Active Positioning and Passive Fixation using Friction in Capacitive Displacement Measurement Applications

van de Ven, Oscar

DOI
10.4233/uuid:3fd41e9c-734e-45f9-aca0-c5e658d13657

Publication date
2016

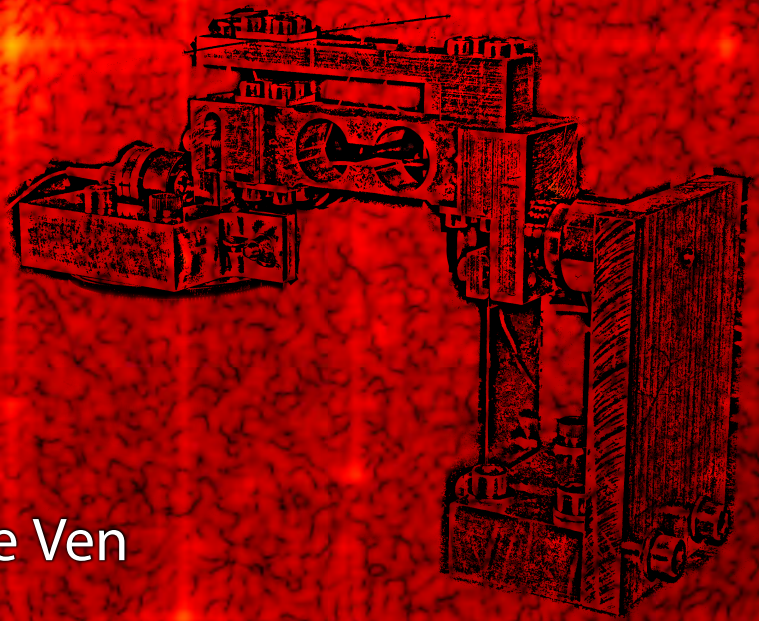
Document Version
Final published version

Citation (APA)
van de Ven, O. (2016). *Active Positioning and Passive Fixation using Friction: in Capacitive Displacement Measurement Applications*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:3fd41e9c-734e-45f9-aca0-c5e658d13657

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Active positioning and passive fixation using friction

in capacitive displacement measurement applications



Oscar van de Ven

Active Positioning and Passive Fixation using Friction

in Capacitive Displacement Measurement Applications

Active Positioning and Passive Fixation using Friction

in Capacitive Displacement Measurement Applications

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op vrijdag 14 oktober 2016 om 12:30 uur

door

Oscar Sebastiaan van de Ven

werktuigkundig ingenieur
geboren te Dordrecht, Nederland.

Dit proefschrift is goedgekeurd door de promotor:

Prof. ir. R.H. Munnig Schmidt

Copromotor: Ir. J.W. Spronck

Samenstelling promotiecommissie:

Rector Magnificus,	Technische Universiteit Delft, voorzitter
Prof. ir. R.H. Munnig Schmidt,	Technische Universiteit Delft, promotor
Ir. J.W. Spronck,	Technische Universiteit Delft, copromotor
Onafhankelijke leden:	
Prof. dr. ir. J.L. Herder,	Technische Universiteit Delft
Prof. dr. ir. P. Breedveld,	Technische Universiteit Delft
Prof. dr. ir. J.P.M.B. Vermeulen,	Technische Universiteit Eindhoven
Dr. S. Nihtianov, MSc,	Technische Universiteit Delft
Prof. dr. U. Staufer,	Technische Universiteit Delft, reservelid

Keywords: Capacitive sensor, aliginment, thermal actuator, friction, pre-sliding, contact, creep, impact, force pulse

Printed by: Ipskamp Printing, Enschede

Front & Back: Background: The spatial frequency spectrum of an interferometer image that was used to measure displacement in three degrees of freedom. Foreground: a picture and a schematic illustration of two positioning set-ups that are treated in this thesis: a Thermal Slider Actuator and a Lorentz actuator that applies a force to the friction contact. The graph on the back shows the measured temperatures of the individual fingers of a Thermal Slider Actuator.

ISBN 978-94-028-0333-4

An electronic version of this dissertation is available at
`http://repository.tudelft.nl/`.

Voorwoord

Ruim vier jaar na mijn eerste dag als promovendus in Delft mag ik beginnen aan het laatste stukje van mijn proefschrift. In dit sluitstuk, dat weliswaar helemaal voorin staat, wil ik een indruk geven van de omgeving waarin dit werk tot stand is gekomen. Eigenlijk begint dat bij de vraag hoe ik bij de Mechatronica onderzoeksgroep van de TU Delft terecht ben gekomen. Mijn fascinatie voor het samenbrengen van mechanica, elektronica en uiteindelijk ook software gaat terug tot ver voor mijn tijd aan de TU, maar pas rondom mijn afstuderen werd de (praktische) mechatronica echt serieus. Ik mocht toen voor Patrice, toen mijn begeleider en nu mijn collega, een volledig geregeld mechanisch systeem opbouwen: een 5 graden van vrijheid haptische parallelle robot waaraan ik ook tijdens mijn promotie nog heb kunnen meewerken. Tijdens mijn afstudeerperiode werd ik benaderd met de vraag of ik het onderzoek naar een thermische actuator voor het nauwkeurig positioneren van een verplaatsingssensor niet voort wilde zetten. Dat daarop uiteindelijk een periode van ruim vier jaar op kantoor, in het lab en op verschillende conferenties volgde moge duidelijk zijn: Dit proefschrift is het tastbare resultaat van deze periode.

De omgeving waarin je als 'jonge onderzoeker' terecht komt heeft natuurlijk een grote invloed op je promotieperiode. Ik had in ieder geval het geluk om op het MSD (Mechatronic System Design) kantoor terecht te komen. Daarmee was niet alleen het mechatronica lab lekker dichtbij, maar nog belangrijker, had ik een hele fijne groep collega's om me heen: Jeroen, Chris, Guido, Ruijun, Rudolf, Arjan, Johan, Jan, Phuc, en Alejandro. Dank jullie voor de goede sfeer, hulp, discussies, koffiepauzes met de PhD én de MSc studenten, uitjes, etentjes, barbecues en vrijdagmiddagborrels.[1] Een paar extra woorden wil ik besteden aan Johan: ondanks dat we aan totaal verschillende projecten werkten hebben we veel en op niveau kunnen discussiëren. Dat is in ieder geval voor mij én voor mijn onderzoek heel waardevol geweest.

Jo speelde in mijn promotieproject de rol van copromotor en dagelijks begeleider. Ik heb daarbij gemerkt hoe fijn het is om iemand in je directe omgeving te hebben die met een enorme ervaring zowel de grote lijnen als de details van complexe problemen doorziet. Dat geldt zeker als je er vrijwel altijd even binnen

[1] For those who prefer English, a translation of the forgoing paragraph: Of course the environment in which you work has a large influence on your time as a young PhD researcher. I at least had the luck of ending up in the MSD (Mechatronic System Design) office. This not only meant that the mechatronics lab was very close, but even more important, that I was surrounded by very nice group of colleagues: Jeroen, Chris, Guido, Ruijun, Rudolf, Arjan, Johan, Jan, Phuc, and Alejandro. Thank you all for the good atmosphere, help, discussions, coffee breaks with both the PhD and MSc students, outings, dinners, barbecues and Friday afternoon drinks.

kunt lopen, hoe druk het ook is. Jo, speciaal bedankt voor je creatieve en pragmatische ideeën, waarvan het helaas niet altijd is gelukt om ze 'eventjes' uit te proberen, je brede interesse en de leuke en leerzame meetings die we de afgelopen jaren hebben gehad. De samenwerking met Rob, mijn promotor, gaat nog wat verder terug. Al tijdens mijn afstuderen mocht ik meewerken aan zijn collegedictaat dat inmiddels is uitgegroeid tot de tweede editie van het mechatronica collegeboek [Munnig Schmidt et al., 2014]. Ik ben blij dat we deze samenwerking, wat betreft mijn eigen onderzoek, en wat betreft het vak mechatronica hebben kunnen voortzetten. Rob, dank voor alle discussies die we hebben gevoerd en voor alle nieuwe inzichten die ik daarin heb opgedaan en dank dat je het daarin niet na hebt gelaten om af en toe te zeggen wat er gezegd moest worden.

Omdat de kern van mijn onderzoek het ontwikkelen van nieuwe actuatieprincipes is geweest, heb ik ook aardig wat tijd in het mechatronica lab doorgebracht, met soms vrij langdurige experimenten. De sprekende verzameling projecten en onderzoekers maakt dit lab een interessante plek. Als je er wat langer rond loopt zie je alle onderzoeken zich in de loop van de tijd ontwikkelen en groeien. Het was daardoor erg leuk om in het lab mee te kijken naar de laatste ontdekkingen, maar het is ook erg leerzaam geweest om mee te denken over de problemen van van mijn mede-promovendi en van een aantal generaties afstudeerders. Mijn speciale dank gaat hierbij uit naar Paul en Martijn, die met hun afstudeerwerk [Ouwehand, 2015, Beek, 2015] zowel praktisch als theoretisch een belangrijke bijdrage hebben geleverd aan de kennis over de pre-sliding- en puls-actuatieprincipes in deze thesis. Deze onderwerpen komen name naar voren in Secties 5.5 en 5.6.

Wanneer je bezig bent met het maken van een opstelling of het uitvoeren van metingen is het goed om te weten dat er genoeg mensen zijn waarbij je aan kunt kloppen voor advies en oplossingen voor allerhande problemen. Mijn dank gaat hierbij vooral uit naar Rob, voor al je adviezen en materialen en voor het brengen van de orde in het lab, naar Harry en Patrick voor alle oplossingen die jullie in no-time hebben bedacht én gerealiseerd, naar Jos voor alle sensoren en andere labapparatuur en naar iedereen van de medewerkerswerkplaats voor alle hulp bij het maken van de onderdelen van mijn testopstellingen. Zoals een lab niet zonder ondersteuning kan, kan een afdeling niet zonder secretariaat. Birgit, Corinne, Eveline, Marianne en Marli, dank voor alles wat jullie in de afgelopen jaren hebben ingevuld, verstuurd, in ontvangst genomen, getekend, geregeld, gecontroleerd, nog een keer geregeld en gedeclareerd. Daarnaast mijn dank voor iedereen, binnen en buiten de afdeling die ik nog niet heb genoemd, maar die wel een bijdrage heeft geleverd aan mijn werk, met name Just, Ron, Anton, Merlijn en Arturo.

Tenslotte wil ik nog mijn dank uitspreken voor al mijn familie en vrienden die interesse bleven tonen in waar ik mee bezig was, maar in het bijzonder mijn ouders, die me altijd hebben ondersteund en ervoor hebben gezorgd dat ik me kon focussen op wat belangrijk was voor mijn onderzoek.

Contents

CHAPTER 1

Introduction and motivation

Permanent alignment is defined as a process that consist of two tasks: positioning the parts to be aligned with respect to each other, and then fixating them in that position. In photo-lithographic machines, sensors have to be permanently aligned in three degrees of freedom, often at an inaccessible location. This means that one device has to perform both tasks. Where the positioning task requires relative motion, fixating requires this motion to be impossible. This thesis presents several solutions to this apparent contradiction and this chapter explains why photo-lithographic machines specifically require permanent sensor alignment.

1.1. Photolithography and the chip industry

Digital technology is becoming increasingly important. Since the 1960s, the era in which Gordon Moore first introduced his well-known growth 'law' [Moore, 1965], the manufacturers of electronic chips or integrated circuits have done whatever they can to keep increasing the complexity and reducing the cost of integrated circuits. In order to achieve this, more components have to be manufactured per chip surface area, and preferably in less time.

Integrated circuits are manufactured on an extremely flat mono-crystalline silicon disk, called a wafer. The circuit is built out of a number of flat layers, as is illustrated in Figure 1.1, of different materials. These materials are either selectively applied to the wafer or applied to the whole wafer and selectively removed, so that the layers together form a three-dimensional structure that contains all the parts of the integrated circuit, such as transistors and wires. After the processing of the layers, the wafer is cut into chips and packaged. This whole manufacturing cycle is depicted in Figure 1.2.

Material is selectively removed or applied on nanometre scale by means of a photolithography process. A layer of photoresist, a light-sensitive material, is applied to the wafer. The illumination process then defines the details of the circuit by selectively shining light on the layer of photoresist. The photoresist on the wafer is then developed and baked, so that it forms a pattern on the wafer which is used for the selective application or removal of material (indicated as 'Processing' in Figure 1.2). The photolithography step defines the patterns in the wafer and is therefore the most crucial step for defining the manufactured geometries: it determines the details within a layer and the mutual alignment between the layers.

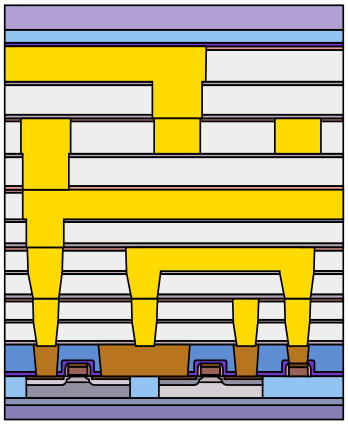


Figure 1.1.: Illustration of the layered structure of an integrated circuit. (This image is based on the work of Cepheiden.)

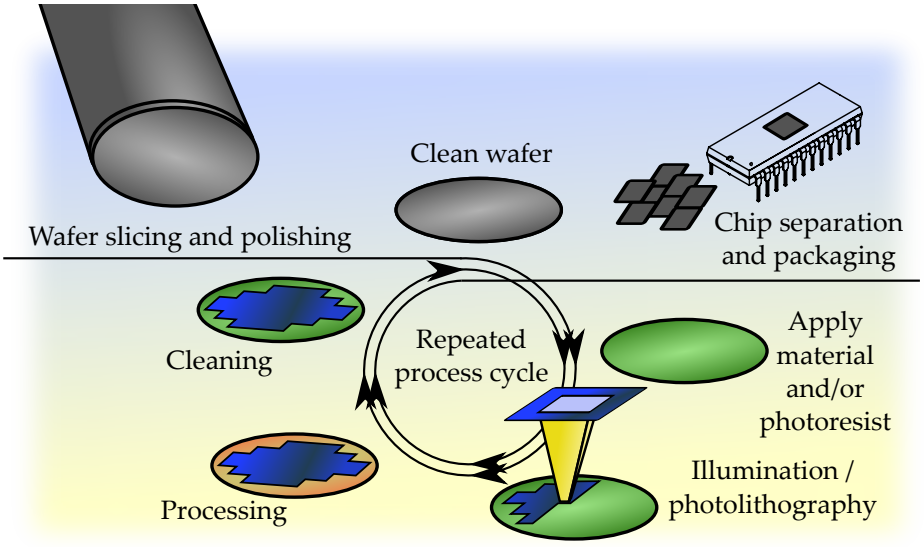


Figure 1.2.: The Integrated Circuit (IC) manufacturing cycle, schematically depicted. Inspired on [Munnig Schmidt et al., 2014].

Figure 1.3.: ASML wafer scanners, on the left a TWINSCAN and on the right an NXT EUV machine (courtesy of ASML).

These two factors are crucial for reducing the size of the details of the integrated circuits and, as they are both determined in the photolithography process, this process is a key enabler for cheaper and faster electronic chips.

Figure 1.3 shows two examples of high performance lithography machines made by ASML, the current world market leader in such machines [ASML, 2014]. The basic principle of the process inside these machines is comparable to developing a picture from an analogue negative. An image of the reticle (the negative) is exposed on the photoresist (the photosensitive layer) on the wafer by means of an optical system, as is shown in Figure 1.4. However, where pictures are magnified, in lithography the reticle imaged four times smaller on the wafer.

The key difference between the illumination of a wafer and a picture, is the accuracy with which the process has to be performed. It is sufficient for a picture to have a resolution of several tens of micrometres, whereas a modern wafer scanner can achieve a resolution well under 30 nanometres [ITRS, 2013]. This imposes stringent norms on the relative position of all the parts of the illumination system, such as the reticle, the lens and the wafer, during the entire illumination cycle, which lasts approximately 20 seconds.

1.2. Displacement measurement of optical elements

As was shown before, the main function of the lithography machine, schematically shown in Figure 1.4, is the accurate imaging of the reticle, through the lens column, on the wafer. The image is projected while the wafer and the reticle move in an opposing scanning motion, and the image is projected many times, in a roster filling the wafer. This requires the reticle and the wafer stage to move during illuminations, but in order to achieve sufficient accuracy the relative positions of the wafer, the lens and the reticle have to be known continuously with an error that is small compared to the details of the pattern.

Many different sources contribute to the total error in the exposure on the wafer

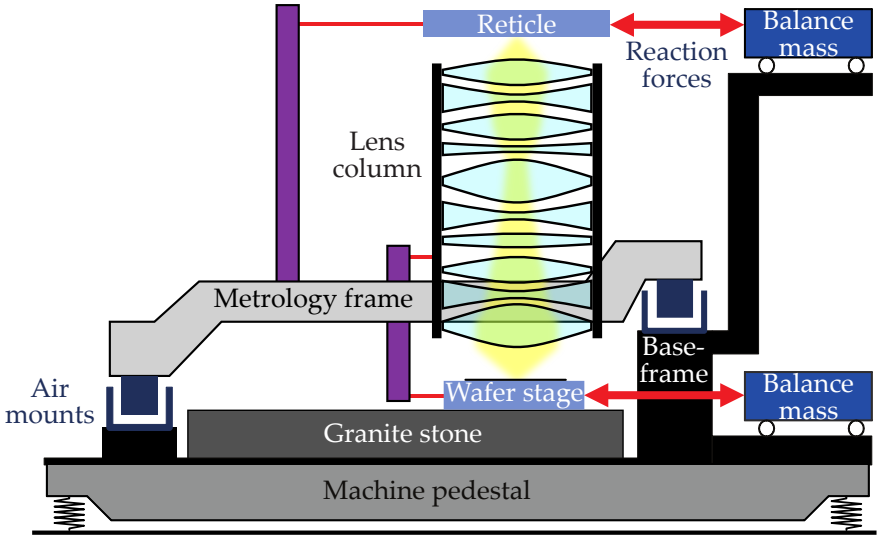


Figure 1.4.: Schematic representation of the architecture of a wafer scanner lithography system. This picture is adapted from [Munnig Schmidt et al., 2014].

and all of them have to be considered for a reduction of the total error. The total error budget includes errors in the lithography process, the scanner, the reticle, the etching process and metrology [Arnold, 2008]. The metrology error is the error in the different measurements that are performed. Many of the relevant measurements are displacement measurements.

The displacement measurement range of the various axes differs significantly. The displacement of the wafer(stage) and the reticle, indicated in Figure 1.4, is in the range of 0.1 to 1 metre, while the motion of the lens column is in the micrometre range. The lens column is actively suspended by the metrology frame. The metrology frame is the reference for the displacement measurements in the machine and it is therefore crucial that it does not deform. This is achieved by using multiple layers of vibration isolation and by limiting the force and the heat that is applied to the frame. External forces and heat sources will however cause displacement of the lens with respect to the wafer or the reticle, which in turn causes imaging errors. These displacements can however be compensated by the other axes when the displacement between the lens and the metrology frame is known sufficiently accurate. This means that the displacement of an object (the lens) that is only moving on a sub-micrometre scale has to be measured.

The relative position of the lens with respect to the wafer and the reticle can be measured directly by means of reference markers on the wafer. During the illumination process, however, these reference markers are not in the field of view. Another measurement system therefore has to take over the measurement between the lens and the metrology frame during the 20 seconds of the exposure, using the measurement through the lens as a displacement reference.

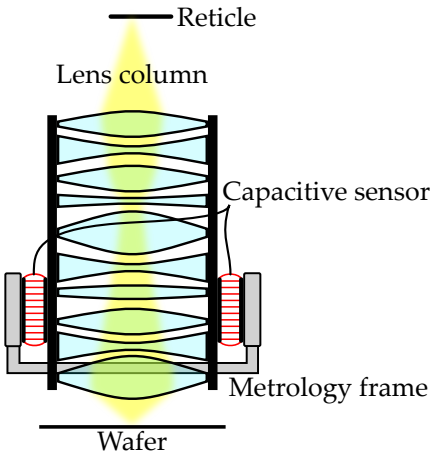


Figure 1.5.: Schematic representation of the lens column system and the metrology frame in which both the lens and the two sensor systems are suspended.

The specifications for the sensor that will perform the lens to metrology frame displacement measurement, originate from the lithography application as derived in [Schieveen and Spronck, 2010]. The key specifications are a measurement range of 2 µm with a 30 pm resolution and a measurement accuracy that is better than 160 pm, within a time-frame of 20 seconds. Based on the precision sensor overviews [Fleming, 2013, Jansen Precision Engineering, 2014], a number of promising high resolution and accuracy displacement sensor concepts can be selected. First of all, optical measurement systems such as interferometers and encoders may be suitable; however, their large complexity and large measurement range causes these systems to be too voluminous and relatively expensive. High resolution can also be achieved with piezoresistive sensors, but this causes a physical connection between the lens and the metrology frame, which is not allowed.

The most suitable sensor technologies are capacitive and eddy-current as they have a high sensitivity in small-range applications, which in turn leads to a high measurement resolution; however, the state-of-the-art sensor systems of these types still show insufficient resolution. For capacitive displacement measurement, the sensitivity can be further increased at the cost of the measurement range by means of both a mechanical and an electronic zoom-in principle [Xia et al., 2010, Schieveen et al., 2009]. The application of a pair of capacitive displacement sensors to the lens column is schematically shown in Figure 1.5.

Generally a capacitive sensor consists of at least two plates made of a conductive material, which are mounted to the objects of which the relative displacement is to be determined. The electric capacitance of these plates, an electric property that is further explained in Chapter 2, corresponds then to their distance. The essence of the electronic zoom-in principle is that not the absolute electrical capacitance, but the capacitance difference with respect to a known and constant 'zoom-in capacitor' is measured. It is therefore possible to measure the capacitance value around

zero, which makes the read-out more efficient and increases the relative sensitivity. The mechanical zoom-in also reduces the range, but does this by bringing the measurement target and the sensor closer together. These principles are described in more detail in Chapter 2 and Appendix A.

Reducing the distance between the electrodes is, however, not trivial. First of all, the required electrode distance is 10 µm, with a maximum distance deviation due to tilt of ± 1 µm. Due to the limited accuracy of (economically viable) manufacturing processes, it is not feasible to achieve this using conventional mounting techniques. Furthermore, the electrode plates will move more than 10 µm due to external vibrations during transport of the machine. Factory-side installation with sufficient accuracy is therefore not feasible and a means has to be provided to reposition the electrodes with respect to each other after installation of the machine and the lens in the clients factory. Re-positioning has to take place automatically, as the sensor will be located at a position deep in the machine that is not accessible for manual adjustment.

1.3. Permanent alignment system

In order to perform a permanent alignment, an automatic alignment system with fixation capability will be included in the lens displacement measurement system of a lithography machine. In the proposed solution, the target electrode will be rigidly connected to the lens, and the measurement electrode will be moved and fixed with respect to the metrology frame. The permanent alignment system will have two distinct modes of operation, which leads to contradictory mechanical requirements:

1. In the *positioning mode,* the measurement electrode has to move with respect to the metrology reference frame in order to position the electrodes with respect to each other, while

2. in the *fixation mode,* the measurement electrode has to be mounted to the metrology frame with sufficient stability, as mechanical drift directly corresponds to a measurement error.

The specific application in the lithography machine leads to several additional *boundary conditions* that define the limitations and opportunities for the sensor system solution:

Alignment time The positioning mode will only be used occasionally, primarily during installation of the machine in which the sensor is incorporated, and occasionally for periodic re-positioning. Because it is only an occasional process, there is no strict time limit for this task.

Position feedback The measurement electrode that has to be positioned with respect to the measurement target, will be used as a displacement sensor when it is positioned properly (during the fixation mode). However, this sensor can also be used to measure the relative electrode position during the positioning mode. Therefore position feedback control is possible.

Passive stability In the fixation mode, position stability, and therefore thermal stability, is important. Active positioning however, requires an actuator action, with the production of heat as a result. Additionally, the position feedback that is described before, cannot be used in the fixation mode. Because of these two reasons, the stability in the fixation mode has to be achieved passively.

Heat production The accuracy and stability requirements prohibit heat production in the fixation mode, in the positioning mode however, these requirements are significantly less strict. It is therefore, to some extent, allowed to produce heat in this mode.

1.4. This thesis

In order to keep improving the performance of our every-day electronics, the size of the details produced in lithography machines need to keep decreasing and thus also the measurement errors within these machines. An important step in improving the measurement resolution and accuracy is the introduction of active sensor positioning that is combined with a good passive stability. Positioning systems that can achieve this form the topic of research of this thesis. The aim is:

> To investigate the motion properties of positioning methods with sub-micrometre scale accuracy, that allow for sub nanometre passive position stability, in the context of capacitive displacement sensor alignment.

The need for limited range, high stability and sub-micrometre resolution alignment is not limited to positioning capacitive sensors. Other applications that share the motion scale and the stability boundary conditions include the alignment of sensors in general, the alignment of optical elements and occasional compensation of slow motions, such as geological mass movement or continental drift. This thesis however focuses on the boundary conditions set by capacitive sensor alignment in the lithographic application.

For the development of a mechatronic system such as the sensor aligner that is the target application in this thesis, it is important to identify the requirements and boundary conditions of the applications, as they define the solution space. Therefore, the positioning and stability requirements that are imposed by the capacitive displacement measurement application are developed in Chapter 2 by analysing the measurement principle. Based on this analysis, a system architecture for the whole measurement system is proposed.

With this set of requirements, different position-and-fix concepts are analysed in Chapter 3. It is first analysed how positioning and holding can be combined fundamentally, concluding that friction contacts are the most promising. Then, a number of concepts that use friction are presented, categorised according to their positioning principle. Finally, a direction for further research is formulated based on these concepts.

The Thermal Slider Actuator (TSA) is selected as the most interesting concept for further analysis, as it seems to be a promising position and hold concept, while not all the required motion properties are yet known. Chapter 4 analyses the TSA motion properties both theoretically and experimentally and identifies which design and control issues have to be taken into account in order to meet the positioning requirements.

The thermal slider actuator contains many friction contacts. An important next step in understanding its behaviour is therefore the study of its typical point-contact friction force-displacement characteristics. From this study, it turns out that the thermal slider contacts move continuously without observing stick-slip effects. This implies that it is also possible to manipulate a friction contact directly and enforce a smooth alignment motion.

The direct positioning capabilities of a single point friction contact are further investigated in Chapter 5. Based on a literature review on the so-called pre-sliding effect and several experimental results, a novel method to control the displacement of a friction contact is proposed. In this method elementary knowledge of the friction phases is used to achieve accurate positioning within the pre-sliding regime. Additionally a third positioning method is proposed that uses short magnetic force pulses to accurately position a friction contact.

Chapter 6 finally summarises the most important findings and recommendations about the Thermal Slider Actuator, the pre-sliding positioning concept and the magnetic force pulse concept. It concludes that either of these principles has the potential to be applied successfully as an alignment system for sub-nanometre resolution displacement sensors and for many other alignment applications.

CHAPTER 2

Capacitive displacement measurement requirements

As was stated in Chapter 1, due to the continuous need for smaller, cheaper and faster electronic devices, the chip-making industry is in a continuous need for lithography equipment that is capable of producing smaller details on electronic chips. The lithography industry in turn requires more precise measurement equipment, among others for micrometre range displacements such as those of optical elements. When aiming for measurement resolutions of around 10 pm, capacitive sensors are a cost effective solution provided that the capacitance can be read out accurately and that the capacitor electrodes can be positioned accurately with respect to each other, prior to measurement. This chapter examines what positioning and stability requirements the capacitive displacement measurement application impose on the 'permanent alignment task'.

The key goal of the capacitive displacement measurement is to achieve the sub-nanometre resolution and accuracy that was stated in Section 1.2. This implies that a high and well-known *sensitivity* and a high *stability* are required. The following sections elaborate how these requirements can be met. In Section 2.1, the relationship between the relative position of the two electrodes and their capacitance is derived. This relationship is used to determine the influence of electrode placement on the conversion sensitivity and stability. Sections 2.2 and 2.3 treat the influence of the electrode distance and orientation respectively. This analysis forms a set of requirements that determines the basis on which positioning and alignment concepts are defined in Chapter 3. Also the conversion from capacitance to a digital number influences the sensitivity and stability of the measurement system. More information on two methods that have been used to improve these properties can be found in Appendix A. With all the parts of the measurement system known, an overview of the proposed measurement system and its specifications is presented in Section 2.5. Section 2.6 contains this chapter's conclusions.

2.1. Capacitance

The physical principle that is used for measuring displacement (or even position) with a parallel plate capacitive sensor originates from the force between electric charges. This interaction force is, in vacuum, given by Coulomb's law for two point

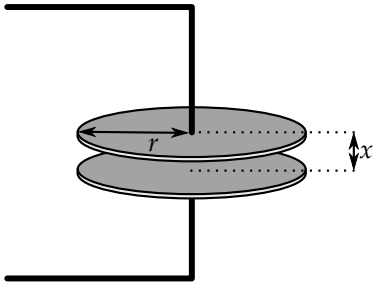


Figure 2.1.: Schematic picture of a finite parallel plate capacitor.

charges [Griffiths, 1999]:

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0}\frac{qQ}{r^2}\hat{\mathbf{r}} \tag{2.1}$$

where $\mathbf{F}$ is the force on charge Q, due to the presence of charge q. ϵ_0 is the permittivity of free space, $\mathbf{r}$ is the displacement vector from q to Q and r and $\hat{\mathbf{r}}$ are the length of $\mathbf{r}$ and the unit vector in the direction of $\mathbf{r}$ respectively. From this relationship the electric force field $\mathbf{F}$ is defined as the force per unit of charge:

$$\mathbf{F} = Q\mathbf{E}. \tag{2.2}$$

The electric field due to n discrete charges q_i at relative positions $\mathbf{r}_i$ is then defined as

$$\mathbf{E} \equiv \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{n}\frac{q_i}{r_i^2}\hat{\mathbf{r}}_i, \tag{2.3}$$

which implies that the electric field is only a function of the distribution of electric charges in space.

By introducing electrically conductive materials, the location and the amount of these charges can be influenced by changing the locations of these conductors, and the charge that flows onto these conductors. Since we want to measure displacement, we can introduce two conductors of which the relative location can be determined by the interaction of the electric charges on them. A convenient way to do this is by measuring not the electric field, but the electric capacitance

$$C \equiv \frac{Q}{\Delta V}, \tag{2.4}$$

where Q is the charge on either one of the conductors due to the potential difference ΔV between the two conductors. The electric potential V is a scalar energy field, defined as the line integral of the electric field

$$\Delta V \equiv \int \mathbf{E}d\mathbf{l}. \tag{2.5}$$

If we use Equations 2.3 and 2.5 in the capacitance definition (Equation 2.4), we see

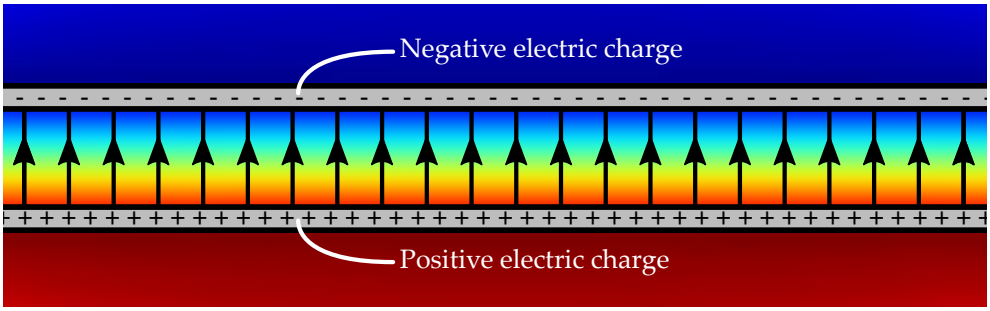


Figure 2.2.: Illustration of the electric field of a pair of charged infinite parallel, plates. The electric field is only non-zero between the plates, as is indicated with the arrows. The colours in the picture indicated the electric potential from high (red) to low (blue).

that the capacitance C is now only a function of the location of the charges and thus the location of the conductors. This implies that the capacitance can be used as a means to determine the mechanical displacement.

The relation between the capacitance of finite plates and their relative distance (x in Figure 2.1) is quite cumbersome to determine analytically [Griffiths, 1999], however, with certain assumptions a reasonable approximation can be derived. This approximation assumes that the field properties are comparable to those of a pair of infinitely large plates.[1] A consequence is that, due to symmetry, the electrical field is directed from the positive plate to the negative plate and exactly perpendicular to the plates between the plates, as is shown by the field arrows in Figure 2.2, and the charge is distributed homogeneously. In practice this is only close to reality when the mutual distance is small compared to the size of the plates and the variation in the distance over the area of the plates is small.[2]

Gauss's law can be applied to determine the magnitude of the electric field for the case of infinitely large plates [Griffiths, 1999], which in this case simplifies to

$$C \equiv \frac{Q}{\Delta V} = \frac{\epsilon_0 A}{x} \tag{2.6}$$

This equation shows that the capacitance C links the electrical domain in terms of charge Q and potential difference ΔV, to the mechanical domain, making it a useful displacement measurement principle. The infinite parallel plate capacitance approximation of two finite parallel plates, can be generalised for non-vacuum situations by replacing the permittivity of vacuum ϵ_0 by ϵ, the permittivity of the

[1] More accurate (and complex) analytical approximations are available, for example in [Heerens, 1986]. The infinite plate model is however currently sufficient due to the large diameter to distance ratio of the plates.

[2] Especially at the edges of a real plate the approximation is not valid, but when the plates are big compared to their distance, the relative area involved in edge effects is small.

medium at hand:

$$C \equiv \frac{Q}{\Delta V} \approx \frac{\epsilon A}{x}, \tag{2.7}$$

The validity of this analysis depends to a large extent on the validity of the infinite parallel plate capacitance model, which does not take into account edge effects and parallelity. For an electrode diameter d of 10 mm and an electrode distance x of 10 µm, the parallel plate model only shows a structural offset of 10 % with respect to a finite elements model. The relationship derived here is the simplest first order approximation of the relationship between displacement and capacitance.

A guard ring, shown in Figure 2.3, is often added to capacitive displacement measurement systems to reduce the influence of external charges. These guard-rings have the additional effect of moving the boundary effects (stray flux) away from the measurement electrode to the guard ring, reducing the difference between the infinite parallel plate model and reality. A finite elements model of an electrode with a guard ring, a diameter d of 10 mm and an electrode distance x of 10 µm, shows a difference of 1 % with respect to the 70 pF found from the analytical model.

2.2. Displacement to capacitance conversion

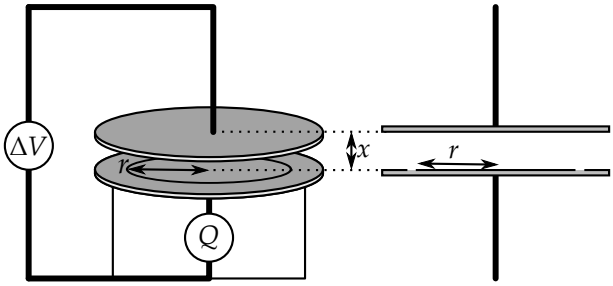
When the electronic part of the sensor measures the capacitance C perfectly, the sensor sensitivity is proportional to the displacement derivative of the capacitance function (2.7), which equals:

$$\frac{dC}{dx} = -\frac{\epsilon A}{x^2}. \tag{2.8}$$

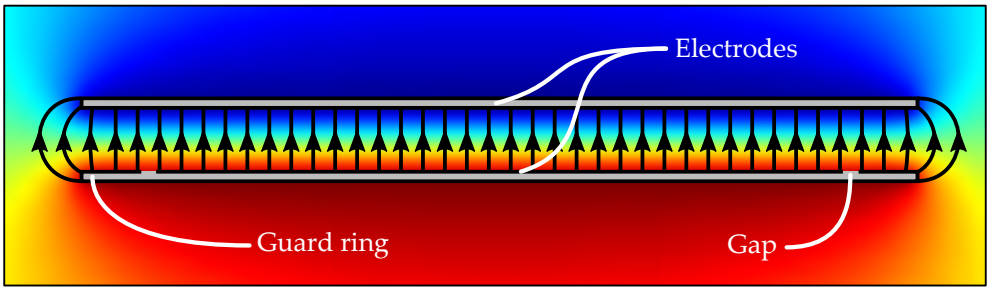
The shape of this derivative as a function of displacement x has the important implication that, in order to increase the measurement sensitivity, an electrode with a large surface area A (large diameter) and a small electrode distance x is preferred.[3]

A large sensor sensitivity implies that the electrodes should be properly positioned with respect to each other because of both the requirement on the electrode area and on the distance. Over the entire electrode area, the distance to the other electrode should not vary significantly with respect to the average distance. This serves two purposes: it prevents mechanical and electrical contact between the electrodes and it prevents deviations in the capacitance-to-displacement characteristic. The latter leads to the most strict positioning specification, which will be elaborated in Section 2.3. When the electrodes are not parallel to each other and either the electrode distance is reduced or the area is increased, the relative distance variation increases. Additionally, the minimum electrode distance should be strictly larger than the expected maximum displacement range that the sensor has to cope with. The current trade-off between area and distance, is an electrode diameter of 10 mm at a stand-off distance of 10 µm. The nominal capacitance is then approximately 70 pF and the sensitivity is around $-7 \cdot 10^{-6}\ \frac{\mathrm{F}}{\mathrm{m}}$.

[3] In this sensitivity analysis, it is assumed that the permittivity ϵ of the dielectric cannot be influenced. This would imply replacing the air or vacuum, for instance with oil or water.



(a) Schematic representation of a parallel plate capacitor with guard ring (bottom electrode).



(b) Electric field and potential of a parallel plate capacitor with a guard ring.

Figure 2.3.: An example of a capacitor with guard-ring configuration. Figure 2.3a shows that a voltage ΔV is applied over the top electrode and the bottom electrode, including the guard ring. The charge Q is only measured at the guarded electrode (the bottom electrode in the figure, excluding the guard ring). This configuration makes it clear that, when the guard-ring is at the same potential as the measurement electrode, most of the edge effects shift to the guard ring (Figure 2.3b). Therefore the presence of the guard ring brings the capacitance of the measurement electrode closer to the infinite plate model of (2.7). Note however that the sensor capacitance does not change when the guard-ring is at a different potential [Heerens, 1986].

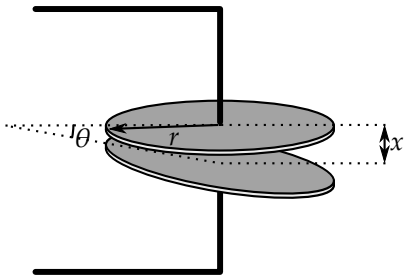


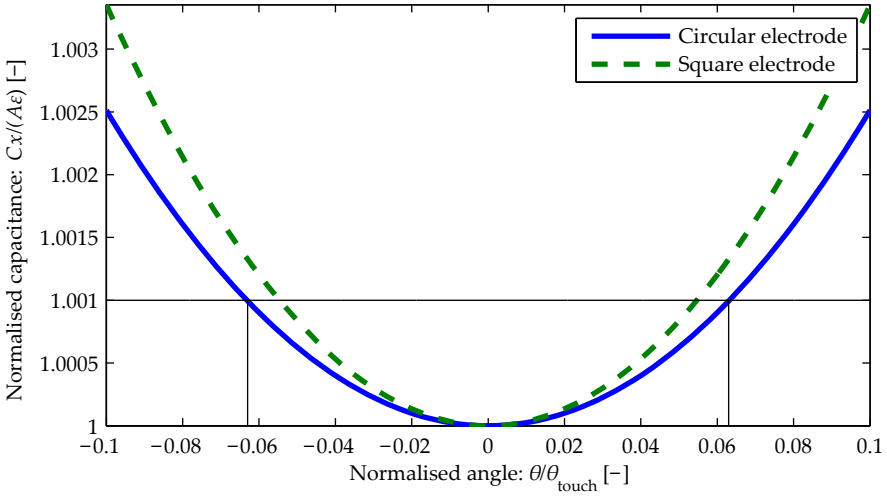
Figure 2.4.: Circular plate capacitor with tilted plates. The parameters are distance x, radius r and relative angle θ.

2.3. Electrode alignment

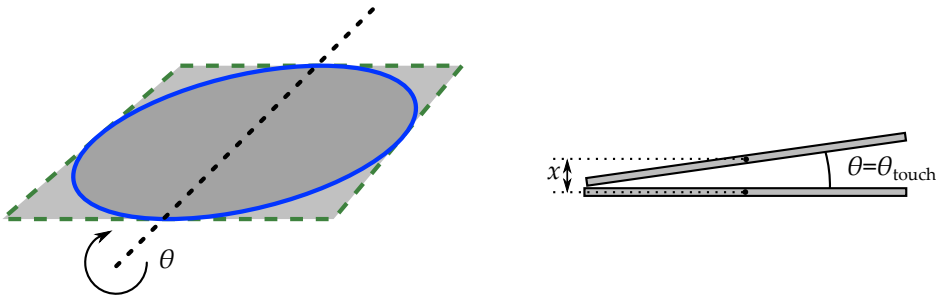
An important factor that influences the displacement measurement accuracy, is to what extent the sensor sensitivity, as noted in Equation 2.8, is known. The sensitivity of the displacement to capacitance relationship is strongly dependent on the relative positions and orientations of the electrodes.

When the electrodes of a parallel plate capacitor are tilted at an angle θ, as is indicated in Figure 2.4, the mutual capacitance changes. It was mentioned in Section 1.2 that the measurement system only has to determine a relative displacement with respect to the position at an earlier moment in time. It therefore seems that a capacitance offset does not have a direct influence on the displacement measurement, however the sensitivity of the sensor is dependent on the distance x as is shown in Equation (2.8), and a sensitivity error corresponds to a gain error in the displacement measurement. To limit this error, either the absolute electrode position (distance and orientation) has to be known with sufficient accuracy, or the sensitivity has to be calibrated in order to determine the magnitude of a relative displacement from a capacitance change.

The tilt angle between the two electrodes (θ in Figure 2.4) is an important factor influencing the displacement to capacitance relationship. The infinite parallel plate model in Equation (2.7) does however not take this angle into account. A first order approximation of the capacitance deviation of a tilted plate can be defined by a series of narrow parallel plates at a varying distance. This method, that neglects the distortion of the electric field with respect to the infinite plate model, is shown for square plates in [Nihtianov, 2008], and is a valid approximation for small angles. It can be adapted to circular electrodes at the cost of losing an easy analytic solution. From a numerical analysis with this method, the normalised graphs shown in Figure 2.5a have been derived. The capacitance is normalised by means of the capacitance of an identical parallel plate capacitor, the angle is normalised by the angle at which the electrodes would touch, defined in Figure 2.5b. Using this normalization, the capacitance converges to the shape that is shown in the graph when the ratio of the electrode diameter and their distance increases. The graph shows clearly that when the capacitor plates are tilted, the capacitance increases at an increasing rate.

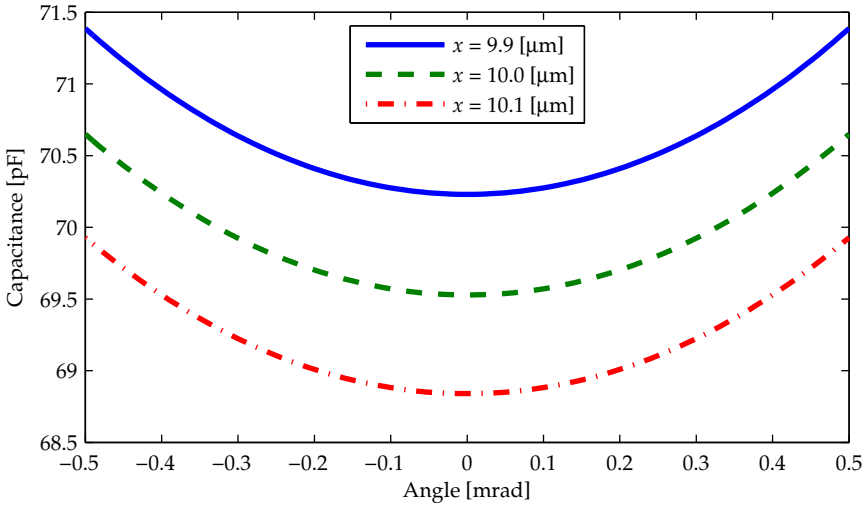


(a) Normalised graph of the influence of the tilt angle on the electrode capacitance. The capacitance is normalised by the nominal capacitance of the parallel plates, that is, at an angle $\theta = 0$. The tilt angle is normalised by the angle at which the edges of the plate touch, as is shown in Figure 2.5b. The approximation only holds for small angles. The boundary for 0.1 % capacitance change due to tilt is indicated for the circular electrode. For an electrode diameter of 10 mm and a stand-off distance $x = 10\,\mu\text{m}$ the maximum allowable angle range is 1.3 mrad. For reference, also the normalised capacitance of a square electrode is included.

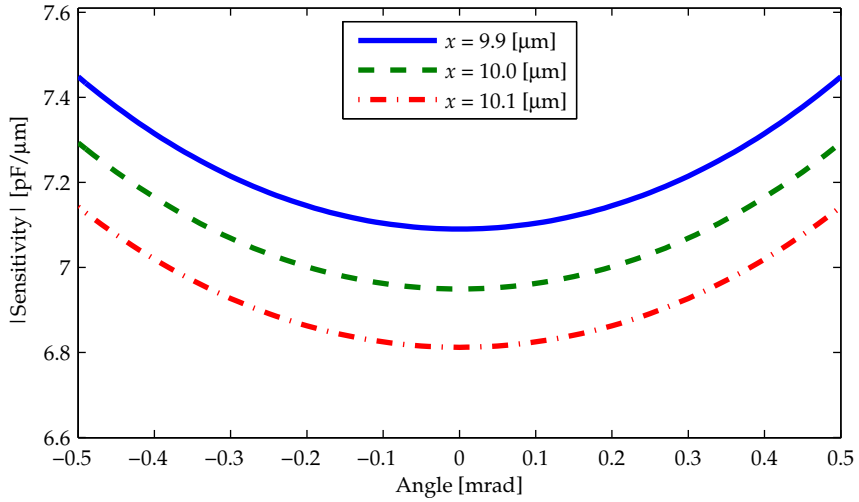


(b) Definitions of the tilt axis and the tilt angle normalization. On the left the dotted line represents the tilt axis for the circular and the square electrode. On the right the electrodes are drawn at a normalised tilt angle of $\frac{\theta}{\theta_{\text{touch}}} = 1$.

Figure 2.5.: The influence of electrode tilt on the capacitance shown in a normalised graph.



(a) Capacitance as a function of the tilt angle θ, for three electrode centre distances x. Electrode tilt increases the capacitance, so that the apparent electrode distance is smaller. The capacitance is approximately 2 % larger at an angle θ of 0.5 mrad compared to zero.



(b) Absolute value of the sensitivity in x-direction as a function of the tilt angle θ, for three distances x. The sensitivity is approximately 5 % larger at an angle θ of 0.5 mrad compared to zero.

Figure 2.6.: Influence of the electrode tilt angle on the capacitance. Circular electrodes with a diameter of 10 mm, and the orthogonal field approximation are used. When the tilt angle is increased from zero, both the capacitance and the sensitivity increase. When observing the sensitivity as a function of capacitance, these two effects partially cancel as is shown in Figure 2.7.

Figure 2.6a shows the capacitance as a function of displacement and tilt angle for a practical electrode pair. However, the measurement goal is not the absolute position, but deriving a displacement from a capacitance difference. The accuracy with which the sensitivity $\left(\frac{dC}{dx}\right)$ is known is therefore a key factor. The slope of capacitance curves shown in Figure 2.6a, is depicted in Figure 2.6b. In practise, this sensitivity can be determined by measuring the total capacitance C, so the relationship between C and $\frac{dC}{dx}$ is critical. When the sensor sensitivity is derived from the absolute capacitance in the presence of electrode tilt, two phenomena play a role simultaneously. When the capacitor tilt angle increases from zero, both the capacitance and the sensitivity increase (in absolute sense), as is shown in Figures 2.6a and 2.6b respectively. When the capacitor plate gap decreases, the absolute capacitance and the absolute sensitivity also increase as is shown in the same figures. A readout system that cannot distinguish between tilt and translation will therefore, based on the capacitance, see a tilted electrode as an electrode with a smaller gap and use the associated larger sensitivity. As these tilt and translation effects work in the same directions, their effects partially cancel out. In order to determine the total sensor gain error due to electrode tilt, both effects have to be combined into the sensitivity for constant capacitance, as is shown in Figure 2.7. The sensitivity variation for constant capacitance is indeed significantly smaller than for constant gap. This becomes clear when comparing Figures 2.7a and 2.6b.

Using the sensor sensitivity variation due to tilt, an estimate of the allowable tilt error can be made. When a displacement error due to tilt of 0.05 nm is allowed (about one third of the measurement error budget) within a motion range of 20 nm, the corresponding relative gain error is ± 0.0025. This results at a nominal distance of 10 µm, in a useful tilt range of ± 0.27 mrad around zero, as is indicated by the black box in the graph of Figure 2.7b. At this tilt angle the electrode is positioned within a 10 ± 1 µm distance. If the tilt range would be around a non-zero tilt angle, the allowed tilt error would be significantly reduced.

Reducing the measurement error due to electrode tilt can be achieved by means of (a combination of) several approaches: auto-aligment, active alignment and tilt compensation. The auto-alignment approach, described in [Xia et al., 2010, Schieveen et al., 2010], uses only linear actuation of the electrode to make the electrodes touch each other so that they automatically align. A linear motion in the opposite direction then separates the electrodes to achieve the required gap. [Xia et al., 2010] showed that this approach is capable of alignment within 80 µrad for a 10 mm diameter electrode, without the need for additional tilt measurement. However, the capacitive measurement electrode to be aligned can also be segmented, so that also the tilt angles can be measured directly. This is shown schematically in Figure 2.8. When this tilt measurement is available, the tilt angle can be measured after the electrode is positioned so that the capacitance to displacement sensitivity can be compensated for tilt. Finally, active closed loop alignment is possible when the tilt measurement is available continuously in the positioning mode. (The different modes were defined in Section 1.2.) This does require an actuator with 3 Degrees of Freedom (DoF) instead of 1, but it does create the possibility to bring the tilt angle close to zero, reducing the uncertainty in the tilt measurement as is

discussed before. A final option is in-situ calibration of the sensitivity, which requires a well defined calibration displacement step.

2.4. Measurement stability

The preceding part of this chapter is mainly concerned with the system requirements originating from the positioning mode of the measurement system. In the second, fixation mode the position stability is important. Stability is here defined as the capacitance deviations when the measurand is at standstill with respect to the metrology frame. The electrical capacitance of the displacement sensor is only dependent on the geometry of the electrically conductive elements in and around the sensor and on the dielectric properties of the material between and around the electrodes. Variations in the dielectric properties can be easily caused by variations in temperature and humidity, which means that these quantities have to be constant, or the variations have to be well-known. This is a boundary condition for environments where sub-nanometre level displacements are of interest in general.

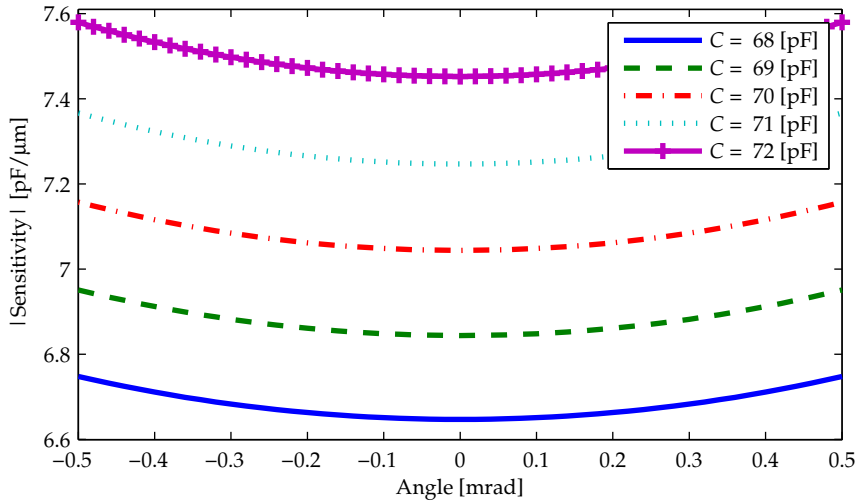
The other factor that compromises the measurement stability is the mechanical stability of the target electrode with respect to the measurand and of the measurement electrode with respect to the metrology frame. It is feasible to permanently mount objects with a good mechanical stability. [Lorenz et al., 2015] shows that no position deviations were observed within an accuracy of 1 to 4 nm for, among others, screw connections over a period of approximately 1 year.

It can be concluded from the foregoing that achieving sufficient mounting stability is feasible. A problem arises however when the capability to displace the measurement electrode with respect to the metrology frame has to be combined with a good stability with respect to the same frame after this alignment has taken place. The majority of this thesis is about the investigation of different solutions to these apparently contradictory boundary conditions.

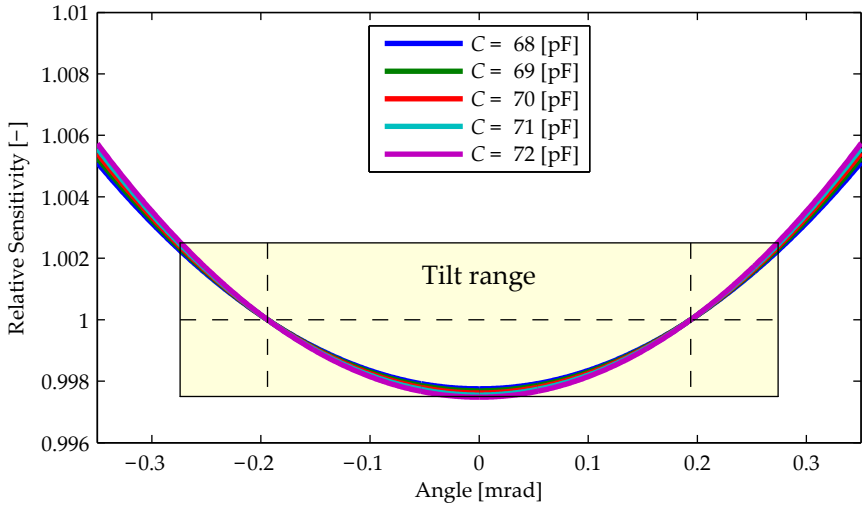
2.5. Measurement system architecture

Placing an object at a 10 µm distance is not trivial and requires either very high machining tolerances or an additional adjustment mechanism. Manufacturing with (sub-)micrometre tolerance levels is mostly not feasible due to technical or financial constraints. Additionally, during transport of the machine, incidental motions of the object to be measured with respect to the measurement electrode occur that are, at 100 µm, significantly larger than the stand-off distance that is required during measurement. In order to avoid damage to the system and in order to achieve the required electrode placement accuracy, the electrode position will have to be readjusted. As the sensor is mounted in an inaccessible location, this is done by implementation of an automatic adjustment system.

Subsystems As the measurement system will fulfil two separate tasks, it will have at least two separate subsystems, one for each task: a mechanical subsystem for



(a) Sensitivity as a function of the capacitance and the tilt angle. On each line the capacitance is constant, which means that for an increasing tilt angle θ, the electrode gap x increases. For instance at 70 pF, the distance increases from 9.9 µm to 10.1 µm between 0 and 0.5 mrad. Note that the sensitivity variation is significantly smaller than for a constant gap distance as is shown in Figure 2.6b.



(b) Relative sensitivity as a function of the capacitance and the tilt angle. The reference value, indicated with the dashed lines, is chosen such that the sensitivity error remains within ±0.25 %. This results in a ±0.27 mrad tilt range.

Figure 2.7.: Sensor sensitivity at constant capacitance.

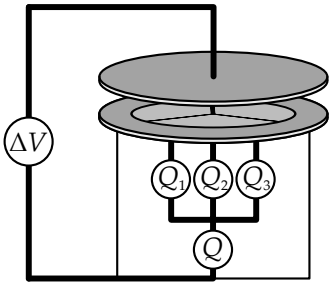


Figure 2.8.: Schematic representation of a segmented electrode with a guard ring that can be used to measure both the electrode distance and the tilt angle.

positioning and fixating the electrode and an electrical subsystem for converting the electrical capacitance to displacement data. An overview of their functions is given in Table 2.1 for three operational phases of the measurement system. The electrical subsystem is developed by the Electronic Instrumentation Laboratory (EIL) at Delft University of Technology as is described in Appendix A and [Ven et al., 2012b]. The content of the mechanical subsystem is the topic of this thesis. The remainder of this section briefly describes how these systems work and communicate with each other and with the controller. The subsystems are also shown in Figure 2.9.

In the measurement phase, the mechanical subsystem (the actuator) is entirely passive and can be disconnected from the controller so that the electrical and thermal interference in minimal. The only electrical connection is to the electronic subsystem, to read out the sensor capacitance. This subsystem is in turn connected to the host that uses the measured data.

In the positioning phase, the actuator is connected to the controller. As the actuator position is changing, the position signal is not useful for the host machine, but it is used by the controller. It is important to note that the measurement requirements in both phases are different. In the measurement phase the measurement electrode is considered to be fixed, in order to determine the target electrode's displacement with a 0.1 nm measurement accuracy. In the positioning phase however, the target electrode is considered to be fixed and the measurement electrode is moved with an accuracy that is better than 1 µm. This is possible because the target electrode and its suspension can be assumed to be stationary on a 1 µm level, but not on a 0.1 nm level.

Also the other requirements on the position measurement are different in the positioning phase than those discussed before with respect to the measurement phase. The position accuracy can be three orders of magnitude smaller, however a value for the absolute distance is required. Also the measurement range increases from 20 nm (short term) for measurement to 100 µm for positioning. The required bandwidth of the position signal during positioning strongly depends on the type of actuator as will become clear in Chapters 4 and 5.

In Table 2.2 the actuator specifications in the two different phases are summarised. The absence of a limit for the thermal load and the long time that is available for alignment are important boundary conditions of the positioning phase that

Table 2.1.: The sub-system functions of the capacitive displacement measurement system. In the positioning phase the measurement electrode is positioned and aligned to the target electrode. The optional calibration phase can be used to determine the sensitivity of the sensor in the current location, so that in the measurement phase an accurate displacement value can be derived from the capacitance change.

Phase	Mechanical subsystem	Electrical subsystem
Positioning	Move the measurement electrode from the transport position to the measurement position & Tilt the measurement electrode to the target electrode within the allowable tilt range	Provide absolute position feedback on the electrode distance and orientation
Calibration (optional)	Move the electrode in a reproducible manner	Monitor the capacitance change so that the sensitivity can be derived
Measurement	Fixate the electrode with sufficient stability	Output accurate displacement data with respect to the displacement at the moment of calibration

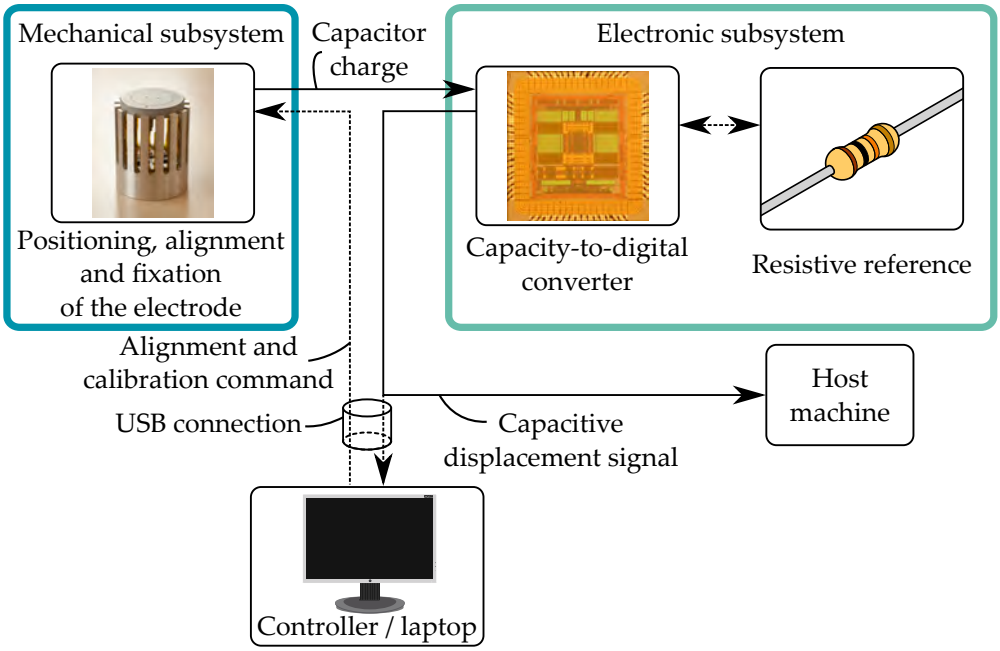


Figure 2.9.: System overview of the capacitive displacement measurement system. This picture is adapted from [Ven et al., 2012b]. The solid arrows indicate continuous communication whereas the dashed arrows indicate occasional or intermittent communication.

will be used in the following chapters. The measurement system properties and specifications that are given in Table 2.2 have been used throughout this thesis as a guideline for the lithographic sensor application.

2.6. Conclusion

In order to achieve capacitive displacement measurement at 10 pm resolution, the measurement electrode has to be positioned within 10 ± 1 µm from the target electrode. A small electrode distance is required, as a reduction quadratically increases the sensor sensitivity (Equation 2.8). When this distance reduction is combined with a requirement for displacement measurement with sufficient accuracy, the uncertainty in the sensor sensitivity due to electrode tilt has to be well controlled. The measurement electrode should therefore be parallel within 0.3 mrad or have the capability to perform a sensitivity calibration. In this thesis the focus is on the first option: aligning the electrodes. As the displacement sensors will be installed at an inaccessible location in a lithography machine, and as occasional realignment is necessary, an automatic alignment system is required.

The alignment system should not compromise the position stability more than 0.1 nm per minute, and therefore the system to be developed should be able to both position and fix the measurement electrode sufficiently accurate and stable. Also the electronic subsystem that converts the sensor capacitance to a displacement value must be able to cope with these levels of sensitivity and stability. When assessing solutions for the mechanical subsystem that performs the position-and-fix task, it is important to take into account that the positioning task will be performed during installation and not during operation. Alignment duration and thermal loads do therefore not cause significant limitations.

Table 2.2.: Guidelines for the measurement system specifications in terms of measurement and positioning. The values given here originate from this chapter, or are derived from earlier work [Schieveen and Spronck, 2010]. These values are used as a guideline throughout this thesis.

Specifications relating to the measurement phase	Value
Measurement accuracy (total error over 2 minutes)	0.16 nm
Sensor position stability (drift over 2 minutes)	0.1 nm
Measurement resolution	0.01 nm
Short term measurement range (2 minutes)	20 nm
Long term measurement range	2 µm
Temperature change during measurement	< 10 mK
Electrode distance	10 ± 1 µm
Electrode alignment error	± 0.3 mrad
Electrode diameter	10 mm

Specifications relating to the positioning phase	Value
Maximum positioning time	1 hour
Positioning range (displacement during transport)	100 µm
Alignment range	0.1 rad
Positioning goal	10 ± 1 µm
Alignment goal	± 0.3 mrad

CHAPTER 3

Position-and-fix using friction

One of the conclusions of Chapter 2 was that a three degrees of freedom (3 DoF), sub-micrometre level positioning system, with picometre scale stability, is required to achieve picometre precision using capacitive displacement sensors. In other words, the measurement electrode has to be properly positioned and after that the position has to be fixed[1]. In order to determine how to combine these two functions, this chapter presents a number of position-and-fix concepts and selects the three concepts that will be further analysed in the chapters that follow.

The number of fundamental methods for fixing an object´s location and orientation is limited. The available fixing methods, and especially their applicability to the sub-micrometre resolution positioning that is required for the sensor alignment application, are analysed in Section 3.1. It will turn out that using friction contacts is the only feasible way in which passive stability and movability can be combined. The classical description of the relationship between force and motion in friction contacts, introduced in Section 3.2, will be used to find several position-and-fix concepts that use friction in Section 3.3. The base concepts are either piezo or piezo-inertia concepts. Additionally, several concepts that can be derived from these piezo(-inertia) concepts, but have a different means of actuation are presented. Finally, in Section 3.4 the concepts that are the most promising for further research are selected and this choice is substantiated.

3.1. Mechanical fixation

There are many different methods with which the position of an object can be passively coupled to that of another object. They can typically be subdivided into three subcategories that are not necessarily mutually exclusive. These are force locking, shape locking and material bonding, depending on whether the presence of a force, shape feature, or bonding agent provides the position fixation. These principles are schematically shown in Figure 3.1, completed with active positioning. All examples in this figure are concerned with a shape lock in horizontal direction. In case of force locking, a force has to be maintained continuously to keep the fixation active. This force can either be a pre-load force, for instance from a spring or from

[1] One could argue that for instance a thermal or piezoelectric stack actuator would be able to perform this task without the need for additional fixation. The combination of range and stability that is required makes this solution however not feasible.

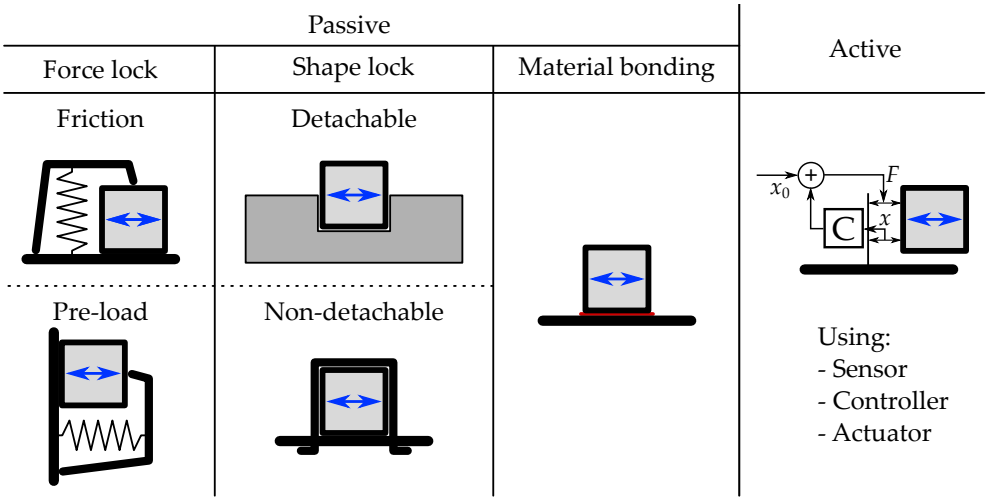


Figure 3.1.: Overview of the different types of position locking in (at least) the horizontal direction. Only the passive principles are considered as solutions in this chapter.

gravity, or a friction force that is generally caused by a pre-load. The first column of Figure 3.1 shows typical force lock mechanisms.

A shape lock, also known as a form lock or positive connection, does not require a force to be continuously present, as the shape of the objects prevents motion in certain degrees of freedom. A shape lock can be detachable, often by moving the object itself or an intermediate object (e.g. a hook) in a not locked degree of freedom. This is not the case with non-detachable shape locks, which are for instance created by plastic deformation.

Material bonding, shown in the third column, relies on a bond on a molecular level. Examples are the use of glue, or phase changes such as is the case with soldering, welding and freezing. Active position locking, which uses a sensor, an actuator and a controller to maintain a position, is added in the last column for completeness. It will not be considered any further as a passive solution is required.

Not all categories of position locking are suitable for this position and hold task. Three important conditions that have to be met are:

1. locking must be possible at different locations that are sufficiently close spaced to achieve the 1 µm positioning resolution,
2. it must be possible to automatically switch between the locked and unlocked state in order to reposition, and
3. the position stability when locked must be sufficient to keep the short term (2 minutes) position drift within 0.1 nm.

Condition 2 limits the applicability of the non-detachable shape lock and the material bonding concept as it is often not possible to (un)lock them without permanent

damage or position drift. Also the position stability (condition 3) of many material bonding substances that can be released, is insufficient due to shrinkage and internal stresses. The detachable shape lock and the pre-load force lock principles do allow unlocking, for instance by moving in an orthogonal DoF and releasing the pre-load force respectively. For these principles, re-locking in different locations does however require mechanical locking- or positioning features that are spaced at at least the positioning resolution. This implies reproducible features on micrometre scale, which is not feasible. Friction force locking does not have these problems. It can be unlocked by removing the contact normal force or by applying a sufficiently large shear force. The locked position resolution is either related to the molecular structure or the surface roughness. Either case leads to sufficient resolution, and the position stability is amongst the highest that can be obtained according to a comparison study of unloaded mounting over a period of more than 1 year at the PTB [Lorenz et al., 2015][2].

3.2. Friction contacts in motion

As it was shown in the foregoing comparison of fixation principles, friction currently appears to be the only feasible way to non-permanently fixate the position of a mover on sub-micrometre position scale with sufficient stability. The next step is to analyse how these friction contacts behave when a motion is induced. Interacting surfaces have been the subject of (scientific) study for centuries in a field that is currently known as tribology. For understanding the concepts in the remainder of this chapter, a description of the relationship between the contact normal force, shear force and shear displacement as it is commonly used for engineering purposes is however sufficient. The focus will be on so-called static friction descriptions, which assumes that the frictional process does not possesses its own dynamics [Wojewoda et al., 2008].

Many more advanced models exists, as is shown in the overview presented in [Olsson et al., 1998]. They for instance include transient effects from the friction dynamics. Anticipating on Chapter 5, where friction dynamics will be introduced, two effects that do not follow from the static description are noteworthy for micrometre-scale motions: a constant friction force is frequently observed during micrometre scale motion, and, also at forces well below the friction limit, relative motion will occur, although on a significantly smaller displacement and velocity-scale.

Static friction descriptions The most basic friction description that will be used here carries the name of Coulomb and states that when a shear force F_s is applied to a moving contact, the two contacting objects experience a friction shear force F_f, as is indicated in Figure 3.2. This friction force is opposite to the motion direction, indicated by the relative velocity v, and this force is limited to a constant: the friction

[2] In this study no noticeable drift was measured over the period of one year for a wrung and a screw connection with a measurement accuracy of approximately 1 nm.

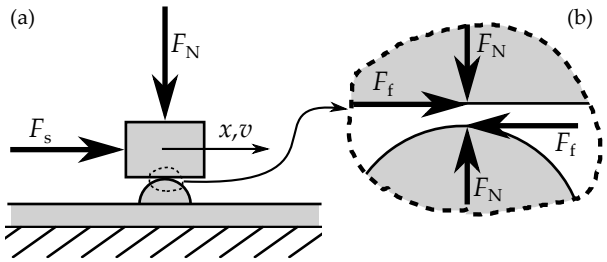


Figure 3.2.: Definitions of the friction contact. Two forces act on the mover: normal force F_N, and shear force F_s, indicated on the left side (a). This results in an internal friction shear force F_f, which under stationary conditions equal to F_s. The friction shear force is indicated on the right (b). For clarity the two contacting bodies are separated. The mover and the contact have a relative displacement x and velocity v.

coefficient μ, multiplied by the contact normal force F_N.

$$F_f = -F_N \mu \operatorname{sign}(v)$$

For a friction contact to start moving, the shear force has to exceed $F_N\mu$. Many variants can be made on this relationship by including additional velocity dependencies, such as the well-known difference between static and dynamic friction coefficients μ_s and μ_d and a velocity dependency described by a Stribeck curve $f(v)$ [Wojewoda et al., 2008]:

$$F_f = F_N \mu_d \left(1 + \frac{\mu_s - \mu_d}{\mu_d} f(v)\right) \operatorname{sign}(v)$$

The additional factors could include both an increase and a decrease of the friction force as a function of velocity. The most important observation is that, in the classic friction theory, a friction contact remains stationary when the friction force remains below a certain limit, and that a (possibly different) friction force limit exists above which relative motion can be present.

3.3. Friction based actuation and fixation concepts

In order to move an object with respect to a base, while it is in contact with this base through friction, either the normal force has to be reduced to zero, so that the contact can be moved without interacting forces, or a force has to be applied to the contact in order to overcome the friction force. A shear force F_s is required in both cases: to provide the mover inertia force and if necessary the friction force. Piezoelectric transducers are commonly used to apply these forces in precision positioning devices, as is shown in [PI, 2008]. [Wulp, 1997] presents an overview of the fundamental principles with which piezoelectric elements can be used to move

a friction suspended object on the nanometre scale. These principles will be used as a basis for the positioning concepts presented here and they are shown in Figure 3.3.

The piezoelectric transducers use the inverse piezoelectric effect, where electrical energy is converted into mechanical energy by means of a piezoelectric crystal [Holterman and Groen, 2013, Munnig Schmidt et al., 2014]. These crystals can be seen as an elastic material of which the neutral length (i.e. the length at which the external force is zero) can be changed by applying an electric potential over the material.

The passive stability of an actuator system is analysed by means of its stability path. The stability path is defined as the path that connects the object to be aligned, to the displacement reference, as is indicated in Figure 3.4 for one of the actuators that will be discussed in the next section. In the case of a friction clamped capacitive displacement sensor (the object to be aligned), the reference is the surface of the metrology frame. The stability path consists of

1. the mover (sensor electrode) material,
2. the friction contact,
3. the base and actuator material and
4. the connection to the metrology frame.

Considering the materials of the base and the sensor, materials that can exhibit bulk creep, such as polymers, should be avoided to prevent sensor position drift, but also the materials thermal expansion must be considered in the material selection and design. It is also important to realise that when elements, such as piezoelectric actuators, are inserted into the stability path this might negatively influence the position stability. An open loop displacement creep of approximately 1 % of the displacement change per logarithmic time decade is mentioned by [PI, 2008, Holterman and Groen, 2013, Jung and Gweon, 2000]. [Munnig Schmidt et al., 2014] specifically advises against the use of open loop piezoelectric elements in (sub-) nanometre stability applications. Finally the friction contact can form a source of instability. This displacement is strongly non-linearly dependent on the contact load as will be further investigated in Chapter 5.

Two final issues that arise when moving friction contacts, are wear and the associated release of particles. As the actuator will only be used for occasional, micrometre scale positioning tasks, the amount of wear and particles will be limited. Also the contacting materials can be chosen such that these effects are minimal [Beek, 2006]. Typical environments in which the alignment mechanism will be used however include ones with low particle requirements. Although the actuator will only be used during installation of a machine and the particle production will be limited, it is still recommended to determine the amount of particles released by the actuator before implementation. The release of particles is however not within the scope of this thesis.

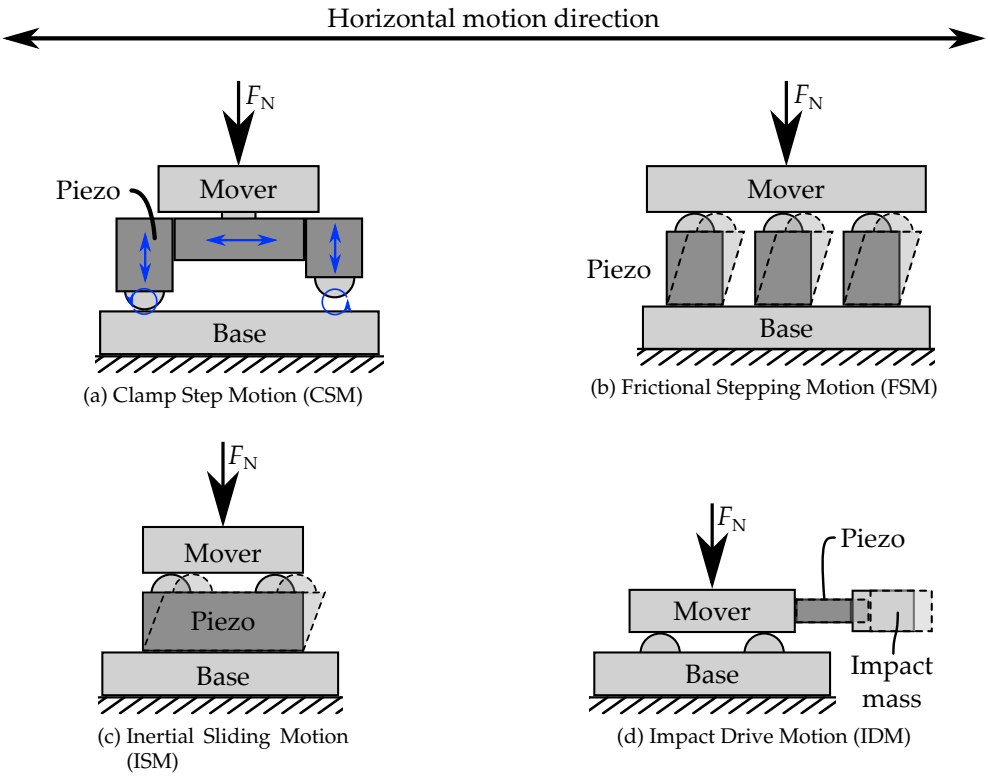


Figure 3.3.: The piezoelectricity based positioning concepts presented in [Wulp, 1997]. The concepts are presented in a simplified form that generates motion in the horizontal direction. The dashed lines indicate another stage in the motion cycle. The CSM and the FSM concepts in Figures 3.3a and 3.3b are in the piezo stepping category (Section 3.3.1). The ISM and IDM concepts in Figures 3.3c and 3.3d both use inertia and are in the piezo inertia category (Section 3.3.3).

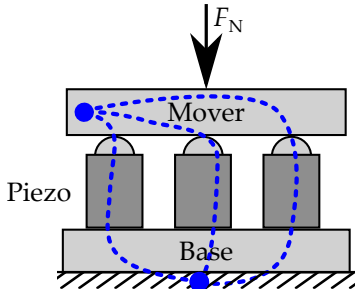


Figure 3.4.: Stability paths of the FSM actuator (see also Section 3.3.1 and Figure 3.3b). The blue dashed line indicates the stability path from an arbitrary point on the fixed world to an arbitrary point on the mover. The paths indicate that the base, the piezos, the contacts and the mover contribute to the stability of the mover.

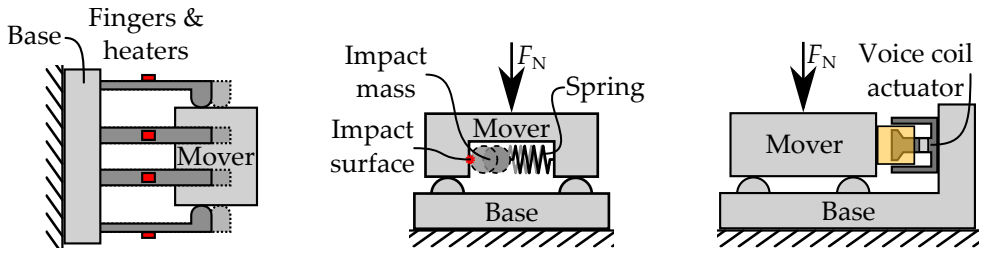
3.3.1. Piezo stepping: CSM and FSM

The piezoelectric concepts that form the core of this chapter are shown in Figure 3.3. The first two of them use a stepping cycle that involves the motion of multiple piezoelectric elements. The *Clamp Step Motion (CSM)* concept in Figure 3.3a has at least two possible contact points between the mover and the base. The contact normal force of one contact is reduced to zero by contracting a vertical piezo element and then the contact is moved by actuation of a horizontal piezo element. During this motion the non-moving point remains in contact to supply the reaction forces. Next the contact normal force is transferred to the other contact using the vertical piezos and the horizontal piezo is moved back.

The *Frictional Stepping Motion (FSM)* shown in Figure 3.3b, also uses multiple piezo elements for motion in one degree of freedom, but instead of using other piezos to reduce the normal force of 2 contact points, many contacts (at least three for a single DoF motion) with comparable and constant normal forces are used. One contact point is enforced in shear direction with respect to the others and due to the smaller total normal force on one contact with respect to the (many) others, it will slide forward while the others remain (approximately) stationary. This is repeated for the other contacts so that after a full cycle the entire mover is displaced and the piezo elements are in their original positions. This means that the cycle can be repeated infinitely. Note that for both concepts the piezo elements are in the stability path, as is illustrated for the FSM concept in Figure 3.4.

3.3.2. The Thermal Slider Actuator (TSA)

The *Thermal Slider Actuator (TSA)* concept, shown in a simplified configuration in Figure 3.5a, can be seen as a variation of the FSM concept discussed in Section 3.3.1. Instead of using piezoelectric elements to load the contacts in shear direction, now thermal expansion is used. Long metallic bodies, referred to as fingers, are individually thermally expanded. Heat is supplied to each



(a) The Thermal Slider Actuator (TSA) uses a propulsion mechanism comparable to the FSM, but uses thermal instead of piezoelectric expansion.

(b) Hammer impact concept (only one direction is possible in the current depiction).

(c) Voice coil force actuation concept. A linear force actuator applies the shear force. This can be any force profile, for instance a pulse or a continuously varying force.

Figure 3.5.: Actuation concepts that use other forms of propulsion. These concepts can be derived from the piezoelectrically driven actuator concepts shown in Figure 3.3. A more detailed descrition of these concepts can be found in Sections 3.3.2, 3.3.4 and 3.3.5.

finger by means of a heating resistor. This concept has been patented by Delft University of Technology [Schieveen and Spronck, 2010] and was first presented in [Schieveen et al., 2009]. Its working principles and several design variables have already been investigated by Jeroen van Schieveen and it showed to be an elegant and relatively easy to implement solution. [Schieveen et al., 2011] addresses several design variables that influence the position stability and the positioning performance. Chapter 4 is a detailed study of the motion properties and specifications of this actuator concept, which completes this earlier research.

3.3.3. Piezo inertia: ISM and IDM

The *Inertial Sliding Motion (ISM)* and Impact Drive Motion (IDM) concepts both use inertia to generate a varying shear force on the friction contacts. With the ISM principle, depicted in Figure 3.3c, the mover is suspended by a piezo element. The piezo is able to move the contact points at different rates. When the piezo moves with a large acceleration in one direction, the contact shear force required to take the mover along is large (due to the mover inertia) and slip will occur. The mover moves therefore less than the piezo. The motion of the piezo in the reverse direction will be at a lower acceleration, such that the mover and the suspension do not slip and move equal distances back.

The *Impact Drive Motion (IDM)* actuator, shown in Figure 3.3d, uses an additional moving mass, next to the mover itself. The piezo element is placed between the two masses so that a force can be exerted on both masses. On the additional mass this will lead to acceleration and deceleration, on the mover this will also lead to a friction force on the contact. Large accelerations in one direction and smaller accelerations in the other, lead to a relative motion in the first direction only. The

IDM concept provides a stability advantage over the ISM, as the piezo element is not in the stability path.

3.3.4. Alternative force pulse actuation

The core of the piezo-inertia concept described in Section 3.3.3, is the application of force pulses to the friction contact(s). These force pulses are generated by a voltage profile on the piezoelectric element(s) and therefore the electrical and mechanical properties of the system imposes limits on the force pulses that can be applied. An example of this is the mechanical resonance of the piezo and the impact/mover mass. Alternative principles to generate force pulses could provide a positioning solutions that does not have these limitations.

One alternative force pulse generator is shown in Figure 3.5b. An impact mass is suspended by a weak spring (as compared to the stiffness of the piezo element). This mass is able to vibrate in its natural frequency at a reaction force level that is below the static friction limit. However, when the mass hits the impact surface, the momentary stiffness will be significantly larger, namely the spring stiffness and the contact stiffness combined. This 'hammer action' results in a short force pulse, that exceeds the friction limit.

Another alternative for applying force pulses is by means of a voice coil-type force actuator as is shown in Figure 3.5c. In this case the reaction forces have to be exerted to the base instead of on the mover. To generate a force pulse, a current pulse has to be generated by an amplifier, or by a dedicated current pulse generator circuit. This solution would have a limited range, however more freedom to choose the force pulse shape is gained. A brief analysis of this alternative impact concept can be found in Section 5.6.

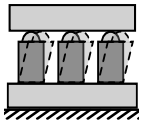
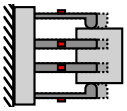
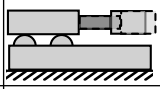
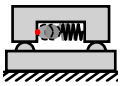
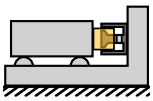
3.3.5. Magnetic force actuation

It was shown in Section 3.3.4 and Figure 3.5c how a friction contact can be actuated by means of force pulses form a voice coil actuator. With the same actuator, a continuous shear force can be applied on the contact as well. In order to apply forces that cause precisely the required motion, a more detailed friction description is normally required. Depending on the specific (non-linear) behaviour of the friction contact, a form of real-time control might be required. A new approach to continuously enforcing a friction contact for precision positioning is presented in Chapter 5. This approach requires minimal knowledge of the friction description, and makes use of the pre-sliding friction regime.

3.4. Conclusion and research direction

Friction clamping is found to be the most suitable solution for combining mechanical fixation with the capability to position an object with (sub-)micrometre accuracy. This apparently contradicting combination of requirements is achieved by exploiting the strongly non-linear behaviour of these friction contacts. Seven basic concepts that

Table 3.1.: Qualitative overview of the positioning concepts that is used by the author. The yellow marked areas contain subjects that are treated in this thesis. Ratings are, from positive to negative: ++, +, 0, - and - -. The items indicated with one or two stars consider continuous positioning using the thermal slider (continuous temperature variation, *) and using the voice coil (motion with a constant force to move within the contact stiffness, **). This is however outside the application boundary conditions and thus outside the scope of this thesis.

	Concept / Property	Piezo stepping: FSM / CSM	Thermal slider actuator	Piezo inertia: ISM / IDM	Impact	Voice coil
						
	Step type	Step-wise & continuous	Step-wise & continuous*	Step-wise & contin. (ISM)	Step-wise	Step-wise & continuous
Stability	Thermal	+	- (finger length)	+	+	+
Stability	Mechanical	-	+ (n. of fingers)	- / ++	+	++
	Resolution	1 nm	100 nm <100 nm*	300 nm / 4 nm	10 nm	10 nm < 1 nm**
	Range	+	∞	∞	Int. hammer ∞ Ext. hammer: -	+
	Speed	+	-	+	+	+
	Parasitic motion absence	+	-	+	+	+
	Volume or mass	0	-	+	-	0
Makeability	Electronics	--	++	-	+	Pulse: 0 Continuous: +
Makeability	Mechanics	--	++	+	1 direction: + 2 directions: -	0
	Robustness	++	++	+	-	+
	Reproducibility or predictability	++	0	+	0	+
	Price (+ low; - high)	--	++	-	0	+
Suitable for:	Sensor alignment	-	+	- / +	+	+
Suitable for:	Continuous positioning	+	-	+ / -	+	+
Suitable for:	6 DOF manipulation	-	+	+	-	0

allow precision positioning of friction contacts have been briefly explained. In the following, the direction of further research will be clarified.

Table 3.1 provides an overview of the seven concepts that were discussed in this chapter, shown in five concept groups. It gives a mostly qualitative evaluation of a concept with respect to the others. Important properties for the sensor alignment task are stability and resolution. Cost is an important argument to explore alternatives to the more conventional piezoelectric concepts. Stability is closely related to the contact type, for instance whether an actuator is in the stability path. The smallest displacement magnitude (resolution), is not only determined by means of the step-size that each concept can achieve, but also by the possibility to move continuously between steps (shown as 'Step type' in the table). The price of a concept is strongly related to the complexity ('Makeability' in the table) and the need for special components such as piezoelectric materials.

Piezoelectric concepts are often chosen for their kilohertz range operating frequency, which allows larger mover velocities with small amplitude actuator movement. This holds for all four piezoelectric concepts that have been presented: The Clamp Step Motion (CSM), the Inertial Sliding Motion (ISM), the Impact Drive Motion (IDM) and the Frictional Stepping Motion (FSM). Within the sensor alignment application, the alignment process is not time critical and therefore a low motion velocity is not problematic. Also heat production is allowed during the alignment process. This opens up the possibility to use alternatives that provide a simpler and cheaper solution. Thermal actuation provides such a solution as it only requires low voltage and current and stable metals can be used as the actuator body. The Thermal Slider Actuator (TSA) potentially provides such a solution that can even be driven from a single USB connection. Its motion properties will therefore be the main subject of further study in Chapter 4.

Also the direct application of forces to a contact by means of a (voice coil) force actuator, either continuously, or step-wise using current pulses, forms a group of new positioning concepts. Exploiting this concept does however require more insight in the (dynamical) properties of friction contacts. The theoretical and experimental investigation of these properties also results in a better insight in the frictional behaviour the friction contacts of the TSA. The friction phenomena that are found in the TSA, even appear to be directly applicable to continuous positioning of friction contacts using friction force control. Therefore the possibilities of direct friction force actuation, both using a continuous force and using force pulses, are further investigated in Chapter 5.

CHAPTER 4

The Thermal Slider Actuator (TSA)

The Thermal Slider Actuator (TSA) is one of the position-and-fix concepts that was introduced in Section 3.3. In order to implement this new actuation principle as a one translation and two tilt degree of freedom (DoF) actuator, the details of its working principle have to be well understood, so that the parameters that influence its performance can be identified and evaluated. This information will then be used to implement a feedback positioning system on this actuator for the sensor alignment application. This chapter describes that process.

In Section 4.1, the TSA is introduced by explaining the working principle and by exploring which motion properties are already known. Grouping of the actuator fingers appears to be an important control variable. The principle of grouping is introduced in Section 4.2. To further analyse the motion properties of the TSA, a thermomechanical model (described in Appendix B) and several experimental set-ups (Section 4.3) have been developed and used. The modelling and experimental results for the motions in the relevant DoF are analysed in Section 4.4. Section 4.5 then shows how these motions can be used to form a closed loop positioning system. The analysis of the Thermal Slider Actuator also leads to several design considerations. These are summarised in Section 4.6. Section 4.7 concludes this chapter by giving an overview of the TSA motion properties that are relevant for the sensor alignment application.

4.1. Introduction to the TSA

4.1.1. Working principles of the TSA

Although the Thermal Slider Actuator (TSA) can be seen as a variation on the FSM piezo concept (Section 3.3.1), the thermal actuation causes the motion behaviour to be significantly different. The Thermal Slider Actuator clamps the mover using a large number of long, compliant elements that will be referred to as fingers. While clamping, it manipulates the mover through thermal expansion of the fingers. The thermal actuation principle and the robustness of the actuation concept for manufacturing tolerances, makes the actuator concept relatively simple and low cost. In most practical cases for the sensor alignment application, the fingers are positioned in a circle, clamping an also circular mover. This structure is commonly referred to as a spring-nest ('verennest' in Dutch).

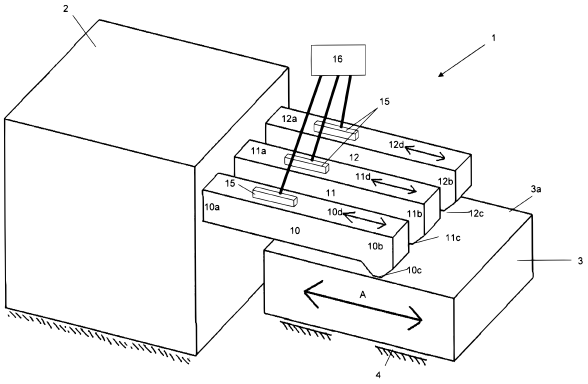


Figure 4.1.: The schematic illustration of the working principle of the Thermal Slider Actuator as depicted in the TSA patent [Schieveen and Spronck, 2009]. It shows a base (2), a mover (3) and three thermally actuated fingers (10, 11 and 12).

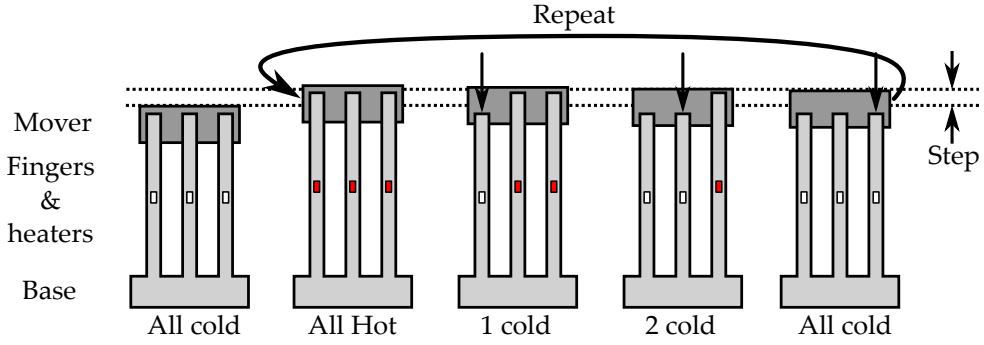


Figure 4.2.: The working principle of a single motion direction Thermal Slider Actuator (TSA) shown on a simplified actuator with the theoretical minimum number of 3 fingers. The upward motion is demonstrated. First all fingers heat up, displacing the mover. When a single finger cools down, it will slide over the electrode surface due to the thermally induced preload, while the electrode is clamped by the other fingers. An arrow indicates the relative motion of the finger with respect to the mover. When all fingers have subsequently cooled down passively, a net displacement of the mover remains. Thermal cycle order can be inverted to achieve a single-dimensional downwards movement. The resistors indicated in red are heating their finger.

The most simplified example that shows the TSA working principle, which is also used in the TSA patent, see Figure 4.1, is flat and single-dimensional. Figure 4.2 shows the motion in one translational degree of freedom with only three fingers, the theoretical minimum. In this situation the mover is free in the vertical direction only and is fixed in all other rotational and translational directions. All fingers are supplied with an electric heating resistor, so that the temperature of each finger can be manipulated individually. When the temperature of a free finger changes, the length of the finger changes proportionally to the thermal expansion coefficient α and the temperature T. When the finger is however fixed, for instance by a friction contact, this thermal expansion changes the force on the finger. These principles are used to manipulate the friction contacts of each finger in a cyclic manner. The temperatures and displacements during such a cycle are indicated in the graphs in Figure 4.3. The states that were indicated in Figure 4.2 are also denoted.

The thermal cycle At the start of a thermal cycle, all fingers are cold. Then the heaters of all fingers are switched on simultaneously and all fingers expand. As the length change of all the fingers is equal, there is no sliding motion in the contacts between the fingers and the mover, and the mover moves along with the finger tips. When the thermal equilibrium is reached, the heater of the first finger is switched off, so that the temperature, and with that the length of the finger decreases. At first this causes only elastic deformation in all the fingers causing the mover to move slightly downwards, but when the shear force on the cooling finger reaches the static friction limit, it starts sliding. During sliding, the mover is stationary. When the second finger cools down, it will also elastically deform the other fingers and start sliding. When this has also happened to the third finger the thermal state of the actuator is identical to the initial cold state. The position of the mover has however made a net motion in the upward direction. The cycle that is shown in this section forms the basis of the actuation of the TSA. It shows that the key to the functioning of this actuator type is the combination of thermomechanical displacement and friction. The two most common variations on this cycle are reversal of the motion direction by reversing the sequence in time and applying the cycle on a varying number of fingers.

The graphs in Figure 4.3 give a qualitative representation of the behaviour of the TSA when the thermal cycle that was described before is applied to an ideal actuator. This is the simplest representation that shows the essentials of the TSA working principle: When the temperature (top graph) of all fingers rises, all fingers (second graph) and the mover (third graph) move equally. Upon the temperature change of a single finger, the elastic displacement of all fingers and the mover is followed by the sliding motion of the finger that is cooling down passively. Ideally the mover is stationary during sliding.

4.1.2. The development of the Thermal Slider Actuator

The TSA working principle was invented and patented [Schieveen and Spronck, 2009] at Delft University of Technology by Jeroen van Schieveen and Jo Spronck in 2008. After the first TSA prototypes, shown in Figure 4.4, had been developed

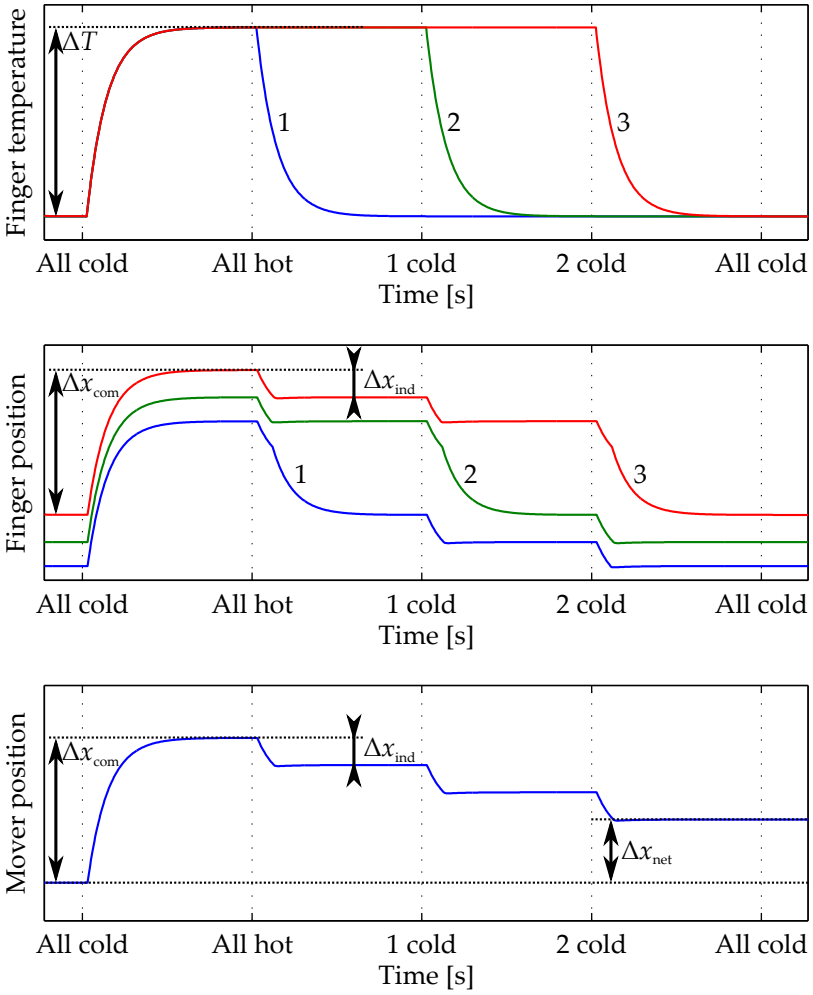
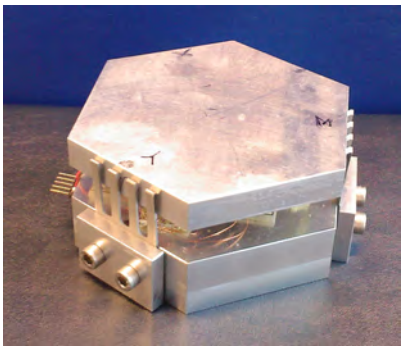
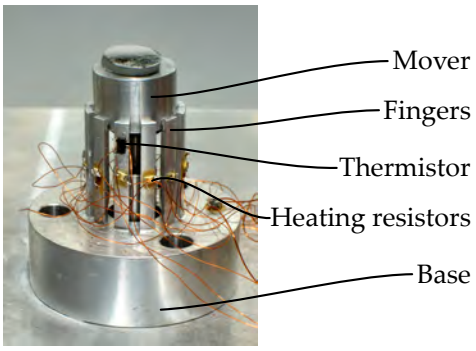


Figure 4.3.: Simulated temperatures and displacements of a 3 finger TSA. The thermal model consists of three elements, one for each finger, and the friction is modelled with only a static friction coefficient. The different lines in the two top graphs indicate the temperature and position changes of the different fingers in the TSA, also a finger number is shown. This simulation represents the behaviour of the simplified actuator shown in Figure 4.2. The step size Δx_{net} and the intermediate displacements Δx_{com} and Δx_{ind} are defined in Section 4.2.2.



(a) The three sided thermal actuator configuration with 3 times 4 fingers, as described in [Schieveen and Spronck, 2009], built by Jeroen van Schieveen and Ruimin Yang.



(b) The first circular TSA prototype built by Jeroen van Schieveen. The base and fingers of this actuator are monolithic. The 12 fingers are heated in pairs opposite to each other.

Figure 4.4.: Early prototypes of the TSA.

and validated, it was shown in [Schieveen et al., 2009] that the circular prototype (Figure 4.4b) could move an object. The actuator has a monolithic aluminium base with 12 fingers of 20 mm length, that each apply a 1 N normal force to the electrode. A thermal cycle was applied to the fingers in opposite pairs where the maximum temperature rise was 2 K (equivalent to 0.9 µm expansion). This results in a net step due to one thermal cycle of over 0.4 µm in approximately 60 seconds.

In 2010 the TSA was used as an open-loop self-aligning sensor head actuator, whereby the measurement electrode was pushed against the target electrode so that they would align themselves by making contact to each other. A predefined electrode distance was then obtained by cooling down all fingers simultaneously from their heated state. This procedure is presented in [Schieveen et al., 2010] and [Xia et al., 2010]. In the presented measurement results, no distinct moment of contact could be detected. However an alignment accuracy better than 80 µrad was achieved.

Although the TSA could be operated entirely in feed-forward, several additions have been made to the actuator system to close the temperature control loop of the fingers. Results of this are presented in [Yang et al., 2010a] and [Yang et al., 2010b], using the actuators depicted in Figure 4.4a and Figure 4.5 respectively. Applying PID control on the temperature does reduce the temperature ripple and the motion speed is slightly improved. The maximum heating power that can be provided to the fingers and the (passive) cooling power remain however important limitations. Continuous measurement of the finger temperatures can be useful in the (passive) measurement mode as well, as it can provide a means to compensate for eventual heating of the actuator.

[Schieveen et al., 2011] describes the analysis of different design and control parameters that influence the performance of the TSA, both in theory and experimentally. The following parameters were identified:

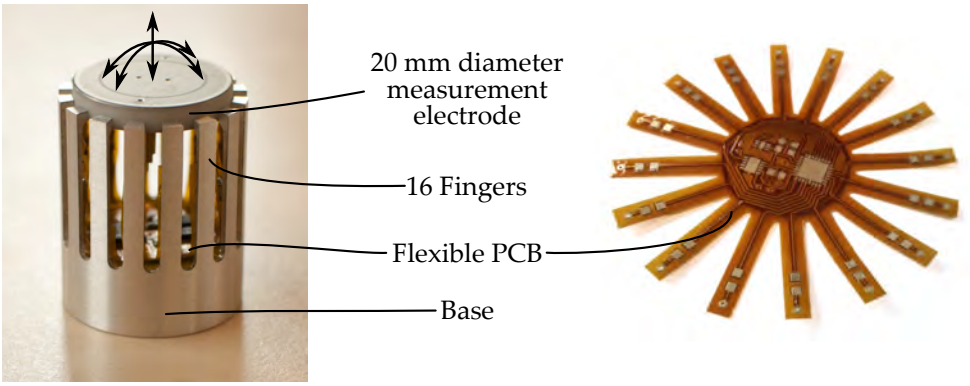


Figure 4.5.: The 16 finger TSA with integrated electronics as designed and made by Jeroen van Schieveen. The heating resistors are mounted on a flexible PCB, which in turn is glued on the inside of the stepper base. In the centre the PCB contains a microprocessor and a power switch/multiplexer, so that only power and serial data has to be provided externally, for instance over a USB connection.

Normal force The normal force that creates the friction between the finger and the mover influences two main system properties. A larger normal force improves the stability (the resistance to external vibrations) as the force needed to move the electrode increases, and it decreases the motion efficiency as more energy (and thus a larger temperature difference) is required to overcome the friction force.

Number of fingers The number of fingers in the actuator influences the stability and the predictability. If the friction (shear) force on one finger exceeds the static friction limit and starts sliding, the sudden position jump is inversely proportional to n^2_{finger} (stability). Also, it is beneficial to increase the number of fingers above the minimum requirement of 3 for a 1 DoF motion system to obtain fixation in the other degrees of freedom, and to average out deviations in frictional properties between fingers (predictability).

Heating cycle Instead op performing the cycle on all fingers (all on and off one by one or vice versa), it is also possible to apply a cycle involving less fingers several times simultaneously. (E.g. heat all, then cool two, then two others, etcetera). Using these groups strongly reduces the efficiency of the cycle, however the time required for a cycle is also significantly reduced. Experimental results show that for a 16 finger TSA (see Figure 4.5) the maximum efficiency is achieved using 8 groups of 2 simultaneously actuated fingers at opposite sides of the TSA, for a heating power of 0.4 W per finger.

Heating power The step size caused by a thermal cycle can be varied by changing the heating power. Below a certain power threshold however, no significant

motion occurs. This is caused by the static friction force that has to be overcome.

Thermal expansion coefficient A larger thermal expansion coefficient α leads, with equal temperature differences, to larger steps. It does however also make the electrode position during the measurement phase more sensitive to environmental temperature changes.

Step temperature difference A free finger (that is, without considering the friction contact) would elongate linearly proportional to the temperature difference, however due to the friction force that has to be overcome to move the finger with respect to the electrode, an additional temperature difference is required. This makes the net displacement per step non-linearly related to the step-temperature. This non-linearity mainly consists of a constant temperature offset.

Temperature control The temperature of the individual fingers can be controlled either open loop or closed loop (that is, involving temperature measurement). By using a PID temperature controller, the shape of the heat-up characteristic can be changed so that the step becomes more efficient. This does however come at the cost of a more complex system, as measurement electronics and temperature sensors have to be integrated. Also the systems becomes less energy efficient due to the added measurement electronics and the added heater amplifiers. As the thermal behaviour reproduces well, similar results could be obtained in feed-forward or using alternative control or optimization techniques.

Cycle control By measuring the displacement resulting from a certain cycle, adaptions can be made to the next cycle in order to reduce the difference between the expected and the real net displacement due to a thermal cycle.

A final development of the TSA is presented in [Khiat et al., 2012], where the TSA is applied on a MEMS (Microelectromechanical systems) scale. Some of the resulting actuators are shown in Figure 4.6. This thermal micro-stepper motor (TMM) was manufactured in a rotational (shown) and a translational variant. Due to the small scale of the MEMS device, the operating frequency can be significantly higher. However, assembly of the mover and the actuator remained problematic. Therefore more optimization is required to make this into a reliable MEMS-scale actuator.

4.1.3. TSA research goals

Many design aspects of the TSA have already been investigated to a large extent, however two major subjects still contain large unknowns. Firstly, only active translational positioning, combined with passive tilt positioning (auto-alignment), has been investigated. This means that the electrodes have to contact each other in order to perform the alignment, which can lead to damage and contamination. To prevent this, also the active alignment properties of the TSA have to be known. Therefore, the thermal cycles that can be used for tilting, their resulting motion, and

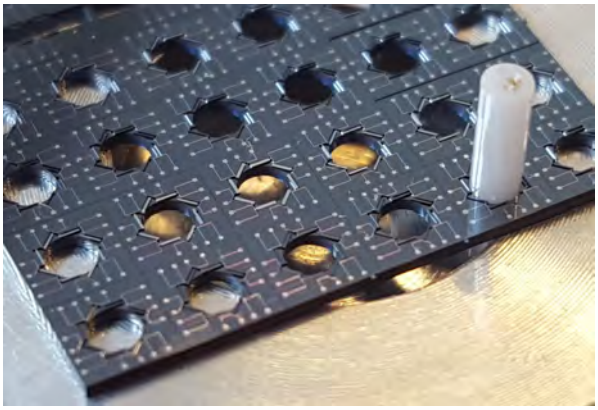


Figure 4.6.: MEMS version of the TSA: the thermal micro-stepper motor (TMM). This actuator provides a rotational motion to a rod of 1 mm diameter, which is friction clamped by 8 fingers. The MEMS actuator is described in more detail in [Khiat et al., 2012].

the way that they are selected, have to be analysed in order to conclude whether active tilting is possible and useful. When active alignment is possible, it can also be used to develop active combined positioning and alignment. This will be done in the remainder of this chapter.

Secondly, the friction principles that govern the motion between the fingers and the mover of the TSA are not clear. In all the foregoing analyses of the TSA, a classical friction model was used. This model only consists of a static friction coefficient μ_s and a dynamic or kinetic friction coefficient μ_d, where generally $\mu_s > \mu_d$. This description mostly results in discrete stick-slip steps during motion. Figure 4.7 shows what the TSA motion would look like using this friction description. Especially the discrete steps with relatively large accelerations of the fingers and the mover are noticeable. However, from all the observations of the TSA using displacement and acceleration sensors, these steps were never seen in practice. Therefore, in order to better understand the behaviour of the TSA, the friction description of the contacts that are typical for this actuator has to be revised. This will be described in more detail in Chapter 5. Conclusions for the TSA can be found in Section 5.4.

4.2. Grouping of the TSA fingers

The principle of grouping determines how many and which fingers of the TSA are operated simultaneously, which has a significant influence on the performance of the actuator. This influence will be shown using a simplified analytical model of the actuator and the actuation cycle.

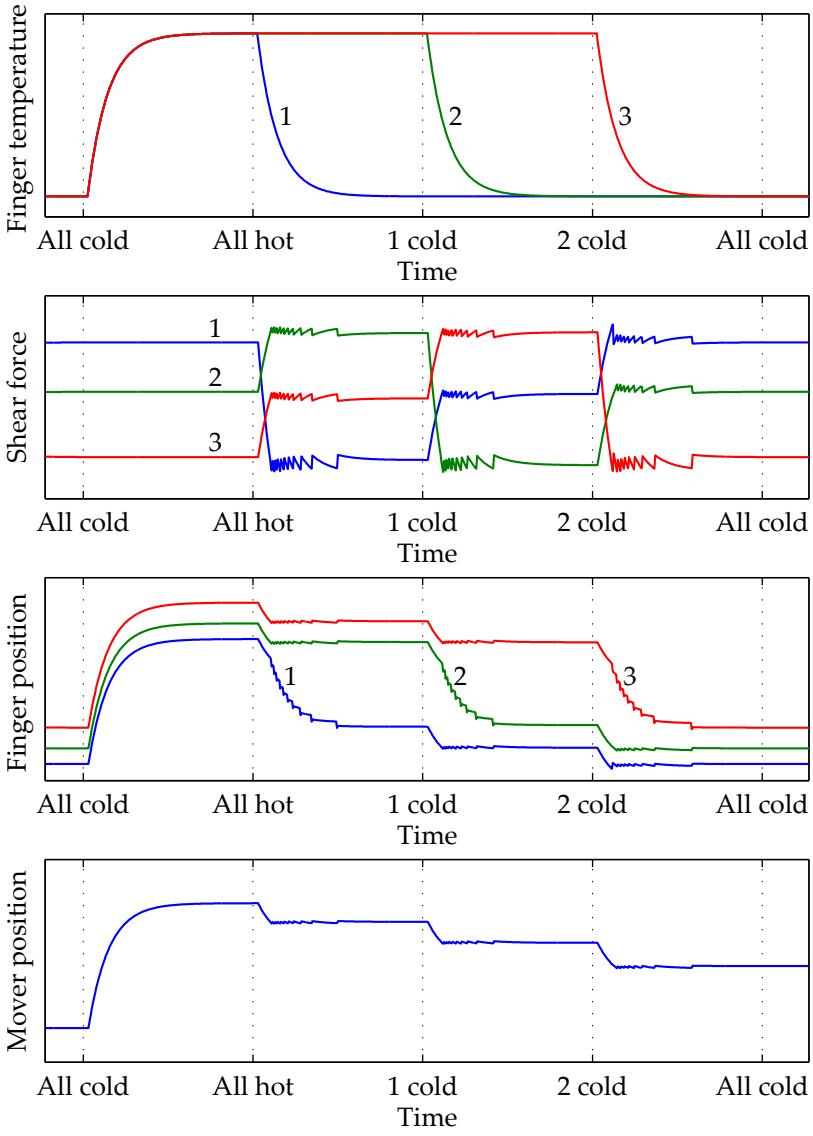
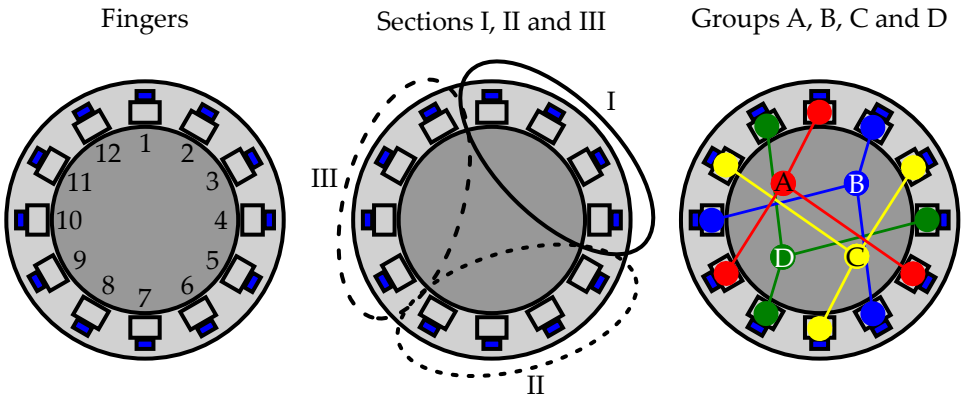
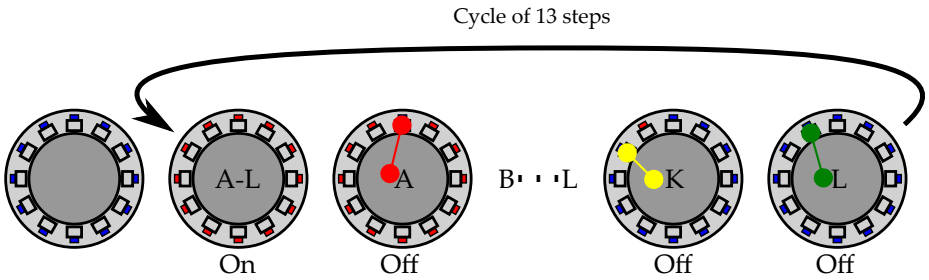


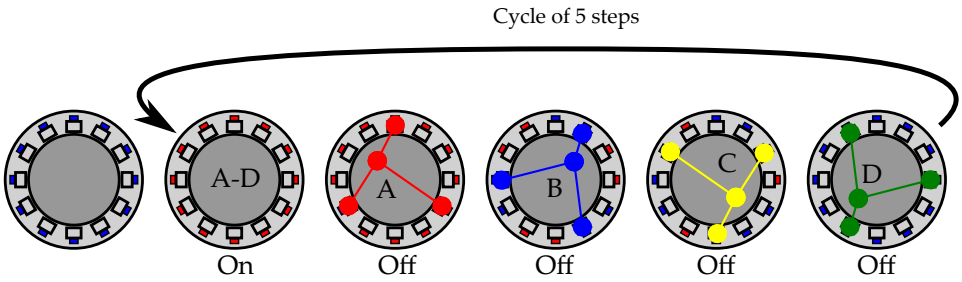
Figure 4.7.: Simulated temperatures and displacements of a simplified 3 finger TSA model. Whereas the simulation shown in Figure 4.3 only uses a static friction coefficient, in this simulation the friction is modelled using a static and a (30 % lower) dynamic friction coefficient. Also a shear-force graph is added to show the variations in the internal forces. The different lines in the three top graphs (shown in a different colour and with a different number) indicate the different fingers of the TSA. Stick-slip steps are clearly visible in this simulation. They were however never observed in practice, which in turn triggered the research described in Chapter 5. The model that is used is described in more detail in Appendix B.



(a) Thermal Slider Actuator: the number of fingers, sections and groups are: $n_f = 12$, $n_s = 3$ and $n_g = 4$.



(b) Grouped in 1 section and 12 groups (A to L): $n_s = 1$ and $n_f = n_g = 12$.



(c) Grouped in 3 section and 4 groups (A to D): $n_f = 12$, $n_s = 3$ and $n_g = 4$.

Figure 4.8.: Examples of grouping of the TSA fingers for translational actuation, shown on the top view of a 12 finger actuator. The coloured circles indicate which fingers are switched in a specific cycle.

4.2.1. Grouping definition

The thermal cycle for a translational motion consists of two phases, which are for the upward (downward) direction: simultaneous heating (cooling) of all fingers and subsequent individual cooling (heating) of these fingers. The number of steps in the individual motion part depends both on the number of fingers in the TSA and on the chosen finger grouping. Grouping is already presented in [Schieveen et al., 2011] where it is shown that grouping influences the actuator efficiency. Grouping implies that the entire thermal cycle is performed simultaneously on different sections of the actuator. The actuator is subdivided in n_s sections of adjacent fingers, as is shown in Figure 4.8a. A TSA cycle is performed on each section simultaneously. All fingers in one section belong to a different group and there are n_g groups. Therefore, for a TSA with n_f fingers it holds that:

$$n_f = n_s \cdot n_g. \tag{4.1}$$

Obviously all terms in (4.1) must be integer. The fingers that belong to the same group, indicated in Figure 4.8, are switched simultaneously. A thermal cycle then consists of first switching all fingers (of all groups) simultaneously and then switching them group by group. The thermal cycle therefore consists of $n_g + 1$ steps. Figures 4.8b and 4.8c illustrate the thermal cycles for two different section and group sizes.

4.2.2. Analytical approximation

The basic translational motion can be approximated by an analytical expression. This approximation assumes that the fingers are the only expanding elements, that the fingers only heat up due to their own heating resistor, that the fingers reach thermal equilibrium, that the thermal expansion of the fingers is sufficient to reach slip and that the frictional behaviour can be approximated by an ideal and constant static friction force. In most practical situations these conditions cannot be met entirely, which means that this model cannot be used to make accurate predictions about the exact actuator displacement. This model does however provide insight in the working principle of the TSA and shows the influence of grouping on the performance.

Common motion Under the aforementioned conditions, the displacement due to the common movement of all fingers (Δx_{com}) equals the thermal expansion of the fingers (Δx_0), as the forces on the fingers do not change:

$$\Delta x_{com} = \Delta x_0 = \Delta T \alpha_{fin} l_{fin}, \tag{4.2}$$

where α_{fin} and l_{fin} are the thermal expansion coefficient and length of a finger respectively and ΔT is the finger temperature change.

Minimum expansion Before discussing the motion of the mover in the second phase, it is important to be aware of the minimum thermal expansion Δx_0 that is required to complete a successful thermal cycle. The criterion is that each finger

must, when actuated, reach the slip-phase and move. The amount of thermal expansion Δx_0 that is required for this depends on the friction shear force prior to the actuation. As the TSA structure is by definition over-constrained, the magnitude of the maximum shear force that can be present is the friction limit: $F_{\mathrm{f,max}} = F_{\mathrm{N}}\mu_{\mathrm{s}}$, where F_{N} is the contact normal force and μ_{s} is the static friction coefficient. The force change that is required to change the shear force from $F_{\mathrm{f,max}}$ to $-F_{\mathrm{f,max}}$ or vice versa, equals

$$\Delta F_{\mathrm{f,max}} = 2F_{\mathrm{f,max}} = 2F_{\mathrm{N}}\mu_{\mathrm{s}}, \tag{4.3}$$

which is the maximum shear force change that could occur. Simulation results confirm that $\Delta F_{\mathrm{f,max}}$ is not only the maximum force change, it is thé force change that occurs during the second phase of the TSA motion when a cycle is repeatedly executed. This effect is shown in the simulation results of Figure 4.9.

The neutral length change of a finger Δx_0 that is required to generate the force change ΔF in the actuated fingers equals:

$$\Delta x_{0,\Delta F} = \frac{\Delta F}{k_{\mathrm{fin}}} \cdot \frac{n_{\mathrm{g}}}{n_{\mathrm{g}} - 1}, \tag{4.4}$$

where k_{fin} is here the longitudinal stiffness of a single finger. The last term in (4.4) takes the elastic motion of the passive fingers due to force of the actuated finger into account. By choosing $\Delta F = \Delta F_{\mathrm{f,max}}$ it follows that the lower limit for Δx_0, so that the TSA will generate a net movement, equals

$$\Delta x_{0,\mathrm{min}} = \frac{2F_{\mathrm{N}}\mu_{\mathrm{s}}}{k_{\mathrm{fin}}} \cdot \frac{n_{\mathrm{g}}}{n_{\mathrm{g}} - 1}. \tag{4.5}$$

Individual motion With the lower limit of the thermal expansion defined, the second phase of the TSA motion cycle can be analysed. In the second phase, the fingers of each group move back subsequently. With each group, also the mover moves back a distance of

$$\Delta x_{\mathrm{ind}} = \frac{\Delta F_{\mathrm{f,max}} n_{\mathrm{s}}}{k_{\mathrm{fin}}(n_{\mathrm{f}} - n_{\mathrm{s}})} \tag{4.6}$$

$$= \frac{2F_{\mathrm{N}}\mu_{\mathrm{s}}}{k_{\mathrm{fin}}\big(n_{\mathrm{g}} - 1\big)}, \tag{4.7}$$

The term $\Delta F_{\mathrm{f,max}} n_{\mathrm{s}}$ (which equals $2F_{\mathrm{N}}\mu_{\mathrm{s}} n_{\mathrm{s}}$) is the total contact shear force change on the mover due to n_{s} fingers simultaneously changing between plus and minus the friction limit. The term in the denominator of equation (4.6), $k_{\mathrm{fin}}(n_{\mathrm{f}} - n_{\mathrm{s}})$, is the mechanical stiffness of the fingers of which the temperature remains constant. As both the numerator and the denominator scale proportional to the number of sections, (4.6) is only a function of the number of groups n_{g} and can be simplified as is shown in (4.7). Note that both the common and the individual motion were already indicated in Figure 4.3.

The first order approximation of the net step Δx_{net} follows then from subtracting

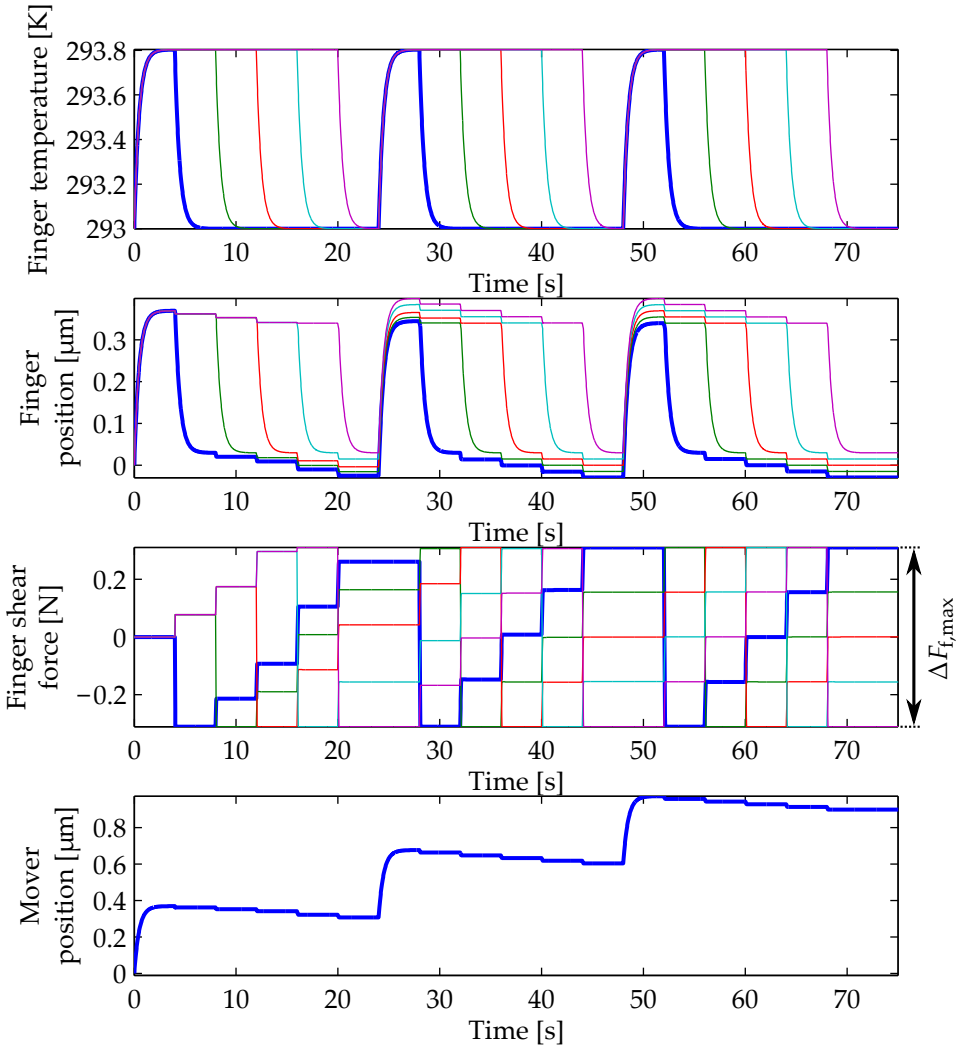


Figure 4.9.: Simulation of three TSA cycles on a 10 finger actuator. It started at zero contact shear force, as is show in the third graph at time equals 0. After the second iteration (at 45 s) the forces follow a stationary pattern, where the contact shear forces are always equally distributed between $-F_{f,max}$ and $F_{f,max}$, where is in this case $\Delta F_{f,max} = 2F_{f,max} = 0.6$ N. For clarity the temperature, shear force and position of the first finger is indicated with a thicker line.

the individual displacement (4.2) n_g times from the common displacement (4.6):

$$\begin{aligned} \Delta x_{net} &= \Delta x_{com} - n_g \Delta x_{ind} \\ &= \Delta x_0 - n_g \frac{2F_N \mu_s}{k_{fin}(n_g - 1)} \\ &= \Delta T \alpha_{fin} l_{fin} - \frac{2F_N \mu_s}{k_{fin}\left(1 - \frac{1}{n_g}\right)}. \end{aligned} \tag{4.8}$$

Using this step-size approximation, also the average motion velocity can be found by dividing (4.8) by the total cycle time $t_{step}(n_g + 1)$. The resulting displacement and average velocity are shown for all translational grouping options on a 24 finger TSA in Figure 4.10. For the parameters chosen, actuation in more groups leads to a smaller net step but, also to a larger velocity. Figure 4.11 shows the same relationship where only the neutral length change Δx_0 was chosen to be 0.4 µm instead of 1.0 µm. This results in a lower step efficiency (defined as the ratio of Δx_{net} and Δx_0). Also for a larger number of sections the relative cycle time decreases less than the efficiency, causing a reduction of the velocity.

4.3. Experimental set-ups

Measurements on the properties of the TSA in 3 DoF positioning are performed on several TSA measurement set-ups. These set-ups contain the following actuators:

- a 16 finger TSA,
- a 12 finger TSA,
- an ASML/PI TSA and
- a simplified heat-sink TSA.

Additionally, all actuators have electronic circuits to control the heaters. These set-ups are briefly described in this section. More information on the displacement measurement set-ups that have been used can be found in Appendix C.

The 16 finger TSA The TSA that is used in most experiments was already shown in Figure 4.5. It consists of a monolithic aluminium anodised body with a 26 mm diameter and a 31 mm height. The 16 fingers have an effective length of approximately 20 mm, a width of 2.7 mm and a thickness of 0.8 mm. The fingers have a tip with an additional thickness of 1.2 mm to clamp a mover of ⌀22 mm. The normal force is approximately 0.3 N. Time constants of 8 s and 90 s have been identified on this actuator while it was mounted to a large steel block. Experimental results with this actuator are shown in Figures 4.16, 4.19, 4.20, 4.21, 4.22 and 4.26 starting on page 57.

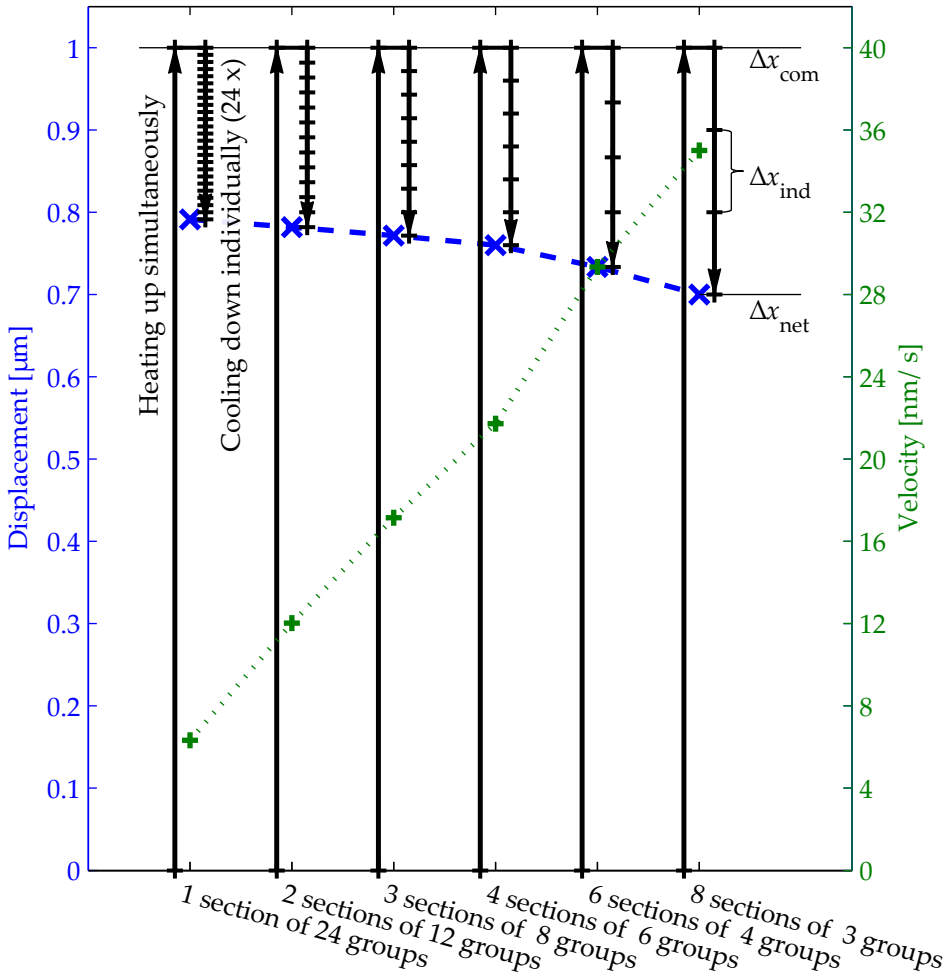


Figure 4.10.: Analytic approximation of grouping of a TSA with 24 fingers, where $F_{f,max} = F_N \mu_s = 1$ N, $k_{fin} = 10^7$ N/m and $\Delta x_0 = 1$ µm (see (4.2)). The different grouping options are indicated on the horizontal axis. Each set of arrows in the graph indicates the motion of the mover of the TSA during one upward cycle. The displacement starts at zero and increases to 1 µm, according to Equation (4.2). Then, when individual groups of fingers cool down the mover moves along as is indicated by the downward arrow. The number of markings in the downward arrow and the magnitude of these steps is dependent on the number of groups that is used, as is described in Equation (4.6). The net displacement is indicated with blue crosses at the end of an arrow. Using less groups leads to a larger net displacement, but because each sub-step takes the same time, a smaller number of sub-steps leads to an increased motion velocity, as is indicated with the green + signs and on the right axis. The displacements that were indicated in Figures 4.3 and (4.8) are also shown at the arrows on the right.

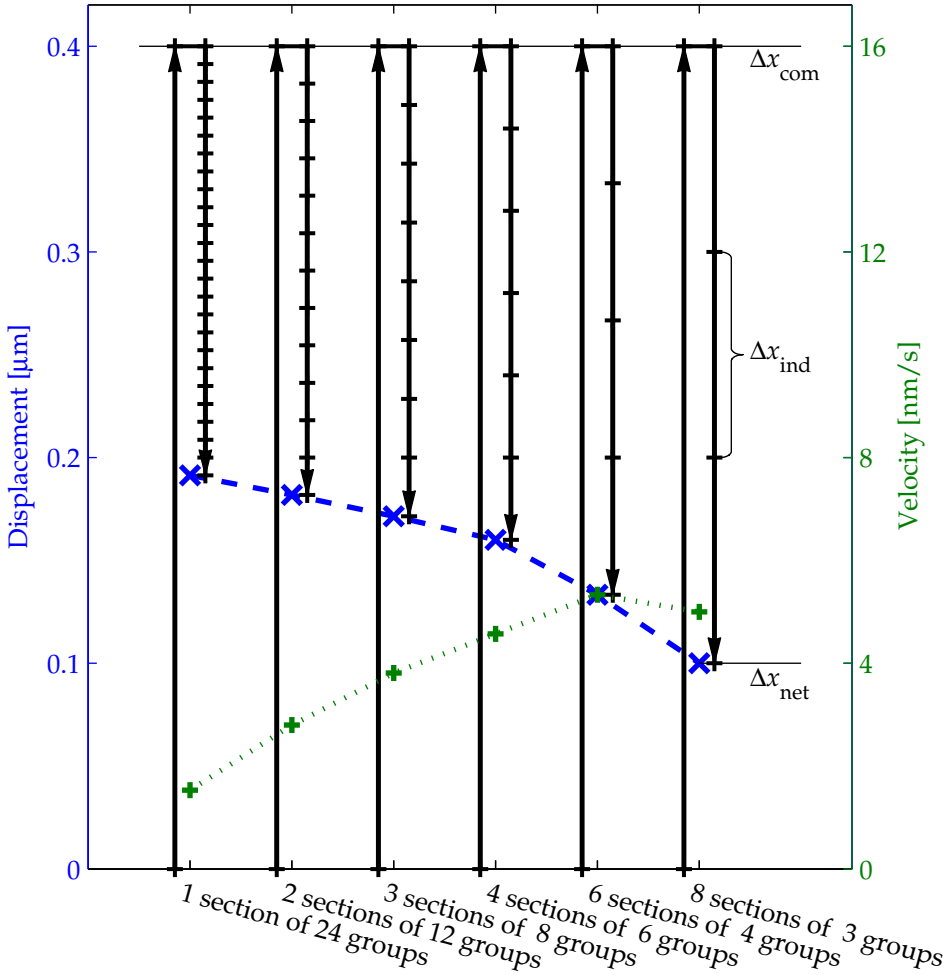
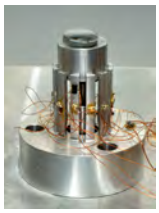


Figure 4.11.: Analytic approximation of grouping of a TSA as was shown in Figure 4.10 (see that figure for a detailed explanation). The TSA also has 24 fingers, $F_{f,max} = F_N \mu_s = 1$ N and $k_{fin} = 10^7$ N/m , but $\Delta x_0 = 0.4\,\mu$m instead of 1.0 µm as the input temperature is 2.5 times smaller. Using more sections still leads to a smaller displacement per step, but in contrast to what is shown in Figure 4.10, using more sections does not necessarily result in a larger average velocity.

The 12 finger TSA Also the earlier actuator shown in Figure 4.4b is used in several experiments as it contains temperature sensors. The fingers of this TSA can however only be controlled in pairs and the actuator is therefore not suitable for tilt actuation (tilt actuation will be discussed in more detail in Section 4.4). It consists of 12 fingers with an effective length of 20 mm, an electrode diameter of ⌀15 mm and a contact normal force of approximately 1.3 N. Time constants of 8 s and 130 s have been identified when mounted on a large aluminium plate. Experimental results with this actuator can be found in Appendix D.



ASML/PI TSA ASML initiated the development of an industrial version of the TSA based on the earlier research of Van Schieveen. Physik Instrumente (PI) designed and manufactured this TSA that is even more compact, and includes suitable mounting features for its use with a capacitive displacement sensor. This actuator is shown in Figure 4.12 and consists of 12 fingers with an effective length of 15 mm, around a ⌀12 mm diameter mover. It is also analysed in several experiments. These experimental results are shown in Figures 4.28, 4.29 and 4.30 starting on page 78.

Simplified heat-sink TSA One important strength of the thermal slider concept is its simple (mechanical) design. Although current demonstrators have a relatively simple design, some steps can be made to further optimise or simplify the design and thus to decrease the system cost. This is done for the actuator that is depicted in Figure 4.13. Some of these changes might compromise system properties such as mechanical stability, it does however show the robustness of the positioning concept with respect to manufacturing processes and tolerances. The expected large volume production cost of this TSA are below €20. Experimental results with this actuator are shown in Figure 4.31 on page 82.

(a) The ASML/PI TSA mounted in its housing. This housing acts not only as a means to mount the sensor and the actuator, but it also reduces the displacement sensitivity to environmental temperature changes, by placing the electrode in the thermal centre. (The housing and the fingers expand equally, reducing the electrode motion due to environmental temperature changes.)

(b) The ASML/PI TSA with the partial flexible PCB and heaters exposed. The actuator is mounted to an aluminium block that acts as a heat sink. The experiments discussed in this thesis use this mounting method to achieve better thermal conduction to the environment.

Figure 4.12.: Pictures of the 12 finger ASML/PI TSA, manufactured by Physik Instrumente.

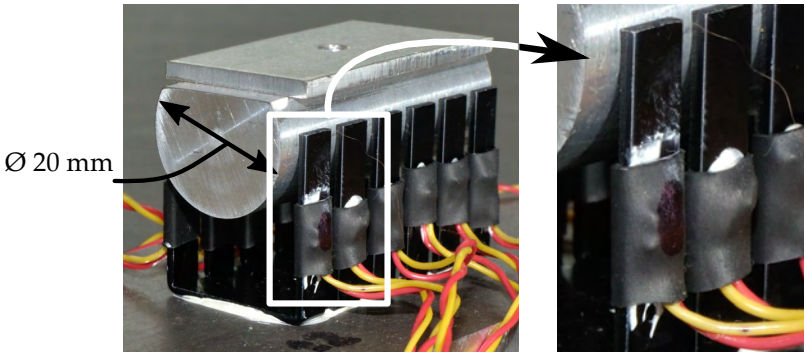


Figure 4.13.: Simplified thermal slider design using a standard heat-sink (an Aavid Thermalloy 504222B00000G). This aluminium heat-sink was selected such that the thermal crosstalk between the different fingers is small. The heating resistors are applied to the finger body using thermally conductive paste and heat-shrinkable tube. The heaters are controlled using an Arduino micro-controller board and a 16 channel LED driver.

Figure 4.14.: Axis definitions of the TSA in three degrees of freedom: tilt angles x and y and translation axis z.

Control electronics Different options are available for controlling the heater power of the individual fingers. The switching frequency is in the sub-Hz range, which allows the application of low-cost electronic components. The thermal process is stable and its properties are sufficiently insensitive to the operating conditions so that they can be controlled in feed-forward. In [Yang et al., 2010b] a flexible PCB is described that is mounted on the inside of the fingers and contains the heating resistors, all the wiring, a power-switch and a microprocessor that controls the switches in pre-defined patterns and communicates with an external controller. By integrating the electronics in the TSA body, a single unit is formed that only requires a power and a data line. It can therefore be controlled from the USB port of a laptop only. Alternatively, the heating resistors can be wired and mounted to the fingers separately, as is shown on the heat-sink TSA in Figure 4.13. The switch and micro-controller are then placed externally, giving more freedom in the selection of the input cycles. In both cases the resistors are glued directly to the anodised aluminium body.

4.4. Three DoF positioning

As was shown in the foregoing, the TSA induces a permanent displacement to the mover by applying thermal cycles to the fingers of the actuator. The basic cycle is explained in Section 4.1.1 on a single dimensional actuator, however, on a three-dimensional actuator many cycle variations lead to displacements in one translational and two tilt directions. These directions are indicated in Figure 4.14. An example of the translational motion of a TSA is given in Figure 4.15. In this section, the parameters that influence the motion in translational and tilt degrees of freedom are investigated by means of both modelling and experiments.

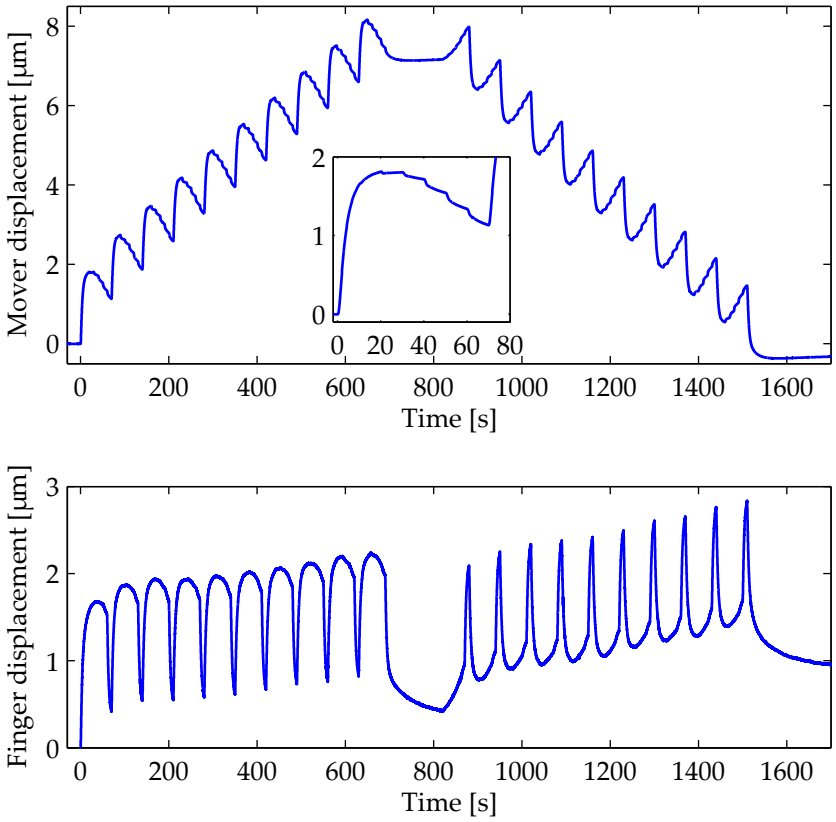


Figure 4.15.: Measurement of the displacement of the mover and one of the 12 fingers with respect to the base of the TSA during 10 upward and 10 downward cycles. The inset graph shows a detail of the mover displacement of the first upward cycle. A comparison of this measurement with a numerical model of this TSA can be found in Appendix D.

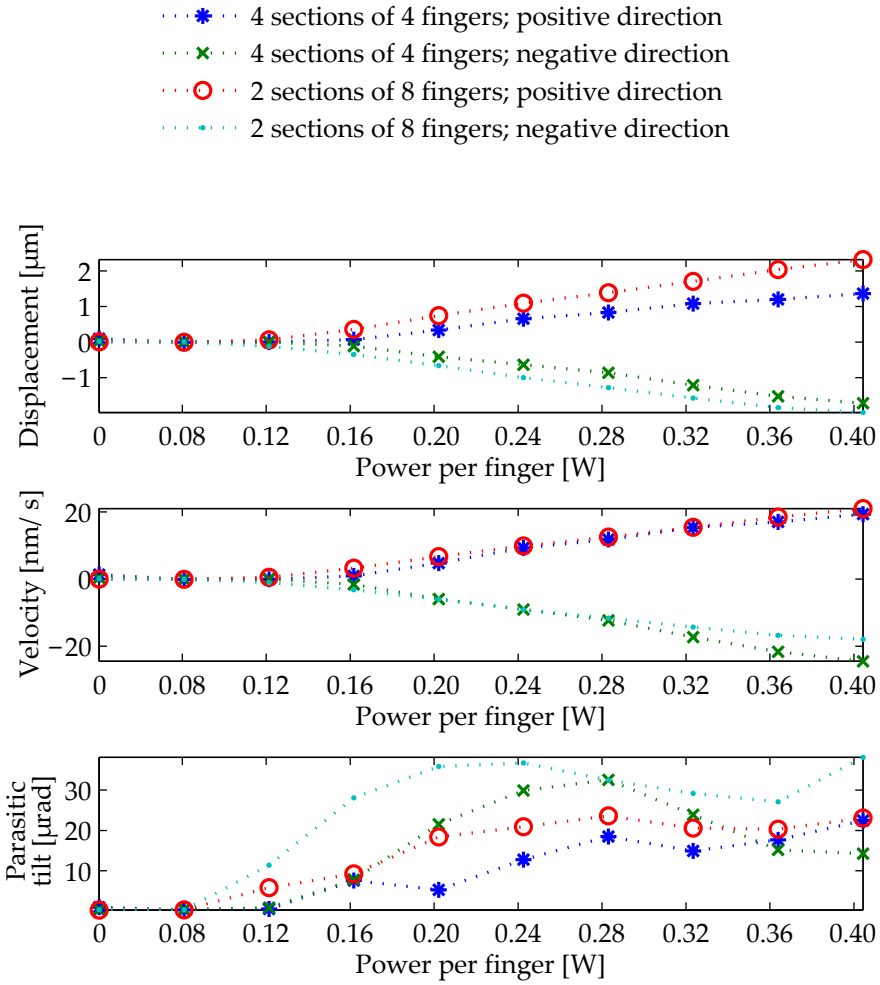


Figure 4.16.: Measured influence of the heater input power and the grouping of the fingers on the displacement efficiency and the parasitic tilt motion. The indicated value is the median of at least 5 repetitions. The standard deviation (σ) of a displacement is typically below 10 % when the power level per finger is above 0.2 W. The positive thermal cycle heats up all fingers simultaneously and cools them down subsequently. For the negative direction the fingers are heated up subsequently and then cooled down simultaneously. The fingers are operated either in sections of 2 or 4 fingers. Grouping is explained in Section 4.2.2 and Figure 4.8.

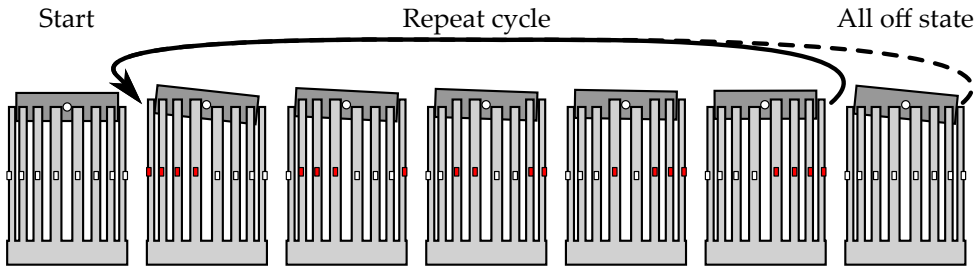


Figure 4.17.: One full tilt cycle that uses all fingers on a 16 finger TSA. The active heaters, that elongate their fingers, are indicated in red. The solid arrow indicates the fastest tilt cycle. The dashed arrow indicates a cycle that includes a state where all fingers are cold. This can be used to easily measure the current tilt angle for control purposes.

4.4.1. Translational grouping

The theoretical influence of grouping on the translational motion cycles of the TSA was extensively discussed by means of an analytical model in Section 4.2. Also [Schieveen et al., 2011] already showed experimentally how grouping influences the motion efficiency. Experiments have shown that groups of at least four fingers should be used to obtain a reliable motion. On a 16 finger TSA, using 2 sections of 8 finger groups increases the step size more than a factor 2, but even larger groups lead to a less efficient motion.

Comparison of the displacement and the velocity of the 2 and the 4 section experiments in Figure 4.16, confirms that changing the number of sections that is used for translational actuation, significantly influences both the step-size and cycle-time. The differences in the average motion velocities are significantly smaller. More specifically, using 2 sections and 8 groups (2 fingers are switched simultaneously 8 times) results in larger steps than 4 sections of 4 (4 fingers are switched simultaneously 4 times)[1]. The cycle for translation in 8 groups does however need 9 sub-steps and thus takes more time than translation using 4 groups, which needs 5 sub-steps. The time that one sub-step requires depends on the thermal time constant and therefore remains the same for both grouping options. The resulting average motion velocities, shown in the second graph, are therefore of comparable magnitude. It becomes also apparent from this graph that there are significant differences between the positive and the negative motion direction. This effect can be explained by the asymmetric shape of the fingers that causes a different friction behaviour during the upwards and the downwards motion and by the thermal effects discussed in Section 4.4.5.

[1] Recall that the definitions of sections and groups can be found in Figure 4.8 on page 46.

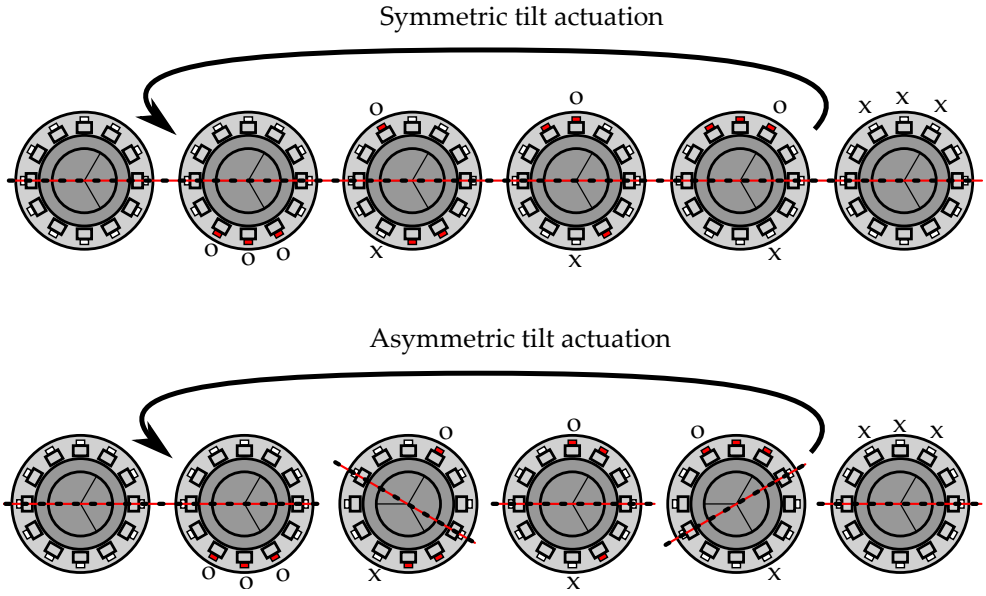


Figure 4.18.: Symmetric and asymmetric tilt actuation. The active heaters are indicated in red and at each cycle step an 'o' indicates switching on, an 'x' indicates switching off. The asymmetric tilt cycle results in parasitic tilt motion (in the other tilt direction) during the cycle, as the fingers that are moving up and downward (heating and cooling respectively) apply a friction torque with respect to the indicated symmetry axes (red and dashed lines).

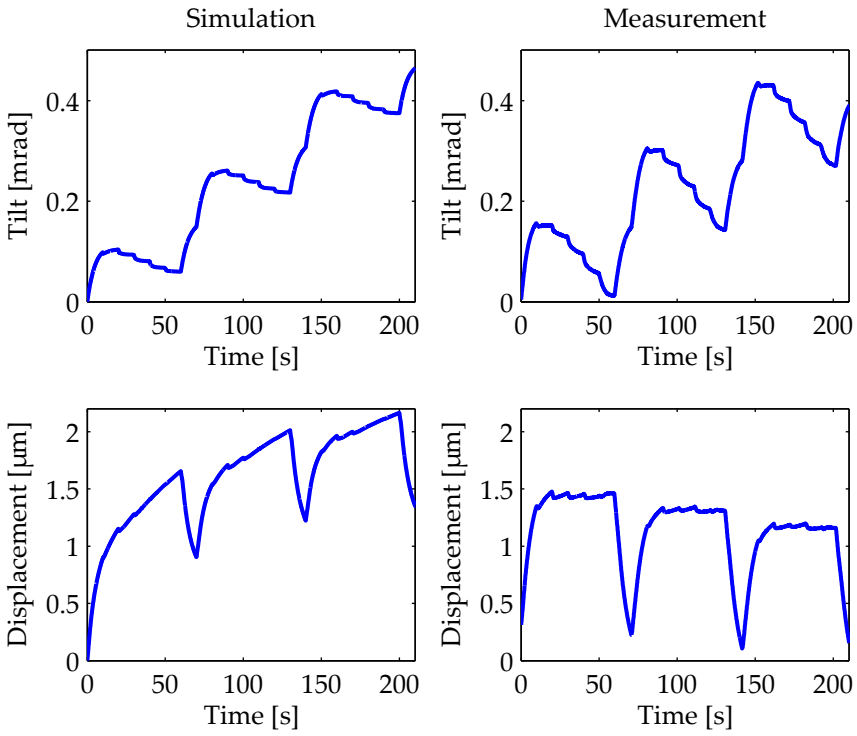


Figure 4.19.: Simulation (left graphs) and measurement (right graphs) results of tilt actuation on TSA with 16 fingers using 2 times 5 fingers for tilt actuation. Positions in micrometres and tilt angles in milliradians. A tilt cycle is performed three times with an all-off state in between (the cycle indicated with a dashed arrow in Figure 4.17). This effect is clearly visible in the dip in the bottom (displacement) graphs. The top (tilt) graphs show a typical tilt motion. The (parasitic) motion in the displacement direction (bottom graphs) can be explained by overall heating of the system (simulation) and by actuator non-uniformities that also cause remanent translations (measurement). This measurement is performed on the 16 finger TSA (Figure 4.5) using the 3-DoF capacitive measurement set up described in Appendix C. A comparison of the measured TSA translational motion with this model can be found in Appendix D.

4.4.2. Tilt grouping

The translation cycle that is described and analysed earlier in this chapter also forms the basis for the cycles that generate a tilt motion. Tilt is here defined as a rotational motion around the axes orthogonal to the translation axis as is shown in Figure 4.14. The most straightforward way to actuate a tilt motion is to simultaneously apply an upwards thermal cycle on one half of the stepper and a downwards cycle on the other half. A few things however, have to be noted. First of all, the fingers that are close to tilt axis (the white circle on the mover in Figure 4.17 and the red/dashed line in Figure 4.18) contribute a relatively small amount to the friction torque around the tilt axis and thus to the total tilt motion. The amount of parasitic motion that is caused by these extra steps is however not necessarily smaller. This is shown later in this section. Because including these fingers as steps in the tilt cycle does also result in a larger number of steps, the tilt-cycle will become more efficient when one or more of these fingers are excluded from the cycle.

Also the order in which the fingers are switched in the step cycle can be relevant. As an example the difference between an asymmetric cycle and a symmetric cycle is shown in Figure 4.18. The asymmetric cycle causes small parasitic motion during a tilt cycle because the tilt axes of the sub-steps are not aligned. This is not the case for the symmetric cycle. This behaviour is verified using the numerical model that is described in Appendix B. The resulting motion of a symmetric cycle is shown in Figure 4.19.

Tilt grouping parameters In the experimental results shown in Figure 4.20, a number of parameters influencing the tilt-cycle is investigated. The first parameter is the number of fingers that is used for tilt actuation. On the 16 fingers TSA that is used in this experiment, 4 to 8 fingers were used to perform a thermal cycle on each side. When the box-plots of identical colours are compared, it becomes clear that using less than 5 fingers leads to unpredictable motions, as the deviations in both motion magnitude (top graph) and directional accuracy (bottom graph) are relatively large. The top graph shows that the step-size when using 6, 7 or 8 fingers does not vary significantly as a function of the number of fingers per tilt-section. This is explained by the aforementioned fact that fingers close to the axis of rotation contribute relatively little to the tilt motion. The time consumed by cycles using less fingers is smaller and therefore the motion speed of these cycles can be larger, as is shown in the second graph. Using a low number of fingers does however come at a cost: the relative variation in the step-size is slightly larger when a small number of fingers is actuated.

Tilt direction variations Figure 4.21 provides a more detailed picture of the tilt motions. The basic tilt cycle that is described before, can be performed in 16 different directions on the 16 finger TSA, by applying the same cycle to different fingers. The picture shows the tilt angles during a series of three tilt steps in all these 16 directions, each direction is indicated with a single colour. The circles indicate the endpoint of a cycle and the line is the tilt-path in the two rotational degrees of freedom. A tilt

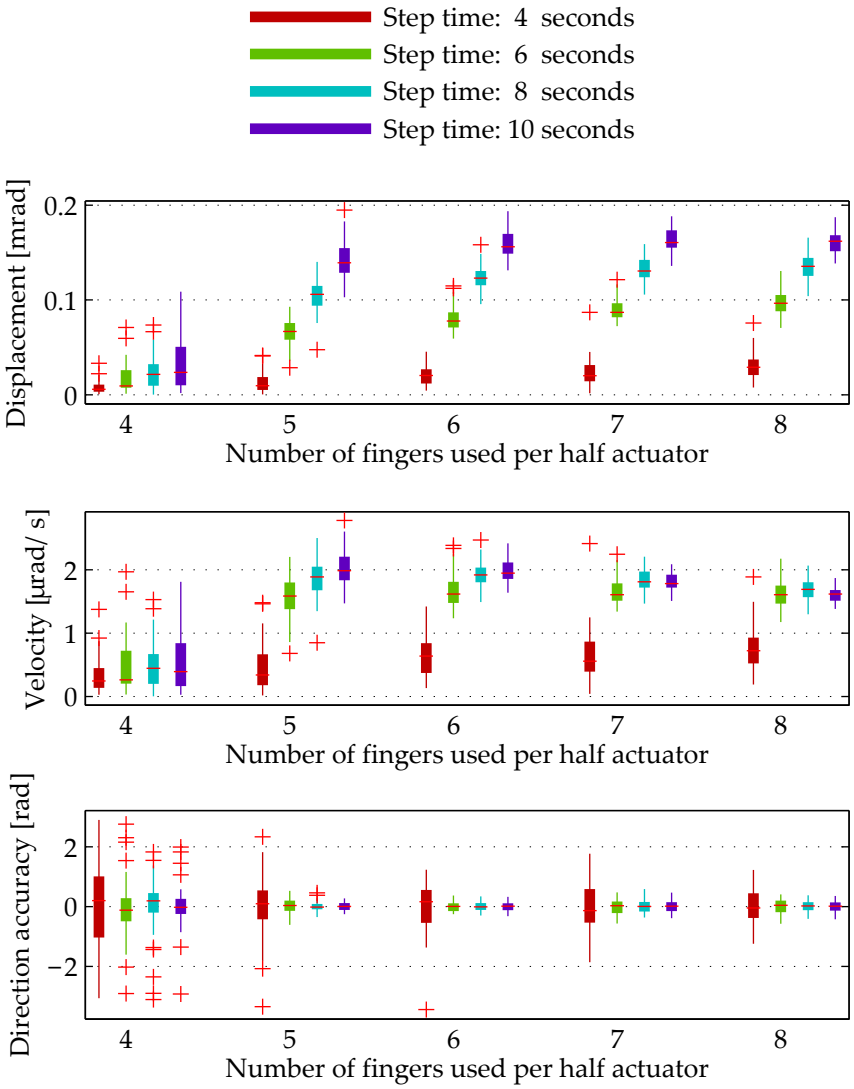


Figure 4.20.: Measurement results that show the influence of the number of fingers used for a tilt motion and the time steps on the motion step size and speed and the direction accuracy on a 16 finger TSA. The number of fingers mentioned in the diagram is the number of fingers that is used in a tilt-section each half of the 16-finger thermal slider actuator (Figure 4.5). Multiple tilt-cycles in 16 tilt-directions are taken into account. A 4 seconds step time or 4 step fingers lead to a relatively large variation on the tilt step size and motion direction. Longer step times and the use of more fingers do increase the step size, but not necessarily the average velocity. This measurement was performed using the PSD measurement set-up, as described in Appendix C.

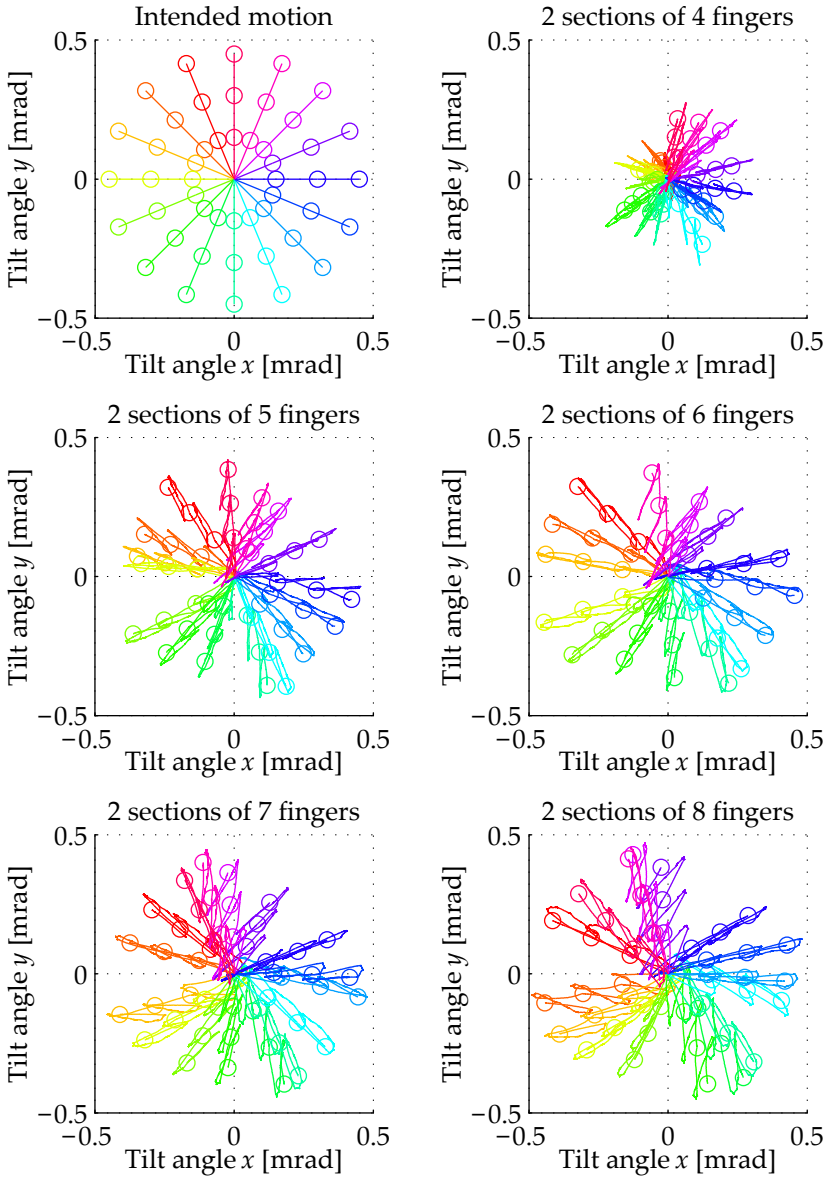


Figure 4.21.: The intended tilt displacement (top-left graph) of the 16 finger TSA compared to those measured using the capacitive 3 DoF measurement set-up. The number of fingers that was used for tilt actuation was varied in the measurements, so that the influence on the tilt motion is shown. As was already shown in Figure 4.20, only significant tilt displacements occur when more than four fingers are used. Also, when the measured motions are compared to the ideal directions, both deviations of the end-points (the circles) and within the motion (the line between the circles) become apparent.

cycle that includes an all-off state (indicated with the dashed arrow in Figure 4.17) was used to show the tilt angle after each cycle.

It can be deduced from the differences between the measured tilt paths and the intended motion directions (shown in the top left graph), and from the variations in the locations of the end-points (the circles), that significant variations are present in the properties of the different fingers. These variations can originate from effects such as effective thermal expansion, normal force and effective friction coefficient. It can however be concluded that the motion paths reproduce within less than 0.1 mrad, and the asymmetry in the tilt motions does not seem to limit the tilt angle that can be reached. Additionally, using more that 5 fingers in a tilt cycle (that is: 5 fingers on each side of the 16 finger TSA perform a translational sequence in opposite direction) does slightly increase the tilt-step-size, but also increases the time required.

4.4.3. Step time

The second variable that is studied in the experiment shown in Figure 4.20 is the heating time for each sub step in the cycle, indicated with colours. Using longer heating times increases the step size significantly (this will continue until the step time is large compared to the relevant thermal time-constants, in this case 8 seconds), however the benefit in motion velocity, shown in the second graph, is less significant.

In the most elementary description, the fingers of the TSA heat up exponentially with a single time constant, as was shown in the simplified simulation at the beginning of this chapter (see Figure 4.3). In reality however, also larger time constants play a role in the thermal expansion of a single finger. This effect can be observed in the measurement of the displacement of one of the fingers of a TSA during several cycles in Figure 4.15: The displacement of the finger is gradually increasing. These larger time constants generally consist of overall heating of all fingers, which does not generate a net displacement of the mover. However, they will contribute to the temperature, and thus to the displacement of an individual fingers. The two most dominant time constants that could be identified on the 16 finger TSA (introduced in Figure 4.5) are 8 and 90 seconds, when it is mounted to a steel block that is large compared to the TSA base. When all fingers are heated simultaneously, these two time constants occur approximately in a 2:1 magnitude ratio.

4.4.4. Heating power

Using the capacitive measurement set-up (Appendix C) an investigation has been done on the motion properties of the thermal slider in the different degrees of freedom. Due to the nature of the friction process between the fingers and the mover, a motion gradually starts when the tangential contact force crosses a threshold value. Therefore there is also a threshold value in the input power needed to initiate a net motion[2]. In Figure 4.16 the results of a series of measurements were shown in which the input power and the type of translational heating sequence is varied.

[2]This also followed from the theoretical analysis of grouping in Section 4.2 that started on page 44.

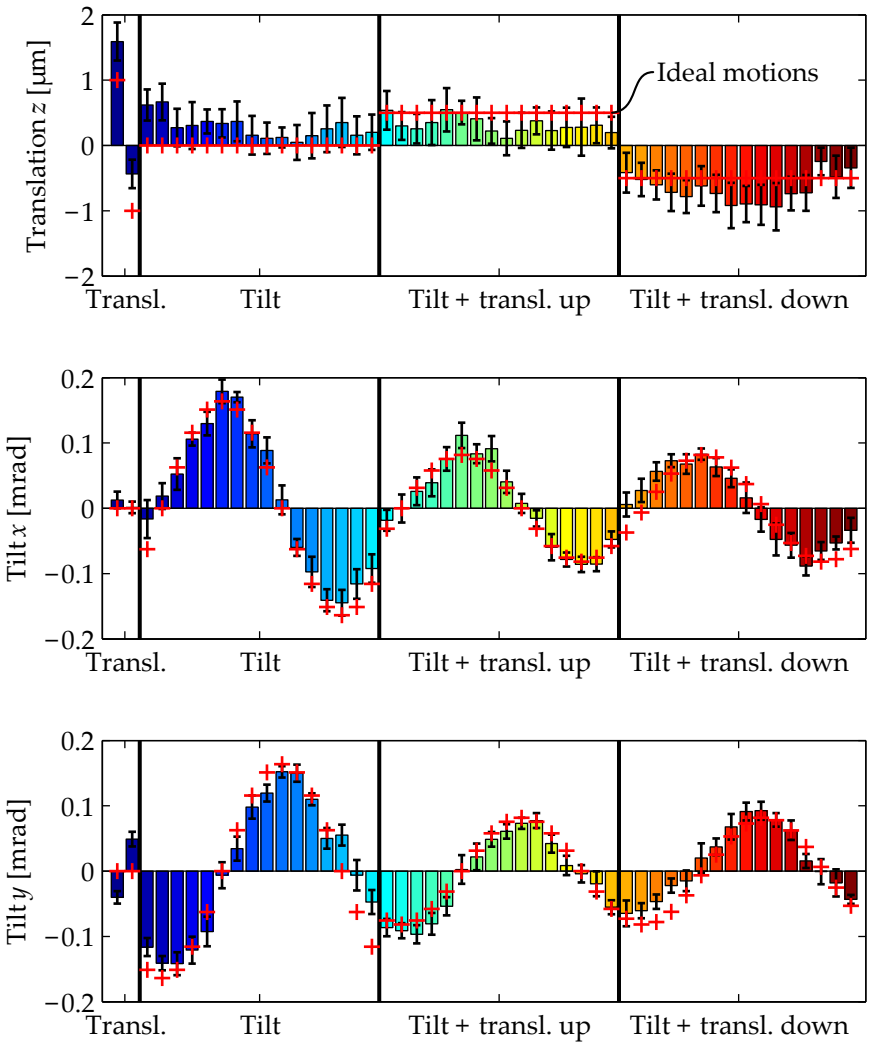


Figure 4.22.: Measured (parasitic) motion and repeatability of the 16 finger TSA in 50 selected directions. A bar in a colour represents the mean displacement due to a step-cycle (vertical axes). The cycles are grouped (horizontal axes) according to their intended motion: Translation, Tilt, Tilt and translation upward and Tilt and translation downward. The three different graphs indicate the three displacement components, one translation and two tilt directions. The black bars indicate the standard deviation σ. The directions of x, y, and z are defined in Figure 4.14. The cycles shown here use a time-step dt of 10 seconds and tilt sections of 7 fingers. The red plus signs (+) indicate the motions that are expected based on an idealised model. The cycles were performed in a random order and 0.24 W was supplied per finger.

It becomes clear from the top graph that there is indeed a power threshold above which significant translational motion occurs. This is in agreement to what is shown in [Schieveen et al., 2011]. Above this threshold the net displacement increases with the power. The heating power also influences the motion repeatability, as the relative influence of the friction contact behaviour becomes smaller when the heating power, and with that the thermal expansion, increases.

4.4.5. Parasitic motion

It was already mentioned in Section 4.4.2 that including fingers close to the tilt axis in a tilt-cycle, only leads to a negligible increase of the motion in the intended direction, but not in a reduction of the parasitic motion in the other tilt direction. A more detailed analysis of the parasitic motions of the 16 finger TSA is shown in Figure 4.22. In these graphs the resulting displacements of a selection of 50 different thermal cycles are shown. Each bar represents the mean displacement of a cycle. Each graph shows a different degree of freedom and each column and colour represents the same cycle. The first two cycles are translational, using 4 sections of 4 fingers. The next group of 16 cycles are tilt cycles that use two sections of 7 fingers on both sides of the tilt axis. The remaining two groups of 16 cycles use one section of 7 fingers on one side of the tilt axis to perform an upward or a downward cycle. This results in a tilt motion combined with an upward and a downward translation. The most significant observations of this analysis are described below.

Translational parasitics The translation magnitude of the translation cycles is significantly larger in the positive direction than in the negative direction and also the pure tilt cycles all result in a net upward motion. One of the causes of this preferred upward motion is the shape of the end of the fingers, which makes the contact normal force change up to 10 % depending on the direction. Additionally, the average heating power influences the overall temperature of the actuator, which can lead to an apparent offset in the translational direction.

The average heating power of the translational, tilt and combined cycles is 2.3 W, 1.8 W and 0.9 W respectively. When a high power cycle is preceded by a low power cycle, the average temperature of the actuator increases, which leads to a positive translational displacement. This displacement can then incorrectly be observed as a permanent displacement due to the step cycle. The average thermal displacement due to the average heating power of the previous cycle is 0.45 $\mu m/W$. This suffices to explain the translational offset shown in the top graph op Figure 4.22, given that in this experiment the cycles were performed in a random order: The expected offsets are 0.48 µm for the translations, 0.25 µm for the tilt motions and –0.15 µm for the combined motion. Using the aforementioned relationship, the mover displacement due to overall heating can be estimated from the average power of the last cycles. This can benefit the positioning performance in two ways. Firstly, the estimated displacements due to a cycle, which will be used for positioning (this will be discussed in Section 4.5), can be compensated for this effect to improve the accuracy of the estimated position after the next cycle. Secondly, a more accurate estimate can be made of the mover position after the TSA is entirely cooled down.

Tilt parasitics The tilt cycles should result in tilt motions in tilt directions evenly spaced in all directions in the $x-y$ plane. This is approximately the case as can be seen when the estimated motions (indicated by the red crosses in Figure 4.22) are compared to the measurements. The combined translation and tilt cycles have tilt displacements comparable to the full tilt cycles, at approximately half the tilt amplitude. Also the directions shown in Figure 4.21 provide a good impression of the tilt directions. This measurement also shows that when the tilt directions deviate significantly, parasitic motions are visible during the cycle. Ideally the paths (lines) are straight between the cycle end points (circles). The deviations from the straight lines indicate that not all fingers behave identically. These deviations can be explained by the contact normal force variations between fingers. For this actuator type, typically the force deviations within an actuator are around 15 % (see also Section 4.6.3). The translational motion of the combined cycles is correctly aimed in the positive and the negative direction respectively, including the translation deviations that were discussed before.

4.4.6. Repeatability

The $\overline{|}$ on top of each bar in Figure 4.22 indicates the standard deviation (σ) on the motion in that direction. The motion cycles were, in this experiment, selected randomly so that the initial conditions of each step vary in a manner that is comparable to closed loop control. Figure 4.9 showed that even in an idealised model, several cycle iterations are needed before an identical thermal cycle results in an identical response. Varying the initial conditions can therefore have a significant impact on the variation of displacement due to a thermal cycle. The first two cycles in Figure 4.22 (translation up and down) show a standard deviation of 0.3 µm and 0.2 µm respectively, including the random initial conditions. Approximately 40 % of the standard deviation in the translational direction can be attributed to the parasitic motion due to the power variation of the preceding thermal cycles (Section 4.4.5). In comparison, the standard deviation of the same two cycles repeatedly executed equals 0.15 µm and 0.06 µm respectively when the first instance is excluded. Also after the first instance the displacement, including its parasitics, keeps converging as is shown in Figure 4.23.

Convergence experiment The two effects that explained the differences in the translational motion of the different cycle types (Section 4.4.5) are also confirmed by the convergence experiment shown in Figure 4.23. The displacement deviation in the first cycle can be partially explained by the initial shear pre-loads in the friction contact, which showed to converge mostly in 1 or 2 cycles (as was shown in Figure 4.9). The first 8 cycles clearly show a decreasing offset of the translation step in the positive direction, which is caused by the lower initial temperature of the actuator. This effect is the parasitic motion that is described in the previous section. The difference between the positive and the negative motion directions that remains in the last 8 cycles is attributed to the sensitivity of the normal force and thus the friction force to the motion direction.

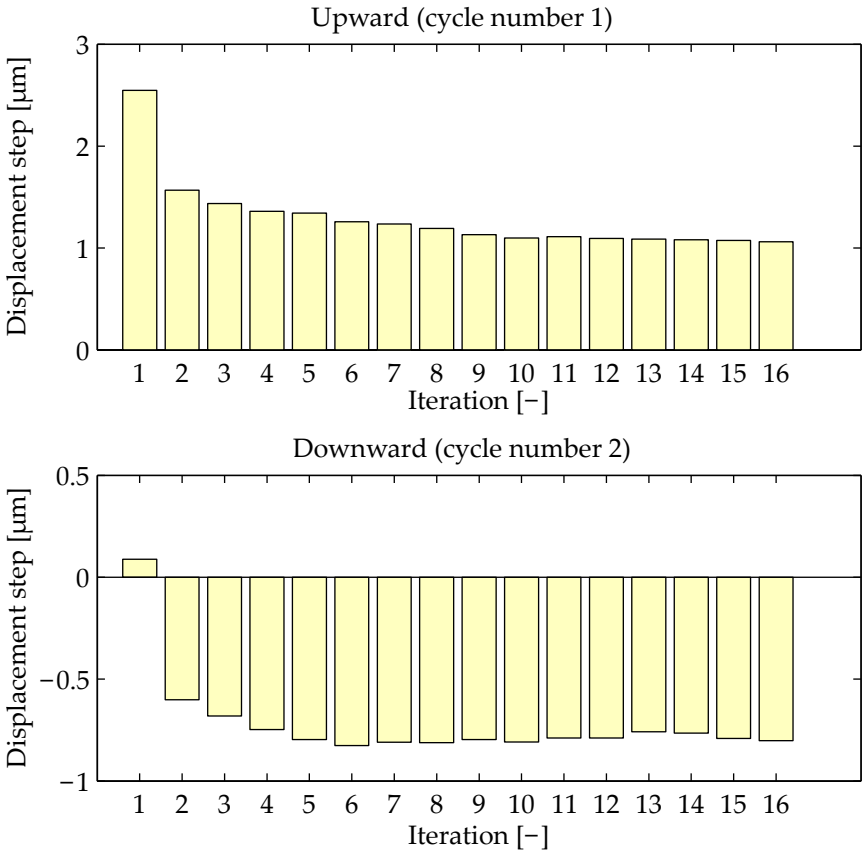


Figure 4.23.: Convergence of consecutive translation cycles with 4 sections of 4 groups on a 16 finger TSA. The top graph shows a series in upward direction, the bottom graph in downward direction. Both cycle series were executed from room temperature and from arbitrary initial contact conditions. The convergence effects from the initial pre-load and thermal condition are visible (iteration 1-8), as well as the friction direction dependency (iteration 9-16). Note however that the initial conditions in this experiment are not necessarily representative for operational conditions. A heating power of 0.24 W was supplied per finger.

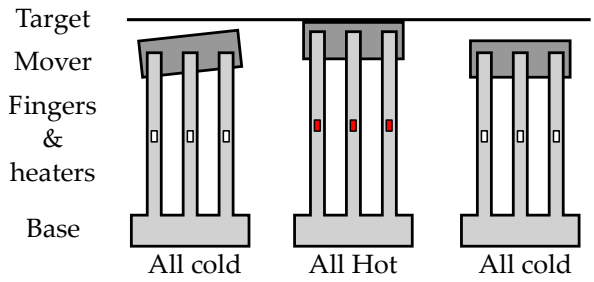


Figure 4.24.: Illustration of auto-alignment using the TSA. The mover is pushed to the target by heating all fingers and by applying thermal cycles. This aligns the surfaces to each other. The required mover-target distance is obtained by cooling down all fingers.

The initial condition that is used in Figure 4.23 is not necessarily representative for use in a positioning system. The measured standard deviation in the more representative case, using the data shown in Figure 4.22, are between 0.15 µm and 0.44 µm in the translational direction and between 7 µrad and 29 µrad in either tilt direction. In a positioning environment it is expected that the majority of the step's standard deviation is caused by the varying initial conditions, as different step-cycles will be used continuously. The remainder of the step-variation is caused by other sources, such as temporal and spatial variations in the friction properties of the contacts between the fingers and the mover.

4.5. TSA closed loop positioning

In the foregoing sections the possibilities of the TSA to displace the mover in three degrees of freedom by means of predefined thermal cycles have been explored. [Schieveen et al., 2010] described an auto-alignment method that only requires translational motion. It pushes the mover towards an alignment target using the regular thermal cycles and then cools down the actuator to achieve the required distance, as is shown in Figure 4.24. The 3 DoF thermal cycles can be used to displace the mover to a predefined position without the surfaces touching each other, so that damage to the system is prevented.

This section describes how the TSA is used as a closed loop positioning actuator and how it performs. First a limited set of thermal cycles is selected (Section 4.5.1) and its displacement properties are determined (Section 4.5.2). The cycle selection method described in Section 4.5.3 then uses this information without any compensation to determine the most appropriate thermal cycle for the positioning goal. Section 4.5.4 experimentally demonstrates this positioning method and shows that an accuracy (1σ) of 0.3 µm and 36 µrad can be achieved.

4.5.1. Thermal cycle set

A set of thermal cycles of, which the resulting displacement is measured, forms the basis of controlling the mover position using the TSA. These thermal cycles have been identified in the previous sections. They are:

- Translational cycles that ideally only generate a translational motion.
- Tilt cycles that use an upward translational cycle on one half of the actuator and simultaneously a downward cycle on the other half. Ideally this results in a pure tilt motion.
- Combined cycles that only use one up- or downward cycle on one half of the actuator. This ideally results in a translational and a tilt motion.

Within these three categories, it is possible to define an infinite number of cycles by varying the parameters that were analysed in the foregoing: grouping, step time and heating power. It was also shown that a significant amount of cross-coupling exists between the degrees of freedom. It is therefore beneficial to identify the real displacements as was done in the experiment shown in Figure 4.22. However, the more steps have to be identified, the more time and effort this process costs. Therefore an operational choice has to be made with respect to the balance between identification effort on one side and performance in terms of resolution, accuracy and speed on the other side. In order to present a representative picture of the positioning capabilities of the TSA with a limited set of cycles, only the following parameters were used for closed loop positioning in this section:

- Translational grouping in 2 sections of 8 groups and in 4 sections of 4 groups. These two options lead to an approximately equal motion velocity at significantly different step-sizes.
- Tilt grouping (for both the pure tilt cycles and the combined cycles) in sections of 7 fingers on one or both sides. This type of cycle leads to relatively well reproducing tilt steps.
- The time between the sub-steps within a cycle (earlier denoted as dt) is 10 seconds. This value is a compromise between motion efficiency and reproducibility. A translational cycle lasts 50 or 90 seconds. The tilt and combination cycles 90 seconds. [3]
- Each finger's heater was supplied with 0.24 W of electrical power. This leads to reproducible steps that are not too large. A larger power would increase the velocity, but also reduce the resolution.

4.5.2. Identified mover displacements

The limited set of thermal cycles can now be used for positioning the TSA mover. However, in order to select the cycles that have to be executed and to reach

[3] The tilt cycle that includes the all-off state, shown in Figure 4.17, is used to be able to monitor the current tilt angle after each cycle.

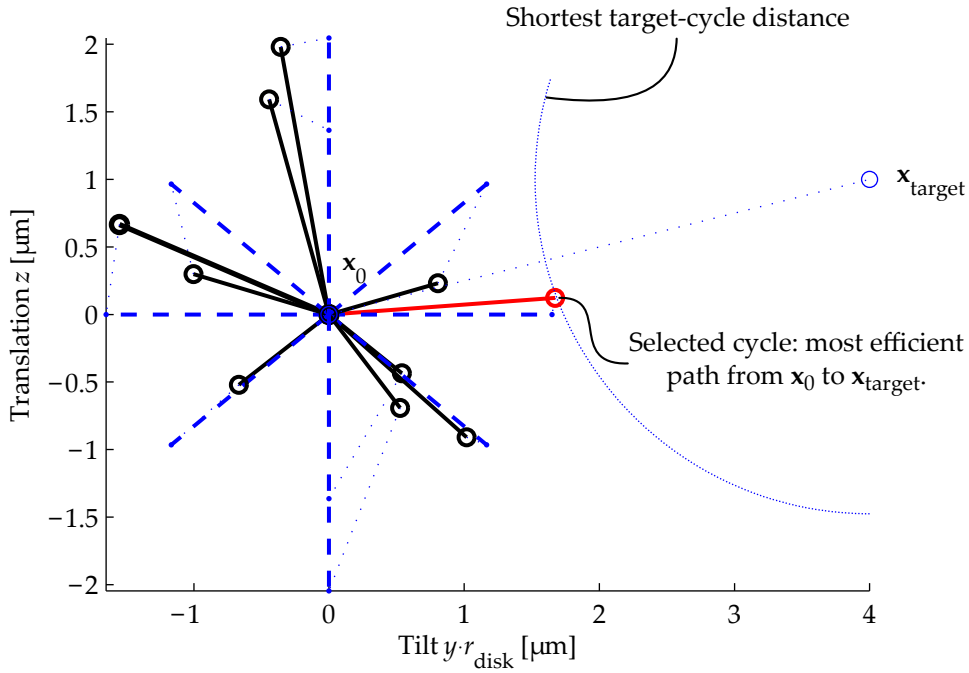


Figure 4.25.: Illustration of the cycle selection procedure for positioning with the TSA. For illustrative purposes the situation is simplified to 2 degrees of freedom: tilt around y and translation along z. The dashed lines represent ideal, estimated displacement steps, the solid lines represent the measured actuator motions. The dotted lines connect the measurement with the corresponding ideal motion. Note that the tilt axis is scaled with the electrode radius (r_{disk}) of 11 mm in order to have the same units on all axes. The positioning goal is to displace the mover from the current location $\mathbf{x}_0$ to $\mathbf{x}_{\text{target}}$. This is achieved by repeatedly selecting the cycle that results in the smallest position error. This is represented by the smallest circle around $\mathbf{x}_{\text{target}}$ that touches an estimated actuator motion. This actuator motion (indicated in red) is selected and executed. The procedure is then repeated.

the positioning goal, information about the resulting motion of each cycle is required. This information can be obtained, either from an idealised model, or from measurement data such as shown in Figure 4.22. The measurement data provides a more accurate description that also includes variations that are present in an individual TSA sample, caused by for instance normal force and friction variations of each finger, at the cost of an increased effort per actuator.

Figure 4.25 shows the mean measurement data that is also used for closed loop positioning using the TSA. For clarity, this figure only shows the displacements that correspond to a selection of the thermal cycles that will be used for positioning. This selection of cycles ideally result only in a displacement in the translation (z) and one tilt (y) direction. These displacements are shown as solid lines starting in the origin. The dashed lines indicate estimates that assume an equal normal force and an equal static friction coefficient at all fingers. The measurements are connected to the corresponding estimates by dotted lines. This overview shows that significant differences between the ideal estimates and the measured displacements exists.

Both the estimated and the measured displacement data can be used for positioning, however the most efficient positioning results are expected using measured displacements that have been obtained before the positioning process starts. When estimates of the displacements are used, reaching the final position may take more cycles and thus more time. Note that the displacement step variations caused by the larger thermal time constants of the actuator are not explicitly taken into account here. Further improvements can therefore be made by including calibration data of combinations of step cycles or information about the average heating power of foregoing cycles.

4.5.3. Position control procedure

The positioning procedure consists of two stages: first, during coarse positioning, the selection procedure selects cycles that move the electrode to its position goal in the most efficient way. When the positioning goal is close compared to the actuator step sizes, the fine positioning procedure continues. This procedure waits for the actuator to approach thermal equilibrium before selecting the next cycle, to take the actuator response due to longer thermal time constants into account. The cycle that leads to the smallest estimated position error is then selected. This procedure is repeated until the error cannot be reduced any further.

Scaling In order to judge which cycle to select and whether the final position is reached, the identified actuator displacements have to be compared to the positioning goal. This is however not directly possible because they have units of both metres and radians and, in order to select the most appropriate thermal cycle, the displacements and angles have to be compared to each other directly. In the computations, the angular dimensions are therefore scaled with an appropriate dimension, resulting in uniform units for all displacements. In this case the radius of the mover (r_{disk}, 11 mm) was chosen as this makes the actuator motion steps in the different directions of comparable magnitude. A larger scaling factor would increases the relative importance of the tilt positioning error. This scaling is also

Table 4.1.: Distinctions between fine and coarse positioning.

Positioning	Cycle selection	Timing	Criterion
Coarse	Velocity vector	No delay	$\mathbf{x}_{\text{target}}$ cannot be reached with one cycle
Fine	Displacement vector	Extra cooling time	$\mathbf{x}_{\text{target}}$ can be reached with one cycle

shown on the horizontal axis of the graph in Figure 4.25. The scaled space is used to select the appropriate motion cycles as follows below.

Cycle selection The cycle selection method is also indicated in Figure 4.25, assuming that the current position $\mathbf{x}_0 = \begin{bmatrix} y \cdot r_{\text{disk}} & z \end{bmatrix}^T$ equals $\begin{bmatrix} 0 & 0 \end{bmatrix}^T$ µm and that the position set point $\mathbf{x}_{\text{target}}$ equals $\begin{bmatrix} 4 & 1 \end{bmatrix}^T$ µm.[4] The displacement step (indicated by the solid lines) that brings the mover the closest to the set point in the scaled view is selected. The circle around the target-point $\mathbf{x}_{\text{target}}$ indicates this distance and the currently selected cycle is indicated in Figure 4.25 in red. Note that the scaling of the axes influences the eventual actuator motion, because a large angular scaling factor increases the contribution of angular position errors to the total position error.

The time duration of the various cycles is thus far not considered. For fine positioning, the method described above is used unaltered, but at larger distances from the target-point (coarse positioning), the displacement estimates are scaled using their cycle times to achieve a more efficient motion. In this way the cycles are selected based on the velocity vector instead of the displacement vector. The various properties of coarse and fine positioning are summarised in Table 4.1. Note that the motion can be made even more efficient by adding larger motions steps, for instance using a larger heating power, to the cycle selection. This is however not currently implemented as the main focus is on precision positioning.

Positioning mode switching The last function of the cycle selection method is to determine when to switch from coarse to fine positioning and when to finish positioning. Fine positioning is selected when the distance from the current location to the target-point ($\mathbf{x}_0 - \mathbf{x}_{\text{target}}$ in the figure) is shorter than the actuator motions. Positioning ends when the estimated magnitude of the position error after any cycle is larger than the current position error. Cooling down of the actuator after positioning may cause the position error to increases over this boundary, even after the positioning process has ended. Then an additional cycle is automatically performed.

[4]Note that in this example the position vector has two elements, so that it can be represented graphically on a two-dimensional plane. A three-dimensional version of this selection method is used in the closed-loop experiment that is discussed in Section 4.5.4.

4.5.4. Positioning results

Figure 4.26 shows a displacement measurement of a closed-loop positioning experiment over a relatively small range. The feasible positioning accuracy of the actuator can be estimated by the magnitude of the used steps. The translation steps are approximately 1 µm and the tilt steps around 100 µrad (or 1 µm when scaled with the mover radius r_{disk}). Therefore the final positioning errors are expected to be smaller than these values. This was the case for the specific measurement shown in Figure 4.26 and in general for the other measurements performed. In the shown example the distance travelled is approximately 11 µm in the displacement direction and 1.4 mrad in the tilt directions. This took about 4000 seconds (more that 1 hour) and 44 cycles.

In the first 3000 seconds of the measurement that is shown, coarse positioning moves the actuator towards the target-point. At 3000 seconds the position error lies within the boundary for fine positioning and a pause of 60 seconds is added to approach thermal equilibrium. After almost 4100 seconds the positioning error is 0.15 µm in translation and 30 µrad in tilt direction, which is within the limit for ending the positioning cycle. The motion that remains after this moment is due to the larger time constants that are present in the system. In this case the displacement while the system cooled down was 0.14 µm and 6 µrad in 400 seconds. No additional cycles were needed as this displacement decreased the final position error. This is however not the case in general, additional steps are required when cooling down results in a sufficiently large position error increase.

Ten more comparable closed loop positioning experiments were carried out. The behaviour is comparable to what is shown in Figure 4.26. In all instances the position converged to the positioning target and the positioning process stopped itself, indicating that the error is at least within the volume that is spanned by all the estimated actuator motions. The final position error was always within 0.7 µm and 50 µrad, with a standard deviations of 0.3 µm and 36 µrad. This is within the displacement and the tilt goal that were defined in Chapter 2.

4.5.5. Closed loop positioning conclusions

The positioning capabilities of the TSA depend on a large number of parameters and the final performance in terms of precision and efficiency is determined by the balance between these parameters. A basic, autonomous positioning process achieves a 1σ positioning accuracy of 0.3 µm and 36 µrad. After positioning, the only displacements that have been observed are caused by changes in the temperature of the environment. The positioning performance in terms of displacement and tilt therefore comply to the requirements: a translational resolution of 1 µm and a tilt accuracy of 300 µrad. Of these two, the tilt accuracy is the most important as this causes the largest systematic error (a gain error) in the capacitive displacement measurement, as was discussed in Chapter 2.

From the current performance, several improvements can be made, especially to the translational accuracy of the positioning system. When the influence of the average heating power of the previous cycle(s) is taken into account, the variation of the translational displacement of the cycles decreases significantly. As this

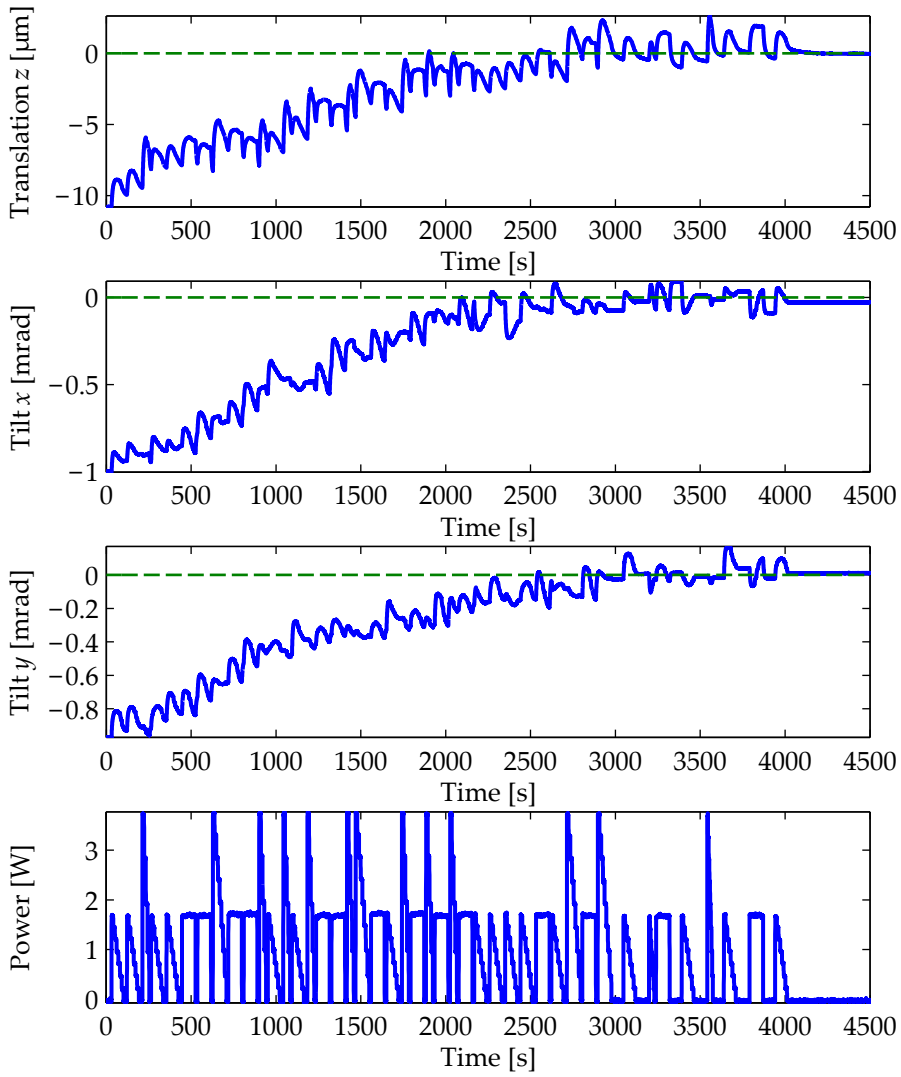


Figure 4.26.: Measurement of three DoF position control of the 16 finger TSA. The positioning target is zero translation (z) and zero tilt in both x and y directions, indicated by the dashed green lines. The initial position error is 10.8 µm in translation and 1.4 mrad in tilt. The final positioning error is 0.15 µm in translation and 30 µrad in tilt direction. The bottom graph shows the total heating power of the actuator. Fine positioning starts after approximately 3000 seconds, as can be seen from the pauses between the cycles.

reduces the smallest displacement step that can be made reliably, it is expected that this improves the positioning accuracy significantly. Additionally, the currently demonstrated velocity is approximately 10 µm per hour. This can be increased significantly, for instance by including larger power thermal cycles, to achieve a movement of 100 µm in one hour. No further efforts in these directions have however been made in this thesis, as increasing the understanding of the low velocity friction contacts (Chapter 5) had a larger priority.

4.6. TSA design considerations

Apart from the analyses that led to the closed loop position control of the TSA, several other relevant design aspects became clear in the course of this research. The aspects that are most relevant for the design of a TSA are summarised in this section, in some cases demonstrated by means of experiments.

4.6.1. Thermal expansion

In order to generate an efficient motion, the length of the clamping fingers should be most sensitive to their temperature. This implies the use of long fingers made of a material with a large thermal expansion coefficient. An absolute theoretical minimum value for the thermal expansion of a finger, was given in Section 4.2.2 as a function of the friction properties as: $\Delta x_{0,\min} = 2F_N\mu_s/k_{fin} \cdot n_g/(n_g-1)$ (Equation 4.5). Although an increase in thermal sensitivity is beneficial for the TSA motion properties, it also makes the mover position more sensitive to changes in the environmental temperature after positioning. This effect can be reduced by mounting the actuator such that the mover is in the so-called thermal centre of the actuator: the location that is neutral for overall temperature changes. An example of such a mount in shown in Figure 4.12a for the ASML/PI TSA.

4.6.2. Thermal design

Cross-talk The TSA working principle is based on differences in thermal expansion between the different fingers. It is therefore important that the thermal leakage, the amount of heat that leaks from one finger to another, is limited. Or in other words, the heater of a single finger should ideally only heat the finger it is attached to. Heat is however also required to flow from each finger to the base[5]. The thermal leakage is mainly determined by the thermal resistance between the fingers and the base, and by the effective thermal capacity of the base[6]. Both the thermal capacity and the thermal resistance have to be carefully considered to achieve a low thermal crosstalk and heating power.

The effect of the thermal resistance between the fingers and the base on the thermal cross-talk between fingers is demonstrated experimentally. Figure 4.27 shows an

[5]The heat flow picture in Appendix B (Figure B.1) indicates the significant heat flows in the TSA.

[6]The effective thermal capacity of the base can also include part of the surroundings when the base and the surroundings are thermally well coupled.

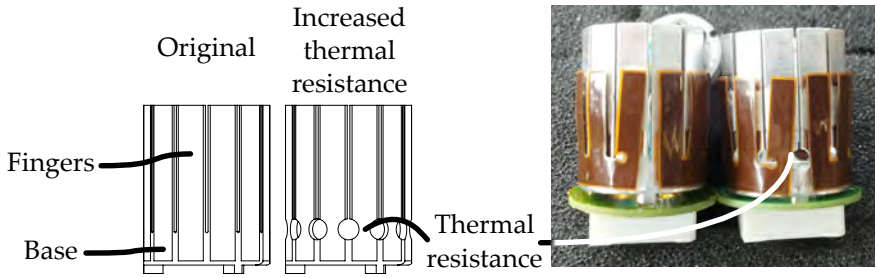


Figure 4.27.: Illustration of a method to increase the thermal resistance between the fingers and the base, applied on a ASML/PI TSA. The picture on the left schematically shows the difference between the original and the adapted actuator, the photograph on the right shows the realised parts (the housing shown in Figure 4.12 was omitted). The modification consists of a width reduction of all fingers close to the base. The thermal effect of this modification is shown in Figure 4.28

adaptation made to the ASML/PI TSA. The finger width is reduced close to the base to increase the thermal resistance. The effect on the thermal expansion of the heated finger and the other fingers of the actuator is shown in Figure 4.28. The finger elongation is increased by a factor 1.5, while the expansion of the first neighbour is only 1.2 times larger. No significant change of the dominant thermal time constant was observed. This adaptation is not ideal in terms of manufacturability, but a similar effect can be achieved by designing the fingers more thin and narrow. These changes to the actuator also affect the mechanical properties of the fingers, which is further discussed in Section 4.6.3.

Environment The thermal properties of the environment to which the TSA is mounted can have a large influence on the thermal behaviour. Several (FEM) simulations and experiments show that a low thermal resistance path from the base to the environment facilitates temperature differences between the fingers, and prevents overall heating of the actuator. Therefore, in practice the base should be chosen to be large with a large thermal conductivity to the environment. When the actuator is supported by a heavy (metrology) frame, this thermal conductivity can increase the effective thermal capacity of the base. This is for instance the case for the simplified TSA (Figure 4.13), where the base consists of plate aluminium with a thermal capacity smaller than that of all the fingers combined. By mounting the base on a 15 mm thick aluminium plate, the thermal expansion of the non-heated fingers is at least 16 times smaller than the expansion of the heated finger, despite the low base-finger thermal resistance. For a heating time of 10 seconds at a heating power of 0.9 W per finger this resulted in a heated finger elongation of 1.7 µm.

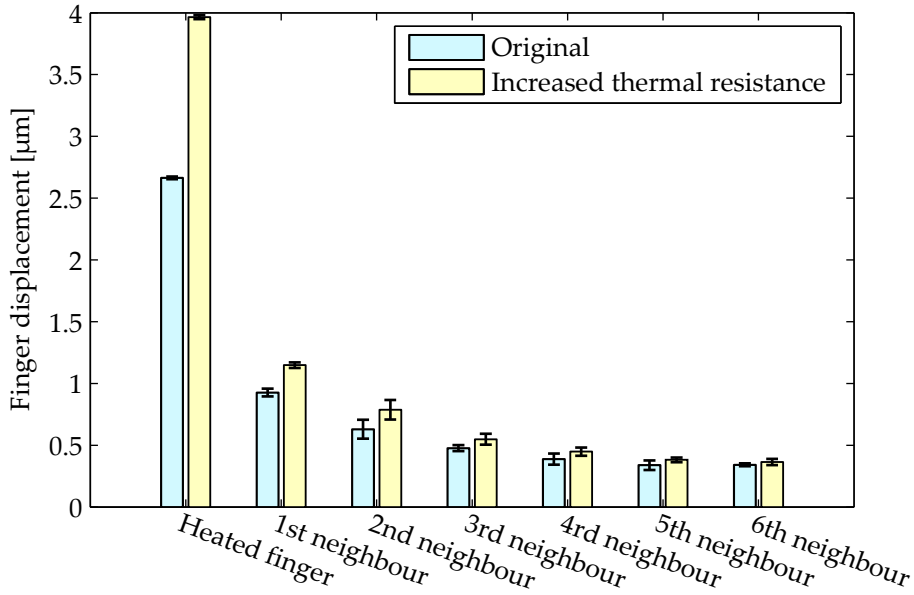


Figure 4.28.: Measurement of the thermal crosstalk between fingers of the ASML/PI TSA for a heating period of 60 seconds. The influence of increasing the thermal resistance between the fingers and the base, see Figure 4.27, is shown. The heated finger of the adapted TSA has a thermal expansion that is approximately 1.5 times larger than the original. The expansion of the first neighbour is only 1.2 times larger. The thermal cross-coupling is therefore significantly reduced.

4.6.3. Friction

The working principle of the TSA is based on the assumption that the friction force limit (the maximum friction shear force in a friction contact) between the mover and each finger are constant, both spatially and in time, and equal. Therefore, in order to work properly and to limit the parasitic actuator motions, the friction force limits have to be close to each other. Generally the frictional behaviour is influenced by the contact normal force and the relevant friction regime.

Normal force The normal force of each contact point is determined by the bending stiffness of the finger and its displacement from the neutral position (the position where the bending force on the finger is zero). The bending stiffness and neutral position are influenced by the dimensions of the finger and tolerances on these dimensions. The fingers of the TSA are long and slender for thermal reasons, as is described in Section 4.6.2. Also this shape causes the stiffness to be large in the motion direction, which reduces the sensitivity for external loads and vibrations, and smaller in the bending direction, normal to the contact. Deviations of the neutral position in the bending direction do however impact the normal forces in the friction contacts, and therefore also cause friction force variations. The circular configuration of the fingers (in a so-called spring-nest) reduces the influence of neutral position variations on the contact normal force, but differences between adjacent fingers do not cancel.

An example of normal force variation can be found in the ASML/PI TSA. Figure 4.29 shows the mover displacement as a result of heating a single pair of fingers on opposite sides of the actuator. When the heated finger-pair has a slightly larger than average normal force, the mover will still remain clamped by the non-heated fingers, but the larger friction force causes a larger than average elastic displacement. This displacement therefore directly relates to the friction force of the heated pair. Figure 4.30 shows the measured neutral position deviations of a comparable TSA sample. The deviations that are shown here agree with the displacement deviations that were shown before, confirming the relationship between neutral displacement variations and friction force variations. Also in the 16 finger TSA samples variations of the neutral finger positions were measured. On a typical sample deviations of maximal 30 µm were found. With a mean elastic displacement of 0.2 mm, this translates to 15 % friction force variations.

In order to reduce friction force variations, the design of a TSA should find a balance between the geometry of the actuator, the required manufacturing tolerances and the option to correct deviations. The parasitic motions that remain, should be dealt with using (feedback) control, as was shown in Section 4.5.3. Preventing finger position variations (which means improved tolerances) seems to be a logical option to reduce the normal force variations. This however complicates the manufacturing process. Reducing the sensitivity for position offset variation is a better option. This can be achieved by choosing a lower bending stiffness combined with a larger elastic displacement to achieve the required normal force. Alternatively, the position deviation can also be compensated by plastically deforming them to the required position, the result of which is shown for the ASML/PI TSA in green in Figure 4.30.

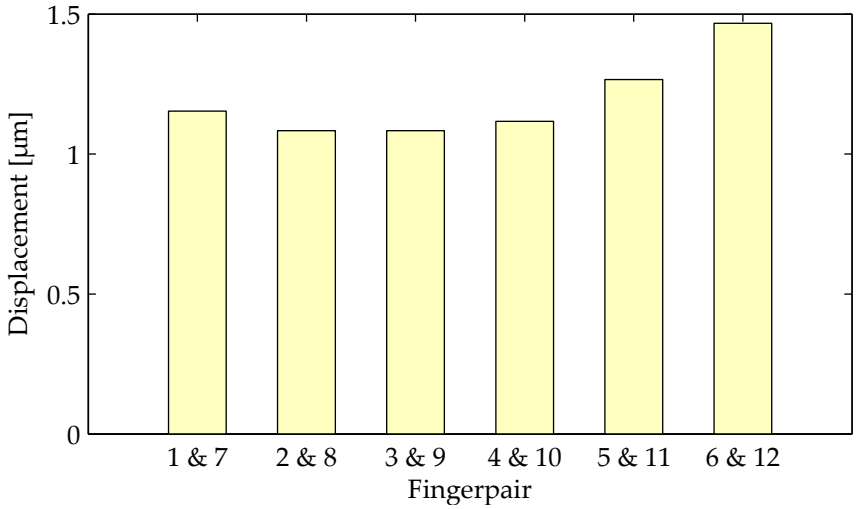


Figure 4.29.: Displacement of the mover of the TSA due to heating of a single pair of fingers. The difference between the finger pairs can be explained by the unequal clamping forces such as shown in Figure 4.30. This effect is measured on an ASML/PI-stepper such as shown in Figure 4.12, although a different TSA sample was used in this measurement.

Friction regime Next to the contact normal force, the friction regime has a large influence on the actuator behaviour. It was mostly assumed in the foregoing that a friction model that consists of only a static friction coefficient would be sufficient to describe the frictional behaviour. Experiments indicate however that also the friction process itself imposes its own dynamic behaviour on the actuator motion. This specific behaviour mainly plays a role in the transition between stick and slip and is referred to as pre-sliding, as will be discussed in more detail in Chapter 5. The existence of this effect means that a strictly force-based boundary between stick and slip does not exists. The fact that in many practical cases an instantaneous transition seems to occur, is explained by the micrometre scale of this pre-sliding effect, which makes its contribution often relatively small. The friction contacts in the investigated Thermal Slider Actuators however move on the (sub-)micrometre displacement- and sub-micrometre per second velocity scale, which is the scale at which the pre-sliding effect is relevant. The influence of the pre-sliding dynamics however seems to be small in practise, and because of the strong non-linearity of the pre-sliding effect, the constant friction force assumption is valid. Note that this is the case because of the specific displacement and velocity scale of the current demonstrators. When the actuation principle is applied at another scale this assumption is not necessarily valid.

Next to its impact on positioning, the friction behaviour also influences the passive mechanical stability of the mover after positioning. The pre-sliding effect influences this stability in two ways. The pre-sliding effect implies that motion of the friction

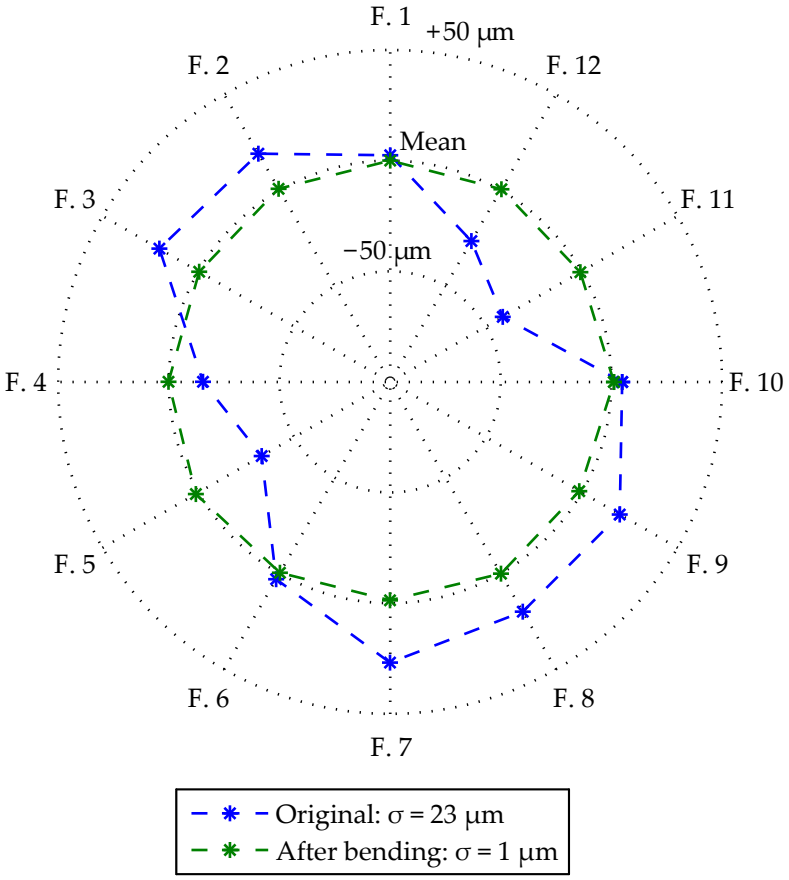


Figure 4.30.: Radial deviations of the neutral position of the fingers of a TSA. These measurements on an ASML/PI-stepper show the position deviations with respect to their mean value. The normal force variation that is associated with the deviation of the original actuator is more than 0.5 N. The measurements for the same actuator after manual correction (plastic deformation) are shown in green.

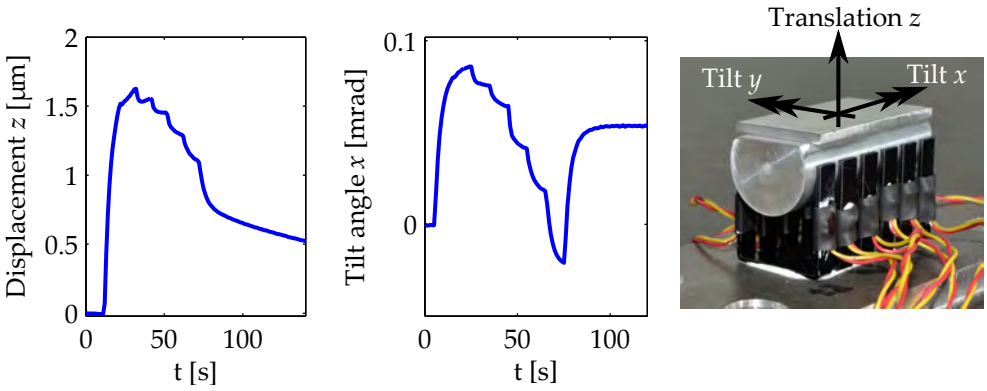


Figure 4.31.: A typical translation (z) and tilt (x) step (not measured simultaneously) of the simplified thermal slider design that was introduced in Figure 4.13. Tilt around the y-axis is relatively ineffective on this actuator due to the location of the fingers.

contacts occurs at lower forces than would be expected based on the static friction coefficient. The effect is however very non-linear and the displacements at forces well below the classical static friction limit will be relatively small. On the other hand, the pre-sliding effect will effectively allow a creep-like motion for continuously applied loads, such as the internal stresses that remain after a thermal cycle. This effect does however reduce over time, as the relaxation motion reduces the internal forces, and the strength of the stationary contacts increases over time. Sudden relaxation of internal forces that occur in the stick-slip regime, have never been observed in the TSA, which confirms that for the TSA demonstrators that have been investigated, the friction contacts move entirely within the pre-sliding friction regime. The absence of stick-slip provides a significant stability benefit for the sensor alignment application, as sudden displacement changes do not occur.

4.6.4. TSA shape

The fingers of a TSA can be arranged in many different shapes, but most actuators that were shown in this thesis have a circular arrangement. The circular shape is beneficial because, due to the actuator symmetry, the response of tilt cycles in different tilt directions is similar. Also with the circular shape a relatively large portion of the fingers is positioned at relatively large distances from the tilt axis, as was shown in Figure 4.18. Therefore an effective tilt motion can be made in all possible directions. When a different type of symmetry is used, significant differences between tilt directions can occur, while the translational motion is largely unaffected.

A good example of an alternative TSA shape is the simplified TSA that was introduced in Section 4.3. The fingers are here positioned in two rows in the x-direction (the axes are indicated in the photograph in Figure 4.31), which makes it more effective for tilt around x than around y. Figure 4.31 shows a typical translation

and tilt cycle response of this TSA for a heating power of 0.9 W per finger and sub-step time of 10 seconds. With 12 fingers in total, a translation step then takes 70 seconds and a tilt step 80 seconds. The cycles for the translation (z) and the tilt around the x-axis generate motions that correspond to their intended directions relatively well. However, the parasitic tilt motion around the y-axis is relatively large and the intended motion generated by an y-tilt cycle is small (these are not shown). This example shows that an alternative TSA shape and design can still result in a working actuator in selected degrees of freedom. This is even the case when the design is strongly simplified and uses low-tolerance of-the-shelf parts, in this case an aluminium heat sink, provided that they have the appropriate thermomechanical properties for a TSA.

4.6.5. Practical number of fingers

All the TSA demonstrators that have been used in this thesis, either have 12 or 16 fingers. Several aspects that are relevant to the selection of the number of fingers of a TSA are considered here.

First of all, there exists a theoretical minimum number of fingers, below which the actuator will not function. For a motion in only a single free DoF, 3 fingers are sufficient, as at least 2 fingers have to be stationary when 1 other is moving. See also Figure 3.5a. In the 3 DoF circular case that is required for the sensor alignment application, simulations confirm that a minimum of 5 fingers is needed. Using the theoretical minimum however has several drawbacks. While in theory all friction contacts might behave identically, in practice the normal force and the other friction properties of each finger can differ significantly. This causes parasitic motions of the actuator, or even an entire failure to move. Using more fingers can to some extent average out these effects, causing the actuator to work more reliably in all directions.

An additional advantage of using more fingers is that more variation in thermal cycles, and thus in motion directions (combinations of degrees of freedom) is possible. This improves the displacement resolution. Also grouping the fingers, for both translation and tilt motion, to give more motion flexibility, is only possible on a larger number of fingers. The main limitation for the number of fingers, is the allowed overall system complexity, as more mechanical features (fingers) and electronic features (switches and control electronics) have to be implemented. Also the thermal cross-talk between the fingers has to remain at an acceptable level and the total heating power might increase. The 16 finger actuator has shown to provide a good balance between complexity on one side and flexibility and resolution on the other.

4.7. Conclusions

In this chapter the Thermal Slider Actuator (TSA) was introduced as a new concept for permanent precision positioning. The TSA uses the opportunities within the boundary conditions of the sensor alignment application to provide a relatively cheap and simple actuator. The boundary conditions provide this space by omitting

a strict limit on the positioning duration, the positioning path (that is: only the final position accuracy is important) and the thermal load during positioning. This set of boundary conditions is however not unique to the sensor alignment task. Positioning and aligning of optical elements seems to be another promising applications field. The TSA concept provides the combination of good stability and actuation properties in three degrees of freedom, while the mechanical design, the electronics and the control can be kept simple.

Several positioning aspects of the TSA were already identified in earlier work. This chapter described the steps that were taken to go from moving in one translational direction to closed loop positioning in three degrees of freedom. This makes it possible to position and align a sensor such that the measurement- and the target-electrode of the capacitive displacement sensor do not need to touch.

Thermal actuation principle The TSA displaces a mover by applying thermal cycles to the fingers that friction-clamp this mover. The large number of these fingers makes the clamping statically overdetermined, allowing thermally induced internal stresses to manipulate the friction contacts. Elementary thermomechanical modelling has provided insight in the internal forces during the motion cycles and these insights helped in the formulation of an analytical approximation of the TSA step-size for translation.

Actuation in one translational and in two tilt directions uses the same basic thermal cycles, however many parameters of these cycles influence the resulting motion. Grouping, which describes to what extent multiple cycles are used simultaneously, showed to have a significant impact on the motion magnitude. Experiments were performed on a 16 finger TSA demonstrator. For both translational and tilt actuation, the mean motion velocity deviations that were observed are small, typically less than 10 %, provided that the grouping sections are sufficiently large for a reliable motion. For tilt actuation 2 sections of 5 fingers are sufficient, while for translations at least 4 sections of 4 finger groups is enough.

The heating power and the heating time per finger also influence the step magnitude. Also for these parameters a minimum value exists below which no, or only an erratic motion occurs. The heating time of a TSA with a shortest time constant of 8 seconds must at least be larger than 4 seconds. The minimal heating power per finger is also dependent on the chosen grouping arrangement as is shown in theory and practise.

The variation of the displacement that is the result of identical thermal cycles varies significantly, for steps in all directions. The convergence of an identical step that is executed repeatedly shows that the majority of the variance is caused by the longer thermal time constants (overall actuator heat-up) and by the initial pre-load state of the friction contacts.

An ideal TSA would show displacements only in the intended directions, however parasitic motions can be significant. It is therefore recommended to identify the real motion directions of each thermal cycle that will be used for closed loop positioning. This was done by executing the randomly selected thermal cycles while monitoring the relative displacements.

TSA position control The TSA can be used for open loop positioning within the limits of the reproducibility of the thermomechanical process: the longer time constants of the thermal system and the unknown initial conditions of the internal forces before each cycle. These uncertainties can be avoided by using an auto-alignment method, but this requires the surfaces to touch each other. Alternatively a closed loop 3 DoF positioning system was developed that applies the afore-mentioned predefined step sequences to the TSA, within a closed loop that automatically selects these sequences. Although closed loop temperature control showed a slight performance improvement, all presented results were obtained without temperature feedback. This shows that the thermal behaviour of the actuator is sufficiently repeatable to achieve sub-micrometre accuracy with a displacement feedback system. This feedback system selects thermal cycles based on the pre-identified actuator motions that result from a fixed collection of cycles. The cycle of which the identified motion would move the closest to the target is selected. Scaling of the different motion axes is used to balance the importance of the tilt and the translational motion. This control scheme has shown to converge to the control target-point within 0.7 µm and 50 µrad. Displacements due to longer time constants (such as in increased temperature of the entire actuator) are also automatically taken care of.

When the positioning accuracy that is obtained with the TSA is compared to the guidelines for the sensor alignment application that were summarised in Table 2.2, the alignment resolution is sufficient to achieve the ± 1 µm and the ± 0.3 mrad goals. The displacement accuracy can however be further optimised. A significant improvement can be achieved by using an estimate of the average actuator temperature and the corresponding mover displacement to improve the estimate displacement of each thermal cycle and with that improve the positioning accuracy.

TSA design As the base design of the TSA consists of a large number of fingers that clamp a mover (in this case a capacitive measurement electrode) and displace it by thermally manipulating the forces in their friction contacts, the key physics of the TSA are thermal expansion and friction. Therefore these are also the most important aspects to consider when designing a TSA. For an efficient motion the thermal expansion is concentrated in the fingers of the actuator and thermal leakage between the different fingers is minimal. This is achieved by carefully selecting the thermal resistance between the fingers (and between the fingers and the base) and the effective thermal mass of the base. The smaller the thermal mass of the base is, the more important are the thermal properties of the surroundings it is mounted to. These properties largely influence the thermal cross-coupling between the fingers, which is important for the effectiveness of the thermal cycles, and the overall heating of the actuator, which causes an error in the measurement of the momentary displacement.

At the specific displacement and velocity scale on which the current TSA demonstrators operate, the pre-sliding effect governs the displacement in the friction contacts, but for modelling purposes an effective static friction coefficient suffices. The contact normal forces influence the cross-talk between the different motion

directions. It is therefore beneficial that the contact normal forces are equal, which should be considered in the design and the manufacturing process. The remaining cross-talk is however taken into account by the positioning algorithm. A simplified TSA design showed the robustness for manufacturing tolerances, as an effective 2 DoF sub-micrometre resolution actuator was made at an estimated cost price below €20 in bulk production.

Many further optimisation steps are possible with the TSA. For instance, both the accuracy and the speed could be increased by optimising the input signals. Additionally the influence of design changes to the actuator and its materials, with different thermal, mechanical and frictional properties, can be further investigated. The focus of the further research has however been on the motion properties and positioning possibilities of low velocity friction contacts that are typically found in the TSA. This is the topic of Chapter 5.

CHAPTER 5

Friction contact manipulation

Friction is a core element of all the alignment concepts discussed in Chapter 3. Also in Chapter 4 it turned out that knowledge of the frictional behaviour of, in this case the Thermal Slider Actuator (TSA), is important to better understand its motion behaviour. A typical example of that is the absence of the expected stick-slip steps in the TSA. The relative motions of the contacts in the TSA are characterised by their micrometre scale displacement at sub-micrometre per second scale velocity. This chapter shows how the micro-scale friction phenomena can be used more directly for positioning purposes.

The relationship between load and displacement of friction contacts has already been studied for centuries and also many observations of micro-scale friction have been made over the years. Section 5.1 presents an overview of relevant literature in this field and discusses the pre-sliding friction phase. These observations in literature lead to a qualitative model of friction contacts that is useful both for understanding the friction contacts in the TSA and for the definition of new friction based positioning concepts. Using a measurement set-up that is introduced in Section 5.2, several experiments demonstrate the behaviour of friction contacts at low velocities in Section 5.3, confirming the validity of the friction model.

The remainder of this chapter concerns the application of this friction behaviour to three positioning concepts. The friction contacts of the TSA are discussed in Section 5.4 and it is concluded that the pre-sliding fiction phase plays an important role in micro-scale friction. The knowledge of this friction phase can also be used to define a new positioning concept: pre-sliding actuation. The working principle of this new positioning concept is explained and experimentally validated in Section 5.5. Section 5.6 describes another method to manipulate friction contact on micro-scale: by means of force pulses. Section 5.7 finally summarises the most important findings about the manipulation of friction contacts in micro-scale positioning systems.

5.1. Friction in micro-scale motion

The force and motion of objects in contact has been studied for centuries and is still an active research area. Work of pioneers from the 15th until the 18th century, such as Da Vinci, Amontons and Coulomb was mainly concerned with gross sliding of dry contacts [Beek, 2006] and the formulation of the basic engineering relationships

that are still in use today. Examples of these relationships are the proportionality of the friction force to the normal load and the independence of friction behaviour on the apparent contact area.

5.1.1. Static and kinetic friction

Coulomb also identified the difference between static and kinetic friction and the time dependency of static friction, the last of which was also confirmed by Morin [Jenkin and Ewing, 1877]. They also see in the work of Morin, and in their own experiments, indications that the transition between static and kinetic friction is continuous:

> "It is in fact highly probable that in those cases in which the static coefficient of friction is greater than the kinetic [...], the latter gradually increases when the velocity becomes extremely small, so as to pass without discontinuity into the former." [Jenkin and Ewing, 1877, p. 526]

But also more recent work such as [Wu and Tung, 2000] confirms this continuity. As is pointed out in [Olsson et al., 1998], many experiments were performed in the 1950's that investigate this transition. Especially the work of Burwell and Rabinowicz has shown several results that are useful for the friction contacts of the precision positioning devices described in this thesis. [Rabinowicz, 1951] explains the decrease of the friction force when motion is initiated on steel-copper interfaces by strong junctions (sub-contact points that originate for instance from the surface roughness) that are replaced by weaker ones. The junction strength is then shown to be dependent on the duration that it is stationary. Also [Sampson et al., 1943] already stated that the kinetic friction coefficient is equal to the initial static friction coefficient, that is, the static friction coefficient when the contact time is minimal. It is currently known that stationary contacts gain strength in time, an effect that is related to the evolution of the real contact area of an interface on timescales ranging from microseconds to hundreds of seconds [Ben-David et al., 2010b]. This effect explains the weakening of contacts due to relative motion and indicates that a sliding surface can be seen as a contact that is stationary for a relatively short duration. [Berman et al., 1996] uses this reasoning to explain the fact that the kinetic friction coefficient is generally lower than the static friction coefficient: The strength of newly formed friction bonds is lower than the strength of the older ones that were broken.

The experiments presented in [Burwell and Rabinowicz, 1953] concern indium and lead, which are significantly softer metals than those mentioned before. With these materials it is shown that the friction coefficient as a function of the steady-state sliding velocity ($\mu(v)$) shows a maximum when observed over a wide range of velocities. This is illustrated in Figure 5.1. They suggest that this maximum is also present in contacts with harder materials, but merely not observed due to the small motion and velocity scale that would be involved. The presence of this maximum (or more accurately the positive $\mu(v)$ relationship) suggests that when a low velocity is enforced on a contact ($^{nm}/_{s}$ scale or smaller for lead and indium), an equilibrium friction force is found. Also [Ruina, 1983] identifies two friction effects that may

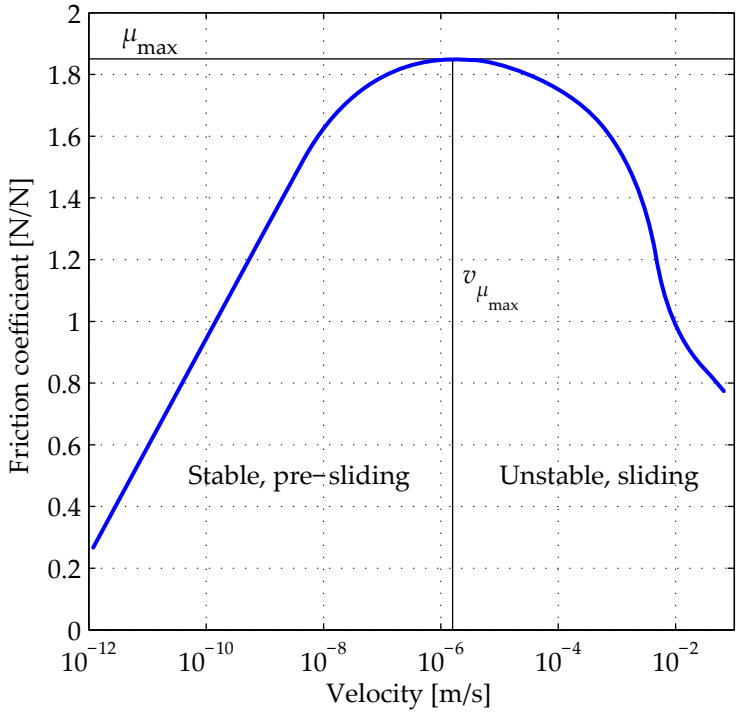


Figure 5.1.: Graph that qualitatively shows the maximum in the $\mu(v)$ relationship, based on the steel-indium graph presented in [Burwell and Rabinowicz, 1953]. This relationship is valid for steady state velocities on this specific material combination. Literature and experimental results in this thesis however indicate that the shape of this relationship is more universally valid, for instance for steel-steel and aluminium-steel combinations. For these other materials it is likely that the maximum force is achieved at significantly lower velocities. The maximum in this graph (μ_{max}) is generally considered to be the static friction coefficient μ_s, however, this relationship shows that a continuous motion also occurs at lower shear force levels.

together form a maximum in the graph of the steady state friction force as a function of the steady slip velocity.

5.1.2. Characteristic friction displacements

[Rabinowicz, 1958] experimentally observes steel-steel contacts and focuses on a characteristic contact displacement of an order of magnitude of 10 micrometres. This distance can be seen as the “minimum resolving power" of the contact: the displacement that has to be travelled in order to cause a change in the friction or shear force[1]. This preceding travelled 10 micrometres also influences the current friction behaviour. Prior to sliding this distance, the friction shear force remains unchanged and therefore also stick-slip effects are not present. Also in [Rice and Ruina, 1983, Ruina, 1983], which both discuss slip stability, a characteristic slip distances is identified that is independent on the slip rate. In this case the distance is 0.3 to 200 µm for the various materials that have been under investigation, and it relates to the travel required to achieve a new, constant shear force.

The time dependent strengthening of stationary contacts can be linked to the characteristic lengths that seem to be present in the friction dynamics on micrometre scale. [Heslot et al., 1994, Baumberger et al., 1995] discuss a scaling factor D_0 in the arguments of the static and dynamic friction coefficient as a function of time and velocity, respectively $\mu_s(t)$ and $\mu_d(v)$, that connects them: $\mu_d(v) = \mu_s(D_0/v)$. This scaling factor D_0 has units of length and is of the same order of magnitude as the contact relaxation length that is encountered in geological tribology. This relaxation length is also encountered in their own experiments on low velocity friction dynamics of Bristol (paper) board. This characteristic distance seemed to be independent on the systems dynamical properties (mass and stiffness), but can be related to the increase of the static friction coefficient in time [Baumberger et al., 1995], and to the typical dimensions of the sub contacts or asperities that comprise a contact [Berman et al., 1996]. These sub-contacts originate from the surface roughness of the contacting materials, which is in the dimensional range of conventional machining techniques such as milling and turning that are mentioned in [Muhs, 2005].

5.1.3. Pre-sliding: Creep in the stick-slip transition

The characteristic displacement discussed in Section 5.1.2, relates to the transition between stick and slip. Many experiments, such as those of [Baumberger et al., 1994, Heslot et al., 1994, Müser, 2008, Yang et al., 2008] show this by means of a creep precursor to slip which again is in the order of micrometres and is independent of the velocity. This precursor is observed in many different types of contacts, including metals, rock and paper. Its observability might however be limited as the dynamics and time scales of the slip processes vary significantly. The creep-like dynamics involved in low velocity friction seem to be qualitatively rather comparable for a wide range of conditions and materials, as is concluded by [Baumberger et al., 1994].

[1] The different forces that act in a contact were defined in Figure 3.2. The friction shear force F_f acts in the motion direction, orthogonal to the normal force F_N.

The creep precursor is one of the phenomena that seems to occur before a contact reaches slip and could therefore be indicated as a pre-sliding effect. There seems to be no general consensus in literature about which friction phenomena are considered as pre-sliding. [Wojewoda et al., 2008] for instance considers pre-sliding to be the mainly elastic part of the displacement that occurs before sliding, while in [Olsson et al., 1998] only the plastic part of the motion before sliding is referred to as pre-sliding. In the remainder of this chapter, pre-sliding is defined as all the displacement that occurs before full sliding, which entails both the elastic and the plastic (or hysteretic/creep) contributions. This plastic part can be seen as sliding on the left, stable side of the graph in Figure 5.1. Elastic displacements are not visible in the steady state velocity shown in this graph. The moment or point where full sliding starts then identifies the end of pre-sliding. Note however that in practical, non steady-state situation this transition from pre-sliding to sliding is a continuous process, and can therefore not always be clearly identified.

The effect of a creep displacement prior to sliding often remains unobserved in practice, not only because the transition to slip can be unclear, but also because the displacements in this creep regime are small compared to slip displacements that are often encountered [Heslot et al., 1994]. In friction modelling for fast and repetitive motion, modelling of this pre-sliding phase is important [Wu and Tung, 2000]. In the micrometre displacement, and (sub-)micrometre per second velocity range, these effects can be dominant. While in rock mechanics the pre-sliding motion is known as "precursory stable sliding" [Scholz, 1990] it shows self-accelerating dynamics at loads that are close to a "runaway threshold" [Heslot et al., 1994].

Although it is suggested by its name, the pre-sliding creep-motion does not necessarily result in full sliding. Many authors have observed that an elastically applied shear load on a contact, smaller than the static friction limit, merely relaxes through this creep-process [Rabinowicz, 1965, Heslot et al., 1994, Baumberger et al., 1994, Baumberger et al., 1995]. It could be clearly identified that on approaching the static friction limit, the relaxation magnitude increases more than linearly, but not exponentially. At shear loads much smaller that the static friction limit creep may also be present, although often practically not measurable due to a limited measurement accuracy [Heslot et al., 1994]. [Baumberger et al., 1994, p. 544] even states:

> "Actually, there is no true stick phase and the elastic stress partially relaxes through a creep process."

which implies that a creep motion is present at any non-zero contact shear load. The graph from [Burwell and Rabinowicz, 1953], shown in Figure 5.1, also suggests that a continuous motion is present at very low force levels.

5.1.4. Sub-contact properties

It is commonly known that macro size contacts, when observed on a micro scale, consists of many micro-sized sub-contacts. These contacts can play an important role in understanding the micro scale motion of macro scale contacts such as those

described above.[2] Next, a qualitative model is developed that is sufficient to understand the frictional behaviour of the TSA and that can be used to develop a force controlled positioning concept.

Friction contact behaviour When the shear force on a contact is increased, starting at zero, a contact passes through three phases that are relevant to low velocity friction. At forces that are small compared to the largest friction shear force ($F_f \ll F_N \mu_{max}$), the elastic behaviour predominates due to the extremely low creep-velocity, indicated on the far left in Figure 5.1. At shear forces that are closer to the maximum, a creep effect becomes more important for the relative displacement within the contact. When a displacement is applied to the contact, (possibly through an additional compliance) the creep process (partially) relaxes the elastic pre-load, as was shown by [Rabinowicz, 1965, Heslot et al., 1994, Baumberger et al., 1994, Baumberger et al., 1995]. When the contact is loaded with an external force, a steady-state sliding (or creep) velocity is approached [Burwell and Rabinowicz, 1953]. When the force is larger than the maximum, the contact runs away to the last of the three phases: sliding [Heslot et al., 1994].

Once the transition to the sliding regime (also known as full-sliding or gross-sliding) is made, many other system properties, such as the overall system dynamics gain importance. Also the transition to sliding itself is a complex dynamic process [Ben-David et al., 2010a, Ben-David and Fineberg, 2011]. This behaviour will not be considered in great detail as it will appear to be not relevant for the TSA and the force control concept that will be introduced later on in this chapter.

Asperity model The description of the pre-sliding motion that will be used for a new friction based actuation concept as well as the explanation of the TSA friction behaviour (Chapter 4) is based on the asperities, or sub-contact, that together form a friction contact. The dynamics of the entire contact (which is a collection of asperities) has been modelled as a whole, for instance in rate-state models [Yang et al., 2008, Rice and Ruina, 1983, Ruina, 1983, Baumberger et al., 1994]. Models that describe multiple sub-contacts however, may lead to more insight in the phenomena behind these dynamics. A model that remains closer to this asperity nature is the Generalised Maxwell-slip model [Al-Bender and Swevers, 2008], which uses friction contacts with varying properties and is able to describe both pre-sliding and sliding. Also according to [Al-Bender and Swevers, 2008] pre-sliding consists of non-linear hysteretic spring behaviour of the asperity contacts. An illustration of such a model is shown in Figure 5.2. These asperities break when they are stretched over a sufficiently large distance. As contact asperities need time to reconnect and gain strength, full sliding occurs when contacts are broken at a larger rate than the rate at which they reconnect. In full sliding, the bonds between contact asperities are thus continuously broken and formed. When contact breaking is mainly displacement dependent and contact forming is time dependent, the velocity dependence of the friction force can be qualitatively understood.

[2] In this context, macro scale refers to typical dimensions in the range of millimetres and bigger. Micro scale refers to micrometre orders of magnitude.

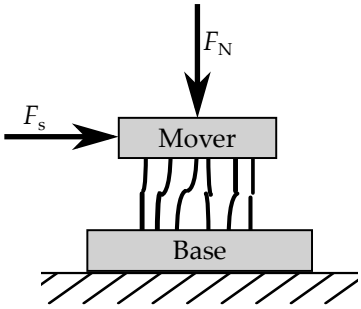


Figure 5.2.: Illustration of an asperity or bristle model of a friction contact. This can be seen as a more detailed view of Figure 3.2.

Reasoning from this concept by [Al-Bender and Swevers, 2008], combined with the continuity of the transition from stick to sliding, the entire (pre-)sliding process can be explained using the interaction of the asperities and the load of the contact. When a contact load is increased relatively fast, the bonds of the asperities will elongate as the contact force increases, resulting in the initial elastic behaviour. When the maximum extension of an asperity is reached, it breaks. This extension is the characteristic displacement or memory length discussed in Section 5.1.2. Not all asperities of a contact do however break simultaneously, as their individual properties (for example strength, location and pre-load force) vary. This explains the gradual decrease from static to dynamic friction that is observed in literature. As the bonds require time to (re)gain their strength, the mover will accelerate. When the mover moves at a sufficiently high velocity, the contact time of an asperity remains so small that the total friction force finds an equilibrium at a relatively low force, generally regarded as the dynamic friction coefficient μ_d. This equilibrium is caused by the balance between the rate at which contact are broken and the rate at which the contact strength grows. Once the effects of these processes are equal and opposite the friction force remains constant and has found an equilibrium.

5.1.5. Low velocity friction model discussion

The model described before can also give a (qualitative) explanation for many of the low velocity friction processes that have been observed in literature and experiment. It can be summarised by the following assumptions:

- A friction contact consists of many asperities that individually have a contact bond or junction between the contacting materials.
- These bonds behave elastically, but not necessarily linearly.
- When the sub-contact elasticity extends too much it breaks.
- The properties of the sub-contacts are randomly distributed.
- Contacts are renewed and grow in strength (logarithmically) in time.

So far all behaviour that is derived from this friction description is in agreement with the experiments that have been referenced in the beginning of this chapter. The question that now arises is: what would this description predict when a contact is loaded at a significantly lower rate, one that is more comparable to what is observed in the experiments with the Thermal Slider Actuator (TSA) in Chapter 4. The main difference lies in the fact that where previously the contacts did not have sufficient time to (re)engage, now the contacts do have this time. Note that the friction force increases with the logarithm of contact time [Ben-David et al., 2010b], which makes the effective contact strength of new contacts sensitive to orders of magnitudes of velocity reductions. Additionally, at reduced velocities the rate at which bonds are broken decreases as well. The combination of these two effects leads to the positive slope of the left side of the graph in Figure 5.1.

Motion at low shear force A first implication of the qualitative model described above, could be that when a force lower than the static friction limit (which can now be considered as the maximum friction force of a contact, as shown in Figure 5.1, but excluding the viscous effects at larger velocities) is applied to a contact, contact junctions will break, reconnect and gain strength. Breaking of these contacts at low forces, follows from the observation that at these low forces, next to the elastic contribution, also a continuous motion occurs that is significantly larger than typical bond dimensions. Additionally, bonds also have to reconnect and gain strength over time, as constant velocity and contact shear forces have been observed. Depending on the balance between the strengthening of the existing contacts and the breaking of the bonds due to the increased load per bond (when a bond breaks, the shear load is distributed over a reduced number of bonds), this process is either stable and self-braking, unstable and self-accelerating, or constant. Or, in other words, the motion velocity either decreases, increases, which is unstable and implies the transition to full sliding, or remains constant. Note that for the stable motion, the motion velocities can be very small due to the strong non-linearity in the friction strength growth, just like the contact motions that have been observed by for instance [Burwell and Rabinowicz, 1953] in Figure 5.1.

Controlled motion A second implication for low velocities could be that the motion (position and velocity) of contacting materials can be accurately controlled by indirectly influencing the combined strength of the bonds of a contact. As an example, the motion velocity can be increased by increasing the contact shear force F_s. This will increase the rate at which contacts break so that a new force balance is found, that is, another steady state point on the left half of Figure 5.1 is reached. The steady state velocity can however not be reached immediately by only changing to a new equilibrium force. An instantaneous force change does however lead to an instantaneous, non-elastic velocity change, which could be used to achieve a continuously controlled velocity.

History dependence Thirdly, due to the time dependency of the friction strength, and thus the history dependency of the contact, the same contact velocity can be achieved at different values for the contact force. The friction force is dependent on

the combined strength of all junctions, while the velocity mainly influences the rates at which contacts break, reconnect and gain strength. It can be imagined that these rates are not necessarily dependent on the total amount of junctions and therefore on the total friction force. A new friction contact for instance requires a smaller force to move at a certain pre-sliding velocity than a contact that has been standing still for a long time.

From the foregoing reasoning, the following conclusions are drawn about the motion of friction contacts on a (sub-)micrometre scale:

- At contact shear forces below the static friction limit relative, non-elastic displacement of a friction contact is possible.
- This non-elastic displacement can be manipulated to permanently position a mover on sub-micrometre scale.
- A continuous (constant velocity) motion can be achieved in the pre-sliding regime by manipulating the contact shear force.

The next sections of this chapter will show that these effects are also observed in experiments and that they can be used to create an accurate positioning system.

5.2. Friction measurement set-up

In order to investigate the usability of the friction behaviour mentioned in the previous section for positioning purposes, an experimental set-up is required that measures the behaviour of a single point friction contact. In this section the measurement set-up that is used in this chapter is presented.

Set-up requirements The measurement set-up, which will be built around a single point friction contact, needs to determine the shear displacement of the contact in time in a sub-millimetre range with nanometre scale accuracy. The contact normal force has to be known for each experiment and the contact shear force has to be continuously variable. A control system then has to set the shear force to a pre-defined value, a time series, or use position information to determine the required shear force in real-time.

Measurement and actuation concepts In order to measure the displacement with a minimal error, the contact position should be measured with a minimal influence of the compliance of the set-up and errors due to misalignment. Ideally the contact only moves along one axis and the contact point, the displacement measurement and the shear force application are on this single axis as well. This axis is show as the measurement axis in Figure 5.3. The normal force should be applied perpendicular to this axis and orthogonal to the contact plane, as this eliminates the cross-coupling between these two forces. As the normal force does not need to be varied continuously, a gravitational load can be used. The shear load can be continuously varied when it is applied by a voice coil force actuator. The motion in the normal- and shear-directions should be free, the others are constrained using

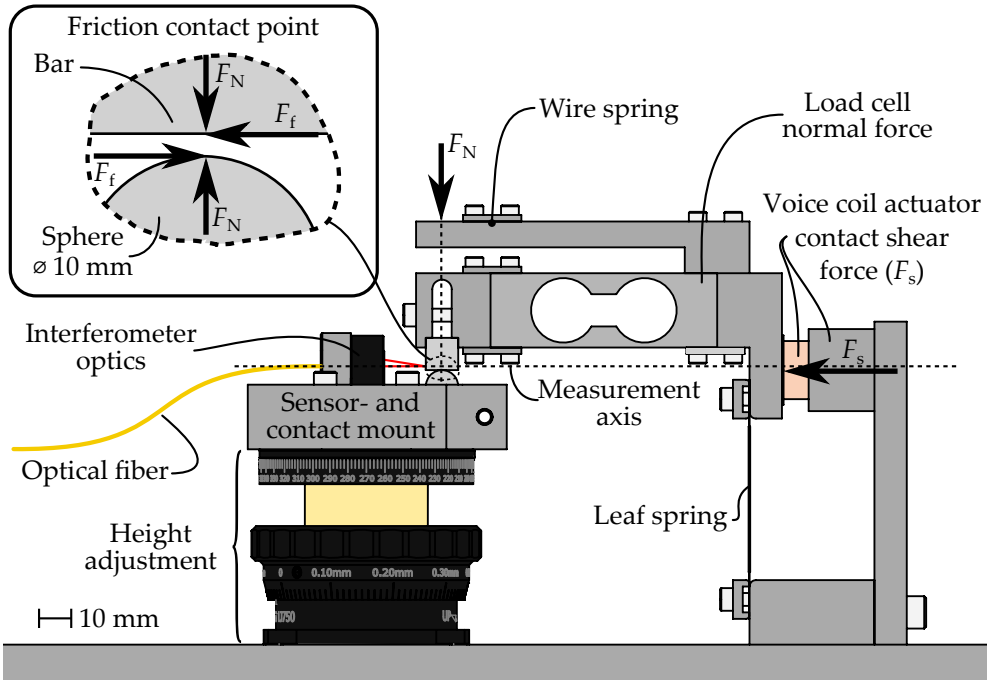


Figure 5.3.: Schematic drawing of the single point friction measurement set-up. The friction contact is as much as possible in line with the shear force actuator and the displacement measurement. The normal force is applied orthogonal to this line. A picture of the set-up is shown in Figure 5.4.

elastic elements, as their parasitic forces (in motion direction) can be kept small and predictable.

Realised set-up The realised measurement set-up is shown schematically in Figure 5.3 and with a picture in Figure 5.4. The contact consists of flat and a 10 mm spherical object, for a circular contact surface. Both interfaces are polished and cleaned. An Attocube FPS fibre coupled laser interferometer performs the displacement measurement and achieves sub-nm resolution over a multi millimetre range. The connection between the contact sphere and the interferometer optics, which is the measurement reference, is stiff. One of the polished surfaces of the contact bar is the measurement target. This results in an approximately 1 mm vertical Abbe offset between the measurement and the contact- and actuation axis, as can be seen in Figure 5.3.

The shear force (F_s) is applied through a voice coil force actuator with a known motor constants. This results in a friction shear force (F_f) in the contact. The normal force (F_N) is applied as a weight and is continuously measured by a load cell. The construction is such that the normal force is applied above the contact point and the leaf spring only supports the mass of the moving part of the set-up in the neutral

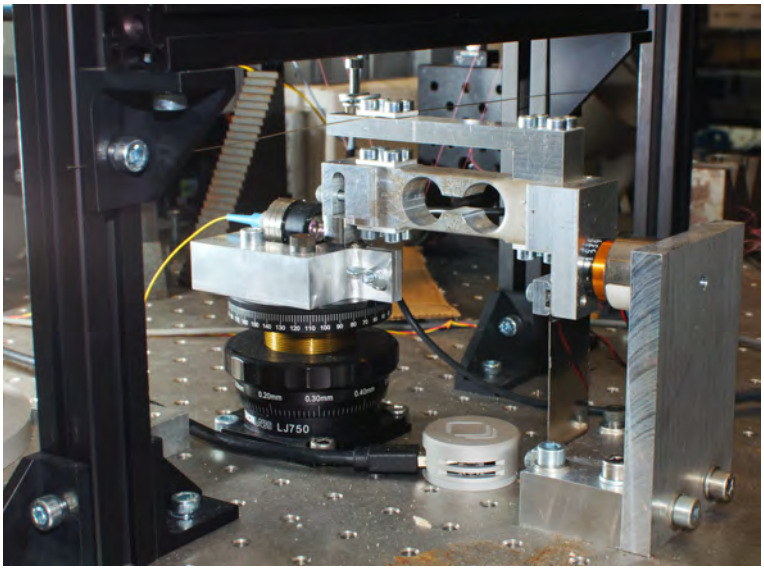


Figure 5.4.: Picture of the single point friction measurement set-up. This set-up is schematically shown in Figure 5.3.

position. The normal force shows a small dependency on the position (2.5 % per millimetre in a typical loading condition) as the contact point moves with respect to the centre of mass of the mover.

The mover is free to move in one shear direction and in the normal direction of the contact point. The moving mass is approximately 250 grams. The effective set-up dynamics are dominated by the contact stiffness and this moving mass, which results in an effective resonance well above 500 Hz. The other Degrees of Freedom (DoF) are constrained by a leaf spring and a wire. Due to the mounting positions of the wire on the mover, its influence on the normal force is measured by the load cell. In the shear direction the parasitic forces have been identified and can be compensated when required. The stiffness in shear direction is below 100 $^{\mathrm{N}}/_{\mathrm{m}}$ in a range of 2.5 mm, which means that for a motion range of 0.1 mm the elastic force varies less that 10 mN.

5.3. Pre-sliding friction experiments

This section describes the results of the pre-sliding experiments that were obtained using the measurement set-up described in Section 5.2. These experiments relate to the conclusions proposed in Section 5.1 about positioning in the pre-sliding regime: Section 5.3.1 demonstrates the plastic relative displacements of a friction contact under small shear loads. Section 5.3.2 shows that continuous positioning is possible using these displacements and in Section 5.3.3 the forces that are required for positioning in the pre-sliding regime are studied.

5.3.1. Pre-sliding displacement

In this section the behaviour of a friction contact at forces below the static friction limit is studied. It was seen in literature that the static friction coefficient of a material combination is not well defined and depends on other factors, such as the contact history [Ben-David and Fineberg, 2011].

Continuously increasing force experiments (Figures 5.5 and 5.6) In the experiments shown in Figures 5.5 and 5.6 the shear force on the friction contact is continuously increased. The top graphs show displacements on (sub-)millimetre scale. On this scale the motion is clearly stepwise, as is expected from the classic theory. However, on a smaller scale the transition is expected to be gradual, which can be seen in the lower graphs of these figures, showing the same data at a larger scale. In all cases the displacement starts to increase linearly with the force (or elastically), as indicated in the bottom graph of Figure 5.6. At increased levels of shear force the displacement gradually changes more-than-elastically at a rate that increases non-linearly with the shear force, towards full sliding. The initial stiffness of a contact and the force around which a contact goes to full sliding, varies significantly. Both of these effects can be related to the contact strength variation caused by the time that each contact is at rest.

Increasing force step experiment (Figure 5.7) The stability of the contact after a plastic pre-sliding motion cannot be judged based on the foregoing experiment, as the contact force is continuously increasing. The response to different levels of constant shear force is shown in Figure 5.7 by means of force steps. At a shear force below 0.3 N only the elastic displacement is visible within the measurement accuracy of several nanometres. According to the literature discussed in Section 5.1 however, at any non-zero contact shear force also a plastic displacement can be expected. When the force is increased, indeed a significant additional displacement is observed during the force application. The middle graph shows a detail of the displacement of the last force step. The displacement after the instantaneous elastic step is increasing, but at a decreasing rate. The contact is thus still regaining strength faster than the contact is broken and the motion is stable. When the force is increased too far (not shown in the graph) the motion does become unstable, showing a response such as shown in Figures 5.5 and 5.6.

In this increasing shear force experiment, the force is brought back to zero between different force steps. This shows two related phenomena. Firstly it shows at each step the amount of elastic displacement that is present in the contact as the instantaneous displacement. Over the course of this experiment no significant difference in the effective stiffness of the contact was detected. The displacement that occurs at constant non-zero shear force, is therefore not an elastic, but a plastic contribution to the total displacement. Secondly, the displacement at the different moments of constant zero-force then show the total plastic deformation of the contact and show that this displacement is indeed permanent and does not move back elastically.

The experiments have thus far shown that also at contact shear forces that do

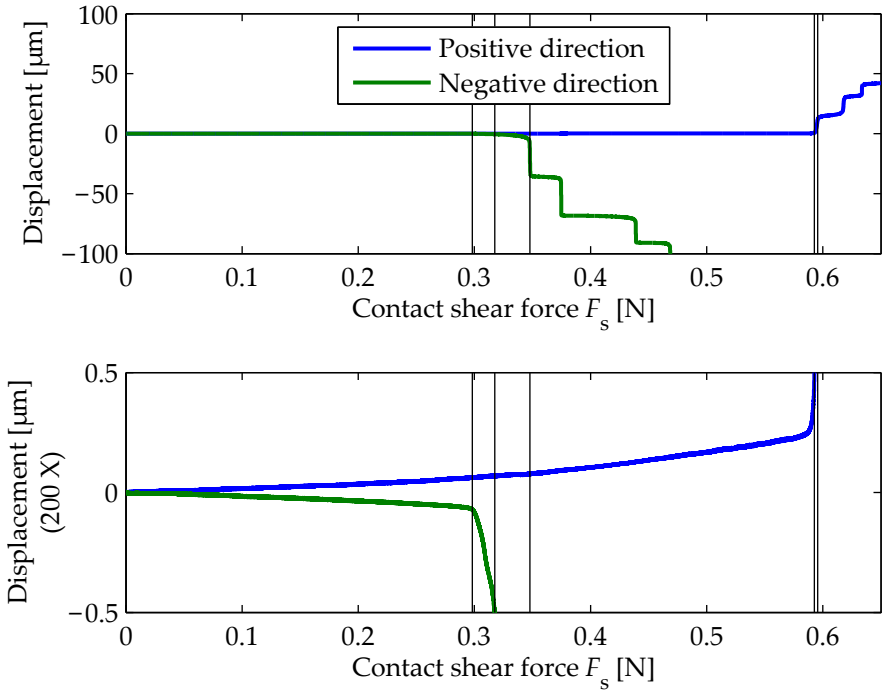


Figure 5.5.: Micrometre scale displacement response of a clean steel-steel friction contact to a contact shear force (F_s) increase from 0 to 0.7 N in 65 seconds. The force at which the contact starts moving plastically varies significantly due to local conditions and history. This experiment shows that although there are strong non-linearities in the friction behaviour, the displacement is always gradual. The black vertical lines in the graph indicate time instants in the graphs at both magnifications. They show that at different displacement scales, different effects are visible. The variation in the absolute shear pre-load between the experiments is smaller that 10 mN. The normal force (F_N) equals 1.7 N.

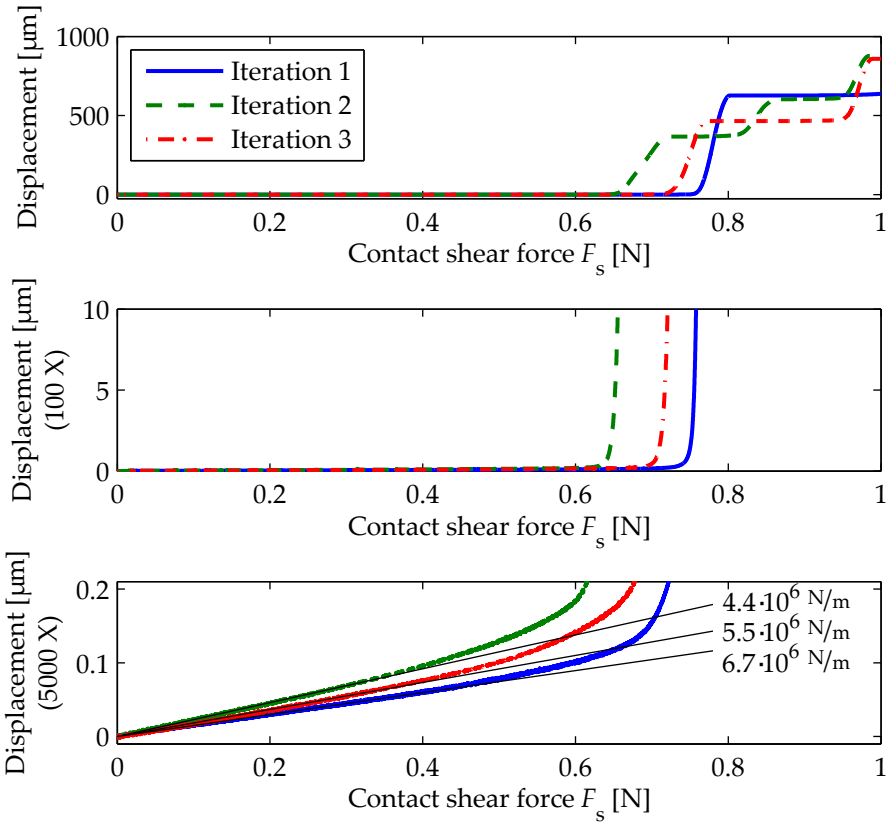


Figure 5.6.: Micrometre scale displacement response of a clean steel-steel friction contact to a contact shear force (F_s) increase from 0 to 1 N in 2 seconds. The experiment is repeated three times to show the effect of contact history. The variation in shear pre-load between the experiments is smaller that 10 mN and the contact normal force (F_N) equals 1.7 N. The stiffness of the contact (the initial slopes indicated in the bottom graph) and the force at which the first full sliding occurs varies significantly depending on the contact history.

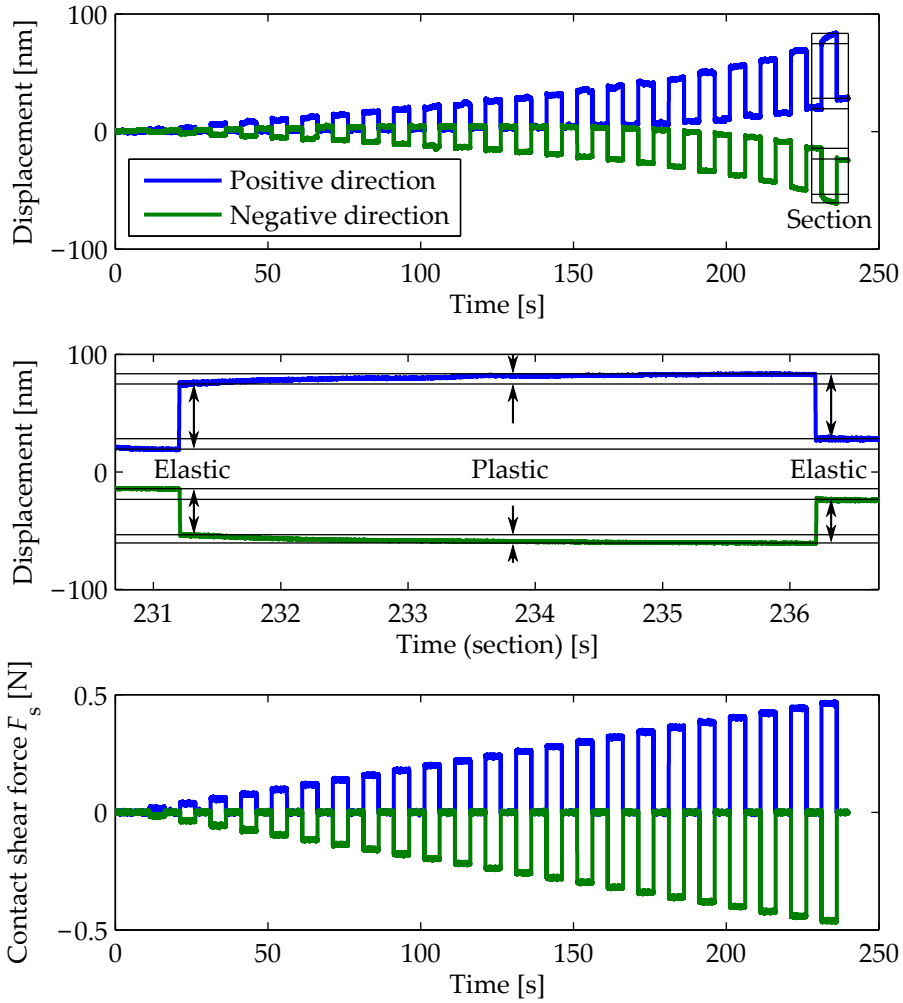


Figure 5.7.: Continuous pre-sliding shown by means of force step responses. The force increment of each step is 20 mN as is shown in the bottom graph. The middle graph shows both the elastic and the plastic part of the pre-sliding displacement of the last step-response, indicated with a black box in the top graph. The plastic displacement of the last steps is 8 ± 1 nm. This experiment shows that plastic (permanent) displacement can be generated by applying a predefined constant contact shear force. A further increase of the contact shear force leads to unstable sliding.

not lead to unstable sliding, generally considered as the static friction limit, relative displacements occur in the contact. For small forces these displacements are mainly elastic and for larger forces a plastic, permanent displacement is added. The velocity of this permanent displacement decreases non-linearly in time, but changes to unstable sliding at sufficiently large shear force. The plastic deformation is not instantaneous and finds, in this case, an equilibrium velocity on a time scale of seconds. The fact that this displacement is plastic and continuous means that these plastic displacements can be used for positioning the contact. This is the subject of the next section.

5.3.2. Continuous pre-sliding motion

The plastic displacements that were been shown in the foregoing all had a limited magnitude. For positioning tasks it is essential that the friction process does not limit the motion range, and that the motion can be stopped at any position. When we look at an instantaneous increase in force, the proposed friction description implies that the contact bonds are instantaneously stretched and a part of them breaks. After that follows a process of continuous breaking an forming of bonds. Both the instantaneous contribution and the transient contribution are influenced by the contact shear force. Also when the force is increased, the stability limit (the maximum in Figure 5.1) is approached. This leads to a gradual approach of the slip phase and thus gives the possibility to significantly increase the velocity. In conclusion, by controlling the friction force it should be possible to move the friction contact at any velocity, provided that the full sliding is avoided. This can be achieved by either limiting the force or the velocity.

Open loop pre-sliding positioning The reproducibility and the stability of the pre-sliding friction process makes it possible to move the contact without the need for a real-time closed loop feedback system. When a stepwise force input, such as shown in Figure 5.7, is used, the zero-force pauses can be used to measure the current displacement. A limitation of this feed-forward method is that the momentary velocity is not exactly known beforehand and the force has to be limited in order to avoid the unstable friction region. Due to the strong non-linearity in the relationship between force and velocity and the uncertainty in the maximum friction force, the safety margin on the force to be applied (to avoid the unstable region) leads to a relatively strong limitation of the velocity that can be achieved. In the foregoing experiment for instance, the instability limit was approached by less than 20 mN (around 95 % of the limit) while the pre-sliding velocity during the step starts at 5 $^{nm}/_{s}$ and decreases towards zero as can be seen in the middle graph of Figure 5.7. In order to achieve larger continuous velocities, either the force limit, which varies significantly, has to be determined experimentally or based on a model, or a form of closed loop control has to be used to approach this limit more closely.

Closed loop pre-sliding positioning The purpose of closed loop positioning in the pre-siding regime is to avoid reaching the unstable sliding regime, while moving the contact within the stable pre-sliding regime. The proposed method to achieve

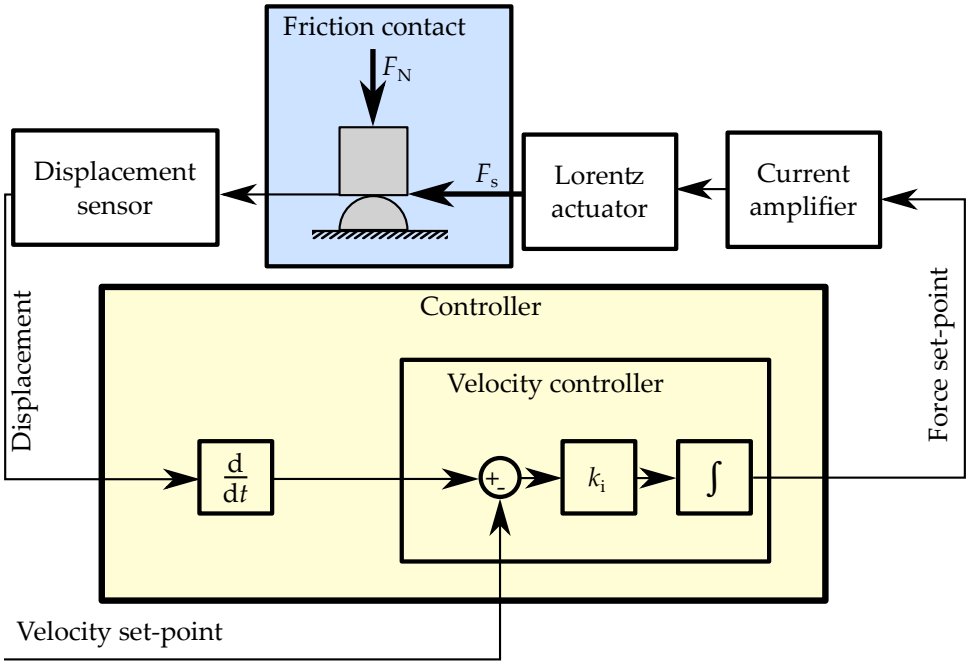


Figure 5.8.: Velocity controller structure. The yellow block represents the controller, the blue block the friction contact.

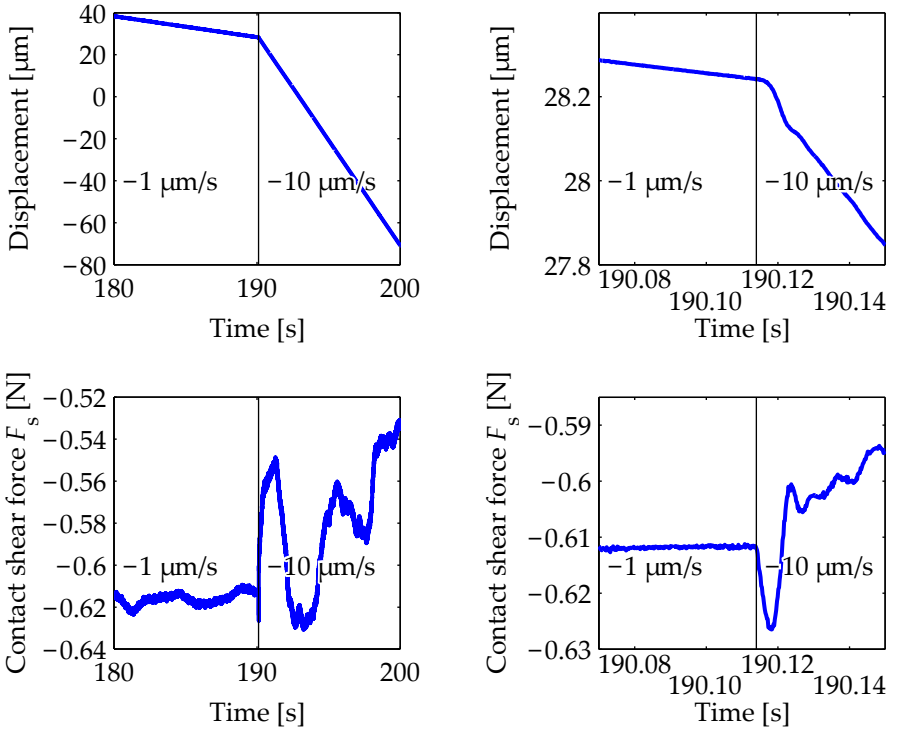


Figure 5.9.: Transient behaviour of a velocity controlled friction contact at a change of the velocity set-point from -1 to -10 μm/s. On the left a velocity change is shown on a 20 seconds scale. The graphs on the right show the same data zoomed in on the velocity set-point change.

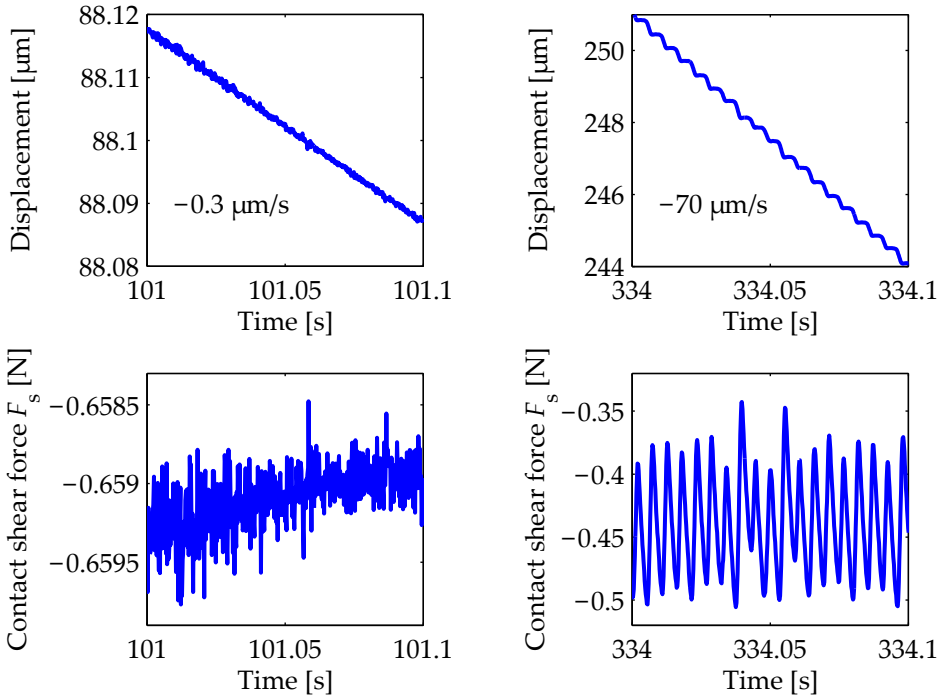


Figure 5.10.: Pre-sliding motion of a steel friction contact at two motion velocities. This measurement is performed using the closed loop velocity controller shown in Figure 5.8. At -0.3 µm/s, the force variation is minimal and the velocity is constant. (This confirms that the displacement is non-elastic.) At -70 µm/s the contact cycles between the stable and the unstable region and therefore shows a significant force and velocity variation.

this, continuously monitors the motion velocity and controls it by controlling the actuation force. When the contact shear force is too large to remain in the stable pre-sliding regime, the velocity increases. This velocity increase can be detected and the contact shear force will then be reduced accordingly to bring the motion to a steady velocity.

Pre-sliding velocity control The velocity regulating behaviour is achieved by means of an integral velocity feedback controller. This structure is shown in Figure 5.8, where the yellow block represents the controller. The inputs of this controller are the measured displacement and the required velocity (the set-point), the output is the required actuator force. The integral feedback on the velocity adapts the contact shear force such that the pre-defined velocity is maintained without the need for a position set-point.

Velocity set-point change (experiment in Figure 5.9) When the velocity set-point changes, the integral action of velocity controller will ideally increase the contact shear force until the required velocity is achieved. In practice the dynamical behaviour of the friction contact has a significant influence on the process of achieving a new velocity. Figure 5.9 shows the displacement and contact force of a velocity set-point change at two different zoom-levels. After the velocity set-point increases, the force initially increases. After that, a new force equilibrium is found. The force at the lower and the higher motion velocity show significant differences. At a larger velocity the variations in the friction force are larger. This can be explained by two phenomena: Firstly, from the perspective of the friction model, at higher velocities the contact is constantly closer to the instability limit. The non-linearity around this point causes a large sensitivity of the motion to force changes. Secondly, at these larger velocities, the changes in local friction conditions will also automatically become larger. This acts as an increase of a disturbance on the already more sensitive system.

Velocity differences (experiment in Figure 5.10) When the motion at various constant velocities is observed, the influence of the non-linearity near the friction stability limit becomes even more apparent. Figure 5.10 shows the displacement and the force for a motion of -0.3 $\mu m/s$ and -70 $\mu m/s$. At the lowest velocity, the deviation from a constant velocity motion is indistinguishable from the 1 nm noise of the displacement measurement, as the friction contact is sufficiently far from instability. At -70 $\mu m/s$ however, a distinct oscillation is present around the overall motion. This oscillation is caused by approaching the friction instability, or even crossing it, as the controller is not sufficiently fast to entirely eliminate the velocity increase. Reaching the instability does not imply that all sub-contacts are broken (for that the friction shear force would become zero), however the contact is loosing strength faster than it is gaining strength. This is a self-accelerating process that causes acceleration of the mover. The integral controller requires time to respond to this velocity increase. At a larger velocity set-point the distance travelled in this response time and the velocity that is reached are larger. Therefore the force and speed variations increase and the motion goes towards a stop and move cycle.

5.3.3. Pre-sliding forces

Although the oscillating behaviour shown in Figure 5.10 shows many features that are similar to the well-known stick slip effect [Beek, 2006], a static-kinetic friction coefficient model does not suffice to explain this behaviour. At different levels of the (average) sliding velocity, the displacement 'step amplitude' and force variation amplitude varies significantly. This confirms that the variation in the friction force can be explained by the friction state moving around the maximum friction coefficient show in Figure 5.1. The key difference with stick-slip is that where inertia is important for the stick-slip dynamics, the friction dynamics are involved in this friction oscillation.

Velocity control (experiments in Figures 5.11 and 5.12) The velocity and force variations that occur for a range of velocity set-points is shown in Figures 5.11 ($\pm 1\,\mu m/s$) and 5.12 ($\pm 100\,\mu m/s$). In these figures the mean and standard deviation (σ) of the velocity and the force is indicated. For low velocities (at least up to $1\,\mu m/s$) the velocity variation, which is caused by the velocity measurement and the friction process, is constant ($5\,\mu m/s$). The force required is approximately constant except for the $\pm 0.1\,\mu m/s$ region. For velocities of $20\,\mu m/s$ and larger (Figure 5.12), the velocity and force variations clearly increase with the velocity set-point, as the friction instability is approached or crossed. Also the mean force that is required for a motion decreases for increasing velocities. An explanation for this effect is that at larger velocities less sub-contacts reconnect and more sub-contacts are broken per unit of time. These effects causes the number of sub-contacts to be lower. When the velocity set-point is set to 0, the actuation force is ill defined, as a range of forces exists over which the velocity can be approximately zero. This explains the non-zero force shown in the bottom graph of Figure 5.12.

The motion range that was achieved using these low-velocity motions is in the millimetre range, in either direction. As these ranges are significantly larger than the contact area, these motions can be continued over indefinite distances without limitations from the friction process. When the contact shear force is brought to zero, the contact displacement remains stable within the measurement accuracy of several nanometres and within several seconds. This process can therefore be used for positioning and permanent alignment with sub-micrometre precision over an unlimited range.

5.4. TSA friction contact behaviour

The general motion behaviour of the Thermal Slider Actuator can be well understood from the theory that was developed in Chapter 4, however an important assumption was made with respect to the behaviour of the friction contacts of this actuator: the assumption that the friction-force does not change during motion. The pre-sliding effect that was introduced from literature in Section 5.1 and the experiments discussed in Section 5.3 suggest that this is indeed a reasonable assumption.

The motion velocity of the friction contacts of the TSA is an important variable for the friction behaviour. In the Thermal Slider Actuators that have been analysed

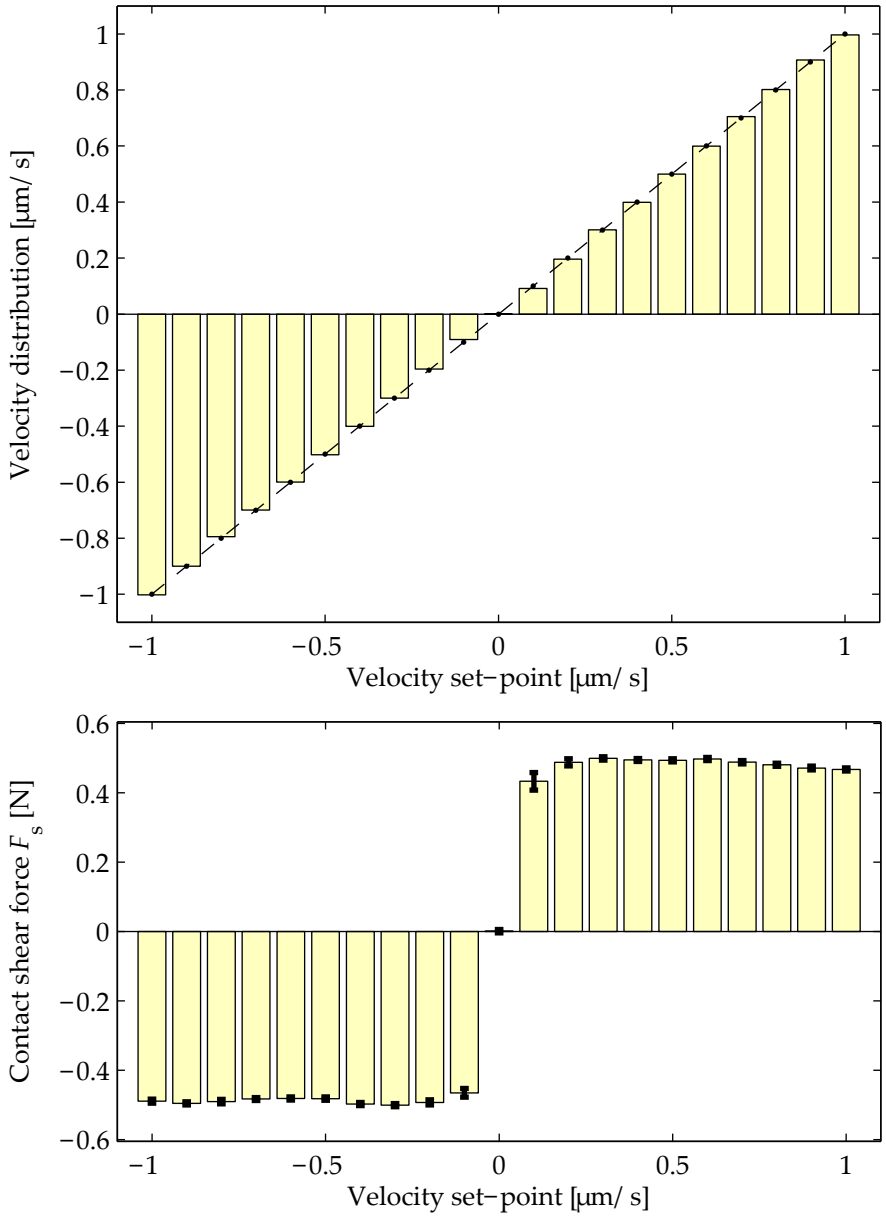


Figure 5.11.: Velocity control over a $\pm 1\,\mu m/s$ velocity range. The bars show the velocity and shear force at each set-point, recorded over approximately 10 seconds. The standard deviation of the velocity signal is constant and $5\,\mu m/s$ in this range. It is omitted in the graph for clarity. The dashed line/dots in the top, velocity graph indicate the velocity set-points. The mean velocity error is always smaller that 10 %. The bottom graph shows the mean contact shear force for each velocity and its standard deviation. This force is approximately constant at velocities above $0.1\,\mu m/s$.

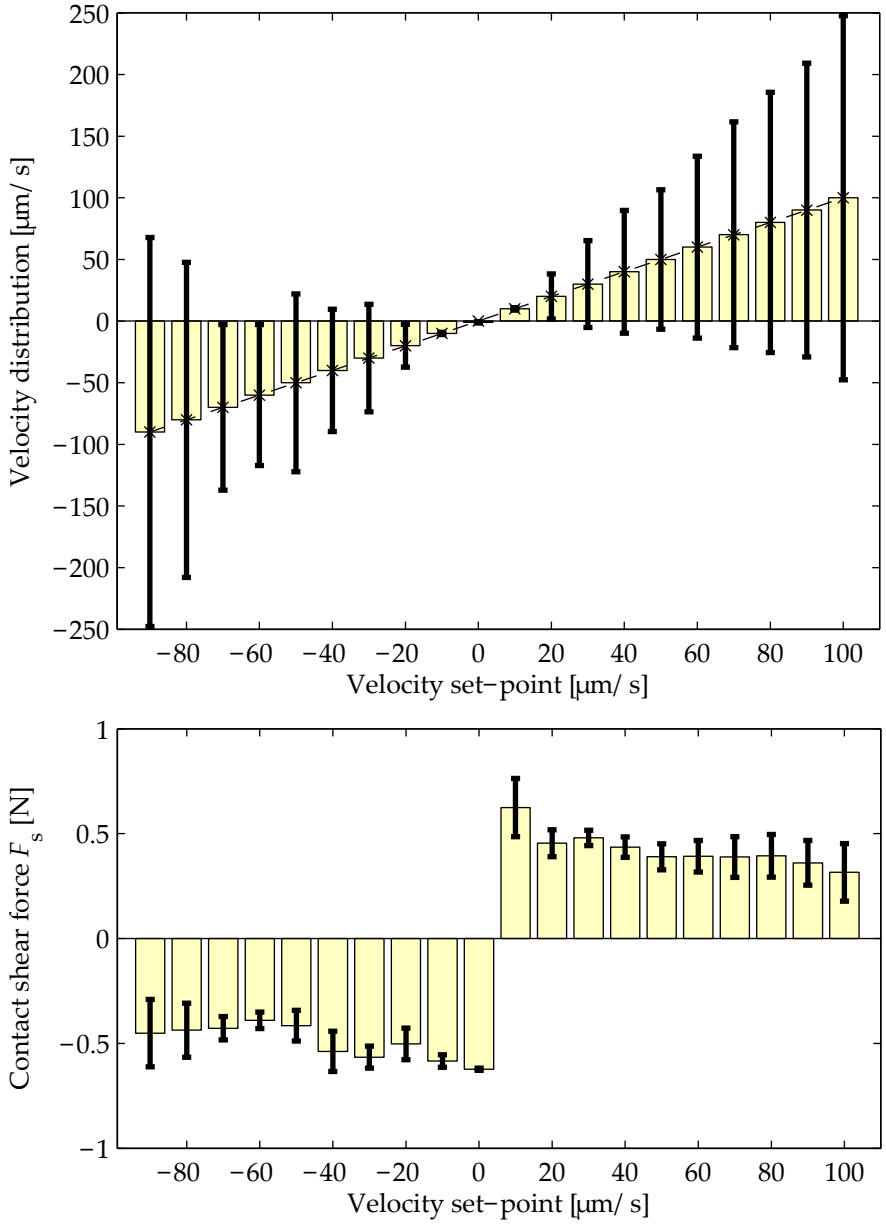


Figure 5.12.: Velocity control over a $\pm 100\,\mu m/s$ velocity range. The mean and standard deviation of the velocity and the contact shear force are shown for different velocity set-points. The dashed line/crosses in the top, velocity graph indicate the velocity set-points. For velocities of $20\,\mu m/s$ and larger the velocity and force variations increase significantly as the unstable sliding regime is approached. Also, in contrast to the lower velocity behaviour shown in Figure 5.11, the shear force decreases for increasing velocities.

in Chapter 4, these velocities are of an order of magnitude of 0.1 μm/s. A typical translational cycle, which was shown in Figure 4.15, for instance has a maximum contact velocity of 0.2 μm/s. Experiments on a steel-steel contact pair in Section 5.3 and similar experiments on steel-aluminium contacts, indicate that a stable pre-sliding motion can be achieved in this velocity range. This means that it is plausible that also the aluminium-steel contacts of the TSA operate at velocities that are sufficiently small to remain in the stable pre-sliding regime, or that the stiffness of the TSA is sufficient to keep it from entering the unstable sliding regime. This is in agreement with the experimental observations that a stick-slip displacement of the order magnitude that can be expected from the classical theory (10 nm on a finger) does not occur in the practice. Due to the large number of fingers the stick-slip motion magnitude on the mover would be significantly reduced, as it averages out over the different fingers, but this displacement would still exceed the stability requirement.

As long as the TSA operates in the stable (pre-sliding) friction regime, the contact shear force will remain constant (see for instance the left graphs of Figure 5.10). It will not (suddenly) change, as would be the case for larger velocities where the stick slip effect or the friction oscillation described in Section 5.3.3 is dominant. This validates the assumption that the friction force in the TSA can in this case be regarded as a constant, as was done in the previous chapter.

5.5. Closed-loop pre-sliding positioning

The experiments in Section 5.3 have shown that an integral control loop on the velocity is able to move a polished, clean steel-steel contact continuously on a micrometre displacement and a sub-micrometre per second velocity scale. In this section the same method will be used to form a pre-sliding closed loop positioning concept. It will be shown that with this method sub-micrometre accuracy positioning is possible, with the potential to reduce the accuracy to several nanometres when the contact elasticity is well estimated.

The contact velocity controller structure, which was shown in Figure 5.8, is able to apply a shear force to the friction contact, such that it will move at the requested pre-sliding velocity. This system is then turned into a position control loop by determining the velocity set-point based on the displacement error. A constant gain is in this case sufficient. This control system is able to move the system down to sub-micrometre per second velocities, however high velocities (towards millimetre per second orders of magnitude) lead to more irregular motion and larger force variation. Therefore, the velocity is limited by the controller. This also ensures that the contact remains moving in the stable pre-sliding regime. The experiment shown in Figure 5.13 demonstrates the behaviour of this position controlled system.

Pre-sliding positioning control (experiment Figure 5.13) The position-velocity control loop is initiated with a displacement set-point change from 0.4 to 0.0 mm, shown as a dashed line in the inset graph. This in turn changes the velocity set-point. With a velocity limit, 50 μm/s in this case, taken into account, the

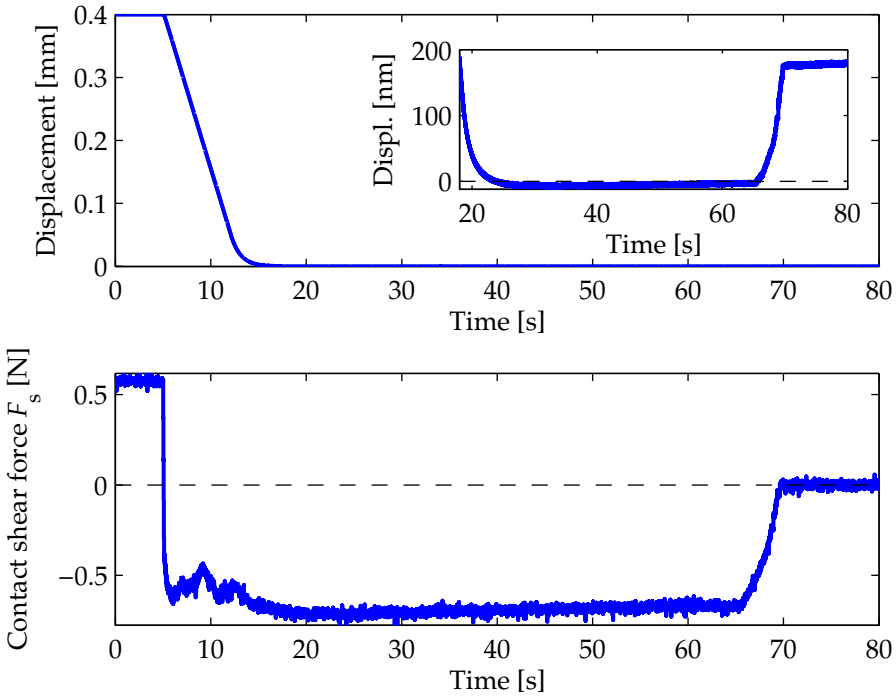


Figure 5.13.: An example of pre-sliding position control using a velocity-position control loop. The positioning goal (indicated by the dashed line) is 0 nm, while the initial position is 0.4 mm. The goal is achieved with a 10 nm error within 30 seconds, however a 180 nm error is introduced when the contact shear force is brought back to 0 (between 60 and 70 seconds). This displacement is shown in the inset graph.

displacement approaches the set-point (0 nm in Figure 5.13). The inset graph shows the displacement when the positioning target is approached within several nanometres. After 30 seconds the positioning goal is reached within 10 nm and the resulting velocity is below one nanometre per second. The contact actuation force is however non-zero, while the goal is to reach this position at zero contact shear force[3]. When the force is reduced to zero, a significant elastic displacement occurs (see the inset graph in Figure 5.13). In this case the effective contact stiffness of $4 \cdot 10^6$ N/m results in an additional displacement of 180 nm, between 60 and 70 seconds. Also including this elastic displacement, the pre-sliding positioning system provides a positioning accuracy that is well within the 1 µm requirement for the sensor alignment application that forms the origin of this thesis. This accuracy can however be reduced to within 10 nm when the measured displacement is perfectly corrected for this elasticity. It should however be noted that, as was shown in Section 5.3, the contact stiffness varies significantly depending on the motion history of the contact. This stiffness therefore changes during the positioning process. In order to achieve the best possible positioning accuracy with this pre-sliding positioning method, the momentary contact stiffness should be determined towards the end of the positioning process. This estimate can then be used to compensate for the elastic displacement. Such a compensation is however not currently implemented.

Pre-sliding positioning with PTFE In additional experiments by Paul Ouwehand, described in [Ouwehand, 2015], it was shown that this positioning method can be used for closed loop positioning of a planar stage. It uses the pre-sliding regime of PTFE-steel contacts to achieve a maximum positioning error of 10 nm during motion tracking at 10 nanometres per second. Additionally it was shown that pre-sliding positioning can be combined with other positioning strategies to achieve fast (millimetres per second) positioning with nanometre scale positioning errors.

5.6. Magnetic force pulse actuator

It was shown in the foregoing part of this chapter that the displacement in a friction contact can be manipulated on a sub-micrometre level by continuously manipulating the contact shear force based on a measurement of the instantaneous motion velocity. The relevant friction phenomenon here is the continuously breaking and reconnecting of the sub-contacts over time. Moving a friction contact does however not require a continuous force, also force pulses can be used effectively for precise manipulation of friction contacts. In view of the qualitative asperity based friction model that was discussed in Section 5.1, each force pulse would break a limited number of sub-contact. The contact than finds a new elastic equilibrium. This section presents mostly experimental work on a new force pulse actuation concept with which sub-nanometre resolution positioning is realised.

[3]Recall from Chapter 1 that for the capacitive displacement sensor alignment application, passive stability is required to avoid any thermal load to the environment.

5.6.1. Impact concept

Applying force pulses to friction contacts is widely used for alignment purposes, either manually or automated. An example of automated force pulse actuation was already shown in Chapter 3 with the piezoelectric Inertial Sliding Motion (ISM) and Impact Drive Motion (IDM) concepts (Section 3.3.3). A means to avoid the use of piezoelectric elements and simultaneously achieve sharper force pulses was suggested in Section 3.3.4 and Figure 3.5b: the use of a physical impact mass, a 'hammer'.

Mechanical force pulses While several authors have constructed impact mass mechanisms [Siebenhaar, 2004, Furst et al., 2010, Liu et al., 2003, Liu and Li, 2010], they remain relatively complex and often allow only actuation in a single motion direction. Displacement steps down to 10 nm are generated by the application of (a series of) pulses coming from a mechanical impact mass that is driven by a piezoelectric and/or a electromagnetic actuator. In these applications found in literature, an intermediate mass (the hammer) is used as a means to apply a well-defined force pulse. This impact mass gains and stores kinetic energy over a relatively large stroke, at a low force level, while this energy is released at a high force level in a short time, generally below a millisecond. This energy storage step is required in order to generate the power that is required for the force pulse, however the examples from literature show that these mechanisms add to the overall system complexity and introduce additional dynamics.

Electromagnetic force pulses An alternative to the use of an intermediate hammer object would be the direct application of a force pulse using an actuator[4]. As voice coil actuators are suitable for applying a well defined force contactlessly over a reasonable range, the single point friction set-up that was introduced in Section 5.2 can be used to test the feasibility of this magnetic force actuator for force-pulse positioning. The set-up transfers the electromagnetic force from the actuator to the friction contact point. Generating force pulses means applying electric current pulses to the inductive load of the voice coil actuator. The methods that were used to generate these current pulses are discussed next.

5.6.2. Generating current pulses

The current pulses that produce the force pulses in the voice coil actuator can be generated by the combination of the real-time control system and the current amplifier that were used in the continuous force experiments described in Section 5.3. Both the speed of the amplifier and the control system limit however the duration of the current pulse to a minimum of 0.2 ms. In the experimental results (in the following section) it will be shown that shorter pulses are advantageous for generating small displacement steps. An alternative means of generating these current pulses is therefore desired. Additionally, as the current pulses become

[4]This can be done provided that the reaction force of the actuator can be applied to the surroundings or another additional mass.

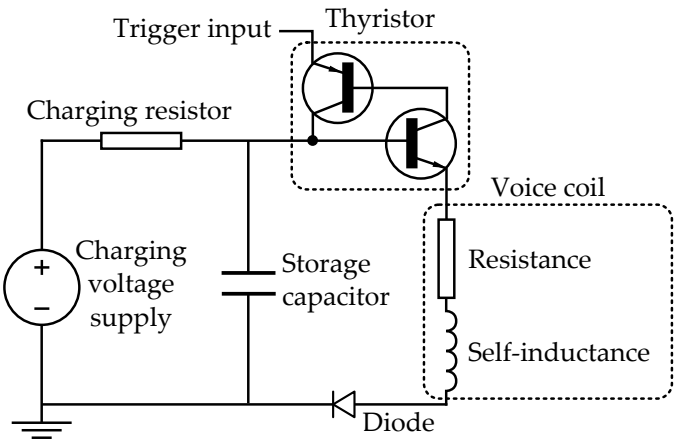


Figure 5.14.: Simplified schematic of the current pulse generating circuit. Energy is stored on the capacitor, which is in turn charged by the voltage supply. The charging resistor limits the charging current. When the thyristor is triggered, it closes the circuit so that current flows from the capacitor through the voice coil, generating an electromagnetic force in the voice coil. The thyristor opens when the current goes through zero, assisted by the diode.

shorter, the instantaneous power that is required increases. Where the mechanical impact hammer overcomes this power requirement by storing kinetic energy in the hammer inertia, it can also be stored as electric potential energy in a capacitor.

Current pulse circuit The circuit that is used for generating short current pulses is schematically shown in Figure 5.14. The energy storage capacitor forms, combined with the self-inductance and the resistance of the voice coil, an electric, damped resonator. When this electric resonance can be excited upon request and stopped after half a period, the resulting current through the voice coil can be used as a current pulse. This pulse shape is shown as the ideal current and force line in Figure 5.15. The pulse duration of this pulse is then determined by the resonator natural frequency, which depends on the capacitor capacity C and the voice coil self-inductance L: $\omega_0 = 1/\sqrt{LC}$. Exciting this resonance is achieved by introducing a thyristor switch in the circuit. This thyristor switch lets current flow from the pre-charged capacitor through the voice coil when it is triggered. The thyristor switches back from the conducting mode to the blocking mode when the current goes through zero, which happens after the first half cycle of the resonance. A measurement of several current pulses for different initial capacitor voltages are shown in Figure 5.15. Due to the non-ideal properties of the thyristor and the diode, a small current flows in the negative direction, after the pulse.

Creating a current pulse that results in a force pulse of a pre-defined magnitude and duration, requires a careful balance between the design of the voice coil and the rest of the electric circuit, as the voice coil parameters (the sizes of the coil and

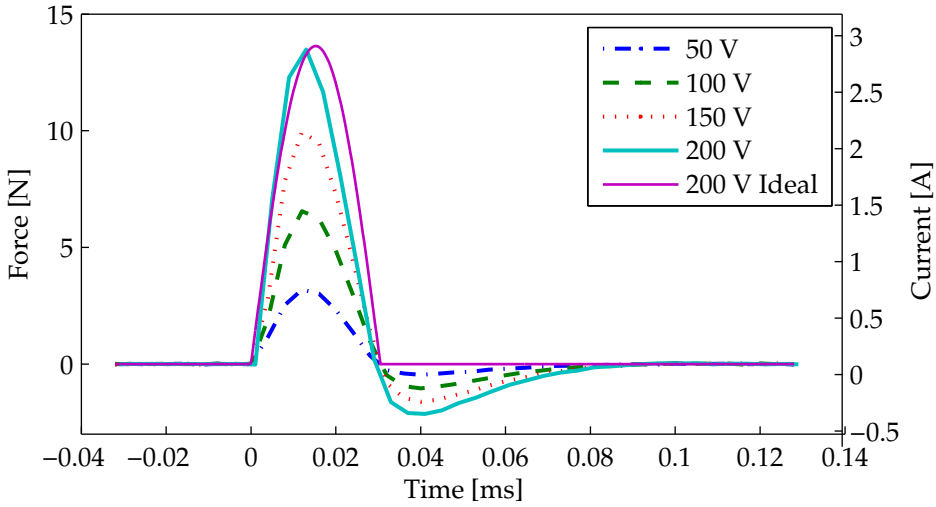
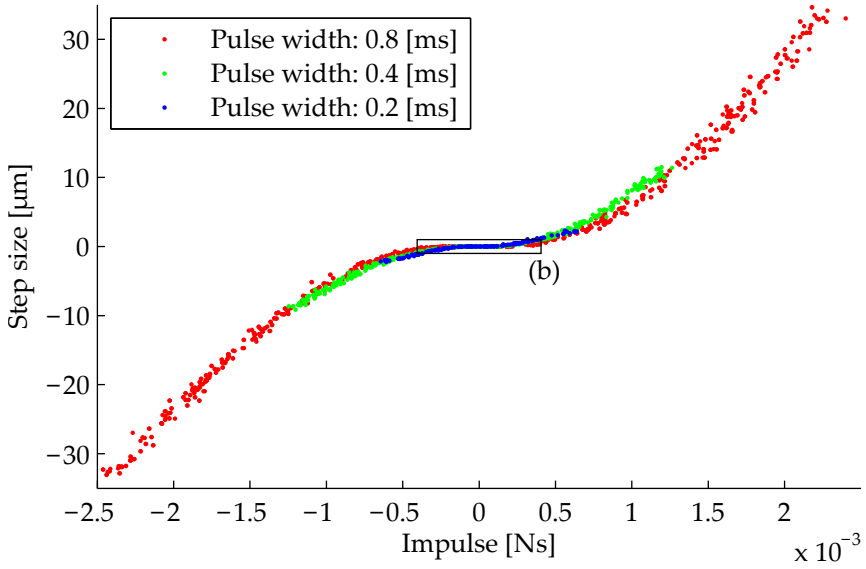


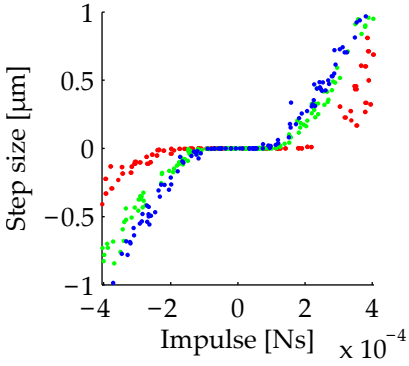
Figure 5.15.: Typical force (left scale) and current (right scale) profile in time of an electromagnetic force pulse generated using the current pulse generator circuit shown in Figure 5.14. The current is measured using a current probe at a sample frequency of 250 kHz. The force is derived from the current measurement. Comparing the 200 V pulse with the ideal (sinusoidal) pulse shape, the measured reverse current after the pulse becomes clear. This is caused by the non-ideal properties of components in the pulse generator circuit.

the stator and the number of windings) influence the motor constant (the current to force ratio) and its electrical properties (resistance and self-inductance). Additionally, these choices impose a required initial charging voltage of the capacitor, as the current pulse magnitude is influenced by this voltage, the capacitance C and self-inductance L.

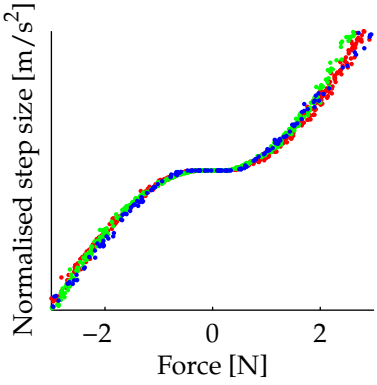
Realised current pulses Short force pulses were generated using a custom voice coil actuator and current pulse generator, designed such that a pulse duration of 0.03 ms, and a magnitude of at least 10 N can be achieved using no more than 200 V on the capacitor. The shape of these pulses is shown in Figure 5.15 for several values of the initial capacitor voltage. This pulse duration lies in the range of commonly encountered pulses from mechanical impact: an order of magnitude of 0.01 ms to 0.1 ms [Beek, 2015]. Longer pulses, in the 0.2 ms to 0.8 ms range, were generated using a current amplifier and a commercially available voice coil actuator. The motion that results from the various force pulses is discussed in the following section.



(a) Step size as a function of impulse.



(b) Magnification of the box in Figure 5.16a.



(c) Scaled step size as a function of pulse force.

Figure 5.16.: The displacement step-size variation of force pulses of different force pulse amplitudes and pulse widths (Figure 5.16a). The current pulses are generated using a current amplifier and have a duration of 0.2, 0.4 and 0.8 ms. The force was limited to ± 3 N. This experiment shows that the effective impulse determines the displacement magnitude. The bottom left graph (Figure 5.16b) is a magnification of the black box in the top graph and shows the differences in step size for small impulse. The bottom right graph (Figure 5.16c) is scaled to show that the force (on the x-axis), and not the impulse, determines whether the contact displaces.

5.6.3. Force pulse positioning experiments

This section provides an experimental investigation of the displacement steps that can be made using the electromagnetic force pulse method. All experiments were conducted on the single point friction measurement set-up that was introduced in Section 5.2 and which was also used for continuous force friction force experiments. The moving mass of this set-up is approximately 250 g and a normal force of 1.7 N and 1.0 N was used for the long and the short pulse experiments respectively.

Long force pulses Figure 5.16 shows the measured magnitude of the displacement steps due to force pulses with a pulse-width of 0.2, 0.4 and 0.8 ms, indicated in different colours. The combination of the current amplifier and the Lorentz actuator is able to generate forces up to a maximum of 3 N, which translates to an impulse range of $\pm 0.6 \cdot 10^{-3}$ Ns, $\pm 1.2 \cdot 10^{-3}$ Ns and $\pm 2.4 \cdot 10^{-3}$ Ns for the three different pulse widths. The displacement steps are shown in Figure 5.16a as a function of the measured impulse that caused them. At impulses larger than approximately $1 \cdot 10^{-3}$ Ns, the displacements follow a linear trend-line within 10 %, also for the different pulse widths. There is however an area around 0 Ns where the displacement steps indistinguishable from zero. The area around zero, indicated with a black rectangle, is shown magnified in Figure 5.16b. In this graph, more significant differences are visible between the different pulse widths. Further investigation of the data of this experiment shows that whereas the effective impulse (the impulse of the net force, shown as the coloured area in Figure 5.17) determines the displacement step magnitude above $1 \cdot 10^{-3}$ Ns, closer to zero impulse, the pulse force determines whether a force pulse results in a net displacement step or not. In this case motion starts around 0.3 N. This can be seen more clearly in Figure 5.16c, which shows the normalised step displacements as a function of the pulse force. The smallest reproducible displacement step that can be generated is around 200 nm.

Short force pulses The observations that the start of the motion is determined by the force pulse magnitude, while the displacement step is a function of the impulse, is confirmed by [Furst et al., 2010][5]. Making reliable displacement steps requires that the pulse force is well above the static friction limit. Small and reliable displacement steps are therefore generated by force pulses of a small impulse and a larger force. The pulse duration should therefore be short to achieve a small positioning resolution with a good reproducibility. This is also illustrated in Figure 5.17. The current pulse generator described in Section 5.6.2 produces force pulses of 0.03 ms. An example of the typical measured displacement during a series of these steps is shown in Figure 5.18. Each force pulse results in a 15 nm net displacement and a 50 nm overshoot. The pulse mainly excites the resonance of the mover mass and the contact stiffness, which is approximately 175 Hz.

The displacements of a series of 0.03 ms pulses of different amplitudes have been analysed. These results are shown in Figure 5.19. As the pulse duration and shape is constant, the impulse and the maximum force are proportional. As could

[5] He assumed the friction force to decrease during motion, but this is not necessarily true for very small steps or short pulses.

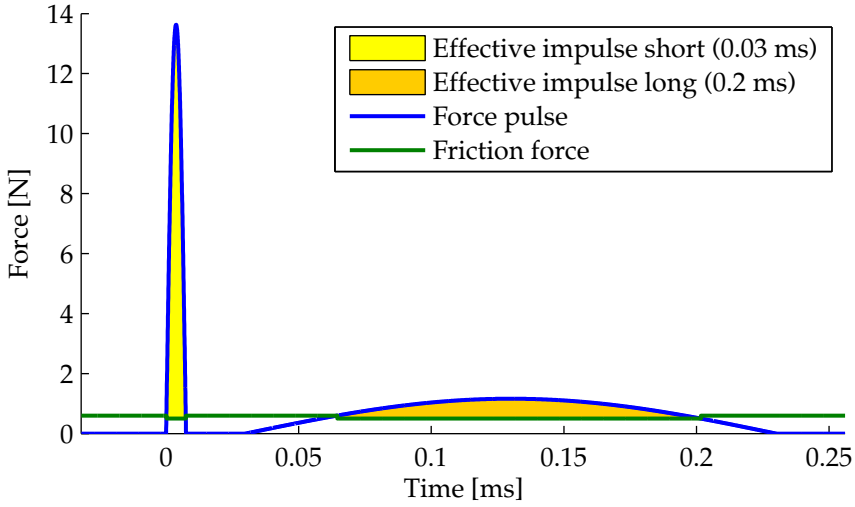


Figure 5.17.: Illustration of the interaction between a force pulse and the friction force for a short and a long pulse width. This graph shows that, when motion occurs if the force pulse magnitude is sufficient, and when the pulse displacement is determined by the impulse (the coloured areas in this graph), a large force and a small pulse width is recommended. The reason is that friction- and pulse-force variations have a relatively small impact on the impulse of the short pulse and thus on the displacement. The net impulse (the coloured area) of both force pulses is equal. Note that for this graph a simple static-kinetic friction coefficient was assumed. The real friction force variation during the pulse might be different. This graph is inspired on [Furst et al., 2010, Siebenhaar, 2004].

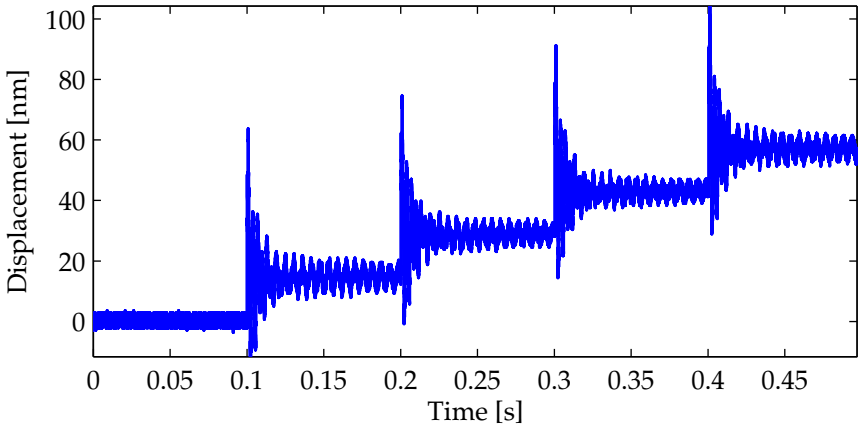


Figure 5.18.: Measured displacement during a series of 4, 4.5 N peak, 0.03 ms electromagnetic force pulses, resulting in results in a 15 nm net displacement with 50 nm overshoot.

be expected from the observations of the foregoing experiment, the displacement magnitude has decreased significantly by using shorter force pulses and reliable displacement steps smaller than 10 nm are possible. For even smaller displacement steps the relative uncertainty becomes larger as the force pulse approaches the friction limit. Using force pulses of automatically determined magnitude, it is however possible to achieve a positioning accuracy that is significantly better than the reproducibility of the displacement steps. It was shown experimentally that such a closed loop force pulse position control system is able to achieve a steady state positioning error smaller than 10 nm using the 0.2 ms pulses, an order of magnitude below the reproducible step size. Position feedback is required for this type of positioning, however only one displacement measurement is required after each force pulse.

Conclusion pulse actuation The experiments in this section show that force pulses can be used for permanent positioning of a friction contact with a resolution of 10 nm and smaller. A minimum force magnitude is required to initiate a step motion and the effective impulse (the impulse of part of the force pulse that is larger than this minimum) determines the magnitude of the step. The smallest reproducible steps are therefore obtained with short pulses. Short, 0.03 ms pulses can be generated using a matched Lorentz actuator and current pulse generator circuit. With these properties the magnetic force pulse actuator is a viable alternative to for instance piezo-intertia and physical impacts actuators, with the potential to generate even smaller displacement steps.

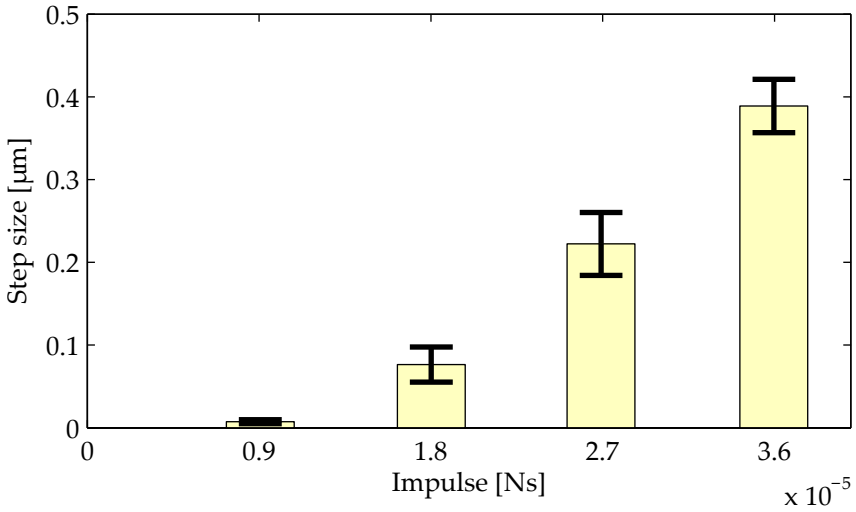


Figure 5.19.: Displacement due to short electromagnetic force pulses. The current pulses were generated using a dedicated current-pulse circuit (see Figure 5.14) and have a duration of 0.03 ms. The mean displacement step and its standard deviation during a single sequence of 180 steps is shown. The smallest step-size is 7 nm with a 2 nm standard deviation.

5.7. Discussion and conclusions

This chapter discussed the micrometre and nanometre scale displacements of friction contacts under shear load and showed how the shear load can be manipulated to accurately position a friction contact. This was all performed in the context of a sensor alignment application that requires sub-micrometre accuracy positioning and sub-nanometre stability after positioning. It was firstly identified which effects are important for the motion of friction contacts on this scale. Both literature and experiments have shown that the so-called pre-sliding friction regime plays an important role on this scale. It was therefore investigated how this behaviour can be used in a continuous, accurate positioning system. Additionally, another method to manipulate the displacement of a friction contact was explored: the use of electromagnetically generated force pulses.

For both positioning methods the focus was on the friction phenomena and on the methods that use these phenomena for positioning purposes. The implications for passive stability and the influence of the choice of materials was not explicitly treated.

Pre-sliding It is often practically assumed that a friction contact is motionless whenever the friction shear force is below a threshold value (mostly referred to as the static friction limit, $F_N \mu_s$). On the scale of micrometres and smaller however, a relative motion is also present at smaller contact shear forces. All these motions that

occur before full sliding are here referred to as pre-sliding. The two main effects are the contact stiffness and a contact creep effect. The contact stiffness is responsible for the elastic part of the motion, the contact creep is a plastic deformation. This plastic deformation is non-linear and history and time dependent. When the shear force on a contact is increased to the static friction limit, the elastic and the plastic motions are dominant at different force levels. The plastic displacement is also responsible for the gradual transition to full sliding that is seen both in literature and in experiments.

The elastic and the plastic pre-sliding behaviour of a friction contact can be explained by a bristle-type friction model. It describes a number of individual sub-contacts, originating from contact asperities. These sub-contacts each deform elastically, break, reconnect and gain strength depending on displacement and time. The motion of a friction contact can therefore be seen as a continuous process of breaking and forming sub-contacts and the rates at which these processes take place can be influenced by the contact shear force. This understanding of the friction phenomena at low velocities and small shear forces can be used to manipulate the displacement of a contact on a sub micrometre level. Additionally it provides insight in the friction behaviour of the Thermal Slider Actuator (TSA): The TSA contact motions are in a velocity- and displacement-range where pre-sliding is relevant and therefore stick-slip is absent in this actuator. This is beneficial for the TSA motion properties and stability.

Controlling pre-sliding As the pre-sliding friction regime allows a continuous motion (without stick-slip), it is suitable for accurate continuous positioning. A limitation of positioning within pre-sliding is that the motion velocity and the friction shear force have to be limited in order to avoid leaving the pre-sliding regime and going into full sliding. When the positioning task at hand is however not time critical, which is the case for the sensor alignment application described in Chapter 2, or when the required motion velocity is sufficiently low, this is not an issue.

It was show for a steel-steel friction contact that the pre-sliding velocity is in the nanometres per second range for forces that approach the static friction limit within several percent. This motion can already be used for open loop positioning at these low velocities. A reliable motion at larger velocities that is still within the pre-sliding regime, can be achieved by actively controlling the contact shear force. In order to maintain a well-controlled (non-elastic) pre-sliding motion, an integrating velocity controller is proposed and implemented. This positioning method results in a smooth motion for low velocities up to $10\,\mu m/s$. At larger velocities, the contact starts to operate around the transition to full sliding, which results in a significant increase in the velocity and force variations when using the same control parameters.

The integral velocity control loop can be used for closed loop positioning by converting the position error to a velocity set-point. This closes a position control loop around the velocity control loop. Using a proportional position-to-velocity conversion and a limit on the velocity set-point, a positioning accuracy in the order of 10 nm was achieved. However, the elastic displacement of the contact can introduce a significant displacement (200 nm order of magnitude) when the contact shear force is removed. Although this positioning accuracy is sufficient for the sensor alignment application (1 µm), it does require continuous position (or velocity) feedback. The

accuracy can be further improved by compensating for the contact elasticity. As the effective contact stiffness is a function of the contact history and therefore varies significantly, it has to be determined experimentally during positioning to achieve the best results.

Electromagnetic force pulse actuation Many positioning concepts exist where force pulses are applied to a friction contact for accurate positioning, however mostly piezoelectric components or mechanical mechanisms are required to generate the force pulse, or to store energy. It was demonstrated that these force pulses can be created directly using a voice coil actuator.

Experiments using a voice-coil set-up confirmed that the magnitude of the displacement step is determined by the net impulse, while the start of the motion is determined by the impulse force. It can then be derived that a short force pulse is beneficial for displacement steps that have to be reliable, small and insensitive to (friction) force variations. Short force pulses however require short current pulses, which require a relatively large instantaneous electrical power. A current pulse generator that uses a capacitor to store energy and a thyristor to start and stop the current at the right moments was implemented and it was shown that reproducible displacement steps down to 10 nm can be generated in open loop. When using force pulses in a closed loop positioning system, which only requires position feedback after each pulse, a 10 nm positioning accuracy is obtained using a pulse that only generates 200 nm repeatable steps. It is therefore concluded that the positioning accuracy of the force pulse concept is well within the sensor positioning requirements.

CHAPTER 6

Conclusions

In this thesis, several observations have been made with respect to the motion properties of sub-micrometre scale accuracy positioning methods, that allow for sub-nanometre passive position stability. These positioning methods are studied in the context of capacitive displacement sensor alignment in lithographic applications. The conclusions that have been drawn in the foregoing chapters about these methods are put in the perspective of this whole thesis in Section 6.1, treating the positioning requirements of the application, and the three novel solutions that were presented, the Thermal Slider Actuator (TSA), pre-sliding friction positioning and magnetic force pulse actuation, in more detail. Additionally Section 6.2 shows several recommendations for additional research that is likely to result in improved performance with respect to the positioning accuracy, the metrology application, and broader application of the presented solutions.

6.1. Thesis conclusions

This thesis has explored the motion properties of three new concepts that can combine positioning in at least three degrees of freedom on a sub-micrometre scale, with sub-nanometre per minute positioning stability after the positioning process is finished. This is required by the benchmark application of this thesis: precise positioning and alignment of capacitive displacement sensors in lithographic applications. It was shown that in order to perform capacitive displacement measurement with an accuracy smaller than 0.2 nanometre, the measurement electrode has to be positioned at a distance of 10 micrometre from the target electrode, with an alignment error that is smaller than 5 milliradians. In this application, the electrode is positioned and aligned when the lithography machine is being assembled, or offline for maintenance, therefore the time duration and the heat input of the positioning process are not critical.

Although the combination of stability and position-ability seems to be contradictory, friction contacts provide a solution, due to the strong non-linearity in relationship between displacement (-velocity) and force. If the shear force on a contact is small, the relative motion is often unobservably small as well. Also, the contacts can attach at virtually any location, so that a sufficiently high positioning resolution and stability can be achieved. The concepts that have been analysed, fit in a broader range of friction based positioning concepts. The following concepts

have been analysed, not only because they comply well to the requirements, but also because of their novelty.

6.1.1. The Thermal Slider Actuator (TSA)

The Thermal Slider Actuator (TSA) uses over-constrained clamping of the mover, so that variation of the internal forces can be used to generate relative motion. The internal forces are varied by means of heat inputs and thermal deformations. Using thermal actuation in a high precision system is possible because the heat input is entirely switched off and can even be disconnected after positioning, when stability becomes important.

Parameters It was shown that the TSA can be used for generating displacements in the three degrees of freedom that are required by the benchmark application. Although the basic principle of actuation with thermal cycles is always the same, many variants are possible. The influence of a selection of variables that is expected to have a large impact on the resulting motion velocity was shown both in simulation and experiment.

The grouping of the fingers, that is how many times a cycle is preformed in parallel on one actuator, is the easiest to influence. When the thermal expansion forces are large compared to static friction force, grouping has a significant influence on the step size and the average motion velocity. Many small groups lead to larger speeds. Small groups however also lead to unpredictable results, due to the relatively large influence of non-uniform friction behaviour, especially in the tilt directions.

A larger heating power leads to more thermal expansion and thus to larger steps. For fine positioning the heating power can be decreased, however, on approaching the earlier mentioned static friction limit, the motion will become more erratic due to the larger influence of friction non-uniformities. At the current scale and speed a maximum power of less than 5 watts is sufficient for repeatable steps, which means that the whole actuator can be controlled and powered over a USB connection.

Finally, it is shown that a larger step-time makes the real steps size approach the theoretical maximum. However, when the step time is significantly larger than the thermal time constant of a single finger, the increase is minimal and the average velocity decreases. Also the time-steps should not be reduced much further than the shortest thermal time constant, as this leads to relatively large displacement step variation and parasitic motions.

Position control It was also shown that the translational and the rotational motions can be combined using a simple control algorithm, to form a closed loop 3 DoF positioning system. Using a selection of thermal cycles, this method has shown to converge in the translation degree of freedom and in both rotational degrees of freedom simultaneously. Using this control loop, a standard deviations of 0.3 micrometres and 36 micro-radians was achieved. The positioning accuracy is thus sufficient for the sensor alignment application. Further improvements can be made to the translational accuracy by taking into account the displacement due to overall heating of the actuator. It was shown that especially in the translational

direction the measured displacement is sensitive to the previous thermal cycle. In the rotational degrees of freedom this issue was less significant.

Friction Friction is a crucial element in the functioning of the TSA. Due to the magnitude- and time-scale of the displacements of the friction contacts, the frictional behaviour of the TSA is governed by a creep-like process that is part of the pre-sliding friction regime. No convincing evidence was found that this friction process has a significant influence on the motion velocity of the TSA, at the observed scale. This specific frictional behaviour does however explain the absence of any stick-slip effects. Also this pre-sliding creep reduces the internal stresses over time, which improves the position stability.

6.1.2. The pre-sliding friction actuator

The transition of a friction contact from stick to slip contains a regime in which the contact itself deforms both elastically and plastically. The plastic deformation of a contact can be used for positioning when the shear force on the contact is carefully controlled. Based on the steady state contact displacement behaviour, two ways in which the plastic pre-sliding displacement can be controlled are defined: an open loop stable, low velocity regime and a regime that is close to the slip instability. In the second regime the achievable velocities are larger, but a closed loop controller is required for accurate positioning.

Feed forward In the first regime plastic displacement can, to some extent, be produced feed-forward. It was shown that for polished steel-steel contacts, this principle can be used to achieve velocities up to 5 nanometres per second, given the experimental conditions.

Closed loop In the second regime closed loop control is necessary for stable and accurate positioning. It was shown that velocities ranging at least from 0.1 to 100 micrometres per second can be achieved using an integral velocity controller on polished steel-steel contacts. With this type of controller the position was shown to converge to the set-point position within 10 nanometres. A new 200 nanometres position error is however introduced when the actuation force is removed. This error is caused by the elastic part of the pre-sliding motion. The contact elasticity, or stiffness, is however history dependent, and therefore either an estimate or a measurement of the contact stiffness is required to successfully correct this error. The presence of this error does however not limit the applicability of the pre-sliding positioning method to the sensor alignment problem.

A qualitative model was used to understand the pre-sliding behaviour, however both continuous positioning methods do not require a complex friction model. This is possible because the controller mainly prevents the contact from reaching the unstable slip regime, for instance by limiting the velocity to 10 micrometres per second. This is possible because time is available for the alignment and positioning process in the capacitive sensor alignment application. A momentary positioning error is therefore allowed during the positioning phase.

6.1.3. The electromagnetic force pulse actuator

Force pulses are frequently used to manipulate the shear displacement of a friction contact, mostly by means of an inertial mass, driven by an actuator such as a piezoelectric element. The newly proposed actuator concept uses a voice coil actuator to directly apply this force pulse to the mover, as this increases the freedom in the shape of the force pulses that can be used, and removes the need for an inertial mass. The current pulses that are required to generate force pulses were created using a current amplifier (pulse widths of 0.2 milliseconds and longer), or by a dedicated current pulse generator that discharges a capacitor using a thyristor (pulse widths of 0.03 milliseconds).

It was confirmed that while the force magnitude of the pulse determines if a contact will initiate a movement, the impulse (in Newton-seconds) determines the displacement magnitude. It was concluded from this that the small displacement steps that are required for a high positioning resolution, can be generated most reliably using large, short force pulses. The minimal displacement steps that have been achieved reliably in open loop are 200 nanometres using the 0.2 milliseconds force pulses and 10 nanometres using 0.03 milliseconds pulses. In a closed loop positioning system a positioning accuracy of 10 nanometres was achieved using 0.2 milliseconds force pulses.

6.2. Research recommendations

The concepts that have been analysed in this thesis provide a proper solution for the sensor positioning task, however, when the positioning performance has to be further increased due to the expected developments of the positioning requirements in the lithography industry, or for a broader application of these positioning methods, it is advisable to further investigate the following.

Thermal Slider Actuator First of all, the step cycles of the Thermal Slider Actuator can be modified depending on the distance to the target. This can be easily implemented in the current control algorithm by including both larger and smaller power steps in the control algorithm. The speed at larger distances will then increase, while lower power and longer pauses will be used at shorter distances to reduce the step-size and the thermal errors, and therefore to make it possible to approach the target more closely.

When the predictability of the TSA becomes increasingly important, the initial conditions of the actuator in terms of temperature and internal stresses can be taken into account when the next cycle is selected. This can be achieved by a complete simulation of the thermo-mechanical system (including friction) for multiple cycle options, which is rather computational intensive, but it was also shown that taking the average heating power of the previous cycle already leads to a significant improvement of the prediction accuracy.

Pre-sliding For improved performance of the closed loop pre-sliding positioning concept, other control strategies can be implemented. However the key to

controlling a friction contact at this (sub) micrometre scale remains to use the behaviour of the contact itself while avoiding strong dependencies on the friction parameters. A robust compensation method for the error induced by the contact stiffness would be a first valuable improvement.

Other applications Finally the concepts described in this thesis were presented in the context of sensor alignment for the lens displacement sensor in a lithography machine. However, it is also worthwhile to consider a broader application range of these actuators. This can be within the lithography environment, for instance for quasi static sensor placement at other locations in the machine, or for positioning individual lenses or mirrors. But the semi-permanent object placement with good passive stability is also of interest for positioning optical elements in for instance experimental set-ups or space applications. For these alternative applications, also other degrees of freedom can be of interest. The TSA and the other friction manipulation concepts can be configured differently and can be stacked to change the (number of) degrees of freedom.

Appendices

APPENDIX A

Capacitance to digital number conversion

Next to the electromechanical part of the sensor system that is described in Chapter 2, also the electronic capacitance readout contributes to the resolution and stability of the capacitive displacement measurement system. [Nojdelov et al., 2007] proposes a capacitance to digital converter based on the Delta-Sigma (also referred to as Sigma-Delta) principle: a Charge Balancing Delta Sigma Modulator. ΔΣ (Delta Sigma) converters are frequently used for analogue to digital conversion and this variant converts the charge of the sensor capacitance C_{sens} into a bit-stream by selectively adding and subtracting the charge of a know capacitor C_{ref} at every clock cycle. This results in a value for the ratio C_{sens}/C_{ref}, which is smaller than zero. The measurement resolution is amongst others determined by the number of additions and the magnitude of C_{ref}. The stability strongly depends on the stability of C_{ref}. Both the resolution and the stability can be improved, by means of an electronic zoom-in and a resistive reference respectively. These two methods are briefly described in this appendix, containing the work that was done by Sha Xia, Ruimin Yang and Roumen Nojdelov within the current and the previous research project at the Delft University of Technology Electronic Instrumentation Laboratory (EIL). It is partially extracted from [Ven et al., 2012b].

Electronic zoom-in In order to improve the electronic capacitance readout resolution the number of clock cycles (and thus addition or subtraction steps) per conversion can be increased. This however comes at the cost of using more system resources or a decrease in dynamic range. Alternatively the resolution can be improved by reducing the reference capacitance C_{ref}. This capacitance can however not be smaller than the maximum effective value of the sensor capacitance. A solution can be found in the notion that the capacitance deviation is significantly smaller than the total capacitance, as the electrode stand-off distance is approximately 10 µm, while the required measurement range is smaller than ± 0.1 µm. This range relates to a capacitance of approximately 69.5 ± 0.7 pF. A capacitance measurement around an offset would therefore significantly reduce the measurement range and thus the required value of C_{ref}. This improves the resolution without a reduction of the dynamic performance.

Capacitance measurement with an offset can be achieved by introducing a third,

zoom-in capacitor C_{zoom}, of which the charge is effectively subtracted from the sensor capacitor charge. Such a technique was already mentioned by [Heerens, 1986], in the use of transformer bridge circuits for the measurement of relative capacitance. It can however also be directly implemented in the Charge Balancing Delta Sigma Modulator. Traditionally, in each clock cycle the following charges are added: $C_{sens} \pm C_{ref}$. In the adapted converter this addition becomes $(C_{sensor} - C_{zoom}) \pm C_{ref}$, where the part between brackets indicates the effectively measured capacitance. More details on this principle and the design of a converter that uses this principle is described in [Xia et al., 2010, Xia and Nihtianov, 2011, Ven et al., 2012b].

Resistive reference The stability of the conversion from the capacitance to a digital value is to a large extent determined by the stability of the reference capacitors, in this case the reference capacitor C_{ref} and the zoom-in capacitor C_{zoom}. Electronic capacitors are however not sufficiently stable on longer time-scales, due to their sensitivity to temperature and humidity. This issue can be addressed by using another element as a reference for the capacitor. [Yang et al., 2011] introduces a method that uses resistance and time as a reference to periodically calibrate the reference capacitors. It is however also possible to entirely eliminate the reference capacitors and use the resistive reference directly in the $\Delta\Sigma$ converter. In [Ven et al., 2012b, Yang and Nihtianov, 2013] the resistive reference is not used for periodic calibration of the reference capacitor, but as the reference itself.

When both the electronic zoom-in and the resistive reference are combined, the capacitance can be converted with a resolution larger that 20-bit within 10 µs. The resistive reference has a 0.5 $\mathrm{ppm/K}$ thermal stability.

APPENDIX B

TSA numerical model

In order to understand and describe the TSA motion as a function of its thermomechanical properties and its input signals, and to find suitable heater input cycles and control rules for motion in multiple degrees of freedom (DoF), the behaviour of three quantities is crucial: temperature, deformation and friction force. The input of the TSA is the heating power of each finger, which causes a changing temperature distribution. The temperature distribution is assumed to be only dependent on this heat input. The temperature distribution can therefore be solved, independently of the elasticity and the friction. It is thus implicitly assumed that the influence of the deformation and the friction on the temperature is negligible. This is a reasonable assumption as the mechanical dissipation is at least six orders of magnitude smaller than the electrical dissipation.

The solution of the thermal problem, the temperature distribution, is the direct input for the combined elastic and friction model. The elastic and frictional model have to be solved simultaneously as there is a strong interaction between them. However, when the model is used to simulate the closed loop positioning behaviour of the TSA, the heat input is varied depending on the displacement output, which partially couples the thermal and the mechanical problem. However, as the displacement is only evaluated at discrete time intervals, between the different cycles, the two problems can be solved independently in time-blocks. This effectively decouples the thermal and the mechanical part again. The two parts of the model are implemented numerically, with constant time steps in MATLAB. The modelling principles are described below.

B.1. Thermal modelling

The thermal part of the model consists of a number of lumped masses that are thermally coupled to each other. These thermal masses, referred to as elements of the model, represent the n_{fin} fingers, the mover, the base and the plate that connects the TSA to the environment, as is indicated in Figure B.1. The temperatures of these elements are written in a vector form:

$$\mathbf{T} = \begin{bmatrix} T_{\text{fin}\,1} & \cdots & T_{\text{fin}\,n_{\text{fin}}} & T_{\text{base}} & T_{\text{plate}} & T_{\text{mover}} \end{bmatrix}^{\mathsf{T}}$$

so that the temperatures as a function of time can be computed using the equation:

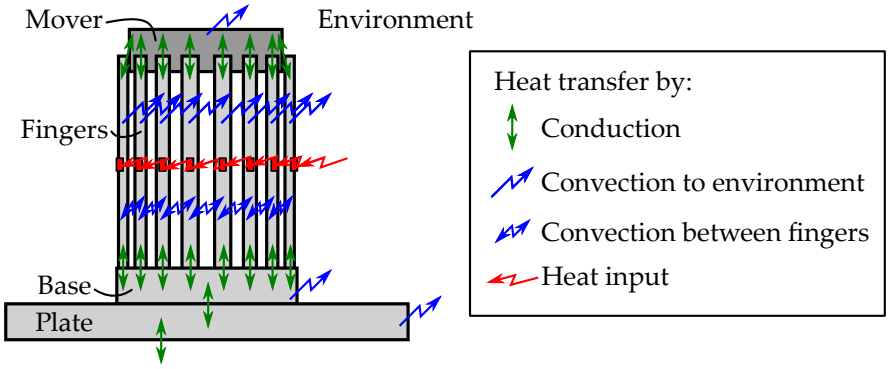


Figure B.1.: Thermal model definitions and heat flows.

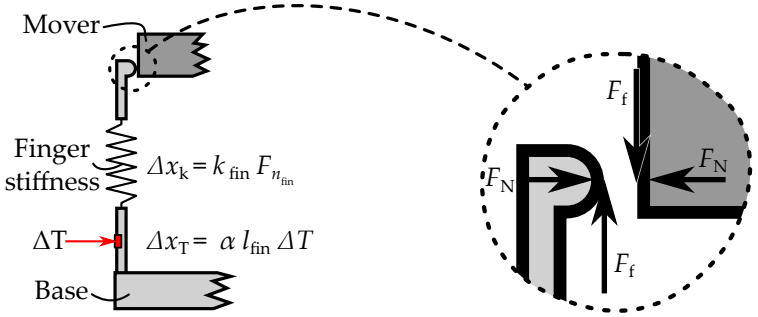


Figure B.2.: Mechanical model overview

$$\mathbf{T}(t_n) = \mathbf{T}(t_{n-1}) + (\mathbf{P}_{in} + \mathbf{K}_{th}\mathbf{T}(t_{n-1}))(t_n - t_{n-1})\mathbf{C}_{th}$$

where $\mathbf{C}_{th}$ is the thermal mass vector, $\mathbf{P}_{in}$ is the input power vector and $\mathbf{K}_{th}$ is the conductivity matrix. Heat conduction paths are indicated in Figure B.1. From the power input, the temperature distribution over time is solved and used as an input of the elastic-frictional model. Note that the number of elements in this thermal model is rather limited, but it is sufficient to describe most TSA behaviour. This choice does however influence the values of the thermal parameters that should be chosen for these elements. This is discussed in Appendix D.

B.2. Elastic-frictional modelling

The elastic part of the model contains the elasticity of the fingers. The (internal) accelerations of the TSA and the moving masses can safely be assumed to be small and therefore inertia effects are ignored. From the temperatures of the fingers T and

Figure B.3.: Axis definitions of the TSA in three degrees of freedom: tilt angles x and y and translation axis z.

the neutral length at ambient temperature $x_{0,T_{\text{amb}}}$, the neutral length of the fingers is computed:

$$x_{0,T} = x_{0,T_{\text{amb}}}(T - T_{\text{amb}})\alpha_{\text{fin}},$$

with α_{fin} being the thermal expansion coefficient of the finger. From the new neutral lengths $x_{0,T,n_{\text{fin}}}$ and the positions of the finger ends $x_{n_{\text{fin}}}$ in the previous time step, the initial compressive force of all fingers is computed.

$$F_{n_{\text{fin}}} = k_{\text{fin}}\left(x_{n_{\text{fin}}} - x_{0,T,n_{\text{fin}}}\right),$$

where k_{fin} is the longitudinal stiffness of a finger. Then the elastic mover displacement, that is the displacement for which no movement of the friction contact is necessary, is computed. For the most common, circular finger configuration this is a motion in three degrees of freedom: one translation and two rotations around the axes orthogonal to the translation axis, as is indicated in Figure B.3. The force that remains due to the over-constrained configuration, after the elastic motion, is the contact shear force, which may cause a relative motion of the friction contact.

Describing the friction-motion relation can be done in many different ways, as was pointed out in Section 3.2. An important observation in Section 4.1.3 was however that discrete displacement changes were never observed in the TSA prototypes and demonstrators. This implies that the sudden decrease in the friction force that is expected due to the stick-slip transition, does not occur in the friction contacts of the TSA at the expected scale. A suitable friction description accordingly shows no (sudden) decrease in the friction force for an increasing contact velocity. The descriptions that have lead to realistic results are: constant friction coefficient, (offset)-linear damping and exponential damping. A constant friction coefficient is used in the results presented in this appendix.

The static friction model displaces the contact a distance Δx_{slip} when the contact shear force of a finger $F_{n_{\text{fin}}}$, exceeds the static friction limit $\mu_{\text{s}} F_{\text{N}}$. Δx_{slip} has a

magnitude such that the contact shear force is reduced to the static friction limit $\mu_s F_N$:

$$\left|\Delta x_{slip}\right| = \begin{cases} \frac{F_{n_{fin}} - \mu_s F_N}{k_{fin}} & : F_{n_{fin}} > \mu_s F_N \\ 0 \end{cases}.$$

The offset linear damping model moves the contact point with a velocity proportional to the difference between the contact shear force and a force limit F_{lim}:

$$\left|v_{slip}\right| = \begin{cases} \frac{F_{n_{fin}} - F_{lim}}{c} & : F_{n_{fin}} > F_{lim} \\ 0 \end{cases}.$$

The offset in this description F_{lim} originates from the experimental observation that a (significant) relative motion becomes visible only when the force reaches a certain value. Effectively this is a combination of static friction and linear damping.

Several sources [Burwell and Rabinowicz, 1953, Baumberger et al., 1994, Heslot et al., 1994] find exponential relations between contact shear force and sliding or creep velocity. That means that the slip-velocity has the form:

$$\left|v_{slip}\right| = e^{a \cdot F_{n_{fin}} + b},$$

where a and b are friction parameters. The slip velocity v_{slip} of these descriptions is converted to a slip displacement Δx_{slip} by multiplying it with the simulation time step.

The contact displacement Δx_{slip} of one finger influences the contact forces of the other fingers and therefore also the force balance of the mover. The computational steps that compute the elastic motion and the slip displacement are therefore run iteratively for each time step. All friction forces then converge to the solution of that time step.

The output of this model is a description of the TSA temperatures and the mover displacements, but also the displacements of the fingers and the internal forces after this displacement. A selection of this data is graphically represented in Figure B.4.

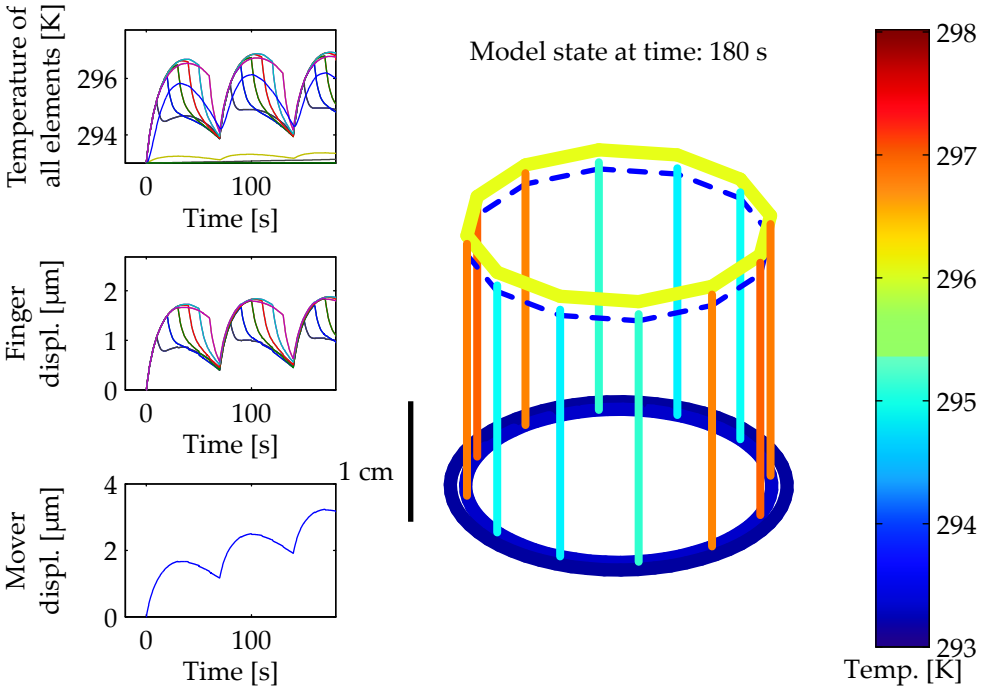


Figure B.4.: Graphical representation of a 12 finger TSA model. On the left side of this figure three separate graphs of respectively the temperatures of all elements in the model (the fingers, the mover, the base and the plate, see Figure B.1), the displacement of the fingers and the displacement of the mover are shown. The schematic drawing in the centre shows the temperatures and deformation of the parts of the TSA, the colours represent their temperatures. Compare this graphic for reference to the picture of a TSA demonstrator in Figure B.3. The blue, dashed ring represents the initial location of the mover. All initial temperatures are 293 K and all displacements are scaled by a factor 500 in the drawing. In the current state, 6 fingers are heated (shown in orange/red, around 297 K). A thermal cycle with 2 sections and 6 groups is used (see Section 4.2 for details on grouping).

APPENDIX C

TSA displacement measurement

In order to evaluate the motion and positioning properties of the TSA, displacement measurement is required in up to three Degrees of Freedom (DoF). In this section a laser interferometric, a capacitive and a PSD measurement system will be discussed. The interferometric measurement uses the interference pattern from a laser to measure the displacement and tilt with a sub-nanometre and micro-radian resolution and a millimetre and 10 mrad range, but with a limited sample rate of a few Hertz and the need for significant offline data processing. The capacitive measurement set-up can be used at kilohertz bandwidths, however the measurement range in translation and tilt direction is limited to ± 25 µm and ± 3.8 mrad. Using a Photo Sensitive Detector (PSD) only the tilt angles can be measured, but in a range of ± 50 mrad.

Interferometric displacement and tilt measurement The interferometric measurement set-up, shown in Figure C.1 on the left side, was already available from earlier experiments. It was however mainly used for translation measurement. The interferometer works according to the Fizeau principle, as is illustrated in the same figure on the right side. Interference is generated by the difference in path length between the reflection from the mover and the bottom of the wedge. A fringe interference pattern, shown in Figure C.2a, is captured on the CCD. The phase of this pattern represents the mover displacement, while the spatial variation period represents the tilt angle. In order to achieve tilt-measurement with sufficient resolution, the spatial fringe frequencies have to be accurately determined.

The horizontal and vertical frequency seen in the interferogram are proportional to the two tilt-angles of the mover disk. These frequencies are determined using a 2-dimensional Fourier transform and an improved frequency interpolation algorithm that increases the tilt resolution from 0.2 mrad by at least a factor 10^3 in post-processing. This processing makes the shape of the tilt motion visible, which was otherwise significantly smaller than the measurement resolution. The displacement resolution after interpolation is better than 1 nm. The interferogram is shown in the frequency domain in Figure C.2b. The locations of the peaks in this image contain the tilt information.

Capacitive 3DoF measurement Although the interferometric measurement method is very suitable for performing displacement measurements in 3 Degrees

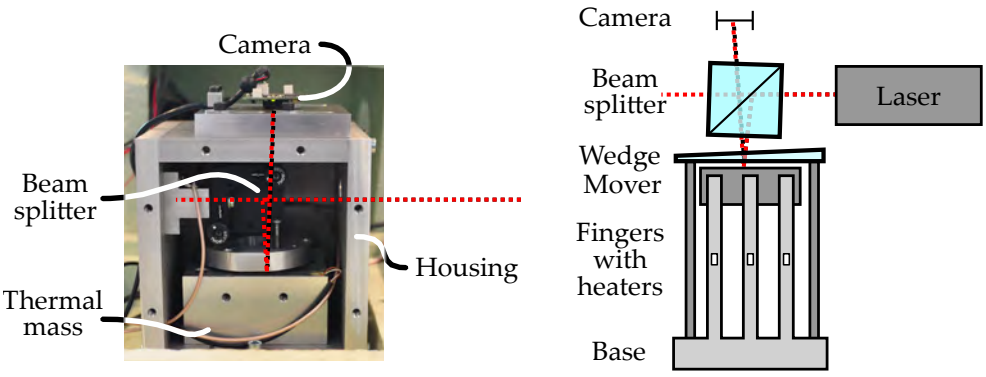
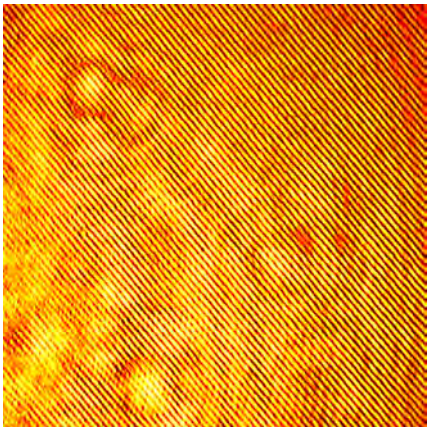
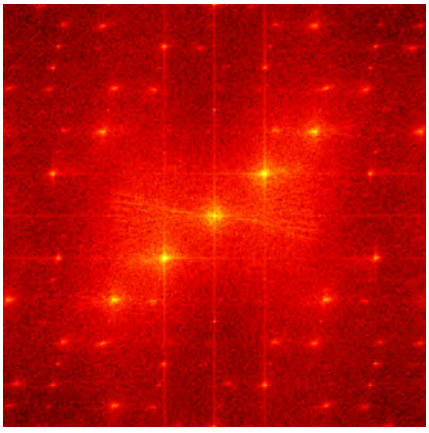


Figure C.1.: The Fizeau interferometer setup. (Setup by Jeroen van Schieveen.) The TSA is mounted within the thermal mass and is of the 16-finger type shown in Figure 4.5. The camera captures the interference pattern of the laser light (shown in Figure C.2a) reflected from the mover top-surface and the wedge bottom-surface. The spatial frequency and phase of the interference pattern are determined by the relative displacement and the tilt angles of the mover with respect to the wedge.



(a) Interferogram of which position and tilt are extracted, captured with the camera and set-up shown in Figure C.1. An initial tilt angle in the reference mirror causes the visible fringes (the sloped line pattern). Deviations in horizontal and vertical frequency correspond to tilt motion, whereas an overall phase change corresponds to a translational motion.



(b) Fourier transform of the interferogram. The locations of the peaks (indicated in yellow) contains information about the tilt angles of the mover. An interpolation process was used to increase the frequency resolution (here shown spatially) and increase the tilt measurement resolution. Note the increased resolution with respect to Figure C.2a.

Figure C.2.: Samples of the interferometer data used for tilt measurement.

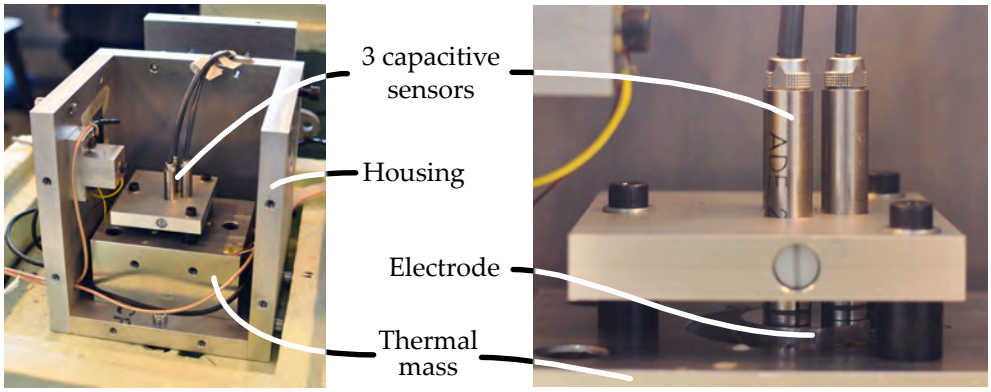


Figure C.3.: Pictures of the capacitive measurement set-up. On the left the aluminium housing and the block that contains the actuator. On the right a detail of the capacitive displacement sensors on top of the mover (in this case a steel electrode) that is clamped by the actuator. The sensors are placed on a circle of 13 mm diameter. The aluminium box was closed on all sides during measurements to reduce the influence of labs temperature variations.

of Freedom (DoF), the data processing needed to convert the interferograms into displacement data is relatively extensive and time consuming. Therefore, a capacitive measurement system is used for experiments that require real-time measurement. The three DoF can be measured independently by either using one electrode with three or more segments, or by three independent capacitive sensors. In this set-up, three separate, calibrated ADE 4810 capacitive probes with a range of ± 25 µm were placed equally spaced on a 13 mm diameter circle, resulting in a tilt-range of ± 3.8 mrad. Due to this small angular range, the capacitance error due to tilt can be safely neglected (see also Section 2.3) and a matrix multiplication suffices for the coordinate transformation from three translations to one translation and two rotations.

Next to the 3 DoF capacitive displacement measurement set-up, also the capacitive readout that was developed for high stability and resolution (see Appendix A and [Xia et al., 2010, Xia and Nihtianov, 2011, Ven et al., 2012b] for more information) has been used for single DoF displacement measurement.

PSD long-range tilt measurement Measurement of tilt angles over a range that is larger than of the interferometric and capacitive measurement set-ups was performed using a Position Sensitive Detector (PSD) as is depicted in Figure C.4. In this set-up the angle of the mover determines the direction in which the laser-beam is reflected and thus the position of the laser spot on the PSD. The resolution of this measurement method is smaller than 0.1 mrad.

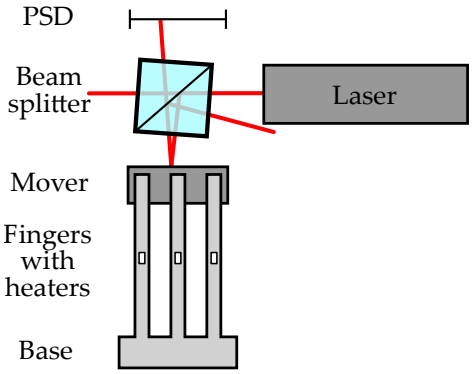


Figure C.4.: Angle measurement using a 2 dimensional PSD (Position Sensitive Detector). The actuator is mounted similarly to the interferometric set-up shown in Figure C.1. The tilt angle is determined from the position of the laser spot on the PSD.

APPENDIX D

TSA model validation

The purpose of the TSA model that is introduced in Appendix B, is to describe the basic behaviour of the actuator, so that it can be used to develop a control method with limited use of time-consuming experiments. In this section the fitness of the model for this purpose will be demonstrated by means of two combinations of measurement and simulation.

Model parameters The model parameters have to be selected for each actuator type. These parameters are friction properties and thermal capacities, resistances and expansion coefficients. They were initially estimated based on the actuator geometry and then adapted to match the behaviour shown in the measurement. The order of the model is kept as low as possible and as a consequence significant corrections had to be made to the initial parameters to find their effective values. They remain however within one order of magnitude. The corrections are especially significant for the thermal parameters (thermal masses and resistances) of the model, as the limited number of elements cannot accurately describe the temperature distribution within the finger and base elements. The fingers for instance use a single element each, while they show relatively large internal temperature differences due to their long and flat shape. The uniform temperature assumption is therefore not valid in the fingers. The model does however suffice to describe the basic behaviour of an actual TSA, although differences between the model and reality in the motion details can be expected. The model parameters used in this Appendix were chosen for the 12 finger TSA (Figure 4.4b).

Overall motion measurement (Figure D.1) The first validation measurement, shown in Figure D.1, contains a series of relatively slow (350 seconds per actuation cycle) translational motions, 10 step-cycles up and 10 down. The model parameters were chosen such that the amplitude and shape of the measured and modelled finger displacement match, as is shown in the middle graph. While the finger displacements do agree, the finger temperature amplitude (bottom graph) shows significant differences between the model and the measurement. This can be explained by the fact that the model assumes a uniform temperature of the finger, while this is not the case in reality. The temperature is measured close to the middle of the finger, resulting in a higher than average measured temperature.

The measured mover displacement (top graph) is strongly dependent on the

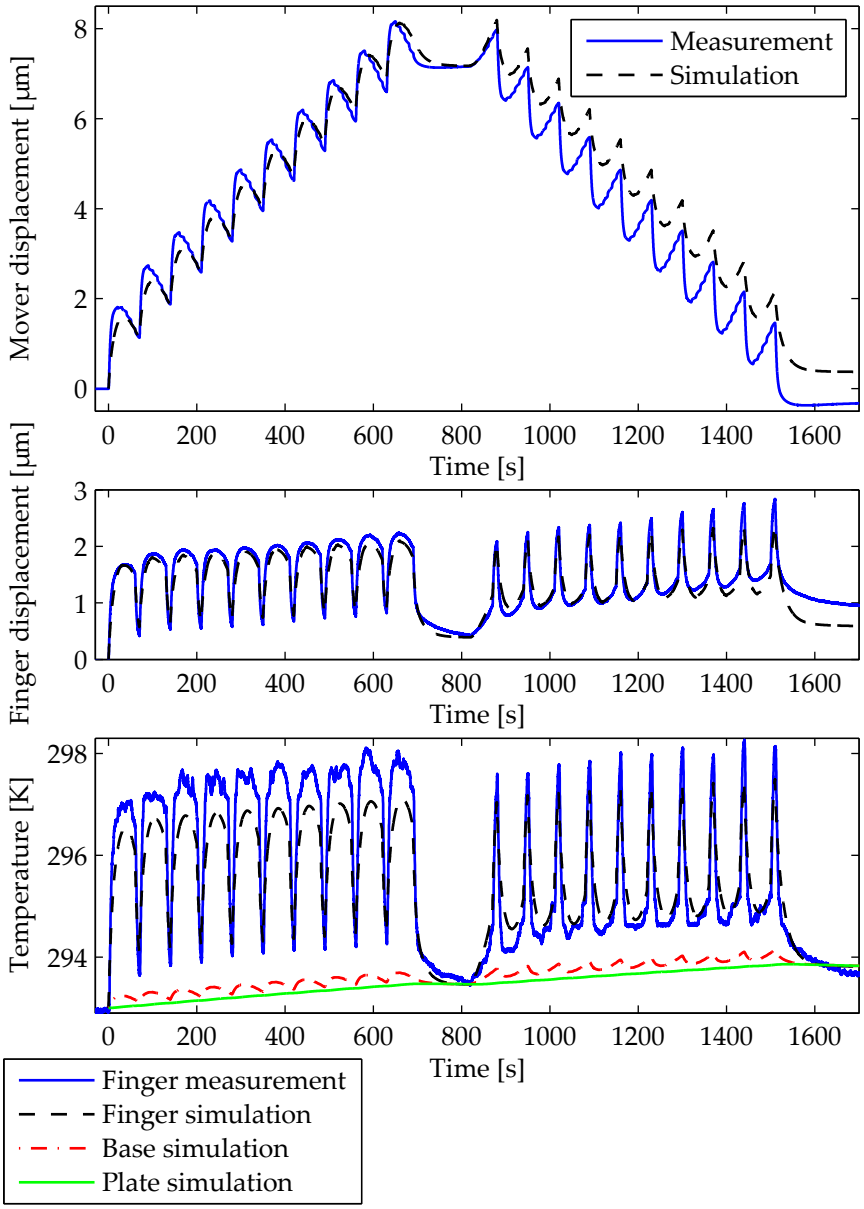


Figure D.1.: Measurement and simulation of the mover displacement, one finger displacement and several temperatures of the 12 finger TSA shown in Figure 4.4b. 10 upward cycles and 10 downward cycles are applied to the actuator. The 12 finger TSA is used in 2 sections of 6 groups (see Section 4.4.1 for more details). Each sub-step lasts 50 seconds, so that one full cycle takes 350 seconds. An additional 100 seconds pause is held between the up- and the downward cycles.

frictional behaviour. The friction model parameters were chosen such that the simulation matches the measurement in terms of the net displacement of the upward motion. Two discrepancies are noticeable here: Firstly the amplitude of the mover motion is slightly smaller in the simulation. This could imply that the constant friction coefficient assumption is not entirely valid. In Chapter 5 the contact friction process is discussed in more detail, but as the current model suffices as it shows the most typical actuator behaviour. Also it is assumed that all fingers behave identically, which is not necessarily the case in reality, mainly due to differences in the normal force and the effective friction coefficient. An additional discrepancy is the difference between the upward and the downward motion. This effect can be explained by the shape of the finger, which is shown in Figure B.2. The width of the fingertip causes a torque that changes the normal force depending on the shear force magnitude and direction. This effect is not included in the model, but it may cause a friction force discrepancy between the upward and the downward motion in the order of 10 %.

Cycle detail measurement (Figure D.2) Figure D.2 shows more details of the same TSA moving downward at a significantly shorter step-time of 70 seconds per cycle. When comparing the simulation (left column) and the measurement (right column) of the mover displacement (top row), the differences in the friction behaviour become clear. It is shown that the transition from stick to slip is more gradual in the measurement (box II) than in the model (box I). (The boxes are in the top graphs in Figure D.2.)

The specific cycle that is applied in this experiment, starts every cycle with a different group of fingers. The measurement of a single finger's temperature and displacement can therefore be used to give an impression of the displacement and temperature of all fingers. Ideally (in the model) all elastic steps within a cycle cause an equal mover displacement, as is shown in box I. In practice these steps are however not equal as can be seen in boxes II and III. Note that the place of these unequal displacements shifts within the cycle, that is, box III is positioned later in the cycle than box II, while the displacement graphs are almost identical. This shift corresponds to the shift of the fingers in the cycle, which implies that the difference in displacement is determined by the properties of the finger pairs, most likely the friction force differences between fingers.

The finger displacements and temperatures, (in the middle and bottom row respectively) do show significant differences between the model and the measurement. This especially holds for the longer thermal time constants. Their overall behaviour and the relationship between temperature and displacement however seems to be consistent. Overall it can be concluded that the model is able to represent the TSA sufficiently for the purpose of controller architecture development. Several aspects that describe the details of the behaviour of a specific actuator are not captured within this model, but these aspects are not required for the analysis of general TSA properties and the controller architecture.

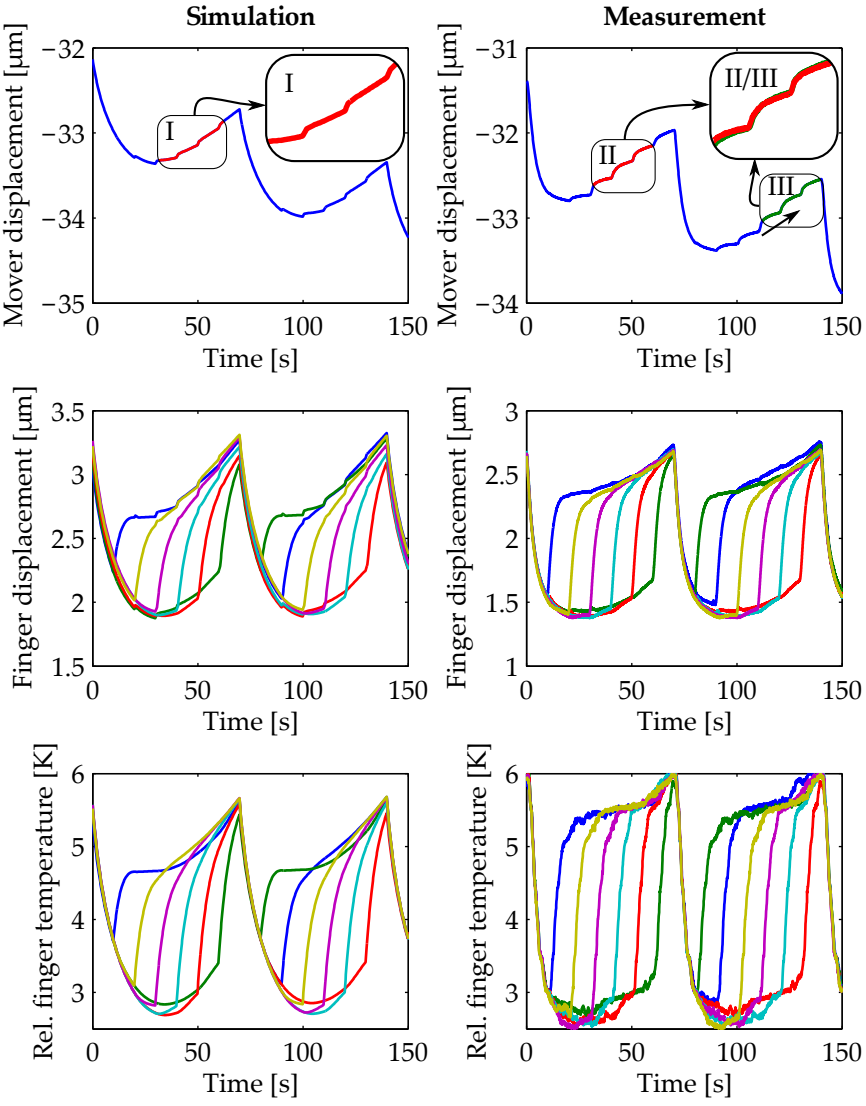


Figure D.2.: Simulation and measurement of two whole thermal cycles. The measurement is based on one finger only, differences between the different fingers of the TSA are therefore not reflected in the finger temperature and displacement measurement. Details of the motion cycles are shown in the measurement and simulation of the mover displacement, finger displacements and finger temperatures of the TSA shown in Figure 4.4b, in two downward thermal cycles. The 12 finger TSA is used in 2 sections of 6 groups (see Section 4.4.1 for more details). Each sub-step takes 10 seconds, so that one full cycle lasts 70 seconds.

Bibliography

[Al-Bender and Swevers, 2008] Al-Bender, F. and Swevers, J. (2008). Characterization of friction force dynamics. *Control Systems, IEEE*, 28(6):64–81.

[Arnold, 2008] Arnold, W. H. (2008). Towards 3nm overlay and critical dimension uniformity: an integrated error budget for double patterning lithography. In Levinson, H. J. and Dusa, M. V., editors, *Proc. of SPIE*.

[ASML, 2014] ASML (2014). Asml company website.

[Baumberger et al., 1995] Baumberger, T., Caroli, C., Perrin, B., and Ronsin, O. (1995). Nonlinear analysis of the stick-slip bifurcation in the creep-controlled regime of dry friction. *Phys. Rev. E*, 51:4005–10.

[Baumberger et al., 1994] Baumberger, T., Heslot, F., and Perrin, B. (1994). Crossover from creep to inertial motion in friction dynamics. *Nature*, 367(367):544–6.

[Beek, 2006] Beek, A. v. (2006). *Advanced engineering design: Lifetime performance and reliability*. Tribos.

[Beek, 2015] Beek, M. A. v. (2015). Electromagnetic force pulse actuator for positioning at nanometer level using stick-slip friction. Master's thesis, Delft University of Technology.

[Ben-David et al., 2010a] Ben-David, O., Cohen, G., and Fineberg, J. (2010a). The Dynamics of the Onset of Frictional Slip. *Science*, 330(6001):211–4.

[Ben-David and Fineberg, 2011] Ben-David, O. and Fineberg, J. (2011). Static friction coefficient is not a material constant. *Phys. Rev. Lett.*, 106:254301.

[Ben-David et al., 2010b] Ben-David, O., Rubinstein, S. M., and Fineberg, J. (2010b). Slip-stick and the evolution of frictional strength. *Nature*, 463:76–79.

[Berman et al., 1996] Berman, A. D., Ducker, W. A., and Israelachvili, J. N. (1996). Origin and characterization of different stick-slip friction mechanisms. *Langmuir*, 12(19):4559–63.

[Burwell and Rabinowicz, 1953] Burwell, J. T. and Rabinowicz, E. (1953). The nature of the coefficient of friction. *Journal of Applied Physics*, 24(2):136–9.

[Fleming, 2013] Fleming, A. J. (2013). A review of nanometer resolution position sensors: Operation and performance. *Sensors and Actuators A: Physical*, 190:106–26.

[Furst et al., 2010] Furst, S. J., Dow, T. A., Garrard, K., and Sohn, A. (2010). Automated part centering with impulse actuation. *Transactions of the ASME*, 132.

[Griffiths, 1999] Griffiths, D. J. (1999). *Introduction to electrodynamics; 3th ed.* Pearson, Boston, MA.

[Heerens, 1986] Heerens, W. C. (1986). Application of capacitance techniques in sensor design. *Journal of Physics E: Scientific Instruments*, 19(11):897–906.

[Heslot et al., 1994] Heslot, F., Baumberger, T., Perrin, B., Caroli, B., and Caroli, C. (1994). Creep, stick-slip, and dry-friction dynamics: Experiments and a heuristic model. *Phys. Rev. E*, 49(6):4973–88.

[Holterman and Groen, 2013] Holterman, J. and Groen, P. (2013). *An introduction to piezoelectric materials and applications*. Stichting Applied Piezo.

[ITRS, 2013] ITRS (2013). The international technology roadmap for semiconductors 2013 edition lithography summary. Technical report, The International Technology Roadmap for Semiconductors.

[Jansen Precision Engineering, 2014] Jansen Precision Engineering (2014). Displacement sensors - concept overview.

[Jenkin and Ewing, 1877] Jenkin, F. and Ewing, J. A. (1877). On friction between surfaces moving at low speeds. *Philosophical Transactions of the Royal Society of London*, 167:509–28.

[Jung and Gweon, 2000] Jung, H. and Gweon, D.-G. (2000). Creep characteristics of piezoelectric actuators. *Review of Scientific Instruments*, 71(4):1896–900.

[Khiat et al., 2012] Khiat, A., Spronck, J. W., Van Schieveen, J., Milosavljevic, S., Wei, J., Estevez, P., Sarro, P. M., and Staufer, U. (2012). Linear and rotational thermal micro-stepper motors. *Microelectron. Eng.*, 98:497–501.

[Liu and Li, 2010] Liu, Y.-T. and Li, B.-J. (2010). Precision positioning device using the combined piezo-vcm actuator with frictional constraint. *Precision Engineering*, 34(3):534–45.

[Liu et al., 2003] Liu, Y.-T., Toshiro, H., Toshiro, H., Fung, R.-F., and Fung, R.-F. (2003). A novel precision positioning table utilizing impact force of spring-mounted piezoelectric actuator - part i: Experimental design and results. *Precision Engineering*, 27:14–21.

[Lorenz et al., 2015] Lorenz, H., Schodel, R., and Beckert, E. (2015). Phase topography-based characterization of thermal effects on materials and joining techniques. *Aplied Optics*.

[Moore, 1965] Moore, G. (1965). Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–5. 1998 reprint.

[Muhs, 2005] Muhs, D. (2005). *Roloff/Matek machineonderdelen: normering, berekening, vormgeving; Tabellenboek*. Academic Service.

[Munnig Schmidt et al., 2014] Munnig Schmidt, R., Schitter, G., Rankers, A., and Van Eijk, J. (2014). *The Design of High Performance Mechatronics*. Delft University Press, 2nd revised edition.

[Munnig Schmidt et al., 2011] Munnig Schmidt, R., Schitter, G., and Van Eijk, J. (2011). *The Design of High Performance Mechatronics*. Delft University Press.

[Müser, 2008] Müser, M. (2008). How static is static friction? *Proceedings of the National Academy of Sciences*, 105(36):13187–8.

[Nihtianov, 2008] Nihtianov, S. (2008). Capacitive sensor surface quality considerations when measuring sub-nanometer displacement. *Proc., Electronics-ET' 08*, pages 15–8.

[Nojdelov et al., 2007] Nojdelov, R., Nihtianov, S., and van Doren, E. (2007). A fast charge-meter for interfacing capacitive sensors. In *AFRICON 2007*, pages 1–6.

[Olsson et al., 1998] Olsson, H., Åström, K., Canudas de Wit, C., Gäfvert, M., and Lischinsky, P. (1998). Friction models and friction compensation. *European journal of control*, 4:176–195.

[Ouwehand, 2015] Ouwehand, P. (2015). Novel nanometer precision planar positioning stage using pre-sliding friction. Master's thesis, Delft University of Technology.

[PI, 2008] PI (2008). *Piezoelectrics in Positioning*.

[Rabinowicz, 1951] Rabinowicz, E. (1951). The nature of the static and kinetic coefficients of friction. *Journal of applied physics*, 22(11):1373–9.

[Rabinowicz, 1958] Rabinowicz, E. (1958). The intrinsic variables affecting the stick-slip process. *Proceedings of the Physical Society*, 71(4):668–75.

[Rabinowicz, 1965] Rabinowicz, E. (1965). *Friction and wear of materials*. Wiley series on the science and technology of materials. Wiley.

[Rice and Ruina, 1983] Rice, J. R. and Ruina, A. L. (1983). Stability of steady frictional slipping. *Journal of Applied Mechanics*, 50(2):343–9.

[Ruina, 1983] Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research: Solid Earth*, 88(B12):10359–70.

[Sampson et al., 1943] Sampson, J. B., Morgan, F., Reed, D. W., and Muskat, M. (1943). Studies in lubrication: Xii. friction behavior during the slip portion of the stick-slip process. *Journal of Applied Physics*, 14(12):689–700.

[Schieveen and Spronck, 2009] Schieveen, J. P. v. and Spronck, J. W. (2009). Actuator and method for positioning an object.

[Schieveen and Spronck, 2010] Schieveen, J. P. v. and Spronck, J. W. (2010). Picometer resolution capacitive sensor for industrial applications. Technical report, TU Delft.

[Schieveen et al., 2009] Schieveen, J. P. v., Spronck, J. W., and Munnig Schmidt, R. H. (2009). Thermal stepper, positioning system for stable micro alignment. *euspen conference proceedings*, II:110–3. Thermal stepper: A high stability positioning system for micro adjustment.

[Schieveen et al., 2010] Schieveen, J. P. v., Spronck, J. W., and Munnig Schmidt, R. H. (2010). Integrated auto alignment and calibration for high resolution capacitive sensor systems. *euspen conference proceedings*, I:188–91.

[Schieveen et al., 2011] Schieveen, J. P. v., Yang, R., Nihtianov, S., and Spronck, J. W. (2011). Performance optimization of a self-alignment system for capacitive sensors. In *Mechatronics (ICM), 2011 IEEE International Conference on*, pages 648 –53.

[Scholz, 1990] Scholz, C. (1990). *The Mechanics of Earthquakes and Faulting*. Cambridge Earth Science Series. Cambridge University Press.

[Siebenhaar, 2004] Siebenhaar, C. (2004). Precise adjustment method using stroke impulse and friction. *Precision Engineering*, 28(2):194 – 203.

[Ven et al., 2014a] Ven, O. S. v. d., Ouwehand, P., Spronck, J. W., and Munnig Schmidt, R. H. (2014a). Controlled pre-sliding for precision positioning. *DSPE conference proceedings*, pages 117–9.

[Ven et al., 2013] Ven, O. S. v. d., Spronck, J. W., and Munnig Schmidt, R. H. (2013). 3 dof positioning of a capacitive measurement electrode using the thermal slider actuator. In *Proceedings ASPE 2013 Spring Topical Meeting MIT Laboratory for Manufacturing and Productivity Annual Summit*, volume 55, pages 58–63. ASPE.

[Ven et al., 2014b] Ven, O. S. v. d., Spronck, J. W., and Munnig Schmidt, R. H. (2014b). Friction controlled precision positioning. *Proceedings of the 14th International Conference of the European Society for Precision Engineering and Nanotechnology*, II:396–9.

[Ven et al., 2012a] Ven, O. S. v. d., Spronck, J. W., Nihtianov, S., and Munnig Schmidt, R. H. (2012a). Contact creep in a thermal micro actuation mechanism. *Proceedings of the 12th International Conference of the European Society for Precision Engineering and Nanotechnology*, 1:478–81.

[Ven et al., 2012b] Ven, O. S. v. d., Yang, R., Xia, S., Schieveen, J. P. v., Spronck, J. W., Munnig Schmidt, R. H., and Nihtianov, S. (2012b). Autonomous self-aligning and self-calibrating capacitive sensor system. In Kamel, M., Karray, F., and Hagras, H., editors, *Autonomous and Intelligent Systems*, Lecture Notes in Computer Science, pages 10–17. Springer Berlin Heidelberg.

[Wojewoda et al., 2008] Wojewoda, J., Stefański, A., Wiercigroch, M., and Kapitaniak, T. (2008). Hysteretic effects of dry friction: modelling and experimental studies. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1866):747–65.

[Wu and Tung, 2000] Wu, R.-H. and Tung, P.-C. (2000). Studies of stick-slip friction, presliding displacement, and hunting. *Journal of Dynamic Systems, Measurement, and Control*, 124(1):111–7.

[Wulp, 1997] Wulp, H. v. d. (1997). *Piezo-driven stages for nanopositioning with extreme stability: theoretical acpects and practical design considerations*. PhD thesis.

[Xia and Nihtianov, 2011] Xia, S. and Nihtianov, S. (2011). Capacitive sensor system for sub-nanometer displacement measurement. In *IEEE Sensors, 2011*, pages 1173–6.

[Xia et al., 2010] Xia, S., Schieveen, J. P. v., Nihtianov, S., and Spronck, J. W. (2010). Concept evaluation of a high performance self-aligning capacitive displacement sensor. In *IEEE International Conference on Industrial Technology (ICIT), 2010*, pages 1575–80.

[Yang et al., 2011] Yang, R., Fekri, A., Nihtianov, S., and Nojdelov, R. (2011). Qualification of a stable capacitive sensor interface based on capacitance-resistance comparison. In *IEEE Sensors, 2011*, pages 1181–4.

[Yang and Nihtianov, 2013] Yang, R. and Nihtianov, S. (2013). A time/resistor-referenced capacitive sensor interface for displacement measurement in the sub-nanometer range. In *IEEE International Symposium on Industrial Electronics (ISIE), 2013*, pages 1–5.

[Yang et al., 2010a] Yang, R., Schieveen, J. P. v., Nihtianov, S., and Spronck, J. W. (2010a). Electronic system for control of a thermally actuated alignment device. In *IEEE International Conference on Industrial Technology (ICIT), 2010*, pages 1581–6. IEEE.

[Yang et al., 2010b] Yang, R., Schieveen, J. P. v., Nihtianov, S., and Spronck, J. W. (2010b). Optimized low-power thermal stepper system for harsh and inaccessible environments. In *IECON 2010-36th Annual Conference on IEEE Industrial Electronics Society*, pages 1779–84. IEEE.

[Yang et al., 2008] Yang, Z., Zhang, H., and Marder, M. (2008). Dynamics of static friction between steel and silicon. *Proceedings of the National Academy of Sciences*, 105(36):13264–8.

Summary

A Dutch translation of this summary can be found on page 159.

Alignment systems typically combine two functions: positioning and fixation. This thesis describes how these two functions can be combined, by means of an application that originates from the lithography industry.

Photo-lithography Lithography is the step in the manufacturing process of electronic chips (Integrated Circuits, or ICs) that defines the shapes and sizes of the components on the chips, by projecting extremely well defined light patterns. As more components have to be fitted on these chips in order to make them faster and more energy efficient, the details on the chips, and with that the light patterns, have to become increasingly smaller, currently between 10 and 20 nanometres. To realise these extremely small dimensions, the machine that defines them has to be significantly more accurate than these dimensions. The allowed errors due to the different components in these machines are defined in an error budget. Individual displacement sensors in the machine have to be accurate up to 0.1 nanometres. This thesis provides a step in achieving this accuracy with capacitive displacement sensors.

Capacitive displacement measurement The measurement that forms the target application of this thesis, is of the displacement of the projection lens column with respect to the reference for all sensors: the metrology frame. When the measurement range is limited to a few micrometres, capacitive displacement measurement provides a cost-effective solution. These sensors determine the distance between two electrically conductive plates by means of their mutual capacitance C, which is defined as the ratio of the charge Q on either one of the plates and their potential difference ΔV. In approximation, for two close parallel plates, this is related to the permittivity of the dielectricum ε and the mutual distance x and the area A of these plates: $C \equiv Q/\Delta V \approx \varepsilon A/x$. This makes the sensor sensitivity inversely proportional to the distance squared, which means that a large sensitivity requires the plates to be mounted close together with sufficient accuracy. In this case the distance must be set to 10 ± 1 micrometres over the entire surface before the machine is used. Additionally the position stability during operation must be within 0.1 nanometres over a period of 2 minutes. The alignment (and the occasional re-alignment) of the sensor has to be performed after it is built into the lithography machine. The sensor is

then inaccessible for manual alignment and therefore an automatic alignment system is required.

Combining stability and alignment in one system seems to be contradictory, however, the fact that alignment and stability do not need to be achieved simultaneously provides an opportunity. When analysing the different fundamental principles that can be used to lock the relative position of an object, it appears that friction locking is the most feasible solution that allows locking and re-locking at the required sub-micrometre position resolution. Friction contacts can combine positioning and locking by exploiting the apparently simple engineering rule-of-thumb that the relative position of two objects that are in contact with each other is fixed as long as the force between them does not exceed a certain level, and that when the force is sufficiently large, relative motion is possible. Several actuation concepts that use these locking properties of friction contacts are known from literature. Most of these concepts use piezoelectrically deformable material. In these actuators, the deformable elements either cause large accelerations so that the friction force exceeds the friction limit, or they have an over-constrained structure so that internal forces in the structure can overcome friction. Next to the known friction working principles, the motion properties of three new actuation principles have been explored in this thesis.

1. The Thermal Slider Actuator (TSA) The piezoelectric concepts could provide a solution that meets the specifications, especially where the mover is not supported by the piezo elements, which would reduce the stability. However, piezoelectric solutions often require high voltage amplifiers and are relatively complex and expensive. The usage of thermal deformation provides a novel concept that could fulfil the requirements at lower cost. Using inertial forces is not feasible with thermal actuation due to the long thermal time constants, however the Thermal Slider Actuator (TSA), developed earlier at Delft University of Technology instead uses thermally induced internal forces in a spring-nest, which is a simple over-constrained mechanical structure that clamps the mover with a large number of fingers using friction. It was already shown that a translational motion (1 degree of freedom) can be generated by applying specific thermal cycles to the different fingers, but the motion properties in all 3 degrees of freedom (DoF) have to be known. These properties are explored using an analytical translation-approximation of the TSA, a thermo-mechanical model and measurements on several TSA demonstrator set-ups.

It becomes clear from the analytical TSA model that the magnitude of the temperature step and the way that the fingers are grouped have a significant influence on the step size and the average motion velocity. When comparing these results to those from the more advanced model and measurements of a real TSA, additional factors that influence its behaviour become clear. The influence of the variation of the friction force (caused by variations in the contact conditions and the normal force) and the thermal behaviour of the actuator and its surroundings are the most significant. The variation in the friction force causes parasitic displacements (in other than the intended direction). The non-ideal thermal properties cause both coupling of the temperatures of the different fingers, resulting in a reduced efficiency, and overall heat up of the system. The thermal cycles for actuation in the tilt

directions can be seen as a translation cycle in positive and negative direction on opposite sides of the actuator. The tilt cycles result in net tilt steps, both in simulation and in reality, although in practice also significant variations of the step sizes and parasitic motions are present.

TSA closed loop positioning Closed loop positioning with the TSA uses a predefined set of thermal cycles. The measured tilt and translational displacements of this set are used for positioning. A simple algorithm selects the most appropriate cycle to approach the positioning goal with a mover position measurement as input. Because different thermal cycles are executed subsequently, the variation in the initial conditions in terms of temperature and pre-load in the fingers causes displacement variations. When a cycle is repeatedly executed, the cycle behaviour and motion converge. In the case of position control different cycles are subsequently executed, so that in practice the step-size deviations can be as large as 20 %. Despite the TSA deviations, the selection algorithm is able to converge to a set-point within 0.3 micrometres and 35 microradians (1σ), which is within the requirements for the sensor alignment application. These results were obtained without taking into account that the average heating power of the previous cycle influences the net displacement of a cycle. Compensating this effect will further reduce the closed loop translation error.

TSA friction The friction of the contact between the fingers and the mover also significantly influences the behaviour of the TSA. Another part of the previously mentioned 'engineering knowledge' entails that in many cases the friction force in a contact suddenly decreases at the moment that a relative motion in the contacts starts. This sudden step is known as the stick-slip transition, an effect that is responsible for a the squealing sound of sliding contacts. This effect, which would cause the nanometre size displacement steps, is however not observed in the TSA. Instead, the pre-sliding effect, which contains a creep-like plastic displacement of a friction contact at forces below the static friction limit, governs the motion of the TSA.

Pre-sliding friction Two phenomena have been observed that apply to friction displacements in the micrometre region. The first observation is that, when the shear load on a contact is increased, the transition from stick to slip is not discontinuous, as would be expected from stick-slip theory, but seems to be a continuous process. The second effect is a characteristic memory length in the order of 10 micrometres, which is significantly smaller than the typical contact dimensions. The friction force only seems to change after travelling this distance. The fact that this typical displacement is also present as a contact creep distance between stick and slip of a contact, relates these two phenomena. This behaviour can be seen as a so-called pre-sliding effect. Pre-sliding is here defined as all the (elastic and plastic) deformation that occurs in a contact interface before unstable sliding.

The relevant pre-sliding effects can, at least qualitatively, be explained using a friction model that is based on the asperities of both interfacing materials. By attributing properties such as elasticity, time-dependent contact strengthening and break-strain to each asperity interface, all the aforementioned contact properties can

be described. The pre-sliding and sliding regime can be seen as stretching and continuously breaking and forming of contacts between asperities. Upon motion the contacts stretches elastically. The longer a contact remains attached, the stronger it becomes (that is: the break-strain increases). When the break limit is exceeded the elastic energy in the contact is dissipated, and it reattaches at another location, again gaining strength in time.

It follows from the observations and the qualitative model that, although the contact can break in an unstable manner (when the contacts are broken at a larger rate than they are strengthening), it is a continuous process that can be stabilised. This for instance holds for the TSA. The motions of the fingers of the TSA are sufficiently slow to make the rate at which the asperities (re)gain strength sufficient to compensate for the asperity bonds that are broken. This explains why the friction coefficient in the TSA can be assumed to be constant and why the earlier mentioned discrete steps will not occur and limit the resulting position stability. On the other hand, this behaviour also means that under small shear loads a transient, or possibly a continuous creep motion may be present. It is however unlikely that the continuous creep velocity will exceed the stability limit of the sensor application, due to the strong non-linearity between the motion velocity and the contact shear load.

2. Pre-sliding positioning A second consequence of the presence of the pre-sliding creep phenomenon at low forces and velocities, is that it must be possible to position a friction clamped mover with a resolution that is orders of magnitude smaller than the characteristic displacement of the friction contact. The creep motion is stable at low velocity and force due to an increasing contact resistance for increasing velocity. In this region the motion is easily controllable and to some extent possible using only feed forward control. At larger motion velocities the contact resistance decreases for increasing velocity. Closed loop control is then required to stabilise the contact dynamics and create a continuous motion.

Enforcing a friction contact with forces that do not lead to full sliding is seen as a new positioning concept for the sensor alignment task. A measurement set-up is designed and built in order to demonstrate that this method of accurate positioning is feasible, and to further investigate the micro scale motion in friction contacts. In this set-up a flat and a spherical object, both polished and cleaned steel, are in mutual contact, the normal force and the displacement are constantly monitored and the shear force is applied using a voice coil actuator.

Using the friction contact measurement set-up, both the open and the closed loop pre-sliding positioning modes are confirmed. In the stable area the open loop motion velocity is relatively small. When the transition to the unstable region is approached within several percent, continuous contact creep is observed with velocities in the nanometres per second range. A velocity control loop that only uses an integral action is used to achieve larger velocities. The controller action prevents that the contact enters the unstable friction region by reducing the contact shear force when the velocity increases. This controller has shown velocities ranging at least from 0.1 to 100 micrometres per second. At the larger velocities (in this case 20 micrometres per second and above) however, the force and velocity variations

become significant, indicating that the system operates around the instability. These experiments prove that it is possible to smoothly position a friction contact on sub-micrometre scale. With an additional closed loop position controller a 10 nanometre accuracy was achieved, however in order to maintain the position with this accuracy when the actuator is switched off, an improved estimation of the contact stiffness has to be implemented.

3. Magnetic force pulse actuation Instead of using the previously mentioned continuous force to accurately position a friction contact, also force pulses can be used to manipulate a friction contact. Most commercial applications use piezoelectric inertia actuators to apply these force pulses. A new method is proposed that eliminates the need for piezoelectric elements and inertial masses: force pulses are applied by an electromagnetic voice coil actuator and a current amplifier or a dedicated current pulse generator.

The experiments with this type of actuation have confirmed that, while the magnitude of a force pulse determines if the contact displaces, the displacement magnitude is a function of the impulse. Short force pulses are therefore favourable for generating small, well defined displacement steps. With 0.2 millisecond force pulses, reproducible displacement steps of 200 nanometres can be generated, while a closed loop positioning accuracy of 10 nanometres was achieved. Using shorter force pulses (0.03 milliseconds), the reproducible (open loop) steps size decreases to 10 nanometres.

Conclusion This thesis has investigated the motion properties of three novel concepts for sub-micrometre accuracy positioning, combined with passive fixation, using friction contacts. All concepts have the capability to position a capacitive measurement electrode sufficiently accurate to perform sub-nanometre resolution displacement measurements of the lens column of a lithography machine. After the alignment procedure the friction contacts of the actuator fixate the electrode sufficiently stable to perform this measurement.

The Thermal Slider Actuator uses thermally induced stress in an over-constrained mechanical structure to displace its friction contacts, and with that the electrode. The specific heating cycles determine to a large extent the motion properties: the net motion step in each of the three directions and the motion velocity. Due to thermal, mechanical and tribological variations within the actuator, significant parasitic motions occur. The positioning algorithm however, uses calibration data to select the appropriate positioning steps and converge to the position set-point.

A friction contact can also be positioned directly by manipulating the contact shear force using a voice coil actuator. With pre-sliding friction positioning the force is controlled by using elementary knowledge of the pre-sliding processes that govern low velocity friction contact displacements. The position can then be changed using feed forward control, although the range and velocity may be limited to avoid reaching the friction instability. When the pre-sliding velocity is controlled by means of feedback, larger speeds can be achieved. Also force pulses can be applied directly to a friction contact for accurate positioning. With both methods positioning accuracies in the 10 nanometre range are feasible.

Samenvatting

Dit is de Nederlandse vertaling van de Engelstalige samenvatting die begint op pagina 153.

Uitlijnsystemen combineren in het algemeen twee functies: positioneren en fixeren. Deze thesis beschrijft hoe deze twee functies kunnen worden gecombineerd aan de hand van een toepassing die afkomstig is uit de lithografie industrie.

Fotolithografie Lithografie is de stap in het productieproces van elektronische chips (ook wel geïntegreerde schakelingen of IC's genoemd) waarbij de vorm en de afmetingen van de componenten op de chip worden bepaald. Dit gebeurt door het projecteren van zeer goed gedefinieerde lichtpatronen. Omdat chips steeds meer componenten moeten bevatten om ze sneller en zuiniger te kunnen maken, worden de details op de chip en dus ook de lichtpatronen die daarvoor nodig zijn steeds kleiner. Op dit moment zijn deze details tussen de 10 en de 20 nanometer. Om deze extreem kleine details te kunnen maken, moet de machine waarmee dit gebeurt nog nauwkeuriger zijn. De groottes van de toegestane fouten van de verschillende componenten in de machine zijn verzameld in een error-budget. Voor individuele verplaatsingssensoren zijn nauwkeurigheden tot 0.1 nanometer nodig om het totale budget niet te overschrijden. Deze thesis is een stap naar het behalen van deze nauwkeurigheid met behulp van capacitieve verplaatsingssensoren.

Capacitieve verplaatsingsmeting De doeltoepassing voor deze thesis wordt gevormd door één specifieke meting: de positie van de projectielens ten opzichte van de meetreferentie voor alle sensoren: het metrologieframe. Omdat het meetbereik voor deze toepassing beperkt is tot een aantal micrometer, kunnen capacitieve sensoren worden gebruikt als een kosteneffectieve oplossing. Capacitieve sensoren meten de afstand tussen twee elektrisch geleidende platen door het meten van de capaciteit C. De capaciteit is de verhouding van de lading Q op elk van de platen en de het potentiaalverschil ΔV ertussen. Voor twee platen die dicht bij elkaar staan is de capaciteit afhankelijk van de diëlektrische constante ε van het materiaal tussen de platen, van de onderlinge afstand x en van het oppervlak A van de platen: $C \equiv Q/\Delta V \approx \varepsilon A/x$. De gevoeligheid van de sensor is hierdoor omgekeerd evenredig met de onderlinge afstand in het kwadraat. Dit betekent dat de platen dicht bij elkaar moeten worden geplaatst om een hoge meetgevoeligheid te krijgen. In dit geval moet de afstand tussen de platen over het hele oppervlak op 10 ± 1 micrometer zijn afgesteld voor de machine wordt gebruikt. Daarnaast mag de positie van de sensor tijdens gebruik niet meer dan 0.1 nanometer variëren in een periode van 2

minuten. Het (her)uitlijnen van de sensor moet gebeuren nadat deze is ingebouwd in de lithografiemachine en omdat de sensor niet toegankelijk is voor handmatige uitlijning is hiervoor een automatisch uitlijnsysteem nodig.

Het combineren van stabiliteit met een uitlijnmogelijkheid in één systeem lijkt tegenstrijdig, maar het feit dat de stabiliteitseis pas geldt na het uitlijnen biedt een mogelijkheid. Als we de manieren bekijken waarop de relatieve positie van een object kan worden vastgezet, zien we dat wrijvingscontacten de beste mogelijkheden bieden om een positie vast te zetten met een resolutie die kleiner is dan een micrometer, terwijl deze verbinding ook weer kan worden verbroken. Het combineren van positioneren en fixeren met behulp van wrijvingscontacten is gebaseerd op de vuistregel dat een wrijvingscontact stilstaat zolang de dwarskracht op het contact onder een bepaalde grens blijft, terwijl het contact wel kan bewegen bij een voldoende grote dwarskracht. Verschillende bekende positioneerconcepten maken gebruik van dit principe. Meestal gebruiken zij piëzo-elektrisch vervormbare materialen, die acceleraties veroorzaken die groot genoeg zijn om de wrijvingskracht te overwinnen. Ook kunnen ze een statisch overbepaald geheel vormen waarbij de interne krachten de wrijvingskracht kunnen overwinnen. Naast deze bekende werkingsprincipes, zijn de eigenschappen van een drietal nieuwe actuatieprincipes verkend in deze thesis.

1. De Thermal Slider Actuator (TSA) Hoewel piëzo-elektrische materialen de stabiliteit van een actuator negatief kunnen beïnvloeden, is het wel mogelijk om hiermee de gestelde specificaties te halen. Voor het gebruik van piëzo's zijn vaak ook hoogspanningsversterkers nodig en daardoor wordt de totale oplossing snel complex en duur. Door in plaats daarvan gebruik te maken van thermische uitzetting en een nieuw positioneerconcept, kunnen we voldoen aan de specificaties bij veel lagere kosten en een lagere complexiteit. De inertiakrachten van thermische actuatie zijn (op deze schaal) niet afdoende, maar de Thermal Slider Actuator (TSA), die in een eerder project aan de TU Delft is ontwikkeld, gebruikt in plaats daarvan de interne krachten in een verennest. Dit is een statisch overbepaald mechanisme dat een object op wrijving inklemt met een aantal vingers. Het was al aangetoond dat deze actuator een translatie kan maken door de temperatuur in de verschillende vingers in een specifiek patroon te variëren, maar om de capacitieve sensor uit te kunnen lijnen is een beweging in 3 vrijheidsgraden nodig. Deze actuatie-eigenschappen van de TSA zijn onderzocht met behulp van een analytische benadering van de beweging, een numeriek model en een aantal meetopstellingen.

Het analytische model van de TSA laat zien dat de grootte van de temperatuurstappen en de manier waarop de vingers gegroepeerd zijn een grote invloed heeft op de stapgrootte en de gemiddelde snelheid van de actuator. Wanneer deze resultaten worden vergeleken met die van het numerieke model en van de metingen, blijken er nog meer factoren van invloed te zijn op het gedrag. Vooral de variatie van de wrijvingskrachten (door variërende normaalkrachten en oppervlakte-eigenschappen) en de thermische eigenschappen van de actuator en zijn omgeving zijn hierbij van belang. Wanneer de wrijvingskrachten niet uniform verdeeld zijn, ontstaan er parasitaire bewegingen (dit zijn bewegingen in andere richtingen dan de bedoelde richting). Verschillende beperkingen in de thermische eigenschappen

van de actuator veroorzaken zowel een koppeling tussen de temperaturen van de verschillende vingers, waardoor de efficiëntie van de actuator kleiner wordt, als voor algehele opwarming van de actuator. Voor actuatie in de kantelrichtingen kan de thermische cyclus voor translatie in positieve richting op de ene kant en in negatieve richting op de andere kant van de actuator worden uitgevoerd. Volgens de theorie en simulaties geeft een dergelijke cyclus alleen een verplaatsing in de kantelrichting, maar in werkelijkheid zijn ook hier de parasitaire verplaatsingen en de variaties in de stapgroottes significant.

TSA gesloten lus positioneren Om met terugkoppeling te positioneren met de TSA, wordt een vaste set thermische cycli gebruikt. De verplaatsingen in de kantel- en translatierichtingen van deze cycli zijn vooraf gemeten en een simpel algoritme selecteert aan de hand van deze informatie en de huidige positie telkens de meest geschikte cyclus om naar het positioneerdoel toe te bewegen. Doordat continue verschillende cycli worden uitgevoerd, begint elke cyclus met een andere temperatuur en voorspanning van de vingers. Als een thermische cyclus verschillende keren wordt herhaald, convergeren het gedrag en de netto verplaatsing, maar omdat verschillende cycli door elkaar worden gebruikt bij het positioneren met de TSA, kunnen de stapgroottes tot wel 20 % afwijken. Ondanks de significante afwijkingen in de beweging van de TSA, convergeert de positie door het positioneeralgoritme binnen 0.3 micrometer en 35 microradiaal (1σ). Hiermee voldoet de TSA aan de eisen die worden gesteld door de sensoruitlijntoepassing. Hierbij is geen rekening gehouden met het opwarmen van de gehele actuator door het (gemiddelde) vermogen van de voorgaande cyclus. Als voor dit effect wordt gecompenseerd kan de positieoneerfout in translatierichting nog verder worden teruggebracht.

Wrijving in the TSA Het gedrag van de wrijvingscontacten heeft een grote invloed op de TSA. Een andere vuistregel met betrekking tot wrijving is dat wanneer een contact begint te bewegen, de wrijvingskracht plotseling afneemt. Deze stick-slip overgang is verantwoordelijk voor het piepende geluid van glijdende contacten. Het stick-slip effect zou verplaatsingsstappen op nanometerschaal moeten veroorzaken in de TSA, maar die zijn nooit waargenomen. Het is gebleken dat het wrijvingsgedrag in plaats daarvan door het pre-sliding effect wordt gedomineerd. Dit effect bevat onder andere kruipgedrag van wrijvingscontacten bij krachten onder de statische wrijvingsgrens.

Pre-sliding wrijving Er zijn twee fenomenen geïdentificeerd die van belang zijn voor het wrijvingsgedrag op micrometerschaal. Om te beginnen zie je dat als de dwarskracht op een contact wordt opgevoerd, de overgang van stick naar slip continue is, en niet discontinue zoals je zou verwachten vanuit de stick-slip theorie. Daarnaast heeft een wrijvingscontact een karakteristieke geheugenlengte van ongeveer 10 micrometer. De wrijvingskracht blijkt pas te kunnen veranderen na het afleggen van deze afstand. Deze twee effecten zijn aan elkaar gerelateerd doordat deze karakteristieke afstand ook terugkomt als de kruip-afstand tussen stick en slip. Dit gedrag kan worden gezien als een pre-sliding effect, waarbij pre-sliding

is gedefinieerd als alle verplaatsing in een contact, voordat het contact instabiel gaat glijden.

Het relevante pre-sliding-gedrag kan, tenminste kwalitatief, worden verklaard door een wrijvingsmodel dat gebaseerd is op de ruwheidspieken van de materialen die in contact zijn. Je kunt de genoemde pre-sliding effecten beschrijven door aan elk ruwheidscontact eigenschappen als elasticiteit, breekrek en het toenemen van de contactsterkte in de tijd toe te kennen. Zowel pre-sliding als sliding bestaan dan uit het vervormen en het continue breken en weer verbinden van de contacten tussen deze ruwheidspieken. Bij beweging rekken de verbindingen elastisch uit. Hoe langer een verbinding niet verbroken is, hoe sterker hij wordt. Als de breekrek overschreden wordt verdwijnt de elastische energie en wordt een nieuwe verbinding gevormd, op een nieuwe locatie.

Het blijkt zowel uit observaties in de praktijk als uit het kwalitatieve model dat, hoewel het breken van een contact een instabiel proces kan zijn (de verbindingen in het contact worden dan sneller verbroken dan ze worden gevormd), het wel een continue proces is. Daardoor kan dit proces worden gestabiliseerd. Dit geldt bijvoorbeeld voor de TSA, waarbij de bewegingen van de vingers voldoende traag zijn. De sterkte van de (nieuwe) wrijvingsverbindingen groeit dan voldoende snel om te compenseren voor de verbindingen die worden verbroken door het bewegen van het contact. Dit is ook de reden dat de wrijvingskracht in de TSA constant kan worden voorondersteld en dat de discrete stick-slip verplaatsingsstappen, die de positiestabiliteit zouden kunnen beperken, niet optreden. Aan de andere kant betekent de aanwezigheid van dit effect dat ook bij kleine belastingen een (continue) kruipbeweging in het contact kan optreden. Het is echter onwaarschijnlijk dat deze beweging een beperking vormt voor de stabiliteitseis van de sensortoepassing, doordat de bewegingssnelheid sterk niet-lineair afhangt van de aangebrachte dwarskracht.

2. Positioneren met pre-sliding Een tweede gevolg van de aanwezigheid van dit pre-sliding kruipfenomeen, is dat het mogelijk moet zijn om een wrijvingscontact te positioneren met een resolutie die ordes kleiner is dan de karakteristieke verplaatsing. De kruipbeweging is stabiel doordat bij lage snelheden de benodigde wrijvingskracht toeneemt voor een toenemende snelheid. Hierdoor is de beweging relatief makkelijk te regelen, tot op zekere hoogte zelfs zonder positieterugkoppeling. Voor hogere snelheden neemt de wrijvingsweerstand van het contact af bij een toenemende snelheid, waardoor een stabiele, continue beweging alleen kan worden bereikt door een extern stabiliserend (regel)systeem.

Het aanbrengen van krachten op een wrijvingscontact zonder dat dat leidt tot onstabiel glijden, wordt gezien als een nieuw positioneerconcept voor het sensoruitlijnprobleem. Om aan te tonen dat dit een haalbaar positioneerconcept is en om de bewegingen van wrijvingscontacten op microschaal verder te onderzoeken, is er een meetopstelling ontworpen en gebouwd. In deze opstelling maken een vlak en een bolvormig oppervlak, beide gepolijst en schoongemaakt staal, contact. De normaalkracht en de verplaatsing worden continue gemeten en een spreekspoel (voice coil actuator) oefent de dwarskracht uit op het contact.

Het is met deze wrijvingsopstelling aangetoond dat positioneren in het pre-sliding

regime zowel met als zonder positieterugkoppeling mogelijk is. In het stabiele wrijvingsgebied is de verplaatsingssnelheid relatief laag. Als de overgang naar het instabiele gebied binnen een paar procent is benaderd, kan een kruipsnelheid van enkele nanometers per seconde worden gehaald. Hogere snelheden zijn gehaald met behulp van een integrerende snelheidsregelaar. Deze regelaar voorkomt dat het wrijvingscontact het instabiele wrijvingsgebied bereikt door de dwarskracht te verkleinen als de snelheid te groot wordt. Hiermee kunnen in ieder geval snelheden tussen de 0.1 en de 100 micrometer per seconde worden gehaald. Voor hogere snelheden (in dit geval 20 micrometer per seconde en hoger) worden de kracht- en snelheidsvariaties wel steeds groter, doordat de wrijvingsinstabiliteit steeds dichter wordt benaderd. Deze experimenten hebben aangetoond dat het mogelijk is om een wrijvingscontact soepel te positioneren op een sub-micrometerschaal. Door een positieregelaar aan het systeem toe te voegen is een positioneernauwkeurigheid van 10 nanometer gehaald. Om de positie met deze nauwkeurigheid te behouden als de actuator wordt uitgeschakeld is het nog wel nodig om een betere inschatting van de momentane contactstijfheid te implementeren.

3. Magnetische krachtpulsactuatie In plaats van een continue variërende kracht, kunnen ook krachtpulsen worden gebruikt om een wrijvingscontact te manipuleren. In commerciële toepassing wordt meestal een piëzo-intertia concept gebruikt om deze pulsen te maken. In een nieuw concept dat hier is geïntroduceerd zijn geen piëzo-elektrische actuatoren en reactiemassa's meer nodig, doordat de krachtpulsen direct worden aangebracht met de combinatie van een elektromagnetische spreekspoel en een stroomversterker of een speciale stroompulsgenerator.

De experimenten met dit type actuatie hebben aangetoond dat de grootte van de krachtpuls bepaalt of een contact wel of niet gaat bewegen, terwijl de grootte van de beweging een functie is van de (kracht)stoot van de puls. Dit betekent dat korte krachtpulsen het meest geschikt zijn voor het maken van kleine, maar goed gedefinieerde stappen. Met krachtpulsen van 0.2 milliseconde is een reproduceerbare stapgrootte van 200 nanometer bereikt, terwijl met positieterugkoppeling een nauwkeurigheid van 10 nanometer is gehaald. Door gebruik te maken van kortere krachtpulsen kan de reproduceerbare stapgrootte (zonder terugkoppeling) worden teruggebracht tot 10 nanometer.

Conclusie In deze thesis is de beweging van drie nieuwe positieoneerconcepten onderzocht waarmee sub-micrometer nauwkeurigheid kan worden gecombineerd met passieve fixatie met behulp van wrijvingscontacten. Het is met al deze concepten mogelijk om een capacitieve meetelektrode voldoende nauwkeurig uit te lijnen om er sub-nanometer resolutie metingen aan de lens van een lithografiemachine mee te kunnen doen. Na de uitlijnprocedure fixeren de wrijwingscontacten van de actuator de elektrode voldoende stabiel om deze meting uit te voeren.

De Thermal Slider Actuator brengt thermische spanning aan in een overbepaald mechanisme om de wrijvingscontacten, en daarmee ook de elektrode, te bewegen. De thermische cycli die worden aangebracht bepalen grotendeels de netto verplaatsing in drie graden van vrijheid en de bewegingssnelheid, maar door

verschillende thermische, mechanische en tribologische afwijkingen in de actuator ontstaan er ook parasitaire bewegingen. Deze bewegingen worden door het positioneeralgoritme meegenomen in de selectie van juiste cyclus om daarmee het positioneerdoel te bereiken.

Een andere mogelijkheid voor het positioneren van een wrijvingscontact is het direct aanbrengen van dwarskrachten, bijvoorbeeld door middel van een spreekspoel actuator. Pre-sliding positioneren maakt gebruik van elementaire kennis van het pre-sliding wrijvingsgedrag om de juiste kracht aan te brengen voor het bereiken van een nieuwe positie. Dit is mogelijk zonder positieterugkoppeling, maar om de wrijvingsinstabiliteit te beperken moet de bewegingssnelheid dan fors worden beperkt. Als met behulp van een meting de bewegingssnelheid wordt geregeld, kunnen significant hogere snelheden worden behaald. Ook met krachtpulsen is het mogelijk om de positie van een wrijvingscontact te manipuleren. Met beide methoden is het mogelijk om een nauwkeurigheid van ongeveer 10 nanometer te bereiken.

Curriculum Vitae

Ir. O.S. (Oscar) van de Ven was born on April 14th 1987 in Dordrecht, The Netherlands. After finishing his pre-univertisty education (in Dutch VWO) in Dordrecht in 2005, he started his Bachelor of Science in Mechanical Engineering at Delft University of Technology, The Netherlands. He obtained his BSc degree (with distinction) in 2008. His final BSc project was the development of a shape-lock mechanism for use in a medical application.

During the Mechatronic System Design specialisation of the Mechanical engineering Master of Science program, he also followed the university Honours Track program. He received his MSc degree in 2011 (with distinction), with a thesis project on the development of the actuator and sensor-system of a novel 5 degree of freedom haptic parallel robot. This thesis was awarded the Young Talent Tata Steel Graduation Price for Mechanical and Materials Engineering by the Royal Dutch Society of Sciences. He was then invited to start a PhD project at the Precision and Microsystems Engineering department of Delft University of Technology, to continue the development of a high precision sensor alignment system. This thesis is the result of that project.

During both his MSc and his PhD projects he contributed to the Mechatronic System Design MSc course. He worked on the lecture notes, in preparation for the course book [Munnig Schmidt et al., 2011], and on several homework assignments. Next to working on this, and other (practical) courses, he guided several MSc students during their final projects.

List of publications

1. Ven, O. S. van de, Spronck, J. W., Nihtianov, S., and Munnig Schmidt, R. H. (2012). Contact creep in a thermal micro actuation mechanism. *Proceedings of the 12th International Conference of the European Society for Precision Engineering and Nanotechnology*, 1:478–81. [Ven et al., 2012a]

2. Ven, O. S. van de, Yang, R., Xia, S., Schieveen, J. P. van, Spronck, J. W., Munnig Schmidt, R. H., and Nihtianov, S. (2012). Autonomous self-aligning and self-calibrating capacitive sensor system. *In Kamel, M., Karray, F., and Hagras, H., editors, Autonomous and Intelligent Systems, Lecture Notes in Computer Science*, pages 10–7. Springer Berlin Heidelberg. [Ven et al., 2012b]

3. Ven, O. S. van de, Spronck, J. W., and Munnig Schmidt, R. H. (2013). 3 dof positioning of a capacitive measurement electrode using the thermal slider actuator. *In Proceedings ASPE 2013 Spring Topical Meeting MIT Laboratory for Manufacturing and Productivity Annual Summit, volume 55*, pages 58–63. ASPE. [Ven et al., 2013]

4. Ven, O. S. van de, Spronck, J. W., and Munnig Schmidt, R. H. (2014). Friction controlled precision positioning. *Proceedings of the 14th International Conference of the European Society for Precision Engineering and Nanotechnology*, II:396–9. [Ven et al., 2014b]

5. Ven, O. S. van de, Ouwehand, P., Spronck, J. W., and Munnig Schmidt, R. H. (2014). Controlled pre-sliding for precision positioning. *DSPE conference proceedings*, pages 117–9. [Ven et al., 2014a]

6. Ven, O. S. van de, Vogel, J. G., Yang, R., Xia, S., Schieveen, J. P., Spronck, J. W., and Nihtianov, S (2016). Autonomous Self-Aligning and Self-Calibrating Capacitive Sensor System for Displacement Measurement in Inaccessible Industrial Environments. *In preparation for IEEE Transations on Industrial Electronics.*

Active positioning and passive fixation using friction in capacitive displacement measurement applications

Sensor positioning is critical for achieving a measurement accuracy in the picometre range with capacitive displacement sensors. This thesis discusses the contradictory requirements that the sensor alignment application imposes: motion capability combined with extreme displacement stability. In addition, it explores three micrometre scale positioning methods that do not compromise the position stability.