DIPLOMARBEIT

INVESTIGATION AND ESTIMATION OF UNCERTAINTIES IN SCANNING PROBE MICROSCOPY

Thomas Zeinzinger

Anerkannt am
Institut für Metallkunde und Werkstoffprüfung
Montanuniversität Leoben
Leoben, im November 1998

Preface

This work has been carried out at the Center for Electron Microscopy, Tampere University of Technology, during the periods of June 1997 to October 1997 and July 1998 to November 1998. Financial support came from the Institute of Materials Science and the Pentti O. Kettunen Foundation.

I want to express my gratitude to Prof. Veli-Tapani Kuokkala and Prof. Toivo Lepistö, for their support and advice. Especially I am very grateful for their almost instantaneous corrections and feedback when writing the work.
Acknowledgment is due to Univ.-Prof. Dipl.-Ing. Dr.mont. Albert C. Kneissl for accepting to be my supervisor at my home university, the University of Leoben, Austria. I would like to express my gratitude also to O. Univ.-Prof. Dipl.-Ing. Dr.mont. Dr.h.c. Franz Jeglitsch, the head of the Department of Physical Metallurgy and Materials Testing at the University of Leoben, for his permission to carry out this work in Tampere.

Peter van Lierop and the technical staff from ThisScientific B.V. and Peter Doppler from Anton Paar GmbH deserve my gratitude for the technical support and the fast repair of their devices.

Special thanks are due to my coworkers at the Center for Electron Microscopy and the Institute for Materials Science for being helpful all the time and creating a fruitful atmosphere. Especially, I wish to thank Kati Rissa and Pirjo Virtanen for their advice during my first steps using the Atomic Force Microscope and Terho Kaasalainen for his elaborate assistance in the Heavy-Lab. Juha Toiminen deserves special thanks for his help in various computer related things. Mari Mäntylä, Tuomas Hakala, Jussi Laurila and Jaakko Keränen I would like to thank for the great atmosphere they created.

Thanks to my friends in Finland and at home to be there for me. My special thanks are due to Gabor who provided a host for the last two weeks and helped me drawing some figures. Imre was a great help with particular C programming problems. Also Agi, Kimmo, Sanna, Santeri, Pit and Stefan earn thanks for the nice time I spent with them and for their help.

Last but not least, I owe warm thanks to my family to provide me the possibility to study. My special thanks to Sandra for her time she spent during the fruitful discussions, her encouragement and her support.

Thomas Zeinzinger
15. 11. 1998, Tampere

Table of Contents

II Experimental Part

Abbreviations

AC	Alternating Current
AFM	Atomic Force Microscope
CEM	Center for Electron Microscopy
DC	Direct Current
DI	Digital Instruments
EBSD	Electron Backscattered Diffraction
EDS	Energy Dispersive Spectroscopy
HOPG	Highly Oriented Pyrolytic Graphite
LED	Laser Emitting Diode
LPCVD	Low Pressure Chemical Vapor Deposition
LSB	Least Significant Bit
MOCVD	Metal-Organic Chemical Vapor Deposition
MS	Microsoft
PS	Postscript
PZT	Lead Zirconium Titanate
SEM	Scanning Electron Microscope
SPM	Scanning Probe Microscope
STM	Scanning Tunneling Microscope
TIFF	Tag Image File Format
TUT	Tampere University of Technology
UHV	Ultra High Vacuum

1 Introduction

The field of Scanning Probe Microscopy is growing so fast that this work can only cover a tiny portion of the whole field. Even though the first Scanning Tunneling Microscope was introduced only in 1981 and the Atomic Force Microscope in 1985, Scanning Probe Microscopy covers now already more than ten instruments. They are all closely related to each other and have at least some parts in common.

The original purpose of this work was to study the capabilities of a NanoScope StandAlone Atomic Force Microscope and to evaluate the possibility of using it "outdoors". Furthermore, a routine was developed and tested to investigate tensile or fatigued samples with the new instrument in combination with other available instruments (SEM, EBSD, Microhardness).

The second part of this work consists of the evaluation of the accuracy and the limitations of the Atomic Force Microscope using the two Atomic Force Microscopes available at CEM/TUT. Although Scanning Probe Microscopes are valued for their accuracy and high resolution, it is quite a problem when bigger scales are involved. Calibration is one of the most important aspects, and this seems to be known by the manufacturers who deliver the instruments usually in a precalibrated state (the mentioned StandAlone was not calibrated). However, it seems reasonable that every user should always calibrate the instrument first and not rely on the promised accuracy so blindly.

Another obvious limitation for the possible resolution of the Atomic Force Microscope is the tip. This problem is described and illustrated to assist novice users in their understanding of the instrument (due to the limited space, some images are only on the enclosed CD). In conjunction with this, a software was developed to estimate the tip shape when the image is captured. Also image simulation from a known surface can be done. Because it is possible to read directly the file format used by the manufacturer, this software can work as a framework when further calculations with the data are desired.

I Theoretical Part

2 Scanning Probe Microscopes (SPM)

2.1 General Operation

All the members of the family of Scanning Probe Microscopes have some basic parts in common, which are schematically illustrated in *Figure 2-1*.
The probe tip (type depends on the used technique) senses the desired interaction with the sample. This interaction will result in a detectable signal which is used in a feedback system. According to this signal the piezoelectric scanner is moved by applying voltage to it. Either the driving voltage on the piezoelectric scanner or the interaction signal is used to construct a 3D-image.

Means of sensing the vertical position of the tip

A feedback system to control the vertical position of the tip

A probe tip

A coarse positioning system to bring the tip into the general vicinity of the sample

sample

A piezoelectric scanner which moves the sample under the tip (or the tip over the sample) in a raster pattern

A computer system that drives the scanner, measures data and converts the data into an image.

Figure 2-1 Schematic of a generalized SPM [1].

As mentioned earlier, it is impossible to introduce the whole field of SPM. Therefore only the first two microscopes in the list below will be discussed in more details. A good introduction for further reading on currently available techniques is given in [1].

- Scanning Tunneling Microscope
- Atomic Force Microscope
- Magnetic Force Microscope
- Lateral Force Microscope
- Force Modulation Microscope
- Phase Detection Microscope

- Electrostatic Force Microscope
- Scanning Capacitance Microscope
- Thermal Scanning Microscope
- Near Field Optical Microscope
- Nanomanipulation
- Nanoindentation

When investigating the sensitivity and accuracy of SPM's, there are two major parts which contribute to the general performance, the scanner and the probe tip. To handle later some frequently used expressions more easily, both are first briefly reviewed. The scanner, which is similar in all SPM's is discussed first and the probe tips are explained when the relevant SPM is treated.

2.1.1 Scanner Basics

The scanner plays one of the major roles in the correct operation of a SPM. It is usually fabricated from PZT (lead zirconium titanate) with various dopants to get a proper piezoelectric behavior. Depending on the amount and type of dopants, piezoceramics can be divided into high sensitivity ceramics ("soft ceramics") and low sensitivity ceramics ("hard ceramics"). Sensitivity means the length change per volt [nm/V].

Piezoelectric ceramics are ferroelectric materials and therefore they show fundamental nonlinearities in their response to applied fields. The higher the electric field strength E [Vm^{-1}] and the higher the piezoelectric sensitivity [nm/V] of the material, the more nonlinear behavior will arise. This will be discussed in Chapter 4 in more details.

To align the electrical dipoles after sintering, a so-called poling process is used. The scanner is heated to about 200°C where the dipoles are aligned with a DC voltage. When cooled down with the DC voltage switched off, a remanent polarization is left. This manufacturing step is expressed in *Figure 2-2* with the path *0cd*. The point *d* is now the operating point, where only lower electric field strengths are applied during use. Exposing the piezoelectric material to relatively low positive and negative fields causes a reasonable hysteresis, expressed by the path *efgde*. The actual hysteresis is not easy to determine because it depends on the magnitude and rate of change of E. As soon as the used fields become higher, the additional influence from the S-shape of curve *aec* in *Figure 2-2* is of importance.
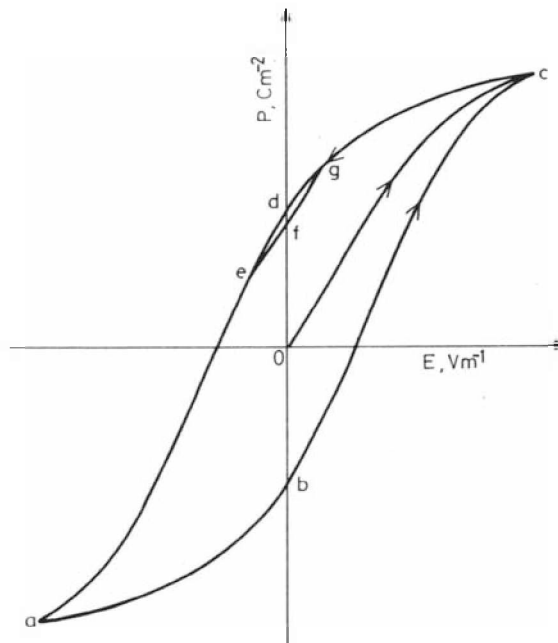
Figure 2-2 Schematic relation between polarization P and electric field strength E for a piezoelectric ceramic [2].

Although a charge driven piezoelectric actuator is suggested to exhibit less non-linear behavior [2], all currently available SPM devices work with voltage driven piezo-scanner. The main advantage is the higher sensitivity of the scanner when voltage is applied. The drawback of higher nonlinearity is compensated with calibration routines.

Lack of use as well as heating will cause the dipoles to randomize again and the built-in remanence will decrease. This will gradually decrease the actual scan size, leaving finally a useless peace of ceramic.

Although there are various designs, one of them is used very frequently in "middle class" SPM's - the tube scanner (*Figure 2-3*). The hollow tube has a metal coating on the outside, which is split into four segments (+x,+y,-x,-y). Applying a bias voltage to opposite electrodes will cause the scanner to bend. Z-motion is made by a separate element independent from the others or by applying voltage to the four other segments with the counter electrode on the inner side of the tube for both cases. Scan sizes offered range from $1\mu m$ up to $200\mu m$ in lateral direction and up to $10\mu m$ in vertical direction.
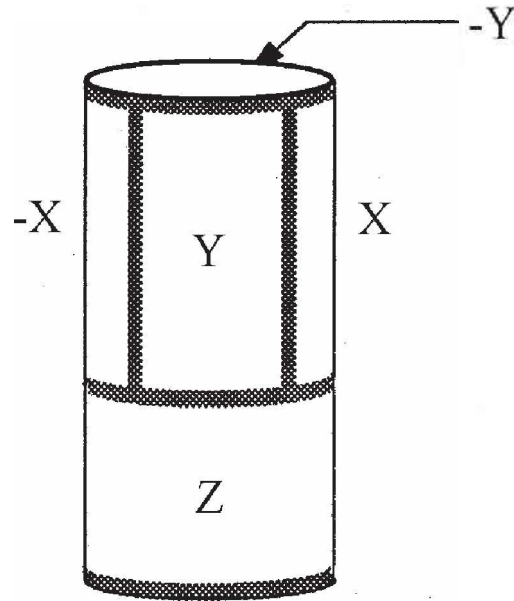
Figure 2-3 General electrode configuration for a piezoelectric scanner [3].

The sensitive and fast responding scanner is driven by the controller electronics (*Appendix A*). The image is stored digitally as interaction values collected during the X-Y scan. The number of points collected per line is (64), 128, 256, 512 or (1024). The number of lines is usually equal to the samples (points) per line.

Other designs like tripod scanner or the beetle concept are not discussed, because they were not relevant to the present work.

Atomic resolution, which highly depends on the scanner, seems to be fairly easy to get with the Scanning Tunneling Microscope (STM) as well as with the Atomic Force Microscope (AFM). Even under ambient conditions, when the surface is covered with a water layer, atoms can be observed with both techniques.

2.2 Scanning Tunneling Microscope (STM)

The STM was the first device which was able to give topographical information on the atomic scale [4]. G. Binnig and H. Rohrer got half of the Nobel price of physics for it 1986, just five years after presenting their invention. The second half was given to Ruska for his development of the electron microscope in 1931(!).
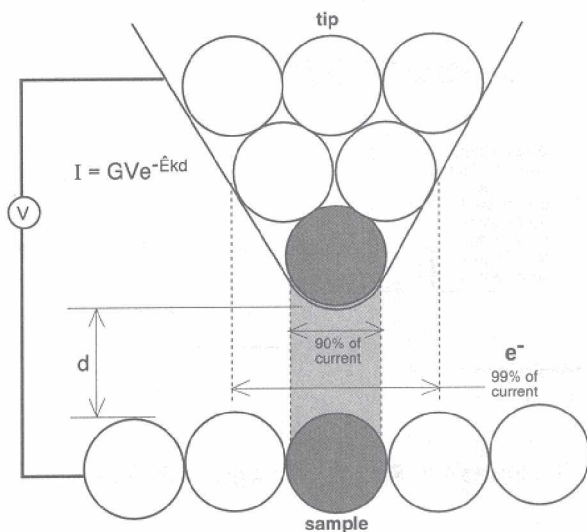
The STM senses the tunneling current between the sample and the tip when a bias voltage is applied. The sample and the tip are only several nanometers apart. There are two possible ways of feedback control - either **constant current** or **constant height mode**. When constant height mode is used, the height is kept constant and the altering current is monitored. In constant current mode, the feedback system controls the scanner motion according to the measured tunneling current so that the current is kept constant. Probes (tips) are mainly made either from tungsten or platinum-iridium by an etching process (more in 2.2.1 and [5]). The image reflects the local density of states which corresponds to the probability that an electron is present at the image points. *Figure 2-4* shows the high tunnel current dependency on the separation. When scanning in atomic scale, the current will be concentrated between the topmost atom of the tip and the sample. In larger scale the image will reflect the surface topography but in atomic scale the image does not necessarily correspond to the topography. For example, an adsorbed helium atom on a jellium surface will decrease the local density of states and therefore appear as a hole in the image [6].

Figure 2-4 Tunneling current dependency for the STM [1].

Samples need a reasonable conductivity for being observable and one should also be aware of the possible problems of surface oxidation. When constant current mode is used, a partial surface oxide scale will cause the tunneling current to drop and the feedback system will apply a counteracting voltage to the piezo so that the current reaches the previous level. This may crash the tip with the sample and according to [7], the tip can curl like a spiral even if it is made from a brittle material like tungsten.

In general, STM is a very good tool for investigating electrical properties for example in semiconductive materials. For topographical imaging it has the limitation that it is not possible to use non-conductive samples or samples with inhomogeneous oxide scales. Therefore the STM is easier to use in UHV conditions, where the nonuniform oxidation can be slowed down.

2.2.1 STM Probe Tips

The above mentioned etching process needs considerable knowledge and expertise. Although tungsten and platinum/iridium are the usual tip materials, it may sometimes be useful to produce tips from gold, molybdenum or other materials. In *Figure 2-5 (a)* some electrochemical etching setups are shown that are used to produce a (hopefully) atomically sharp tip *(b)*.
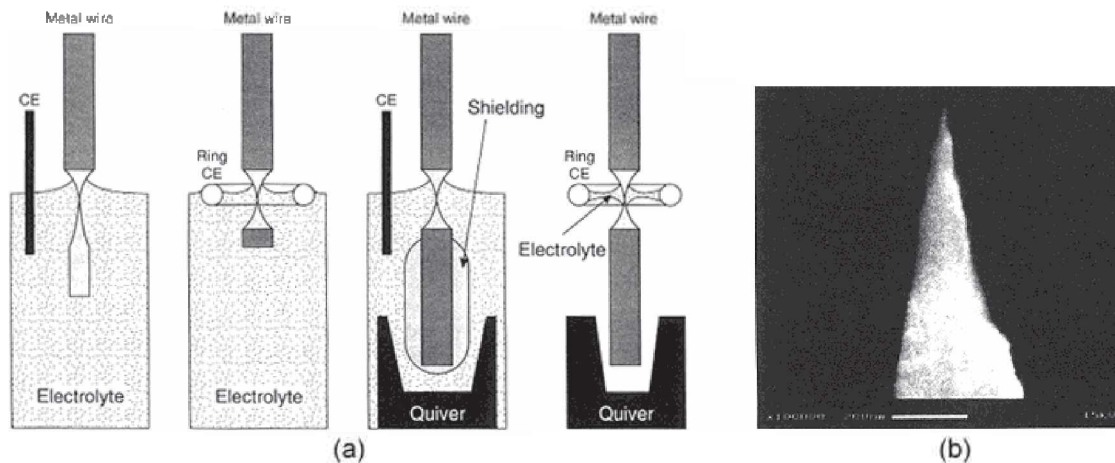

Figure 2-5 (a) Various setups for electro-chemical etching (CE = counter electrode),(b) SEM micrograph of a platinum-iridium tip, the white bar indicates 200nm [5].

Even though it does not look dramatically difficult to produce own tips, one should consider that there are lots of influencing parameters, which can affect the tip shape and the apex radius. That is, whether to use AC or DC voltage or a combination of both, the frequency of AC voltage (few Hz to 1000 Hz), the applied voltage amplitude (1 to 30 V), the used material, the motion of the wire material (rotating or not), the concentration of the electrolyte, *etc..* Some basic routines are given in Table 2-1.

Table 2 - 1 Suitable materials for STM tips and the recommended electrolyte and voltage type. The star denotes frequently used materials[5].

wire material	electrolyte	concentration of electrolyte	voltage type
Au*	HCl	30 vol%	DC or AC
Ir	$CaCl_2/HCl/H_2O$	60%/4%/36%	AC
Mo	KOH	5 vol%	AC
Ni*	H2SO4	30 vol %	DC or AC
Pt (Pt/Ir)*	KCN	20 vol%	AC
	$CaCl_2/HCl/H_2O$	60%/4%/36%	AC
	NaCN/NaOH	3 M/1M	AC
Re	H3PO4/H2O2		DC
W*	NaOH	1 M to 4 M	DC or AC
	KOH	1M	DC or AC

2.3 Atomic Force Microscope (AFM)

The AFM was invented in 1985 by G. Binnig, Ch. Gerber and C.F. Quate [8]. Samples for the STM had to be reasonably conductive but with the AFM this limit vanished. The probe consists of tips which are on the downfacing side of flexible cantilevers. Commercially available tips are mainly made from silicon or silicon-nitride. A brief introduction to AFM probe microfabrication is given in 2.3.1 and a comprehensive treatment can be found in [9].

When the probe is approaching the sample, the cantilever bends according to the increasing influence of the acting forces. In *Figure 2-6* the different forces are illustrated. The closer the tip is to the specimen surface, the higher is the bending force.
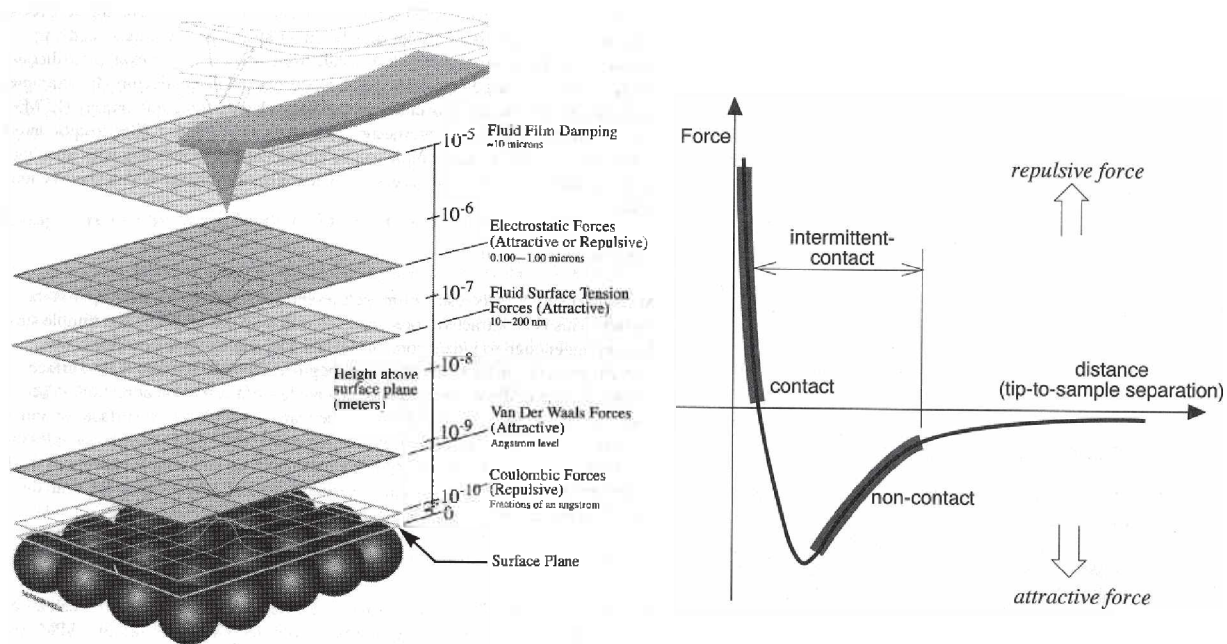
Figure 2-6 (left) Sample-Probe boundary layers [10]. (right) Force-Distance dependency [1].

One can choose between different operation modes:

1. **Contact Mode:** The tip "touches" the sample all the time while scanning. The force acting onto the surface is so high (10^{-7} to 10^{-6} N) that it is possible to move atoms. Also image artifacts can be caused by tip-sample interactions, which can not be explained easily or predicted in advance.

2. **Tapping Mode:** The tip "taps" the surface by having the cantilever oscillating when scanning. This is equal to the intermittent contact expression in the figure above. The changes in the oscillation amplitude are monitored and used for the feedback control as well as for the image construction.

3. **Lift Mode:** When an oscillating tip is lifted, only the van der Waals forces are acting, which are comparably low (10^{-12} N). However, due to the long distance between the tip and the sample, features are less distinguishable (somehow blurred).

Nowadays tapping mode seems to be the most widely used imaging method.
The bending of the cantilever (contact mode) or the change in the amplitude (tapping and lift mode) is usually measured with a laserdiode/photodiode system. It measures the signal with a split photodiode via a mirror (*Figure 2-7*), or the reflected beam interferes with the emitted beam and this is used to define height variations (*Figure 2-8*).
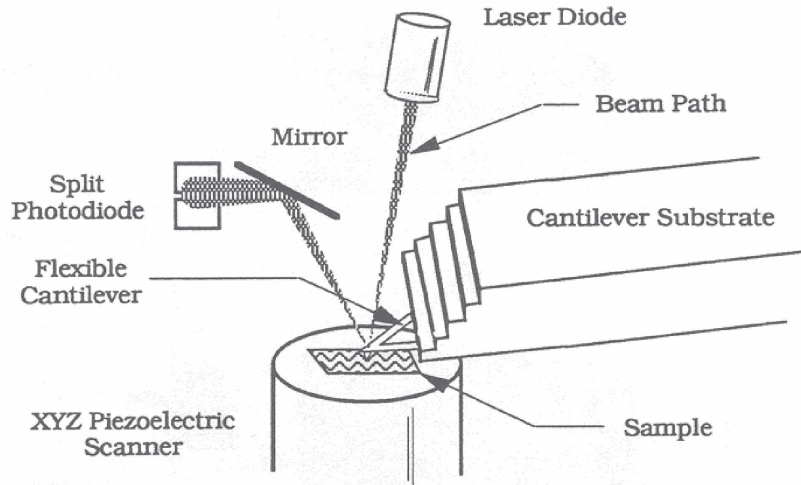
Figure 2-7 Optical sensing system of the AFM [11].

Like constant current
and constant height mode for STM, the AFM offers **constant force** and **constant height mode**. In constant force mode the feedback system tries to keep the deflection (or amplitude) of the cantilever constant. In constant height mode the scanner is at constant height and only the deflection (or amplitude) of the flexible cantilever is monitored.

Figure 2-8 Cantilever deflection detection system [12].

2.3.1 AFM Probes

The term probe comes from the fact that the various SPM techniques not necessarily need a sharp tip to probe the specimen. To cover all of them by one word, probe is generally used.

The AFM probe consists of two parts - the tip and a flexible cantilever. Quality is highly dependent on the tip but also on the cantilever.

For example, when using tapping mode a rather stiff cantilever with high resonance frequency is used, which is important if good resolution is desired.

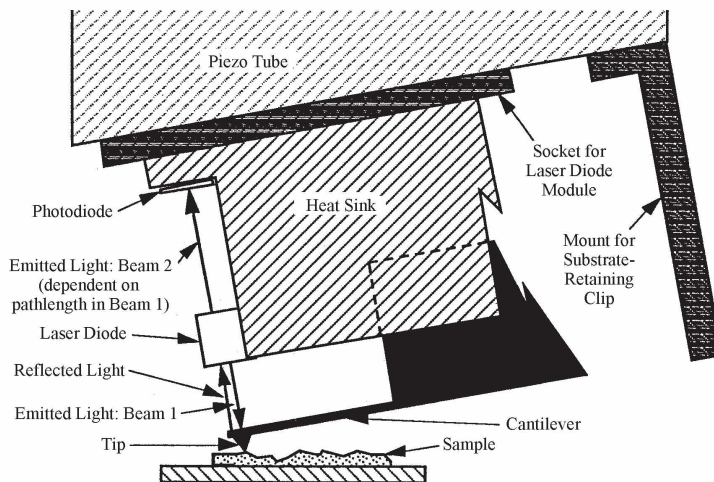For tips the important factors are the apex radius and the aspect ratio (both dependent on the type of the tip material). Mostly used materials are silicon and silicon-nitride, but also diamond, Nb, W, Ir, Ta and SiO_2 have been produced in reasonable quantities. They have all in common that the production procedure is based on a (100) Si wafer and relies on conventional batch fabrication techniques.

✷ *Cantilever Fabrication*

Cantilevers are made either from SiO_2 or Si_3N_4 by a thermally grown film or a LPCVD deposited film, respectively. The cantilever shape is defined by a masking and selective etching process with KOH, which self terminates at the (111) planes (*Figure 2-9 a*).

✷ *Miniature Cones*

When employing the described cantilever fabrication process, cones can be deposited onto them by evaporating material through an orifice from a point source located above the cantilever (*Figure 2-9 b*).

Figure 2-9
(a) Cantilever fabrication using masking techniques and selective etching [9].
(b) Miniature cones deposited through an orifice (typ. 3µm diameter) [9].
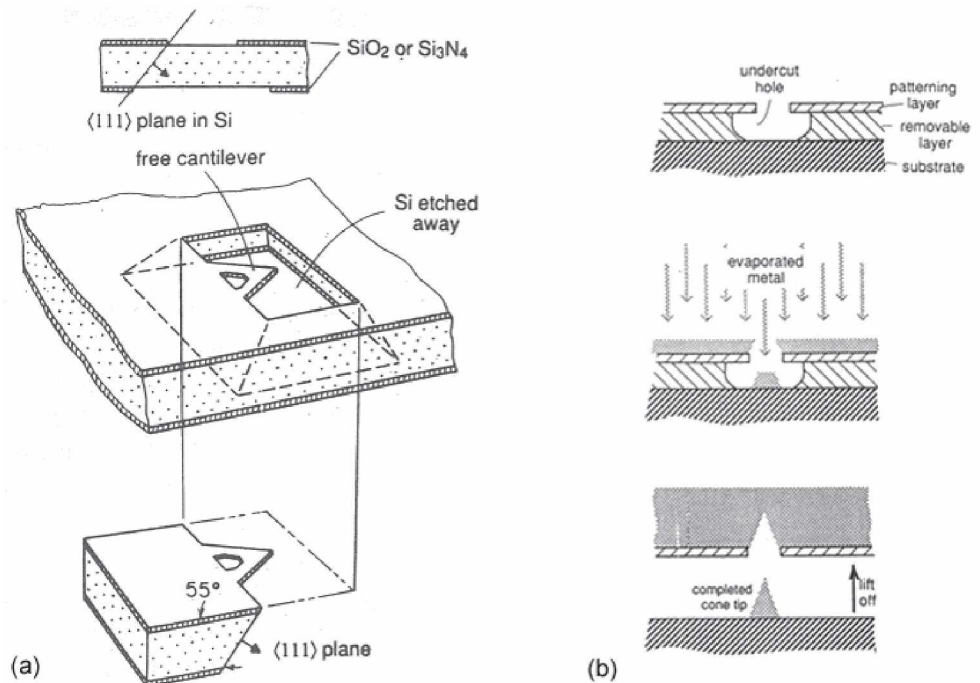
✳ Si_3N_4 Cantilevers with Integrated Si_3N_4 Tips

The anisotropic nature of KOH when etching Si can be used to make a small pyramidal etch pit, which can be used as a mold for other materials. Si_3N_4 tips and cantilevers can so be fabricated in one deposition step. Then the cantilever is patterned and the film is anodically bonded to a glass plate. All the Si is etched away and the back side is coated with metal (Au) for good reflection of the laser beam towards the detector (*Figure 2-10 a*).

✳ SiO_2 Cantilevers with Integrated Si Tips

First a square Si_3N_4 mask is patterned on the surface of a (100) Si wafer where the corners point in ⟨110⟩ directions. Anisotropic dry etching is used to form vertical sidewalls under the square. Then thermal oxidizing forms a SiO_2 layer and after selectively removing the Si_3N_4 (dry etch), a wet etching step with KOH is applied. After removing all the SiO_2, four protruding Si tips are left. The cantilever is produced by oxidizing everything again and patterning according to the demanded shape (three tips are sacrificed) (*Figure 2-10 b*).
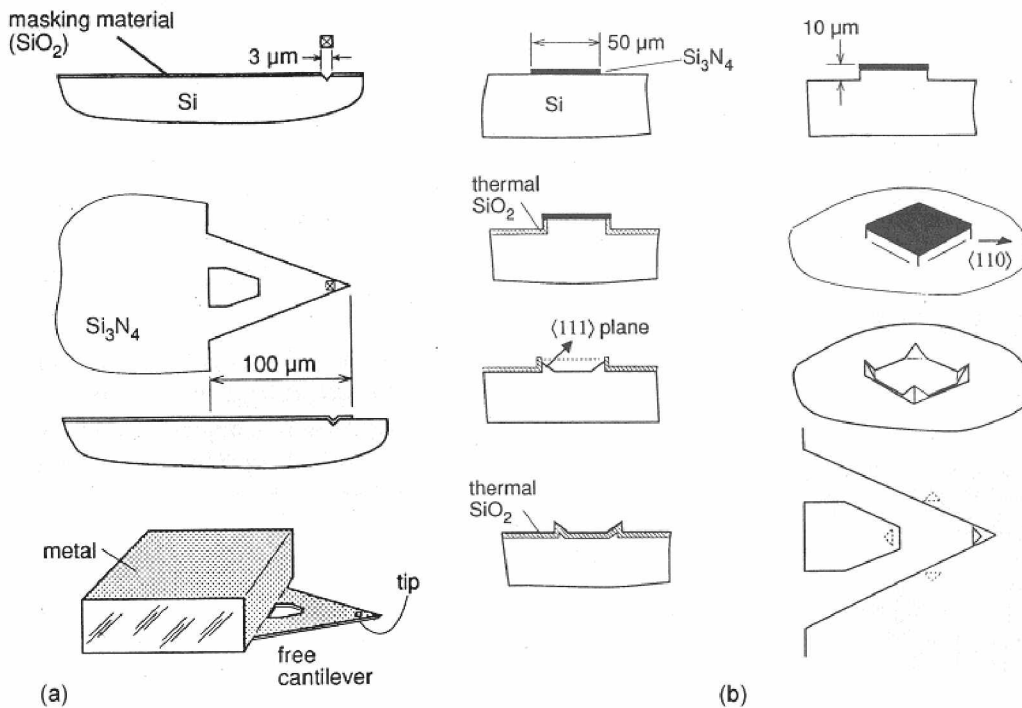
Figure 2-10 (a) Fabrication of Si_3N_4 cantilevers with integrated Si_3N_4 tips [9]. (b) Fabrication of SiO_2 cantilevers with integrated Si tips [9].

When evaluating the quality of the tips produced by different processes, it is apparent that the Si tips are best in the respect of the apex radius (< 300Å) and aspect ratio (note - tip is asymmetric). Si_3N_4 tips are least expensive, have a larger apex radius (≥ 300Å) and a poor aspect ratio (sidewall slope 55°). Cones produced by evaporation through an orifice are less sharp (≥ 500Å) and the aspect ratio depends on the material (*e.g.*, Nb cones - high aspect ratio, Si cones - low aspect ratio).

3 Probe Induced Limitations

3.1 Introduction to Probe Artifacts

Soon after the invention of STM it became apparent that the tip-shape is contributing intensively to the image [13]. When simultaneous tunneling happened from more than one atom on the tip, the Si(111)7×7 reconstruction was given as a double image. An interesting way of improving the resolution is presented in [14]. It was shown that applying a bias voltage of -10V caused a needle formation on the tip which in turn improved the resolution of the image. How important the actual tip shape is for STM is comprehensively treated in [15].

Atomic resolution with AFM was first reported around 1987 and it was rather something unexpected. The less defined interaction of an AFM (especially on air) did not give raise to believe in this future prospect. The observed atomic flat surfaces were later explained by the fact that the force acting between the probe and the sample was high enough to move atoms. Nowadays force can be reduced so that atoms stay at their original place when tapping mode is used.

Tip convolution or tip imaging is another important part in AFM images. As soon as sample features become sharper than the tip, the image will represent both the tip and the sample. In *Figure 3-1* this is illustrated by means of a conical and a pyramidal tip. Whenever steep slopes are involved, like in the semiconductor industry, the AFM image quality is defined by the aspect ratio of the tip [16].
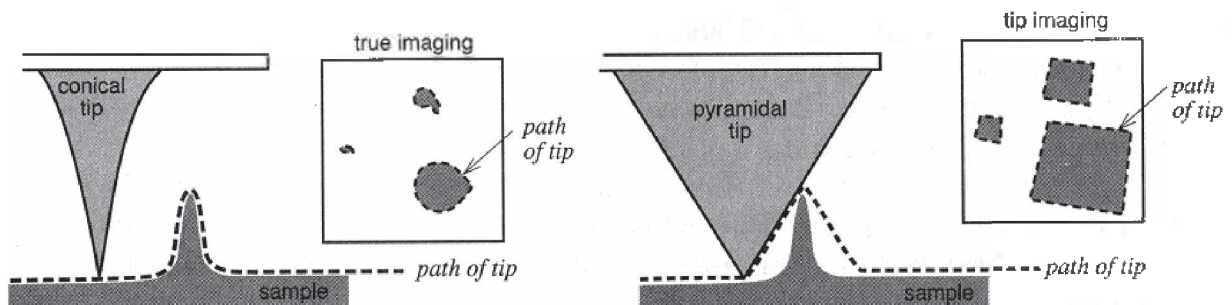

Figure 3-1 Tip convolution dependence on the tip shape [1].

Further problems are caused by blunt tips and the lateral shift when the true contact point is not the same as the apparent one (which is in principle the same as the slope problem). This is illustrated in *Figure 3-2*.
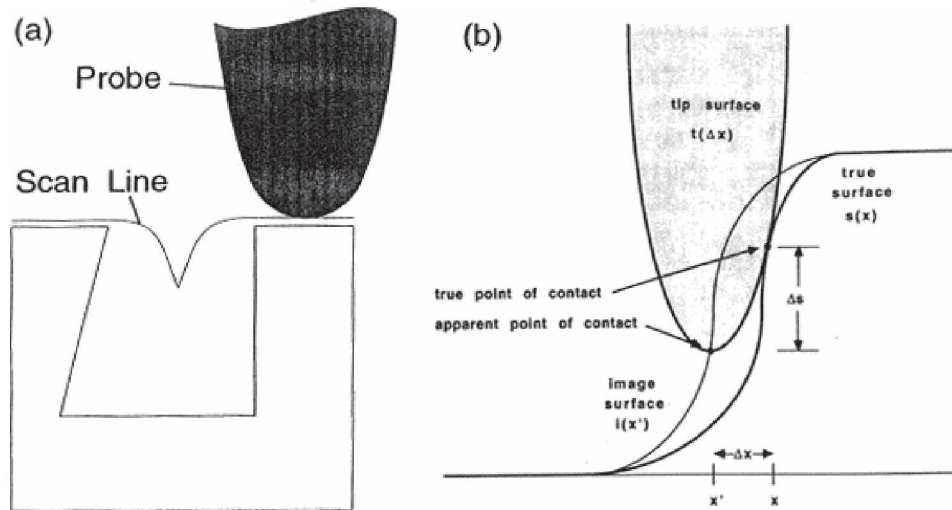
Figure 3-2 (a) Limitation due to blunt tip [16] and (b) lateral shift [17].

3.2 Direct Tip Imaging

When sharp features are imaged the shape of the tip is reflected and the actual tip radius can be estimated. This is quite useful, for example, when assessing whether the tip has to be changed or not. Pin-like features can, for example, be prepared by depositing copper-films on quarz with MOCVD [18]. The tip image is shown in *Figure 3-3,* which also illustrates immediately the drawback of this method. The most interesting part of the tip - the apex - is not resolved as well as needed due to the finite size of the copper spikes on the surface. The second drawback is that the applied horizontal force is high enough to bend the spike-like features and this will lead to distortion of the given tip shape.

It is obvious that if a special characterizer is used, the actual tip shape can change. The same applies to investigatons of the tip in a SEM. The best results of the tip shape estimation are therefore obtained when the characterization of it is based on the image itself.

Figure 3-3 Tip image obtained when a pyramidal tip is scanned across a sharp pin [18].

3.3 Computational Estimation Routines

SPM's are all computer-controlled and therefore it is natural that tip estimation procedures can easily be included in the software. Since 1991, several articles have been published and commercial software for tip estimation is also available[1].
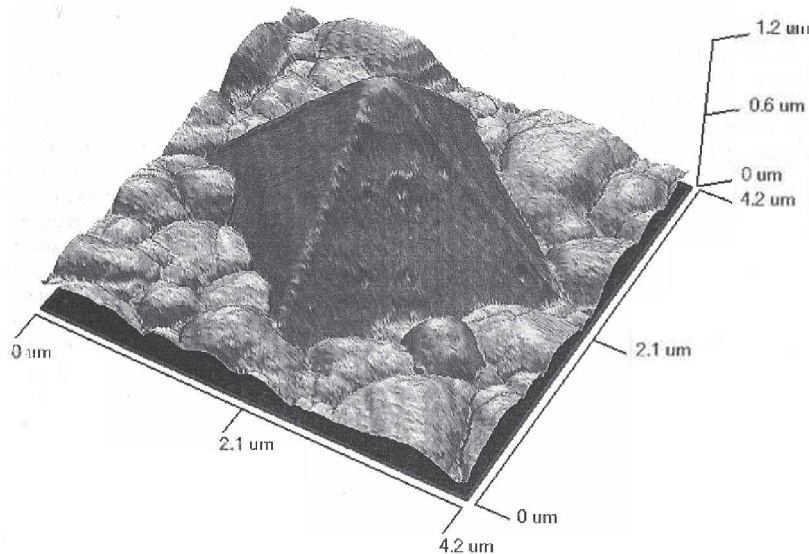
3.3.1 Mathematical Morphology

Mathematical morphology is the branch which offers the tools for the computational treatment of SPM tips. The theory presented in the following is based on the work of J.S. Villarrubia and can be found in his publications [19]-[22]. The so-called envelope technique, developed by D.J. Keller [17], [23] is practically equal to Villarrubia's technique.

First, two important expressions have to be defined:

Surface is the true surface of the sample, meaning the shape of it as if one would look with the eye through an ideal magnifier.

Image is the 3-dimensional result when an AFM is used to investigate a surface.

Morphology is a branch of set theory and therefore objects are denoted as sets. The function $s(x,y)$ describing the sample surface has its counterpart in a set S, which contains all the sample points. When describing a SPM surface of a sample, it is of no concern how the bulk material beneath the surface looks like and S can be defined by $S = \{(x, y, z) \mid z \leq s(x, y)\}$. Correspondingly, $s(x,y)$ is called the *top* of S, also denoted T(S). An object which consists of a single-valued *top* and all the points beneath it, is called an *umbra*. Upper case letters are from now on reserved for the sets and the surface of the set is denoted by the corresponding lower case letter or the T() notation.

To make the notation a little bit more familiar to the reader, some important operations are presented in *Figure 3-4*. In *(a)* the translation of a set A, by a vector **d**, is determined by adding **d** to every element of A:

$$A + \mathbf{d} = \{\, \mathbf{a} + \mathbf{d} \mid \mathbf{a} \in A \,\} \tag{1}$$

If A is an umbra, the notation in terms of its top is:

$$T[\, A + \mathbf{d}\,](x, y) = a(x - d_x, y - d_y) + d_z \tag{2}$$

where $\mathbf{d} = (d_x, d_y, d_z)$.

---

[1] Digital Instruments [32]

In *(b)* the union of two umbras is covered by the whole shaded area, independently of the orientation. The top of the union is therefore the maximum of the two tops.

$$T[A \cup B](x, y) = \max [a(x, y), b(x, y)] \qquad (3)$$

The crosshatched area is in turn obviously the minimum of the two tops.

$$T[A \cap B](x, y) = \min [a(x, y), b(x, y)] \qquad (4)$$
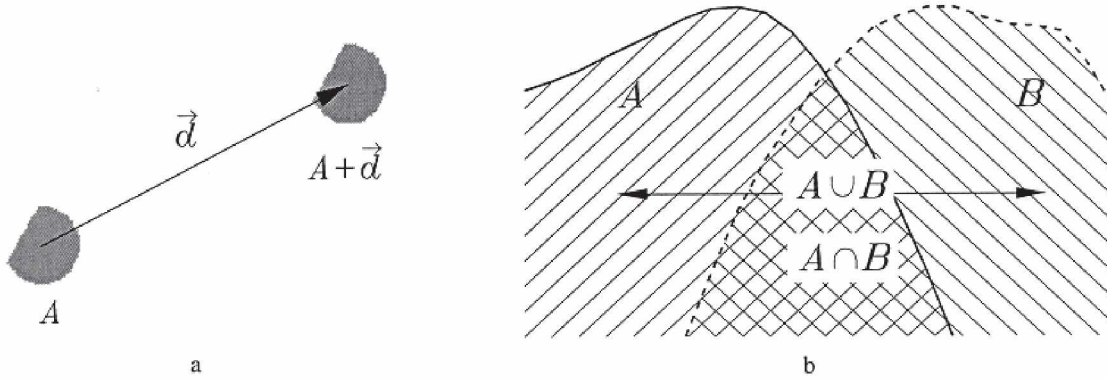
Figure 3-4 Basic set operations. (a) Translation of a set by a vector. (b) Union and intersection of sets, and their relationship to the maximum and minimum of the tops of the set [22].

3.3.1.1 Image Simulation by Dilation

The afore mentioned tip convolution can also be expressed as a dilation of the surface, where protrusions become broader and blunter, and pits or crevices become smaller and tend to disappear. Treating dilation in terms of sets is defined as:

$$A \oplus B = \bigcup_{\mathbf{b} \in B} (A + \mathbf{b}) \qquad (5)$$

which resembles the maximum of the two tops from the corresponding umbras A and B. When only the function of the top is needed, it can be also written as:

$$T[A \oplus B](x,y) = \max_{(u,v)}[a(x-u,y-v)+b(u,v)] \qquad (6)$$

Considering now a topographical image - we have two sets that contribute to the final image (I), the surface (S) and the tip (T). In *Figure 3-5* the image top i(x, y) is constructed from the surface top s(x, y) and the tip top t(x, y). Keeping the focus on the tip apex (x', y'), the dilation A⊕B is the space which is described by the translation of the tip apex, when the two sets do not overlap.

Figure 3-5 Image formation in topographical images [22].
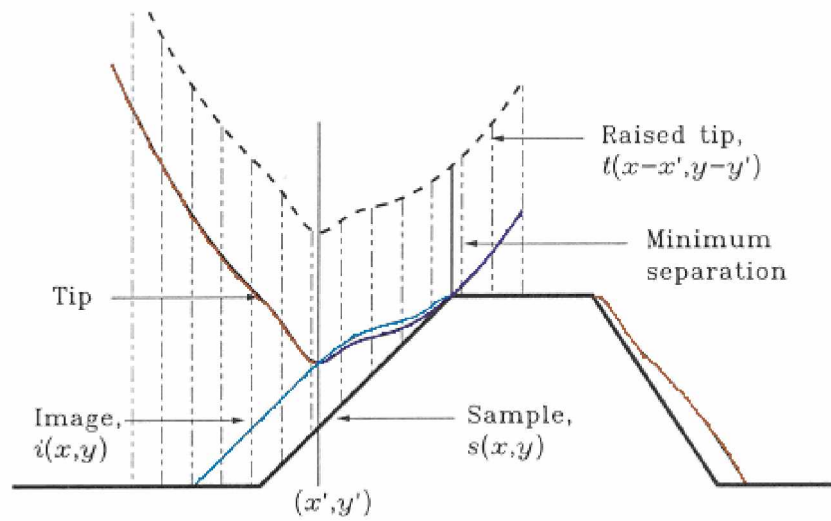
From *Figure 3-5,* it seems obvious that the image top can also be constructed by reflection of the tip through the origin - note the corresponding colors. *Figure 3-6* illustrates this in an even more obvious way.
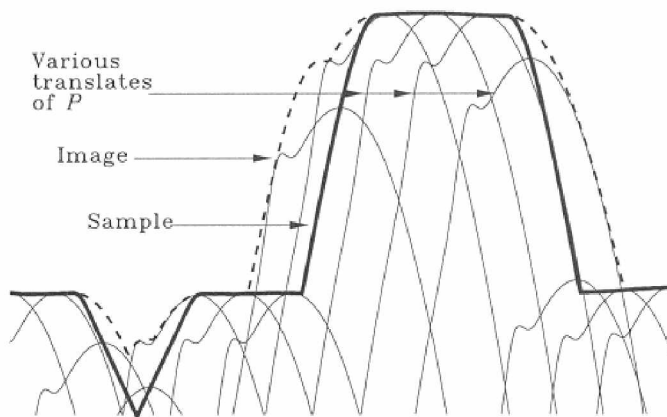
Figure 3-6 Formation of an image by dilation [22].

It should be noted that the reflection through the origin means that x, y and z have a change in sign. If the dilation given in *Figure 3-6* is expressed as a formula, we have

$$I = S \oplus P \quad (\text{where } P = -T) \tag{7}$$

with the respective tops i, s and p. In other words, the sample (S) dilated by the reflected tip (P) equals the image (I).

The consequence of eqn. (7) is that there is a tool to simulate the image one will obtain from a known surface with known tip characteristics. Especially, when molecules are imaged (like DNA-chains), it is of great interest to know how the image should look like. Generally, the steeper and smaller the surface features are, the more affected is the final image by tip convolution.

3.3.1.2 Surface Reconstruction by Erosion

As we have seen, the image produced by AFM is not the true topographical surface and therefore it seems reasonable that one would like to know at least an estimate of it. The mathematical operation for this is

$$S_r = I \otimes P \qquad (8)$$

where $\otimes$ denotes erosion and is defined by

$$A \otimes B = \bigcap_{\mathbf{b} \in B} (A - \mathbf{b}) \qquad (9)$$

It first looks that erosion is the inverse operation of dilation, when the two definitions are compared. Unfortunately, erosion equals "only" the best possible surface estimate S_r, which is equal to the surface only if the apex was always in contact with the surface. As soon as features become steeper than the tip, the true contact point is not any more the apex and after eroding the image, parts of the surface can not be resolved (*Figure 3-7*). As a result, S_r is the *least upper bound* of the actual surface, *i.e.*, no upper bound smaller than S_r is acceptable.
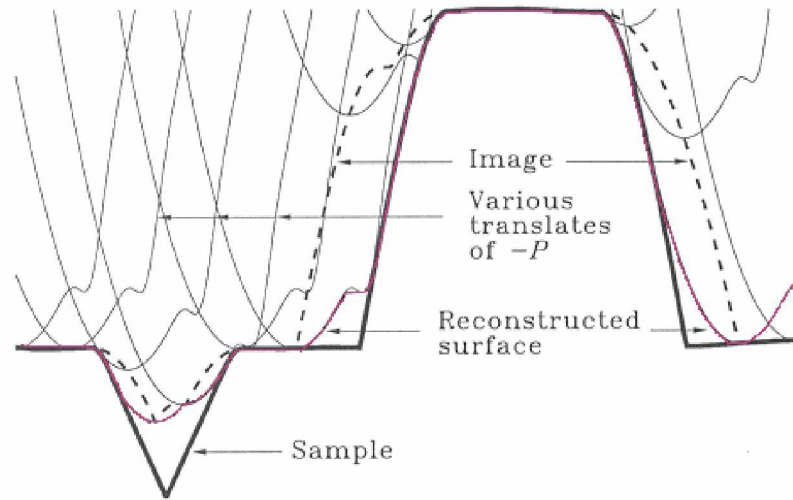
Figure 3-7 Geometrical interpretation of erosion. The specimen surface is the thick continuous line and the dashed line describes the image. The best estimate of the surface is given in color, which is the deepest penetration of the tip [22].

When the image and the surface are given (*e.g.*, when a known characterizer is used), the erosion routine can be used to construct the tip.

$$P_r = I \otimes S \qquad (10)$$

The image (I) eroded by the surface (S) results the outer bound of the tip (P_r). As already explained for S_r, the outer bound of the tip, P_r, is only equal to the real tip at points where P touched the surface (S). This is illustrated in *Figure 3-8*, with three structuring elements on the surface.

On the left, there is a sharp (r_1) feature with height h_1, in the middle a blunt (r_2) feature with the same height and on the right side a tall (h_2), blunt feature. The tip reconstruction (thin line) is basically shaped after the left and right feature, but due to the smaller sidewall of the right, the reconstruction shows a step. When this tip is used to reconstruct the surface (by dilation), it causes a poor approximation for the outer features, while the middle one is nearly equal to the real surface.

Figure 3-8 Tip and surface reconstruction depends on the structuring elements on the surface [22].

In *Figure 3-9* the two discussed operations are once again show graphically.

Figure 3-9 Simulation of dilated and eroded images of a cylindrical rod stud: (b) dilated and (c) eroded by a (a) parabolic tip and in (d) are the corresponding cross-section profiles - (1) cylinder, (2) dilated and (3) eroded [24].
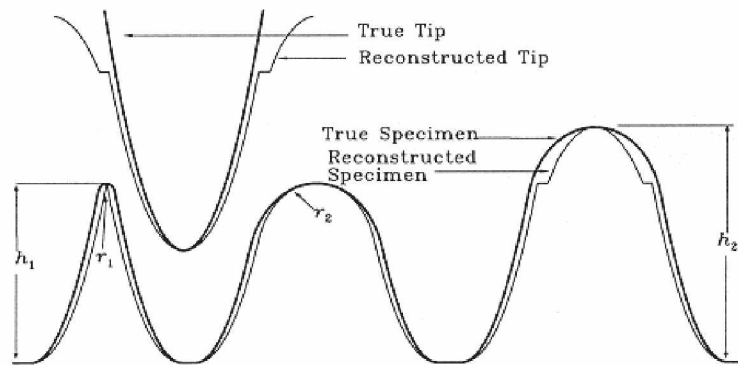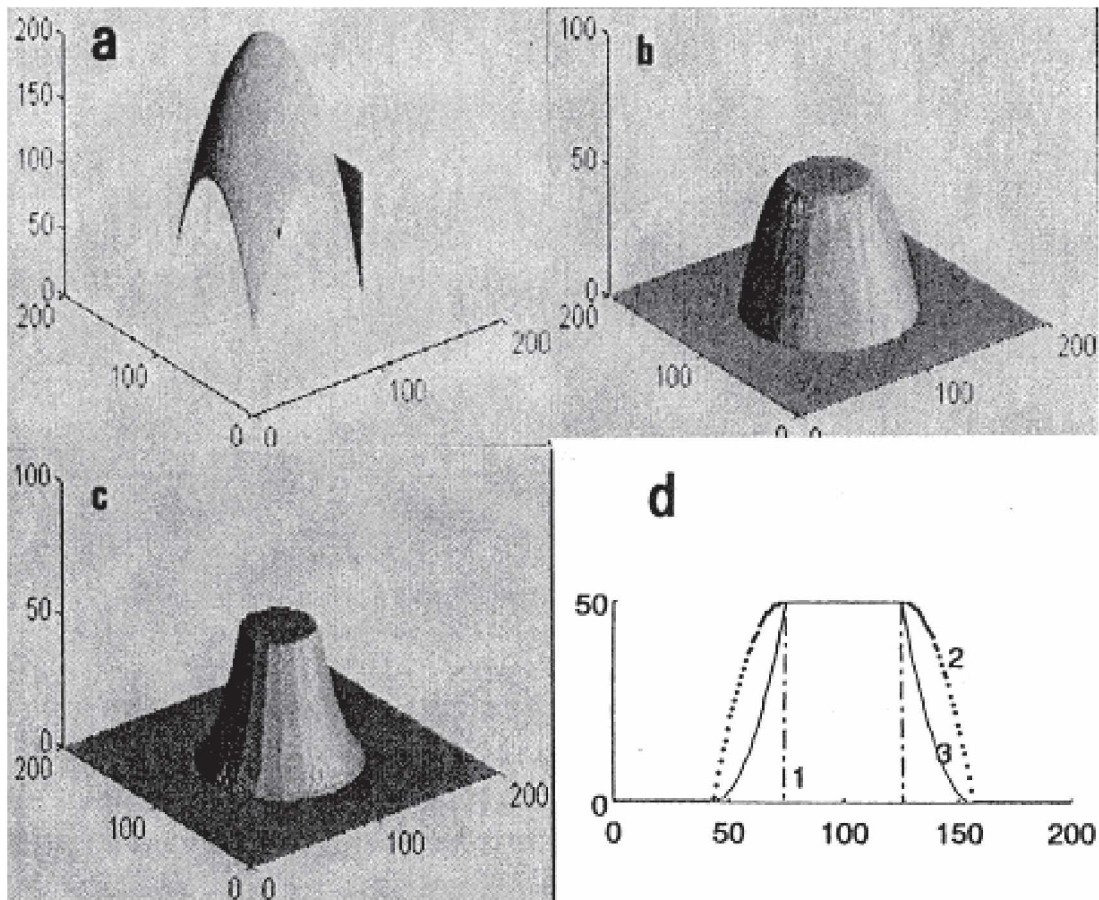
3.3.1.3 Certainty Maps

It was shown that the erosion only gives a best estimate of the true surface. Consequently, the knowledge of locations where $s_r = s$, is of great interest. This can be expressed as so called certainty maps.

The inputs for the certainty map calculation are the tip and the eroded surface reconstruction. When the tip is brought into contact with the reconstructed surface, there are two possible scenarios presented in *Figure 3-10*. In tip position 1 there is only one contact point and $s_r = s$ is valid for this point, which will be expressed by white color. In tip position 2, multiple points are touched and that means that it is not known which of them is from the real surface - therefore all of them are handled as unresolvable, denoted as black.

Figure 3-10 Two scenarios - in position 1 there is one contact point and in position 2 two points are touched at the same time [22].
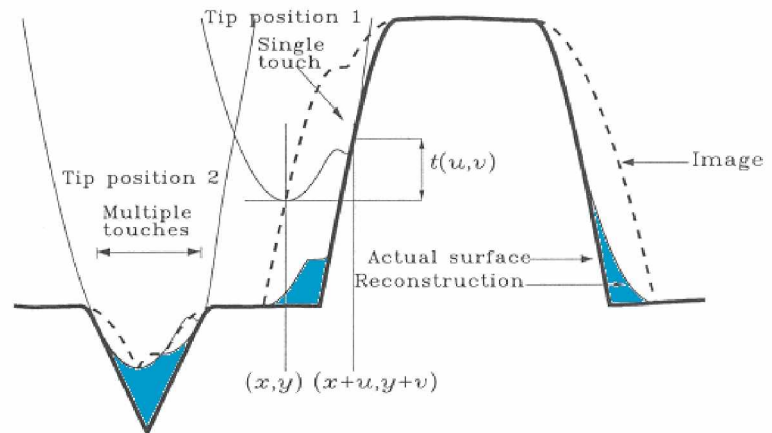
3.3.2 Use of Characterizer

The crucial thing when using a characterizer is the accuracy of the characterizer itself. When the tip shape should be revealed, the feature of the characterizer has to be always smaller than the tip-feature.

One way is to use polystyrene latex balls [25], which are very uniform in size and can be purchased with a diameter ranging from 10 to 100 nm. When all the geometrical information is considered, the apex radius of the probe can be estimated. This is unfortunately not the result one may want to get but more information is needed. Latex balls have another limitation, since sometimes they appear as cones in the image, although being round, when contact imaging is used [26].

Another, more reliable, way is to use a characterizer with a known shape silicon surface. Even the scanner calibration grating can be used for this purpose. In *Figure 3-11 (a)* a favorable shape is presented, because when imaging under ambient conditions the probe can be flexed due to the adsorbed waterlayer (*Figure 3-11 (b)*).
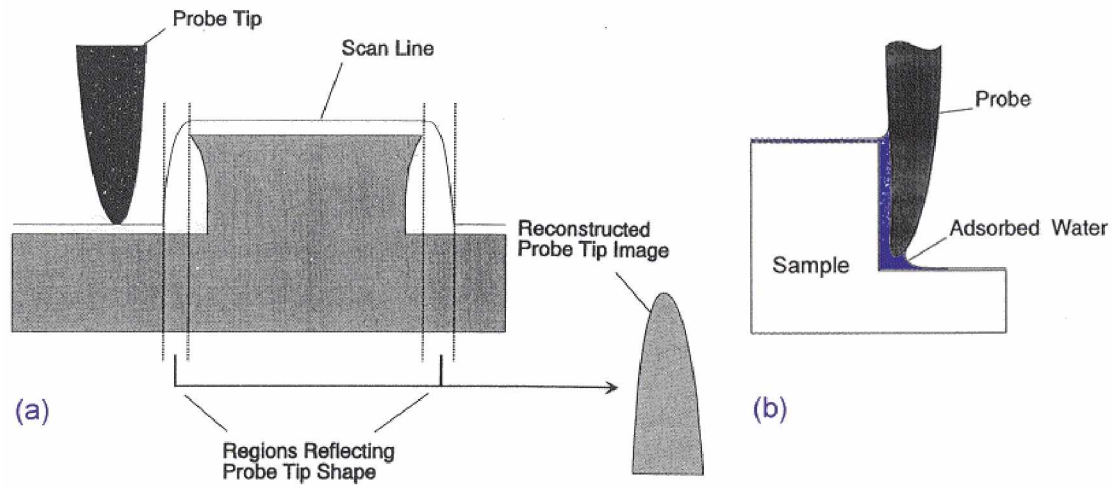
Figure 3-11 (a) Preferable shape of defined surface and the (b) water layer problem [16].

The known surface and the resulting image contain the information needed to calculate the tip shape with the erosion routine. One has to remember that the given tip-shape is equal to the real tip-shape only if all the tip surface points have touched the characterizer.
However, it is still uncertain weather the tip-shape is still the same or not when later sampling a surface, due to possible intermediate degradation of the tip.

3.3.3   Blind Tip Reconstruction

This is the field where the tip shape is revealed without using any special characterizer. The expression "blind reconstruction" means that the tip-shape is calculated from an *unknown* characterizer.
The image can be seen as the dilation of the tip and this can be used as its own characterization tool. Every image can be imagined as a set of columns. As soon as the slope between two columns is higher than the slope of the corresponding tip, there will be some tip-imaging. This information can be collected together and the resulting tip can be reconstructed. Some considerations have to be taken into account to apply an iteration routine. The discussed tools of mathematical morphology (dilation and erosion) were used to develope the iteration algorithm, which is used to estimate the tip [19].

$$P_{i+1} = \bigcap_{x \in I} [(I - x) \oplus P_i^*(x)] \cap P_i \qquad (11)$$

where $P_i^*(\mathbf{x})$ is defined as,

$$P_i^*(\mathbf{x}) = P_i \cap (\mathbf{x} - I) \qquad (12)$$

Iteration will converge to the result P_r.

$$P_r = \lim_{(i \to \infty)} P_i \qquad (13)$$

It can be shown that each iteration of eqn. (11) produces a result smaller than or equal to the preceding one, but that each P_i remains larger than the actual tip. This resembles the best possible tip estimation.

3.3.3.1 Choosing the Initial Tip

When starting the iteration, one needs to define the initial tip and to give the measured image. The most primitive and bluntest tip consists of zeros. Consequently, when running the calculation more refinement has to be done in comparison to a nearly correct starting estimate. However, as soon as the initial tip is too sharp, the accuracy will decrease, because the algorithm does not work in both ways, *i.e.*, only reducing of the outer bound takes place.

It is also not trivial how big the initial tip should be chosen. When the tip is as big as the image, only one contact point has to be calculated. Such a tip size is a far too conservative point, because actual tips are sharp enough not to have a large area that contributes to the image formation.
To get the structuring information from different image locations, the optimal size of the tip should be estimated according to the topography of the specimen. Suppose the height variation is about 100nm, then the tip should be higher than this. Taking now a parabolic $(z=x^2/(2r))$ or pyramidal (35° sidewall angle) tip with an apex radius of 40nm, the height will exceed 100nm when x>90nm. Then even a vertical step of 100nm would not be invisible as structuring element for the tip.

The effect of the initial size of the tip and the resulting width of the blind tip estimate is illustrated in *Figure 3-12 (a)*. The image (thick line) was simulated by dilation of the surface (thin line) with a parabolic tip (given in the width representations). In *(b)* the width (arbitrary units) is given as a function of the tip to image dimension ratio.
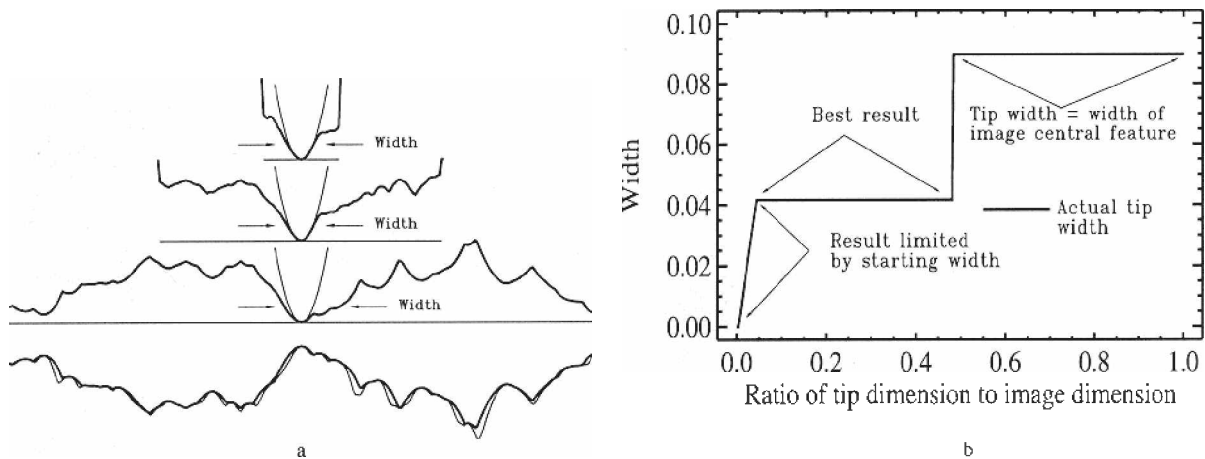

Figure 3-12 Effect of the initial tip size on the resulting blind tip reconstruction [22].

3.3.3.2 Noise Effects

It is evident that every electronic instrument shows some sort of noise in its output, and so does the SPM, too. In real pictures, this leads to problems for the algorithm. One might expect that the random noise cancels itself out, but unfortunately the presented algorithm rather exaggerates noise. With the aid of *Figure 3-13* the problem of reconstructing a noisy image is shown. The upgoing spike on the left side of the figure has nearly no effect on the reconstruction, because the erosion will take the minimum of the two sets. The reflected tip will be used and therefore the spike looks like a narrow cleft that the relatively blunt tip can not fully enter. In contrast to this, a downgoing spike will cause a lower estimate compared to the real surface. When noise is present like that on the right hand side, the estimation of the surface is generally lower than the real one (dashed line). The actual lower estimation is dependent on the frequency characteristics of the noise and the tip shape.

Figure 3-13 Effect on the surface reconstruction (red colored), when an overlay of noise (standard deviation, σ, indicated) is present in the image [22].
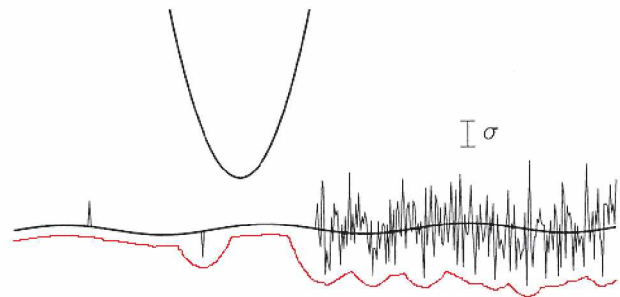
To reduce the influence of noise, several procedures are suggested and the outcome may vary a lot.

☆ **Image Processing**
This term stands for a vast number of different operations, like flattening, neighborhood averaging (with or without weighting), medial filtering, *etc.*. Care has to be taken not to alter real surface features when performing one of these operations, because the reconstruction depends on them. For example, neighborhood averaging smooths steps and therefore known real features have to be excluded. Median filtering is better in this respect and should be preferred.
Filtering before reconstructions is not equal to filtering after it. The filtering before the reconstruction will obviously yield a better result, because of the described problems of the algorithm.

☆ **Less Precise Algorithm - Threshold Value**
To make the algorithm less susceptible to noise a so called threshold parameter can be added. When comparing the current tip estimate with the previous estimate from another image location, all the different tip points would be changed when thresh=0. When the threshold parameter is different from zero, only tip points are changed where the difference between the two estimations exceeds the threshold value. Further they are changed to a value which is the new one plus the threshold value.

The convenient way to define the real "touching" points with certainty maps is also heavily limited by the effect of noise (*Figure 3-14*).
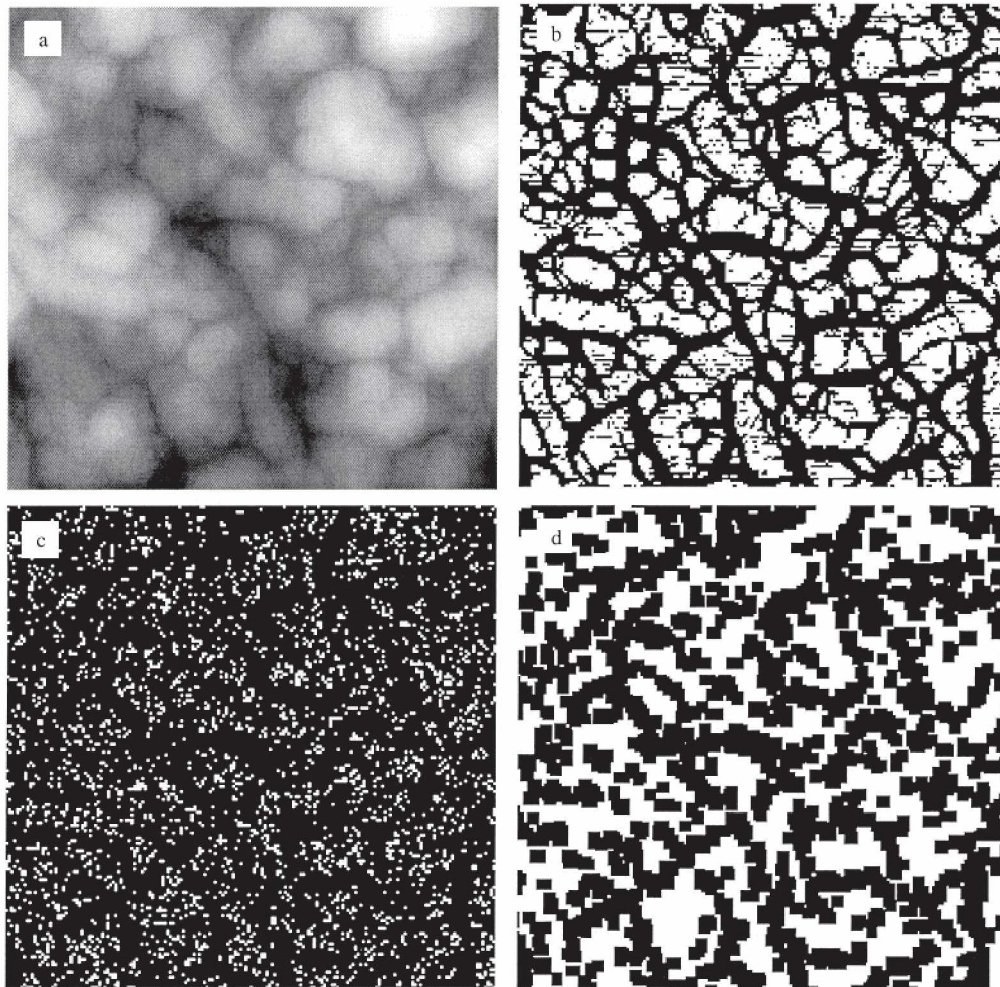
Figure 3-14 (a) An image simulated with a parabolic tip. (b) The certainty map upon reconstruction of the noiseless image. White areas are those scored as recoverable. (c) The certainty map upon reconstruction of image + noise. (d) Closing small gaps between pixels in (c) as an aid to visualize areas with a higher density of points [22].

3.3.3.3 Other Limitations

In addition to noise, also other uncertainties can cause the algorithm to fail. Scanning too fast can introduce feedback overshoots, especially on step like features. Varying friction of the tip or degradation of the tip during imaging can cause problems. In the atomic level, the standard model may fail because of compressibility (AFM) or work function (STM) changes.

4. Scanner Induced Limitations

The utilization of piezoelectric materials in the scanner is what makes atomic imaging possible. Unfortunately, these materials behave well only when used for atomic scale scanning and as soon as the covered area becomes larger, the scanner shows severe nonlinearity. If the scanner is not properly calibrated, image distortion is easy to observe when a known characterizer is used. One should be aware not to assume that the scanner nonlinearity decreases when small scale imaging is done. If the scanner is not calibrated for the used scale, the calibration routine may loose its validity and this results in an increase of relative uncertainty. Accuracy problems have to be treated according to the application and the special demands. No general way can be applied and therefore knowledge about flaws and how to work around them is essential for interpreting images in a correct way.

The SPM companies have solved the problem of nonlinearity quite well and the instruments are sold precalibrated. Therefore it is sometimes forgotten that piezoelectric material is not easy to monitor in its movement. Severe problems may arise if a large scan size scanner is first used to spot an area of interest and then to go for a more close look by zooming in. This can lead to huge uncertainties (*e.g.*, 30%) which are not expected. The demand to develope scanners with bigger and bigger scan ranges is limited by the appropriate calibration routines which are valid for the largest scan size as well as for the atomic scale. So far, it is usually better to have several calibrations which apply for the used scan size and not to rely on the validity of the calibration routine over the full range.

To complicate the topic further, it should always be kept in mind that every scanner has its own "personality". The calibration for one scanner is not applicable for any other (even with similar scan size) scanner. Because of the large number of possible calibration routines, the chapter after introduction of nonlinearity is devoted to different calibration schemes introduced.

4.1 Motion of the Scanner and its Nonlinearity

4.1.1 Scanner Motion and Image Formation

Care should be exercised in presentations where scanner motion and captured images are mixed. This leads easily to misinterpretations because the image resembles the voltage applied to the scanner (constant force mode) while the scanner movement is measured relative to something else.

Scanners are either reverse or forward biased during production. A reverse biased scanner will contract as soon as a positive voltage is applied, and the forward biased will extend. In the following, only reverse biased scanners will be discussed. When height variations are presented, the graphs apply only to constant force mode, because in constant height mode the scanner does not move at all in Z-direction.

In *Figure 4-1*, a line profile is given where
  (a) is an imaginary surface profile which is tracked by an AFM
  (b) is the scanner motion according to the surface profile
  (c) is the voltage applied to the scanner and equal to the surface profile
     because of the reverse biased scanner

(a) surface profile

(b) scanner movement

(c) voltage applied to scanner

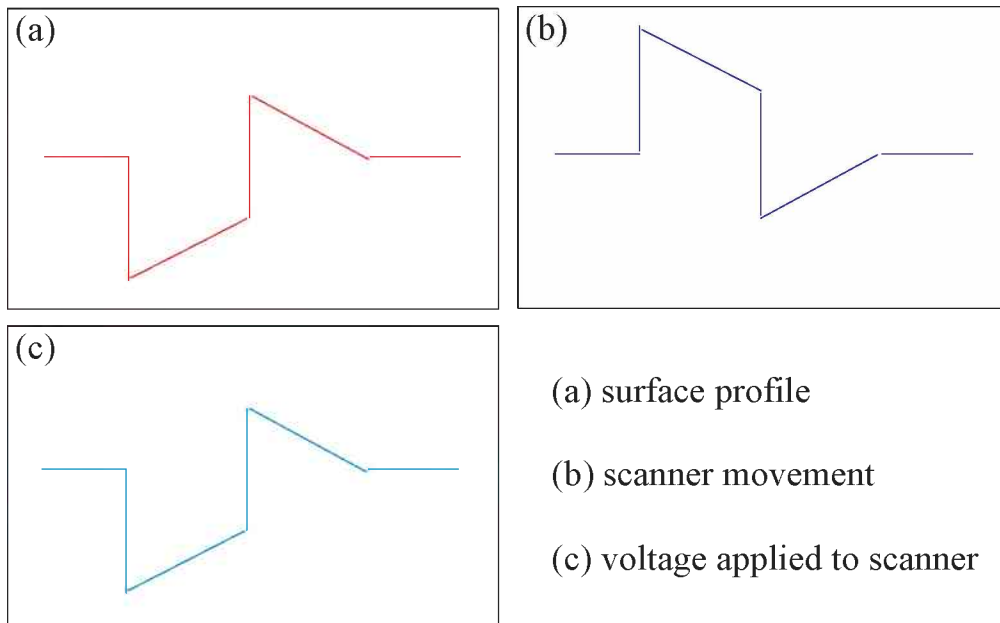*Figure 4-1 Illustration of surface, scanner movement and applied voltage profiles in AFM.*

To maintain constant deflection (constant force), the feedback system has to react very fast by applying voltage to the scanner when one of the above steps is traversed. This voltage is used to reconstruct the image and this poses a problem, because the motion is not linear with the voltage and is also time dependent.

4.1.2 Intrinsic Nonlinearity

The actual response of piezoelectric material is usually given as nm/Volt and is referred to as sensitivity. While on small scale the sensitivity is quite linear, on large scale deviations from this linearity can occur. The reason for this was already presented in *Figure 2-2*. According to this, the sensitivity of the piezo is a function of the applied voltage.

In *Figure 4-2* this is given again in a more schematic way. The more the voltage on the piezo deviates from zero Volts, the shorter is the movement of it.
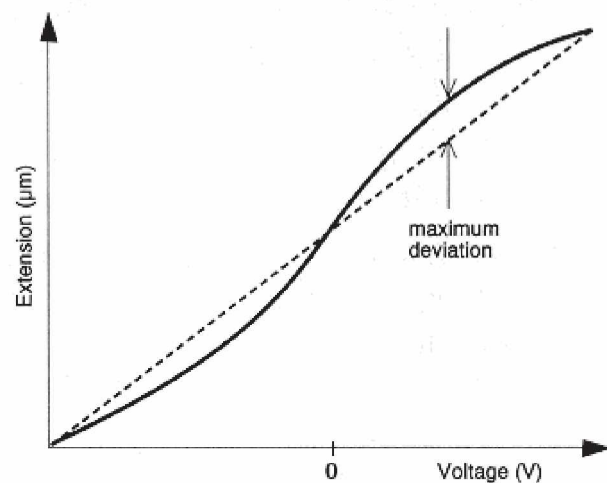
Figure 4-2 Deviation from linear behavior caused by intrinsic nonlinearity [1].

Intrinsic nonlinearity is a phenomenon which is quite hard to measure, and to avoid any other effects, special care has to be taken in the experimental setup. In *Figure 4-3,* where a curve for vertical movement of a scanner is presented, slight S-shape of the curve can be observed, manifesting intrinsic nonlinearity. (The reason for the hysteresis will be discussed in the next chapter). Z-motion was accomplished by applying voltage to all four segments of the scanner in a very slow stepwise way, using 0.3 V/s corresponding to 27nm/s.
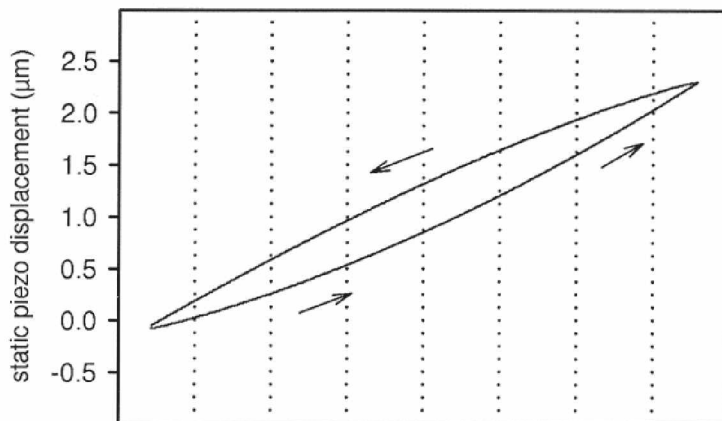
Figure 4-3 Intrinsic nonlinearity and hysteresis of a tube scanner, when the amplitude of the static displacement is monitored [27].

In practice, intrinsic nonlinearity should be treated in connection with other non-linearity effects.

4.1.3 Hysteresis

As already mentioned when *Figure 2-2* was discussed, hysteresis is due to the basic properties of ferroelectric materials. When voltage is applied to the piezo, starting from zero and going up to some value and then back to zero, hysteresis is easily observed. The actual amount of hysteresis, as it is expressed in *Figure 4-4,* is again dependent on time. The experimental curve in *Figure 4-3* is an example how hysteresis and intrinsic nonlinearity affect the motion of a piezoelectric scanner.

Figure 4-4 Hysteresis of a scanner [1].
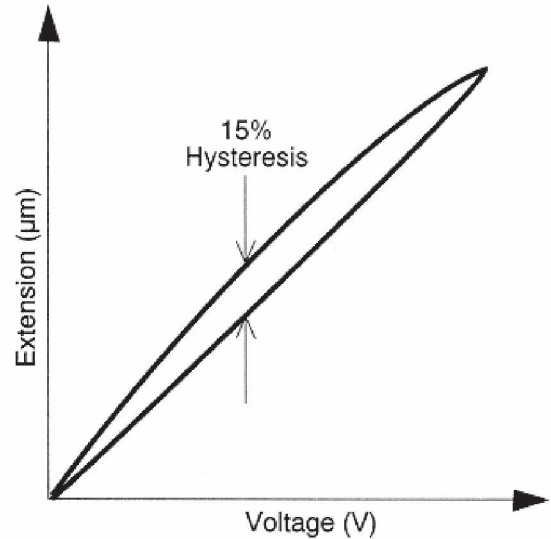
Considerable effort has been made to make tube scanners suitable for SPM techniques. An example of the effects of the piezo material properties is given by the two curves in *Figure 4-5*. It is apparent from the curves that *(a)* has a much smaller hysteresis. However, the applied voltage for material *(b)* is much smaller for the same extension, which means that it has a higher sensitivity. If now the maximum required scan size of the piezo is rather small, material *(a)* is better, but as soon as the required scan size is increased, material *(b)* is needed to reach the desired extension.
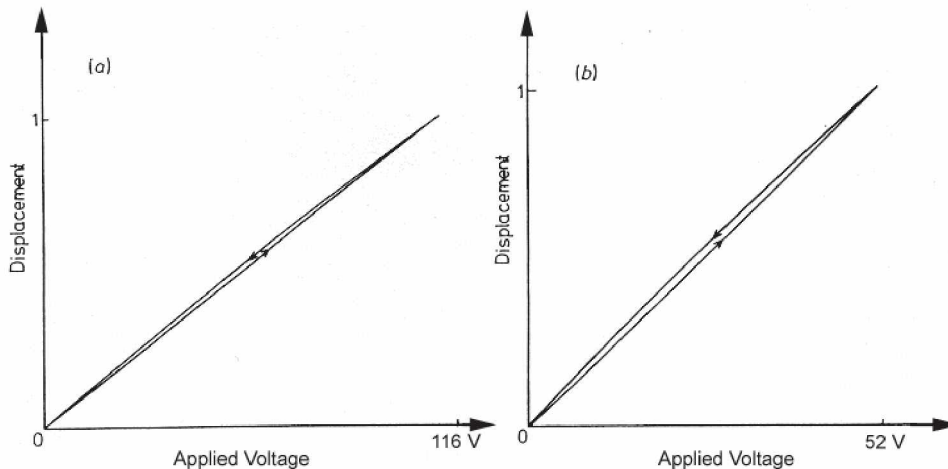
Figure 4-5 Hysteresis loops of (a) a PZT-4 tube and (b) a PZT-5H tube [28].

4.1.4 Creep

Creep is the time dependent displacement of piezoelectric materials and can easily be observed when taking SPM images. The easiest way to observe this is to use the Zoom command which allows one to jump to the location of interest. However, due to creep it may happen that the targeted location is missed. The actual displacement can be divided into two parts, an instantaneous response (Δx) when a voltage step is applied and a slight creep component (Δx_c) that increases the displacement further with time (*Figure 4-6*). Fast response to Δx is the reason why piezoelectric materials are suitable for SPM applications but the creep rate and also the time how long creep is present (can be up to 15 minutes) must be taken into account. Sometimes a characteristic time (T_{cr}) is used to express how long creep (usually given in percent) occurs after application of a voltage.
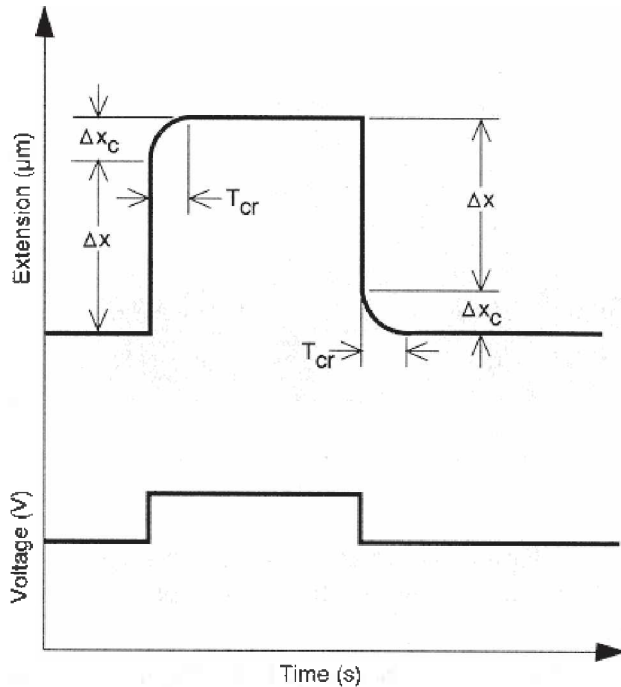
Figure 4-6 Scanner extension due to a voltage step - the instantaneous extension is followed by a creep component [1].

The effect of creep to images is quite easy to explain when considering a step like feature. As soon as the tip reaches it, the cantilever will bend and the feedback control will put a voltage to the piezo which makes it contract. This voltage corresponds to the full step height, but the following creep (further contraction) forces the feedback to reduce the applied voltage on the scanner to keep the deflection constant. As mentioned earlier, the image is constructed by using the voltage applied to the piezo, and therefore it will look like the dashed line in *Figure 4-7*.

Figure 4-7 The shaded area is the surface cross section and the dashed line shows the captured image trace [1].
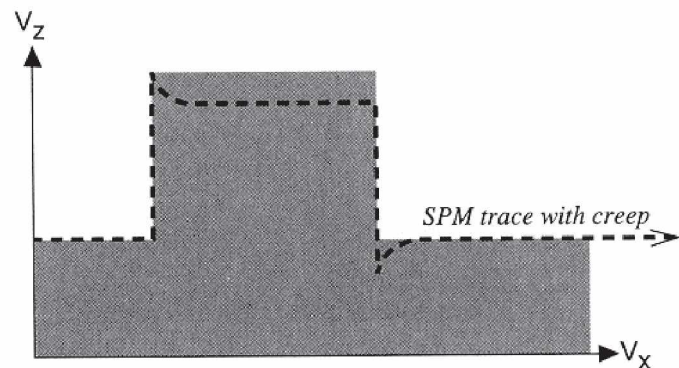
Considering the hysteresis, creep starts to be important when the voltage steps become bigger. The maybe surprising effect is that the hysteresis will become smaller with creep present. Difficulties arise from the different time scales used to present the hysteresis curve and the effect of creep. To get one hysteresis curve like that in *Figure 4-3*, it may take more than 10 minutes, while a creep curve like that in *Figure 4-7* may take only 1/10 of a second. The actual effect of hysteresis in an image would lead to an image shown in *Figure 4-8*, but due to rather small hysteresis and the always present creep, this effect is generally not observable in real images.
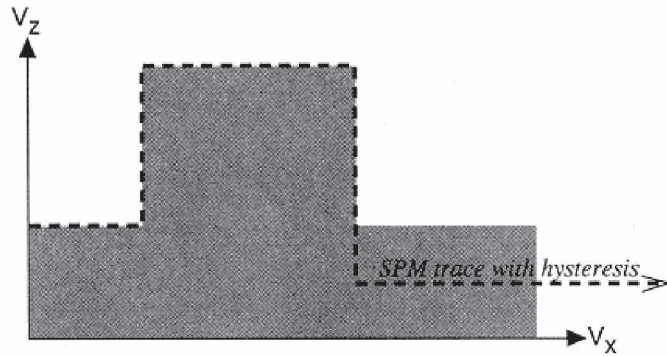
Figure 4-8 Theoretical effect of hysteresis on a step, when creep is not present [1].

It seems that in real images creep contributes most strongly to image flaws, while hysteresis is less important and intrinsic nonlinearity becomes important only when high extensions are used. Although already visible at an extension of 2.5μm (like in *Figure 4-3*), the really strong S-shape is seen only in extensions which are close to 10μm. That is also the actual limit for the vertical scale of the scanner. Both hysteresis and intrinsic nonlinearity are maybe masked by the influence of creep, when real imaging is done.

When investigating these nonlinearities, it whould be easiest to watch the scanner dilation only, but with a tube scanner, the lateral movement is due to the bending of the tube. Only the vertical motion is made by dilation and therefore only the vertical uncertainty can be corrected after investigating the contributing influences separately. In contrast to this, in lateral directions the problem is so complex that it is impossible to develop experiments which whould reveal the nonlinearities separatly. The effects of all contributing nonlinearities have to be put into a complex model to enable the simulation of the correct position of the scanner when voltage is applied to it. The possible methods to correct the vertical nonlinearity are presented in the calibration section 4.2.

4.1.5 Cross Coupling

Cross coupling includes several factors that cause Z-motion although there is no topographical reason for this on the surface. For a scanner having no separate Z-piezo, coupling of the X- and Y-voltage will cause the tube to elongate or to contract. This is the reason why it is better to have an additional piezo-element for z-motion. This coupling is generally neglected, probably because of the difficulty to get proper correction factors.

Another important factor is geometrical, which depends on the way the scanning is accomplished, that is, whether the sample or the probe is moved. When the probe is scanned over the surface, it will describe an arc as shown in *Figure 4-9*. Moving of the sample describes in principle the same curve, but results in a reversed image, as demonstrated in *Figure 4-10*.
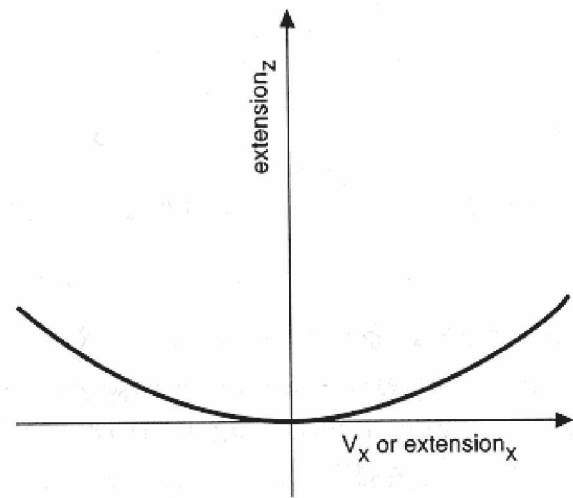
Figure 4-9 Geometrical effect from the bending of the tube scanner [1].

The bending of the tube in *Figure 4-10 (a)* moves the sample and produces an additional height variation. The flexible cantilever will deflect and the feedback makes the scanner contract to get back to the original deflection. In the image this will look like a positive curvature. When moving the probe with the scanner (*Figure 4-10 (b)*), the deflection of the cantilever is reduced due to the lift off. The scanner has to extend to keep the deflection constant and therefore the image will have a negative curvature.

(a)

Image trace

(b)

Image trace
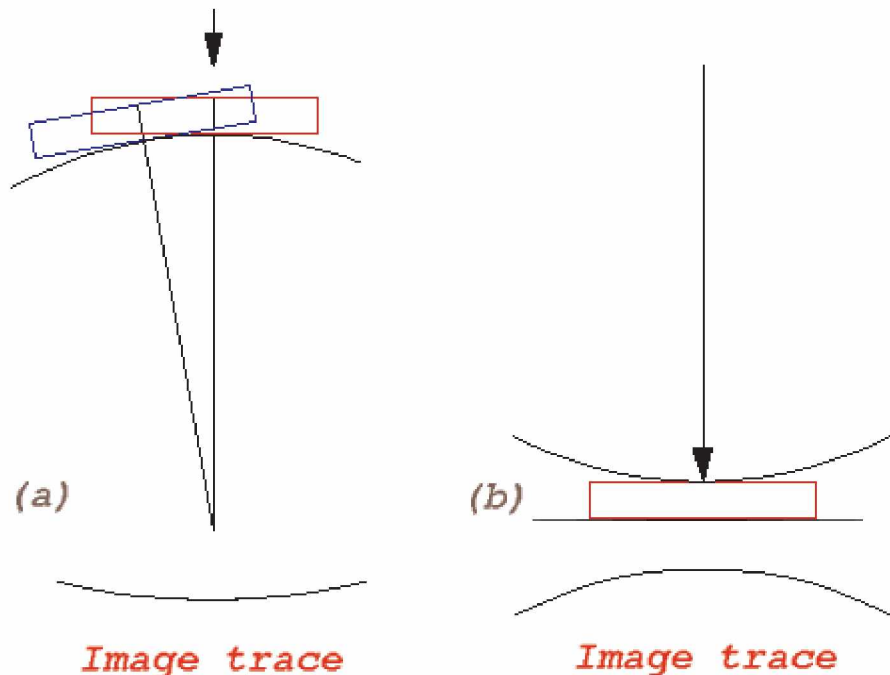
Figure 4-10 Changes in the image due to the geometrical effect of tube bending when (a) sample is moved (tip offset for clarity) and (b) probe is moved.

4.1.6 Aging

It was already mentioned that the scanner has to be used on a regular basis in order to keep the dipoles aligned. When the scanner is not used for a longer period of time, the possible maximum extension will decrease, as illustrated in *Figure 4-11*. Regular use will increase or keep the reachable extension constant.

Even small decrease in the performance can have a big influence on the value of the scanner. For example, if the lateral scan range is rather big (more than 100μm), a decrease in the vertical scan range from 5μm to 3.5μm will cause problems in the alignment of the instrument relative to the sample. If they are not parallel, the vertical range may be exceeded just by the misalignment and the tip looses contact to the sample.
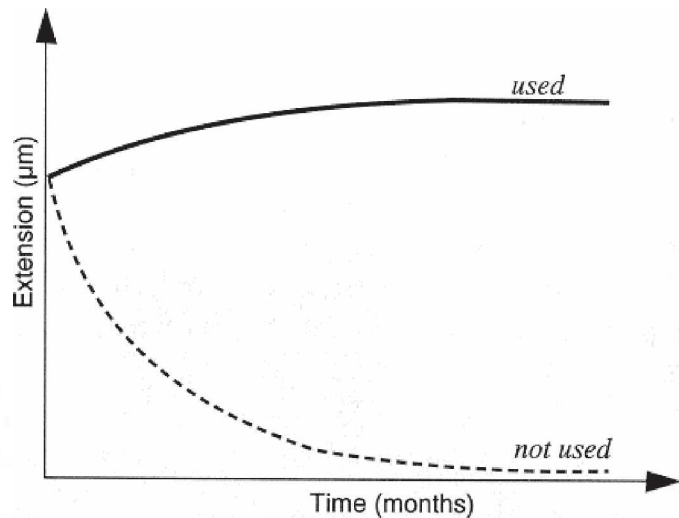
Figure 4-11 Dramatical loss of aligned dipoles due to aging [1].

4.2 Ways of Correction

There are many ways to enhance the image quality and to produce more accurate images. In the following, three different possibilities, with increasing price tags, are discussed.

4.2.1 Recalculation of Captured Images

This is the method typically used only for small scale imaging. As soon as the scan size increases, it is difficult to work with it, because the image already shows distortion during capturing and it is more difficult to spot the area of interest. Nevertheless, this way has definite advantages, because no special controller units have to be designed and recalculation routines are freely available from people who have already built and worked with self-built SPM's, where the calibration had to be developed by themselves.

Calibration routines always need known structures to define parameters, which can be used for recalculation. When SPM's are built for small scale imaging it is usually no difficulty to resolve atoms (otherwise the instrument is quite worthless) and this can be used as a known structure.

Regularly used structures are highly oriented pyrolytic graphite (HOPG) or mica. When imaging one of these structures the result may look distorted, but because of the known lattice constants, the real shape can be restored.

To define the actual distortion, Fourier transformation offers a practical way to make this recalculation more automatic. In *Figure 4-12 (a)*, the image of HOPG does not show the typical hexagonal structure. For a distorted image the Fourier transformation gives a skewed ellipse instead of a circle. After recalculation (*Figure 4-12 (b)*), the hexagonal structure is clearly visible and the Fourier peaks form the desired circle.

(a)

(b)
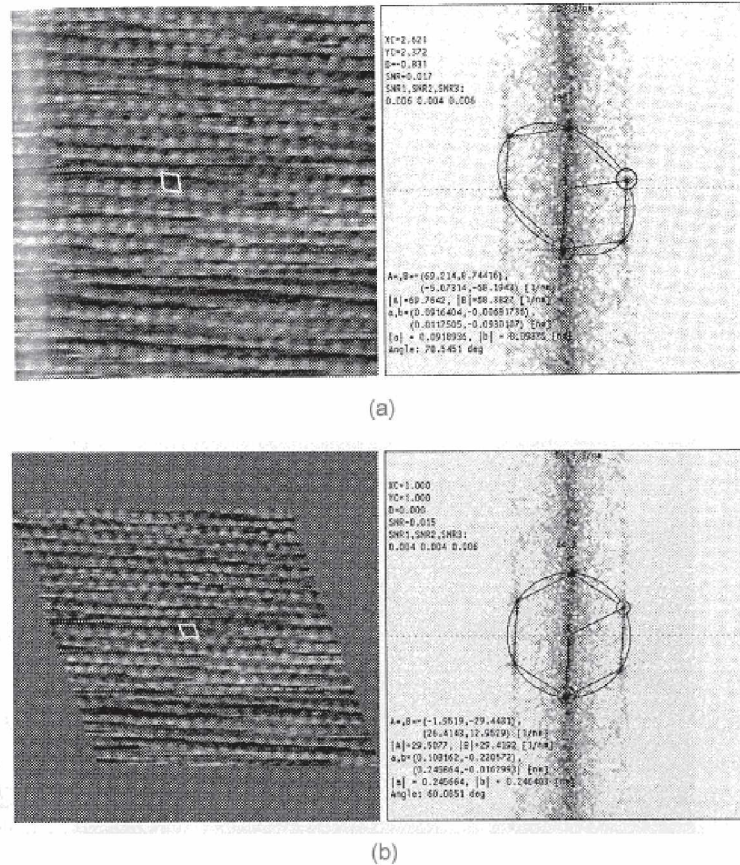
Figure 4-12 Image of HPOG (a) heavily distorted image (left) and its Fourier transformation (right), (b) recalculated image (left) and the corresponding Fourier transformation [29].

These recalculation routines are usually very specific and try to eliminate the most disturbing effects on a specific problem. Whenever a different scan rate or scan size is used, compared to the one used for evaluating the calibration parameter, the routine may loose its validity. Then the calibration has to be carried out again or even a new routine has to be developed.

4.2.2 Real-time Software Correction

Real-time software correction is the most usual correction technique implemented to SPM's.

Every try to calibrate a scanner just by observing the movement after applying linear voltage ramps to it, with neglecting the time-scale, will lead to a limited validity calibration routine. Changes in scan size as well as scan rate will affect the time-scale and therefore, when neglecting the time, a parameter look-up table has to be defined for every possible scan size and scan rate combination.

The concept of a look-up table is used for commercial as well as for self built instruments. A schematic illustration of a look-up table is given in *Figure 4-13*.
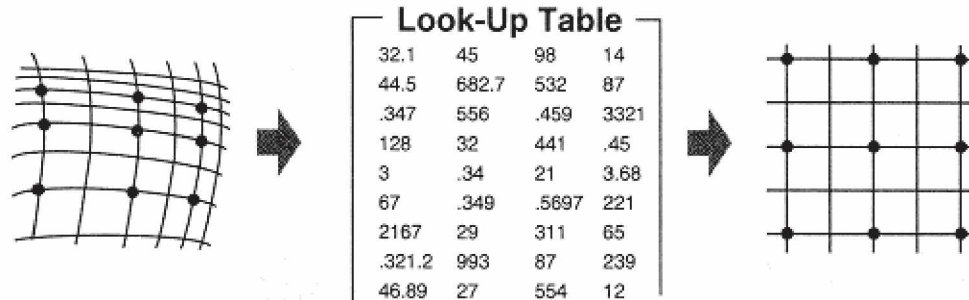
Figure 4-13 Software correction, using a look-up table [1].

The second, more difficult approach is to use a model which includes the time-scale.

Companies have developed quite complex calibration routines which are used to control the voltage applied to the scanner. The following linearity correction is used in Digital Instruments [30] SPM's and it is probably valid for other systems in a similar manner.

In this method the scanner is driven in a linear way by applying a nonlinear voltage curve (further called scan waveform) to it. To define the necessary parameters, the user has to scan a known structure (silicon grating) and adjust three parameters for the X- and Y-axes. Autocalibration routine is activated to scan the known structure several times and to capture images at different scan sizes, scan rates and scan angles. From these images the user has to measure the grid spacings and correct manually the measured values according to the known dimensions of the grid. The program extracts from this information appropriate correction parameters which are stored in a parameter file.

When the scan waveform is applied to the scanner, it consists of a linear ramp added to a decaying exponential of varying magnitude and argument. The following formulas illustrate one part ("linearity correction") of Digital Instruments' calibration routine. It is clear, that the full calibration routine is a secret of the manufacturer.

The three parameters mentioned for either axis are:

$$Mag0 \text{ ---- Correction magnitude}$$
$$Mag1 \text{ ---- Correction magnitude derate}$$
$$Arg \quad \text{ ---- Correction argument}$$

Arg determines the exponential decay of the nonlinear term. *Mag0* and *Mag1* determine the amount of the exponential term to use, and how that amount varies with the scan size.

The used formulas are:

$$B = S * (Mag0 - Mag1 * \ln(S))$$

$$C = \frac{1}{1 + \dfrac{B}{Arg} * (1 - e^{-Arg})}$$

$$x' = C * (1 + B * e^{-Arg*t})$$

$$x = C * (t + \frac{B}{Arg} * (1 - e^{-Arg*t}))$$

B - A derated correction magnitude
C - A normalization constant
S - The scan size
x' - The derivative of the scan waveform
x - The scan waveform
t - The independent "time" variable of the scan

An example how the applied scan waveform could look like is given in *Figure 4-14*. The normalization is applied so that S=1 corresponds to the full scan size (440 volts), x=1 and t=1 corresponds to the scan having made it once across (time for one scan line).
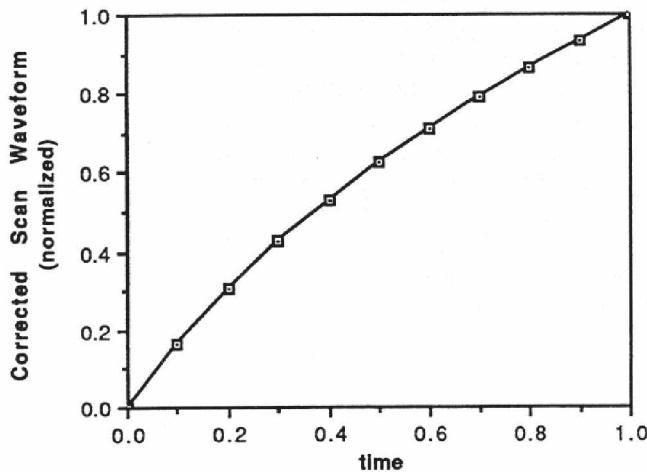
Figure 4-14 Example for a corrected scan waveform [3].

The important point in these formulas is that time has a major influence on the voltage applied to the scanner to get a linear movement.

The calibration routine discussed so far is only for the X-Y plane and does not include the Z-axis. However, if cross-coupling is present, it should be removed from the images before evaluating the X-Y plane calibration factors.

When the maximum voltage is applied to the Z-piezo (440V), the extension varies depending on the used scanner. To correlate the applied voltage value to specific metric dimensions the already introduced sensitivity value (nm/V) has to be determined. The most reliable way for this is to image a known feature and to measure its height. The measured height is then compared to the real one and the sensitivity value is appropriately changed .

The calibration standard delivered with the instrument is a platinum-coated silicon grid with vertical features of 180 ±3 nm. When one desires greater accuracy, an appropriate standard has to be chosen. Unfortunately, the Z-axis calibration is very sample-dependent and it is not easy to control Z-piezo dynamics because the Z-piezo does not move at a constant rate, as the X- and Y-axes do during scans [31].

Digital Instruments recommends that for instrument stabilization reasons several frames (Up and Down scans) should be made before the image from the 180nm deep pits is taken. The Z Center Position should be zero and the scan size should not be bigger than 40μm (depends on the type of scanner). After capturing, the image should be Flattened and X-Plane Fitted with first order functions. By invoking the Bearing command the user should draw a rectangle around 1-4 pits and after activating a second arrow, the depth can be measured.

$$Z-sensitivity_{new} = \frac{180nm}{measured\ depth} * Z-sensitivity_{old} \tag{14}$$

The new Z-sensitivity should be checked weather it gives correct depth values, otherwise eqn. (14) has to be applied again until the desired accuracy is reached.

The change in the sensitivity of the scanner is corrected by a Retracted and Extended Offset Derating factor. When measuring the 180 nm steps at Z Center Positions, 100 and -100 Volts, the depth will be somewhat different. With eqn. (15) these two derating factors can be calculated.

$$offset\ derate_{new} = (1 + offset\ derate_{old}) * \frac{180nm}{measured\ depth} - 1 \tag{15}$$

The two Z Center Positions, 100 and -100 Volts, have to be adjusted with the motor for the NanoScope AFM and by hand for the StandAlone AFM.

4.2.3 Real-time Hardware Correction

Metrology applications need the best possible accuracy which usually cannot be guaranteed by software correction. Additional sensors attached to the SPM are used to monitor the scanner movement. This information is sent to the feedback control which in turn controls the voltage applied to the scanner to reach the desired position. In *Figure 4-15,* such a setup is illustrated.
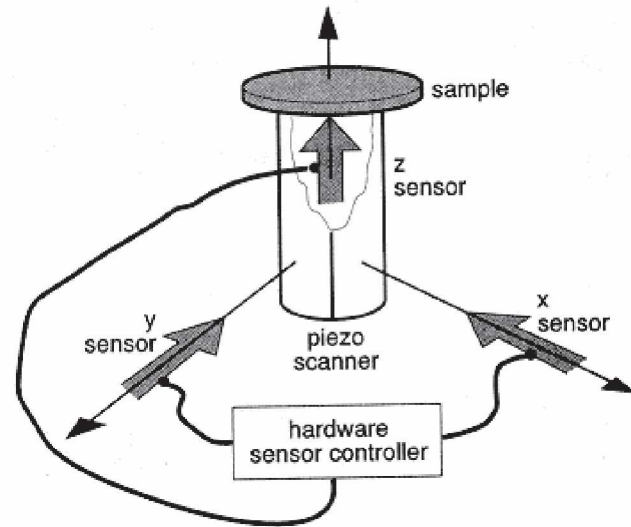
Figure 4-15 Schematic hardware setup with external detectors [1].

The sensors themselves have to be stable in operation and properly calibrated. Furthermore, they should not exert any force to the scanner, which may affect the motion, and also the noise level should be as low as possible. Therefore the sensors are usually capacitive in nature, although there might be the problem with a proper alignment of the capacitor plates. Also frequently used are optical sensors, where interference or a direct signal (*Figure 4-16*) is used to monitor the scanner motion. The performance limit for optical systems lies in the relatively high noise level of the used LED's.
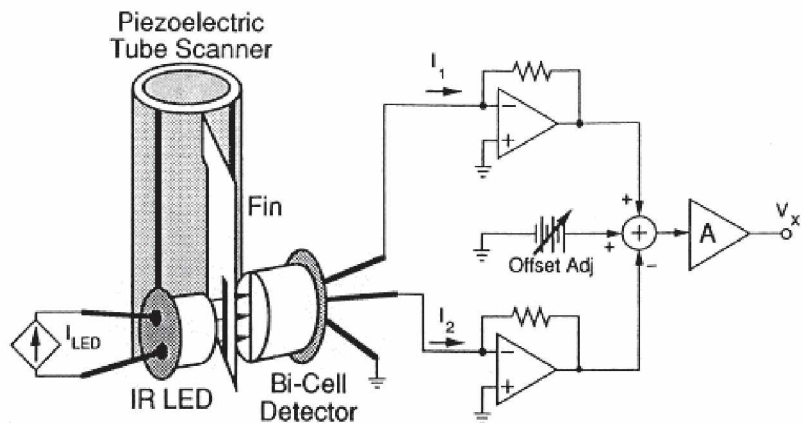
Figure 4-16 Optical displacement sensor with a split photodiode illuminated through a slit in a fin (only one of two perpendicular fins is shown) [32].

A common problem for all hardware correction systems is that they violate the Abbé principle (*Figure 4-17*). Before flexing the tube, the point to measure is at position x_0. When deflecting the tube, the position monitor will give a value corresponding to x_1, but the actual point position is x_2. This introduced extra motion, x_2-x_1, has to be corrected using numerical methods.
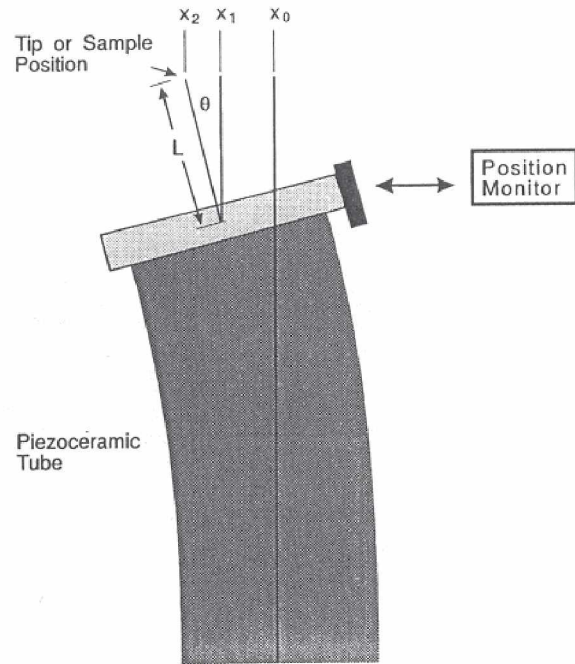
Figure 4-17 The Abbé principle violation due to the flexing of the tube scanner [16].

Systems with hardware correction are in principle, best as far as the accuracy is concerned. Their only drawback is the considerabye high price. So far there has to be a really good reason for purchasing such a system.

II Experimental Part

5 Instrument Description

5.1 NanoScope E-Controller and Computer System

The SPM system discussed in this work is based on the NanoScope E-Controller from Digital Instruments [30].

Possible operating modes with the E-Controller are:
- ✓ Contact AFM (with a fluid cell also in liquid environment - possible at CEM)
- ✓ STM
- ✓ Lateral Force Microscopy (special AFM head needed - not at CEM)

The computer system is, in principle, the same for all Digital Instruments Nano-Scope Scanning Probe Microscopes except for some special (or additional) features for some of the offered imaging modes. It consists of a Pentium (120MHz) with two graphic cards (for the two 15" monitors) (*Figure 5-1*) and the necessary controller card plugged in the ISA slot of the computer. The operating system for the used software (Version 4.22b4) is MS-DOS 7.0. In *Appendix A* is the block diagram for the STM and AFM.

Figure 5-1   Picture of the controller and the computer system.

5.2 NanoScope AFM and STM

Nanoscope AFM/STM is one of the best selling devices on the rapidly growing SPM market. It is constructed for both AFM and STM use and therefore offers an inexpensive way to get both imaging modes simply by changing the head.
The AFM head in combination with the E-Controller is intended for contact mode only. With the additional top view system (camera plus TV-screen) the positioning of the laser on the tip is easier and the regions of interest are also easier to find. The scanner (Type E) has a maximum scan size of $13{\times}13\mu m$ and a vertical range of about $3.2\mu m$. To get the full range of motion, $\pm220V$ is applied to either axis. Damping of vibrations is accomplished with a very heavy rubber damped table and the silicon rubber legs of the AFM stage (*Figure 5-2*).
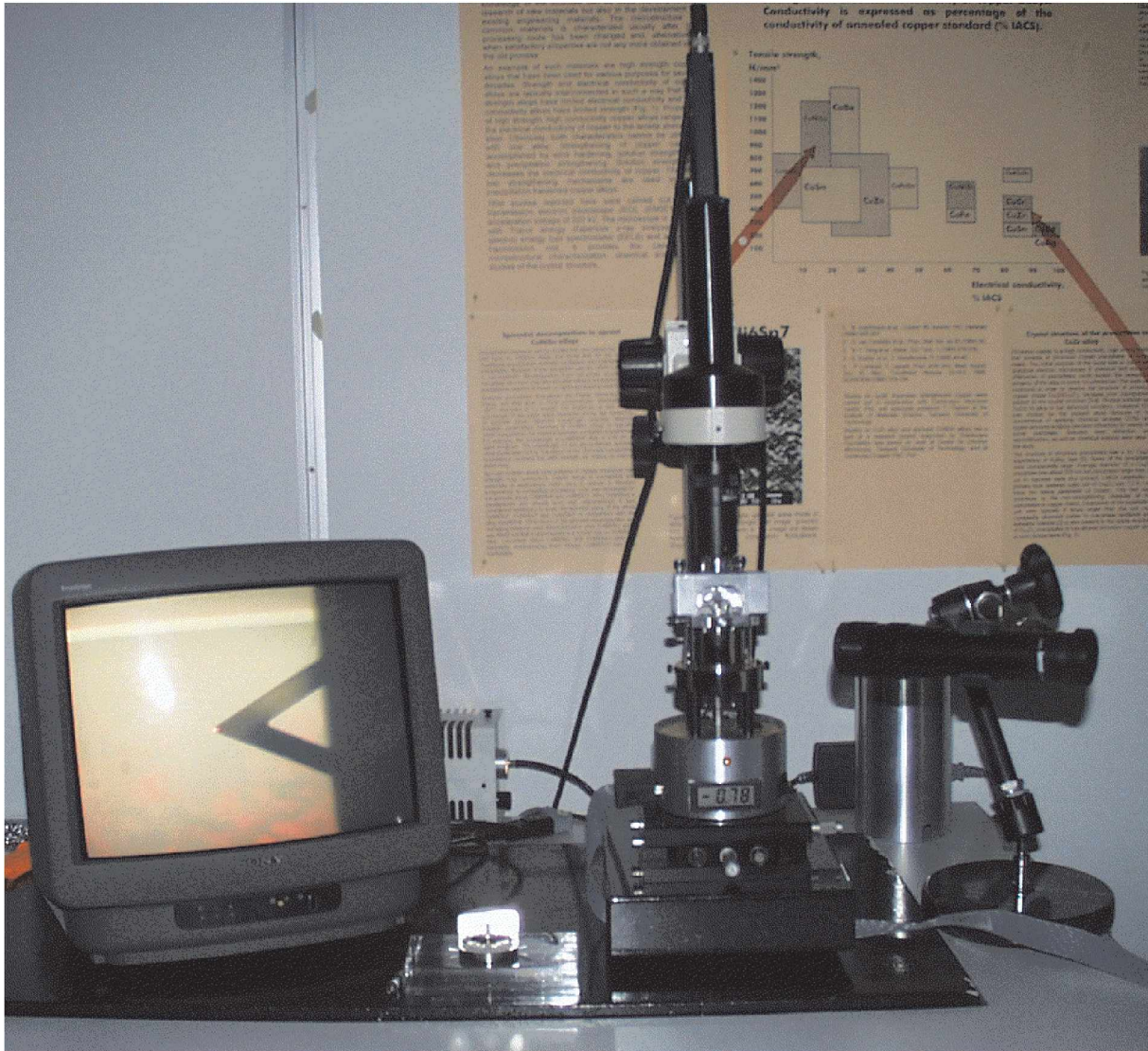
Figure 5-2 Picture of the Nanoscope AFM/STM and the Top View System. The AFM head is mounted on the instrument and the STM head is lying on the table.

The optical sensing system used in Nanoscope AFM works with a split photo-diode as described in *Figure 2-7*. The coarse approach is done with two adjustment screws in the base (*Figure 5-3*) and with a motor driven screw. During this, the optical head is lowered towards the sample, and special care should be addressed to their parallel alignment. The final tens of micrometers are approached by the motor, initiated by a software command. As soon as the probe touches the sample and bending is sensed, the motor stops and the scanner starts to move the sample according to the preset values.
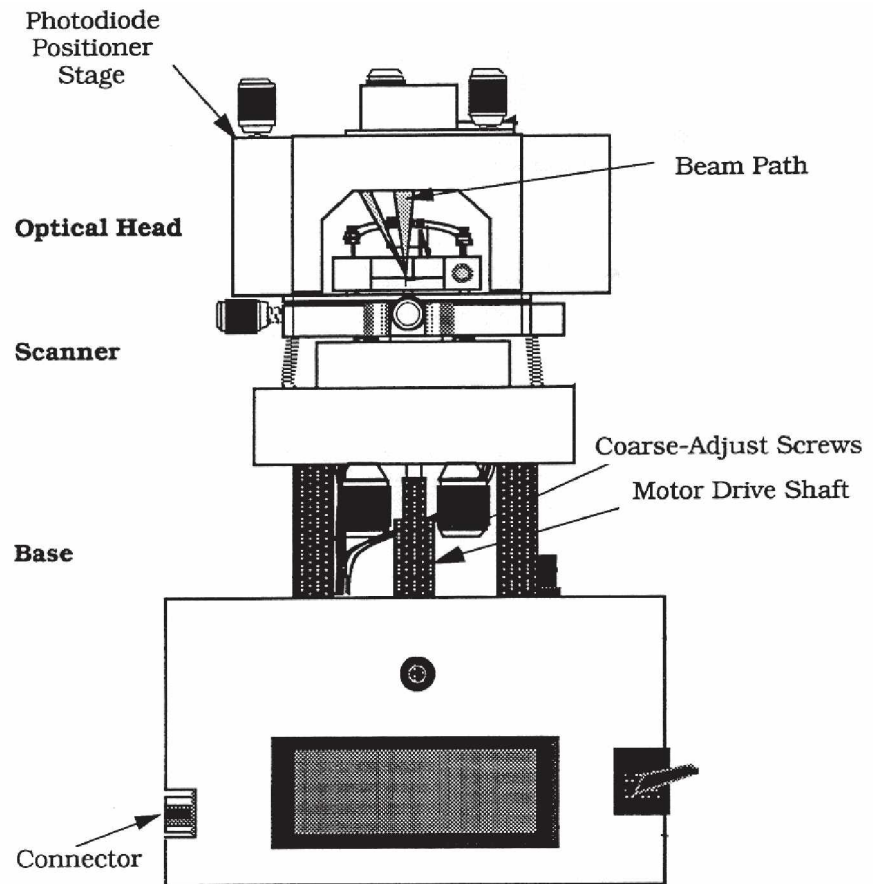
Figure 5-3 Front view of NanoScope AFM with its major parts [11].

5.3 NanoScope StandAlone AFM

Digital Instruments built this type of AFM only till 1995 and it is not further supported by current software developments. The present model is from 1991 and is equipped with a J-Scanner ($130 \times 130 \mu$m scan size and vertical range of about 3.5μm). Also here, to get the full range of motion, ± 220V is applied to either axis.

StandAlone AFM is designed to be placed on samples which are too big for the standard Digital Instruments AFM. That means larger than 15mm in diameter and higher than 5mm. It is a very compact instrument and the manufacturer talks even about portability, but unfortunately it needs the whole computer and controller system also and therefore portability is a little bit overemphasized.

However, large samples do not have to be cut into smaller pieces and it can therefore be seen as an approach to nondestructive testing.

StandAlone AFM was initially intended to be the major target of the present work.

At the beginning of the work, various hardware and software problems had to be solved in collaboration with the manufacturer, which delayed actual research work considerably. The instrument had to be shipped for repair back to the Netherlands before it could be properly operated using software version V4.22b4.

Figure 5-4 Cross section view of the StandAlone AFM [12]. Photos about this instrument will be presented later.

Height-Adjustment Screws

Piezo Tube

Scanner Shell

Cantilever and Deflection-Detection System

Piezo Shield
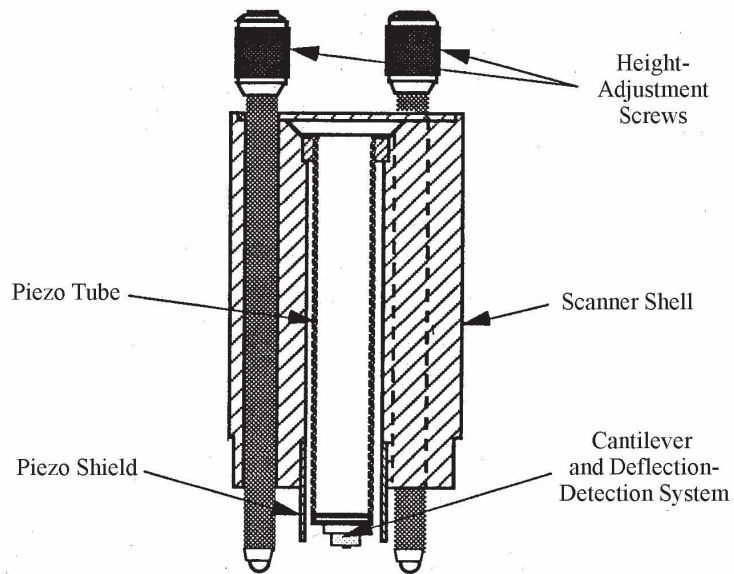
For the coarse as well as for the fine approach screws going through the shell are used. One of them has a fine thread and is used for the final tens of micrometers, while the other two have a coarse thread. The instrument is placed onto the sample or the sample has to be placed between the screws. The software offers a so called force calibration mode which is used for engagement of the StandAlone AFM. In this mode, the scanner moves the Deflection Detection System (which works in the way described in *Figure 2-8*) only vertically. When the operator lowers the shell using the screws, the contact with the sample will result in interference fringes shown on the computer monitor. Switching to normal imaging mode will start the lateral scanning in the pre-set scan range.

The exact step by step routine is further described in the rewritten manual supplied on the enclosed CD.

6 Development of an Investigation Routine

The use of several different instruments and research routines to study a given problem is always encouraged, because this frequently leeds to more profound understanding of the research subject.

At the CEM, there are several good instrument standing close to each other, and therefore it is just obvious that one starts to think about their combined use. For example the normal AFM has the limitation of the specimen size, but with the StandAlone AFM, also bigger specimens can be investigated. In addition with SEM (Philips XL30), equipped with EDS, EBSD, μ-Indenter 5 [49] and a big motorized stage, one can supplement the AFM studies by obtaining the chemical composition, orientational date and surface hardness profiles of the specimen.

The idea was to investigate samples from deformation experiments with the StandAlone AFM in combination with orientation measurements using the EBSD system. The motorized stage allows controlled indentations, which can be used for marking locations of interest and using them to find the same area with the StandAlone AFM.

The first main objectives were to design specimen holders for the StandAlone AFM as well as for the SEM.

6.1 StandAlone (Sample) Holder

Although the StandAlone AFM is designed to be placed on large samples, experience shows that it needs a very gentle hand not to destroy the deflection detection system. For safe operation of the device, the following points for a holder-design should be considered:

⇨ The StandAlone should stand on its own support plate so that the instrument and the sample are not coupled when one of them is moved.
⇨ It should be possible to watch the coarse approach towards the sample surface with the monocular.
⇨ The holder should be usable for a wide variety of specimen sizes (height as well as lateral dimension).
⇨ It should be possible to move the sample in a defined way.
⇨ Samples for the normal AFM should be usable without any change.
⇨ Flat tensile or fatigue test samples should be rigidly mountable and it should be easy to use also other geometries.

6.1.1 First Holder Concept

The first design fulfilled already all the necessary points from the list above. Also, due to the rather high position of the StandAlone AFM, it was easy to watch the approach with the monocular in a operator-friendly way.

The original 12mm acrylic plate was too thick for the instrument to reach the sample surface and therefore an opening was made so that only 2mm of material was left to work as a support onto which the AFM screws were placed. The distance between the AFM support-plate and the sample was defined by four spacer pipes (different heights are available) and the sample was placed on a micrometer stage. Positioning was therefore easy, and even turning, which is useful for aligning the calibration grating, was possible. Old samples, prepared for the normal AFM are either magnetic themselves or are placed on a magnetic stub. Therefore a magnet was immersed into the micrometer stage. For the flat strip specimens a special clamping system was made, which itself was mounted on the micrometer stage.

All the major parts of the holder were made from acrylic glass (*Figure 6-1*).

Figure 6-1 First design for the StandAlone (Sample) Holder.

Nevertheless, obtaining good images was a problem with this particular design. When the instrument was working, the Z Center did not stay constant, which means that the distance between the instrument and the sample was not constant. Therefore the microscope had to be readjusted with the screws during the operation. First the AFM was suspected to cause this malfunction, but soon it became clear that the acrylic glass was the reason for this behavior. The temporary solution (before building a new holder) was to replace the 2mm acrylic support with a steel plate (*Figure 6-2*). Unfortunately, there was no significant improvement, implying that the whole holder was "waving" unpredictably. Therefore a new holder had to be built.

Figure 6-2 The steel plate improved holder. Top-front view towards the opening for the microscope shell (red) as well as for the piezo shield (blue). Three microscope positioning holes can be seen on the steel plate.
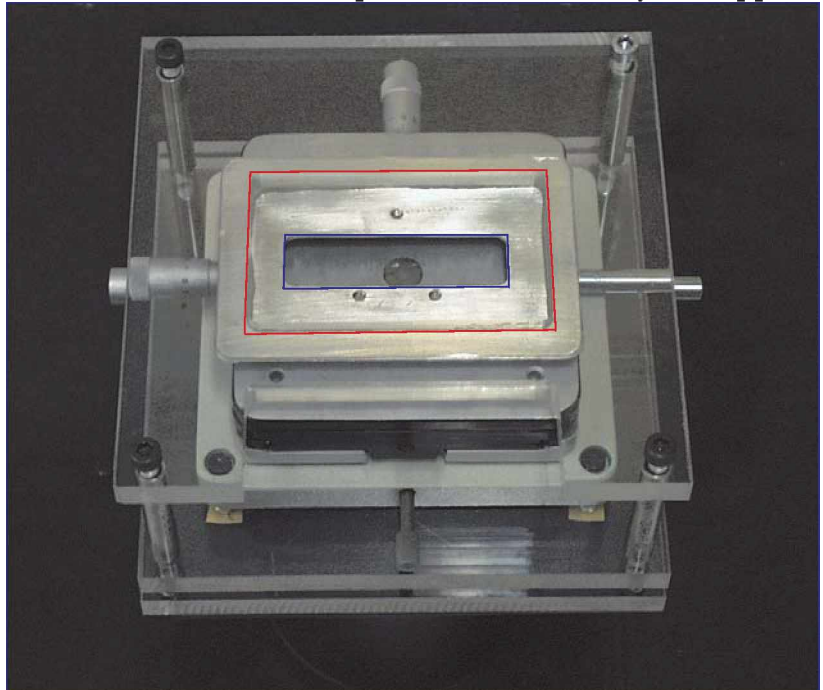
6.1.2 Final Holder Concept

Because of the mentioned problems, the structure of the new holder had to be more rigid and also the damping properties had to be considered more thoroughly.

Sometimes it is recommended to place the AFM onto a support which hangs on rubber-bands. This concept was rejected because the StandAlone has to be engaged by hand, and therefore it is too easy to destroy something just by hitting this support. Also stacks made of plates with viton spacers in-between have been proposed. When one takes a look at the normal AFM setup, it is apparent that sufficient damping is obtained by relatively low effort. The instrument stands on a granite plate with very soft silicon rubber legs.

This damping method was taken as a recommendation when building the new holder. For the legs, different trademarks of silicon were bought, of which one was suitable for making the desired legs. The silicon was molded into pipes and also into petri dishes to give them a defined shape. The material for the support plate was a 15mm thick aluminum plate. The micrometer stage was screwed to a granite plate.

The two designed holders were basically similar but made of different materials. The four main parts in *Figure 6-3* are described briefly from bottom to top .

- Bottom Plate: Ceramic (filter) plate with silicon-rubber legs.
- Granite Plate: placed onto the ceramic plate on silicon-rubber legs.
- Micrometer Stage: screwed to the granite plate.
- Aluminium Plate: screwed to the granite plate, the distance is defined by the spacer-pipes. To avoid the possible sliding of the StandAlone on its support plate, several small holes are made to predefine the positions on the plate (as already in the steel plate - *Figure 6-2*).

Figure 6-3 Final holder with improved damping.

The bottom plate was added to improve the damping capacity and also to lift the instrument a little bit. This allows again a quite convenient way to follow the approach of the probe towards the sample surface with the monocular.

When investigating the damping properties of the holder, it seemed that there was still a $30\mu m$ wave going through the image (*Figure 6-4*). It could be shown that this particular wave did not depend on the holder or the StandAlone AFM, because also when using the normal AFM, the wave was visible, although not so strong due to the far smaller maximum scan size (*Figure 6-5*).
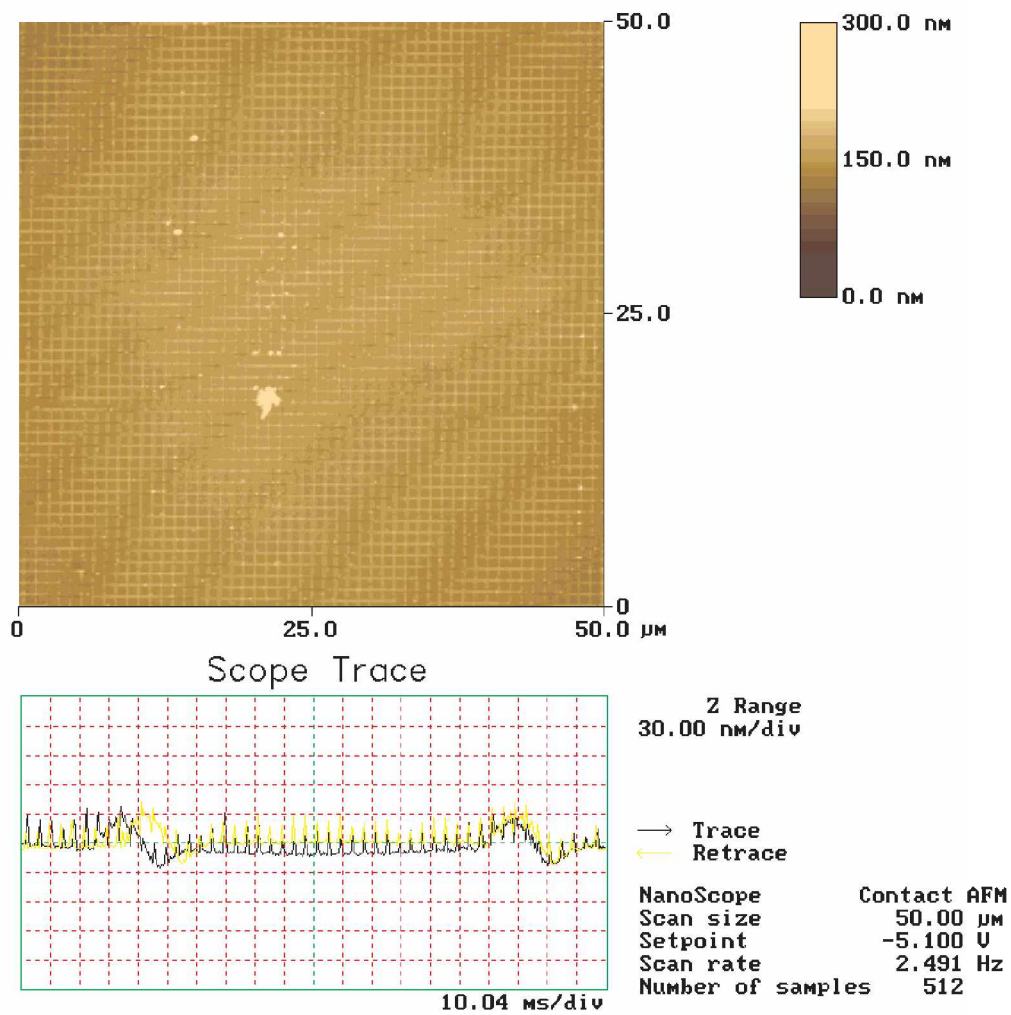
Figure 6-4 Calibration grid with an overlay of a ~30µm wave. In the scope trace below the wave is clearly visible.
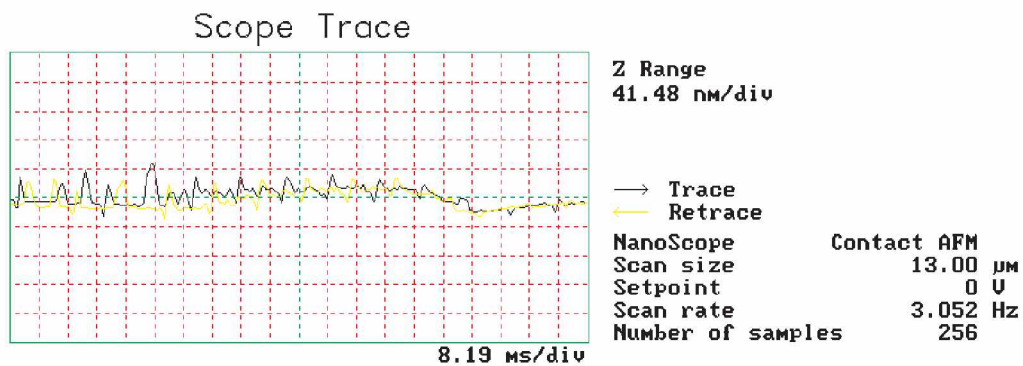
Figure 6-5 The same wave as in Figure 6-4 but imaged with the normal AFM. Note the changed Z-range and the different scan size.

However, it was not possible to find the reason for the vibration. Only some suggestions can be made, like a normal mode frequency of the building, caused by some machine or the construction work done nearby or an instrument's own normal mode frequency. Also a tip-sample interaction, like charging could be possible.

Luckily, the vibration disappeared by itself and recent checking showed that nothing has changed since then.

After getting rid of the 30μm wave, a 30nm wave appeared in the images. This was again something totally unexplainable, but it could be eliminated by putting the legs of the holder into a high-viscosity oil (*Figure 6-6*). This trick is used also for the normal AFM.
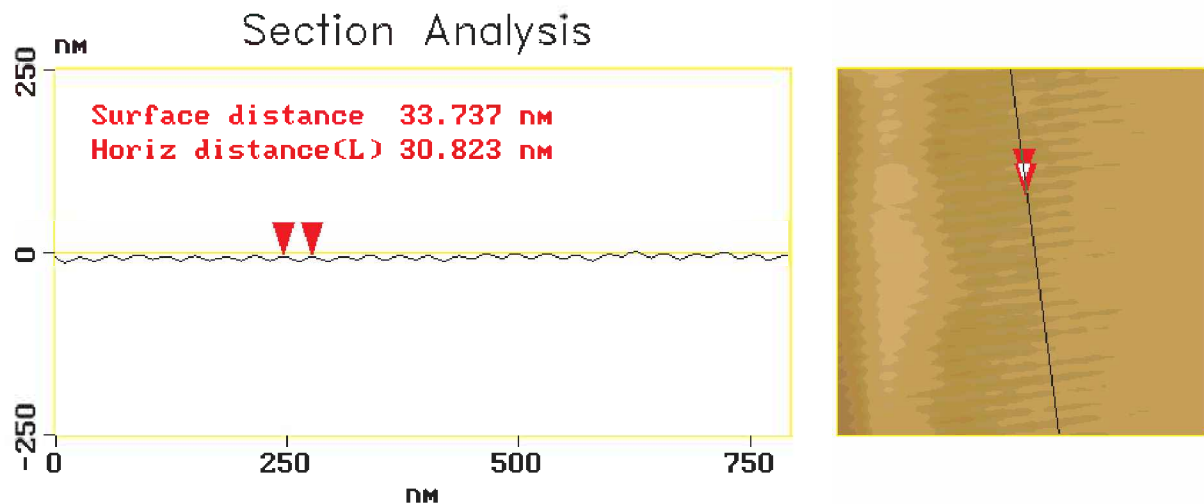
Figure 6-6 StandAlone Holder, without oil below the legs (wavelength ~30nm).

When working on the vibration and damping problems of an AFM, the probably fastest approach is trial and error, even though it would be interesting to calculate some damping properties. Depending on the environment, every operator will find himself in a different position to damp the system properly. For industry, Digital Instruments even sells a fully rubber damped table with a cover to get rid of air vibrations. In this respect, it is somewhat uncertain that one really can use the StandAlone in a normal "outdoor" environment.

Now that the StandAlone AFM and its holders were working, the next step was to design a SEM holder for the sample, allowing a combined use of the μ-Indenter and the EBSD system.

6.2 Special SEM Sampleholder

The combined use of the EBSD system and the μ-Indenter is not very straightforward and easy. For EBSD measurements, the sample stage has to be tilted to ~70 degrees or a pre-tilt holder has to be used to reduce this angle. In contrast to this, the μ-Indenter allows a stage tilt of just about 35 degrees for not hitting the arm of the indenter. Evidently, microhardness indentations have to be done at a zero degree tilt.

To be able to use both instruments without removing the μ-Indenter, a special sampleholder was designed (*Figure 6-7*). It is possible to mount flat tensile or fatigue samples and it allows a zero degree (with the fork) as well as a 35 degree tilt (by removing the fork). When using other geometries, only the clamping has to be changed according to the sample. The five holes on the rear side are for the stubs when silicon is used for calibration of the EBSD system [33].

Figure 6-7 SEM sampleholder with pre-tilt possibility.

With this holder it is now possible to mark areas of interest with the μ-Indenter and define afterwards the orientation of the grains within the marked area. After conducting experiments like tension or fatigue tests, it is still possible to find the marked area and to investigate it by different techniques, like SEM, EBSD and also the StandAlone AFM.

6.3 Routine Evaluation

As long as the sample is only investigated in the SEM (before and after deformation), it is fairly easy to find the once marked area again, but when using the StandAlone AFM, it becomes more difficult. An important question is whether it is possible to find the markings within a reasonable time or not, which is especially important when a metal is used, showing often strong nonuniform oxidation.

Lacking a top view system, like for the normal AFM, one has to guess the position of the indentations by looking from the side. Additional light may assist, but it still needs a good eye to estimate the approximate location.

Therefore a minimum of 500×500μm area should be covered with indentations, were the maximum distance between indentations is 100μm. This means that on the scan area (130×130μm) of the StandAlone AFM there should be at least one indentation, if the marked area was not missed completely. Still it is not possible to recognize the center of the marked area, and therefore on the side away from the middle, the indentations should have a "satellite" indentation which defines the direction towards the center. When engaging the StandAlone AFM (inside the marked area), these "satellite" indentations will guide the user as to how to move the micrometer stage of the StandAlone holder. Unfortunately, the AFM has to be raised for moving the sample and afterwards engaged again.
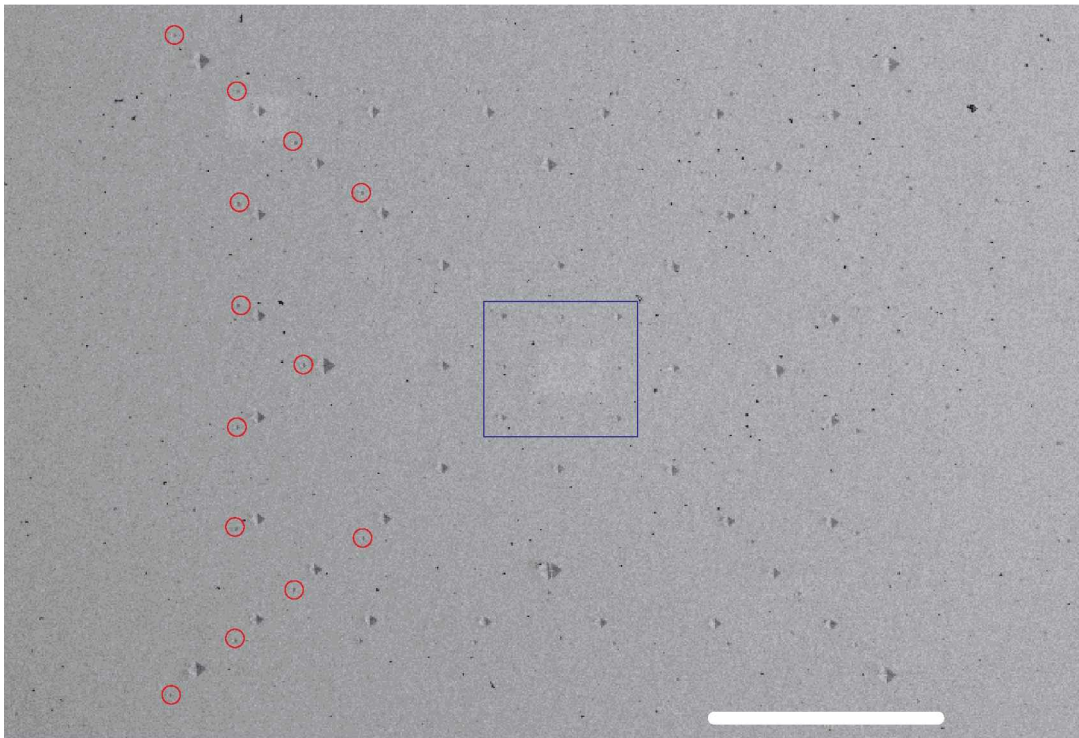
Figure 6-8 SEM micrograph of indentations made with the motor-driven stage and the μ-Indenter. Some of the "satellite" indentations are in the red circles. The blue rectangle refers to the size of the image in Figure 6-10 taken with the StandAlone AFM. The white bar indicates 200μm.

To make an indentation grid like that in *Figure 6-8* it takes several hours, and therefore it is probably better to define a grid which guides to more than one interesting area. It could be, for example, done like that in *Figure 6-9*, where rectangle 5 represents the one in *Figure 6-8* and all the others are positioned on sites where the amount of needed indentations increases much. Dividing of them could be done by several big indentations according to the number of the area. As soon as one of the locations is found, the known arrangement of them allows one to move easily to the next location.
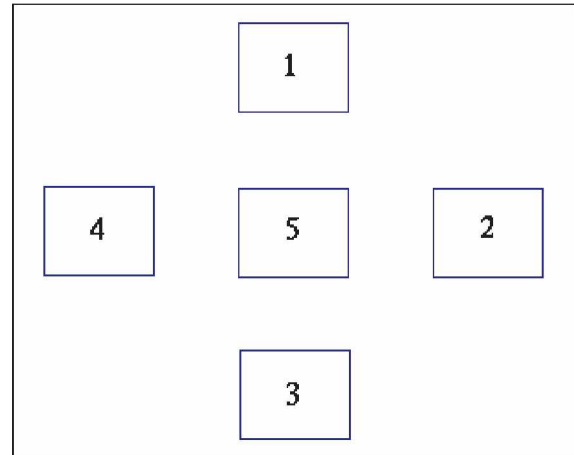
Figure 6-9 Suggested arrangement of indentation areas.

Whenever such a grid is used for positioning the StandAlone AFM, one should exactly know how long useful information can be gathered under ambient conditions. The surface properties (scale formation) should be investigated before starting to make time consuming indentation grids. *Figure 6-10* is an example of copper oxidation, taking place on one of the grid specimens. Although it takes hours to come to this stage, oxidation becomes a problem already after 30 minutes.
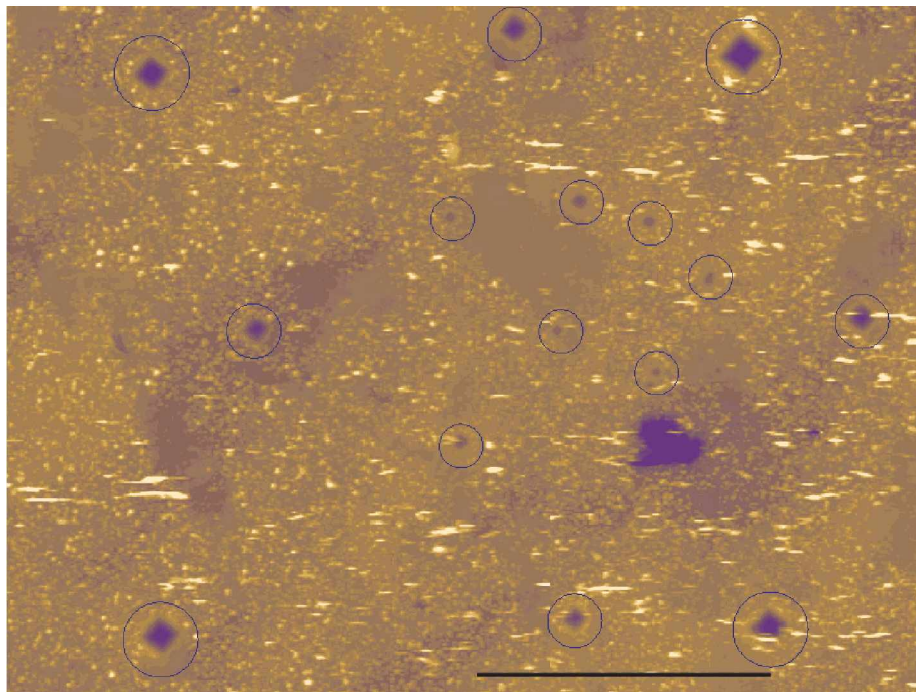
Figure 6-10 StandAlone AFM image (combined from four images) with severe surface oxidation on copper. Indentations are in blue colored circles. The black bar indicates 50μm.

7 Software Development

7.1 Programming Language, Operating System and Software Considerations

The program developed in this work is written in the C programming language, which is probably the best language concerning the execution speed and resource management. Many mathematical libraries are available for free and also the previously discussed tip estimation library can be downloaded as a public domain source code.

The operating system one chooses to use is mainly a matter of taste and the possible drawbacks or advantages of each of them will not be discussed in the present work. Thus, Linux was chosen mainly because of three reasons: it is free of charge, it has one of the best C-compilers and it is a multitasking operating system. It should be emphasized that, as long as no operating system specific routines are used, the program is portable to nearly every operation system. Usually file handling and directory listing is a specific operation and has therefore to be changed if the operating system is changed.

The current software version is still completely textbased and does not have any graphical output. A definite improvement would be a graphical user interface (GUI) and output in different graphical formats (TIFF, PS, etc.). It would allow operations which cannot yet, though implemented to the algorithm, be used. The whole architecture of the software is well structured to allow an easy implementation of further improvements of the tip estimation algorithm, where already suggestions have been made [34]. The direct support of Digital Instruments' (DI) files is complete in a sense that the program reads all the relevant information from single and double AFM files. This is not the usual case in other programs (SPIP, [35]), where only the image information is read. Reading of the whole information is needed if the program is to be extended with new recalculation routines, based on the nonlinearity measurements on the scanner.

7.2 DI File Format

The image information stored is specific to every manufacturer. The file format used by Digital Instruments is a mixed ASCII/Binary format which does not have any general specification, and several format changes have already been made since 1987, when DI introduced their first AFM. New software releases are always backward compatible only and therefore changes to an old file in the new software can only be stored in the new format.

Generally the files have a header which contains all the information relevant to the image. The header ends with a "Ctrl-Z" character and after this some "padding" data is stored till the file is 8 Kbytes long (4 KBytes for old file-formats). After this the image data is stored in a 2-byte binary (LSB, "little endian") form. For the two bytes (2^{16}) per pixel the 2's complement is used and therefore the values are ranging from -2^{15} (-32767) to 2^{15} (32768). Depending on the number of image points, the files will vary in size from about 41K to 532K per image.

Digital Instruments offers the possibility to store up to three images at once, either from trace and retrace or even different imaging modes. The arrangement of the relevant parameters and the data for the corresponding images can be seen in *Figure 7-1*.
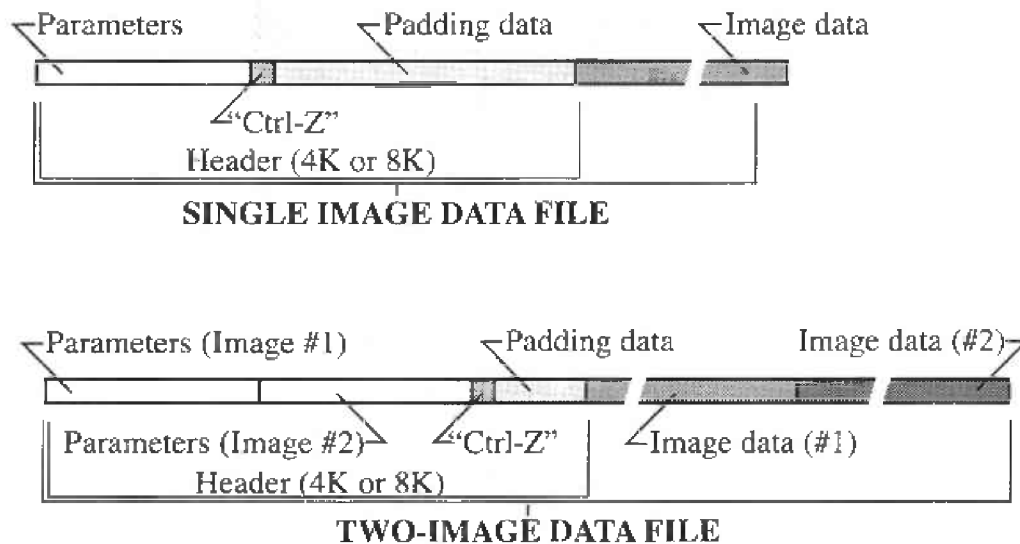

Figure 7-1   General file format of DI files for single and double images [10].

To interpret the image data correctly, the header is crucial and has to be carefully analyzed, but before going into the details of the header, the offered export filters in the NanoScope software are first discussed.

7.2.1 DI Data Export

Digital Instruments offers two different export filters, a TIFF Export and an ASCII Export.

The TIFF Export is used to convert the files for third-party image processing software (e.g., Photoshop, CorelDraw, The Gimp, etc.). Rendering of 3-D images is best done in the NanoScope software itself while filtering and adjustments to get a good printout (contrast, brightness, color, etc.) are better done in other image processing software. Especially the printing in the described system is a critical point, because there are only a few printers supported and printing is very slow. Therefore it is always better to print out from another program.

The ASCII Export is needed if the user wishes to export data to third-party spreadsheet programs (e.g., Excel, Lotus 1-2-3, Mathematica, Mathlab, etc.), because none of them is able to read the DI files just as they are. Matrix based mathematical programs, such as Mathematica and Mathlab, can be programmed to read the files directly. The ASCII Export window (*Figure 7-2*) offers several possibilities to choose from. *Format* defines the delimiter character used between the data, **space** for *ASCII* and **comma** for *Spreadsheet*. One can further choose if the header should be saved or skipped, where the latter is only sensible when the important parameters from the header are already known. The *Number of Columns* should be set according to the used samples per line and defines the location where a carriage return and line feed character is added to the exported file. The *Destination DOS Directory* and *Destination File* are straightforward and do not need further explanation.
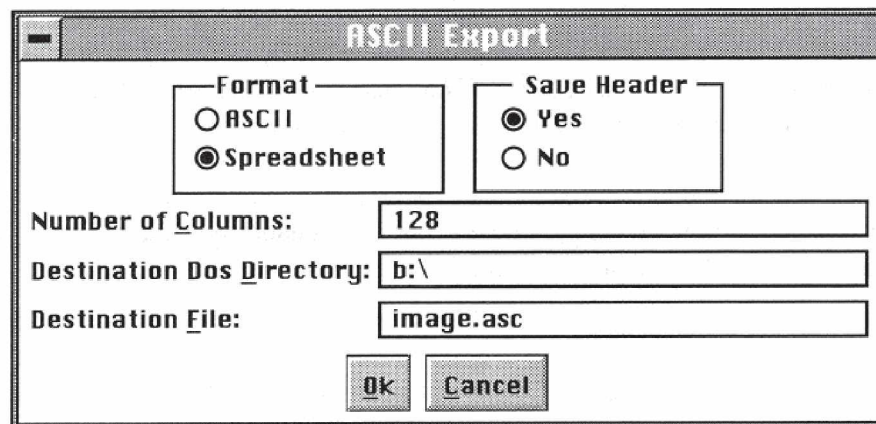
*Figure 7-2
NanoScope
ASCII Export
Window [10].*

One should note that several programs have limitations in the number of columns or the largest possible matrix size. Excel has a limitation of 256 columns and is therefore not suitable for calculations on 512×512 images, but even on 256×256 images the limitation becomes quite disturbing. The Mathlab Student Edition has an even more serious limitation with a maximum matrix size of 128×128. If one of these programs has to be used, one can downscale the images in the NanoScope software to make them suitable for calculations, but information loss should be taken into consideration.

7.2.2 DI File Header

As mentioned above, a deep understanding of the header content is needed to interpret the stored data. Digital Instruments mentions that the total amount of different parameters (from several instruments) is about 2000, but luckily only 120-150 are relevant to the NanoScope AFM file. In *Figure 7-3* the most relevant parameters are given, while the whole set of parameters from this file can be found in Appendix B.

Figure 7-3 Some of the parameters in one particular file. The space between lines indicates that there was something left out.

The parameters are arranged in several "lists", which can have different names according to the used type of instrument.

*File list: \Data length gives the header length (in bytes) in the file (parameters + padding data).

*Afm list: The \Operating mode can be one of the various offered modes (Scope, Force, Image, etc.) and the \Scan size, \Samps/line, \Lines and \Scan rate specify the used settings during image capturing. \Sample period specifies the time (tens of microseconds) between data collection operations in the image mode. \Z atten. describes the used resolution of the Z-scale (can be defined by the operator). The lowest value 16384 limits the Z voltage to a maximum of 55 Volts, while the maximum value 65536 corresponds to the full Z-range of 440V. The \AFM mode shows the used mode during capturing (contact, tapping, etc.).


```
File    Edit    Options                    Help

\*File list

\Data length: 8192
\*Afm list
\Operating mode: Image

\Scan size: 4000 nm

\Samps/line: 512
\Lines: 512
\Scan rate: 4.06901
\Sample period: 160

\Z atten.: 65536

\AFM mode: Contact

\*Microscope list

\Piezo size: E
\Z sensitivity: 8.12985
\Z polarity: Forward
\Retracted offset der: 2.7
\Extended offset der: 7.6

\*Controller list

\Z max: 220

\*AFM image list
\Data offset: 532480
\Data length: 524288

\Plane fit: 4719 3533 -5045.75 2
\Frame direction: Down
\Samps/line: 512
\Number of lines: 512

\Scan size: 4000 nm
\Line direction: Trace
\Image data: Height
\Z magnify image: 1.52474
\Z scale: 336.23 nm (6160)
\Realtime Planefit: Line
\Offline planefit: Full
\Highpass: 0
\Lowpass: 0

                          INS  Line: 38 Col: 1
```

***Microscope list:** \Piezo size specifies the kind of piezo used. \Z sensitivity is important for the Z-scaling of the image. Every used voltage value on the Z-piezoelement will be multiplied by this factor to get a metric scale. \Z polarity specifies whether a positive or negative voltage causes an extension. \Retracted offset der. and \Extended offset der. specify a factor used to compensate the change in sensitivity with increasing retraction and extension, respectively.

***Controller list:** \Z max specifies the maximum allowed voltage for the Z-piezoelement. The whole list contains several hardware specific limitation values, like maximum voltages for the scanner elements or maximum auxiliary output voltages.

***AFM image list:** \Data offset specifies the starting location (in bytes) of image data in the file and \Data length specifies their length in bytes. \Plane fit - The first three numbers specify the numbers used for recalculation of the data, the first number is the X coefficient, the second number the Y coefficient and the third number is the Z offset. The fourth number specifies what kind of planefit was used. \Frame direction is important because the data is stored in a different way when the image was captured during an Up or a Down scan. \Samps/line, \Number of lines and \Scan size are equal to the values in the *Afm list. The \Line direction specifies what data was captured, either when scanning from left to right or in the opposite direction. The same applies here as for the \Frame direction, *i.e.*, data is stored differently for the two possibilities. \Image Data specifies the interpretation of the captured data (height, phase, deflection, etc.). \Z magnify image is a NanoScope software related parameter and is used for rendering the image. \Z scale is used to calculate the correct Z-values of the data in the file. The first value is the full Z-range of the data and the value in braces reflects the same but as a fraction of 65536. \Realtime Planefit specifies the planefit operation used during capturing of the data and does not change the data to be stored. \Offline planefit specifies the planefit operation done with the stored data. The used constants can be checked from the \Plane fit parameter. \Highpass and \Lowpass are outlining the strength of the corresponding filters used for the captured data.

All of these given parameters are important with respect to the data, because they are either important when interpreting the data, or they are needed to find out the changes made in the original measured data.

7.2.3 DI Image

The two byte numbers stored can be interpreted in either volts (applied to the scanner) or metric units for an AFM. Working with Volt is generally easier and should therefore be preferred. Calculation of height values can be done with eqn. (16), while to get voltage values only the Z sensitivity factor has to be left out of the formula.

$$Height \ (nm) = \frac{Data \ point}{65536} \left[Z \ scale * \frac{Z \ atten}{65536} * \frac{2 * Z \ max}{65536} * Z \ sensitivity \right] \qquad (16)$$

7.3 Tip Estimation Library

The Public Domain Library from J.S. Villarrubia for the tip estimation will not be explained in further details because this has already been presented in [22]. Only the additional routines will be discussed here.

7.3.1 Memory Allocation for Floating Point Numbers

When dealing with the tip estimation, desired speed and accuracy affect the selection between the use of floating point numbers or integers. For speed, the library can only handle integer numbers. Nevertheless, the actual stored data is based on rather small voltage values and when these are calculated one will get floating point numbers. After this they can be scaled up to values where the loss of decimal numbers, when changing to integer values, is negligible. However, when voltage values have to be handled, additional functions are needed to store the data.

In the following the first routine allocates memory for floating point numbers and the second routine frees the given space again.

```
double **alloc_db_matrix(ysiz,xsiz)
long ysiz,xsiz;
/*Allocates an floating point matrix of dimension [ysiz][xsiz] using an
   array of pointers to rows. ysiz is the number of rows. xsiz is the
   number of columns.*/
{
 double **mptr;      /* points to allocated matrix */
 long   i;           /* counter */

 /* Allocate pointers to rows */
 mptr = (double **)malloc((double)ysiz*sizeof(double*));
 if (mptr == NULL) {
   printf("Error: Allocation of mptr failed in alloc_db_matrix\n");
   exit(1);
 }
```

```
/* Allocate rows */
for (i=0;i<ysiz;i++) {
  mptr[i]=(double *)malloc((double)xsiz*sizeof(double));
  if (mptr[i] == NULL) {
    printf("Error: Allocation of mptr[%ld] failed in alloc_db_matrix\n",i);
    exit(1);
  }
} /* Done. Return result. */
return mptr;
}


void free_db_matrix(mptr,ysiz)
  double **mptr;
  long  ysiz;
/* Frees memory allocated with alloc_db_matrix */
{
  long i;
  for (i=0;i<ysiz;i++) free(mptr[i]);
  free(mptr);
}
```

7.3.2 Parabolic and Pyramidal Tips

The usual way is to start the tip estimation with an array of zeros with a defined tip apex inside the array. When the tip shape is already known, which is usually the case, the tip estimation will speed up when the start estimate is close to the real tip. For a parabolic starting tip the algorithm is straightforward, but for a pyramidal starting tip a new algorithm had to be developed. These algorithms are too long to be shown in full in the present work and will therefore be only briefly explained.
The parabolic tip calculation uses eqn. (17).

$$z = \frac{x^2 + y^2}{a^2} \tag{17}$$

where x and y are the discrete image points of the tip and a is a constant which is dependent on the Z-Scale used for the DI-file.

A pyramid has a four fold symmetry and this is made use of in the algorithm. The only limitation is that it is not possible to use a mixed even-odd matrix, because the apex has to be on a discrete image point. As soon as the apex has only two nearest neighbor points, the algorithm fails. The user has to define the tip matrix size and the angle of the pyramid. With this and the Z-Scale from the file, several height values h(i) are calculated. Starting from the apex, the matrix points are addressed with the necessary values h(i). As soon as one square is ready, the next bigger and higher one is made. All the calculated points are then reversed, so that the apex is the highest point of the pyramid (*Figure 7-4*).
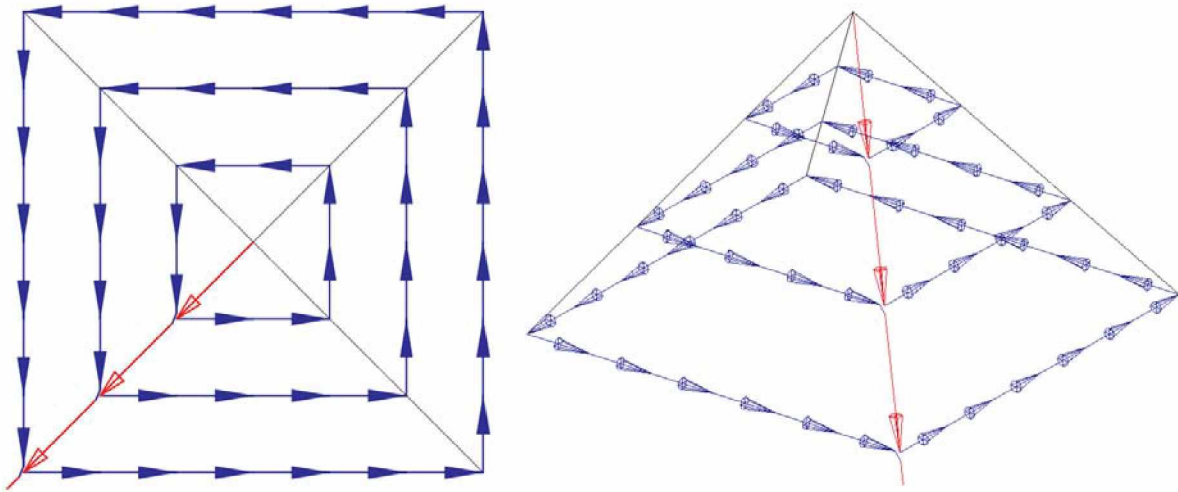
Figure 7-4 Schematic illustration how the pyramid calculation algorithm works.

7.4 DI File Filter

The filter library for DI files contains several functions, of which two are discussed here. The main function is the reading of the parameters for further calculations or just to view the parameters, because the NanoScope parameter view does not show every value stored. Reading out the values is a long list of "if"/"else if" calls and only one of them is given here as an example:

```
else if(!strncmp(&Buff[1],"Scan size",8) &&
        !strcmp(headline[i],"Afm list\r\n")) {
    while( (Buff[Idx]!=':') && (Idx<BUFFERSIZE) ) {
      Idx++;
    }
    Idx+=2;
    para_value_double[0]=atof(&Buff[Idx]);
    }
```
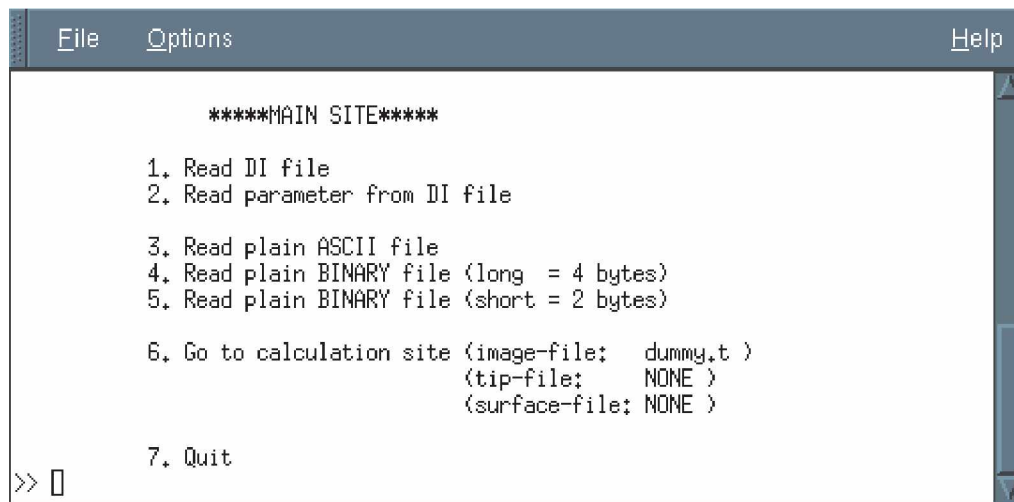
Here the read parameter string (stored in "Buff[1]") is compared to "Scan size" and the stored string in "headline[i]" is compared to "Afm list". When both are equal the "Idx" variable is increased until the colon in the parameter string is reached. "Idx" is then increased by 2, and the floating point value is stored in "para_value_double[0]".

A small but important function called "two_images" is used to spot double-image files. It is so far not possible to have more than one image loaded to the program at the same time, and therefore the user will be asked which one of the two images is to be loaded. If the second set of data is a surface related to the image, one will have to load them one by one from the file.

7.5 Main Program

The appearance of the program is divided into two main "windows", where the "Main Site" (*Figure 7-5*) is for loading files in different formats and to get the parameter view for DI files, while the "Calculation Site" (*Figure 7-6*) is for calculations on the loaded data.

```
 File    Options                                             Help


            *****MAIN SITE*****

       1. Read DI file
       2. Read parameter from DI file

       3. Read plain ASCII file
       4. Read plain BINARY file (long  = 4 bytes)
       5. Read plain BINARY file (short = 2 bytes)

       6. Go to calculation site (image-file:   dummy.t )
                                 (tip-file:     NONE )
                                 (surface-file: NONE )

       7. Quit
>> []
```

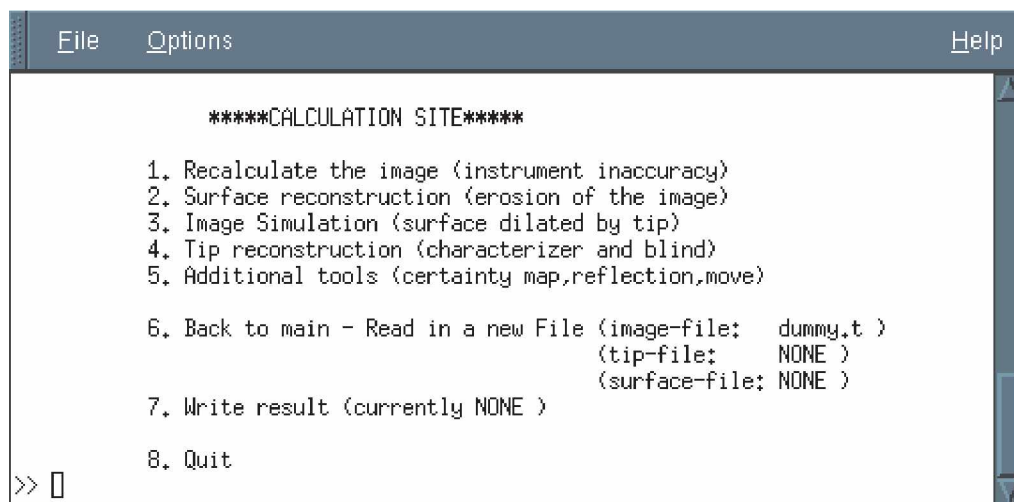*Figure 7-5 Main Site of the program for loading different file formats and to switch to the parameter view.*

```
 File    Options                                             Help


            *****CALCULATION SITE*****

       1. Recalculate the image (instrument inaccuracy)
       2. Surface reconstruction (erosion of the image)
       3. Image Simulation (surface dilated by tip)
       4. Tip reconstruction (characterizer and blind)
       5. Additional tools (certainty map,reflection,move)

       6. Back to main - Read in a new File (image-file:   dummy.t )
                                            (tip-file:     NONE )
                                            (surface-file: NONE )
       7. Write result (currently NONE )

       8. Quit
>> []
```

Figure 7-6 Calculation Site of the program for starting different routines with the loaded data.

It should be noted that point 1. in the Calculation Site, "Recalculation of the image", contains so far only a function which calculates the radius of curvature resulting from cross-coupling. An assumption was made that the scanner will follow a sphere, which is true if there is no coupling between the axes. However, if coupling is present it might be better to approximate the cross-coupling with an elliptical paraboloid.

8 Investigation of Scanner Nonlinearity

8.1 General Notes

The investigations on scanner nonlinearities were focused mainly on the Z-scale of the scanner, because there the calibration is rather poor. An assumption was made that Digital Instruments' calibration routine works fine for the X-Y plane, because otherwise the whole instrument calibration routine would have to be developed from the scratch, which would be a huge task.
The calibration was carried out first for the StandAlone AFM and the precalibration was checked for the normal AFM.

The focus was to work on "easy to run" and wide range spanning height calibration possibilities. As soon as special tools or equipment not available at CEM would have been needed, the idea had to be rejected. Such rejected ideas are briefly discussed in the following.
Latex balls can be used for height calibration, as well as for tip radius definition. The latex surface can be prepared in monolayers by depositing the latex suspension onto a (inclined) mica or silicon surface [36][37]. Available sizes range from 60 to 1000 nm and therefore a wide calibration range can be covered.
Gold particles dried on a mica surface [38] with a size from 11 to 40nm reveal the quality of the tip and uncertainties of the used Z-calibration.
The already discussed way of *tip imaging* has also been shown to be suitable for Z-calibration [39]. Here, an InP surface after processing by means of reactive ion etching was used to obtain tip images. The known shape of Si_3N_4 tips was then used for the calibration parameter evaluation.
In [40] a method using a *"calibration piezo"*, set under the AFM tip in place of the sample was presented. This is applicable for a wide Z-range but needs a reliable method to measure the "calibration piezo" in advance.
An *interferometry* based height calibration of the AFM was presented in [41], where only a mirror is needed additionally. The fixed wavelength of the laser is used to span a calibration range from several tens of nm up to several μm.
In the nanometer and subnanometer range several methods have been discussed, like *mica etch pits* [42], *patterned SiO_2* surfaces formed on a silicon wafer [43] or a stepped perovskite crystal of *$SrTiO_3$* [44].

8.2 Intrinsic Nonlinearity Measurements on Indentations

Indentations discussed earlier in the investigation routine can be used for height calibration due to the defined shape of the indenter. They are easy to make and further the whole height scale of the AFM can be covered with one indentation grid.

8.2.1 Material Selection

The shape of indentations depends heavily on the material, and therefore a material that preserves the inverted shape of the indenter should be chosen. The characterization of the surface displacement of the material during indentation can be done with a so called surface displacement factor β (after K. Zeng *et.al.* [45]), which is defined by A_{ideal}/A_{true}, where A_{ideal} is the "ideal" projected contact area and A_{true} the "true" projected contact area during loading, respectively. In *Figure 8-1* the surface displacement for $\beta>1$, $\beta=1$ and $\beta<1$ can be seen.
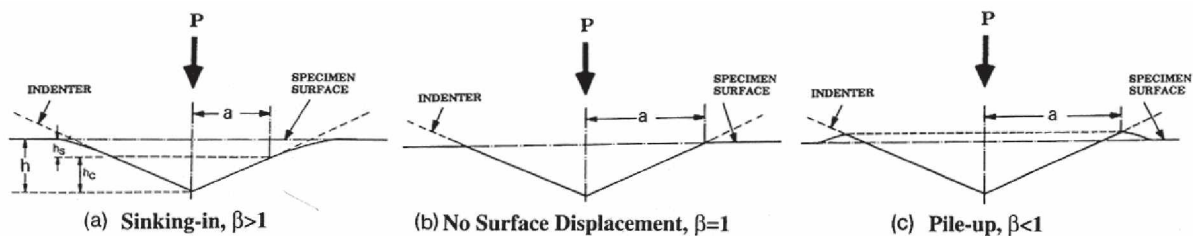

Figure 8-1 Schematic drawing of the surface displacement, with (a) sinking-in, (b) no surface displacement and (c) pile-up [45].

Especially for brittle materials like ceramics, β is usually larger than one, but also when hardening takes place during indenting, this is valid. When considering indentations for AFM height calibration purposes, β should never be larger than one. Finally, to get $\beta \leq 1$, the material has to be ductile and should not work harden much during the indentation. The elastic recovery of materials can be easily studied with load-displacement curves (*Figure 8-2*). When increasing the load to maximum (P_{max}), a maximum displacement (δ_{max}) can be measured. Releasing the load will end in a residual displacement (δ_{res}), where the total work (W_t) of indentation can be divided into elastic (W_e) and plastic (W_p) works. The contact stiffness (S) can be used to calculate the hardness, Young modulus and the contact area. For copper (*Figure 8-2 b*), W_e is very small and $\beta \leq 1$ and therefore it can be used for the discussed purpose.
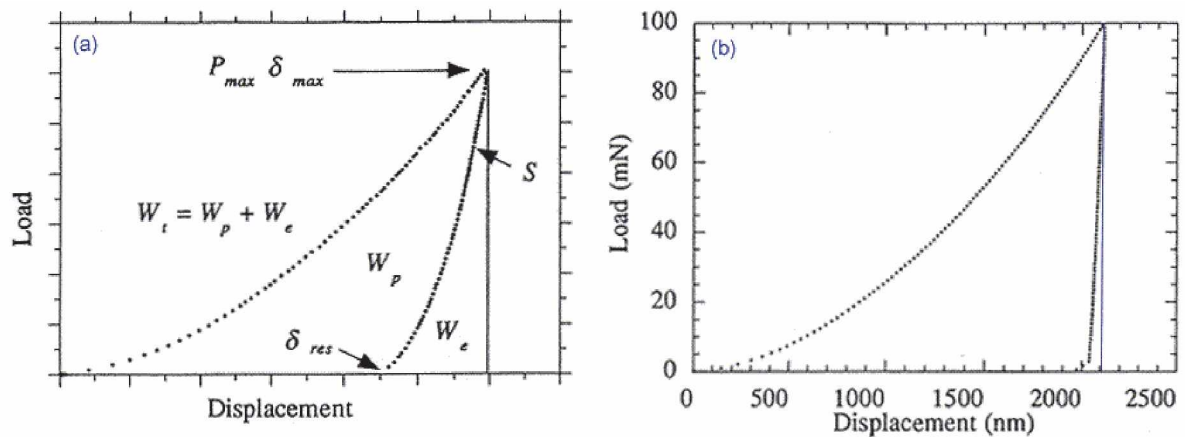
Figure 8-2 (a) General load-displacement curve and (b) from copper [46].

8.2.2 Measurements on Indentations

Measurements on indentations should be as accurate as possible to make them suitable for calibration. Not only the already mentioned factors (elastic recovery and pile-up's) but also other influences determine the suitability.

When using imaged indentations for calibration, one uses the known shape of the indenter to calculate the real depth or to compare the angles. Unfortunately, the real depth and the angles are influenced by the uncertainties of the indenter and the AFM tip, the elastic recovery, pile-up's and the asymmetry of the indentation. The properties of the used indenter tip were:

<div align="center">

Tip radius: 250nm

Face angle: 136° ±15'

</div>

When operating with depths one will have to deal with the indenter tip radius, pile-ups, asymmetry of the indentation and elastic recovery. The indenter tip radius is already approximately known and can be more accurately measured with
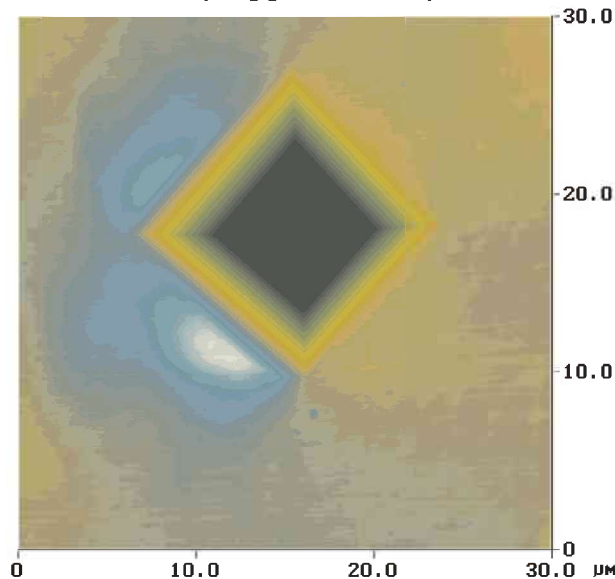
the AFM. Pile-up's are more difficult to treat because they do not regularly happen with the used equipment and their shape varies. A typical example is given in *Figure 8-3*.

Figure 8-3 Pile-up on one side of an indentation.

These asymmetrical pile-ups indicate that the force applied during indentation is not evenly distributed. This could be caused by the construction of the indenter itself, because the indenter arm describes an arc during indentation. Several tests showed that the shape of the indentation is very sensitive to tilts of the microscope stage, but pile-up's could not be avoided.

Another factor which is due to this is the regularly appearing asymmetrical shape of the indentation. While the indentation in *Figure 8-3* is fairly symmetrical, the indentation in *Figure 8-4* has a strong asymmetry. This introduces a deviation from the ideal depth when only the medium diagonal length is considered for depth calculations.
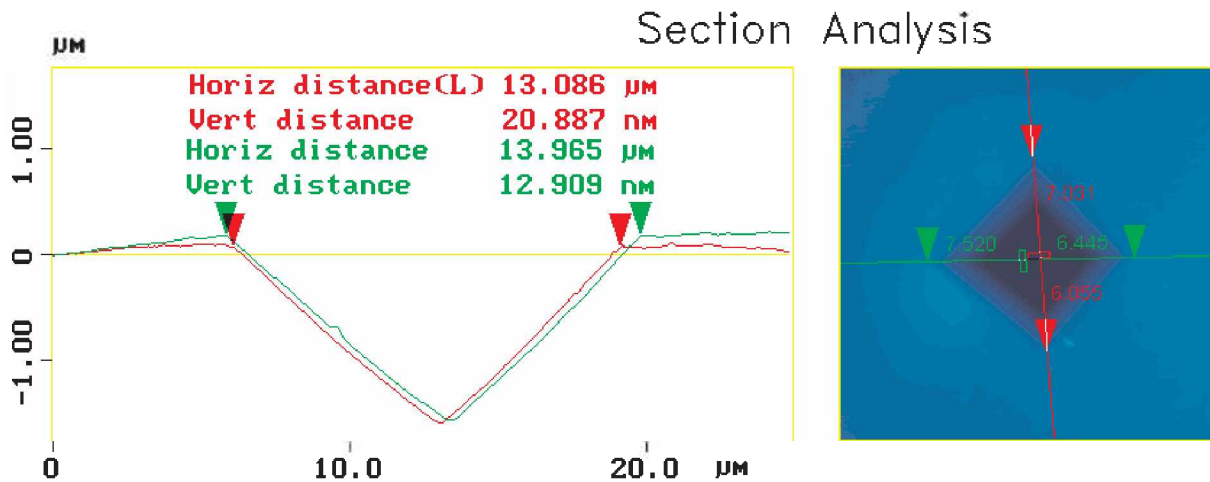
Figure 8-4 Asymmetry of an indentation. In the left picture the total lengths of the diagonals are given and in the right image they are divided into their contributing parts.

The AFM tip itself may also raise problems in the correct interpretation of the depth. It did happen not just once that a new tip was already in a bad condition or it degraded during the various raising and engaging of the instrument to reach indentations of other sizes.

The elastic recovery of the material must be treated when using the indentation angles for calibration. One could imagine that, when reducing the force of the indenter, the elastic recovery could change the face angle, while the angle in the indented edges stays rather constant. This was examined by making section analysis on indentations (*Figure 8-5*). If there would be measurable bowing, the green arrow would not lay on a straight line between the black and the red arrow. The given lengths in *Figure 8-5* are always from the black arrow to the respective color.

Section Analysis

Surface distance   9.771 µM
Horiz distance(L) 9.609 µM
Vert distance   17.014 nM

Surface distance   4.863 µM
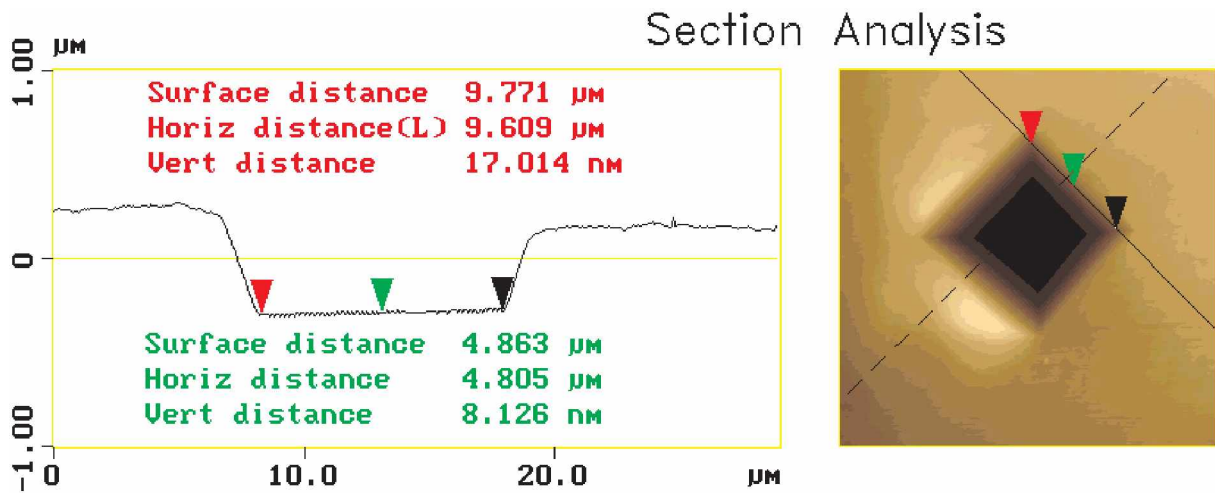Horiz distance   4.805 µM
Vert distance   8.126 nM

Figure 8-5 Measurement to determine the possible bowing on the face surface of the indentation.

The measurement suggests that there is no bowing, and therefore it can be assumed that it does not make any difference weather the face angle or the edge angle is used for calibration purposes. Still, it is more likely that the face angle changes [47] and therefore the edge angle should be preferred for measurements. To find out where the elastic recovery happens it would need further investigations, which are also of interest for the exact height calculation. For brittle materials with $\beta > 1$, the tip is last in contact with its apex and that means that the whole indentation bounces back. This change would be crucial for an accurate height calibration.

8.2.3 Experimental Observations

Various experiments were carried out to determine the change in sensitivity from indentations. At this time it was not possible to analyze the images by exporting the data and to do the calculations in a third-party program. Thus, one indentation was imaged with several settings of sensitivity. The obviously irregular pile-up's did not allow a correct measurement of the depth and therefore the edge angles were taken. The asymmetry of the indentations was handled by averaging the angles of the edges.

Unfortunately, the Z Center of the instrument was not precisely enough controlled and therefore the result of the measurements varied a lot. The small edge angle of 15.96° may also have contributed to the strong variation of the data. However, to get statistically valid data, further measurements would be needed and the outcome is still uncertain because of the vast amount of factors disturbing the "ideal shape".

8.3 Intrinsic Nonlinearity Measurements on Steps

The method presented in 8.2 may provide a direct and easy-to-repeat route for sensitivity determination, but the proof of its feasibility has to be delayed until further studies can be carried out. Another way to get an idea about how the sensitivity changes with increasing voltage is to use the DI recommended procedure for the Retracted and Extended Offset Derate determination. When the 180 nm step of the calibration grid is imaged at different Z Center Positions the actual height varies. The actual measured height is further dependent on the used scan rate and therefore it is of interest how big this influence is.

8.3.1 Experimental Setup

The experiments were carried out on the NanoScope AFM because it is much easier to adjust the Z Center with the motor than with the hand on the StandAlone AFM. After engagement of the tip the Z Center was adjusted to zero and the instrument was left on until stabilization. The Z Center was readjusted and further care was taken to keep it constant. Then four images for different frequencies where captured, and after this the Z Center was increased and decreased in the following way: 0 40 80 120 100 60 20 0 -40 -80 -120 -100 -60 -20 0 Volts. The new Z Center was adjusted within one scan after the last image captured.

It is important to note that measurements have to be carried out on the same pit, because every pit will have a slightly different depth. Due to the use of the motor the targeted pit will move (the tilt changes) and therefore the above presented series taken is the maximum possible. The recommended image processing routine for Z calibration was applied to every image (Flatten, X-Planefit, Bearing).

8.3.2 Results in Different Environments

Two sets of data were captured, one at 2.44 Hz and another at 4.88 Hz (Table 8-1). To keep eventual changes in the tip-sample interaction small, the four images with faster scan rate were stored directly after capturing the slow scan rate images. In *Figure 8-6,* the measured step height at different Z Center Positions are drawn. The used scan rate was 4.88 Hz and the colors outline the path how the series was taken. It is apparent that the standard deviation on a certain Z Center Position was very high. This was accounted to tip-sample interactions under ambient conditions. The scan rate influence can be seen from *Figure 8-7*, though the expected increase of the absolute step height with increasing scan rate can not be seen, which is probably due to the high standard deviation.

Table 8-1 Average step heights ± s.d. for two different scan rates.

| Scan rate:	2.44 Hz			Scan rate:	4.88 Hz	
Z Center [V]	Average Step Height [nm]	STDEV		Z Center [V]	Average Step Height [nm]	STDEV
0	189.00	0.36		1	185.83	2.42
38	185.19	1.69		44	179.42	0.98
80	174.99	2.90		80	174.20	0.35
119	174.38	3.37		120	175.27	0.46
99	174.15	0.75		99	171.78	0.44
59	177.94	0.45		59	177.03	1.26
19	184.37	3.73		19	187.17	0.34
0	182.46	1.71		1	185.04	0.96
-43	182.88	4.25		-42	183.01	2.34
-82	188.17	0.56		-79	181.24	0.64
-121	182.80	4.39		-119	179.28	0.70
-100	181.37	1.47		-96	182.08	1.22
-59	178.55	1.64		-63	180.95	0.32
-20	183.68	4.69		-22	185.14	3.12
1	180.28	2.73		0	189.21	3.40
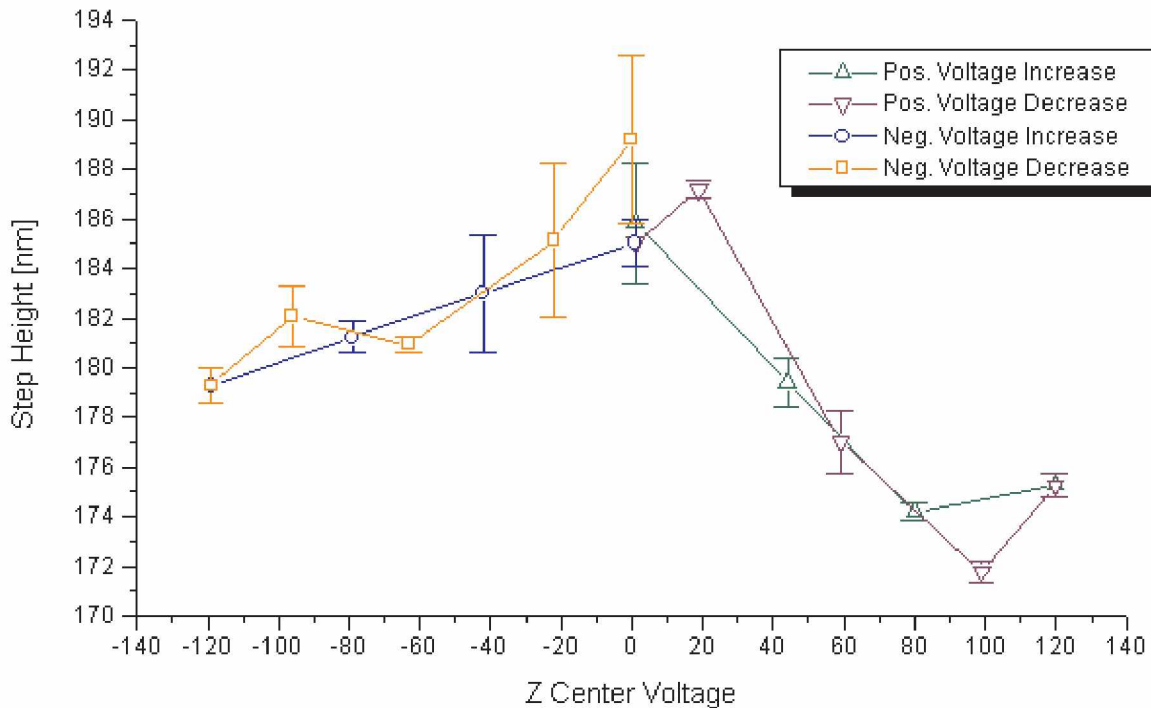
Figure 8-6 Measured step heights at different Z Center Voltages for a scan rate of 4.88Hz in ambient environment.
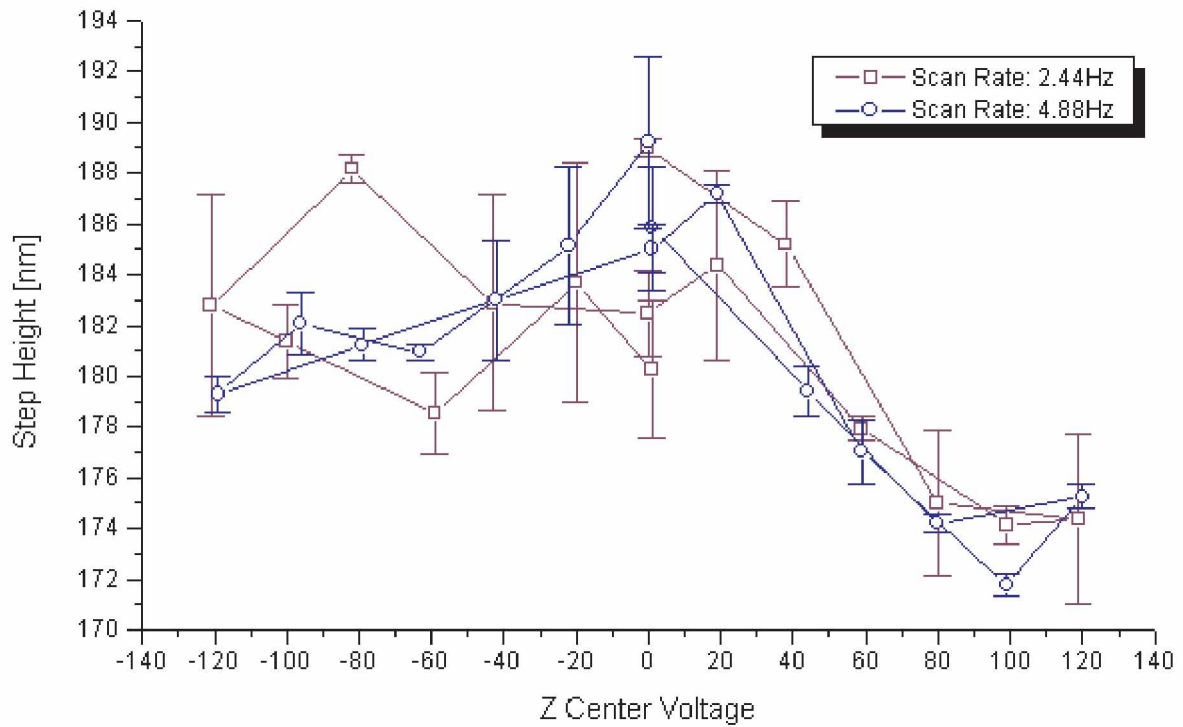
Figure 8-7 Step height measurements for two different scan rates at various Z Center Positions.

To reduce the significant influence of the ambient environment, the same experiment was carried out in liquid environment. The investigation of the scan rate influence was extended by two frequencies, 1.25 Hz and 7.5 Hz, but at 1.25 Hz only a few measurements were taken on the same pit, and therefore they are not listed in Table 8-2. One should note that the absolute values are not comparable to the previously presented ones, because the same pit of the grid was not imaged.

Table 8-2 Step height measurements in liquid environment at three different scan rates.

| Scan rate:	2.44 Hz		Scan rate:	4.88 Hz		Scan rate:	7.5 Hz	
Z Center [V]	Average Step Height [nm]	STDEV	Z Center [V]	Average Step Height [nm]	STDEV	Z Center [V]	Average Step Height [nm]	STDEV
			6	180.23	0.69	6	179.88	1.50
46	178.17	0.95	46	179.02	0.55	46	179.81	0.16
80	173.87	0.60	80	175.85	1.09	80	176.83	1.41
115	169.35	1.36	115	170.84	0.23	115	170.86	0.36
98	170.72	1.67	98	171.99	0.78	98	172.72	0.46
51	176.60	0.79	51	177.75	0.64	51	177.71	0.12
5	180.07	1.05	7	180.10	0.42	7	180.92	0.22
-40	180.69	0.26	-38	181.65	0.10	-38	182.02	0.10
-82	178.72	0.57	-80	179.82	0.29	-79	181.22	0.35
-121	177.09	0.72	-120	179.21	0.72	-120	179.65	0.68
-102	177.80	0.31	-102	179.79	0.86	-100	179.53	0.38
-80	178.15	0.66	-80	179.85	1.62	-81	178.38	0.52
-42	180.80	0.95	-38	181.91	0.31	-37	183.24	0.14
-2	180.97	0.72	0	182.61	0.40	-1	182.76	0.30
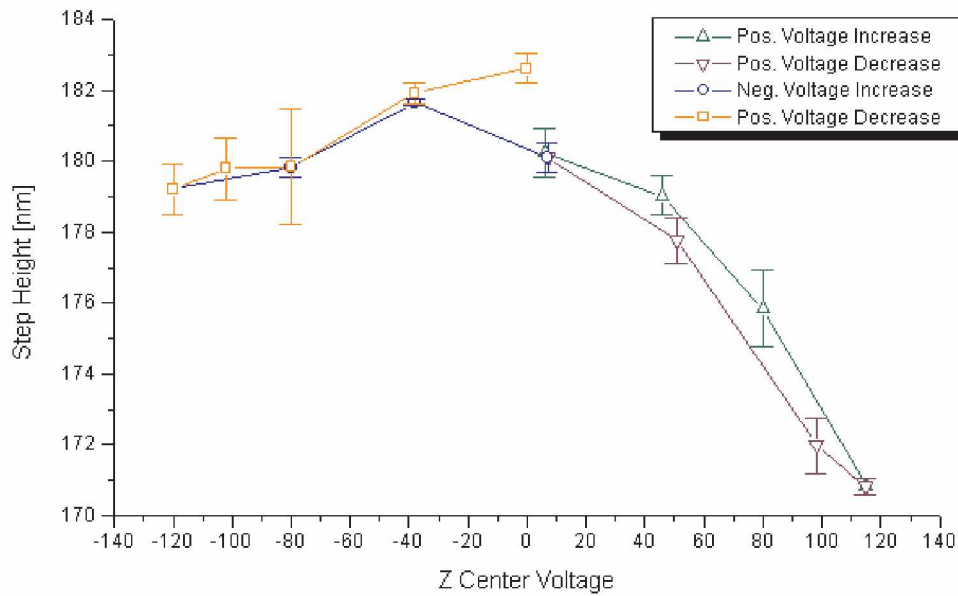
Figure 8-8 Measured step heights at different Z Center Voltages for a scan rate of 4.88Hz in liquid environment.
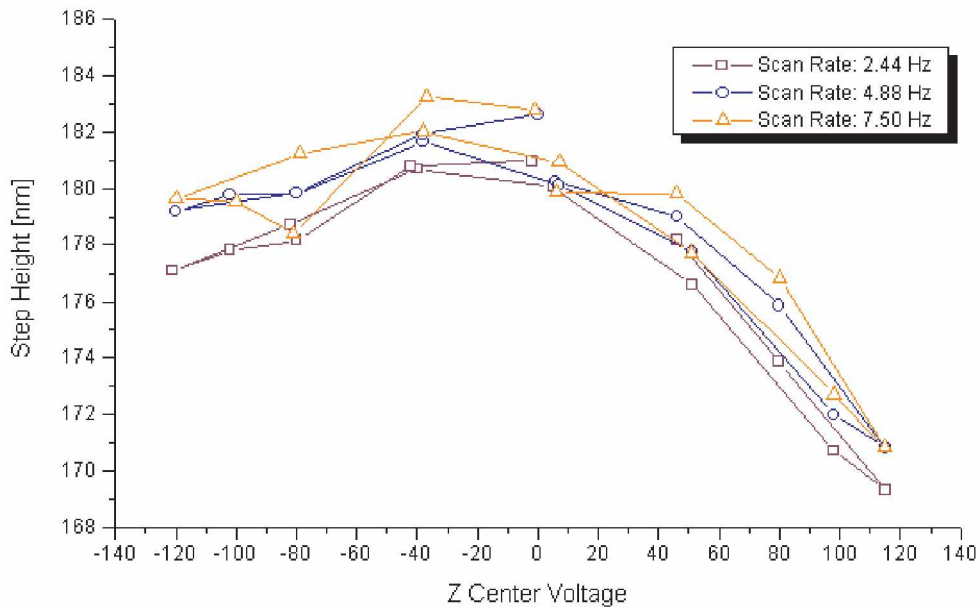

Figure 8-9 Step height measurements for two different scan rates at various Z Center Positions. The standard deviation bars are left out for clarity.

The measurement in liquid environment showed much less variation and also resulted in the expected increase of the step height when increasing the scan rate. The increase is lower from 4.88 to 7.5 Hz than from 2.44 to 4.88 Hz. For scan rates of 4.88 and 7.5 Hz the curve did not close when the negative Z Center Voltage was reduced, which can so far only be accounted to some creep effect. Considering that **DI** applies a linear correction for changes in the Z Center Voltage and does not take into account the influence of the frequency, the resulting uncertainty is clearly seen.

8.4 Creep Measurements on Steps

Creep of the piezoelectric material becomes more severe when high voltage steps are applied. This, however, does not happen on the X-Y plane during normal scanning but only when the Zoom command is used. Unfortunately this can be hardly quantitatively measured without removing the scanner from the instrument. On the other hand, the creep on the Z-axis is easier to trace and also of more relevance, because whenever there is a step on the surface a voltage step will be applied to the Z element of the scanner. There are two possible ways of introducing steps, either with the motor of the NanoScope AFM or by imaging a step on a patterned surface like the calibration grid.

8.4.1 Motor Induced Steps

The special type of stepper-motor used in the instruments from DI has to be considered when steps are introduced with it. It is a 400 step/rev motor, operating in half step mode. This means that two consecutive steps correspond to 1/400 of a revolution, but the half steps may not have the same size. Therefore there should be always an even number of steps to induce a defined height. Furthermore, the backlash from the drive should be removed before the initiation of steps.

The first set of experiments was carried out with an even multiple number of the smallest step size (19 nm/step). It turned out that the time to reach the desired Z Center Position was far too long whenever high Z Center Voltages had to be adjusted. Thus, a second set of experiments was carried out where the desired Z Center Position was induced with only two steps. To investigate small voltage steps as well as large changes, the following series was chosen (2 times) 19, 38, 57, 76, 95, 190, 285, 380, 475, 570, 665 and 760 nm/step. After inducing one step in the upper third of the image frame, the change in the Z Voltage was monitored (*e.g.*, during a Up scan). In *Figure 8-10* the change in metric units due to creep is outlined with the green arrows.

Section Analysis

Left panel:
Horiz distance(L) 46.875 nm
Vert distance   1.115 µm
Horiz distance   1.531 µm
Vert distance   101.24 nm

Right panel:
Horiz distance(L) 39.063 nm
Vert distance   1.247 µm
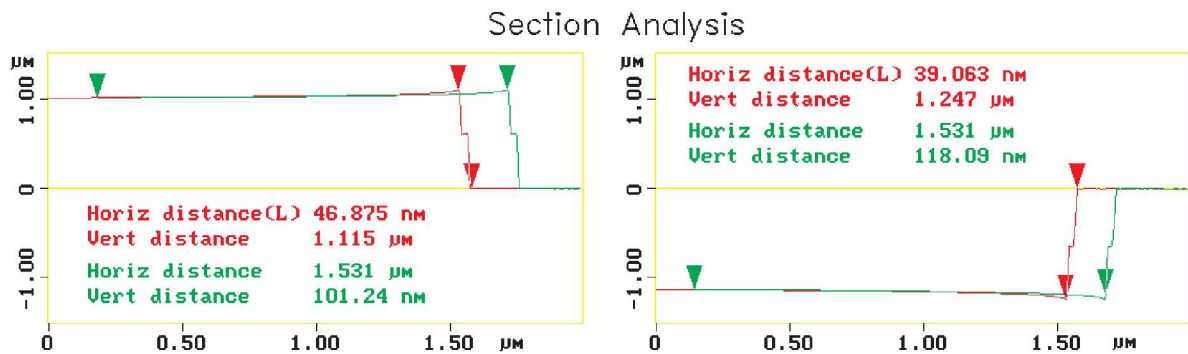Horiz distance   1.531 µm
Vert distance   118.09 nm

Figure 8-10 Positive Z Voltage change (left) and Negative Z Voltage change (right) due to creep. Horizontal offset of the curves for clarity.

When the image was fully captured, the Z Center Voltage was set back to zero with the motor during the next frame (Down scan) and after a stabilization frame (Up scan) the next step was induced (during the Down scan) towards the opposite direction of the previous Z Center Voltage.

The whole experiment was carried out with the focus on creep modelling and therefore the resulting curves will be discussed there in further detail.

8.4.2 Calibration Grid Steps

During the experiments with the stepper-motor it became already apparent that the fast response of creep after voltage steps can not be investigated with it. Only sharp surface features are suitable for this kind of an investigation. Although various heights of steps would be needed to study the behavior in detail, the calibration grid with 180 nm deep pits already gives a good idea of the creep rate.
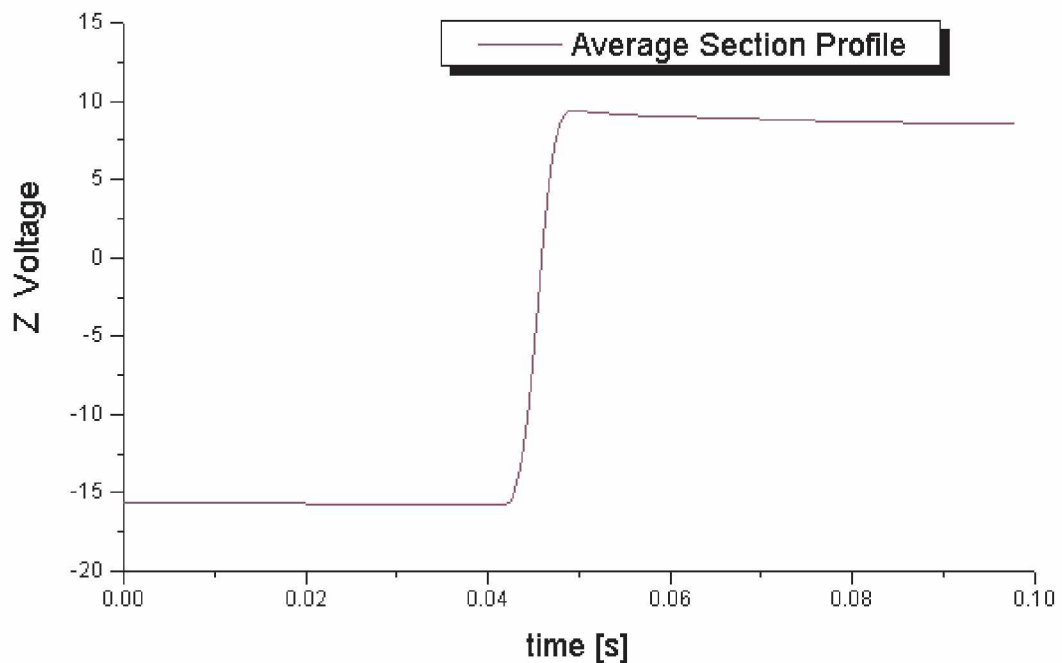

Figure 8-11 Trace profile of a traversed pit of the calibration grid from left to right. The curve is constructed out of 256 average image points. The scan rate was 4.06 Hz.

Table 8-3 The important values from the curve in Figure 8-11.

	elapsed time [s]	Z Voltage
Start of the Image	0 s	-15.59 V
Minimum before the Step	0.041 s	-15.79 V
Maximum after the Step	0.049 s	9.35 V
End of the Image	0.096 s	8.55 V

Assuming that the voltage difference from the "Start of the Image" to the "Minimum before the Step" is due to the sample tilt, the same magnitude could be expected also after the step. The measurement shows a reduction of the voltage by 0.6 Volts (9.35-8.55-0.2) from the "Maximum after the Step" till the "End of the Image" due to creep. The voltage change is therefore 2.4% of the full step height. This is considerably lower compared to the stepper-motor induced steps where the creep for the big steps was more than 10%. Note that the time scales are not equal.
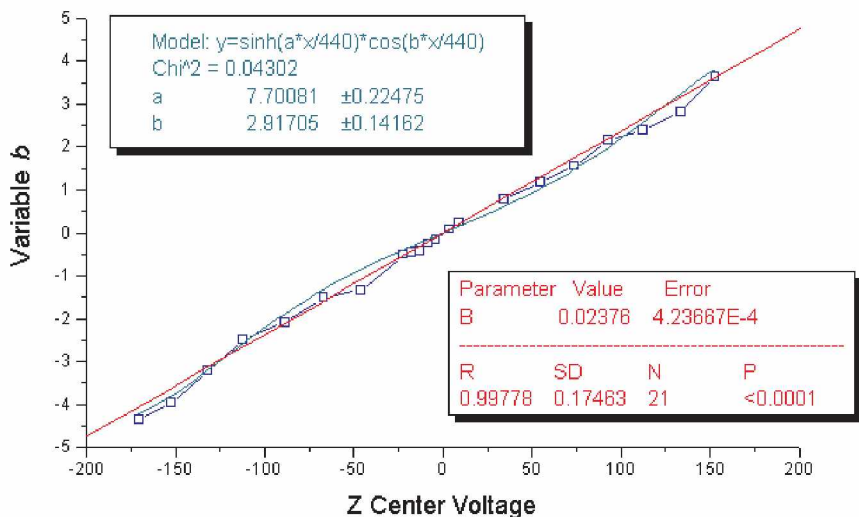
8.4.3 Creep Modelling

The two introduced methods for measuring the creep can be used to find a function which could describe this behavior. Several functions were tried to fit to the measured curves over the whole scale of steps. First different variations of exponential decays where used, but the fit was quite poor, especially when the whole scale was considered. After some trial and error, eqn. (18) showed to match the observed behavior fairly well.

$$y = y_0 - b * \ln(t+c) \tag{18}$$

where y_0 is the height of the step in Volts, b is a voltage step dependent variable, t is the time and c is a constant, set to 1 so that the second term will become zero at time=0.

This equation shows another important behavior, $i.e.$, for $b>0$ y becomes larger than y_0 with increasing time and for $b<0$ y becomes smaller than y_0. Therefore, eqn. (18) can be used without any changes for positive and negative voltage steps. The evaluation of b was done by fitting eqn. (18) to every image of the 24 steps made with the stepper-motor and by plotting the value over the step magnitude. The apparently linear behavior in *Figure 8-12* was first fitted with a straight line but due to the offset at zero volts the match was not good for small voltage steps. The function shown in *Figure 8-12* was also tried but finally an ordinary straight line (red curve) through the origin showed the best fit. The result of fitting over the whole range is shown in *Figure 8-13*.

Figure 8-12 Though also nonlinear functions where used (dark cyan) for fitting, the b-variable shows fairly linear behavior (red line).

Model: y=sinh(a*x/440)*cos(b*x/440)
Chi^2 = 0.04302

a	7.70081	±0.22475
b	2.91705	±0.14162

Parameter	Value	Error
B	0.02376	4.23667E-4

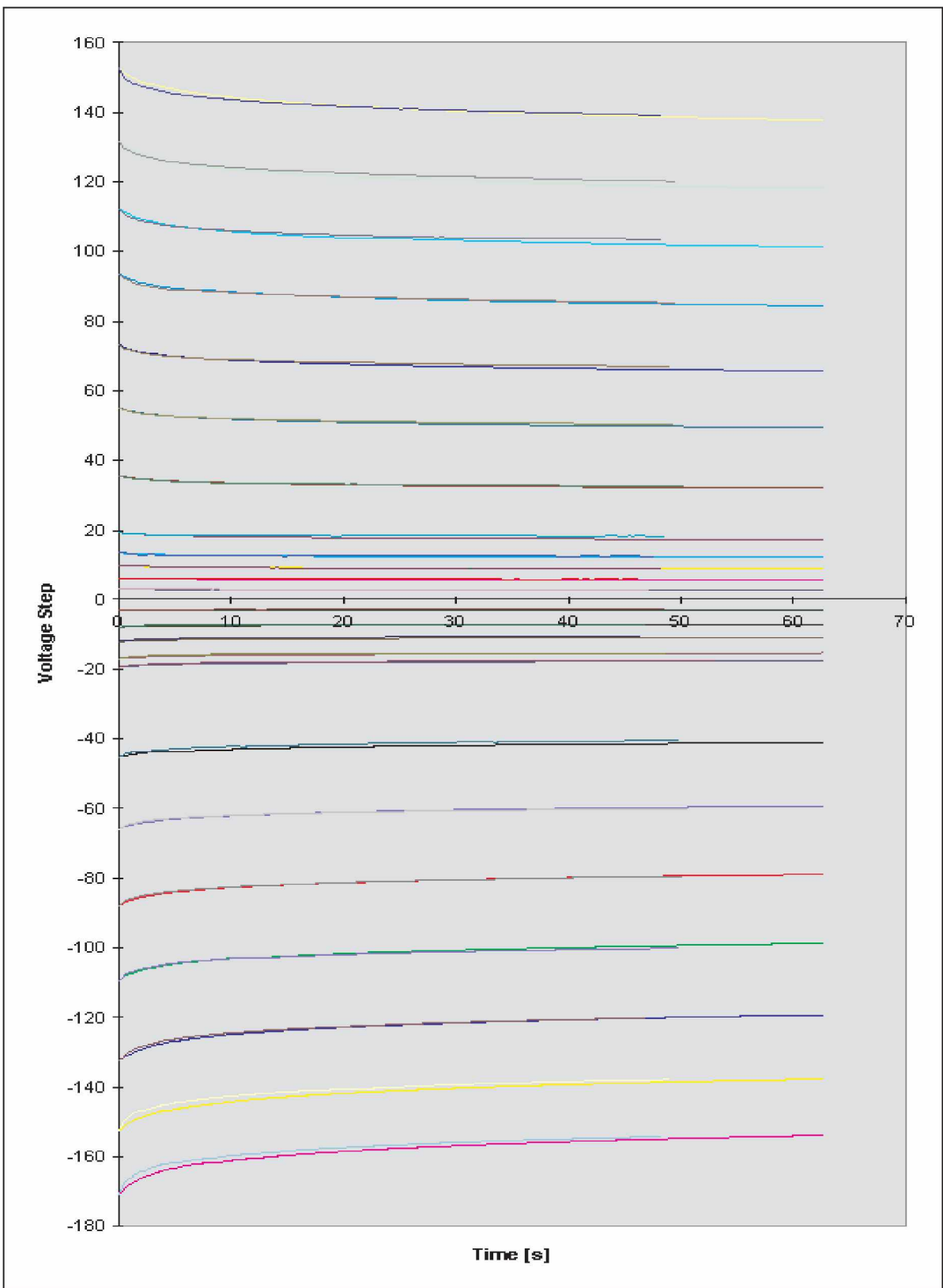| R	SD	N	P
0.99778	0.17463	21	<0.0001

Figure 8-13 Comparison of simulated and measured motor initiated steps. The simulated curves are ranging over 60 seconds while the measured data end at about 50 seconds.

The model of creep after a step of the calibration grid is still incomplete because no general matching function could be found. Eqn. (18) is not sensitive enough because at very short times the fit will always look like a straight line. This can be solved by eqn. (19) at the expense of a good fit for the motor induced steps.

$$y = y_0 - b * \sqrt{\ln(t+1)} \qquad (19)$$

It should be once more emphasized that the response of the scanner, with respect to creep, happens very fast. The exact timescale for a Trace and Retrace scan over a step are given in *Figure 8-14*. *Figure 8-15* shows the magnified sections after the step with the respective colors.
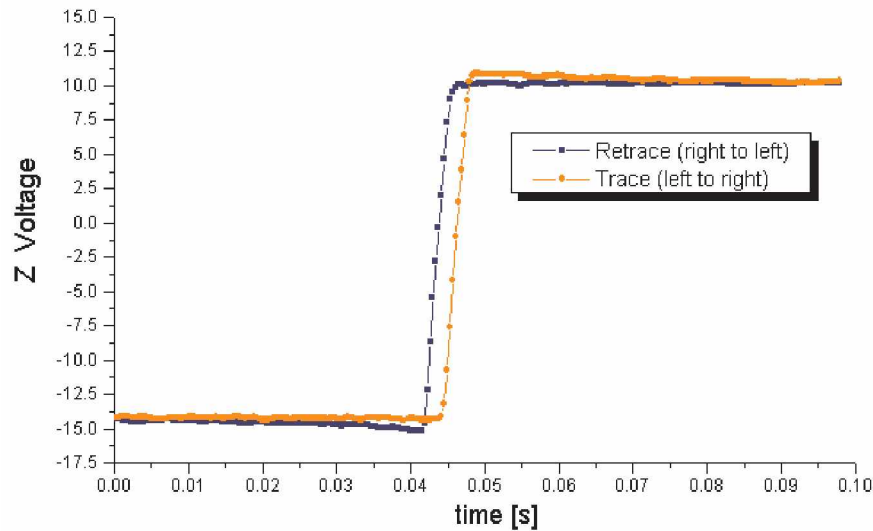

Figure 8-14 Section of one particular Trace and Retrace where the creep can be clearly seen after traversing the step.
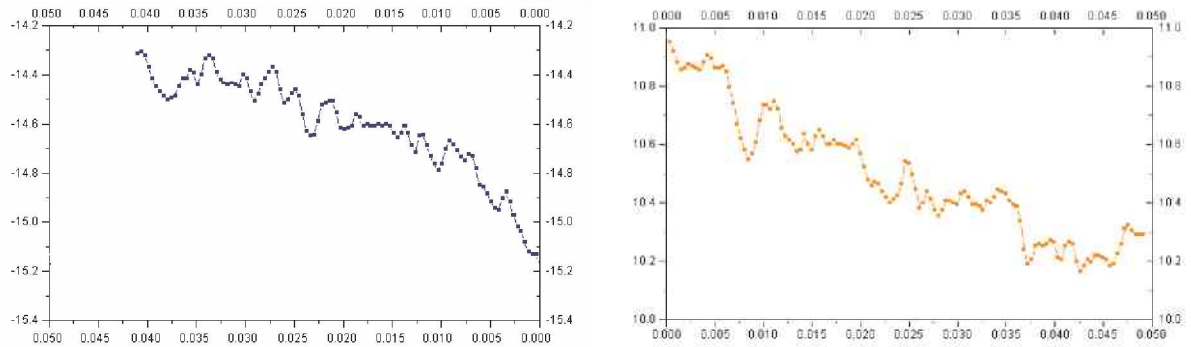

Figure 8-15 Magnified view of Figure 8-14 with the corresponding colors.

8.5 Nonlinearity Measurements on Smooth Samples

Some of the discussed nonlinearities, like hysteresis or cross coupling, can not be investigated with the methods presented so far. Therefore another possibility to investigate the nonlinearities of the scanner was considered where smooth, tilted or untilted, samples are imaged. Use of inclined surfaces in the Z sensitivity determination has been described in the literature [48], but the intrinsic nonlinearity calculation presented there should dealt with caution.

8.5.1 Experimental Setup

The samples for the experiments were made from a parallel ground steel plate with further milling for the inclined sample. Onto it a piece of silicon (from a silicon wafer) was glued. Steel was used because it is easy to machine and it is magnetic, which is needed for mounting the sample to the instrument. Silicon provides a surface flat enough for the experiments to be presented.

All tests were carried out only with the NanoScope AFM, except the cross coupling experiments, where the StandAlone AFM was also used. The cross coupling experiments with the StandAlone needed a zero degree surface, but when investigating intrinsic nonlinearity or hysteresis, the tilt of the sample should be large enough to reach the upper limit of the Z scale within the maximum scan size. Therefore a tilt of 15 degrees was chosen, where the full Z range (for the NanoScope AFM) is reached below 13 μm scan size.

The inclined sample was mounted in the instrument so that the fast scan direction had to traverse the full height magnitude. Naturally, the magnitude depends on the scan size when the tilt of the sample is constant. Alignment of the sample was done so that the height difference between start and end of the image on the slow scan direction was less than 30nm when imaging with the maximum scan size.

It was already mentioned that careful Z Center Voltage control has to be maintained to get meaningful results. When the instrument was started, it showed strong changes in the Z Center Voltage during normal operation. Sometimes it took up to 30 minutes until it stabilized. Due to changes in scan size and scan rate the Z Center Voltage had to be several times readjusted by withdrawing and engaging the tip again.

Double images (Trace and Retrace) where captured as a following series: 0.35, 0.7, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11 μm, each of them captured at scan rates of 1, 2, 4, 8, 16, 32, 49 and 98 Hz. All these 208 images had a resolution of 256×256 image points because of MS Excel's limitation.

8.5.2   Intrinsic Nonlinearity and Hysteresis
          Measurements

The image data was averaged along the slow scan direction and the standard deviation was calculated. It will not be given in the following figures because it was very low and further the graphs would lose their clarity.

Figure 8-16 shows data for all the used scan sizes (0.35 - 11 μm) at 4.06 Hz. Evidently the Trace and Retrace curves do not match at the left and right side of the image, which will be further called the Offset.
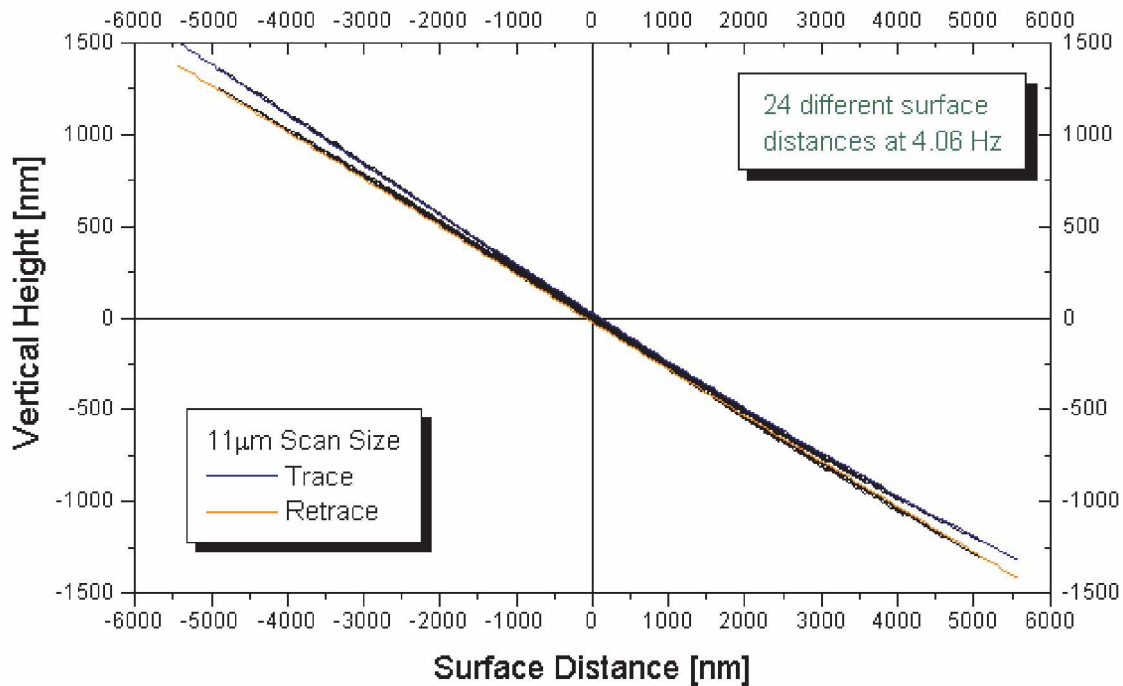

Figure 8-16 Vertical heights for 24 different scan sizes on a 15 degree inclined sample. Only the Trace and Retrace at the highest used scan size are given in color.

It can be clearly seen that all the curves are basically within the two colored curves, representing the largest scan size. Only at surface distances between 1500 and 5000 slightly lower values than the orange curve were measured. Furthermore, the curve shows that intrinsic nonlinearity is also present, resulting in the slight S-shape of the curve.

Representations of hysteresis curves usually show a closed shape, which is somehow misleading because it is not the usual case. Experiments by M. Heyde *et.al.* [27] required up to ten up and down cycles to get a closed hysteresis loop. J. Fu [48], who used a similar method for hysteresis determination, made the loop closed by recalculating the data. The end of one image was assumed to have the same height as the start of the next (considered as Offset). For full closure at the other side, the difference was treated as creep where each data was applied with a straight line linear function.

The method used here (Figure 8-17) only applies an Offset correction between Trace and Retrace images and leaves the amount of non-closure visible.
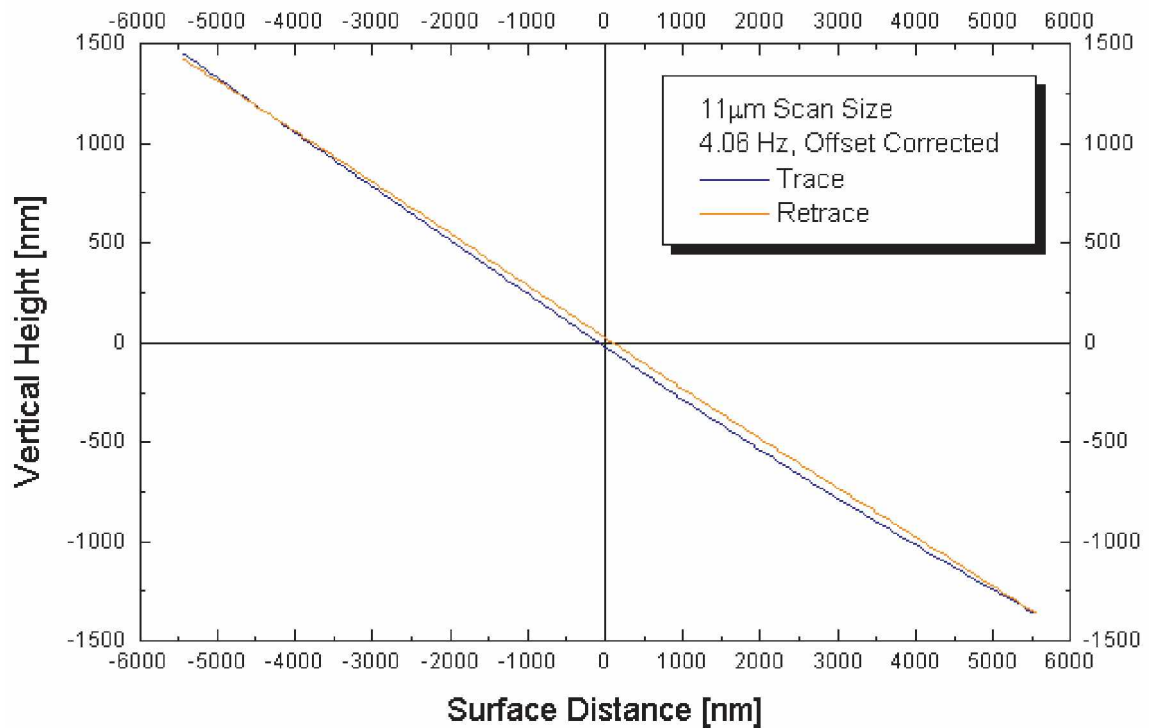

Figure 8-17 Offset corrected hysteresis curve for 11μm scan size and 4.06 Hz scan rate. The Retrace data was increased and the Trace data was decreased by 50.02 nm of magnitude.

Not considering the removed Offset, the curve in Figure 8-17 shows a hysteresis of exactly 2% of the full vertical height. This means that, when traversing a step, the step height given by the instrument depends on the direction of capturing. If the instrument was calibrated for Z voltage changes from high to low, the step would appear smaller when imaged from low to high Z voltages, due to the hysteresis.

The measured hysteresis with the NanoScope AFM is not seen during normal Z calibration because it will totally vanish in the standard deviation of measured heights, which was observed with the intrinsic nonlinearity measurements on steps (8.3).

Though the amount of the hysteresis is important, it is probably of more interest when the S-shape of the curve is investigated in more detail. This is better not to do on one particular curve because the exact S-shape is not known. Furthermore, problems may arise from the Offset. Therefore the data of corresponding Trace and Retrace images were averaged to define a common Z=0 point for both. From there, angles to the outermost image points where calculated for all the used scan sizes and scan rates.

Due to the Offset, the angle at the beginning of the Trace and at the end of the Retrace are higher than at the opposite image location. Figure 8-18 shows the angles for three scan rates and all the scan sizes shown. The expression "stressed" means that the angle was calculated for the last captured image points of the corresponding trace. A "relaxed" angle means that the first points of the image were used for the calculation.
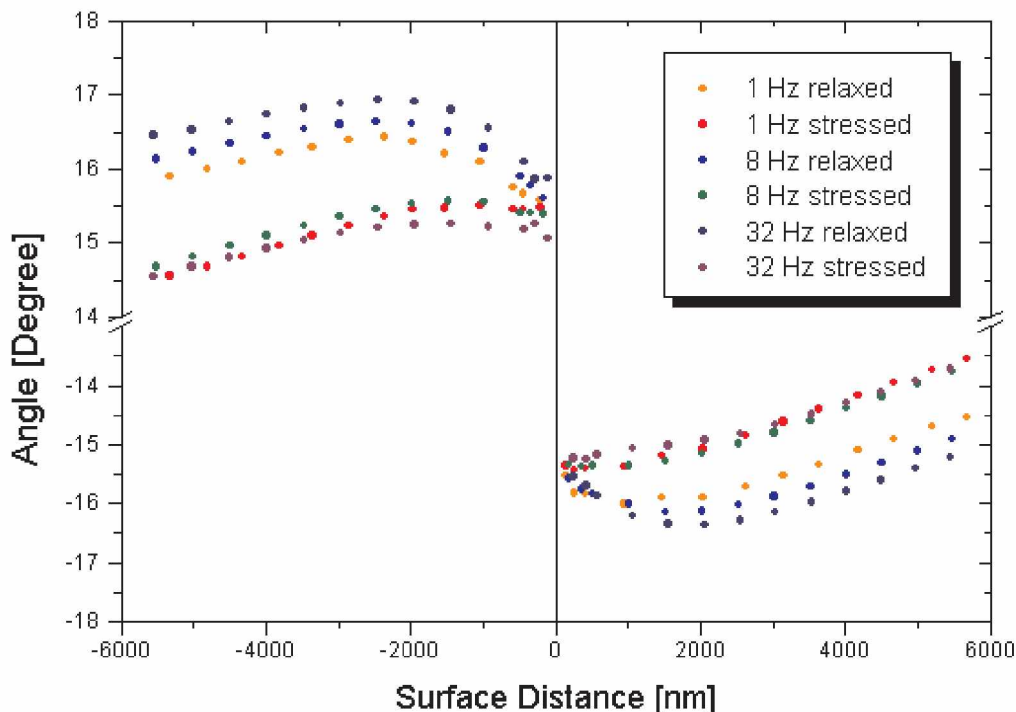

Figure 8-18 Calculation of angles from a Z=0 point to the last image points of a respective scan size. Only three different frequencies are shown for clarity.

However, it should be noted that the same color for image points at positive and negative surface distances only means that the same frequency was used but the points belong to two different captured images.

An increase in the calculated angles means an increase in the sensitivity of the Z piezoelement. For the "stressed" angles with negative surface distance, only a slight increase of the first view angles is visible, while all the others are decreasing. In contrast to this, a rather strong angle increase can be seen for the "relaxed" angles.

The split of the curves, especially for the "relaxed" angles is due to the calculation method used, because due to the increase of the Offset at higher scan rates the Z=0 point is not anymore very close to the Trace and Retrace curves, which will increase the angle. However, the Offset seems to take place mainly at the "relaxed" angles.

To visualize this behavior better, the whole sets of "relaxed" angles (Figure 8-19) and "stressed" angles (Figure 8-20) are given with the scatter plots fitted together.

The exact incident angle was not calculated, but this should not be very difficult by applying a fitting function to the quite well defined curves.
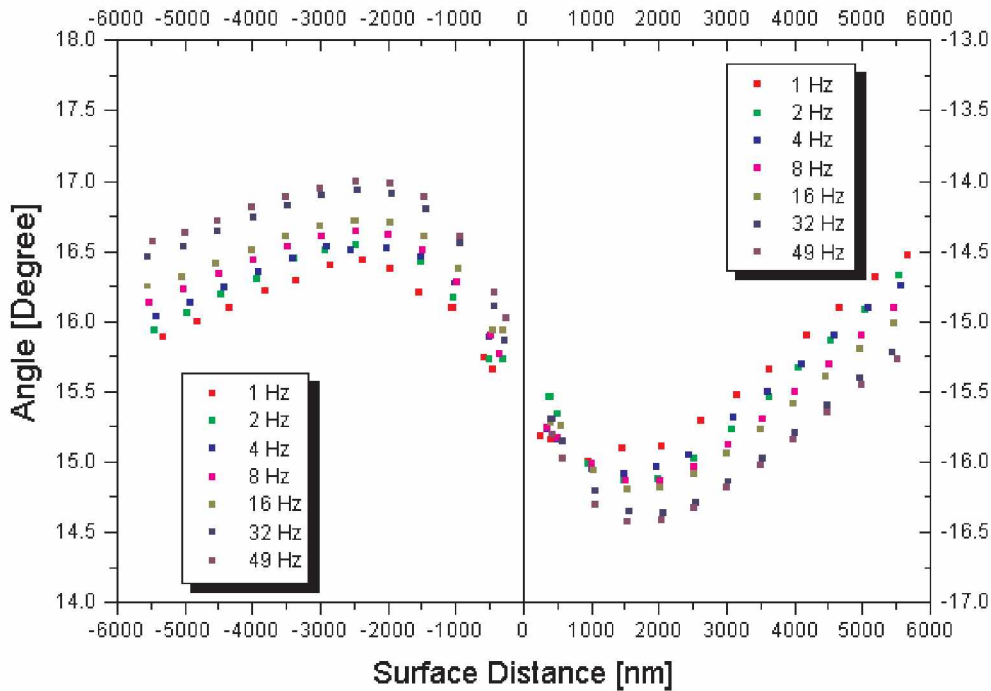

Figure 8-19 Positive and negative "relaxed" angles fitted together. Note the different Y scale, where the scale on the left applies to the negative surface distance angles and on the right to the positive ones.
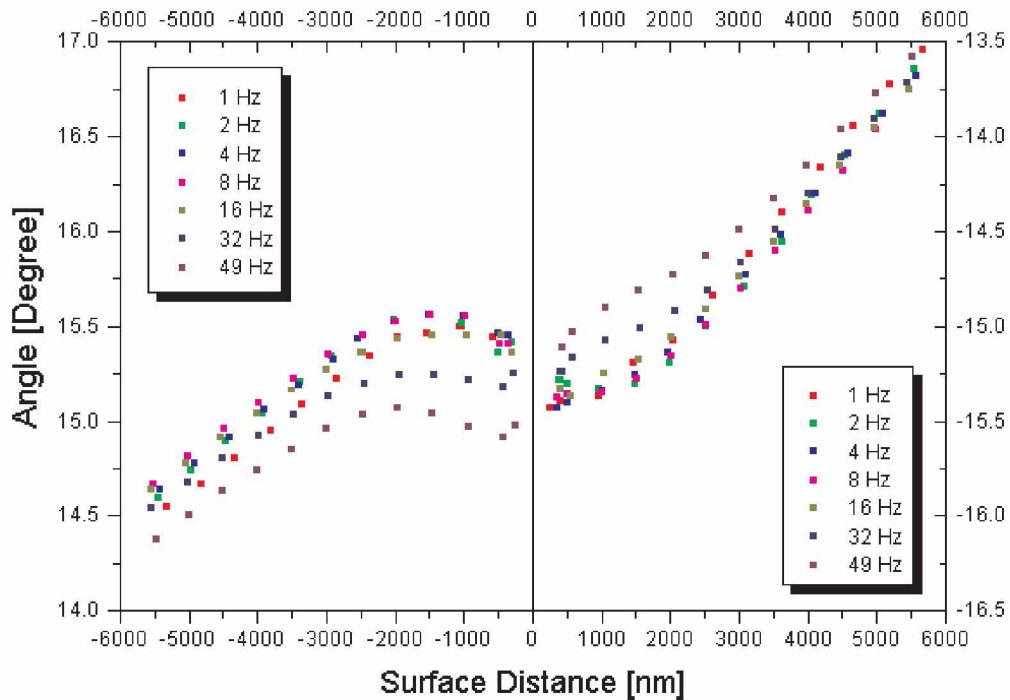

Figure 8-20 Positive and negative "stressed" angles fitted together.

The difficult part in calculating the exact sensitivity values from data like this lays in the increasing Offset, because the angle is very sensitive to changes in it. For considering the Offset it seems that it is necessary to know the exact timescale of the imaging, because the Offset could be caused by creep.

8.5.3 Creep Measurements

The time spent on one trace will depend on the scan rate, but the time for capturing one image value was always the same (16 μs) for the images and data presented here. The maximum scan rate of 98Hz means that only one image value is used for the image point, but as soon as the scan rate is reduced to *e.g.*, 49Hz, the final image point will be averaged from two collected image values.
The increase of the Offset from Trace to Retrace images with higher scan rates is shown in more detail in Figure 8-21.
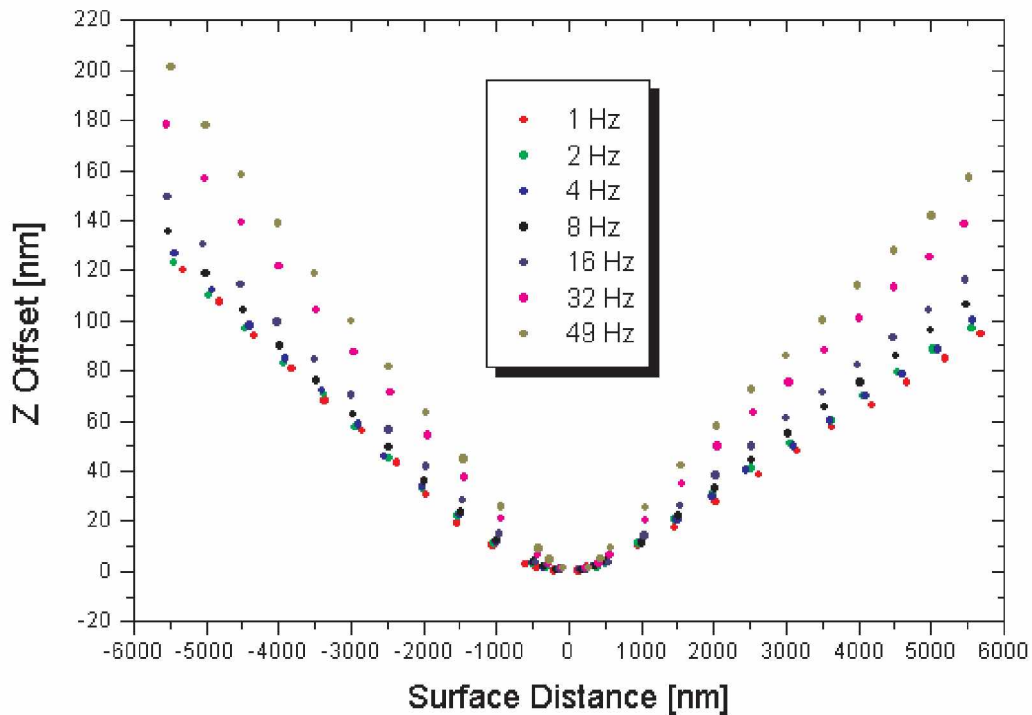

Figure 8-21 Z Offset values between all the Trace and Retrace images taken at several scan rates.

It is unlikely that even at a scan rate of 1Hz creep alone would cause the observed Offset. Unfortunately, the manufacturer has not so far given an answer to the question, whether there is an additional known influence or not. The difference in the Offset between slow and fast scan rates is probably due to creep, and by considering the involved timescale this could be a very powerful method to calculate variables for creep simulation.

8.5.4 Cross Coupling Measurements

The magnitude of cross coupling for the NanoScope AFM is negligible, because even over the whole scan size of 13μm no significant bowing of a flat surface could be observed. The StandAlone AFM was also tested to see whether it behaves in a similar manner. The Retracted and Extended Offset Derate was set to zero and a flat sample was imaged without any Real-time planefit and Offline planefit. The parallel alignment of the StandAlone AFM with the surface was accomplished during operation with the screws of the instrument. The adjustment takes some time, but after this the instrument showed an interesting behavior. The Up and Down scan frames could not be adjusted so that the pole would have stayed constant in the middle. In Figure 8-22 and Figure 8-23 this is given with the corresponding cross section profiles.
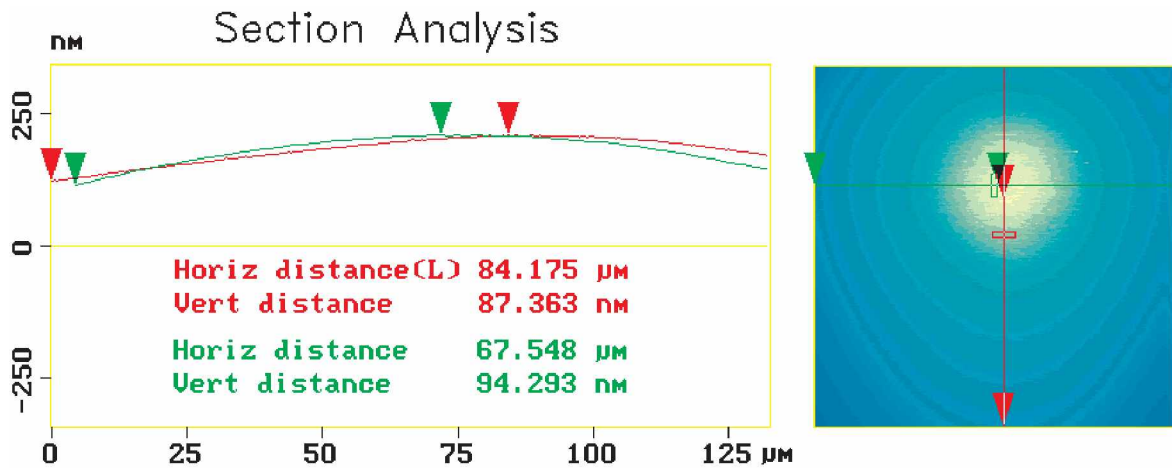

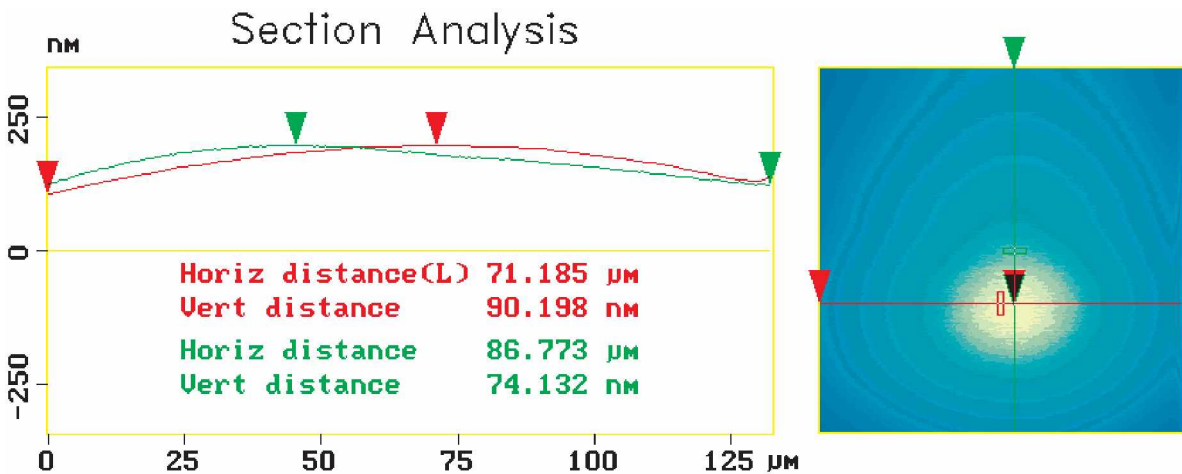*Figure 8-22 Down scan with the StandAlone AFM on a flat silicon sample.*

Figure 8-23 Up scan with the StandAlone AFM, captured directly after taking the image for Figure 8-22.

This behavior raised a question if there could be some sort of regularity behind it, and therefore the two images where substracted. The resulting "wave" from this operation is shown in Figure 8-24 with the heights outlined.
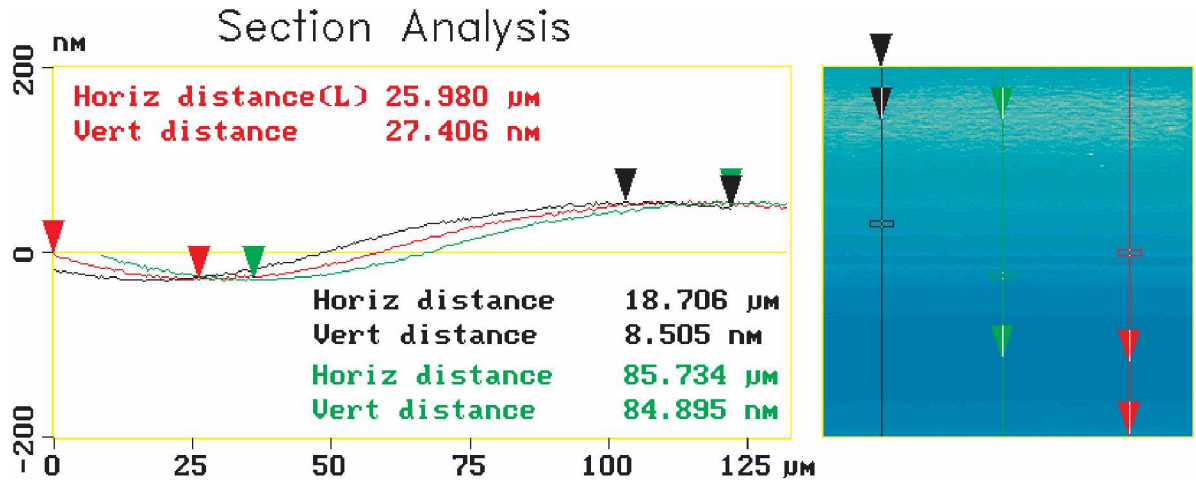
Figure 8-24 Section analysis of the substracted images, which reveals a quite strong wave. The perpendicular direction to the drawn lines were perfectly flat.

Compared to the scan size this wave is small, but still the regularity is rather surprising. To make sure that this behavior is a real effect, several series were investigated, which all showed the same outcome. An explanation for this could not be found, but at least this could be the reason why it becomes rather difficult to obtain stable imaging conditions in the nanometer scale.

9 Summary

The original aim of this work was to initiate the use of the newly purchased StandAlone AFM and to study its capabilities and performance. Actual start was delayed due to various instrument and software related problems. However, when these where solved, the focus was placed on an investigation routine of combined use of the StandAlone AFM and the SEM with its EBSD and μ-Indenter. The μ-Indenter was considered as a "link" between the SEM and the StandAlone AFM, providing the possibility to define areas of interest with a grid of indentations. Due to the limited stage tilt when the μ-Indenter is mounted in the SEM, a pre-tilt holder was made so that the μ-Indenter does not have to be removed when the EBSD is used. To locate the desired area with the StandAlone AFM, the sample has to be moved independently from the StandAlone AFM. This required another holder, where the StandAlone AFM is placed onto a support-plate and the sample is mounted to a micrometer stage underneath. The whole setup was tested and proved already to be valuable for other users.

When everything worked as expected, the AFM instrument accuracies were considered to bear important issues. The less generally known uncertainties of the piezoelectric scanner where investigated by easily accessible possibilities, *e.g.*, indentations, step features and (un)tilted smooth surfaces. While indentations seem to have too many factors that contribute to the final measurements, step features and smooth surfaces are very powerful tools for investigations. All the nonlinearities of the scanner can be measured and modelled with them. For creep, a modelling suggestion was presented and a new method to calculate variables for it was proposed. The outlined experiments to define the scanner nonlinearities can be used as a guide to develop a full image recalibration routine. This could have a major impact on the interpretation of AFM images.

The regularly observed phenomenon of tip convolution (or imaging) is the second major issue for AFM images. Recently adapted routines from Mathematical Morphology have proven to allow blind tip reconstruction and image simulation. Therefore a program was written that can read DI files as well as other formats. It also includes a freely available C-library for the calculations. The architecture of the program was planned so that it is easy to include further developments, either in tip reconstruction and image simulation or image recalculation based on scanner nonlinearities.

10 References

[1] R.Howland, L. Benatar, *A Practical Guide to Scanning Probe Microscopy,* Park Scientific Instruments.

[2] C.V. Newcomb, I. Flinn, *Electr. Lett,* 1982, 18 (11), 442.

[3] Digital Instruments, *Nanoscope III STM Instruction Manual*, 1992, V2.2, 16.

[4] G. Binnig, H. Rohrer, Ch. Gerber, E. Weibel, *Phys. Rev. Lett.,* 1982, 49 (1), 57.

[5] *Procedures in Scanning Probe Microscopies,* John Wiley & Sons, 1998.

[6] N.D. Lang, *Phys. Rev. Lett.,* 1985, 55, 230.

[7] Personal conversation with Prof. Anders Thölen, Chalmers TH, Sweden.

[8] G. Binnig, C.F. Quate, Ch. Gerber, *Phys. Rev. Lett.,* 1986, 56 (9), 930.

[9] T.R. Albrecht, S. Akamine, T.E. Carver, C.F. Quate, *J. Vac. Soc. Technol. A,* 1990, 8 (4), 3386.

[10] Digital Instruments, *Command Reference Manual,* 1995, V4.10.01, 1-17.

[11] Digital Instruments, *Nanoscope III AFM Instruction Manual*, 1992, V2.2, 5.

[12] Digital Instruments, *Nanoscope III StandAlone Multimode AFM Users Manual*, 1993, V0.01, 3.

[13] S. Park, J. Nogami, C.F. Quate, *Phys. Rev. B,* 1987, 36 (5), 2863.

[14] S. Heike, T. Hashizume, Y. Wada, *J. Vac. Sci. Technol. B ,* 1996, 14 (2), 1522.

[15] M. Tsukada, K. Kobayashi, N. Isshiki, H. Kageshima, *Surf. Sci. Rep.,* 1991, 13, 265.

[16] J.E. Griffith, D.A. Grigg, *J. Appl. Phys.,* 1993, 74 (9), R83.

[17] D.J. Keller, *Surf. Sci. ,* 1991, 253, 353.

[18] F. Atamny, A. Baiker, *Surf. Sci.,* 1995, 323, L314.

[19] J.S. Villarrubia, *Surf. Sci.,* 1994, 321, 287.

[20] J.S. Villarrubia, *J.Vac. Sci. Technol. B,* 1996, 14 (2), 1518.

[21] J. Schneir, J.S. Villarrubia, T.H. McWaid, R. Dixson, *J. Vac. Sci. Technol. B,* 1995, 14 (2), 1540.

[22] J.S. Villarrubia, *J. Res. Natl. Inst. Stand. Technol.,* 1997, 102, 425.

[23] D.J. Keller, F.S. Franke, *Surf. Sci. ,* 1993, 294, 409.

[24] S. Dongmo, M. Troyon, P. Vautrot, E. Delain, N. Bonnet, *J. Vac. Sci. Technol. B,* 1996, 14 (2), 1552.

[25] C. Odin, J.P. Aime, Z. El Kaakour, T. Bouhacina, *Surf. Sci.,* 1994, 317, 321.

[26] Personal conversation with Micha Semmler, ETH-ITOe Zürich, Switzerland.

[27] M. Heyde, H. Sturm, K. Rademann, *Surface and Interface Analysis,* to be published.

[28] R.W. Basedow, T.D. Cocks, *J. Phys. E: Sci. Instrum.,* 1980, 13, 840.

[29] J.F. Jørgenson, L.L. Madsen, J. Garnaes, K. Carneiro, K. Schaumburg, *J. Vac. Sci. Technol. B,* 1994, 12 (3), 1698.

[30] Digital Instruments, Inc., Santa Barbara, CA 93103, www.di.com.

[31] R.C. Barrett, C.F. Quate, *Rev. Sci. Instrum.,* 1991, 62, 1393.

[32] Digital Instruments Support Note No.: 217, Rev. D, *Fine Calibration for the SPM - a User' Guide,* 1996.

[33] Sandra Zitz, Master Thesis, Dec. 1998, University of Leoben (Austria) and Tampere University of Technology (Finland).

[34] J.S. Villarrubia, *Proc. SPIE 3332,* 1998, 10.

[35] J.F. Jørgenson, *SPIP-the Scanning Probe Image Processor,* http://www.dfm.dtu.dk/spip/

[36] Y. Li, S.M. Lindsay, *Rev. Sci. Instrum.,* 1991, 62 (11), 2630.

[37] D. Alliata, C. Cecconi, C. Nicolini, *Rev. Sci. Instrum.,* 1995, 67 (3), 748.

[38] S. Xu, M.F. Arnsdorf, *Journal of Microscopy,* 1994, 173 (3), 199.

[39] F. Jensen, *Rev. Sci. Instrum.,* 1993, 64 (9), 2595.

[40] H.M. Brodowsky, U-C. Boehnke, F. Kremer, *Rev. Sci. Instrum.,* 1996, 67 (12), 4198.

[41] M. Jaschke, H-J. Butt, *Rev. Sci. Instrum.,* 1995, 66 (2), 1258.

[42] l.A. Nagahara, K. Hashimoto, A. Fujishima, D. Snowden-Ifft, P.B. Price, *J. Vac. Sci. Technol. B,* 1994, 12 (3), 1694.

[43] T. Ohmi, S. Aoyama, *Appl. Phys. Lett.,* 1992, 61 (20), 2479.

[44] S.S. Sheiko, M. Möller, E.M.C.M. Reuvekamp, H.W. Zandbergen, *Phys. Rev. B,* 1993, 48, 5675.

[45] K. Zeng, E. Söderlund, A.E. Giannakopoulos, D.J. Rowcliffe, *Acta mater.,* 1996, 44, 3, 1127.

[46] S.V. Hainsworth, H.W. Chandler, T.F. Page, *J. Mater. Res.,* 1996, 11, 8, 1987.

[47] K. Zeng, A.E. Giannakopoulos, D.J. Rowcliffe, *Acta metall. mater.,* 1995, 43, 5, 1945.

[48] J. Fu, *Rev. Sci. Instrum.,* 1995, 66 (7), 3785.

[49] Anton Paar GmbH, A-8054 Graz, Austria, www.anton-paar.com/ap/.

Appendix

Microscope   Controller   Computer Workstation
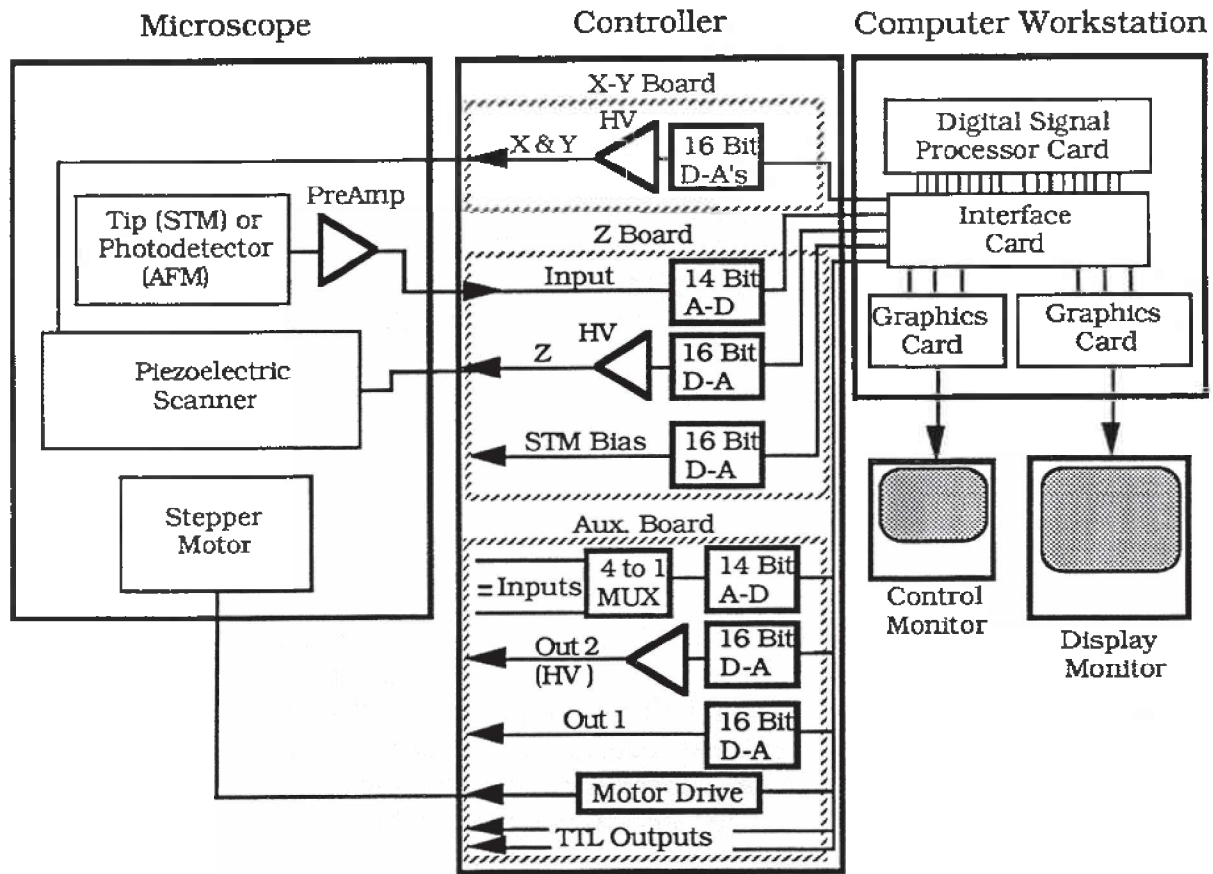
Appendix

Digital Instruments File Header Parameters

*File list
\Version: 0x04220104
\Date: 11:37:53 AM Fri Aug 08 1997
\Start context: OL2
\Data length: 8192
*Afm list
\Operating mode: Image
\Non-square scan: No
\Scan size: 4000 nm
\Aspect ratio: 1:1
\X offset: 1818.98 nm
\Y offset: 4515.66 nm
\Rotate Ang.: 0
\Samps/line: 512
\Lines: 512
\Scan rate: 4.06901
\Sample period: 160
\Units: Metric
\Color Table: 2
\Y disable: Enabled
\Z atten.: 65536
\Input attenuation mode: 1x
\Scope dualtrace: Dual
\Scope slope: Average
\Step size: 40
\Auto X Sep: 0
\Auto Y Sep: 0
\Auto pattern: Linear
\Auto number: 2
\Retrace mode: Lift
\Retrace height 1: 300 nm
\Retrace height: 0 nm
\Z modulation: Enable
\AFM mode: Contact
\Interleave mode: Disabled
\Parameter select: Main

\Int. gain: 375.7
\Prop. gain: 360.896
\Smart gain: 0
\Int. gain 2: 93.312
\Prop. gain 2: 64
\Smart gain 2: 0 linked
\Analog 1 inter: 0 linked
\Feedback type: Lin
\Detect sens.: 0.0414625
\Offset: 0
\Setpoint: -4509.64
\Setpoint 2: -15876.1 linked
*Microscope list
\Equip spec: 1 1021 871066996
\Serial number: 2794E
\Id: AFM E
\Start context: AFM
\File name: afme.par
\Piezo size: E
\Z sensitivity: 8.12985
\Z polarity: Forward
\Retracted offset der: 2.7
\Extended offset der: 7.6
\Piezo cal: 440
\X sensitivity: 21.5023
\X derate: 0.0274947
\X mag: 1.15
\X mag1: 0.630792
\X arg: 2.93515
\X round: 0.2
\Orthogonality: 0.0174533
\Y sensitivity: 21.8893
\Y derate: 0.0352517
\Y mag: 1.35
\Y mag1: 0.6
\Y arg: 2.5

\In sensitivity: 0.237 V
\In polarity: Forward
\Allow rotation: Allow
\Allow in attenuation: Allow
\Bias derate: 0
\Motor sensitivity: 19
\Motor direction: Reverse
\X slow sensitivity: 22.016
\X slow derate: 0.0361
\Y fast sensitivity: 20.936
\Y fast derate: 0.0291
\X slow-fast coupling: 0.0711884
\X slow-fast coupling derating: 4.02187e-005
\Y slow-fast coupling: 0.0754647
\Y slow-fast coupling derating: 2.77354e-005
\Fast cal freq: 2.44141
\Slow cal freq: 4.76837
\Xs-Yf coupling: 0.00568439
\Xs-Yf coupling derating: -8.95765e-006
\Ys-Xf coupling: 0.0122403
\Ys-Xf coupling derating: 4.69271e-006
\Is Defl: Yes
\Linearization mode: Normal
*Controller list
\X max: 220
\Y max: 220
\Z max: 220
\Bias max: 10
\Ana1 max: 10
\Ana2lv max: 12
\Ana2hv max: 220
\In1 max: 10
\In2 max: 10
\Sample max: 6553.6
*AFM image list
\Data offset: 8192
\Data length: 524288

\Start context: OL
\Data type: AFM
\Plane fit: 4808 3518 -4964 2
\Frame direction: Down
\Samps/line: 512
\Number of lines: 512
\Aspect ratio: 1:1
\Scan size: 4000 nm
\Line direction: Retrace
\Scan line: Interleave
\Image data: Height
\Z magnify image: 1.55361
\Z scale: 321.831 nm (6048)
\Realtime Planefit: Line
\Offline planefit: Full
\Highpass: 0
\Lowpass: 0
*AFM image list
\Data offset: 532480
\Data length: 524288
\Start context: OL
\Data type: AFM
\Plane fit: 4719 3533 -5045.75 2
\Frame direction: Down
\Samps/line: 512
\Number of lines: 512
\Aspect ratio: 1:1
\Scan size: 4000 nm
\Line direction: Trace
\Image data: Height
\Z magnify image: 1.52474
\Z scale: 336.23 nm (6160)
\Realtime Planefit: Line
\Offline planefit: Full
\Highpass: 0
\Lowpass: 0