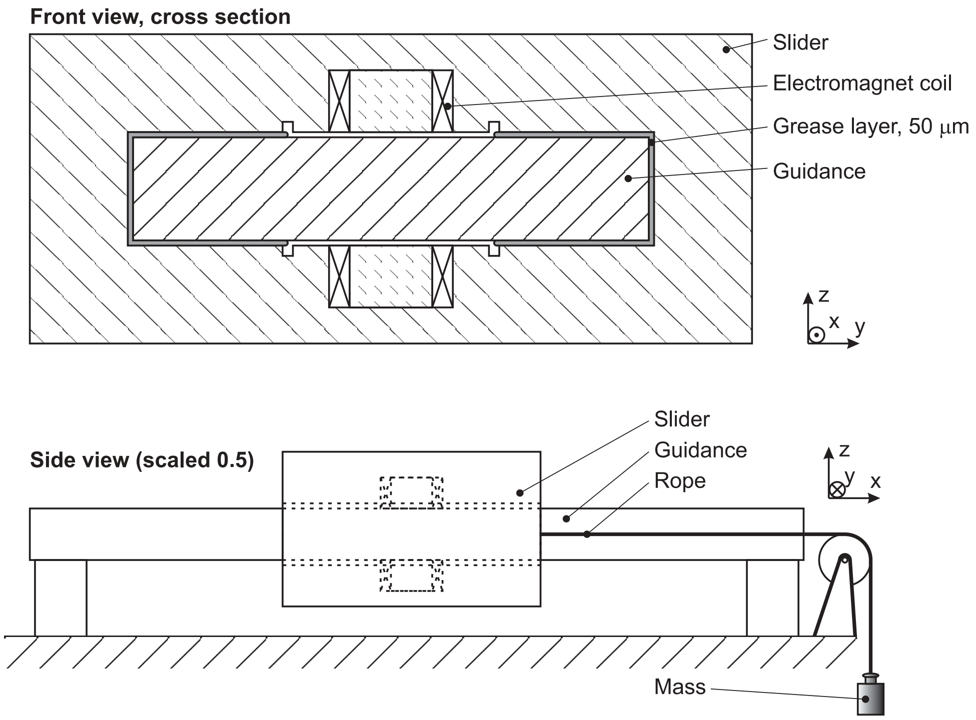


Figure 4.29: Sketch of the test set-up to examine the bearing properties and the dynamical behavior of a high-viscosity sliding bearing. Two electromagnets provide a bi-directional correction force in the z-direction. The guidance dimension is about $300 \times 100 \times 15\ mm^3$.

by the fact that in the z-direction the squeeze effect determines the behavior instead of the shear effect. Because the force, required for squeezing a grease film, is inversely proportional to h^3, while the shearing force is inversely proportional to h, the squeeze damping will be much higher dan the shear damping.

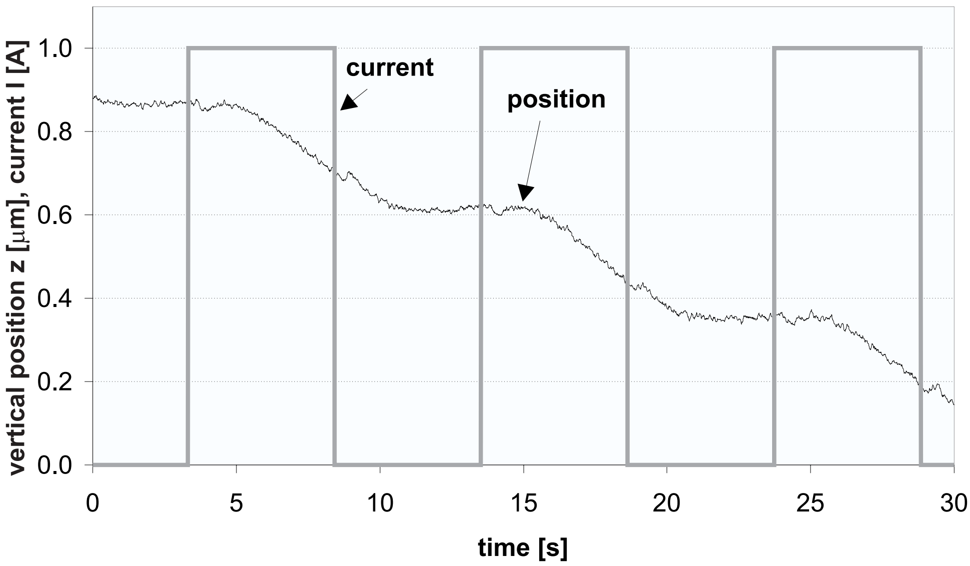


Figure 4.30: Typical time response of the slider in the vertical z-direction as a result of switching the coil current on and off (1 A resp. 0 A). A certain elastic response is visible after switching the current on, having a small time constant. After a longer time a relatively constant velocity $v_z \approx 60$ nm/s is achieved due to the viscous shearing effect and the squeeze film effect. The same time constants dominate the behavior after switching off the current again.

In figure 4.31 a comparison has been made between the open-loop and the closed-loop time response in the z-direction during slow movement, about 50 μm/s, in the x-direction. In the closed-loop situation, the low-frequency position deviations (< 1 Hz) appear to be eliminated. The high-frequency position deviations are slightly amplified because of the simple controller design, but they remain in the same order of magnitude. The reason that there is no large difference between the two curves is that the open-loop signal already has very small errors, just above the noise level of the sensor.

A possible configuration for a straight motion system, controlled in five degrees of freedom, is given in figure 4.32. At least ten electromagnets are necessary for five bi-directional forces, of which only six are shown in this cross section.

Given the dynamical behavior as represented in figure 4.30, it is clear that a system like this test set-up, where the actuator forces directly affect the grease film

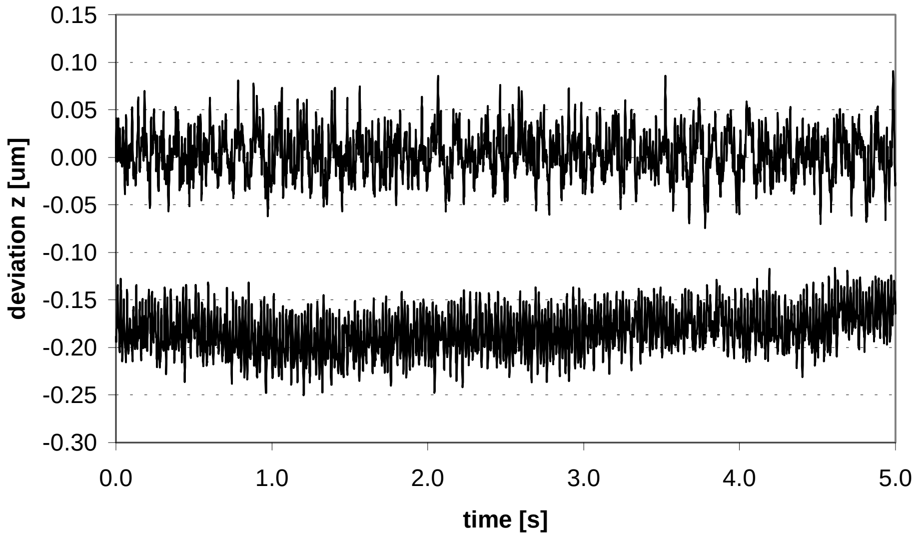


Figure 4.31: Open-loop (lower graph) and closed-loop (upper graph) time response in z-direction. In the closed-loop situation, the low-frequency position deviations ($<$ 1 Hz) appear to be eliminated; the high-frequency position deviations are not reduced but even slightly amplified because of the simple controller design, but they remain in the same order of magnitude.

thickness in the bearing, cannot have a high bandwidth. Therefore, it may only be suitable for quasi-static systems.

For systems requiring a higher bandwidth, a fast fine stage is a better solution. By stacking an accurate fine stage on top of the coarse stage, the low bandwidth of the coarse stage does not influence the performance of the fine stage. The coarse stage limits the maximum velocity in the direction of travel.

Summarizing the properties of the coarse stage utilizing a high viscosity grease bearing, a list of advantages and disadvantages (in high-precision positioning) can be formulated. They are given in table 4.8.

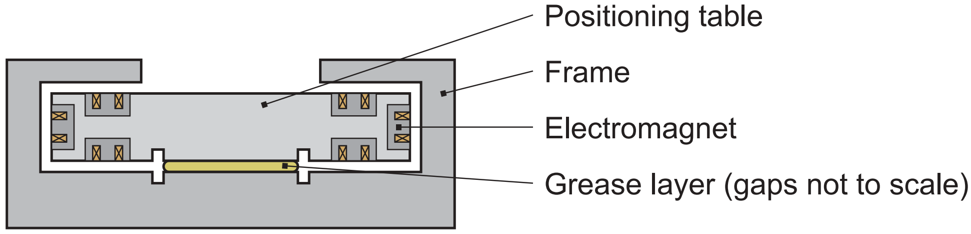


Figure 4.32: Possible configuration for an active linear motion system using a grease bearing to provide very high dynamic stiffness. The electromagnetic actuators provide low-frequency straightness error compensation in five degrees of freedom.

Advantages	**Disadvantages**
High dynamic stiffness	Low velocity applications only
Strong spatial averaging gives smooth motion	Not suitable for UHV; HV possible with dedicated grease
Sub-μm straightness without very accurately machined parts	

Table 4.8: List of advantages and disadvantages for a coarse stage using high viscosity grease.

4.5 Concluding remarks with regard to feedback controlled straight motion systems

1. Open-loop straight motion systems are subject to straightness deviations due to manufacturing errors, bearing irregularities, external forces and vibrations, wear and thermal effects.

2. As a consequence of the increasing requirements on accuracy, feedback control becomes inevitable. The positioning performance of a closed-loop system depends on mechanical properties to a considerably lesser extent.

3. Feedback controlled straight motion systems are able to eliminate at least the quasi-static straightness errors.

4. The performance (in particular the accuracy and bandwidth) of feedback controlled straight motion systems is mainly limited by the sensor resolution and bandwidth and the internal dynamical behavior of the mechanical system.

5. Motion systems based on piezo inertial sliding motion (ISM) are suitable for high-precision positioning applications in multiple degrees of freedom. The choice of materials for the bearing surfaces is not very critical, although it can limit the allowed preload and the maximum velocity, and it will influence long-term repeatability due to wear.

6. High viscosity grease bearings are suitable for low-velocity, high-accuracy applications. They can be used as a coarse stage for an additional fine-stage having a high mechanical bandwidth. With the application of a dedicated grease, they can be used in high vacuum applications as well.

Chapter 5

Zero stiffness actuators in straight motion systems

In precision straight motion system design an interesting trend is visible. Every time a step forward is made towards higher accuracy, there will appear some other, nearer or more remote barrier. This shifting of bottle-necks can be represented as follows:

Mechanical straight motion systems: In general, in mechanical straight motion systems the bottle-necks are the *manufacturing accuracy*, resulting in deterministic errors, and the *mechanical properties*, concerning stochastic errors, e.g. due to disturbance forces. These errors are in the micrometer range, several micrometers up to several tens of micrometers. It appears to be very difficult to realize motion with sub-micrometer straightness accuracy.

Feedback controlled straight motion systems: By adding a control loop to the mechanical system, the actual position can be maintained relative to a stable straightness reference, as was described in chapter 4. Thus, the earlier problems in the mechanical domain are transformed into the electrical and control engineering domain. The performance here is limited by the *sensor resolution*, the *quality of the straightness reference* and the *closed-loop bandwidth* related to the disturbance forces.
By using a sensor system of sufficient quality (i.e. a sensor having a measurement accuracy below the required positioning accuracy, a bandwidth at least ten times the desired mechanical bandwidth, etc.), one encounters the limit of *environmental disturbances*, of which *temperature* and *support vibrations* are the most important. Moreover, not only the support vibrations make demands on control bandwidth, but also the vibrations that are generated internally due to e.g. bearing irregularities, internal motors, etc.

Vibration isolated straight motion systems: Because of these disturbing external and internal vibrations, new actuators have to be developed which make the straight motion system less sensitive to these vibrations. Up to now, most effort has been made in the field of stiff actuators, high-bandwidth control loops and accurate modeling.
A new approach, presented in this chapter, is based on making the actuator's inherent stiffness and damping equal to zero. This results in a *mechanical decoupling* of the supported object with respect to the ground—actually, with respect to everything that is 'under' the actuator, in the block scheme shown in figure 5.1—while maintaining most other properties. This approach wil appear to be advantageous to realize closed-loop controlled positioning systems that are even more independent on mechanical parameters than the feedback controlled positioning systems described in chapter 4. In this chapter, the theory and practice of the new zero stiffness actuators will be described.

5.1 The principle of zero stiffness

5.1.1 The reason for zero stiffness

Before describing the principle of zero stiffness, first some explanation will be given about the reason why actuators having no stiffness are advantageous above the high-stiffness actuators that are widely applied in a variety of positioning systems. To understand this, first consider the block scheme that was already presented in figure 4.3, which is shown here in figure 5.1 in a slightly modified way. The essential difference is that the straightness reference is *isolated* from the world's vibrations.

Actually, the system depicted here applies a metrology frame (the principle is mentioned earlier, in section 2.4.5), in order to separate the 'force path' and the 'position path'. When the straightness reference would be fixed to the world with a high stiffness, then it would be practical to design a positioning system having a stiffness (dynamic stiffness, c.q. control bandwidth) as high as possible. However, the metrology frame principle places the straightness reference outside the force path, so that there can not be a closed force path between the table and the straightness reference. In order to make a virtual position coupling between the table and the straightness reference, forces have to be exerted on the table. The reaction forces are conducted via the force path to the world. This principle is illustrated schematically in figure 5.2.

When a high position accuracy is required, which is mostly the case for this kind of systems, mechanical disturbances from the world necessitate vibration isolation of the positioning system. Environmental vibrations (which behave stochastically) and internal vibrations caused by the bearing system (which are partly deterministic and partly stochastic), will cause position errors of the table.

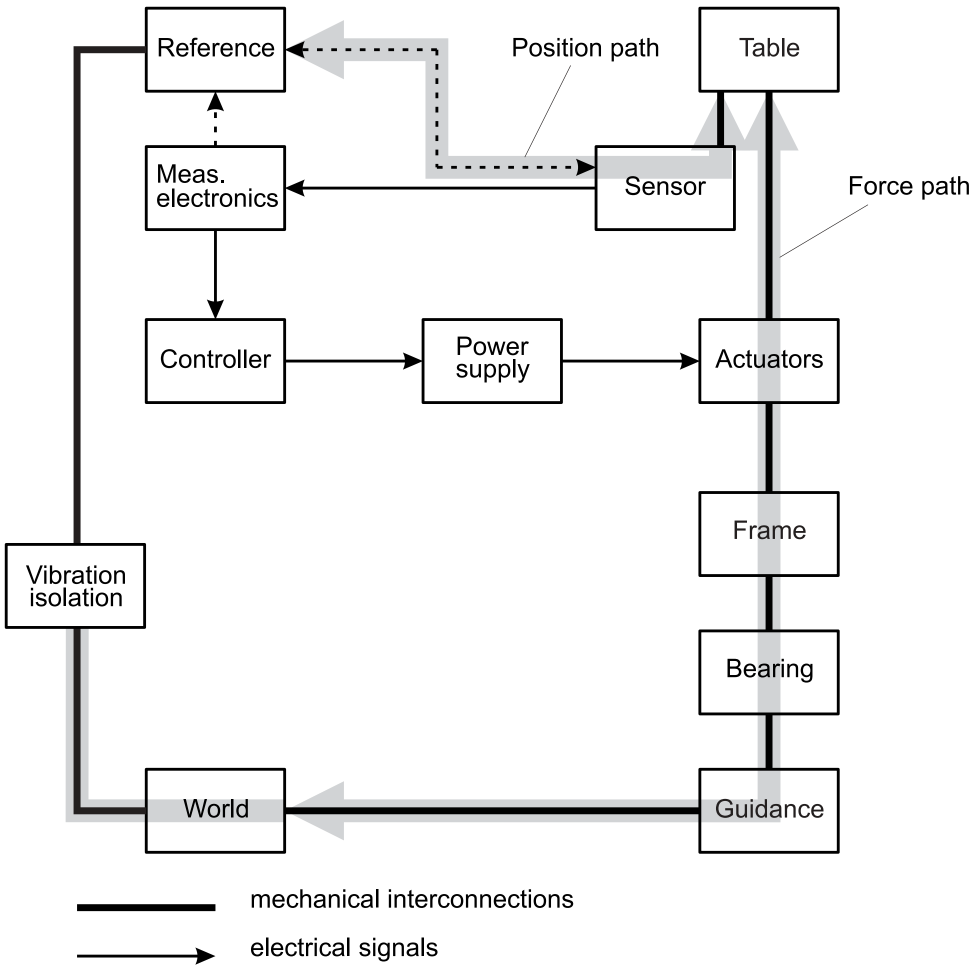


Figure 5.1: Schematical representation of a closed-loop controlled positioning system, in which the straightness reference is isolated from the world's vibrations. Because a stiff positioning system also conducts the vibrations of the world and the bearing, it is advantageous to apply actuators having zero stiffness. In that case, the forces of the actuator system are conducted to the world. At the same time, the position of the table is not influenced by the position deviations of the frame, but can be coupled with a virtual stiffness to the straightness reference.

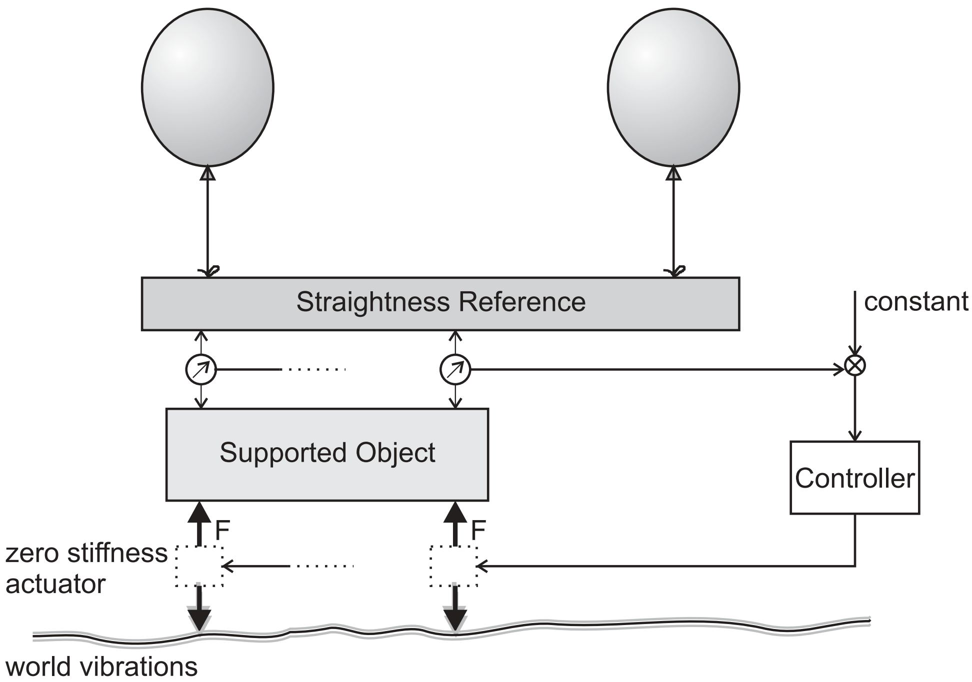


Figure 5.2: Schematical drawing of a positioning system, in accordance with the block scheme in figure 5.1. The supported object is able to follow the straightness reference at a constant distance, while being undisturbed by the vibrations of the supporting world, because its actuator stiffness is equal to zero. N.B.: for clarity, only two of the five sensors and one of the five controllers has been shown.

The common way to *reduce* these errors, by control techniques as described in the chapters 3 and 4, will not be sufficient any more if an accuracy in the nanometer range is required. This is because it is very difficult to realize a practical control system that is able to reduce stochastical external disturbances in the micrometer range with a factor 1000, over a sufficiently large frequency range. In the case that the system's behavior is known very well, a repeatable reduction factor in the order of $10 \ldots 50$ is possible in practice. The only other way to get rid of the influence of these disturbances is to *prevent* them from generating position errors by adequate isolation of the parts that have to be positioned accurately.

In most of the current applications vibration *isolation* has been realized using low stiffness (for example, by long springs or air pods) or vibration *reduction* using passive absorbers or active control theory. However, these systems often tend to be bulky and expensive. Examples of passive and active vibration control systems are summarized in e.g. [26] and [5]. Also in [46] and [47] some versions of electromagnetic vibration isolating systems are described where maximum vibration reduction factors of $10 \ldots 100$ have been achieved. The resulting position stability is in the order of micrometers.

One important remark must be made. With very few exceptions, the passive and active vibration isolation systems described in the literature are based on force generating devices (e.g. springs, electromagnets, and, to a smaller extent, pneumatic suspensions) having a certain *inherent stiffness* which compromises the isolation properties of the complete system.

The larger the inherent stiffness of these actuators is, the more control effort has to be made (in terms of e.g. control bandwidth, knowledge of the physical model and sensor resolution) to reduce vibrations from the outer world. A better approach is to use force generating devices having *zero stiffness* (or almost zero stiffness). In that case vibrations of the outer world are not (or virtually not) transmitted through the bearing c.q. actuator system.

Zero (or almost zero) stiffness actuators can be distinguished by the following three categories of applications:

1. Gravity compensation. The simplest application is the generation of a static bearing force, which must be applied to an object to compensate for a static gravity force $F = mg$, without adding stiffness or damping, to make the supported object virtually 'free floating'.

2. Vibration isolation. Because an actuator unit having zero stiffness is inherently insensitive to the relative distance between the supporting frame and the supported object (within a certain range), vibration isolation is obtained. Actually, this is the same as gravity compensation, with another purpose.

3. Vibration isolated positioning systems. This is an extension to vibration isolation; here the vibration isolating properties are incorporated in the actuators of a feedback controlled positioning system. An application example can be

found in positioning systems in which an object (e.g. a manipulator table) has to be positioned relative to a target (e.g. a lens), where neither the actuator force nor any other disturbing force is allowed to act on the target. This is the case in e.g. IC lithography machines.

The difference between the third category and the first two has important consequences for the control system applied, as will become clear in section 5.4.3.4 and 5.4.3.5.

In order to integrate vibration isolating properties in positioning systems, the approach adopted in this thesis is to design *actuators* having essentially zero stiffness. Thus, the feedback controlled positioning systems as described in chapter 4 can become vibration isolated at the actuator level instead of the world / guidance level (see the block scheme in figure 4.3 and 5.1).

To clarify the principle of zero stiffness, first consider figure 5.3, showing the elementary transfer function of a second order spring-mass-damper system. As an input, the disturbing displacement z_w of the world is assumed, so that its equation of motion becomes

$$m\ddot{z}_m = c(\dot{z}_w - \dot{z}_m) + k(z_w - z_m) \tag{5.1}$$

where k is the stiffness of the suspension and m the mass of the suspended object.

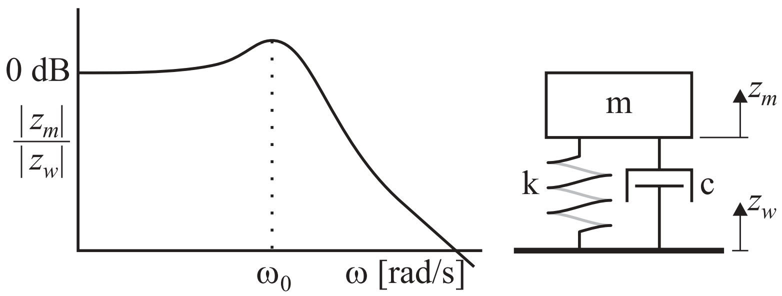


Figure 5.3: Elementary spring-mass-damper system with the transfer function for disturbing displacements. To minimize the sensitivity for external vibrations, ω_0 has to approach zero. This means, either m has to approach infinity (which is impractical), or both the stiffness k and the damping c have to approach zero.

Equation 5.1 can be written as

$$m\ddot{z}_m + c\dot{z}_m + kz_m = c\dot{z}_w + kz_w \tag{5.2}$$

which gives, Laplace transformed:

$$(ms^2 + cs + k)z_m = (cs + k)z_w \tag{5.3}$$

resulting in the transfer function describing the transmissibility of world vibrations z_w to the position z_m of the supported mass

$$\frac{z_m}{z_w} = \frac{cs + k}{ms^2 + cs + k} \tag{5.4}$$

or

$$\frac{z_m}{z_w} = \frac{\frac{c}{m}s + \frac{k}{m}}{s^2 + \frac{c}{m}s + \frac{k}{m}} \tag{5.5}$$

In obtaining vibration isolation, the goal is to minimize the ratio $\frac{|z_m|}{|z_w|}$. Together with equation 5.4 this gives

$$\frac{z_m}{z_w} \to 0\, \forall\, s \Rightarrow m \to \infty \vee (c \to 0 \wedge k \to 0) \tag{5.6}$$

This means that two ways are available to obtain optimum isolation:

1. $m \to \infty$: although mathematically correct, this is not preferable in practice, due to the limited load capacity of the bearing.

2. $c \to 0$ and $k \to 0$: this method will be described further in this chapter.

In figure 5.3 the effect of the conditions $c \to 0$ and $k \to 0$ can be pointed out graphically: With decreasing damping c the resonance peak will become sharper and higher; besides that, the slope at the right of the resonance peak will turn from -20 dB/dec (high c) to -40 dB/dec (low c). With decreasing stiffness k the resonance peak will shift to the left, i.e. ω decreases. Therefore, the vibration transmission will decrease for all frequencies at the right side of the resonance peak. When the limit $c = 0$ and $k = 0$ is reached, the vibration transmission will be zero as well, according to equation 5.4; this means that the mass can be considered as a free floating object; the relative distance between the object and the world does not have any influence on the accelerations of that object.

Realizing $k \approx 0$ is equivalent to constructing a *position independent force*. That means, to create a force $F \neq 0$ with $\partial F/\partial z = 0$ (in figure 5.3: $z = z_m - z_w$).

Summarizing:
To realize a *position independent bearing force* three conditions must be fulfilled.

1. Force equilibrium: $F = m \cdot g$. The bearing *force* must always be equal to the weight of the suspended object, as is common in all bearing systems. But for a position independent bearing force this condition is not sufficient.

2. Neutral equilibrium: $\partial F/\partial z = k = 0$. The force equilibrium has to be neutral, the *stiffness* of the bearing force must be zero in the force equilibrium point. That means that the bearing force must be independent to the position of the bearing element relative to the supporting world.
3. Zero damping: $\partial F/\partial \dot{z} = c = 0$. The *damping* has to be zero as well to prevent disturbing forces due to the velocity of the vibrating support.

From now on, when using the naming 'zero stiffness bearing' or 'zero stiffness actuator' for vibration isolation, also *zero damping* is meant to be included, unless otherwise specified. Furthermore, when referring to 'position independent force' also 'velocity independent force' is meant. In general, the damping effect will be substantially less than the stiffness effect.

5.1.2 The realization of zero stiffness in practice

As was mentioned before, many active systems for vibration *isolation* actually only perform vibration *reduction*, because they are based on non-zero stiffness bearings or actuators. In that case, the eventual performance will be limited by the quality of the control loop and the knowledge of the mechanical model. In section 5.2 a theoretical study [33] is mentioned briefly, showing that controller requirements are very high in order to achieve satisfying properties. To achieve better accuracy, it will be necessary to overcome the inherent mechanical coupling with the environment.

There are several options to achieve an *inherent* zero stiffness bearing force relative to a support (see also figure 5.4 through 5.9):

1. Combination of mechanical springs with positive and negative stiffness, illustrated in figure 5.4. Disadvantages are the small assembly tolerances and the limited bandwidth due to the internal resonant frequencies of the so-called *slinky modes* of the springs (usually between 10 Hz and 100 Hz for plate springs). Examples are given in [16], [18] and [56].

2. Plate spring combinations having linearly changing bending energy E_b with a displacement z, so that the force $F_z = \frac{\partial E_b}{\partial z}$ is constant and the stiffness $c = \frac{\partial F_z}{\partial z} = 0$. An example of curved triangular plate springs is given in figure 5.5.

3. Rolamite[1]–like rolling contact systems as displayed in figure 5.6. Virtually any desired force–displacement relation can be achieved by geometrical variations of the metal band, the preload or of the guidance parts [76]. Actually, this can be regarded as an extension of the former principle dealing with curved plate springs.

[1]Rolamite is a registered trade mark.

4. Counterweight balancing as shown in figure 5.7. This method does not meet the needs of most bearing systems and does not give vibration isolation, although the suspended object is in a neutral equilibrium. In other words, counterweight balancing is suitable for gravity compensation but not for vibration isolation.

5. Lorentz actuator, also known as 'voice coil actuator'. The Lorentz force is defined as $\vec{F_L} = \oint_C I d\vec{l} \times \vec{B}$, where C is the current-carrying conductor, $d\vec{l}$ is the length of a small part of the conductor and $\vec{B}$ is the magnetic flux density at the location of $d\vec{l}$. When B is constant for all elements $d\vec{l}$, as is approximately the case in most voice coil actuators, this can be written as $F_L = B \cdot l \cdot I$ so that $\partial F/\partial z = 0$[2], provided that $\partial B/\partial z = 0$[3], see figure 5.8. Unlike the principles mentioned above, this one is contactless and therefore the internal eigenfrequencies will not affect the quality of the vibration isolation. However, for larger bearing forces, energy dissipation and thus thermal problems become substantial. Still this option can be interesting for some applications, as will be described in section 5.3.

6. Permanent magnet bearing unit, as depicted schematically in figure 5.9. This option uses compensation of positive and negative stiffness and is also contactless. A new construction element based on this principle will be described in section 5.4. The design of a linear motion system using such a zero stiffness permanent magnetic bearing element will be described in section 5.5.

7. Electrostatic bearing unit, analogous to the permanent magnetic bearing unit mentioned above. Instead of repulsing and attracting magnets, repulsing and attracting electrostatically charged objects might be applied, resulting in an analogous behavior. This variant will be left out of consideration in this chapter; its practical implementation will show large similarity to the permanent magnetic case, although the generated forces will be considerably smaller.

[2]In practice, there is a side effect due tot he presence of ferromagnetic yoke parts. Besides the Lorentz force a reluctance force will be generated, which is proportional to $\frac{I^2}{d^2}$, where I is the current through the coil and d is the distance between the coil and the yoke.

[3]In practice, there is the additional requirement that the generated current I is independent on the voltage V because in the voice coil an induced voltage will be generated proportional to the velocity $\partial z/\partial t$. This implies the use of powerful and stable current amplifiers with a clear claim on a high output impedance.

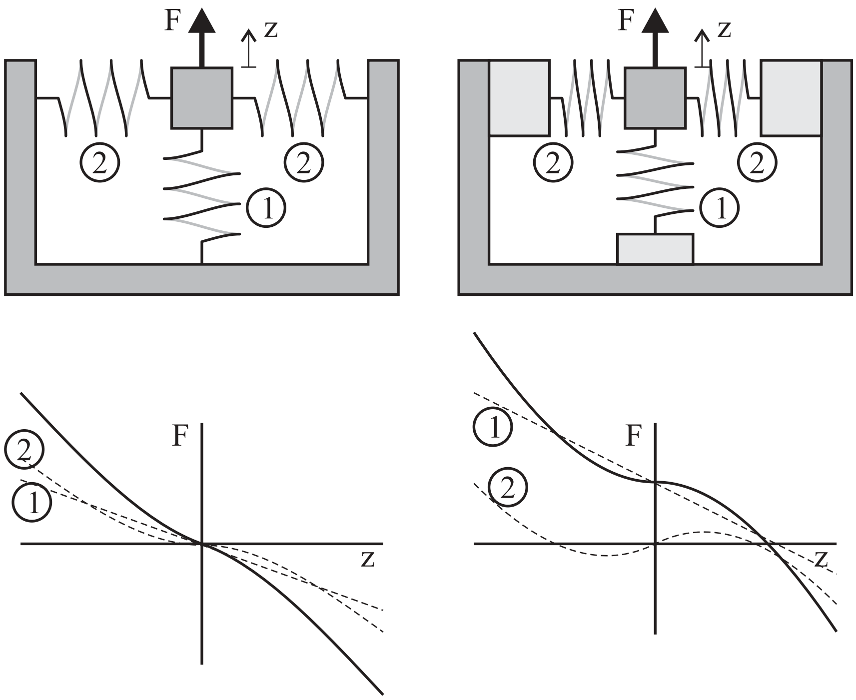


Figure 5.4: Combination of springs with positive and negative stiffness. F is the force delivered by the springs. Left pictures: all springs are unloaded. Right pictures: the preload of the vertical spring 1 determines the net bearing force F. The preload of the horizontal springs 2 causes a negative stiffness in the vertical direction. This negative stiffness must be tuned to match the positive stiffness of spring 1. In the two lower graphs the corresponding force–displacement relation is shown for the springs 1 and 2 with a dashed line, and for the total force with a solid line.

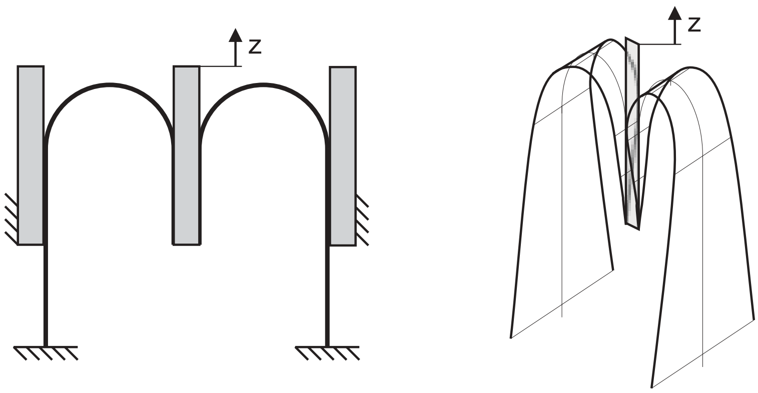


Figure 5.5: As a result of the displacement z the curved plate springs will deform, but with a constant bending radius. Due to their triangular shape the stored bending energy E_b will vary linearly with the displacement z. This causes the force $F_z = \frac{\partial E_b}{\partial z}$ to be constant, so that the stiffness $k_z = \frac{\partial F_z}{\partial z} = 0$. Left: cross section; right: 3D view.

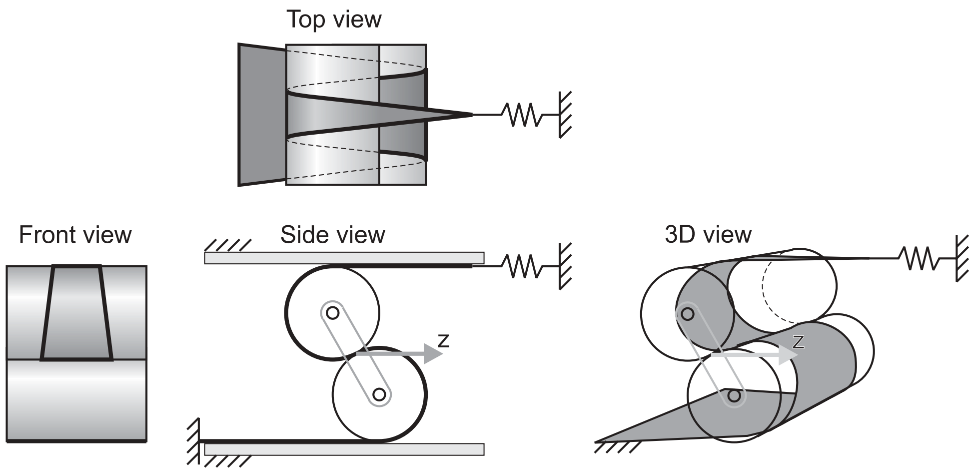


Figure 5.6: Rolamite rolling contact system. By choosing the geometry of the metal bands, the magnitude of the preload force or the shape of the guiding parts (the latter shown only in the side view), virtually every force–displacement characteristic can be achieved. This principle uses the elastic energy stored in the triangular springs.

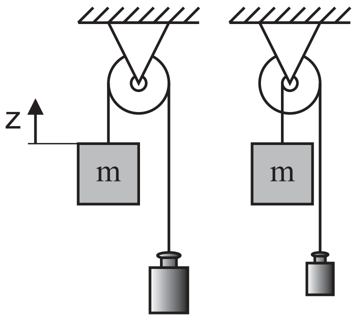


Figure 5.7: Counterweight balancing. The weight of the mass, m, is balanced by a counterweight so that a neutral equilibrium exists for the vertical position z of the mass. However, there is no isolation for vibrations from the supporting world, although counterweight balancing can be regarded as a way of gravity compensation.

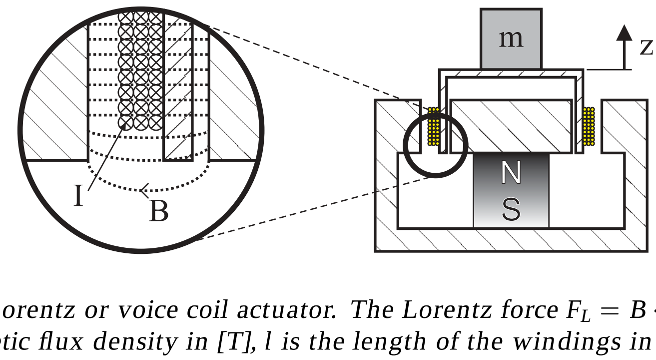


Figure 5.8: Lorentz or voice coil actuator. The Lorentz force $F_L = B \cdot l \cdot I$, where B is the magnetic flux density in [T], l is the length of the windings in the air gap in [m], and I is the current through the windings in [A]. The force is independent on z as long as $\frac{\partial B}{\partial z} = 0$. This actuator is contactless and does not introduce additional eigenfrequencies, except for its own rigid-body eigenfrequencies. However, when suspending loads statically, heat dissipation can become a problem.

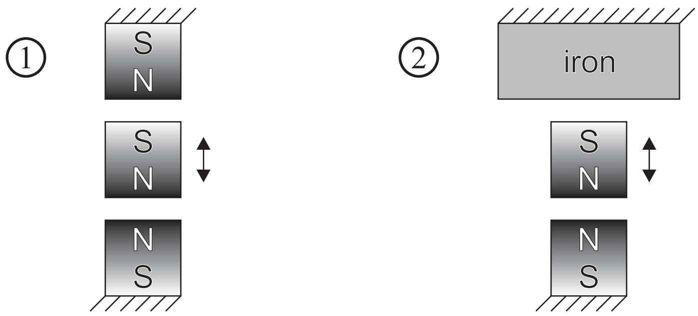


Figure 5.9: Zero stiffness permanent magnetic bearing. An upward force with positive stiffness is exerted on the middle —suspended— magnet by the lower magnet having opposite polarity. The upper magnet (1) or the upper iron plate (2) provide an attracting force on the suspended magnet, upward as well, with negative stiffness. In a certain point the net stiffness will be zero.

5.2 A stiffness compensated reluctance actuator

As was announced already above, this section will only shortly describe the considerations in using an actuator *without* zero stiffness behavior, where a control system is applied to suppress the unwanted properties and enforce the desired behavior, i.e. a position independent force.

In [33] a mainly theoretical study is described on the stiffness compensation of reluctance actuators. A basic reluctance actuator is shown in figure 5.10. It consists of an iron E-shaped core with a coil wound around the middle yoke part and another iron yoke part placed opposite to it with a small (sub-mm) air gap. The goal of this research was to investigate the possibility of using such a reluctance actuator as a zero stiffness actuator.

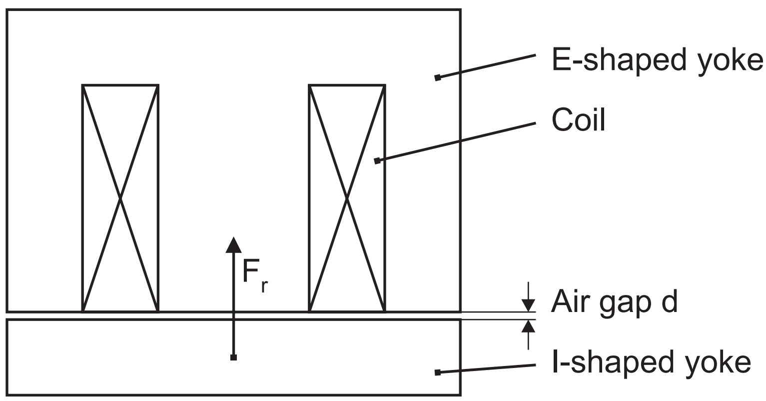


Figure 5.10: Sketch of a basic reluctance actuator. It consists of an iron E-shaped core with a coil wound around the middle yoke part and another I-shaped iron yoke part placed opposite to it with a small (about 1 mm*) air gap.*

The advantage of reluctance actuators above e.g. Lorentz actuators (that already have zero stiffness characteristics) is their *energy efficiency*, expressed as the ratio between maximum force and coil volume in [N / m^3].

The disadvantage of reluctance actuators in zero stiffness applications is their *inherent negative stiffness*. The fact is that the relation between the attracting reluctance force F_r and the air gap or distance d is strongly non-linear. This will make non-linear control algorithms necessary, unless a coarse linearization is applied. But this would limit the performance too much.

Assuming no stray flux, the reluctance force is given in [3] by

$$F_r = \frac{\frac{\mu_0}{\mu_{r,air}} n^2 I^2 A}{(\frac{2d}{\mu_{r,air}} + \frac{\ell_{iron}}{\mu_{r,iron}})^2} \tag{5.7}$$

in which n is the number of turns of the coil, I the current through it, A the pole surface, μ_0 and μ_r are the magnetic permeability of vacuum and the relative magnetic permeability of the medium (iron or air), respectively, d is the air gap and ℓ_{iron} the length of the flux path through the iron yoke.

Because $\mu_{r,iron} \gg \mu_{r,air}$ this may be simplified to

$$F_r = \frac{\mu_0 \mu_{r,air} n^2 I^2 A}{4d^2} \tag{5.8}$$

from which it is clear that the reluctance force F_r is inversely proportional to the air gap d squared. This is a highly non-linear characteristic.

At first sight it does not seem a big problem to embed such an actuator in a control loop, such that the non-linear behaviour is cancelled and the desired closed-loop behaviour (c.q. zero stiffness or a position independent force) is achieved.

However, in practice this is not simple. Firstly, the control algorithm leads to complicated equations which necessitate the use of fast Digital Signal Processors (DSP's). Secondly, the parameters that are neglected to simplify the model equations (such as stray flux, non-parallellity of the yokes etc.) still have influence on the final performance of the closed-loop actuator system.

When it is assumed that the negative stiffness of a certain reluctance actuator is in the order of magnitude $10^3 \ldots 10^4$ N / m and the desired stiffness is as low as $10 \ldots 100$ N / m, then the original actuator characteristics have to be reduced with roughly a factor 100. That means that the model parameters have to be known with less than 1% uncertainty. This appears to be difficult to achieve in practice. Advanced control techniques might be a solution for this problem, but this is not investigated within the scope of this thesis.

5.3 A six degree-of-freedom Lorentz manipulator

By replacing the actuators in figure 5.1 by actuators having zero stiffness the problems due to external vibrations of the support and velocity induced guidance errors will be eliminated. In the ideal case a force controlled system results that is fully decoupled from environmental position disturbances. Actually, the actuators form a vibration isolation system, although they are able to exert control forces on the table.

To examine the possibilities of zero-stiffness actuators in a feedback controlled positioning system, a 6-DoF Lorentz stage has been built and tested. The choice for Lorentz actuators is motivated already in section 5.1.2. The heat dissipation when a static force has to be delivered, which is an unfavorable property of Lorentz actuators, has to be kept as small as possible by choosing a low table mass. For proving the feasibility of the vibration isolating principle this will not be a major problem.

The experimental six-DoF Lorentz manipulator has to meet the following specifications:

1. Vibration reduction of at least 40 dB at all frequencies
2. Purely passive damping above 100 Hz
3. Position stability in the order of the sensor resolution, i.e. 0.1 μm
4. 'Moving magnet' configuration, no cables attached to the table
5. Stroke ± 0.5 mm in x, y and z, 10 mrad in $\varphi_{x,y,z}$

All experimental results in this section are taken from [6].

The experimental stage consists of six Lorentz actuators (voice coils) of which the magnet and yoke parts are connected to the suspended table so that it can float contactless, without connecting wires. Figure 5.11, 5.12 and 5.13 give an overview of the experimental set-up. For the horizontal actuators NdFeB magnets have been used with dimensions $\varnothing 9.75 \times 2$ mm. The vertical actuators each use three of these magnets, because they have to provide for a control force *and* a bearing force.

The coils are made using 0.45 mm diameter copper wire. The vertical coil holders have 170 windings, the horizontal ones 80, which results in a coil resistance of 0.67 and 0.30 Ω, respectively, and an inductivity of 233 and 26.7 μH, respectively.

Each Lorentz actuator has its own axial position sensor (an eddy current sensor), placed on the ground plate and measuring the distance to the table contactless.

The six sensor signals represent six local positions of the suspended table. A coordinate transform has been applied to translate these signals into six orthogonal coordinates (x,y,z,φ_x,φ_y,φ_z). This makes it possible to design a controller consisting of six independent controllers, one for each coordinate, instead of applying much

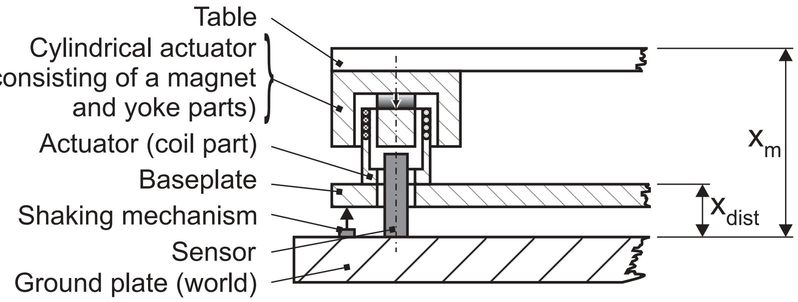


Figure 5.11: Schematical and simplified scheme of the six-DoF Lorentz manipulator. Only one actuator is shown; in reality, the table is supported by six actuators, three in the vertical plane and three in the horizontal plane. The table position is measured relative to the world (ground plate) while the actuators are mounted on an intermediate frame (base plate) that is moveable relative to the ground plate by means of a shaking mechanism, for testing purposes.

more complex multivariable control techniques. In [6] a more detailed description of the system is given.

Due to the constant Lorentz force a perfect 'passive' vibration isolation is expected. However, in practice there are two effects that cause a small non-zero effect:

1. The magnetic field is not perfectly homogeneous, so that a position-dependent force deviation occurs;

2. Besides the Lorentz force, there appears a reluctance force between the coil (acting like a magnet) and the ferromagnetic yoke. This reluctance force causes a small negative stiffness that is dependent on the current through the coil, analogous to equation 5.8.

For these two reasons, the vibration isolation for lower frequencies will be slightly less effective. To overcome these non-idealities, active control will be necessary in the lower frequency range.

In figure 5.14 the block scheme of the closed-loop controlled system is shown, for one degree of freedom. The actuator model has been composed of two parts: a part depending on the input current and a part depending on the position of the table relative to the baseplate. The transfer function of the actuator can be written as

$$x_m(s) = G_1(s)x_{dist}(s) + G_2(s)i(s) \tag{5.9}$$

in which x_m is the position of the suspended mass relative to the imaginary, rigid and immovable world, and x_{dist} is the position of the baseplate with respect to the world. From simple mechanical considerations it follows that the transfer function

Figure 5.12: Overview of the test set-up for the six degrees of freedom Lorentz stage. Clearly visible are the six cylindrical yoke parts holding the magnets, mounted on the star-shaped, freely floating table. Three vertical actuators provide forces in z, φ_x and φ_y, three horizontal actuators act in the x, y, φ_z plane. Three sensor holders for the horizontal position sensors are visible, the three vertical sensors, located under the baseplate, are not visible. The baseplate is hinged, so that vibrations relative to the lower ground plate can be introduced by means of a lever (top right) and a small motor equipped with a cam shaft (not shown).

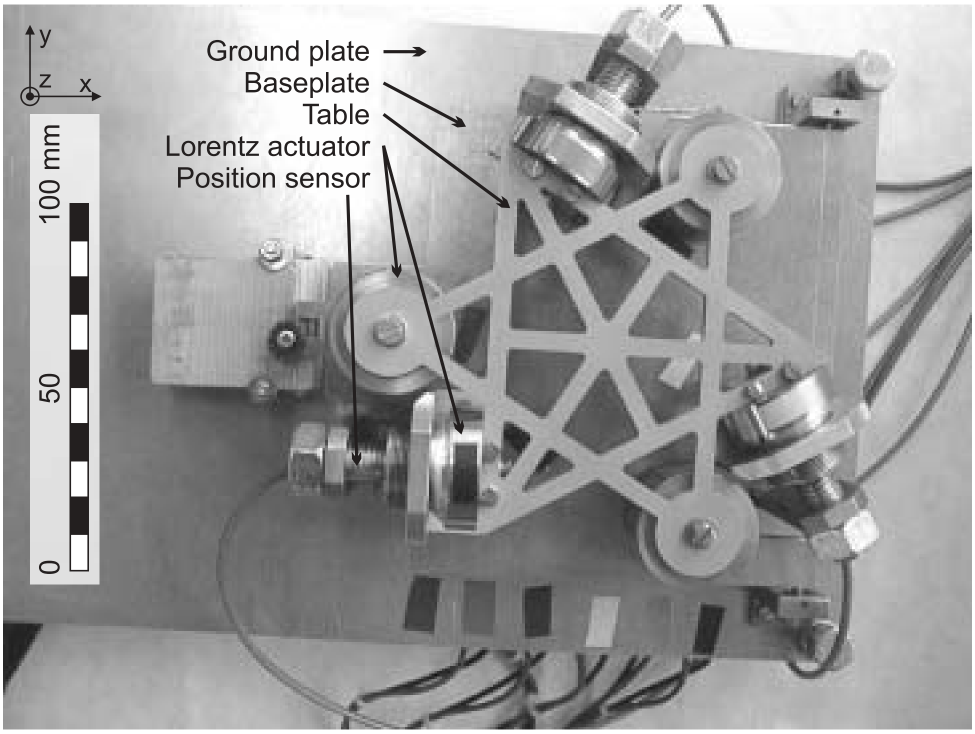


Figure 5.13: Photograph of the test set-up for the six degrees of freedom Lorentz stage. Top view. Note that the difference between the ground plate and the base plate is not very well visible.

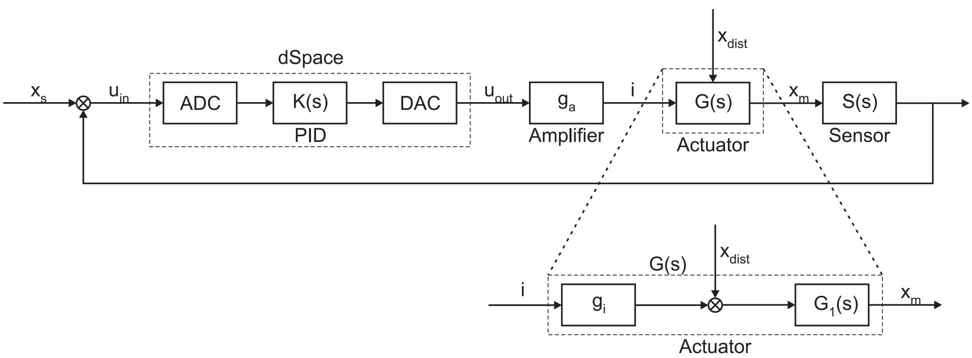


Figure 5.14: Block scheme of the closed-loop controlled system, for one degree of freedom. The actuator, shown in more detail in the lower figure, delivers a force that depends not only on the current i, but also on the relative position x_{dist} between the baseplate and the table.

$G_1(s)$ will look like

$$G_1(s) = \frac{\omega_1^2}{s^2 + \omega_1^2} \qquad \text{where} \qquad \omega_1 = \sqrt{\frac{k}{m}} \tag{5.10}$$

But the actuator stiffness k is a function of both $x = x_m - x_{dist}$ and i. Therefore, the expressions relating the actuator stiffness k to the position and the current have to be examined.

In figure 5.15 a sketch is given of the actual and linearized curves of the force as a function of position and current.

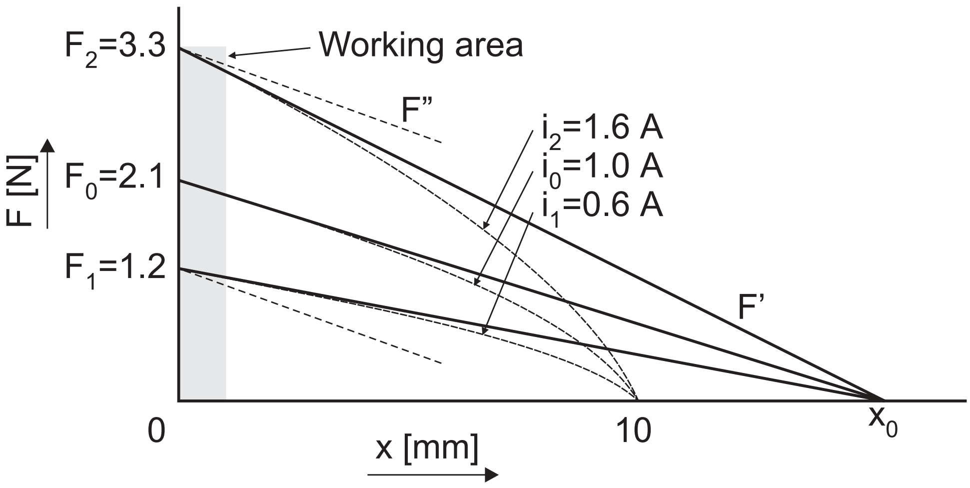


Figure 5.15: Schematic representation of the measured Lorentz actuator force as a function of position and current. The dashed curves represent the actual force–displacement curves for three different currents. The solid lines show the linearization for the position x, the dashed straight lines depict the linearization of the solid lines for the current i. In this graph, the values for x, i and F are only indicative, in order to give an idea of the order of magnitude.

It appears that in the working range the force–position relation can be linearized without large errors. When in figure 5.15 this linearization is executed, the stiffness k becomes a function of only the current i:

$$k(x,i) \approx k'(i) = k_0 \frac{i}{i_0} \tag{5.11}$$

which gives for the force

$$F(x,i) \approx (F_0 - k_0 x)\frac{i}{i_0} \tag{5.12}$$

The point x_0 where the force is zero, is defined as

$$x_0 = \frac{F_0}{k_0} \tag{5.13}$$

which is independent on i. When this point x_0 is sufficiently far to the right of the working area, the linearized forces can be approximated by parallels to $F'(i = i_0)$, which leads to the linear function F''

$$F(x,i) \approx F''(x,i) = F_0 \frac{i}{i_0} - k_0 x \tag{5.14}$$

eventually resulting in the equation of motion

$$m\ddot{x} = -mg + F_0 \frac{i}{i_0} - k_0 x \ \text{ or } \ ms^2 x(s) = \frac{F_0}{i_0} i(s) - k_0 x(s) \tag{5.15}$$

Since $x = x_m - x_{dist}$ and the effect of x_{dist} already is taken into account in the transfer function $G_1(s)$, this can be written as

$$x_m(ms^2 + k_0) = i \frac{F_0}{i_0} \tag{5.16}$$

which finally gives $G_2(s)$

$$G_2(s) = \frac{x_m}{i} = \frac{\frac{k_0 x_0}{m i_0}}{s^2 + \frac{k_0}{m}} = g_i \frac{\omega_2^2}{s^2 + \omega_2^2} \tag{5.17}$$

with

$$g_i = \frac{x_0}{i_0} \quad \text{and} \quad \omega_2 = \sqrt{\frac{k_0}{m}} = \omega_1 \tag{5.18}$$

so that the actuator transfer function can be written as

$$x_m = G_1(s)(x_{dist} + g_i i) \tag{5.19}$$

which is shown in the block scheme in figure 5.14. From measurements it can be concluded that the (parasitic) stiffness of the actuators is 420 N / m (vertical, all three actuators) and 1100 N / m (all horizontal actuators), respectively.

5.3.1 Controller design for the six-DoF Lorentz manipulator

The transfer function of the system has been measured using a spectrum analyzer. First, a preliminary controller has to be found that stabilizes the system in order to make it possible to measure the system's open-loop transfer function. The performance of this preliminary controller is not important, as it does not influence the measured transfer function.

After identifying the relevant parameters of the system, a choice can be made for the controller. This implies a tradeoff between steady-state error rejection, bandwidth and robustness.

With standard loop shaping techniques the required values of the PID controller have been determined. Starting from the open-loop transfer function, the following criteria have to be met:

1. Increasing gain at decreasing frequency, to ensure zero steady-state error.
2. Sufficiently large gain margin (at $-180°$) and phase margin (at the 0 dB crossing point) to guarantee stability and robustness.
3. Rapidly decreasing gain at high frequencies to decrease the sensitivity to measurement noise.
4. 0 dB crossing point at the desired bandwidth.

This has led to the implementation of six independent PID controllers meeting the requirements concerning stability, bandwidth and robustness. The closed-loop transfer function of the controlled system in the vertical direction is shown in figure 5.16 as an example.

5.3.2 Measured results of the six-DoF Lorentz manipulator

The most important parameter in evaluating the performance of the manipulator is the achieved vibration reduction. Figure 5.17 shows the achieved transmissibility of vibrations, which can be considered as the ratio beween the vibration amplitude of the baseplate and the vibration amplitude of the suspended table. Both distances might be measured relative to the ground plate, which is assumed to be rigid and stable. However, in the experimental set-up, another approach has been followed. The disturbance rejection function can also be derived from signals that are available in the controller during closed-loop operation.

The disturbance rejection function $D(s)$ is defined as

$$D(s) = \frac{x_m}{x_{dist}} \tag{5.20}$$

Because $x_m = G(s)(x_{dist} + g_i i)$ and $i = g_a g_s g_e K(s) x_m$, equation 5.20 can be written as

$$D(s) = \frac{x_m}{x_{dist}} = \frac{G(s)}{1 + g_t G(s) K(s)} \tag{5.21}$$

in which $g_t = g_a g_i g_s g_e$. Because the sensor, including the conditioning electronics, has a bandwidth of approximately 670 Hz, its transfer function has been assumed to be equal to $g_s g_e$ (consisting of two gain factors, representing the sensor part and the electronics part), in the frequency region of interest.

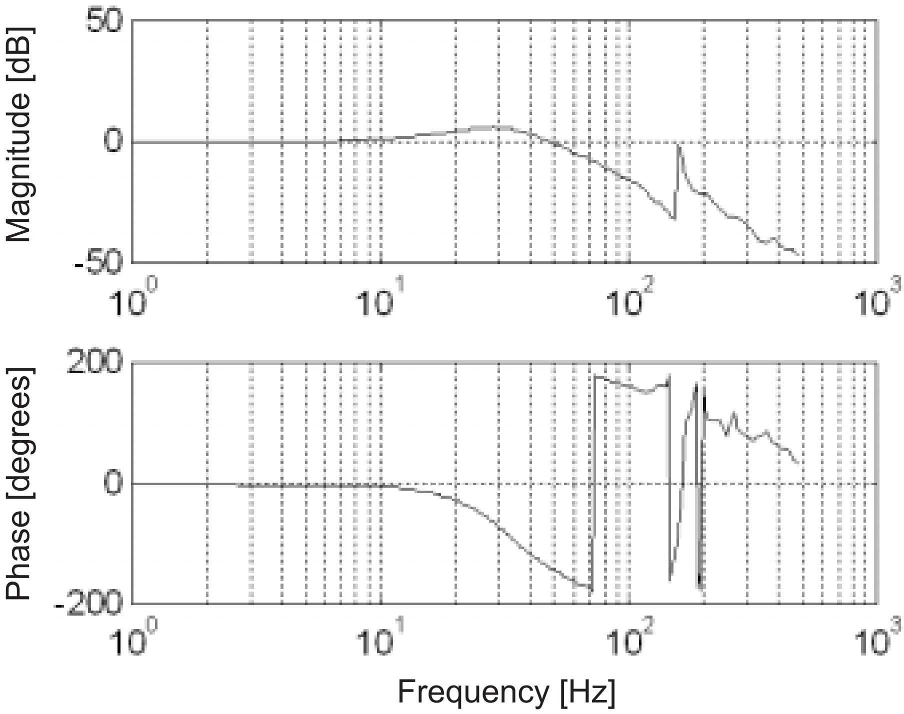


Figure 5.16: Measured closed-loop transfer function for the vertical coordinate (z) of the six-DoF Lorentz manipulator. Input is the current through the three vertical coils, output is the vertical table position, while the base plate is rigidly connected to the ground plate. The bandwidth of this system appears to be about 30 *Hz. In order to achieve a good vibration isolation, the bandwidth has to be large compared to the eigenfrequency, which is related to the parasitic stiffness of the actuators.*

The actual measurement of the disturbance rejection function has been done by determining the frequency response function between u_{out} and u_{in}. From figure 5.14 it can be seen that

$$u_{in} = g_s g_e x_m \tag{5.22}$$

and

$$u_{out} = \frac{1}{g_a g_i} x_{dist} \tag{5.23}$$

Now the disturbance rejection function $D(s)$ can be written as

$$D(s) = \frac{x_m}{x_{dist}} = \frac{u_{in}}{g_t u_{out}} \tag{5.24}$$

The two measured signals—the sensor input signal and the scaled power amplifier output—now are connected to a spectrum analyzer, which is able to measure the transfer function which represents the disturbance rejection function.

At low frequencies (< 1 Hz) the vibration reduction appears to be more than 35 dB. For high frequencies (> 200 Hz) the passive isolation provides a reduction of more than 60 dB. In the middle region, the passive isolation is less and also the controller's bandwidth can not be made wide enough (i.e. the gain can not be increased far enough) to achieve 40 dB reduction at the mid-frequency range, for stability reasons. Note that the figure only shows the vertical coordinate, but it is illustrative for the other degrees of freedom as well.

The disturbance rejection function does not completely meet the required specifications; this has several reasons:

1. The non-zero stiffness of the Lorentz actuators, because of a non homogeneous magnetic field in the actuator. This may be solved by fine-tuning the magnetic circuit of the actuators.
2. The non-zero stiffness of the Lorentz actuators, caused by the parasitic reluctance force.
3. The ground plate is assumed to be rigid and stable, but in practice this may be less ideal.

Other results achieved are listed in table 5.1. From these results it can be concluded that not all requirements, as listed on page 142, are met. The reason that the position errors are still about 0.8 μm, with corresponding angular errors, is the use of other sensors (for practical reasons), which had a resolution of only 0.5 μm.

The lowest eigenfrequency of 5 Hz seems relatively high for a vibration isolation system, but it has to be read in combination with the table mass, which is as low as 0.425 kg. From these values the effective stiffness k_f can be calculated:

$$\omega = \sqrt{\frac{k_f}{m_{table}}} \Rightarrow k_f = m_{table} \cdot \omega^2 = 0.425 \cdot (2\pi \cdot 5)^2 = 4.2 \cdot 10^2 \text{ N / m} \tag{5.25}$$

Note that this lowest eigenfrequency is the frequency related to the disturbance rejection. Therefore, the passive *eigenfrequency* of the system has to be as low as possible. The *bandwidth* of the closed-loop controlled system has to be as high as possible, in order to realize a small tracking error between the suspended table and the immovable world. In a real positioning application, the 'immovable world' might be replaced by the reference object to be followed. In that case, the suspended table will have a high (virtual) stiffness towards the reference object and a very low stiffness towards the supporting frame or vibrating world into which the reaction forces of the actuator are led.

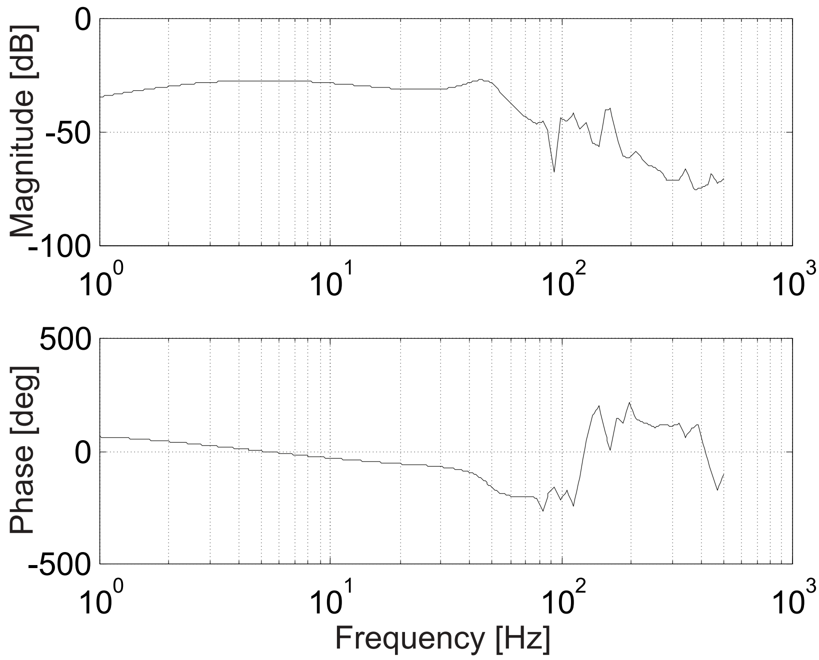


Figure 5.17: Disturbance rejection function of the six degrees of freedom Lorentz stage, showing the ratio between the vibration amplitude of the baseplate and the vibration amplitude of the suspended table. This graph only shows the vertical (z-)direction but it is illustrative for all six degrees of freedom.

From the experimental 6-DoF Lorentz manipulator the following conclusions can be drawn:

1. Lorentz actuators principally are suitable for passive vibration isolation, when the maximum delivered force and the power dissipation are not limiting its application. For a table of 425 g a lowest eigenfrequency of 5 Hz has been achieved. This is relatively low, taken into consideration the low mass of the table. The eigenfrequency can be decreased even more by fine-tuning

Parameter	Value	Unit	Remark
Vibration reduction	40	dB	$f < 1$ Hz
	> 25	dB	full frequency range
	60	dB	$f > 200$ Hz
Stroke	> 0.7	mm	< 1 mm, in x, y, z
	7 mrad	in $\varphi_{x,y,z}$	
Position stability	< 0.8	μm	Close to sensor resolution
	< 34	mrad	Close to sensor resolution
Table mass	0.425	kg	
Lowest eigenfrequency	5.0	Hz	vertical
	8.1	Hz	horizontal
Maximum force	6.15	N	constant
	9.75	N	peak
Maximum acceleration	5.8	m/s^2	for a 200 g load
Power dissipation	3	W	no load
	6	W	200 g load
	22.3	W	peak

Table 5.1: Measured results of the six-DoF Lorentz manipulator.

of the magnetic circuit in the Lorentz actuators, in order to decrease their stiffness.

2. The vibration reduction for frequencies below 1 Hz is at least 40 dB, above 200 Hz even more than 60 dB is achieved. In between there is a controller-induced resonance peak at 30 Hz where the reduction is 25 dB.

3. The moving magnet configuration can lift its own weight (4.17 N) and a payload of about 2 N. Larger loads are possible with larger magnets and coils.

4. Heat dissipation is 1 W per actuator for bearing the table mass. The generated heat can be transferred to the table by radiation or convection and—for larger amounts of heat—cause thermal deflections or even affect the properties of the permanent magnets. If these problems become critical, forced coil cooling must be applied.

5. The manipulator is relatively easy to control, using six independent PID controllers. Optimal control strategies may improve the system's performance, i.e. decrease the vibration transmissibility.

5.4 A new permanent magnetic bearing element

Although the Lorentz actuator set-up, described in the previous section, gives good results, its main disadvantage is the power dissipation. This disadvantage makes it difficult to apply it more generally. To generate a constant (bearing) force a constant current I through the coil is needed which results in a power consumption $P = I^2R$. There are several techniques to optimize the behaviour of Lorentz actuators towards maximum force, minimum heat, etc. [57].

Especially in precision positioning applications, heat sources have to be avoided as much as possible. Therefore a force generating element (i.e. a bearing element) is desired which does not require energy to maintain a static force.

In this section, the theory and practice of *permanent magnet configurations* as zero stiffness bearing elements will be described. To understand the possibilities and impossibilities, first the most relevant properties of permanent magnet materials will be enumerated, after which (in section 5.4.2) some theory will be treated with respect to positioning with permanent magnets.

5.4.1 Properties of permanent magnetic materials

5.4.1.1 Permanent magnetic properties: the BH-curve

The most comprehensive way to characterize permanent magnets is by means of the so-called BH-diagram. In this diagram the relation between the magnetic field strength H [A / m] and the magnetic induction or flux density B [T] is stored [54].

After the first magnetization, in which the field strength H is increased from zero to a certain maximum and then decreased again to zero, there remains a certain induction B. This $B_{H=0}$ is called B_r, the remanent induction. From this moment the magnet is ready to be used.

When a sufficiently large demagnetizing—negative—field strength H is applied, the induction B will decrease from B_r to zero, following the BH-curve in its second quadrant. This $H_{B=0}$ is called H_c, the coercitive field strength. When H increases again towards zero, B also increases, but now follows a line having the slope equal to the slope at B_r.

This leads to an important property of permanent magnets: if the BH-curve in the second quadrant is non-linear, a certain demagnetization will occur when the magnet is exposed to external magnetic fields.

In practice, in a magnetic circuit an equilibrium point—the working point—will be found for the B and H, depending on the magnetic reluctance. For static configurations (e.g. in Lorentz actuators) it is favorable to choose the magnet dimensions in such a way that the working point lies there where the product $B \cdot H$, representing the total stored magnetic energy, is at a maximum. In that case, the required magnet volume will be minimal.

5.4.1.2 History of permanent magnetic materials

Although since a long time materials with magnetic properties are known—already in the first centuries of this era they were applied in compasses—only in the last few decennia permanent magnetic materials have been developed that are useful as engineering elements. Only around 1930 magnets became strong enough to lift their own weight when attracting each other. From then, different alloys were developed continuously, e.g. ceramic magnets (Bariumferrite, Strontiumferrite, etc.), or magnets containing iron (Fe) and elements like cobalt (Co), nickel (Ni) and aluminum (Al), showing a rapid increase in magnetic energy, as is depicted in figure 5.18.

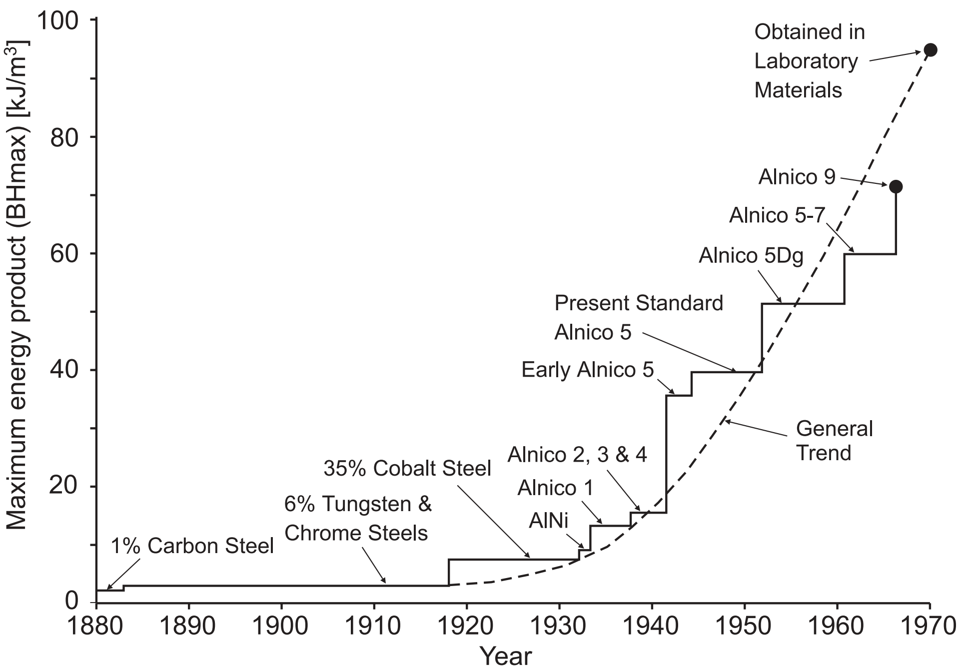


Figure 5.18: This rather old picture shows the development of permanent magnetic materials. Since about 1930 it is possible for permanent magnets to lift their own weight by attraction. From 1983, with the development of rare earth alloys, it is also possible to lift a magnet's own weight by repulsion without demagnetization. The modern magnet energy product of NdFeB magnets can be as high as 400 kJ/m^3*, which is slightly less than would be expected, based on extrapolation of the fitted exponential growth curve.*

A new breakthrough came around 1980 with the commercial availability of the so-called rare earth alloy permanent magnets. The most important types are Samarium-Cobalt ($SmCo_5$ or Sm_2Co_{17}, commonly abbreviated in the industry as SmCo) and Neodymium-Iron-Boron ($Nd_2Fe_{14}B$, commonly abbreviated as NdFeB).

They both have some remarkable characteristics:

- B_r is about $1 \ldots 1.47$ T, which is the highest of all commercially available permanent magnet materials.
- The BH-curve's second quadrant shows a nearly linear shape, meaning that there is virtually no demagnetization with varying applied magnetic field strengths.

There are also some differences between SmCo and NdFeB magnets:

- temperature resistance: typical Curie temperatures are 800°C and 350°C, the maximum operating temperatures are 250°C and 150°C for SmCo and NdFeB, respectively.
- NdFeB magnets are less expensive than SmCo magnets (the main elements Nd and Fe are abundant)
- NdFeB magnets are mechanically stronger than SmCo magnets, although they are more brittle.
- NdFeB magnet's specific mass is about 13% less compared to SmCo magnets

Some relevant numbers have been listed in table 5.2.

Parameter	**Unit**	**$SmCo_5$ sintered**	**Sm_2Co_{17} sintered**	**SmCo bonded**	**$Nd_2Fe_{14}B$ sintered**	**$Nd_2Fe_{14}B$ bonded**
B_r	T	$0.82 \ldots 0.95$	$0.9 \ldots 1.15$	$0.6 \ldots 0.8$	$1.1 \ldots 1.44$	$0.4 \ldots 0.68$
H_c	kA/m	$565 \ldots 692$	$692 \ldots 820$	$414 \ldots 485$	$844 \ldots 955$	$318 \ldots 460$
$(B \cdot H)_{max}$	kJ/m^3	$130 \ldots 175$	$160 \ldots 260$	$56 \ldots 120$	$240 \ldots 400$	$38 \ldots 80$
temp.coeff. B_r	%/°C	-0.05	-0.03		-0.1	
T_c	°C	800	800	800	350	350
α	ppm/°C				< 2	
Hardness	(Vickers)				750	
Corrosion res.	$+/-$	$+$	$+$	$+$	$-$	$-$
Machinability	$+/-$	$--$	$--$	$-$	$-$	$-$

Table 5.2: Some typical properties of SmCo and NdFeB permanent magnets.

Due to their linear relation between H and B, these rare earth metal magnets are virtually not demagnetized when magnetic fields are applied on them, up to the maximum coercitive field strength H_c. This also means that two of these magnets repelling each other cannot demagnetize each other. This property greatly facilitates the assembly of magnetic circuits and it also substantially enlarges the design possibilities. Thanks to these magnetic properties many new permanent magnetic configurations, like the ones designed in this chapter, have become possible.

5.4.1.3 Magnetic properties: ageing and corrosion

An important property of NdFeB magnets is their sensitivity to temperature and corrosion. The temperature effect is reversible, provided that (a) the working point remains on the linear part of the BH–curve, which becomes smaller at higher temperatures, and (b) the Curie temperature is not exceeded. The temperature coefficient of the remanent induction B_r is typically $-0.13\%/^\circ$C. The corrosion effect is irreversible. Without additional measures the material degradation causes the remanent induction to decrease.

It is recommended not to apply NdFeB magnets in the following environments [40]:

1. In acidic, alkaline or organic solvents, unless the magnet is sealed hermetically.
2. In water or oil, unless hermetically sealed, or a limited life is acceptable.
3. In electrically conductive liquids such as electrolyte containing water.
4. In hydrogen containing atmospheres, especially at elevated temperatures.
5. In environments containing corrosive gases such as Cl, NH_3 or NO_x.
6. In the path of radio-active rays.

Coating the NdFeB magnets with a metal layer (Ni, Zn, Al, Ti etc.) or with a resin film effectively prevents corrosion. Apart from that, the decay of remanent induction depends on the difference between the operation temperature and the Curie temperature of the magnet material. An illustration: a NdFeB magnet, exposed to a temperature of $130 \pm 5^\circ$C and a relative humidity of $90 \pm 5\%$ during 1000 hours (6 weeks), lost less than 5% of its maximum flux.

5.4.2 Earnshaw's Theorem

In 1842 Earnshaw published his famous article 'On the nature of the molecular forces which regulate the constitution of the luminiferous ether' in the Transactions of the Cambridge Society of Philosophy [14]. Very briefly summarized this proves:

> *For a pole placed in a static field of force it is impossible to have a position of stable equilibrium when an inverse square law relates force and distance.*

Because magnetic fields do have an inverse square law relating force and distance, this implies that purely passive permanent magnetic bearings never can be stable. A summary of the theory can be found in appendix A.

Nearly a century later this matter regained attention and interest, around the time that the development of new permanent magnet materials started to accelerate. In 1939 Braunbek worked out the practical consequences of Earnshaw's theorem [9], concluding a.o.:

> *Stable levitation is impossible in electrostatic or magnetostatic fields when all the present materials have $\mu_r \geq 1$ or $\varepsilon \geq 1$, but possible if materials of $\mu_r < 1$ or $\varepsilon < 1$ are introduced. From this, it follows that (a) stable levitation in static electric fields is impossible because materials with $\varepsilon < 1$ do not exist and (b) that only with the help of diamagnetic materials ($\mu_r < 1$) or superconductors ($\mu_r = 0$) stable permanent magnetic levitation can be realized.*

However, not everybody knew about Earnshaw's theorem. When around 1940 permanent magnetic materials improved so far that repulsion became possible with much less demagnetization, this led to several enthusiastic but incorrect proposals, formulated in patents, of which the most important and representative are summarized in table 5.3 [22].

Besides the impossibilities there are, fortunately, also several possibilities to achieve suspension[4] by means of electromagnetic fields. The principal methods can be described by the following classifications [22]:

1. Levitation by magnetic attraction or repulsion forces using magnets of fixed strength and ferromagnetic materials. Earnshaw's theorem excludes the possibility of stable levitation.
2. Electromagnetic levitation by attraction or repulsion forces with active control of the current through the coil.
3. Electromagnetic levitation by repulsion forces due to eddy currents induced in the levitated object.
4. Electromagnetic levitation utilizing the force acting on a current-carrying linear conductor in a magnetic field.
5. Magnetic levitation by repulsion forces using diamagnetic materials.
6. Magnetic levitation by repulsion forces using superconductors.
7. Electric levitation by attraction forces with active control of the electrode potential.

The first mentioned item is of particular interest, because it concerns passive systems. The word *stable* is crucial. After all, when vibration isolation is desired, not a

[4] Some authors make a difference between *suspension* for hanging objects and *levitation* for objects supported from below.

Snell, 1940 British Patent 539409	*Bearing arrangements* Levitation of a shaft between two magnetostatic conical bearings
Neal, 1941 US Patent 2323837	*Magnetic apparatus* A platform containing vertical bar magnets with N poles down should be levitated over a fixed set of vertical bar magnets with N poles up
Schug, 1943 US Patent 2436939 British Patent 583298	*Magnetic coupling and bearing* similar to Snell, 1940.
Viazemsky, 1946 British Patent 625900	*Scientific measuring instruments of the mirror type* Levitation of an axially magnetized shaft by repulsing magnets beneath, beside and at the ends.
Heidenwolf, 1947 British Patent 655429	*Apparatus for measuring temperature* Contains magnetically levitated spindle for temperature sensing instrument.
Goldschmidt, 1948 British Patent 642353 US Patent 2704231	*Magnetic bearings* Two-material radial bearing. Since in magnetic bearings constructed from permanent magnets the flux density is greater near the magnetic neutral zones than near the poles, two materials may be advantageously employed for fabricating the magnets. Material a for the neutral zones would be one giving its maximum energy product at high flux densities but having a smaller coercive force than material b for the poles, which would have a high coercive force but yield maximum external energy at lower flux densities. Radial bearings with mutual repulsion between axially or radially magnetized inner and outer rings are indicated. It is asserted that the supported member can be levitated by some of the magnet arrays, in contravention of Earnshaw's theorem.

Table 5.3: Overview of some representative patents that contradict Earnshaw's theorem and therefore are impossible to realize.

stable but a *neutral* equilibrium has to be achieved, that means, just between stable and unstable. *The situation of neutral equilibrium in all degrees of freedom is not excluded by Earnshaw's theorem.*

This fact provides a theoretical basis to develop permanent magnetic bearing elements that have zero stiffness in at least one degree of freedom.

5.4.3 Modeling of 1-DoF permanent magnetic zero stiffness elements

In this section we will describe a new permanent magnetic bearing element having zero stiffness[5]. Firstly, the principle is explained based on a theoretic example and finite element analysis, successively paying attention to bearing force, stiffness, damping, stability and control properties. Subsequently, a practical test set-up with its measured performance will be presented.

5.4.3.1 Bearing force

For simplicity, first consider a permanent magnetic configuration consisting of a fixed support magnet exerting an upward repulsing force on a floating magnet and another fixed magnet above that floating magnet resulting in an upward attracting force, as is depicted in figure 5.19. Note that the upper magnet can be replaced by an iron plate, resulting in essentially the same behavior, although the actual force and stiffness will be different.

Based on the principle sketched in figure 5.19, a radially symmetrical set-up has been designed to calculate the bearing force by means of a finite element analysis (FEA)[6]. Figure 5.20 shows the model with one quarter cut out to show the cross section and the flux density pattern.

These calculations of the vertical force have been carried out for several values of the height z of the middle element in the air gap, so that a force–displacement graph can be plotted [52]. In figure 5.21 this graph shows that indeed there is a point where the gradient of the force–displacement curve is zero. This point will be called *working point* from now on, having z_{wp} as a vertical coordinate.

The force–displacement characteristics of the bearing element, as depicted in figure 5.21, can be influenced in three basic ways:

1. Adjustment of the air gap (upper, lower or both). This can be of importance for drift control and quasi-static disturbing displacements. In figure 5.22 the typical behavior is shown when the upper and lower gap are varied symmetrically.

[5]Dutch patent, international patent pending [51].

[6]In this thesis, electromagnetical finite element calculations have been done using Maxwell 2D and / or 3D Field Simulator software, which is a product of Ansoft Corporation, Pittsburgh, PA, USA.

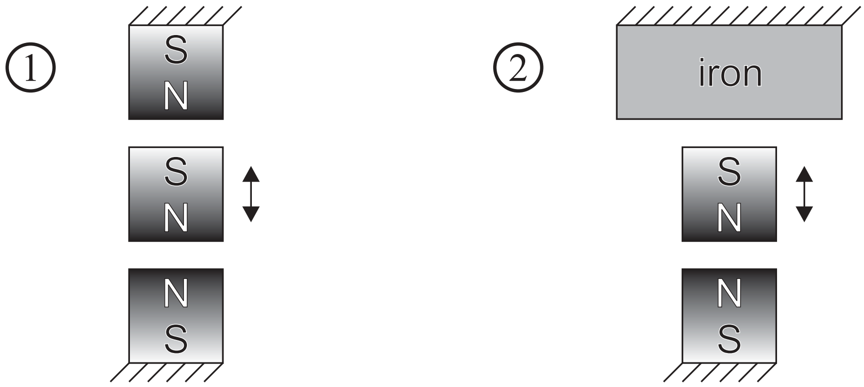


Figure 5.19: Principle of the zero–stiffness bearing unit using permanent magnets. The attractive force (above) and repulsive force (below) create negative and positive stiffness, respectively, and cause the total stiffness to be zero in a working point. In (2) the upper magnet has been replaced by an iron plate, resulting in essentially the same behavior.

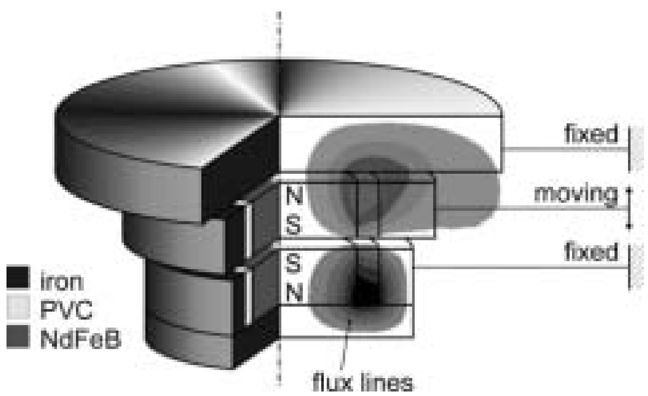


Figure 5.20: Model of a practical configuration, based on the principle sketch (2) of figure 5.19, used for a finite element analysis. One quarter has been cut out to show the cross section and the flux density pattern. The nominal upper and lower air gaps both are 2 mm.

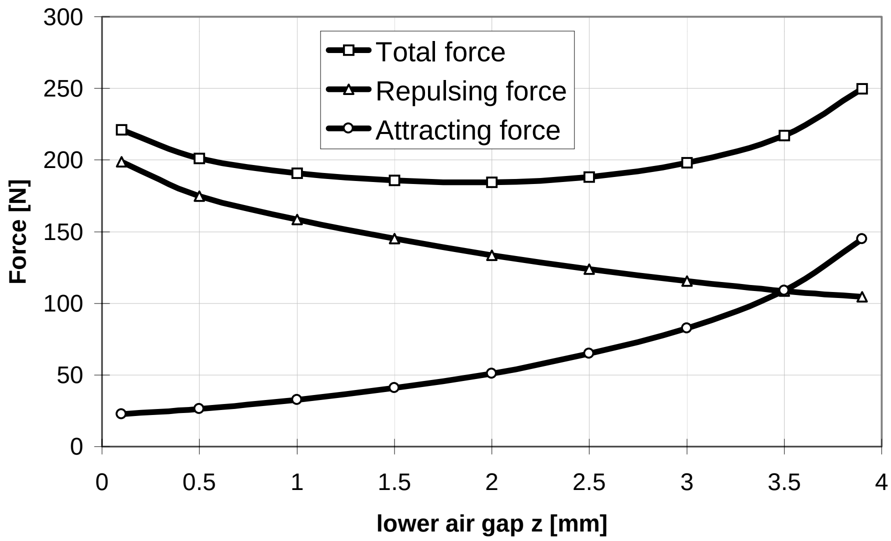


Figure 5.21: When the attracting and repulsing forces of the permanent magnets are added up, the total force curve shows a minimum where the first derivative—the stiffness—is zero. The values in this figure are obtained from a Finite Element Analysis using the model shown in figure 5.20. At $z = 2$ *mm, the moving part is just centered between the two fixed parts. The height of the working point* $z_w p$*, where the stiffness* k_z *is equal to zero, is slightly lower than* 2 *mm.*

2. Modification of the magnetic field. In practice, it is very difficult to predict the consequences of any possible field variation at any possible height. This prediction is necessary to build a model for a control loop.

3. Addition of an external force. For this option, the Lorentz force appears to be the most suitable alternative because it is contactless and it adds only force, no stiffness.

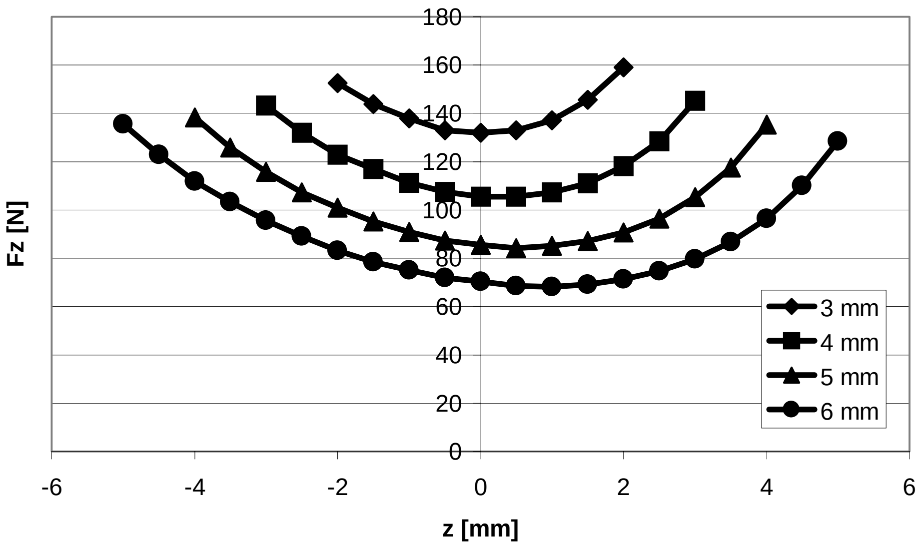


Figure 5.22: Typical behavior of a zero stiffness bearing element when the upper and lower air gap are varied symmetrically (N.B.: in the legend the value of one air gap is displayed). With smaller air gap the bearing force in the working point increases and the slopes become steeper. This means that the range in which the stiffness has acceptable small values will be smaller.

From figure 5.22 it appears that the location of the working point, z_{wp}, varies with the gap height. In figure 5.23 the calculated position z_{wp} and corresponding bearing force F_{wp} in the working point are shown. It appears that the working point location as well as the bearing force in the working point can be linearized with good approximation. Therefore, the implementation of an air gap adjustment facility will not cause severe problems in the controller.

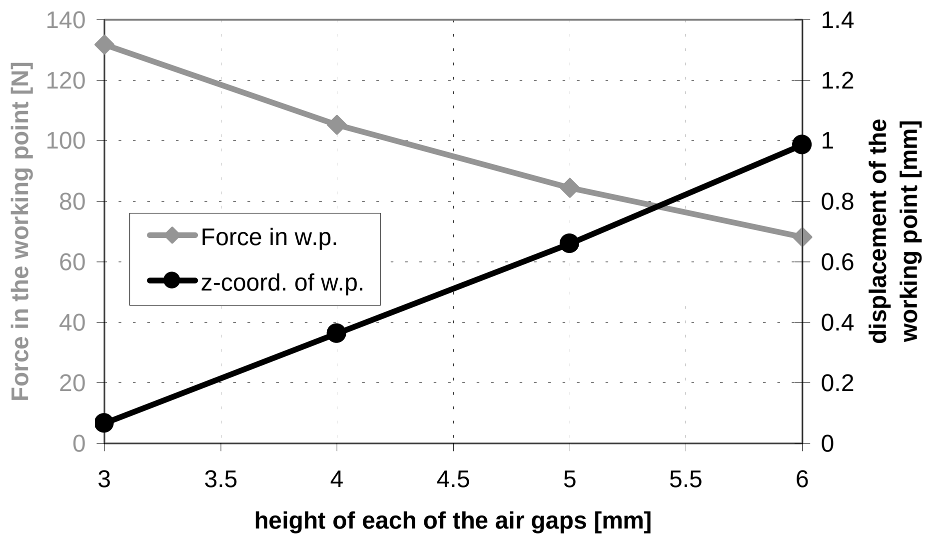


Figure 5.23: This graph shows two effects of an air gap variation, derived from figure 5.22. In the simulation the upper and lower air gap are increased simultaneously, relative to the nominal middle position. Firstly, the force in the working point will decrease with enlarging air gaps. Secondly, the location of the working point relative to the center will shift upward with enlarging air gaps. The relation between the air gap height and the position z_{wp} appears to be linear with good approximation. The bearing force as a function of z_{wp} can be linearized as well, although a third order polynomial fit gives slightly better results. Therefore, the implementation of an air gap adjustment facility will not cause severe problems in the controller.

5.4.3.2 Stiffness

By applying a polynomial curve fit[7] through the measured force–displacement curve we can approximate its derivative, the stiffness. It appears that the stiffness, approximated by a third order polynomial over the whole range of the height z, can be linearized by good approximation in a small region ($\pm 5\,\mu$m) around the working point z_{wp}. This stiffness is positive ($\partial F/\partial z < 0$) for $z < z_{wp}$, negative ($\partial F/\partial z > 0$) for $z > z_{wp}$ and crosses zero at $z = z_{wp}$. That means that for downward movements there is a stabilizing upward force and for upward movements there is a destabilizing force, in the upward direction as well. In the working point z_{wp} the force equilibrium is *neutral* and *marginally stable*.

Some additional remarks have to be made concerning the *radial* stiffness, which has been left out of consideration so far. Earnshaw's theorem (appendix A) states a.o. that

$$k_x + k_y + k_z = 0 \tag{5.26}$$

For radially symmetric permanent magnet configurations having axial stiffness $k_{ax} = k_z$, the radial stiffness $k_{rad} = k_x = k_y$ follows that

$$k_{rad} = -\frac{1}{2}k_{ax} \tag{5.27}$$

This last equation implies that when an object is suspended axially by a permanent magnet bearing under zero stiffness conditions as described above, the radial stiffness will be zero as well.

However, this rule does not hold in the presence of soft magnetic material. To understand this, consider the following situation: A permanent magnetic bearing consisting of two repulsing magnets, one fixed to the world and the other one suspended above it (part A of figure 5.24). The axial stiffness is positive, the radial stiffness is negative and will be half of the axial stiffness in magnitude, according to the above rule. Now, connect another magnet to the world, located above the suspended magnet and attracting it. The resulting force (fig. 5.24 D) will be upwards while generally there will exist a point where the axial stiffness is zero. In this point the radial stiffness should be zero as well, according to equation 5.27. This can be understood by the fact that two attracting magnets having negative axial stiffness will have a positive radial stiffness resulting in a centering force (fig. 5.24 B).

But when we replace the upper attracting magnet by a ferromagnetic (e.g. iron) plate, things become different. The ferromagnetic plate results in a net attracting force as well, now generated by the field of the suspended magnet only. This attracting force has a negative stiffness in the axial direction, but *no* radial stiffness, provided that the plate is wide enough (fig. 5.24 C). There will be a net negative radial stiffness in the point where the axial stiffness is zero (fig. 5.24 E), which contradicts equation 5.27. This effect is confirmed by the theory discussed in appendix A.

[7] In this case a fourth order polynomial gives good results. The deviation from the measured points is in the same order as the expected uncertainty in the FEA results.

The conclusion of this imaginary experiment is that equation 5.27 is valid for magnet configurations only, without the presence of ferromagnetic material.

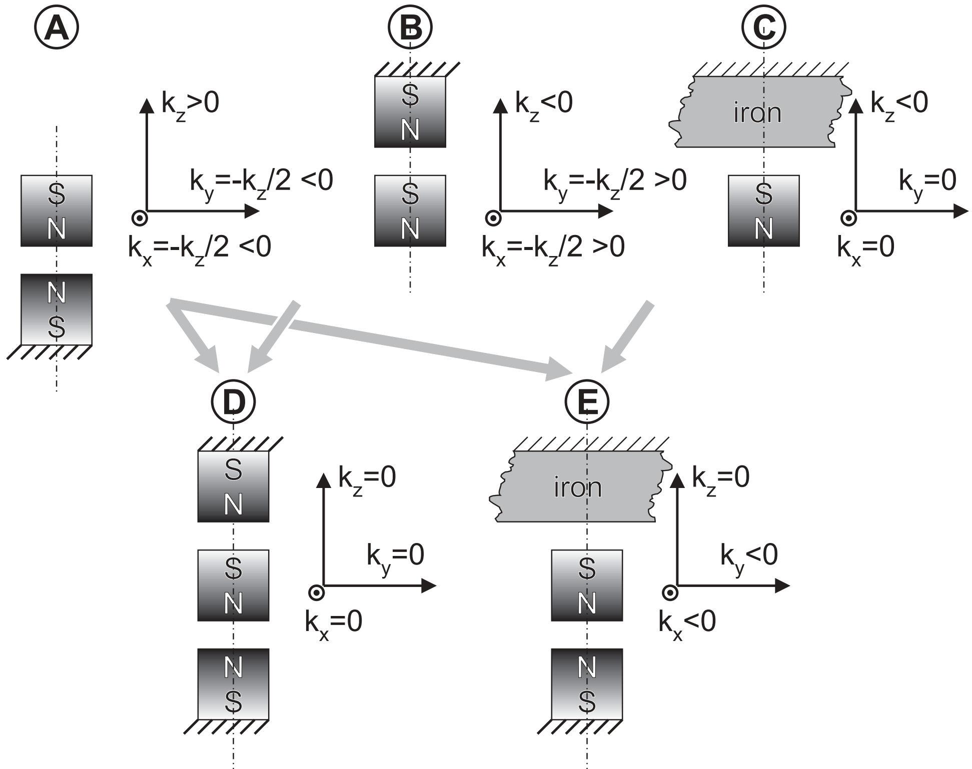


Figure 5.24: Comparison of the lateral stiffness for bearings with rotational symmetry. In both cases (D) and (E) there is a point where the vertical stiffness is zero, but lateral zero stiffness in the working point is only possible when no ferromagnetic parts are present.

In appendix A, also one section is dedicated to the *rotational stiffness* around the x-, y- and z-axes, which have been left out of consideration in Earnshaw's theorem. These stiffnesses, however, are important when designing a suspension element that has to operate in six degrees of freedom.

5.4.3.3 Damping

The internal damping of the permanent magnetic bearing element can only be caused by *eddy currents* and *air friction*, the latter in the case that the moving part of the axial bearing element according to figure 5.19 is guided by air bearings, see section 5.4.4.

With regard to the *eddy currents*, there are three reasons that allow us to assume that their influence will be small in this type of bearing element:

Firstly, the magnitude of the eddy current damping force is proportional to the square of the magnetic field variation, $(\delta B)^2$. The eddy current power loss can be approximated by [48]

$$P = \frac{1}{2\rho}(\omega \cdot \delta B)^2 L^3 A \tag{5.28}$$

in which ρ is the specific resistance in Ωm, ω is the vibration frequency in rad/s, L is the air gap in m, δB is the variation of the flux density in T through the eddy current area, A, in m^2. The flux density variations δB depend on the geometry and are

$$\delta B = \frac{3 \cdot \delta z \cdot h^3}{8L^4} B_r \tag{5.29}$$

where δz is the vibration amplitude in m, h the magnet thickness in m and B_r the remanent induction of the magnet in T.

For small position deviations around the working point the force virtually does not change, and because the force is directly related to the magnetic field strength this field will not change as well. However, this is not completely true in the case of a configuration like in figure 5.19(2). Although the total force does not change, the upper and lower components of that force do, so that eddy currents will be generated in the ferromagnetic parts (i.e. the upper iron plate). Because the specific electrical resistance of NdFeB is high, the eddy current loss within the magnets will be very small.

Secondly, the magnitude of the eddy current damping force is proportional to the square of the vibration frequency. Fortunately, the vibration amplitude decreases with increasing frequency under normal laboratory conditions.

Thirdly, in order to prevent the generation of eddy currents irrespective of above considerations, radial slits may be applied in the appropriate parts, reducing the eddy currents with one or two orders of magnitude.

In the description of the practical test setup in section 5.4.4 a quantitative estimation will be made for the upper limit of the influence of the eddy currents.

The damping caused by *air bearings* will depend on the relative velocity between the suspended magnet and the support. The shear stress between air layers is given by the dynamical viscosity η of air and the velocity profile through the layer thickness h:

$$\tau = \eta \frac{\partial v}{\partial h} \tag{5.30}$$

which can be simplified, under the assumption of a linear velocity profile, to

$$\tau = \eta \frac{v}{h} \tag{5.31}$$

with which the damping force F_d becomes

$$F_d = \tau A = \frac{\eta v A}{h} \tag{5.32}$$

when A is the bearing area. Now, the damping coefficient c can be expressed as

$$c = \frac{F_d}{v} = \frac{\eta A}{h}. \tag{5.33}$$

In section 5.4.4.2 it will be shown that for the practical test set-up the maximum damping coefficient and the maximum damping force will be relatively small.

5.4.3.4 Stability and modeling

The stability of the bearing element is related to the stiffness. Positive stiffness means stability, negative stiffness means instability. In section 5.4.3.2 it has been shown that the stiffness in the working point causes a neutral and marginally stable equilibrium.

This stability situation can be visualized by a ball, rolling on a curved path in the presence of gravity. In this analogon the horizontal movement of the ball corresponds to the z coordinate of the magnetic bearing element. In figure 5.25 this is depicted for three situations:

1. The weight of the supported object slightly exceeds the minimum bearing force: $m \cdot g > F_{wp}$. Then there exist two intersections of the line $F = mg$ with the bearing force curve, of which one (A) is stable and the other (B) is unstable.

2. The weight of the supported object exactly matches the minimum bearing force: $m \cdot g = F_{wp}$. Then the points (A) and (B) coincide in a neutral equilibrium.

3. The weight of the supported object is less than the minimum bearing force: $m \cdot g < F_{wp}$. Then there is no force equilibrium possible, the object will accelerate upwards.

In this analogon the additional active controller used for bearing force adjustment can be considered as a rotating and / or translating mechanism providing neutral equilibrium for the rolling ball, that is, (A)=(B). However, a *neutral* equilibrium is only possible in the point of inflection of the curved path, so the goal of the controller is to manipulate the curve in such a way that for a changing mass or an external disturbing force (resulting in horizontal ‘forces’ in this analogon) the ball is supported again in a new point of inflection.

To achieve sufficient vibration isolation, the supported object has to be kept as close as possible in the working point. There are two principal possibilities to realize this:

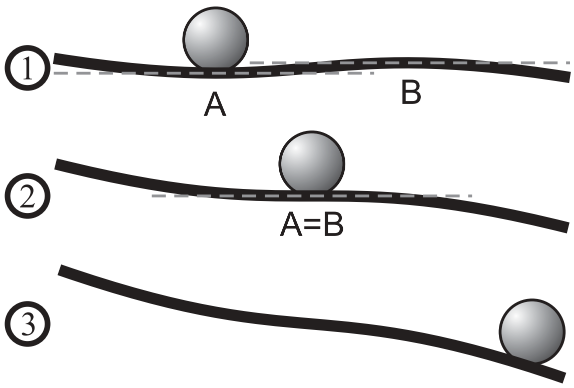


Figure 5.25: Analogon for the stability in the working point. The horizontal ball movement corresponds to the displacement z. 1. bearing force is smaller than weight → stability in A, instability in B; 2. bearing force equals weight → neutral equilibrium only in the working point; 3. bearing force is larger than weight → instability.

1. Make the support follow the supported object. Drift will be a problem because the supporting force will increase both below and above the working point, so that the suspended object always will tend to move upwards. Therefore, the application of exclusively this method will not be sufficient for the stabilization of the bearing element, although it may be of use in systems which do not allow small position deviations from the working point.

2. Exert a control force on the supported object using one of these three methods (which are already mentioned before):

   (a) Adjustment of the air gap

   (b) Modification of the magnetic field

   (c) Addition of an external force

The latter three possibilities may create an additional virtual stiffness. From this point of view, method 1 is preferable. However, a zero stiffness bearing using this method will be unstable, because every positive or negative deviation from the working point results in an upward force causing the suspended object to drift in the upward direction without a possibility to get it back down. This problem can be solved only by choosing the operation point slightly below the optimal working point so that a small positive stiffness is created. The required distance from the working point depends on the expected maximum disturbance force (on the suspended object) or displacement (of the support).

If method 2 were used, a compromise has to be found between position accuracy and net virtual stiffness. When a very fast and accurate controller is implemented that keeps the suspended object tightly in the working point, also a large (virtual) stiffness is introduced between the support and the suspended object, thus effectively destroying the inherent vibration isolation capabilities. On the other hand, when the controller bandwidth is too low, drift will cause position deviations from the working point resulting in a certain (real) stiffness proportional to the position deviation. Actually, there must be found an optimum choice for the contributions of the *virtual* stiffness (caused by position feedback of the controller) and *real* stiffness (caused by position deviations from the working point).

Therefore, the optimum strategy to stabilize this zero stiffness bearing is to maximize the proportional *acceleration* feedback while making the proportional *position* feedback as low as possible (and just as high as needed), combined with a damping just large enough to prevent the bearing from large amplitudes. In table 5.4 the parameters are listed which may be used to realize this.

	position	**velocity**	**acceleration**	**jerk**
$I(z)$	$P(z)$ $I(\dot{z})$	$D(z)$ $P(\dot{z})$ $I(\ddot{z})$	$D(\dot{z})$ $P(\ddot{z})$	$D(\ddot{z})$
	$\Downarrow$ virtual stiffness	$\Downarrow$ virtual damping	$\Downarrow$ virtual mass	

Table 5.4: Effect of possible feedback parameters in combination with standard PID controllers on the closed-loop behaviour of the zero stiffness magnetic bearing.

From table 5.4 it can be concluded that a velocity sensor as described in section 2.1.2 would be suitable when used with a large D, small P and I for the v-signal and a small P for the x-signal. Of course, these requirements for the best possible vibration isolation properties have to be matched with control criteria dealing with closed-loop stability.

Assuming that for the force correction an external force will be used (such as the Lorentz force, which is very suitable for this task), the following model can be derived to serve as a starting point for the controller design.

Because the bearing system will behave like a (non-linear) mass-spring-damper system, the first approximation for the equation of motion is

$$m\ddot{z} + c(\dot{z} - \dot{z}_w) + k(z - z_w) = F_{ext} \tag{5.34}$$

in which z is the absolute height relative to the imaginary immovable reference world and z_w is the height of the vibrating world to which the bearing is attached.

The parameter m is known or can be measured. The parameter c can be estimated from the damping considerations in section 5.4.4.2. This will have a small value, but it may be not negligible. In the following simulations, however, the damping coefficient has been assumed to be equal to zero.

Subsequently, the relation between position and bearing force has to be determined from measurements or simulations to give an estimate for the stiffness. A fourth order polynomial appears to give a good approximation for the force curve, so the stiffness will be approximated by a third order polynomial. From measurements it is observed that in a small region around the working point (several micrometers) this relation between position and stiffness can be linearized[8] with an uncertainty of less than 100 ppm. The stiffness k now becomes

$$k = \kappa \cdot (z - z_w) \tag{5.35}$$

where κ is a constant with dimension N / m^2.

The equation of motion now becomes

$$m\ddot{z} + c(\dot{z} - \dot{z}_w) + \kappa(z - z_w)^2 = F_{ext} \tag{5.36}$$

which is non-linear. This equation can be transformed to a Simulink[9] block scheme which is shown in figure 5.26. Here, the inputs 1 and 2 are the external force (e.g. a Lorentz force for control) and a disturbing displacement of the world, respectively. Output 1 is the relative height $z - z_w$ measured by an air gap sensor, output 2 the absolute position z with respect to the immovable reference world and output 3 the absolute acceleration $\ddot{z}$.

Around the modeled block scheme of the zero stiffness bearing element given in figure 5.26, a controller has been designed, as is shown in figure 5.27. The bearing element is represented here by a block named ZSMB, for Zero Stiffness Magnetic Bearing. The controller consists of four parts:

- The inputs for the model are generated internally. Firstly, the block named 'Floor vibrations' generates low-pass filtered (second order, at 10 Hz) white noise with an amplitude of 10 μm. Secondly, the setpoints for the position and acceleration (named 'z' and 'acc', respectively) both are zero in this case, but they could be the outputs of another model. Thirdly, there are two blocks named 'Measurement noise'. They create a white noise signal representing random noise on the measured position and the acceleration, which can be used to test the controller's stability.

- The block named 'Controller' contains the control parameters. It will be shown in figure 5.28.

- The block named 'ZSMB' contains the model of the Zero Stiffness Magnetic Bearing, which was shown in figure 5.26.

[8] Approximating the force curve with a second order polynomial will give a linear (first order) stiffness–position relation directly, but then the location of the working point will be calculated less accurately.

[9] Matlab and Simulink are trade marks of The MathWorks, Inc.

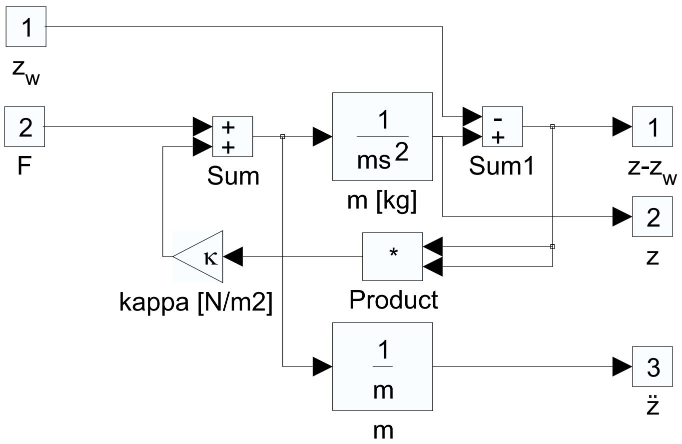


Figure 5.26: Simulink block scheme of the bearing element's equation of motion in the z-direction (assumed there is no damping). The inputs 1 *and* 2 *are the disturbing displacement z_w of the world and the external force F (e.g. a Lorentz force for control), respectively. Output* 1 *is the relative height $z - z_w$ as it is measured by an air gap sensor, output* 2 *is the absolute height z with respect to the immovable reference world and output* 3 *is the absolute acceleration $\ddot{z}$.*

- The outputs of the model (vectors containing a signal value for each time step), which are stored in the Matlab workspace, where they can be processed further. The variables are: 'time' (the time vector), 'vibr' (the floor vibration signal), 'pos_real' (the real position relative to an imaginary, immovable world), 'pos_rel' (the measured position relative to the vibrating support), 'acc' (the absolute acceleration) and 'acc_meas' (the measured acceleration, which is equal to the absolute acceleration plus some measurement noise).

Besides these elements there are some emergency stops to facilitate the iterative simulation process towards a satisfying controller.

Using this controller, simulations have been carried out to obtain information about the closed-loop behavior. By means of the position dependent proportional gain, the real stiffness of the bearing element has been eliminated (or, in practice, will be substantially reduced). Therefore, the suspended object of the bearing element behaves like a free mass. By adding a PD controller, in the closed-loop situation the object will behave like a simple mass–spring–damper system. Thus, the necessary proportional gain can be calculated, in order to make the open-loop transfer function (in a Bode plot this transfer function will be a straight line having a -40 dB / dec slope) cross the 0 dB line at ω_{bw}, the desired bandwidth. After that, a lead-lag filter has to be added in order to add a certain amount of damping (in the Bode plot: change the slope at 0 dB from -40 to -20 dB / dec and increase the phase).

In this simulation, a (very low) desired bandwidth of 0.1 Hz has been assumed.

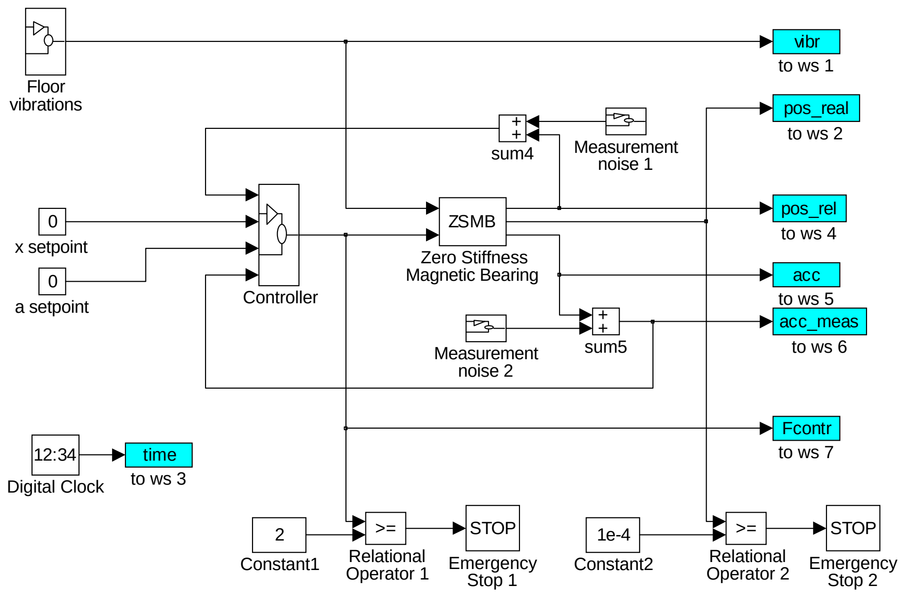


Figure 5.27: Simulink block scheme of the entire controller structure for the z-direction of the zero stiffness bearing element. The two main parts are the ZSMB block, containing the model of the Zero Stiffness Magnetic Bearing, and the controller block which is shown in figure 5.28.

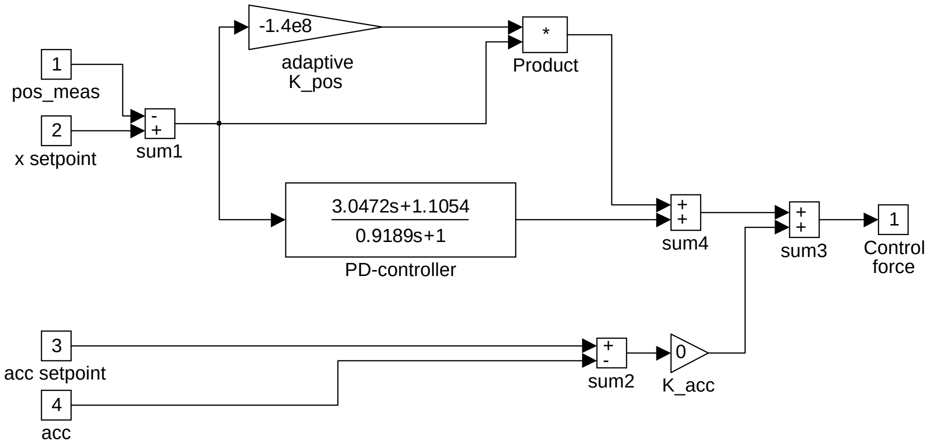


Figure 5.28: Simulink block scheme of the controller, used in the model shown in figure 5.27. The gain factor 'Adaptive K_pos' multiplied by the measured position results in a proportional gain (equivalent to a stiffness) which eliminates the real stiffness of the bearing element. The PD-controller is a lead–lag filter that is applied to create a sufficiently large phase margin around the 0 *dB crossing point in the open-loop transfer function. Using this type of control, the suspended object can be regarded as a simple mass–spring–damper system, having a natural frequency which can be made equal to any desired value.*

With the help of the following Matlab script, the parameters for the PD controller are calculated.

```
% Matlab script, calculating the PD parameters for
% the control of the zero stiffness magnetic bearing

nump=[1];          % \  process transfer function, including stiffness
denp=[2.8 0 0];    %  > compensation: the suspended object behaves like
                   % /  a free mass

m=2.8;             % actual mass of the floating object
w0=0.1*(2*pi);     % desired 0 dB crossing point of the open-loop tf
Kp=m*w0^2;         % necessary proportional gain to realise w0
phi=30*(2*pi/360); % desired phase margin at w0
T2=sqrt(-(1+sin(phi))*(-1+sin(phi)))/((1+sin(phi))*w0);
T1=1/(w0^2*T2);    % T1 and T2 are time constants of a phase lead/lag

numc=Kp*[T1 1]     % \_ controller transfer function
denc=[0 T2 1]      % /

num_ol=conv(numc,nump); % \_ open loop transfer function
den_ol=conv(denc,denp); % /
figure(1);
bode(num_ol,den_ol);    % bode plot of the open loop tf

[num_cl,den_cl]=cloop(num_ol,den_ol,-1);
figure(2);
bode(num_cl,den_cl);    % bode plot of the closed loop tf
```

In figure 5.29 the positions are shown, as a result of the simulation using these calculated values. It is clearly visible that only the low-frequency position deviations are transmitted through the bearing. The cross-over frequency, determining this behavior, can be chosen arbitrarily. However, for decreasing cross-over frequencies the amplitude will increase, so that a compromise has to be found between isolation frequency range and absolute amplitude. Therefore, a bearing or actuator system like this will not achieve a much better performance than conventional vibration isolation systems.

Mainly based on this approach, the controller for the experimental zero stiffness *linear* bearing has been modeled. This is described in section 5.5.2. In the next section 5.4.3.5, it will be shown that there is another way to control such systems, which is more suitable to motion systems.

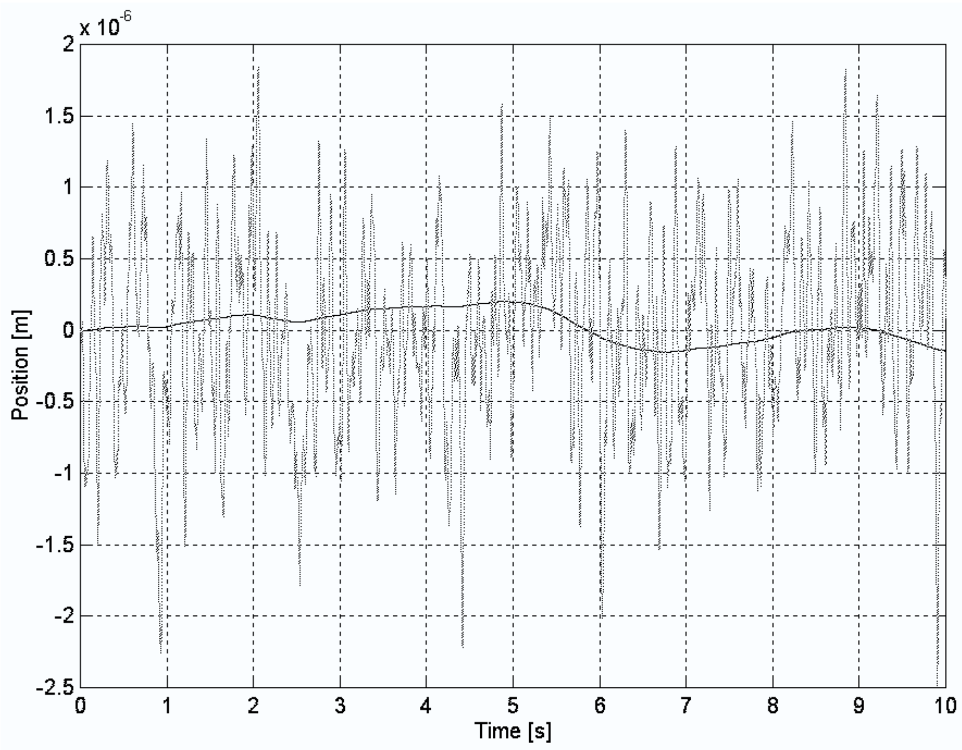


Figure 5.29: Simulated position output of the controller for the z-direction. Displayed are the vibration of the world, z_w, (noisy) and the absolute position relative to an imaginary, immovable world, z (slowly varying). Only low-frequency position deviations of the support are followed.

5.4.3.5 Considerations with regard to control

In controlling a positioning system containing zero stiffness actuators, it is important to consider the position reference used and the intended use of the actuator. This has been mentioned earlier in section 5.1.1, in the category list on page 131. The difference between systems that only have vibration isolation as a goal, and systems that have to position an object by means of vibration isolating actuators, has been visualized in figure 5.30.

In this figure, the upper part represents a sketch of the positioning system, The middle part shows the block scheme of a system that is used merely for vibration isolation, which means that there is no position signal relative to an external reference. The lower part shows the block scheme for a positioning system, where the position of the supported object relative to the supporting world as well as relative to a position reference is available. First, consider the system for vibration isolation only, which is marked with (1) in figure 5.30.

In a vibration isolation system, two parameters have to be optimized: Firstly, the desired setpoint tracking (z_m = constant) can be written as

$$\frac{za_m}{z_{sp}} = 1 \tag{5.37}$$

Besides that, it is desired that the vibrations of the world, z_w, do not affect the absolute position of the supported object, z_m:

$$\frac{z_m}{z_w} = 0 \tag{5.38}$$

It is easily seen that equation 5.38 only can be fulfilled for each z_w when $z_m = 0$. This implies that the desired setpoint tracking in equation 5.37 can not be realized for real values of z_{sp}. Therefore, a compromise will have to be found between absolute position and isolation. This agrees with intuition, which says that for a decreasing stiffness of the supporting system (necessary for improved vibration isolation) the amplitude of the supported object will increase. It is not possible to decrease that amplitude without affecting the vibration isolating performance.

However, it is very well possible to decrease the amplitude *relative to an external reference*. To understand this, consider the vibration isolated positioning system (marked with (2) in figure 5.30). Here, the two desired properties are:

$$\frac{z_{ref} - z_m}{z_{sp}} = 1 \tag{5.39}$$

for optimal setpoint tracking, and

$$\frac{z_{sp} - z_{ref} + z_m}{z_w} = 0 \tag{5.40}$$

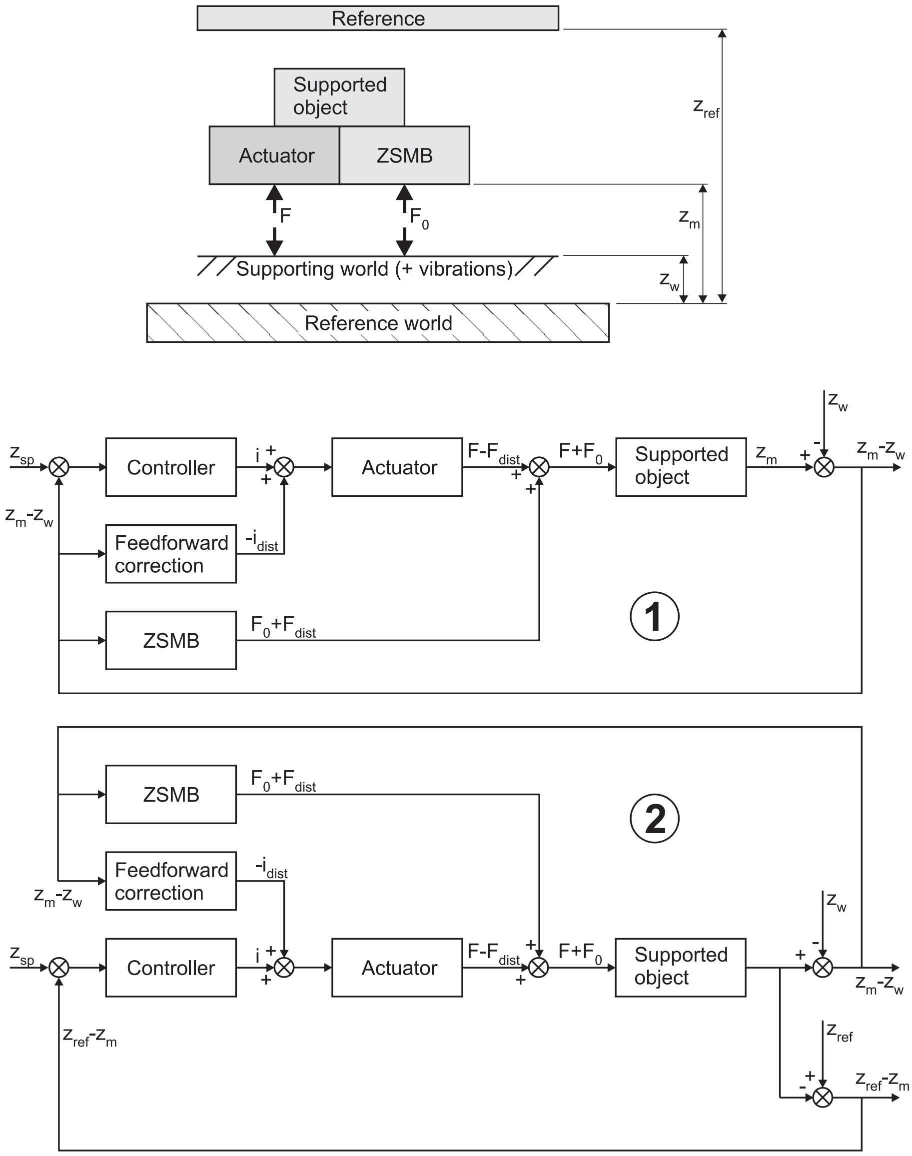


Figure 5.30: This figure shows the difference between systems which merely have vibration isolation as a goal (marked with 1), and systems that perform a certain positioning task by means of vibration isolated actuators (marked with 2). The first category does not have a reference signal, only the relative position between the supported object and the supporting world, $z_m - z_w$, is present. As a result, a compromise has to be found between quality of isolation and position amplitude. In the second case, the positioning task (z_m has to follow z_{ref}) can be separated from the vibration isolating task.

for optimal vibration isolation. From equation 5.39 follows that the setpoint $z_{sp} = z_{ref} - z_m$. This does not contradict with equation 5.40, which substitutes to

$$\frac{z_{ref} - z_m - z_{ref} + z_m}{z_w} = \frac{0}{z_w} = 0 \quad \forall z_w \tag{5.41}$$

This means that in this case the isolation is optimal when the tracking is optimal. Now, the suspended object can be positioned relative to the reference object, while the vibrations from the supporting world virtually do not affect the positioning accuracy.

5.4.4 Experimental 1-DoF permanent magnetic zero stiffness bearing

Based on the ideas and theory described in the previous sections, an experimental set-up has been built to investigate the behavior of a zero-stiffness bearing element in practice [48]. A sketch of the practical set-up is given in figure 5.31, a photograph is shown in figure 5.32. Vertical motion (i.e. the z-direction) has been studied around the zero stiffness working point. For simplicity, only one degree of freedom has been considered. Radial stability is provided by four air bearing pads. Although this is not the optimum solution in a kinematical sense, it is geometrically simple and sufficient for this purpose. Furthermore, angular stability has been achieved by locating the center of gravity below the point where the bearing force acts.

5.4.4.1 Bearing force and stiffness

The bearing force of the experimental 1-DoF bearing element as a function of the air gap height has been measured using a force sensor. The results are displayed in figure 5.33. The irregularities of the curve must be due to mechanical properties of the force sensor fixture, the inaccuracy in the position measurement and the influence of the air bearings, because there is no reason to assume that the actual force curve is *not* smooth. The force on the vertical axis is the rest force after choosing the suspended load in such a way that it nearly compensates for the bearing force in the working point, which is about 40 N.

The first derivative of this measured force is the stiffness, of which the most interesting part (around the working point) is shown in figure 5.34. This curve is a fitted first order polynomial through the numerically calculated stiffness data in this small area. From this graph it is clear that a stiffness $k_z < 1$ N / m is achieved in a region of about 1 μm. This corresponds to an eigenfrequency as low as 0.5 rad / s.

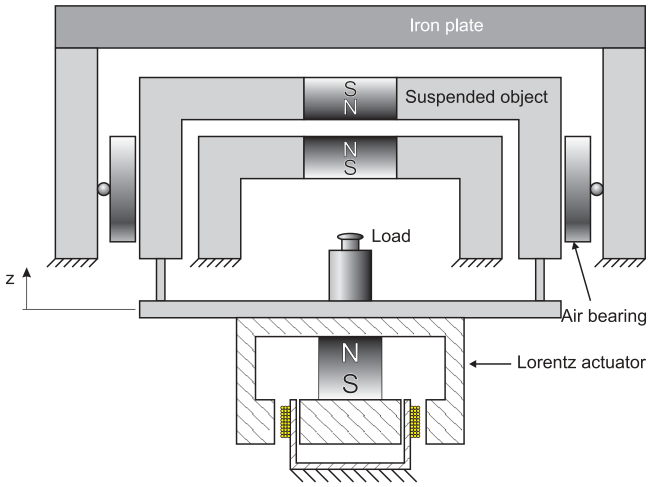


Figure 5.31: Schematical cross section of the experimental test set-up that is used to investigate the practical behavior of a zero–stiffness bearing element in the vertical (z-)direction. The load is chosen such that the bearing force just is in equilibrium with the weight. The Lorentz actuator is used during feedback control, exerting a force on the suspended object without adding stiffness.

5.4.4.2 Damping

First, consider the damping due to *air friction* in the lateral bearings. In the test set-up, having four circular air bearings of 40 mm diameter and a nominal air gap of 30 μm this gives for the damping coefficient (see also section 5.4.3.3):

$$c = \frac{\eta A}{h} = \frac{2 \cdot 10^{-5} \cdot \pi(\cdot 40 \cdot 10^{-3})^2}{30 \cdot 10^{-6}} \approx 3.4 \cdot 10^{-3} \text{ Ns/m}. \tag{5.42}$$

If there is a relative displacement z present with amplitude $\hat{z}$ and frequency ω,

$$z(t) = \hat{z} \cdot \sin(\omega t) \tag{5.43}$$

then the corresponding relative velocity $v = \dot{z}(t)$ will be

$$\dot{z}(t) = \hat{z} \cdot \omega \cdot \cos(\omega t) \tag{5.44}$$

so that the maximum damping force $F_{d,max}$ becomes

$$F_{d,max} = c \cdot v = 3.4 \cdot 10^{-3} \cdot \hat{z} \cdot \omega. \tag{5.45}$$

From equation 5.45 it is clear that even for rather extreme vibrations of, say, 10 μm amplitude at 600 rad/s, this maximum force will be very small, in the order of 10^{-5} N.

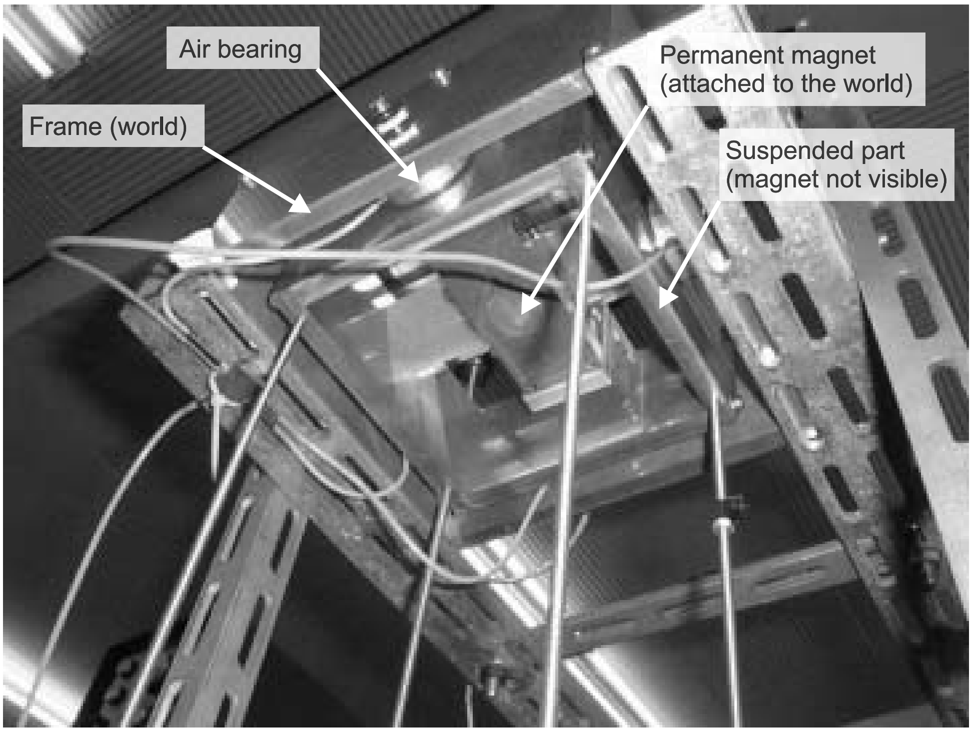


Figure 5.32: Photograph of the experimental zero stiffness bearing test set-up, seen from below. Visible are two of the four air bearings, the lower magnet (see also figure 5.31) which is attached to the frame, and the suspended part that is a frame containing a second magnet (not visible here). Four rods are attached to the suspended part, carrying a platen (not visible here) on which the load can be placed, in order to make an equilibrium with the bearing force.

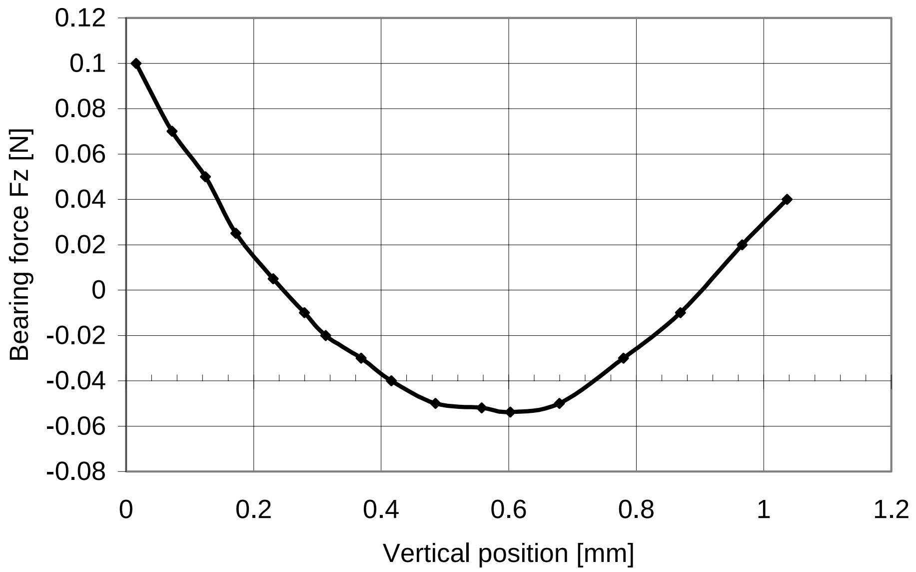


Figure 5.33: The measured vertical bearing force F_z of the one-DoF zero stiffness bearing element. The irregularities are induced by the measurement inaccuracy; the actual force curve should be very smooth. The measured force is the rest force after making the suspended weight approximately equal to the bearing force. The actual bearing force is about 40 *N.*

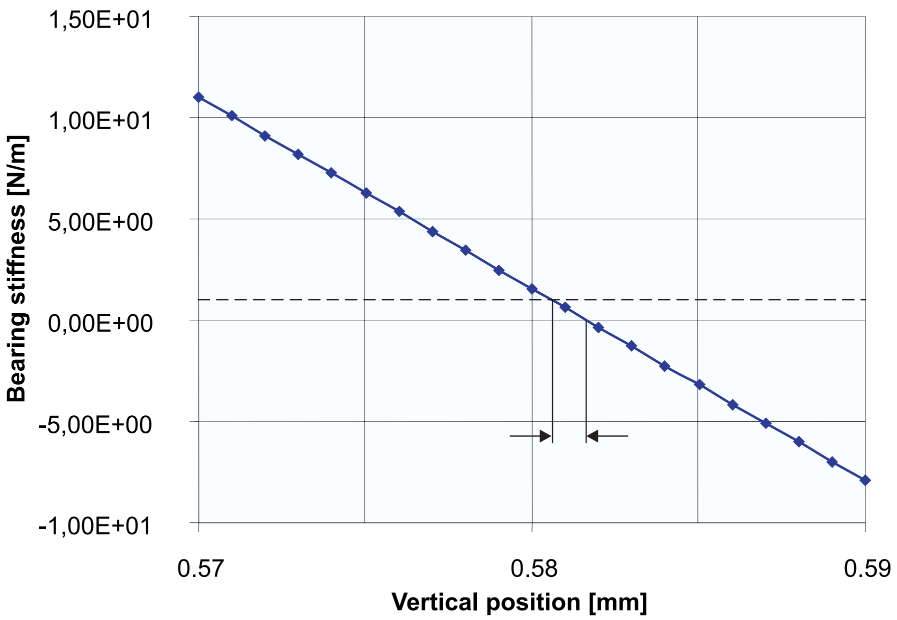


Figure 5.34: Linear approximation of the stiffness of the one-DoF zero stiffness bearing element, derived from the measured bearing force shown in figure 5.33.

Second, consider the damping as a consequence of *eddy currents* induced in the upper iron plate of the test set-up.

The flux density variations depend on the geometry and can be approximated by

$$\delta B = \frac{3 \cdot \delta z \cdot h^3}{8L^4} B_r \tag{5.46}$$

where δz is the vibration amplitude in m, h the magnet thickness in m, L the air gap in m and B_r the remanent induction of the magnet in T. When δz is in the order of magnitude 10^{-5} m and h and L are in the order of magnitude 10^{-2} m, δB will be about 10^{-3} T.

The eddy current power loss now can be approximated by [48]

$$P = \frac{1}{2\rho}(\omega \cdot \delta B)^2 L^3 A \tag{5.47}$$

in which ρ is the specific resistance in Ωm, ω is the vibration frequency in rad / s, δB the variations in the flux density in T through the eddy current area A in m^2.

Assuming vibrations in the order of 1 Hz, an estimated eddy current area A of order 10^{-4} m^2 and $\rho = 1 \cdot 10^{-7}$ Ωm, the eddy current power loss will become about $5 \cdot 10^{-10}$ W. Because

$$P = F \cdot v = (c \cdot v) \cdot v \Rightarrow c = \frac{P}{v^2} = \frac{P}{\dot{z}^2} \tag{5.48}$$

the damping constant $c \approx 10^{-2}$ Ns / m. This damping constant is an order of magnitude larger than the damping constant due to air friction, but even here the maximum disturbance force as a result of eddy currents will be very low, in the order of 10^{-4} N.

5.4.4.3 Control

In this test set-up only the position information has been used as an input for the stabilizing controller. Figure 5.35 shows a block scheme of the test set-up structure.

In order to design a control system for this experimental set-up, first a model of the zero stiffness bearing has to be built. For each of the blocks in the scheme of figure 5.35 the transfer function will be estimated.

Zero stiffness bearing

With regard to the stiffness, a simplification has been made to keep the model linear. A constant stiffness $k_z = 1$ N / m has been assumed, being approximately equal to the effective stiffness at $z = z_{wp} - 1$ μm.

With regard to the damping, based on the estimations as described in section 5.4.4.2, $c = 0.03$ has been assumed for the damping coefficient.

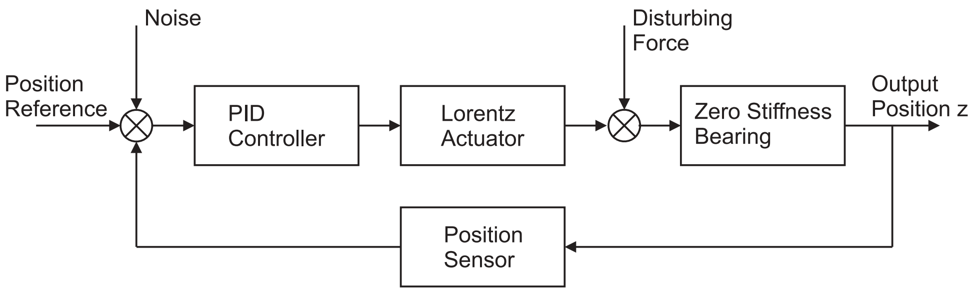


Figure 5.35: Block scheme of the structure of the one-DoF zero stiffness bearing element. In this test set-up only the position z has been used as a controller input.

Together with the mass $m \approx 4$ kg this gives an estimated transfer function H_{ZSMB} for the zero stiffness magnetic bearing element in the z-direction:

$$H_{ZSMB} = \frac{z(s)}{F_{dist}(s)} = \frac{1}{4s^2 + 0.03s + 1} \tag{5.49}$$

which results in a second order behaviour (open-loop), with a resonant frequency of about 0.5 rad / s.

Lorentz actuator

Because the Lorentz actuator's electrical time constant $\tau = \frac{L}{R} \approx 1.7$ ms, the bandwidth $\omega_b = \frac{1}{\tau} \approx 93$ Hz is sufficiently high for not disturbing the dynamics of the mechanical system. Therefore, the transfer function has been approximated by

$$H_{actuator} = \frac{F_L}{V_{act}} = 0.8 \ \ [\mathrm{N/V}] \tag{5.50}$$

Position sensor

As a position sensor an eddy current sensor has been used with a flat response up to 20 kHz. Its sensitivity is 11.2 V / mm. To reduce high-frequency noise, the sensor output has been filtered using a low-pass filter with a cross-over frequency of 100 Hz (resulting in a time constant $\tau = \frac{1}{2\pi \cdot 100 \text{ Hz}} \approx 1.6 \cdot 10^{-3}$). This gives for the sensor transfer function H_{sensor}:

$$H_{sensor} = \frac{11.2 \cdot 10^3}{1 + 1.6 \cdot 10^{-3} s} \tag{5.51}$$

PID controller

The controller has to be designed in such a way that the closed-loop system meets the following requirements:

- Stiffness less than 1 N / m
- Position deviation $|z - z_{wp}| < 1\ \mu$m
- The transmission of noise to the output position has to be minimized

Note that this experimental set-up is an example of a vibration isolating system according to figure 5.30. The goal is to achieve a low stiffness, while trying to keep the supported object close to the optimum working point.

After some calculations, the transfer function for the controller is set to

$$H_{controller} = 3.6 \cdot 10^{-6} \frac{s + 0.11}{s + 6.58} \tag{5.52}$$

which stabilizes the zero stiffness bearing by adding a (virtual) stiffness and damping. The problem in this experimental set-up (where only *position* feedback has been applied) is that the control action is too weak to keep the maximum position error within 1 μm. Increasing the gain to achieve the 1 μm position error limit increases the eigenfrequency to 10 Hz, which is too high to meet the requirements. From this it can be concluded that position feedback only is not sufficient to control the zero stiffness bearing element. Therefore, an acceleration feedback and possible also velocity feedback will be necessary. Because a velocity feedback introduces virtual damping and an acceleration feedback introduces virtual mass, the latter is preferable (see also table 5.4 on page 169).

5.4.4.4 Summary of results

From the investigation of this experimental 1-DoF zero stiffness bearing the following conclusions can be drawn:

1. Zero stiffness in bearing elements is feasible using permanent magnets. The measured test results are in accordance with the theory and the simulations. The test set-up showed a bearing force around 40 N in the working point and a stiffness $k_z < \pm 4$ N / m for $\delta z < \pm 5\ \mu$m. This is comparable to pneumatical vibration isolation systems.

2. The main advantage of using magnets in a bearing element is that the bearing has no mechanical contact and that there are no disturbing internal eigenfrequencies. The additional advantage of using permanent magnets is that there is no external power supply required, and that there is no heat-generating dissipation of energy, except for possible eddy current losses.

3. The achieved zero stiffness working point is marginally stable. Therefore, a controller is needed to keep the bearing element in its working point.

4. In practice, the position feedback cannot be made arbitrarily small because it is related to the magnitude of the disturbing forces that have to be compensated for. This causes a too high virtual stiffness. Therefore, acceleration feedback might be considered to create also a virtual mass (which results in smaller amplitudes and a lower eigenfrequency). Possibly also a certain velocity feedback will be desird in order to realize a virtual damping. Eventually, a compromise has to be found between sufficient vibration isolation and allowable disturbance amplitude.

5.5 A zero stiffness linear motion system

After realizing zero stiffness in permanent magnetic bearings in one degree of freedom, the next step is to extend the motion range of the bearing in one dimension. This would enable a long range displacement in one degree of freedom (e.g. the x-direction) while maintaining the vibration isolation properties in most of the other degrees of freedom.

In this section, first the theoretical modeling of these linear bearing elements will be discussed. After that, the design and test will be presented of an experimental straight motion system that has zero stiffness in both the travel direction (x) and the vertical direction (z).

5.5.1 Modeling of linear permanent magnetic zero stiffness elements

The radially symmetrical set-up as depicted in figure 5.9 can also be considered as being a cross section of a linear bearing, where the upper and lower parts connected to the world have infinite length and the middle, floating, part has a length in the same order of magnitude as its width.

It will be plausible that there will be a zero stiffness working point in this configuration as well. However, there is at least one significant difference.

When recalling figure 5.24 and appendix A, comparing the configurations with and without ferromagnetic material, it is clear that in this linear case always $k_x = 0$, provided that the upper and lower magnet arrays are sufficiently long. That means that, because $k_x + k_y + k_z \leq 0$, here $k_y \leq -k_z$. For this setup, when the geometry is optimized for zero stiffness in the vertical direction ($k_z = 0$), there will be a negative or zero stiffness in the lateral, horizontal, direction ($k_y \leq 0$). When in the magnet configuration only materials having $\mu_r = 1$ are applied, then $k_y = 0$. The more ferromagnetic material (having $\mu_r \gg 1$) is applied, the more negative k_y will be.

However, an iron yoke is desired to minimize stray fields and to shield the bearing element from disturbing external electromagnetic influences. To minimize the lateral stiffness effect mentioned above, it is recommended that the suspended magnet ‘sees’ as little iron as possible. Therefore, to reduce lateral stiffness, it is better that also the upper part is provided with a magnet array, although for the generation of the attracting part of the vertical force an iron yoke would have been sufficient.

Based on these considerations and using magnet dimensions that are standard available, the experimental model as described in section 5.5.2 has been designed.

A possible method to improve the vibration isolation properties also in the lateral y-direction, is to change the geometry in such a way that the lateral stiffness k_y becomes equal to zero in the same point where the vertical bearing stiffness k_z is zero. This is illustrated in figure 5.36. Using a FEA model, as displayed there, as an example, the force–displacement characteristic has been calculated in a certain vertical (z) range and a certain horizontal (y) range simultaneously, resulting in a matrix of values for $F_y(y,z)$ and $F_z(y,z)$. After that, the minimum value for every F_z can be approximated at $y = 0$ by fitting a sixth order polynomial through the data. Subsequently, at the calculated height z_{wp} the corresponding $F_y(y, z = z_{wp})$ can be approximated by interpolating the known values for $F_y(y,z)$. After that, the distance between the lower two magnets is increased symmetrically, so that the lateral stiffness will become less negative (although the vertical stiffness will change as well). Then the FEA procedure is repeated to get other sets of $F_{y,z}(y,z)$. This iterative process will lead to a value for the magnet distance for which the vertical working point coincides with the point where the lateral stiffness is equal to zero.

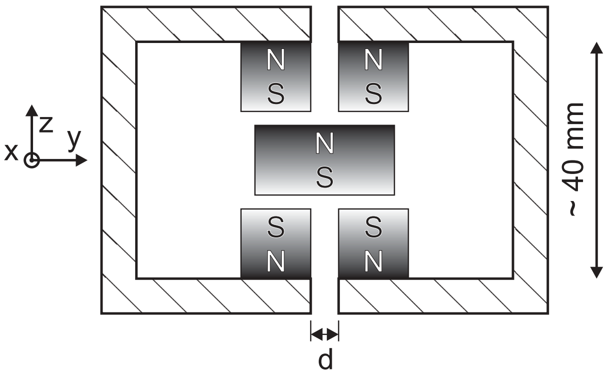


Figure 5.36: Example of a linear permanent magnet configuration (cross section), used in an FEA in order to investigate whether it is possible to make k_y equal to zero in the working point where $k_z = 0$. The center magnet moves in y (left/right) and z (up/down), while the forces F_y and F_z are calculated. The horizontal distance between the lower two magnets, d, is increased from 4 mm to 6, 8, 10 and 12 mm, as shown in figures 5.37 and 5.38. The air gap above and below the center magnet in its central position is 2 mm in this picture.

The results of these calculations can be found in figures 5.37 and 5.38. Except for the obvious outliers and inaccuracies, caused by the limited accuracy of the FEA, the figures give a good impression of the behavior of the linear bearing element. Indeed it appears that with an increasing gap between the lower magnets the lateral stiffness in the working point changes from negative to positive. Compare for instance the graphs for $d = 4$ mm and $d = 8$ mm in figure 5.38.

These results seem to contradict the theory as discussed in appendix A about lateral stiffness and $\mu_r \neq 1$. Based on this theory the lateral stiffness k_y is expected to be negative when both k_z and k_x are equal to zero and when there is high-μ_r material near the force generating magnets. A possible explanation is that in this configuration the ferromagnetic parts have sufficient distance to the suspended permanent

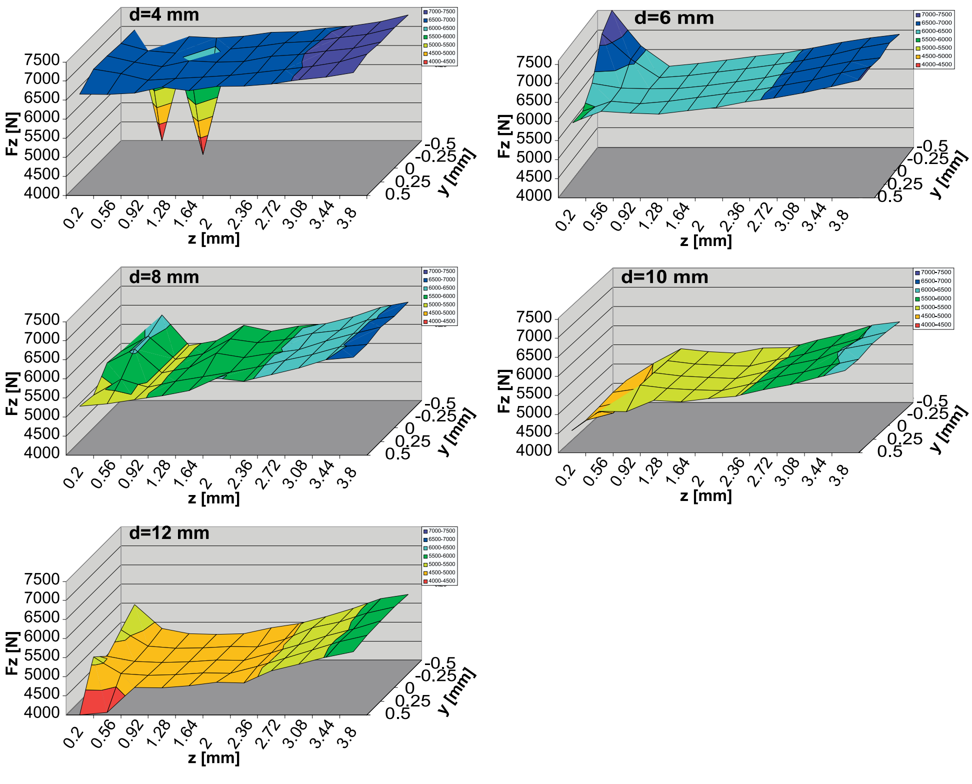


Figure 5.37: Values for the vertical bearing force $F_z(y,z)$, obtained from the FEA. The values for d refer to the horizontal distance between the lower two magnets, as depicted in figure 5.36. The scales are identical; the forces are displayed in N per meter length perpendicular to the cross-section. For each d there appears to be a point z_{wp} at $y = 0$ where $\frac{dF_z}{dz} = 0$.

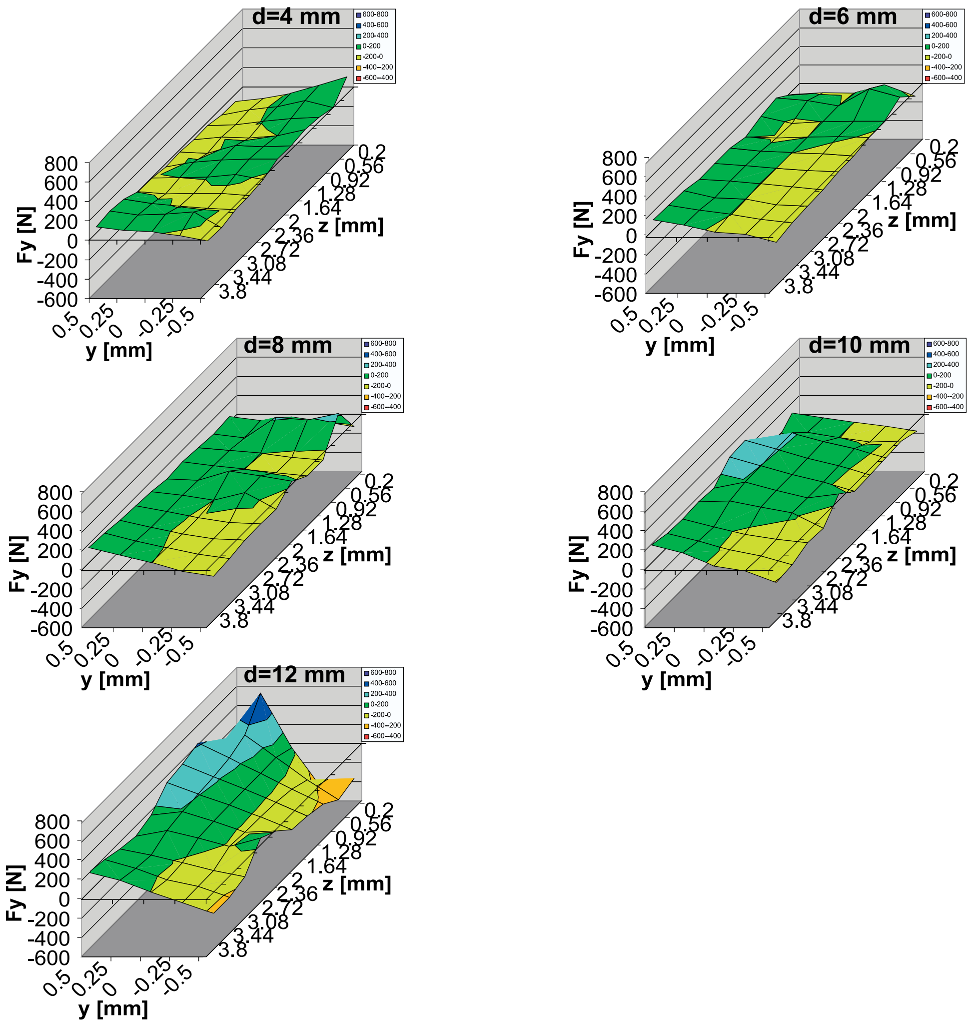


Figure 5.38: Values for the lateral force $F_y(y,z)$, obtained from the FEA. The values for d refer to the horizontal distance between the lower two magnets, as depicted in figure 5.36. The scales are identical; the forces are displayed in N per meter length perpendicular to the cross-section. It appears that for $d \approx 6$ mm the lateral stiffness becomes equal to zero while $z = z_{wp}$.

magnet, so that their influence on the stiffness will be small.

5.5.2 Experimental linear permanent magnetic zero stiffness bearing

Based on the principle of permanent magnetic zero stiffness as described in section 5.4 and the considerations in section 5.5, a linear version of this bearing element has been designed. In this section, an outline of this design will be presented. Unfortunately, due to lack of time, the straight motion system designed in this section has not been completely realized. Therefore, experimental data is not available, except for the permanent magnetic bearing unit, which has been built and tested. At the Laboratory for Micro Engineering, research activities and experiments on this subject are continued.

A sketch of the bearing principle is given in figure 5.39. The guidance base consists of four linear magnet arrays kept together by an iron yoke. The moving part of the bearing element contains an aluminum frame and two magnets. All magnets are $30 \times 8.5 \times 2$ mm^3 NdFeB magnets. In figure 5.40 a cross-sectional view of the experimental set-up is given. An overview sketch of the bearing system is shown in figure 5.41.

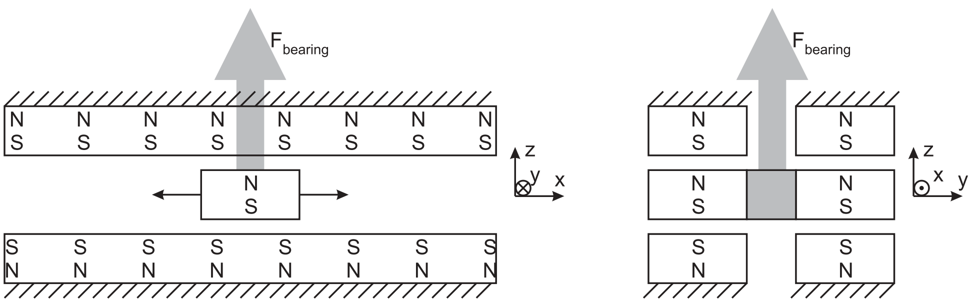


Figure 5.39: Principle sketch of a permanent magnet configuration for application in a linear bearing system. Similar to the configuration as sketched in figure 5.9, there is a working point where the vertical stiffness, k_z, is equal to zero, while a net vertical force is generated. However, in this linear configuration such a working point exists for every x. Because, when the guiding magnet arrays are sufficiently long, no force is generated in the x-direction, the horizontal stiffness k_x is equal to zero. In the y-direction the stiffness depends on the presence of a possible iron yoke (which is not shown here, but in most cases this will be desired for magnetic shielding purposes). When all materials near the bearing element have $\mu_r = 1$, also the lateral stiffness, k_y, will be equal to zero.

Based on a finite element analysis it is estimated that the lateral stiffness (in the y-direction) is negative but rather small, in the order of -10^3 N/m. To overcome

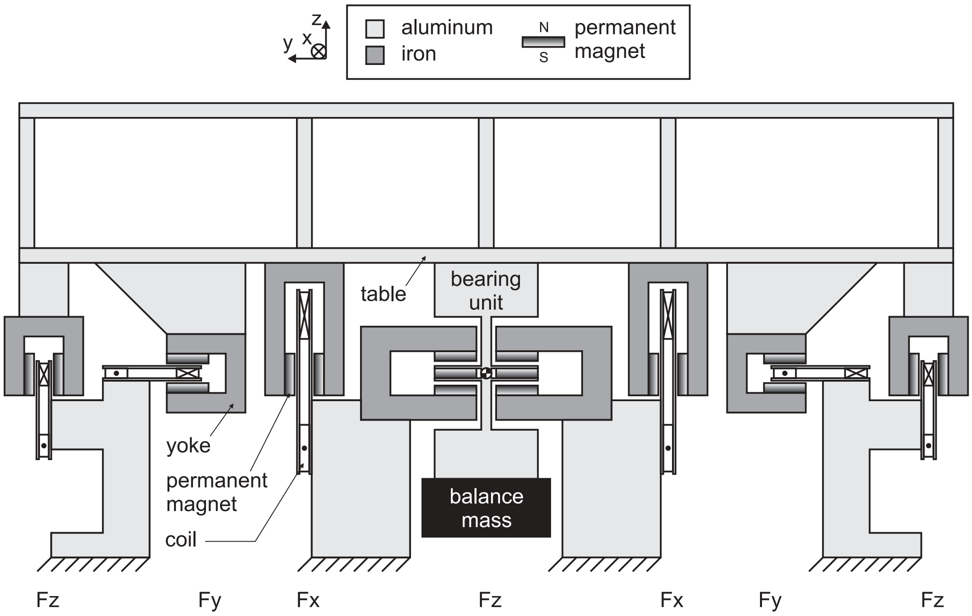


Figure 5.40: Cross section through the yz-plane of the experimental linear zero stiffness magnetic bearing. The bearing element is located in the center. At the sides, twelve yokes (of which six are visible in this cross section) have been added to generate Lorentz forces needed for propulsion and control in all six degrees of freedom. These yokes, together with the balance mass, have been located in such a way that all resulting force vectors go through the center of gravity. The necessary sensor system has been omitted in this figure. In order to realize a vibration isolated positioning system according to figure 5.30, the sensors should measure the position relative to a straightness reference that has no position coupling to the world. This aspect is left out of consideration here.

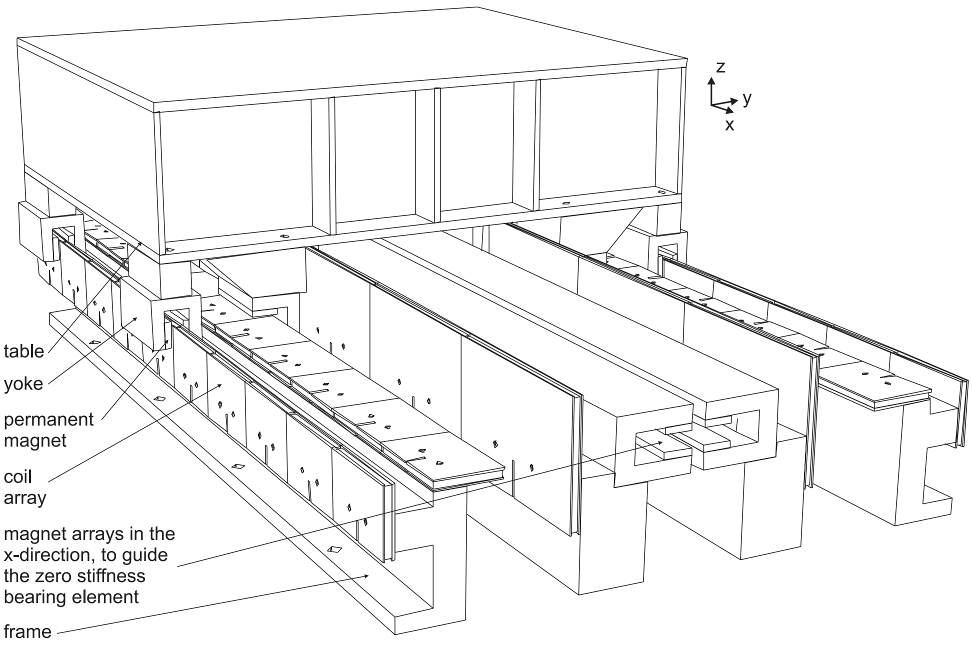


Figure 5.41: Overview sketch of the linear zero stiffness magnetic bearing system. Clearly visible are the magnet arrays and the coil arrays in the x-direction and some of the C-shaped Lorentz actuator yokes attached to the moving table.

this negative stiffness, active control will be necessary. To avoid the generation of a virtual stiffness coupling by the actuators needed for this active control, a Lorentz type of actuator is the most suitable, as is already discussed in the previous sections. Because the bearing element has no inherent stability in any degree of freedom (all stiffnesses are equal to zero or slightly negative), a 6-DoF control has to be applied.

To enable unlimited travel range in the x-direction all control force generating elements (Lorentz actuators) have to be designed in such a way, that either an arbitrarily large number of coils can be linked (as is the case for the y- and z-actuators), or commutation is possible. Commutation will only be necessary in the x-direction, when a travel range larger than the coil pitch is desired. This is explained further in figure 5.42.

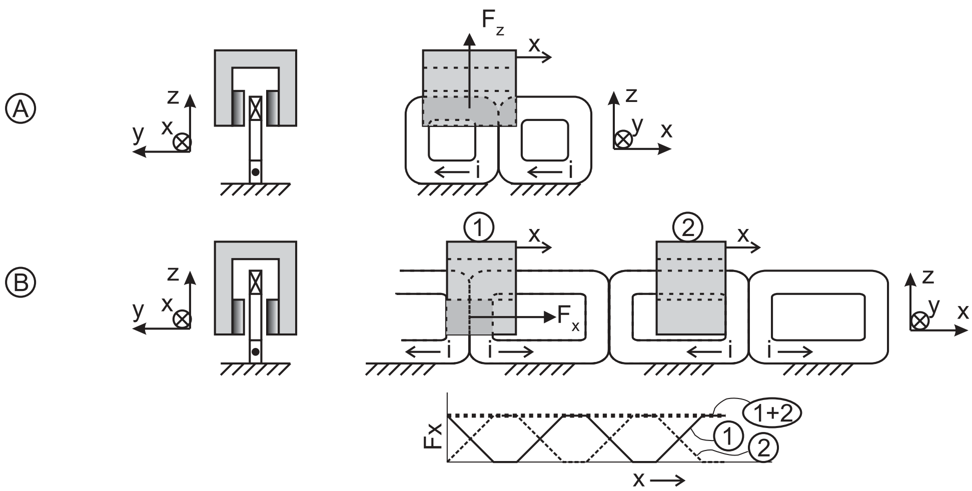


Figure 5.42: Schematic drawing of the coils and yoke parts for the generation of a Lorentz force. A. Front view and side view of an actuator generating a force in the z-direction. By choosing the yoke width equal to the coil width, the force will be constant and independent on x. In this configuration, an arbitrarily large number of coils can be used to create any desired travel length in the x-direction. An identical configuration (but rotated 90 *degrees around the x-axis) is used for the y-direction. B. Front view and side view of an actuator generating a force in the x-direction. Because the force generating parts of the coils are perpendicular to the direction of travel, it is necessary to switch the direction of the current through the coils dependent on the x-position (i.e. to apply commutation), in order to prevent the generated force from changing sign. Moreover, the additional second yoke must be shifted half of the coil pitch in order to give a constant resultant force, as is illustrated in the lower graph.*

As a sensor system, a single-beam laser straightness sensor has been chosen, which has been presented in section 2.2.2.2. Together with an optional measurement system for the x-direction, this leads to a 6-DoF controlled linear motion system that is virtually isolated from external vibrations.

5.5.2.1 Modeling of the linear zero stiffness bearing

In section 5.4.3.4 the theoretical modeling of a zero stiffness permanent magnetic bearing has been discussed. Based on the model structure derived there, a model for this experimental linear version has been made. The goal is to build a model that predicts the parasitic forces and stiffnesses, in order to be able to compensate for these undesired effects in the final controller.

The eventual closed-loop positioning system has two important properties: the *tracking performance* and the *disturbance sensitivity*. The first one has to be as high as possible, the second one has to be as low as possible. The disturbance sensitivity can be minimized by minimizing the effective stiffness. This is realized by means of feedforward corrections in the actual controller, which are described hereafter. The part of the controller dealing with tracking performance is left out of consideration here.

First, the bearing forces of the linear bearing element have been measured. In figure 5.43 a sketch of the measurement set-up for the y-direction is given. The force in the z-direction has been measured in a similar way. The figures 5.44 and 5.45 show the outcome of the measurements of the force in the (vertical) z- and (transversal) y-direction of the central bearing element, from which the stiffness k_y and k_z in the working point can be derived.

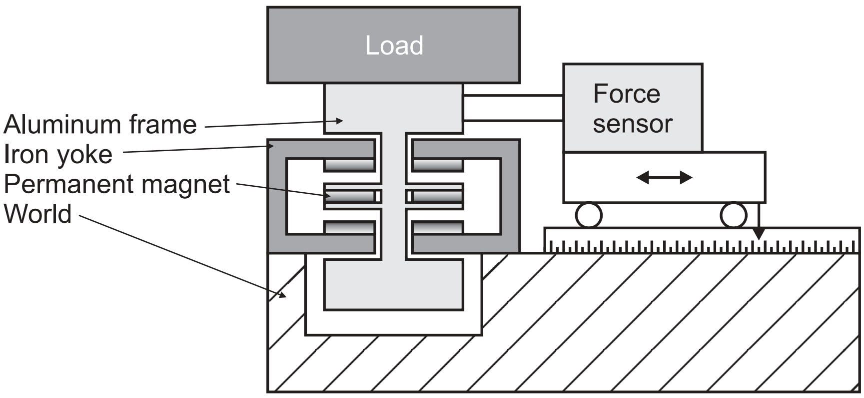


Figure 5.43: Schematical test set-up for measuring the force, generated by the linear bearing element, in the transversal y-direction. The force in the z-direction has been measured in an analogous way.

Due to the relatively long stroke in the x-direction, the stiffness k_x can be assumed to be zero, at least near the center of the linear guidance. At the point where $k_z = 0$

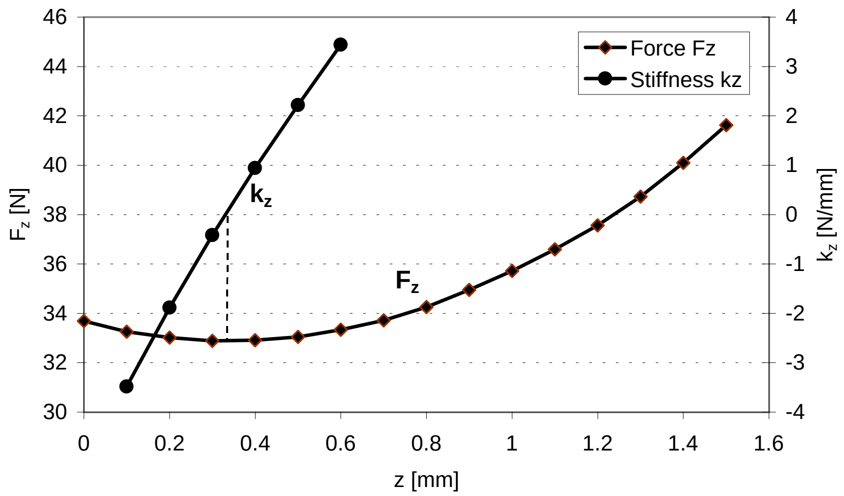


Figure 5.44: The measured force F_z and the stiffness k_z of the linear zero stiffness bearing element, as presented in the figures 5.40 and 5.41, in the z-direction, with y centered.

Figure 5.46 gives an overview of the entire compensator system block scheme as it is modeled in Simulink. To the left are three groups of four inputs, representing the current from each quadrant of the three QPDs that form the detector part of the single laser beam straightness sensor (see section 2.2.2.2). They are labeled green, red and blue, respectively, while the four quadrants in turn are labeled red, green, blue and yellow. The output of the individual quadrants of each QPD results in a local (x, y) pair, representing the location of the laser spot's mean intensity. These (x, y) pairs are converted to orthogonal coordinates y, z, φ_x, φ_y and φ_z, representing the location and orientation of the moving table's center of gravity, in- order to construct a feedforward compensator that consists of five independent elements instead of one multivariable structure. To the right are the five outputs for each of the five degrees of freedom that have to be addressed. After these outputs, another transformation matrix (not shown in figure 5.46) has to be placed, which converts the control signals to the actual driving signals for each of the actuators. The controller for the x-direction, the direction of propulsion, has been left out of consideration here, because it is less interesting from zero stiffness bearing point of view. The reason that the x-signal still is present in the calculations is that the measured signals for the rotations φ_y and φ_z depend on x. Hereafter the main parts of the compensator system will be described.

The five parallel conmpensator blocks for each of the five degrees of freedom will

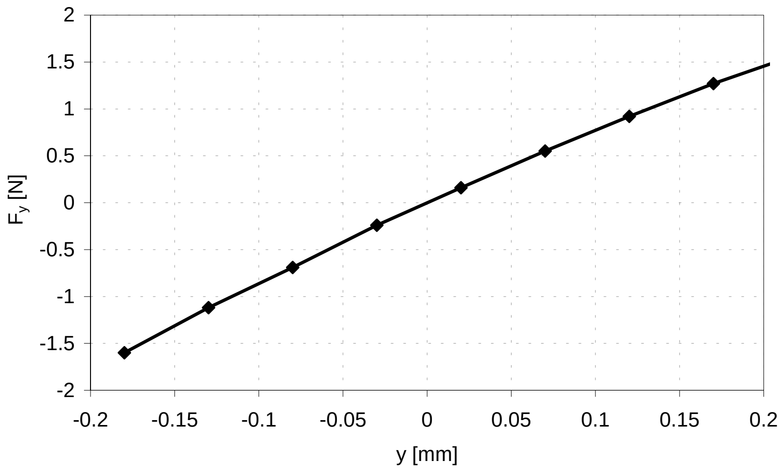


Figure 5.45: The measured force of the linear zero stiffness bearing element, as presented in the figures 5.40 and 5.41, in the y-direction, at $z = z_{wp}$. The value of z_{wp} was determined from figure 5.44. The graph shows that the lateral stiffness is nearly linear, $k_y \approx -7 \cdot 10^3\ N/m$.

be shortly described here (except the controller for the x-coordinate, which is less interesting from zero stiffness bearing point of view). An overview of the five block schemes is given in figure 5.49. To clarify the derivation of the influence of the angular deviations, in figure 5.48 some sketches are drawn.

Cntr_y: The compensator for the y-direction has been designed in such a way that the negative stiffness k_y is eliminated by a positive stiffness. Therefore, the compensating force equals

$$F_y(y) = -k_y \cdot y \tag{5.53}$$

Cntr_z: The block for the z-direction consists of a 'feedforward compensator', which provides a correction force of magnitude

$$F_z(z) = -\kappa \cdot z^2 \tag{5.54}$$

so that the non-linear behavior of the bearing element in the z-direction is eliminated in first order.

Cntr_Rx: A rotation around the x-axis does not result in a moment, but only in a parasitic force that works in the z-direction. This force can be approximated in first order by

$$F_z(\varphi_x) = 2 \cdot \kappa(z - z_{wp})^2 = \frac{\kappa b^2}{2} \cdot \varphi_x^2 \tag{5.55}$$

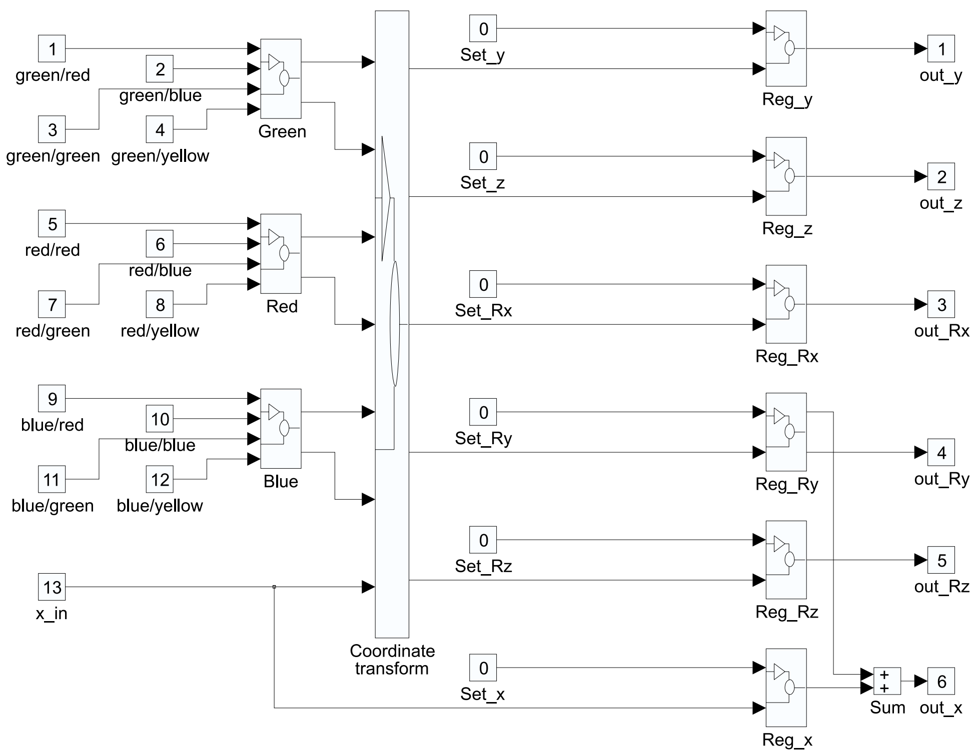


Figure 5.46: Overview of the entire block scheme of the feedforward compensation, necessary for the position controller to minimize the vibration transmission. To the left are three groups of four inputs, representing the quadrants of the three QPDs that form the detector part of the single laser beam straightness sensor. To the right are the five outputs for each of the five degrees of freedom that have to be addressed.

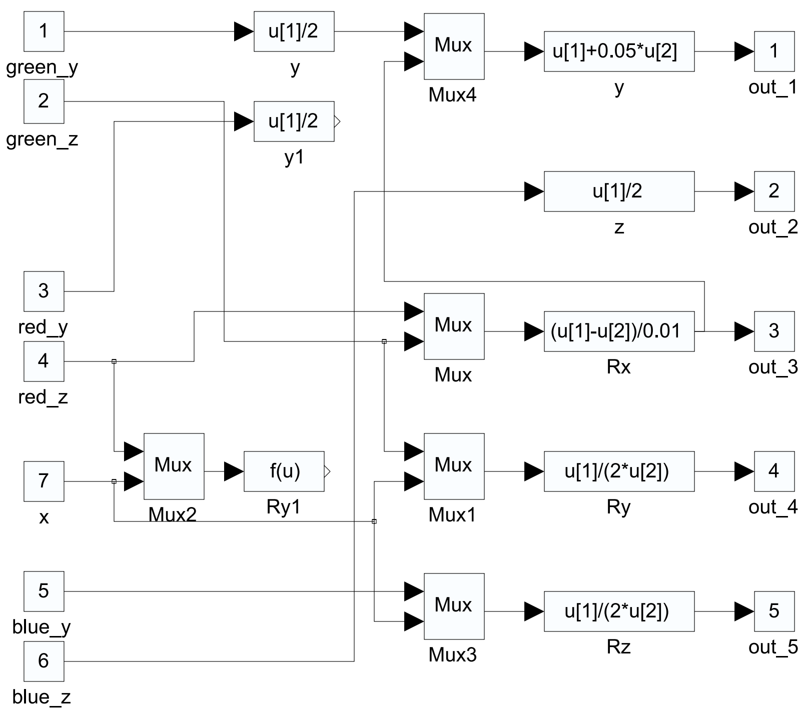


Figure 5.47: The coordinate transform, part of figure 5.46, that converts the outputs of the QPDs to the actual coordinates of the moving table. Note that in the block schemes the name 'Ri' has been used instead of φ_i. The outputs 'y1' and 'Ry1' are redundant and might be used for verification purposes.

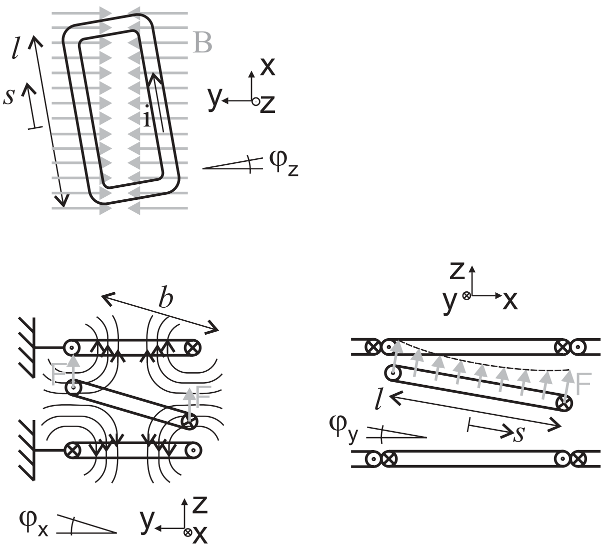


Figure 5.48: Simplified sketches of one half (in the y-direction) of the permanent magnetic linear bearing element. The figures refer to the description of the feedforward compensator equations in Cntr_Rx, Cntr_Ry and Cntr_Rz.

in which b is the width of a single magnet in the y-direction.

Cntr_Ry: The compensator for φ_y resembles that for φ_x. Also here, a force in the z-direction is generated as a result of a rotation φ_y. This force is approximated by

$$F_z(\varphi_y) = 2\cdot\cos\varphi_y \int_{s=-\frac{l}{2}}^{\frac{l}{2}} \kappa(s\varphi_y)^2 ds = 2\cdot\cos\varphi_y\kappa\varphi_y^2\left[\frac{1}{3}s^3\right]_{-\frac{l}{2}}^{\frac{l}{2}} = 2\cdot\frac{\kappa l^3}{12}\cdot\varphi_y^2\cos\varphi_y \tag{5.56}$$

where l is the length of a magnet in the x-direction. Besides that, a force in the x-direction is generated, which is equal to

$$F_x(\varphi_y) = \frac{F_z(\varphi_y)}{\tan\varphi_y} = 2\cdot\frac{\kappa l^3}{12}\cdot\varphi_y^2\sin\varphi_y \tag{5.57}$$

In practice, the increasing effect on F_z due to the quadratic force curve ($F_z = \kappa z^2$) will partially outweigh the decreasing effect due to $\cos\varphi_y < 1$.

Cntr_Rz: Because the lateral stiffness k_y is not equal to zero, a force in the y-direction will be generated as a result of a rotation around the z-axis. This force is approximated by

$$F_y(\varphi_z) = \int_{s=-\frac{l}{2}}^{\frac{l}{2}} s^2\varphi_z k_y ds = k_y\varphi_z\left[\frac{1}{3}s^3\right]_{-\frac{l}{2}}^{\frac{l}{2}} = \frac{k_y l^3}{12}\cdot\varphi_z \tag{5.58}$$

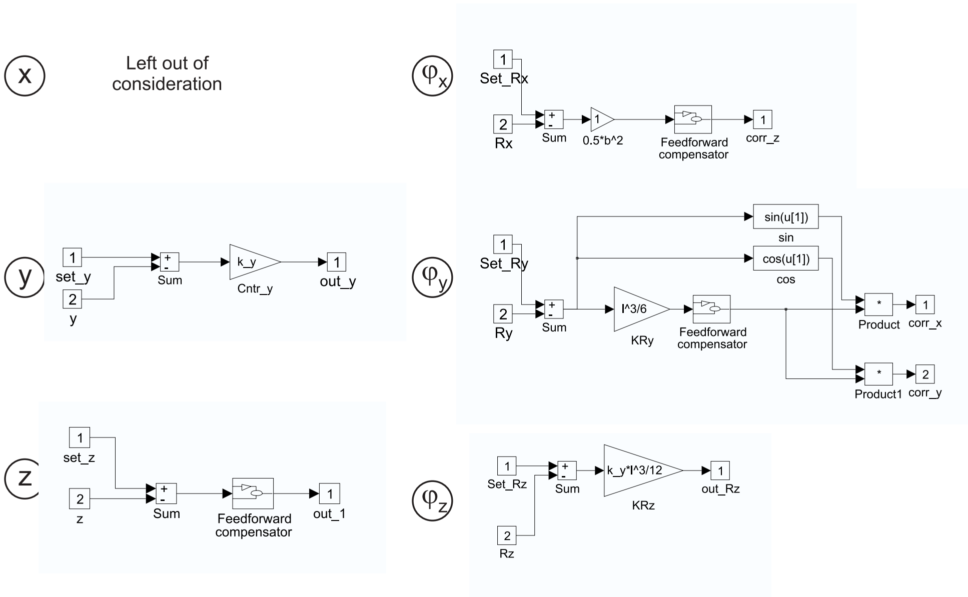


Figure 5.49: Overview of all five parallel controllers (parts of figure 5.46) for the five degrees of freedom.

In figure 5.50 the block scheme of the feedforward compensation is given. It simply generates a signal for the correction force in the z-direction, being equal to $F_{corr} = \kappa \cdot z^2$. Here it is assumed that the force–displacement curve of the vertical bearing force is parabolic, which is a fair first order approximation. By applying this feedforward compensation, the largest part of the non-linear effect will be eliminated, which simplifies the other controller parts.

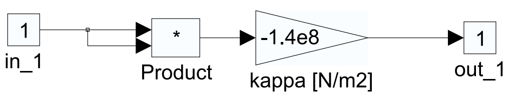


Figure 5.50: Block scheme of the feedforward action added to some of the controllers described above. The force–displacement curve of the bearing element in the z-direction is here assumed to be quadratic.

In table 5.5 a summary of the most important parameters is shown. Due to lack of time the experimental set-up has not been finished completely, so that only a part of the desired measurement values could be collected. However, that part of the data obtained matches the expectations based on calculations and simulations. At the Laboratory for Micro Engineering, one of the currently running projects continues the research in the field of linear positioning, using the permanent magnetic zero stiffness principle presented in this section.

Parameter	Value	Unit	Remark
Stroke	30	mm	Unlimited when commutated
Propulsion force	8	N	4 magnet yokes
Stiffness k_x	0	N/m	Neglecting edge effects
Stiffness k_y	$-7 \cdot 10^3$	N/m	Without controller
	...	N/m	With controller
Stiffness k_z	0	N/m	In the working point
	70	N/m	$\delta z = 5\ \mu$m
Eigenfrequency f_z	...	Hz	Vibration amplitude ... μm
Maximum position error	...	μm	
Vibration transmissibility	...	dB	$f <$ ... Hz

Table 5.5: Calculated and/or measured data from the linear zero stiffness permanent magnetic bearing. Of the system designed in this section only the permanent magnetic bearing unit has been built. Because of this, some values (that should have been measured in a practical test set-up) are missing.

5.5.3 Concluding remarks with regard to zero stiffness bearings and actuators

1. By means of zero stiffness bearings or actuators, vibration isolating properties can be incorporated in feedback controlled positioning systems.

2. Electromagnetic reluctance actuators are not specifically suitable for zero stiffness positioning. Their application requires a large amount of control effort.

3. Lorentz actuators essentially have zero stiffness. Therefore, they are suitable for zero stiffness positioning applications. Their largest disadvantage is the dissipation of energy, and thus the generation of heat, when a static force has to be generated.

4. The concept of *permanent magnetic* zero stiffness bearings presented in this chapter, is feasible. The advantage with respect to Lorentz actuators is that permanent magnetic bearings do not dissipate energy while generating a static force.

5. The permanent magnetic zero stiffness bearing requires active control. The neutral equilibrium having marginal stability has to be maintained by means of position feedback and (preferably) acceleration feedback. When used merely as a vibration isolating system, a compromise has to be found between stability and vibration transmissibility, just like conventional vibration isolating systems. When used as a *vibration isolated positioning system*, following an external reference, the tracking performance and the isolation quality both can be optimized at the same time.

6. The concept of the zero stiffness bearing element can be extended to a linear motion system, so that a contactless straight motion system can be designed, having vibration isolating properties in one or more degrees of freedom.

Chapter 6

Conclusions and recommendations

6.1 Conclusions

In this chapter, general concluding remarks will be made. Besides these general conclusions, most sections in this thesis also have a paragraph containing a more detailed evaluation.

Section 6.1.1 contains the main conclusions of the research described in this thesis. The sections 6.1.2 and further give the less important—although relevant—concluding remarks, grouped according to subject, mainly following the chapter division. The numbers in the margin refer to the corresponding sections.

6.1.1 Main conclusions

1. A modular approach in designing and improving straight motion systems is effective. Using relatively simple means—commercially available actuators, a simple sensor system, mechanical parts having tolerances of about 0.1 mm, standard feedback control theory—the positioning accuracy of an experimental straight motion system has been decreased from typically 10 μm to less than 1 μm. Straightness error compensation by means of feedback control enables operation of straight motion systems in the sub-micrometer range. Using only conventional technology, this sub-micrometer performance would have been possible only under very strict conditions with regard to e.g. mechanical load and thermal influences.

2. The straightness performance of feedback controlled positioning systems is determined to a large extent by the sensor system and the accompanying

straightness reference. Therefore, it is necessary to pay much attention to the choice of both the sensor and the reference. Calibration techniques like straightedge reversal are useful when the straightness reference has a long-term form stability better than the required accuracy tolerance.

3. In closed-loop high-precision positioning, it can be very advantageous to aim for *force* control instead of the usual *position* control. This requires position (and possible velocity and acceleration) feedback, while ideally the stiffness coupling—relating position and force—is eliminated. Essential in this approach is the use of zero-stiffness actuators, because they incorporate the separation between force and position. The application of non-zero-stiffness actuators would make very high demands on the controller performance. By means of zero-stiffness actuators, it becomes possible to create a virtual stiffness between two independent objects, which both can be isolated from environmental vibrations.

4. When regarding the development from simple mechanical guidances via high-end (open-loop) mechanical guidances towards advanced active (closed-loop) straight motion systems, the following types of systems can be distinguished.

Simple open-loop systems: In the most straightforward open-loop straight motion systems, position errors are mainly caused by

(a) guidance imperfections
(b) bearing irregularities
(c) external forces
(d) thermal expansion
(e) wear
(f) external vibrations

High-quality open-loop systems: By spending much effort to reduce guidance imperfections, bearing irregularities and thermal deformations, higher accuracy can be achieved. However, the influence of *external forces* and *vibrations* remains, and there is no real-time information about the actual straightness errors. Therefore, to guarantee a high accuracy, strict conditions have to be met, with respect to vibrations and external forces.

An example of an improved open-loop straight motion system is given in section 4.4, where a high-stiffness passive bearing has been presented.

Closed-loop systems, quasi-statical operation: This kind of systems result when a straight motion system from one of the previous categories is augmented by a feedback controlled reduction of straightness errors. In these systems, the control loop eliminates or reduces the influence of the mechanical errors. The demanding problems will be transferred from the mechanical domain towards the electronics and control engineering

domain. The *sensor resolution* is the most important limiting factor for the positioning accuracy.
In this thesis, this has been shown in section 4.2. The reduction factor of the disturbances also depends on the actuator stroke, because at low frequencies virtually every position error occurring within the control range can be eliminated. The 25 μm actuator stroke together with the 0.2 μm sensor resolution gives a maximum reduction factor 125.

Closed-loop systems, dynamical operation: Like in the previous—quasi-statical—category, the sensor resolution is the most important limiting factor, but now also the *control bandwidth* (limited by *internal structural dynamics*) and *environmental disturbances* c.q. vibrations become important.
In this thesis, for the specific straight motion system described in section 4.2, it appears that also the stacked piezo actuator's hysteresis (being a part of the internal structural dynamics) has a negative influence on the system's closed-loop bandwidth.

Closed-loop systems, zero-stiffness: A further improvement in positioning accuracy can be realized when the mechanical coupling to the world is eliminated by means of *zero stiffness actuators*, so that a force controlled positioning system results, insensitive to the position of each actuator relative to the world. In that case, only the *internal error sources* limit the performance, but the errors that are generated within the positioning system itself often are small or well known.
Applying the permanent magnetic zero stiffness bearing principle, presented in chapter 5, a very compact suspension system can be realized, having a stiffness that is virtually equal to zero.
When a permanent magnetic zero stiffness bearing element is added to a zero stiffness actuator like a Lorentz actuator, a versatile actuator results, which combines vibration isolation with control force generation. This type of actuators is very useful in high-precision positioning applications, where external position disturbances (i.e. vibrations) would have a too large disturbing influence on the positioning of two objects relative to each other.

6.1.2 Conclusions with regard to sensors

1. Position sensors are of crucial importance in feedback controlled positioning systems. In many applications, the performance of the sensor limits the performance of the entire closed-loop system.

2. In most cases measuring straightness is realized by measurement of positions in five degrees of freedom, so that five *position* sensors form one straightness sensor. This implies the application of a coordinate transform algo-

rithm, when a decoupled controller—consisting of five independent simple controllers instead of one complex multivariable controller—is desired.

3. In this thesis two new, low-cost straightness sensors have been presented. The first one is a capacitive sensor, consisting of five electrodes, a multiplexer and oscillator ASIC, resulting in a compact sensor unit measuring five degrees of freedom quasi-statically, having a resolution in the order of 0.2 μm. The second one is an optical sensor, consisting of three quadrant cell photodiodes together with several glass components, enabling the detection of five degrees of freedom of a moving object, using only one laser beam. This optical sensor also can have a sub-micrometer resolution.
   Besides these two straightness sensors, also a novel sensing principle has been investigated, or rather a new combination of two capacitive sensing methods. The application of a position sensor having a low-pass characteristic and a velocity sensor having a high-pass characteristic results in a position and / or velocity sensor having a wide bandwidth ($> 1kHz$), good accuracy (sub-μm) and low cost.

4. Capacitive sensors are suitable for high-precision straight motion systems. Using simple electrode structures, straightness deviations can be measured relative to a reference bar. Due to the spatial averaging, the roughness of the reference bar does not influence the quality of the straightness measurement.

5. The most important requirement to straightness references is *form stability*. A known form stability enables calibrating the remaining straightness errors of the reference itself, by means of e.g. reversal techniques. However, even without such a calibration, absolute *repeatability* can be achieved. Preferably, a straightness reference is included in a metrology frame, so that no external forces act upon it.

6. There is a trend visible concerning miniaturization of position sensors towards the IC level, resulting in integrated sensor elements together with all necessary signal conditioning electronics in a relatively small volume, which can be produced in large quantities for a low price. In the near future very compact, low-cost and accurate sensor systems can be expected to become available.

6.1.3 Conclusions with regard to current technology

1. From literature research it appears that there are very few straight motion systems that work in the sub-micrometer or nanometer range. For open-loop systems, sub-micrometer straightness only can be realized under very strict conditions. In particular, it appears to be very difficult to obtain sub-micrometer accuracy in combination with a travel range in the order of 100 mm.

2. None of the systems currently known meets all the specifications required for operation on nanometer scale with a travel range in the order of 100 mm.

3. To enable working in the accuracy range of 1...100 nm—below current manufacturing accuracy—feedback control appears to be indispensable, in order to cope with mechanical disturbances and deformations. By means of feedback controlled straightness error compensation, a system's positioning performance can be made virtually independent on mechanical error sources including manufacturing inaccuracies, external forces, thermal effects and wear.

6.1.4 Conclusions with regard to feedback controlled straight motion systems

1. The application of a straightness error compensating feedback control system increases the complexity and cost of the total system, but results in a net improvement of the performance. In the experimental system described in this thesis, which is based on a coarse bearing system having straightness errors in the order of 10 μm, the straightness errors of the positioning table are reduced to less than 0.5 μm (which is close to the sensor's resolution) relative to a straightness reference.

2. Stacked piezo actuators are suitable for precision positioning purposes. However, the performance may be limited due to hysteresis and/or amplifier properties like current limitation or slew rate.

3. The current level of knowledge in the field of control engineering is sufficient for precision positioning purposes.

4. By means of *electromagnetic bearing technology* it is possible to integrate several parts of a feedback controlled straight motion system. The guidance can be simple, without special surface machining, while the bearing and the actuator system can coincide, as well as the carriage and the top table (see the block scheme in figure 4.3 on page 80). The sensor system, which is indispensable in electromagnetic bearings, can be chosen as good as is needed for the desired accuracy.

5. From earlier research it was known that the piezo inertial sliding motion (ISM) principle is suitable for high-precision positioning and for multi-degrees of freedom positioning with a relatively long stroke in one or more degrees of freedom. Although, when using the piezo ISM principle, best results can be achieved applying hard bearing surfaces with wear resistant coatings, this sliding motion also can be realized applying only low-cost standard materials like steel and glass for the bearing surfaces. Due to the less optimal conditions for friction and wear, the allowed load will be less. Also the maximum velocity achievable will be lower, in the order of 50 μm/s.

6. A simple sliding bearing using high viscosity grease as a lubricant, is very suitable as a high-stiffness base for a low-velocity straight motion system having nanometer resolution. The experimental high viscosity grease bearing, presented in this thesis, combines high dynamic stiffness (typically 10^9 N / m), virtually zero static stiffness, high damping (typically 10^5 Ns / m), strong spatial averaging (the straightness errors are 10 to 100 times smaller than the form errors of the mechanical components) and smooth motion. It enables open-loop sub-micrometer straightness at low travel velocity ($\ll$ 1 mm / s).

6.1.5 Conclusions with regard to zero stiffness systems

1. Actuators having zero stiffness or near-zero stiffness are suitable for closed-loop controlled positioning systems that require vibration isolation. When applying zero stiffness actuators, the stiffness properties of the closed-loop system are fully determined by the control system.

2. Bearing irregularities can be regarded as external disturbances (vibrations), acting on the actuator system, that depend on the travel velocity. These irregularities do not influence the positioning performance when *vibration isolating actuators* are applied.

3. Actuators having a stiffness much larger or much smaller than zero, can be applied in vibration isolation systems, but this is not an optimal choice in both cases. The requirements to the controller are high, due to the model-based stiffness compensation that is necessary to eliminate the actuator's own stiffness, which hampers proper vibration isolation.

4. Lorentz actuators are suitable for vibration isolation of relatively lightweight objects. They can have a stiffness that is essentially zero. The most important drawback is the heat dissipation when a statical force is required.

5. In this thesis, a new permanent magnetic zero stiffness bearing has been presented. It generates a statical force, while the stiffness is equal to zero in a working point. Around the working point, the stiffness is approximately proportional to the displacement from the working point. Thus, within a certain working range, a force can be generated that is virtually independent on the relative position between the suspended object and the world. No external energy is required and no heat is generated, which is an advantage compared to Lorentz actuators. This type of bearing element is marginally stable, so that closed-loop operation is required.

6. By means of the *permanent magnetic zero stiffness* configuration presented in chapter 5, it becomes possible to integrate vibration isolation properties in a compact contactless bearing element. Besides nearly ideal vibration isolation

this also results in a redistribution of mechanical couplings: the force coupling to the world is maintained whereas the position coupling is removed. This offers interesting possibilities from the control engineering point of view, since a virtual position coupling to any desired reference object can be created by means of an additional control system.

7. Earnshaw's theorem does not conflict with zero stiffness bearings using permanent magnets, because it allows the situation that there is a neutral equilibrium in all degrees of freedom. Only in the case that ferromagnetic material is present within the magnetic field, negative stiffness will occur in one or more directions.

8. It appears to be possible to design a permanent magnetic bearing element according to the principle introduced in this thesis, in such a way that it has a long stroke in one degree of freedom—enabling straight motion—preserving the (zero) stiffness behavior in all degrees of freedom.

6.2 Recommendations

1. Further research should be done in the field of precision positioning using the zero stiffness principle for the integrated bearing and actuator system. The most important application will be the virtual coupling (by means of force control instead of position control) between an object to be positioned and a reference object, while both objects can be isolated from external vibrations. The most demanding subject of research will be the optimization of the (nonlinear) control system, in order to achieve zero stiffness properties in a larger working range.
   A possible industrial application is a so-called wafer scanner for IC lithography, where both the wafer stage and the reticle stage could contain a very accurate and stable positioning system based on zero stiffness actuators. Besides that, the active vibration isolation of the wafer scanner's metrology frame, having a mass larger than 1000 kg, might be improved by means of permanent magnet bearing units.

2. The long-term behavior of NdFeB permanent magnets in a continuous repulsing situation should be investigated.

3. There is a growing industrial interest in high accuracy *linear* and *planar* positioning systems having a high degree of inherent vibration isolation. Therefore, the possibilities to combine the zero-stiffness permanent magnetic bearing (as described in section 5.4) with the planar active magnetic bearing concept (as described in section 3.6.3), should be investigated.

4. The possibilities to exploit the high dynamic stiffness and straightness error averaging of the grease bearing, as proposed in section 4.4, should be investigated further.

5. Much effort has still to be spent in the development of compact and low-cost (IC-size) position sensors, since they are often the limiting factor in feedback controlled positioning systems.

Bibliography

[1] F.R. Aalbers, *High speed precision rotation utilizing the squeeze film bearing technique*, Master's thesis, Delft University of Technology / École Polytechnique Fédérale de Lausanne, July 1997.

[2] M. Abramowitz and I. Stegun, *Handbook of mathematical functions*, Dover Publications Inc., New York, 1965.

[3] F. Auer, *Combined electromagnetic suspension and propulsion for positioning with sub-micrometer accuracy*, Ph.D. thesis, Delft University of Technology, November 1995.

[4] J.W. van Beek and F.K. Koster, *Schut's atlas der geometrische meettechniek [Schut's handbook of geometrical measuring technique]*, Schut's Im– en Exporthandel BV, Groningen, 1990, in Dutch.

[5] R.J. Bernhard, *The state of the art of active-passive noise control*, Noise-Con 94; Proceedings of the 1994 National Conference on Noise Control Engineering. Progress in Noise Control for Industry. Noise Control Found, New York, NY, USA, May 1994, pp. 421–428.

[6] Ph. Blum, *Passive 6 degree of freedom vibration isolation with Lorentz actuators*, Master's thesis, Delft University of Technology / École Polytechnique Fédérale de Lausanne, February 1998.

[7] T. Boltshauser, C.A. Leme, and H. Baltes, *High sensitivity CMOS humidity sensors with on-chip absolute capacitance measurement system*, Sensors and Actuators, B: Chemical **B15** (1993), no. 1–3, 75–80.

[8] M.H.W. Bonse, *Capacitive position transducers*, Ph.D. thesis, Delft University of Technology, December 1995.

[9] W. Braunbek, *Freischwebende Körper im elektrischen und magnetischen Feld*, Zeitschrift für Physik **112** (1939), no. 11–12, 753–763.

[10] J.C. Compter, *Principal methods for coils and magnets*, Philips CFT Report nr. CTB595-98-4001, 1998.

[11] Fan Disc Corporation, *http://www.fandisc.com/mci.htm*, internet web page, 1997.

[12] R. van Dalen, *Scanning tunneling microscopie als verplaatsingsinstrumentatie [STM as a displacement transducer]*, Literature research, Delft University of Technology, 1994, in Dutch.

[13] J. van Drecht, *Relaxatie-oscillator*, Patent application 9101076, 1991.

[14] S. Earnshaw, *On the nature of the molecular forces which regulate the constitution of the luminiferous ether*, Transactions of the Cambridge Philosophical Society **7** (1842), 97–112.

[15] R.W.J. Egmond, *De kalibreerbaarheid van capacitieve sensoren [Calibration aspects of capacitive sensors]*, Master's thesis, Delft University of Technology, December 1996, in Dutch.

[16] J. van Eijk, *On the design of plate spring mechanisms*, Ph.D. thesis, Delft University of Technology, 1985.

[17] W.T. Estler, *Calibration and use of optical straightedges in the metrology of precision machines*, Optical Engineering (1985), no. 24, 372–379.

[18] A. Alabuzhev *et al.*, *Vibration protecting and measuring systems with quasi-zero stiffness*, Applications of vibration, Hemisphere Publishing Corporation, New York, 1989.

[19] D.J. Jendritza *et al.*, *Technischer Einsatz neuer Aktoren*, 2 ed., vol. 484, Expert Verlag, 71272 Renningen-Malmsheim, 1998.

[20] C.J. Evans, R.J. Hocken, and W.T. Estler, *Self–calibration: reversal, redundancy, error separation, and 'absolute testing'*, CIRP Annals **45** (1996), no. 2, 617–634.

[21] S. Futami and A. Furutani, *Nanometer positioning using AC linear motor and rolling ball guide*, Bull. Japan Soc. of Prec. Engg. **24** (1990), no. 3, 184–185.

[22] P.J. Geary, *Magnetic and electric suspensions; a survey of their design, construction, and use*, (SIRA survey of instrument parts no. 6), Chislehurst : British Scientific Instrument Research Association, 1964, SIRA research report R 314.

[23] V. Gopalakrishnan, *Capacitance meters & transistor tester*, Modeling, Measurement & Control A: General Physics, Electronics, Electrical Engineering **53** (1993), no. 3, 37–39.

[24] W.Chr. Heerens, *Application of capacitance techniques in sensor design*, J. Phys. E: Sci. Instrum. (1986), no. 19, 897–906.

[25] I.J.B. van Hees, *Ontwerp van een teruggekoppelde rechtgeleiding gebaseerd op piezo inertial sliding motion [Design of a feedback controlled straight motion system based on piezo inertial sliding motion]*, Master's thesis, Delft University of Technology, February 1997, in Dutch.

[26] K.P. Heiland, *Recent advancements in passive and active vibration control systems*, SPIE **1619** (Vibration control in microelectronics, optics, and metrology) (1991), 22–33.

[27] C. Hoffrogge, R. Mann, and H-J. Rademacher, *Geradheitsprüfung von Linealen mit einem 2-Flächen Verfahren*, Meßtechnik (1972), no. 9, 263–266.

[28] G. de Jong, *Smart capacitive sensors; physical, geometrical and electronic aspects*, Ph.D. thesis, Delft University of Technology, April 1994.

[29] A.H.J. Jongepier, *Systemen voor ultraprecisie rechtgeleidingen in vacuüm [Ultra precision straight motion systems in vacuum]*, Literature research, Delft University of Technology, September 1997, in Dutch.

[30] H. Kawakatsu, H. Bleuler, T. Saito, T. Yahiro, M. Watanabe, and R. Siegwart, *Laser sensing and new stator designs for a linear slider*, Proceedings of the Fourth International Symposium on Magnetic Bearings, ETH Zurich, August 1994, pp. 189–192.

[31] K. Kondo and K. Watanabe, *Switched-capacitor capacitance measurement circuit with the vernier scale*, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences **E76-A** (1993), no. 7, 1139–1142.

[32] D. Landheer and A.W.J. de Gee, *Tribotechniek A; inleiding dynamische contactverschijnselen [tribology A; introduction to dynamic contact phenomena]*, TU Delft, August 1991, Lecture notes, in Dutch.

[33] M.H.A. Leenders, *Position independent force using reluctance actuators*, Master's thesis, TU Delft, September 1997.

[34] A. Leissa and A. Saad, *Large amplitude vibrations of strings*, Journal of Applied Mechanics **61** (1994), 296–301.

[35] S. Li, C. Liang, and D. Fan, *The optimum EST — A new method for on-line measurement of the straightness of precision machine tools and machined workpiece*, Proceedings of the 7th International Precision Engineering Seminar, May 1993, pp. 451–459.

[36] X. Li, *Low-cost smart capacitive sensors for position and speed measurement*, Ph.D. thesis, Delft University of Technology, May 1997.

[37] K. Lindsey and S.T. Smith, *Precision motion slideways*, International Patent, October 1998.

[38] K. Lindsey, S.T. Smith, and C.J. Robbie, *Sub-nanometre surface texture and profile measurement with* NANOSURF *2*, Annals of the CIRP **37** (1988), no. 1, 519–521.

[39] P.B. Lock, *Design of a permanent magnetic gravity compensator*, Master's thesis, TU Delft, 2001, to be published.

[40] Magnetic Component Engineering, Inc., `http://www.magneticcomponent.com/AboutNdFeB.htm`, Internet document, 1999.

[41] B.K. Marlow, D.C. Greager, R. Kemp, and M.B. Moore, *Highly sensitive capacitance measurement for sensors*, Electronics Letters **29** (1993), no. 21, 1844–1845.

[42] K. Martin, *A voltage-controlled switched-capacitor relaxation oscillator*, IEEE Journal of Solid-State Circuits **16** (1981), no. 4, 412–413.

[43] A. Molenaar, *Aangedreven magneetlager*, Dutch patent application nr. 1005344, 1997.

[44] A. Molenaar, *A novel planar magnetic bearing and motor configuration applied in a positioning stage*, Ph.D. thesis, Delft University of Technology, October 1999.

[45] R. Müller, W. Andreasch, H. Bleuler, H. Kawakatsu, B. Rupp, and M. Schwab, *High precision linear magnetic bearing with laser sensing and reluctance motor*, Proceedings of the Third International Conference on Motion and Vibration Control, Chiba, September 1–6 1996, pp. 100–104.

[46] K. Nagaya, M. Atsumi, and M. Endoh, *Simplified vibration isolation control using disturbance cancellation and velocity feedback controls for a magnetic levitation table*, Internation journal of applied electromagnetics in materials (1993), no. 4, 115–122.

[47] K. Nagaya and M. Ishikawa, *A noncontact permanent magnet levitation table with electromagnetic control and its vibration isolation method using direct disturbance cancellation combining optimal regulators*, IEEE Transactions on magnetics **31** (1995), no. 1, 885–896.

[48] W. Niesing, *Ontwerp nulstijfheid-lager [Design of a zero stiffness bearing]*, Master's thesis, Delft University of Technology, March 1998, in Dutch.

[49] G.J.P. Nijsse, *Verbetering van een rechtgeleiding door terugkoppeling van positiefouten in vijf vrijheidsgraden [Improvement of a straight motion system by 5DoF position error feedback]*, Master's thesis, Delft University of Technology, June 1995, in Dutch.

[50] G.J.P. Nijsse, C. Hengst, and J.W. Spronck, *A versatile, simple and low-cost capacitive sensor for straightness measurement*, Proceedings of the 9th International Precision Engineering Seminar / 4th International Conference on Ultraprecision in Manufacturing Engineering, Braunschweig, May 1997, pp. 78–81.

[51] G.J.P. Nijsse and J.W. Spronck, *Magnetic bearing element*, Dutch patent application nr. NL1007127, International patent application nr. PCT / NL98 / 00558, September 1997.

[52] G.J.P. Nijsse, J.W. Spronck, M.H.A. Leenders, and H.F. van Beek, *A new zero-stiffness magnetic bearing for vibration isolation*, Proceedings of the 12th annual meeting of the American Society for Precision Engineering, vol. 16, October 1997, pp. 503–506.

[53] C.H.T. Pan and P.H. Broussard jr., *Squeeze-film lubrication*, Gas bearing symposium, University of Southampton, April 1967.

[54] R.J. Parker, *Advances in permanent magnetism*, Wiley, New York, 1990.

[55] J. de Pee, *Ontwikkeling van een gecombineerde capacitieve positie–snelheidssensor [development of a combined capacitive position–velocity sensor]*, Master's thesis, Delft University of Technology, March 1998, in Dutch.

[56] D.L. Platus, *Negative-stiffness-mechanism vibration isolation systems*, SPIE **1619** (Vibration control in microelectronics, optics, and metrology) (1991), 44–54.

[57] A. Portegies, *Optimaal ontwerp van Lorentz actuatoren [Optimal design of Lorentz actuators]*, Literature research, Delft University of Technology, March 1998, in Dutch.

[58] A. Post, *Ontwerp van een rechtgeleiding met een zeer lage eigenfrequentie [Design of a straight motion system having a very low eigenfrequency]*, Mikroniek (1996), no. 1, 15–18, in Dutch.

[59] R. Reed, *A glass reference surface for quality control measurements*, International Journal of Mechanical Sciences **8** (1966), 703–715.

[60] P.C.J.N. Rosielle and E.A.G. Reker, *Constructieprincipes [Construction principles]*, vol. 1, Eindhoven University of Technology, section Precision Engineering, March 1997, Lecture notes, in Dutch.

[61] M.J.L. Sanders, *A new linear active magnetic suspension configuration; a permanent magnet biased non-coplanar circuit*, Master's thesis, Delft University of Technology, 1997.

[62] Ph. Schmiechen and A. Slocum, *Analysis of kinematic systems: a generalized approach*, Precision Engineering **19** (1996), no. 1, 11–18.

[63] A.H. Slocum, *Precision machine design*, Prentice-Hall, Inc., New Jersey 07632, 1992.

[64] A.H. Slocum, N.R. Kane, P.A. Scagnetti, and C. Brunner, *Design of self-compensated, water-hydrostatic bearings*, Precision Engineering **17** (1995), no. 3, 173–185.

[65] S.T. Smith, K. Lindsey, and C.J. Robbie, *Sub-nanometre surface texture and profile measurement with nanosurf 2*, Annals of the CIRPAnnals of the CIRP **37** (1988), no. 1, 519–522.

[66] J.W. Spronck, M.H.W. Bonse, and G.J.P. Nijsse, *A 5 degree-of-freedom capacitive position sensor for straightness measurements*, Proceedings of the American Society for Precision Engineering, vol. 12, ASPE, 1995, pp. 35–38.

[67] J.W. Spronck, P.F.A.M. van Putten, and G.J.P. Nijsse, *Design of an ultra low speed straight motion system based on a high viscosity fluid bearing*, Proceedings of the American Society for Precision Engineering, vol. 16, ASPE, 1997, pp. 373–376.

[68] H. Tanaka, K. Tozawa, H. Sato, M. O-hori, and H. Sekiguchi, *Application of a new straightness measurement method to large machine tools*, Annals of the CIRP (1981), no. 30, 455–459.

[69] J. Tapson and J.R. Greene, *Improved capacitance measurement by means of resonance locking*, Meas. Sci. Technol. **5** (1994), 20–26.

[70] F.N. Toth, *A design methodology for low-cost, high-performance capacitive sensors*, Ph.D. thesis, Delft University of Technology, January 1997.

[71] K. Tozawa, H. Sato, and M. O-hori, *A new method for the measurement of the straightness of machine tools and that of machine works*, Jounal of mechanical design **104** (1982), 587–592.

[72] J. Uchikoshi, S. Shimada, N. Ikawa, and A. Komura, *Straightness measurement using laser beam straight datum*, Proceedings of the SPIE **2576** (1995), 315–322.

[73] B. de Veer, *$XY\phi$ manipulator met sub-micrometer nauwkeurigheid gebaseerd op elektromagnetische reluctantie actuatoren [Sub-μm $XY\Phi$ manipulator based on electromagnetic reluctance actuators]*, Master's thesis, Delft University of Technology, 1994, in Dutch.

[74] M.S. Virdee, *Non-contacting straightness measurement to nanometre accuracy*, International Journal of Machine Tools & Manufacture **35** (1994), no. 2, 157–164.

[75] H. Wan, Z. Pan, and J. Liang, *On-line measurement and compensation straightness of machine tools and machined work*, Advances in Instrumentation **42** (1987), no. 2, 635–642.

[76] D.F. Wilkes, *Rolamite: a new mechanical design concept*, Springfield, Clearinghouse, 1967.

[77] H. van der Wulp, *Piezo-driven stages for nanopositioning with extreme stability*, Ph.D. thesis, Delft University of Technology, December 1997.

[78] J. Yamaguchi, *Measurement and control of straight motion accuracy using liquid surface for absolute reference*, Proceedings of the 7$^{\text{th}}$ International Precision Engineering Seminar, May 1993, pp. 467–478.

[79] J.-P. Yonnet, *Passive magnetic bearings*, Proceedings of the 3$^{\text{rd}}$ international workshop on rare earth–cobalt permanent magnets and their applications, San Diego, CA (1978), 241–251.

[80] E.H. Zaaijer, *Nanosensoren [Nanosensors]*, Literature and market research, TU Delft, August 1996, in Dutch.

[81] F. Zhu, *Development of capacitive position transducers through a mechatronic approach*, Ph.D. thesis, Delft University of Technology, April 1992.

Appendix A

Earnshaw's theorem and related matter

A.1 Permanent magnetic fields

For the field, generated by permanent magnets, the Laplace equation holds for the magnetostatic energy W:

$$\nabla^2 W = 0 \tag{A.1}$$

which also can be written as

$$\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2} + \frac{\partial^2 W}{\partial z^2} = 0 \tag{A.2}$$

A magnetic object in such a field can be in a state of *equilibrium* only if the net force acting on it is zero, i.e. if the resultant forces in the x-, y- and z-direction are equal to zero. This can be written in terms of energy as follows:

$$\frac{\partial W}{\partial x} = 0 \wedge \frac{\partial W}{\partial y} = 0 \wedge \frac{\partial W}{\partial z} = 0 \tag{A.3}$$

For a *stable* equilibrium there is an additional requirement:

$$\frac{\partial^2 W}{\partial x^2} > 0 \wedge \frac{\partial^2 W}{\partial y^2} > 0 \wedge \frac{\partial^2 W}{\partial z^2} > 0 \tag{A.4}$$

which clearly is in contradiction with equation A.1. This result is known as Earnshaw's theorem, who first published this in 1842 [14]. In words, this theorem states:

> *For a pole placed in a static field of force it is impossible to have a position of stable equilibrium, when an inverse square law relates force and distance.*

In 1939, Braunbek [9] worked out the practical consequences of Earnshaw's theorem for magnetic bearings.

Two important notes have to be made:

1. The theorem also prohibits the situation of an *unstable* equilibrium in all directions, unless soft magnetic material (i.e. material having $\mu_r > 1$) is present (see section A.2).

2. The theorem *does not exclude* the situation that there is a *neutral* equilibrium in all directions. This aspect will appear to be very important for zero stiffness applications.

In terms of F_x, F_y and F_z, the Laplace equation from A.1 is equivalent to

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 0 \tag{A.5}$$

When stiffness is defined as

$$k_x = -\frac{\partial F_x}{\partial x} \tag{A.6}$$

then equation A.5 can be rewritten as

$$k_x + k_y + k_z = 0 \tag{A.7}$$

In the case of a rotational symmetric configuration around the z-axis, the radial stiffness k_{rad} is

$$k_{rad} = k_x = k_y \tag{A.8}$$

from which follows, in combination with equation A.7, that

$$k_{rad} = -\frac{1}{2}k_z \tag{A.9}$$

A.2 Permanent magnetic fields and ferromagnetic material

In the presence of soft magnetic or ferromagnetic material, e.g. iron, having a relative magnetic permeability $\mu_r > 1$, the behavior described above is substantially different [79]. The induced magnetization in the soft magnetic material results in a decreased magnetic energy, because that material ‘absorbs’ a part of the field energy to align its magnetization parallel to the applied field. Now, the stiffnesses from equation A.4 can be rewritten as

$$k_i' = \frac{\partial^2 W'}{\partial i^2} \tag{A.10}$$

in which $i \in \{x, y, z\}$ and $W' < W$ due to the energy loss, so that

$$k_i' < k_i \tag{A.11}$$

and equation A.7 becomes

$$k_x + k_y + k_z < 0 \tag{A.12}$$

In the special case that there is rotational symmetry around the z-axis, equation A.9 becomes

$$k_{rad} < -\frac{1}{2}k_z \tag{A.13}$$

In an analogous way it can be shown that in the case of $\mu_r < 1$ the equations A.7 and A.12 become

$$k_x + k_y + k_z > 0 \tag{A.14}$$

which means that in that case a stable equilibrium is possible. Materials having $\mu_r < 1$ are diamagnetic materials like water, most organic materials, graphite etc., where μ_r is slightly smaller than 1, or superconducting materials, which can be considered as having $\mu_r \approx 0$.

These considerations lead to the conclusion that for the realization of a bearing element having zero stiffness in all degrees of freedom, only hard magnetic material ($\mu_r = 1$) may be used. In practice, most permanent magnetic materials have a μ_r slightly larger than 1, so that the sum of the translational stiffnesses will be slightly less than zero.

When zero stiffness is required in one direction only, soft magnetic materials ($\mu_r > 1$) may be used—for example, to increase the bearing force, or to shield the bearing element or the environment against disturbing magnetic fields—at the expense of stability in the other directions.

A.3 Angular stiffness

In all the above discussion the angular stiffness is left out of consideration. A similar approach can be made for rotations as well. However, Earnshaw and Braunbek [14, 9] finished their calculations after proving that even the first requirement for mechanical stability could not be satisfied. They did not pay attention to the moment equations. Nevertheless, since we want to realize a neutral equilibrium in *six* degrees of freedom, the moment equilibrium is of equal importance as the force equilibrium. Therefore, some attention will be paid here to the angular stiffness. This will be done in a qualitative way, without paying much attention to the theoretical calculations.

When considering the rotational stiffnesses, a clear example is a small permanent magnet, placed in a homogeneous magnetic field. It is obvious that there are no *forces* working upon the magnet, since the magnetic energy W remains constant for any Δx, Δy and Δz and thus $F_x = \frac{\partial W}{\partial x} = 0$ (which holds analogously for y and z). But, although there are no net forces on the magnet, there may occur a net *moment*. This moment or torque T depends on the direction of the magnetization $\vec{m}$ relative to the magnetic field $\vec{B}$ and the magnitude of the field. When $\vec{m}$ and $\vec{B}$ both are in the x, y plane, the moment can be expressed as

$$T_z = mB \sin\varphi \tag{A.15}$$

In general vector notation it is written as follows:

$$\vec{T} = \vec{m} \times \vec{B} \tag{A.16}$$

which is equal to zero if the direction of $\vec{m}$ is parallel to $\vec{B}$. It can be seen easily that there will be a *positive rotational stiffness*—a torque T is generated that works opposite to the angle φ—when the angle φ between $\vec{m}$ and $\vec{B}$ is smaller than 90°, and a *negative rotational stiffness* when that angle is larger than 90°.

From this thought experiment it can be concluded that the rotational stiffness is independent on the translational stiffness. This is only partly true, since this example considers the situation that the translational stiffnesses all are equal to zero, so that indeed there is no relation possible. But it is important to realize that, when all translational stiffnesses are equal to zero, the rotational stiffnesses are *not* necessarily equal to zero.

Now let us make the example a little more difficult by assuming that the field is *not homogeneous*. Consider a permanent magnet, fixed to the world, having a magnetization vector in the positive z-direction. Now another permanent magnet is brought in this magnetic field from below, having the same magnetization vector, such that both magnetization vectors are aligned axially. By representing the second magnet by a current loop, it is easily seen that there will be an attracting force between the two magnets. By rotating the second magnet, as is shown schematically in figure A.1, the magnitude and the direction of the force vectors changes, so that a moment is generated with a sign opposite to the rotation angle. Thus, in this case a positive rotational stiffness is present. It can be made plausible that when the radial stiffness is positive—as is the case in the figure—also the rotational stiffness will be positive.

When in figure A.1 the lower magnet is turned upside down, it will be repelled by the fixed magnet in the axial direction. This repelling force has a positive stiffness.

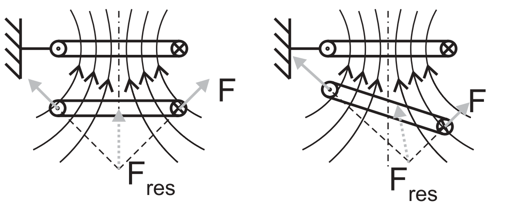


Figure A.1: Schematic visualization of the force acting on a permanent magnet (represented by a current loop), and the moment that is generated when the lower magnet is rotated. It appears that there is a positive rotational stiffness in this case, because the sign of the moment and the angle are opposite.

The radial forces will have negative stiffness, as has been explained earlier. In this case, the rotational stiffness will be negative as well.

In figure A.2 it is clarified that in the case that the translational stiffnesses are equal to zero (as is the case in the permanent magnetic zero stiffness actuator), also the rotational stiffness will be equal to zero. In this figure, it is assumed that the magnetic field only has a radial component in the region where the middle (suspended) current loop is situated. It appears that when the angle is equal to zero, also all moments are equal to zero. When the middle magnet is subject to a small rotation, still there is no moment. This is because all small force elements, that are sketched in figure A.2, occur in pairs that are symmetrical with respect to the circle's center. Instead of a moment, a (parasitic) horizontal component of the force appears.

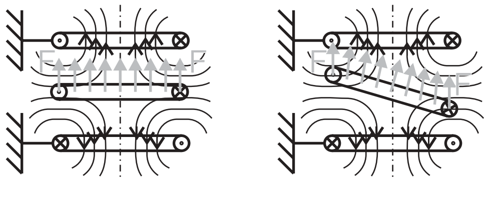


Figure A.2: Schematic visualization of the force acting on a permanent magnet (represented by a current loop), and the moment that is generated when the middle magnet is rotated, in the geometry that is used to illustrate the zero stiffness principle. It is assumed here that the magnetic field is directed purely radial. It appears that when the axial and radial stiffnesses both are equal to zero, also the rotational stiffness will be equal to zero. However, a rotation of the middle magnet will cause a horizontal force.

Summarizing the previous paragraphs, some remarks can be made with regard to combinations of permanent magnets aiming at near–zero stiffness in all six degrees of freedom:

- Only materials having $\mu_r = 1$ may be applied

- It is preferable to make use of (horizontal, rotational) symmetry as much as possible
- When the sum of the translational stiffnesses $k_x + k_y + k_z = 0$, this does *not* necessarily imply that the sum of the rotational stiffnesses $k_{\varphi_x} + k_{\varphi_y} + k_{\varphi_z} = 0$ as well
- Rotations can introduce parasitic translational forces

A.4 Example of analytical calculation of magnetic forces

In this section we will pay some attention to the derivation of approximations for the forces and stiffnesses in permanent magnet configurations. To reduce the analytical problem to the core, the magnetic configurations are simplified by assuming that each magnet is represented by a current loop. For the magnetic field and the mutual forces between these current loops, analytical equations can be formulated. By trying to find an exact solution for the force equations, one encounters complicated elliptic integrals [2] that are difficult to handle. Fortunately, solutions have been found by means of *numerical* analysis [10].

Two Matlab scripts to calculate (numerically) the K- and E-integral, necessary for the calculation of the magnetic forces between cylindrical magnets, are printed in figure A.3. In analytical form they are written as

$$\begin{aligned} K(k) &= \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{1 - k^2 \cdot \sin(\nu)^2}} d\nu \\ E(k) &= \int_0^{\frac{\pi}{2}} \sqrt{1 - k^2 \cdot \sin(\nu)^2} d\nu \end{aligned} \qquad \text{(A.17)}$$

Using these two integrals, the resulting magnetic field of any combination of current loops can be calculated. Subsequently, the force on one or more specific current loops can be determined, for example in the equations below, which give the force between two circular conductors with radii r_1 and r_2, carrying a current I_1 and I_2, and having an axial distance d:

$$k = 2 \cdot \sqrt{\frac{r_1 r_2}{d^2 + (r_1 + r_2)^2}}$$

and

$$F_{axial} = \frac{\mu_0 I_1 I_2 d}{\sqrt{d^2 + (r_1 + r_2)^2}} \left(\frac{r_1^2 + r_2^2 + d^2}{(r_1 - r_2)^2 + d^2} \cdot E(k) - K(k) \right) \qquad \text{(A.18)}$$

In order to predict the forces and stiffnesses of magnetic configurations suitable for vibration isolation (see figure 5.19), a Matlab tool has been written [39], as a part of

```
function [K] = kint(a)

a = sqrt(1 - a^2);
if a > 1-1e-10,
   a = 1-1e-10;
end;

if a < 1e-13,
   K = 2 * ln(2) - log(a);
else
   h = 1;
   while abs(h - a) > h * 1e-12,
      m = a + h;
      a = sqrt(h * a);
      h = m / 2;
   end;
   K = pi / m;
end;
```

```
function [E] = eint(a)

a = sqrt(1 - a^2);
if a < 1e-13,
   E = 1;
else
   m = 1 + a;
   a1= 1 + a^2;
   c = a1;
   m0= 1;
   b = 2 * a * m;
   while abs(m0 - a) > m0 * 1e-13,
      a1= b / m + a1;
      a = sqrt(a * m0) * 2;
      b = (c * a + b) * 2;
      c = a1;
      m0= m;
      m = a + m;
   end;
   E = (pi / 2) * (a1 / m);
end;
```

Figure A.3: Two Matlab scripts to calculate (numerically) the K- and E-integral, necessary for the calculation of the magnetic forces between cylindrical magnets.

a master's thesis, which is to be published in 2001. Using this tool, the influence of axial and radial deviations on the forces and moments can be approximated. This is done by replacing each magnet by a number (e.g. 10 or 20, depending on their height and the desired accuracy) current loops, resulting in sets of equations like A.18. Also the influence of a rotation around the lateral axis can be calculated, as well as the axial and radial stiffness. An impression of the user interface of the tool (under construction) is given in figure A.4.

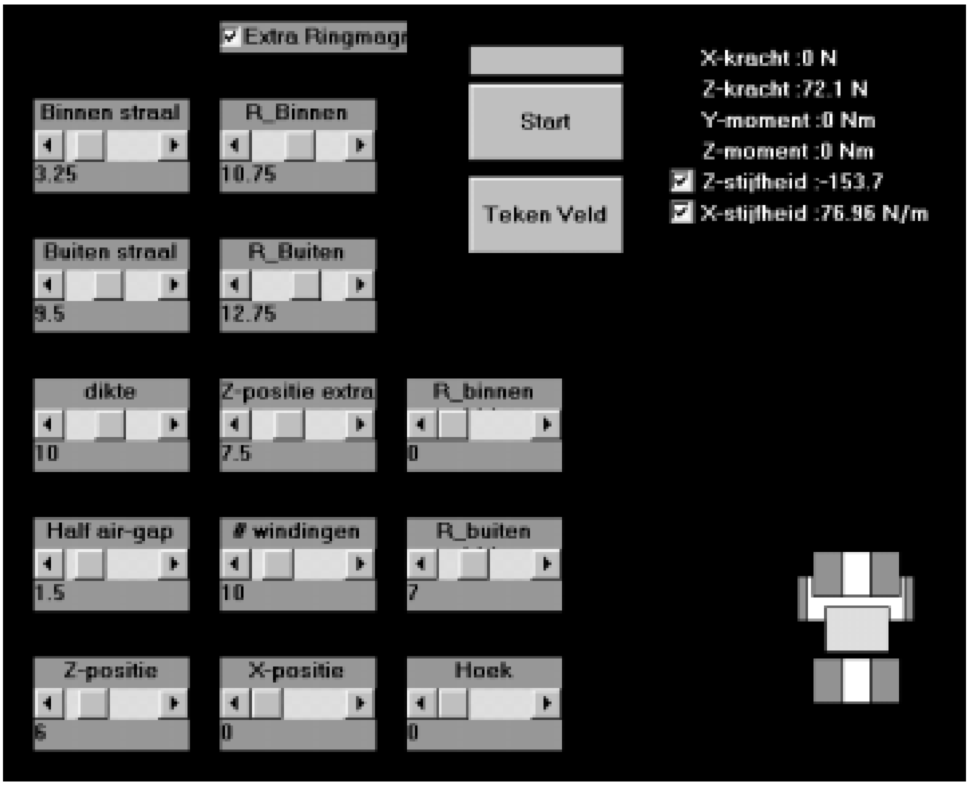


Figure A.4: This figure shows the user interface of a Matlab tool (under construction [39]) that is used to calculate the fields, forces and stiffnesses of permanent magnetic configurations like the one sketched in the bottom right corner. There, the middle (lighter colored) magnet is the suspended magnet, the other ones are fixed to the world.

By separating the magnetic field (see figure A.5), in which the supported magnet floats, into a radial and an axial part (see the figures A.6 and A.7), a tool is created to optimize the geometry of the magnets. When the outer contour of the supported magnet is located in a region of the field where the axial component is either equal to zero or exactly symmetrical relative to the supported magnet, the axial stiffness will be zero. Thus, by drawing lines of equal magnitude for the axial as well as the radial field component, an optimum geometry (outer contour) can be drawn for the supporting magnet. Currently, the research in this field focuses on the practical application and the control engineering aspects of these magnetic configurations.

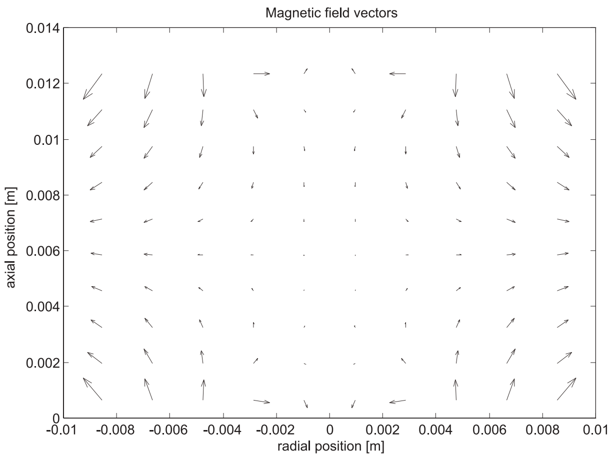


Figure A.5: Plot of vectors representing the magnetic field between the three fixed magnets as shown in figure A.4. The vertical axis only shows the region between the upper and lower magnet. [39]

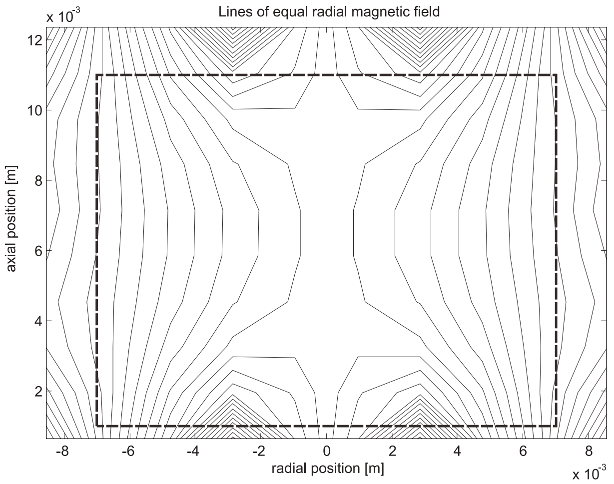


Figure A.6: Plot of lines having equal radial components of the magnetic field between the three fixed magnets as shown in figure A.4. The geometry of the supported magnet is shown with a dashed line. The radial component of the field generates the axial component of the force. When the lines in this figure are straight and vertical at the outer cylindrical contour of the supported magnet, this means that the axial force does not change with a radial or axial displacement. [39]

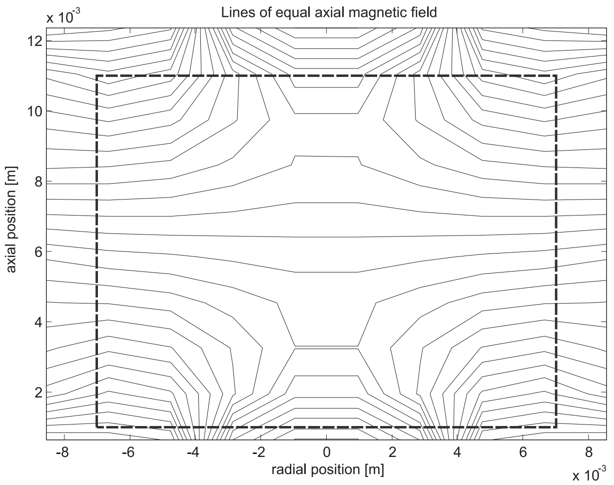


Figure A.7: Plot of lines having equal axial components of the magnetic field between the three fixed magnets as shown in figure A.4. The geometry of the supported magnet is shown with a dashed line. The axial component of the field generates the radial component of the force. When the lines in this figure are straight and horizontal at the outer cylindrical contour of the supported magnet, this means that the axial force does not change with a radial or axial displacement. [39]

Abstract

This thesis deals with straight motion systems. A modular approach has been applied in order to find ways to improve the performance. The main performance parameters that are considered are position accuracy, repeatability and, to a lesser extent, cost.

Straight motion systems

Because of the increasing requirements to positioning systems, concerning accuracy, repeatability and velocity, it becomes more and more difficult to meet these requirements when using merely mechanical means. The errors in position occurring as a result of e.g. external forces, wear or thermal expansion, are the reason that open-loop straight motion systems can only operate in the sub-micrometer range when very strict conditions are applied. Only if the environmental parameters, together with their influence, are known accurately or limited strictly, their effect can be kept small.

In this thesis, some representative straight motion systems are investigated, distinguished by the type of bearing applied. The main properties, advantages and drawbacks of these straight motion systems are summarized, and the suitability for the operation in the sub-micrometer range is considered.

When the required accuracy becomes smaller than the manufacturing tolerances achievable, active compensation of errors becomes indispensable. On the other hand, by the application of active error compensation, the requirements to the manufacturing accuracy can decrease.

Measurement of straightness

A wide variety of sensors is available for the measurement of straightness errors. Measurement of straightness, in five degrees of freedom, is done in practice mostly by measuring five relative distances, after which a coordinate transform converts these distances to the desired straightness parameters. This measurement of straightness can be done either ‘off-line’, in order to calibrate a straight motion system, or ‘on-line’, for closed-loop control purposes.

An analysis has been done on the available types of sensors to see whether they are suitable for real-time straightness measurement in closed-loop systems.

The highest resolution, in the nanometer or even the sub-nanometer range, can be achieved using capacitive sensors or laser interferometers. The lowest cost can be realized using IC photonic sensors, mainly due to the high degree of integration, so that a minimum of additional electronics is required. It is expected that sensors based on other measurement principles (e.g. inductive or capacitive) will be available at comparable cost, as soon as they can be manufactured in sufficient quantities, using IC-technology.

In this thesis, a number of newly developed sensor systems is presented: an integrated capacitive straightness sensor, a combined capacitive position–velocity sensor and an optical straightness sensor utilizing only one laser beam. Prototypes of these systems have been built, tested and used.

In the field of position sensors two interesting trends are visible. Firstly, the increasing *integration* of the necessary electronics into the sensing element housing, resulting in a substantial decrease of wiring, shielding and separate conditioning electronics, and mostly in better performance as well. Secondly, the increasing *miniaturization* of sensor elements and the accompanying electronics in ASICs[1], so that a sensor system, consisting of a sensing element, conditioning electronics and wiring, can be replaced by only one IC from which the measurement signal directly is available in the format required by the control system.

Besides the straightness sensors, also the straightness *references* are regarded. They determine the eventual accuracy of the measurement. Depending on the measurement needs, the required *form accuracy* and the *stability* of the straightness reference may vary.

Feedback controlled systems

When measuring straightness errors with a high-bandwidth straightness sensor, it becomes possible to compensate in real-time for the straightness errors measured, by means of an additional table on top of the motion system under test, which is moveable by five independent actuators, for five degrees of freedom. This additional table is positioned in such a way, that it follows the straightness reference. In other words, the deviations with respect to the straightness reference are reduced as much as possible.

By means of this feedback of straightness errors, the problems and bottle-necks shift from the mechanical domain towards the electronics and control engineering domain, where they are better solvable. The eventual performance of a feedback controlled system is not limited any more by fabrication tolerances, bearing properties, thermal influences or external loads, but is mainly limited by the resolution of the sensor applied.

[1]ASIC: Application Specific Integrated Circuit

The capacitive straightness sensor, developed within the Laboratory for Micro Engineering, appears to work well in a quasi-statical test set-up. However, the measurement time is too long to serve also in dynamical operated systems. This problem is solvable; new versions of the sensor electronics, having measurement times substantially below 1 ms, are under development.

In the quasi-statically operated experimental set-up of a straight motion system using the capacitive straightness sensor mentioned before, a position resolution below 0.5 μm has been achieved for the translations and a resolution below 20 μrad for the rotations. This is only slightly above the sensor's resolution.

To investigate the dynamical performance of the feedback controlled straight motion system, photonic sensors have been used. It appears that the eventual performance is restricted mainly by the sensor resolution, but also by the internal dynamical properties of the mechanical system, the control bandwidth and the frequency spectrum of the external disturbances. Besides that, the hysteresis of the stacked piezo actuators, applied in the experimental set-up, appears to be an important drawback.

One of the most important disadvantages of piezo actuators, their limited stroke (in the sub-micrometer or micrometer range), can be overcome by applying an other driving method: when a sawtooth-shaped motion profile is imposed upon a shear mode piezo element, an object can be propulsed by means of the friction force between the piezo element and the object. Because of the inertia of the moving object, the influence of the fast retracting movement of the piezo element is small. This driving method is known as the Inertial Sliding Motion (ISM) principle. Two extensions to the existing knowledge have been investigated on the basis of a newly developed six-DoF manipulator, which has a long stroke in several degrees of freedom. Firstly, it is researched whether it is possible to apply standard materials like steel, aluminum and glass as contact surfaces, from cost reduction point of view. This appears to be possible, although the design margins are smaller, in particular those concerning preload, and the velocity achievable is relatively low, in the order of 50 μm / s. Secondly, the possibilities are investigated for reducing the large number of high voltage amplifiers (the tested six-DoF manipulator contains 15 piezo elements in 12 groups). It appears to be possible to use only two amplifiers, one for each of the two directions of motion, by distributing their output signal by means of high voltage switches.

In the case that the straightness errors of a linear guidance are smaller than the stroke of the (piezo) actuators, and this linear guidance satisfies a number of requirements concerning maximum straightness error, smoothness and stiffness, then it is sensible to build a feedback controlled system, in which the guidance can be regarded as the coarse stage, and a (piezo) manipulator on top of it functions as a fine stage. A possible guidance for such a system, utilizing high-viscosity grease as a lubricant, has been investigated. It appears that the grease lubricated sliding bearing has almost ideal properties for the application as a coarse stage at (very) low velocities, in the order of 100 μm / s. The dynamical stiffness is extremely high, the statical stiffness is low, the spatial averaging results in a high smoothness of the motion, and the maximum straightness errors are in the sub-micrometer range. By adding (for example) a fast monolithic piezo manipulator as a fine stage to this coarse stage, an accuracy in the nanometer range over a relatively long stroke, in the order of 100 mm, can be achieved.

Zero stiffness actuators

When aiming at a position accuracy in the sub-micrometer range, also vibration problems become important. Not only the external vibrations of the world, but also the vibrations that are generated by e.g. the bearing imperfections may cause straightness errors. Therefore, a completely different approach has been applied. Instead of aiming for the highest possible open-loop stiffness of the positioning system relative to the ground, an actuator stiffness equal to zero is proposed. A zero stiffness actuator can exert a force on the suspended object, but the position of the object relative to the actuator does not influence that force. If done perfectly, this will result in a perfect vibration isolation.

First, the vibration isolating properties of Lorentz actuators have been studied. Lorentz actuators are the most well-known actuators of which the force does not depend on the relative position. The test set-up consists of a table, supported in six degrees of freedom by means of Lorentz actuators. It appears that indeed a good vibration isolation (40 dB) can be realized, utilizing Lorentz actuators. A shortcoming in the experimental set-up is that the stiffness is not exactly equal to zero, so that in the frequency range roughly between 5 . . . 50 Hz only 25 dB vibration reduction has been achieved. By means of geometrical fine-tuning this effect is expected to be substantially reduced.

After that, research has been carried out on the possibility to overcome one of the most important disadvantages of Lorentz actuators, namely the power consumption and accompanying heat generation, when a static force has to be generated. A *permanent magnetic bearing element* has been designed and built, which is able to generate a static bearing force that is without mechanical contact, without stiffness and without power consumption.

This permanent magnetic bearing element is marginally stable. This means that at a disturbance in one direction a positive, stabilizing, stiffness occurs; at a disturbance in the other direction a negative, destabilizing stiffness occurs. Therefore, a feedback control is necessary. It appears that, using only position signals in the feedback control loop, no satisfying operation is possible. A feedback of the acceleration, combined with a position feedback and a velocity feedback, gives better results.

The permanent magnetic zero stiffness bearing principle has been patented in the Netherlands; an international patent application is pending.

The zero stiffness bearing results in a decoupling of the position, while preserving the force coupling to the world. This means that the position of such an object

(e.g. a specimen holder), supported by a zero stiffness bearing, may be coupled to the position of another object (e.g. a lens on which no external forces are allowed) by means of a control loop, resulting in a virtual stiffness relative to each other. The forces, needed for positioning the moving object, cause a reaction force on the world, while the vibrations (position errors) of the world do not influence the positioning accuracy of that object. This concept offers possibilities in the design of precision positioning systems, where otherwise vibrations from the outer world would have a too large disturbing influence.

An experimental zero stiffness linear motion system is the subject of ongoing research. The central part of this linear motion system is an extended version of the permanent magnetic zero stiffness bearing element, which now has an unlimited motion range in one degree of freedom. The propulsion force (in the x-direction) and the control forces in the other five degrees of freedom are generated by Lorentz actuators. With this system, a linear motion can be realized that is essentially decoupled from the position deviations of the world.

Samenvatting

Dit proefschrift heeft rechtgeleidingen als onderwerp. Een modulaire aanpak is toegepast om manieren te vinden ter verbetering van de prestaties. De belangrijkste parameters die zijn beschouwd in dit onderzoek zijn positienauwkeurigheid, reproduceerbaarheid en, in mindere mate, kosten.

Rechtgeleidingen

Door de steeds hoger wordende eisen aan positioneersystemen op het gebied van nauwkeurigheid, herhaalbaarheid en snelheid, wordt het steeds moeilijker daaraan te voldoen met behulp van zuiver mechanische middelen. De optredende positiefouten als gevolg van bijvoorbeeld externe krachten, slijtage of thermische uitzetting maken dat rechtgeleidingen zonder actieve foutcorrectie slechts onder zeer strenge voorwaarden in het sub-micrometergebied kunnen werken. Alleen als de omgevingsparameters met hun invloeden nauwkeurig bekend of begrensd zijn kan hun effect worden beperkt.

In dit proefschrift wordt een aantal representatieve rechtgeleidingen beschouwd, gegroepeerd naar de toegepaste soort lagering. Daarbij wordt onderzocht wat de belangrijkste eigenschappen, voordelen en nadelen zijn en in hoeverre de toegepaste lagering geschikt is om in het sub-micrometergebied te werken.

Wanneer de vereiste nauwkeurigheid kleiner wordt dan de haalbare fabricagetoleranties, blijkt actieve foutcompensatie onvermijdelijk. Anderzijds kunnen door het gebruik van actieve foutcompensatie de eisen aan fabricagenauwkeurigheid lager worden.

Rechtheidsmeting

Voor het meten van rechtheidsfouten is een grote verscheidenheid aan sensoren beschikbaar. Het meten van rechtheid, uitgedrukt in vijf vrijheidsgraden, komt in de praktijk meestal neer op het meten van vijf afstanden, waarna met een coördinatentransformatie deze afstanden worden omgezet in de gewenste rechtheidsparameters. Deze rechtheidsmeting kan ‘off-line’ worden uitgevoerd, voor het kalibreren van rechtgeleidingen, of ‘on-line’, in closed-loop regelsystemen.

In dit proefschrift wordt een overzicht gegeven van de beschikbare soorten sensoren, waarbij wordt aangegeven of ze geschikt zijn voor *real-time* rechtheidsmeting in actieve systemen.

De hoogste resolutie, in het nanometer- of zelfs het sub-nanometergebied, kan worden behaald met behulp van capacitieve sensoren of laserinterferometers. De laagste kostprijs, enkele guldens, kan worden gerealiseerd met IC photonic sensoren, wat vooral wordt veroorzaakt door de vergaande integratie zodat nauwelijks additionele elektronica benodigd is. De verwachting is dat sensoren gebaseerd op andere principes (bv. inductief of capacitief) tegen vergelijkbare kosten beschikbaar komen zodra ze in voldoende aantallen in IC-vorm kunnen worden gefabriceerd.

In dit proefschrift wordt een aantal nieuwe sensorsystemen gepresenteerd: een geïntegreerde capacitieve rechtheidssensor, een gecombineerde capacitieve positie–snelheidssensor en een optische rechtheidssensor die gebruik maakt van slechts één laserbundel. Van deze sensorsystemen zijn prototypes gemaakt, getest en gebruikt.

Op het gebied van positiesensoren zijn een tweetal belangwekkende trends zichtbaar. Allereerst de toenemende *integratie* van de benodigde elektronica in het sensorelement, wat resulteert in een sterke afname van de hoeveelheid bedrading, afscherming en losse verwerkingselektronica, en tevens meestal in betere prestaties. Daarnaast de toenemende *miniaturisatie* van sensorelementen met de bijbehorende elektronica in ASICs[2], waardoor een sensorsysteem, bestaande uit een sensorelement, verwerkingselektronica en bijbehorende bedrading, kan worden vervangen door één IC waaruit direct het meetsignaal beschikbaar is in de vorm die het regelsysteem nodig heeft.

Naast rechtheidssensoren zijn ook de bijbehorende rechtheidsreferenties beschouwd. Deze bepalen de uiteindelijke nauwkeurigheid van de meting. Afhankelijk van het soort meting kunnen de vereiste *vormnauwkeurigheid* en de *stabiliteit* van de rechtheidsreferentie variëren.

Teruggekoppelde systemen

Wanneer rechtheidsfouten worden gemeten met behulp van een rechtheidssensor met voldoende hoge bandbreedte, dan wordt het mogelijk om een real-time compensatie te maken voor de gemeten fouten. Een dergelijke compensatie kan worden uitgevoerd door middel van een extra positioneertafel bovenop de beschouwde geleiding, die ten opzichte daarvan kan worden bewogen met behulp van vijf onafhankelijke actuatoren. Deze extra tafel wordt zodanig actief gepositioneerd, dat zijn beweging de vorm van de rechtheidsreferentie volgt. Met andere woorden, de rechtheidsfouten ten opzichte van de rechtheidsreferentie worden zoveel mogelijk gereduceerd.

[2]ASIC: Application Specific Integrated Circuit

Door middel van het real-time meten en terugkoppelen van rechtheidsfouten verschuiven de problemen en bottle-necks van het mechanische domein naar het elektronische en regeltechnische domein, waar ze beter oplosbaar zijn. De uiteindelijke prestaties van het teruggekoppelde systeem zijn niet meer bepaald door fabricagenauwkeurigheid, lagereigenschappen, thermische invloeden of externe belastingen, maar worden hoofdzakelijk begrensd door de resolutie van de gebruikte sensor.

De binnen Micro-techniek ontwikkelde capacitieve rechtheidssensor blijkt goed te voldoen in quasi-statische systemen. Voor systemen met hogere bandbreedte is de meettijd nog te lang. Dit probleem is oplosbaar; nieuwe versies met meettijden aanzienlijk kleiner dan 1 ms zijn in ontwikkeling.

In de quasi-statische experimentele opstelling van een rechtgeleiding met toepassing van de genoemde capacitieve rechtheidssensor wordt een positieresolutie behaald kleiner dan 0.5 μm voor de translaties en 20 μrad voor de rotaties, wat slechts weinig verschilt van de resolutie van de sensor.

Voor het bestuderen van het dynamisch bedrijf van de teruggekoppelde rechtgeleiding worden photonic sensoren gebruikt. Hier blijkt dat de uiteindelijke prestaties eveneens worden begrensd door de sensorresolutie, maar daarnaast ook door de interne dynamische eigenschappen van het mechanische systeem, de bandbreedte van de regeling en het frequentiespectrum van de externe verstoringen. Ook de hysterese van de gestapelde piëzo-actuatoren blijkt een belangrijk knelpunt te zijn.

Een van de belangrijkste nadelen van piëzo-actuatoren, de beperkte slag (in het sub-micrometer of micrometer gebied), kan worden ondervangen door een andere wijze van aandrijven: wanneer een *shear-mode* piëzo-element een zaagtandvormig bewegingsprofiel wordt opgelegd kan daarmee een te verplaatsen object worden voortgeschoven door middel van de wrijvingskracht. Door de traagheid van het aangedreven object is de invloed van de snelle teruggaande beweging van het piëzo-element gering. Deze manier van aandrijven staat bekend als het Inertial Sliding Motion (ISM) principe. Twee uitbreidingen op de bestaande kennis zijn onderzocht aan de hand van een nieuw ontwikkelde 6-DoF manipulator met lange slag in meerdere vrijheidsgraden. Eerst is nagegaan of ook met standaard materialen als staal, aluminium en glas kan worden geconstrueerd, uit oogpunt van kostenreductie. Het blijkt dat dit mogelijk is, maar dat de ontwerpmarges kleiner zijn, vooral wat de voorspanning betreft, en dat de maximale snelheid relatief laag is, in de orde van 50 μm/s. Daarnaast is onderzocht wat de mogelijkheden zijn om het grote aantal hoogspanningsversterkers (de geteste 6-DoF manipulator bezit 15 piëzo-elementen in 12 groepen) te verminderen. Het blijkt mogelijk te zijn om slechts twee versterkers te gebruiken, voor de heengaande en teruggaande beweging, en hun uitgangssignalen met behulp van hoogspanningsschakelaars te verdelen over de betreffende piëzo-elementen.

Wanneer de maximale rechtheidsafwijking van een rechtgeleiding kleiner is dan de slag van een bepaalde (piëzo-)actuator, en deze rechtgeleiding bovendien voldoet aan een aantal eisen betreffende *smoothness* en stijfheid, dan is het zinvol om een teruggekoppeld regelsysteem te bouwen, waarin de basis-rechtgeleiding de *coarse stage* vormt en een daarop geplaatste (piëzo-)manipulator als *fine stage* fungeert. Een mogelijke basisgeleiding voor een dergelijk systeem, gebruikmakend van vet met hoge viscositeit, is onderzocht. Het blijkt dat het vetgesmeerde glijlager welhaast ideale eigenschappen heeft voor toepassing als basisgeleiding voor (zeer) lage snelheden, in de orde van 100 μm/s. De dynamische stijfheid is extreem hoog, de statische stijfheid is laag en door de sterke *spatial averaging* zijn de *smoothness* van de beweging groot en de maximale rechtheidsfouten klein, in het sub-micrometergebied. Door een dergelijke basisgeleiding aan te vullen met bijvoorbeeld een snelle monolithische piëzo-manipulator als *fine stage* is een nauwkeurigheid in het nanometergebied over een relatief grote slag, in de orde van 100 mm, haalbaar.

Nul-stijfheid actuatoren

In het streven naar een positioneernauwkeurigheid in het sub-micrometergebied wordt ook het trillingsprobleem meer en meer belangrijk. Niet alleen de trillingen van buitenaf, maar ook de trillingen die worden opgewekt door bijvoorbeeld lagerfouten kunnen leiden tot ontoelaatbare rechtheidsfouten. Voor dit probleem wordt een volledig andere benadering voorgesteld. In plaats van te streven naar een zo hoog mogelijke open-loop stijfheid van het positioneersysteem ten opzichte van de vaste wereld, wordt een actuatorstijfheid gelijk aan nul voorgesteld. Een stijfheidsloze actuator kan wel een kracht uitoefenen op een te ondersteunen object, maar de relatieve positie van dat object ten opzichte van de ondersteunende actuator is niet van invloed op de grootte van de kracht. Hiermee wordt een volledige trillingsisolatie gerealiseerd.

Eerst is onderzoek gedaan naar de trillingsisolerende eigenschappen van *Lorentz-actuatoren*, de meest bekende actuatoren waarvan de kracht in principe niet afhankelijk is van de relatieve positie van het bewegende ten opzichte van het stilstaande deel. De testopstelling bestaat uit een tafel die in zes vrijheidsgraden was ondersteund door in totaal zes Lorentz-actuatoren. Het blijkt dat inderdaad een goede trillingsisolatie (40 dB) kan worden gerealiseerd met behulp van Lorentz-actuatoren. Het knelpunt hierbij is het niet volledig nul zijn van de stijfheid van de Lorentz-actuatoren, waardoor in het middenfrequentiegebied, ordegrootte 5...50 Hz, slechts 25 dB trillingsreductie werd bereikt. Door *fine-tuning* van de actuatorgeometrie kan de trillingsreductie nog verder worden vergroot.

Daarna is onderzoek gedaan naar de mogelijkheid om een van de belangrijkste nadelen van Lorentz-actuatoren te ondervangen, namelijk het energieverbruik—en overeenkomstige warmteproductie—wanneer er een statische kracht moet worden geleverd. Hiertoe is een *permanentmagnetisch lagerelement* ontworpen dat een statische draagkracht levert die zowel contactloos als inherent stijfheidsloos is en geen energie dissipeert.

Omdat het permanentmagnetische lagerelement in principe randstabiel is (voor verstoringen naar de ene zijde ontstaat een positieve, stabiliserende, stijfheid en voor verstoringen naar de andere zijde een negatieve, destabiliserende stijfheid), is een terugkoppeling noodzakelijk. Het blijkt dat met uitsluitend positieterugkoppeling geen goed functioneren mogelijk is. Een versnellingsterugkoppeling gecombineerd met een positieterugkoppeling en eventueel een snelheidsterugkoppeling geeft wel het gewenste resultaat.

Op het in dit proefschrift beschreven permanentmagnetische basisprincipe van een stijfheidsloos lagerelement is een Nederlands patent verleend en een internationaal patent aangevraagd.

Het stijfheidsloos lageren zorgt voor een positie-ontkoppeling, met behoud van krachtkoppeling, ten opzichte van de 'vaste wereld'. Dat betekent dat van een dergelijk stijfheidsloos gelagerd object (bijvoorbeeld een preparaattafel) de positie regeltechnisch kan worden gekoppeld aan de positie van een ander object (bijvoorbeeld een lens waarop geen krachten mogen worden uitgeoefend), zodat een onderlinge virtuele stijfheid ontstaat. De krachten die nodig zijn om het aangedreven object te positioneren veroorzaken reactiekrachten op de vaste wereld, terwijl de trillingen (positiefouten) van de vaste wereld niet van invloed zijn op de positioneernauwkeurigheid. Dit concept biedt nieuwe mogelijkheden voor het ontwerpen van precisie-positioneersystemen waar de trillingen vanuit de omgeving anders een te grote storende invloed zouden hebben.

Een vervolgonderzoek wordt uitgevoerd naar de eigenschappen van een stijfheidsloze rechtgeleiding. Het centrale lagerelement daarvan is een uitbreiding op het permanentmagnetische stijfheidsloze lagerelement, waarbij de bewegingsvrijheid in één richting onbeperkt is gemaakt. De aandrijfkracht (in de x-richting) en de regelkrachten in de vijf andere vrijheidsgraden worden geleverd door Lorentz-actuatoren. Hiermee is een lineaire beweging te realiseren die in hoofdzaak ontkoppeld is van de positiefouten van de vaste wereld.

Acknowledgements

Many persons have contributed to the realization of this thesis. In particular, I would like to thank:

Professor Van Beek, for his supervision, extensive writing suggestions, thorough proofreading and many valuable discussions and advices;

Jo Spronck, for the countless technical and non-technical discussions, and for being a source of inspiration and creative ideas;

Marcus Bonse, my Master's thesis supervisor and predecessor in this Ph.D. project;

Lex Molenaar, for being a pleasant colleague and also a business associate in establishing the 'Mechatronisch Ontwerpbureau StoneLift', which resulted in the world's first magnetically levitated work of art;

All students and trainees of the Laboratory for Micro Engineering who carried out projects that were valuable for this thesis: (in alphabetical order) Torrence Belfroid, Philipp Blum, Abdelouahid Boujnane, Rogier Egmond, Ian van Hees, Casper Hengst, Aernout Jongepier, Peter Lock, Willem Niesing, Bas Ouwerkerk, Joost de Pee, Pascal van Putten, Warner Venstra, Jeroen de Vroege, Erik Zaaijer;

The Micro-techniek employees for much practical assistance: Frans van Rongen, Els Koets and Johan Buis;

Furthermore, I would like to thank my wife Jojanneke, and our daughters Marjolein en Hannah, because they always help me remember that there are even better things than enjoying my work...

Curriculum vitae of the author

Gert-Jan Nijsse was born in 1971 in Goes. There he attended primary and secondary school. He followed the Atheneum program at the Voetius Scholen Gemeenschap. In 1989 he started his studies at the Delft University of Technology, Faculty of Mechanical Engineering, where he graduated in 1995 at the Laboratory for Micro Engineering.

After his graduation, he took the opportunity to continue the experimental research in the field of mechatronics, by doing a Ph.D. project having feedback controlled straight motion systems as a main theme. The project was finished in 1999, leading to the thesis appearing in 2001.

Gert-Jan currently is employed at Philips CFT (Centre for Industrial Technology), department of Mass Products and Technology, where he is a member of the Drive Systems group.