

THE TIP-SAMPLE INTERACTION IN ATOMIC FORCE MICROSCOPY AND ITS  
IMPLICATIONS FOR BIOLOGICAL APPLICATIONS

Thesis by  
David Baselt

In Partial Fulfillment of the Requirements  
for the Degree of  
Doctor of Philosophy

California Institute of Technology  
Pasadena, California

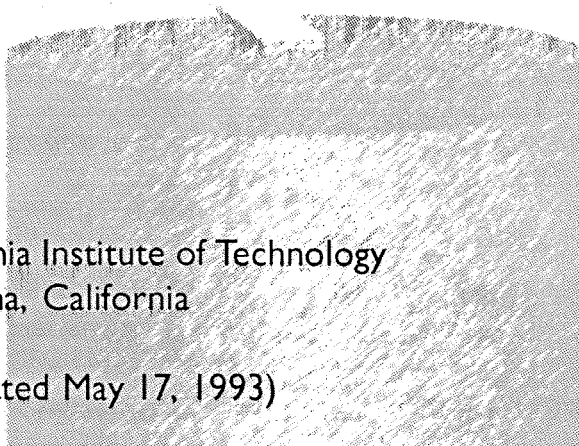
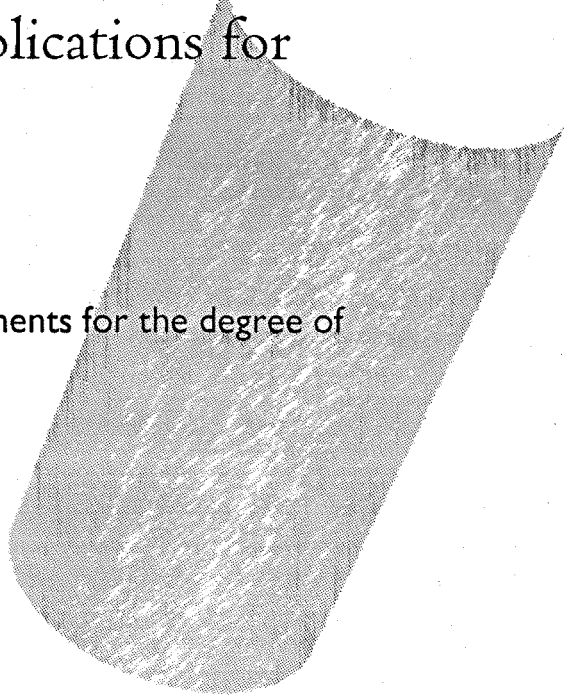
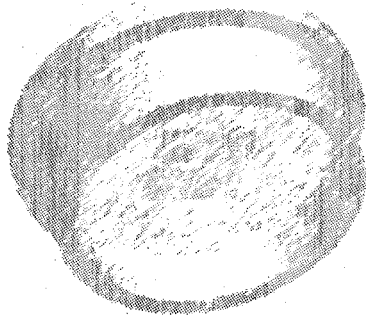
1993

(Submitted May 17, 1993)

# The tip-sample interaction in atomic force microscopy and its implications for biological applications

Thesis by David Baselt

In partial fulfillment of the requirements for the degree of  
Doctor of Philosophy



California Institute of Technology  
Pasadena, California  
1993  
(Submitted May 17, 1993)

*D e d i c a t e d t o J L C † M C M L X X V I I I*

Copyright © 1993 by David Baselt  
All rights reserved

Typeset in Centaur 12/22 with titles and captions in Gill Sans

# Table of Contents

Introduction	1
Part 1. Construction	
<i>Introduction</i>	3
Chapter 1. Principles of AFM	
<i>Introduction</i>	4
<i>General concept and defining principles of AFM</i>	4
<i>Alternative imaging modes</i>	12
<i>AFM and biology</i>	15
<i>References</i>	17
Chapter 2. The stage	
<i>Introduction</i>	19
<i>Vibrational and thermal stability</i>	20
<i>Scanner design</i>	28
<i>Cantilever design</i>	30
<i>Optical lever design</i>	33
<i>Alignment mechanisms</i>	36
<i>Tip-sample approach</i>	38
<i>The scanned-cantilever AFM</i>	39
<i>References</i>	45
<i>Appendix: AFM design tables</i>	46
Chapter 3. The control system	
<i>Introduction</i>	56
<i>Hardware components</i>	59



<i>PC-DSP communication</i>	60
<i>Timing considerations</i>	61
<i>Self-optimizing PID feedback</i>	62
<i>Scanning</i>	70
<i>Tip-sample approach</i>	74
<i>Modulation measurements</i>	75
<i>Spectroscopy</i>	77
<i>The user interface</i>	81
<i>References</i>	83

## Chapter 4. Image processing

<i>Introduction</i>	84
<i>Review of standard SPM image processing algorithms</i>	85
<i>Newly-developed SPM image processing algorithms</i>	94
<i>References</i>	99

## Part 2. Application

<i>Introduction</i>	100
---------------------	-----

## Chapter 5. The lateral meniscus force

<i>Introduction</i>	101
<i>Contact force and cantilever deflection</i>	102
<i>Methods</i>	104
<i>Observations when scanning over steps</i>	106
<i>Effect of boron impurities</i>	115
<i>Atomic resolution lateral deflection images</i>	116
<i>References</i>	119

## Chapter 6. Contrast mechanisms in lateral force microscopy

<i>Introduction</i>	120
<i>Methods</i>	121
<i>Coupling with X, Y, and Z</i>	122
<i>Coupling with the error signal</i>	126
<i>Meniscus effects</i>	130
<i>Slope-dependent contact force effects</i>	131
<i>The differential lateral deflection image</i>	133
<i>References</i>	134

## Chapter 7. Hardness imaging

<i>Introduction</i>	135
<i>Measurement of sample compression</i>	136
<i>Slope coupling</i>	137
<i>Z coupling</i>	139
<i>Nonlinear elastic behavior</i>	139
<i>References</i>	141

## Chapter 8: Float-polished quartz

<i>Introduction</i>	142
<i>Float polishing</i>	143
<i>Curvature removal</i>	144
<i>Tip shape artifacts</i>	145
<i>Evaluation of surface finish</i>	147
<i>References</i>	149

## Chapter 9. Collagen

<i>Introduction</i>	150
<i>Type I collagen</i>	151
<i>Sample preparation</i>	152
<i>Instrumentation</i>	153
<i>Results</i>	153
<i>Discussion</i>	162
<i>Conclusion</i>	166
<i>References</i>	168

## Part 3. Nearfield Scanning Optical Microscopy

<i>Introduction</i>	170
---------------------	-----

## Chapter 10. Nearfield Scanning Optical Microscopy

<i>Introduction</i>	171
<i>General principles</i>	171
<i>The stage</i>	174
<i>Preliminary results</i>	177
<i>References</i>	178

# Standard terms and abbreviations

A	Sample hardness signal (for AFM modulation images)
ADC	Analog-to-digital converter
AFM	Atomic force microscopy or microscope
AFM-3	The AFM I used for the work presented in this thesis
BPG	Boronated pyrolytic graphite
C	Cantilever deflection signal (AFM) or tunneling current (STM)
D	Derivative feedback gain
DAC	Digital-to-analog converter
DNA	Deoxyribonucleic acid
DOS	Disk operating system (a personal computer standard)
DSP	Digital signal processor
E	Error signal (difference between C and S)
FFM	Frictional force microscopy (synonymous with LFM)
H	Z modulation amplitude (for hardness images)
HOPG	Highly oriented pyrolytic graphite
I	Integral feedback gain
I/O	Input and output
kw	Kilowords (information storage capacity)
LFM	Lateral force microscopy (synonymous with FFM)
ms	Milliseconds
mV	Millivolts
Mw	Megawords (information storage capacity)
nm	Nanometers
nN	NanoNewtons
ns	Nanoseconds
NSOM	Nearfield scanning optical microscopy
P	Proportional feedback gain
PC	Personal computer (DOS-based)

pm	Picometers
pN	PicoNewtons
PROCESS	The image processing software I have developed
S	Feedback setpoint
SCAN	The control system software I have developed
SPM	Scanned probe microscopy
STM	Scanning tunneling microscopy or microscope
V	Volts
X	X piezoceramic voltage (fast raster direction)
Y	Y piezoceramic voltage (slow raster direction)
Z	Z piezoceramic voltage (normal to sample surface)
$\mu\text{m}$	Micrometers

# Abstract

This thesis describes the construction of an atomic force microscope and its application to the study of tip-sample interactions, primarily through the use of friction and hardness (elasticity) imaging.

Part one describes the atomic force microscope, which consists of a scanned-cantilever stage (chapter 2); a versatile digital signal processor-based control system with self-optimizing feedback, lock-in amplifier emulation (for hardness imaging), and macro programmability (chapter 3); and image processing software (chapter 4).

Part two describes a number of results that have helped to characterize the tip-sample interaction and the contact imaging modes used for its study. Meniscus forces act laterally as well as normally, and that they vary with position (chapter 5). Friction measurements couple with scanner position and feedback, and the meniscus effects friction images (chapter 6). Sliding of the tip over the sample surface introduces slope-dependence into hardness measurements (chapter 7). Dull tips can create prominent topography artifacts even on very flat surfaces (chapter 8).

In an investigation of collagen fibrils, AFM has revealed the characteristic 65 nm banding pattern, a second, minor banding pattern, and microfibrils that run along the fibril axis. The distribution of proteoglycans along the fibrils creates a characteristic pattern in friction images. Although imaging in water reduces interaction forces, water can also make biological samples more sensitive to force. However, for robust biological samples imaged in air, tip shape presents a greater obstacle than tip-sample interaction forces to obtaining high-resolution images. Tip contamination increases tip-sample friction and can occasionally improve resolution (chapter 9).

For a separate project I have designed a general-purpose nearfield scanning optical microscope (chapter 10).

# Introduction

I have furthered our experimental understanding of how the tip-sample interaction in AFM affects biological structures

In the broadest sense, I have carried out the research presented in this thesis to improve the methods that we use to observe and understand biological structure. My interests lie in the field of cell biology — in the dynamics of cellular molecules — and currently available techniques severely limit such studies. Clearly a method that provides nanometer-resolution movies of living cellular components would provide great benefits.

The scanned-proximity probe techniques developed over the past decade, and atomic force microscopy (AFM) in particular, promise to provide such a capability. Several people, including myself, have demonstrated that AFM can obtain images rivaling the best results from electron microscopy — but on living rather than dead samples. However, the technique still requires significant refinement. As in electron microscopy, imaging artifacts tend to obscure features and limit resolution.

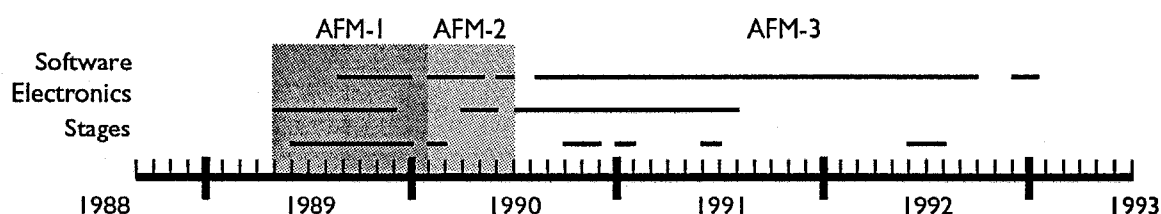
Unlike electron microscopy, though, the actual imaging process rather than the sample preparation process tends to create these

artifacts. AFM works by dragging a tip over the sample, and the forces that the tip exerts on the sample can adversely affect biological structures. Thus, my work has focused on finding, understanding, and controlling AFM imaging artifacts. I have done this by extending the capabilities of the AFM to provide more information about the tip-sample interaction.

Extending the capabilities of an AFM required making extensive modifications to it. Therefore, I needed a particularly versatile instrument that I had full understanding of. Since commercial instruments did not meet these standards, I constructed my own. Figure 1 shows that doing this took up most of my time at Caltech. Part one of this thesis presents the lessons learned about the effective construction and use of AFMs, and describes the novel capabilities I have implemented.

In part two I describe a number of experiments, focusing on the tip-sample interaction rather than what the experiments have revealed about the structure of the various samples imaged.

Figure 1. AFM development timeline. I spent my first year at Caltech studying the theory of AFM design. I began experiments in August 1991 but continued developing software to meet requirements discovered in the process. I did not begin doing experiments full-time until 1993.





# *Part I* **Construction**

## *Introduction*

---

Part one conveys working knowledge of atomic force microscope design and use

Though conceptually simple, atomic force microscopy requires many refinements to achieve its full potential. Part one of this thesis describes these refinements, presenting details concerning the stages, control systems, and image processing software that I have created.

My principal contributions to the field of AFM design include the development of the scanned-cantilever stage, self-optimizing feedback, and elasticity imaging. Although other force microscopists developed these same techniques simultaneously, I have created significant new refinements for each.

# Chapter I. Principles of AFM

## *Introduction*

---

Chapter I presents background information that the other chapters assume knowledge of

For those not familiar with the technique, chapter 1 reviews the fundamentals of atomic force microscopy. Microscopists other than myself developed all of the concepts presented here; although I did implement the friction and elasticity imaging modes independently, I was not the first to report them.

## *General concept and defining characteristics of AFM*

---

Scanned-proximity probe microscopes provide very high resolution images of various sample properties

The atomic force microscope is one of about two dozen types of scanned-proximity probe microscopes. All of these microscopes work by measuring a local property — such as height,

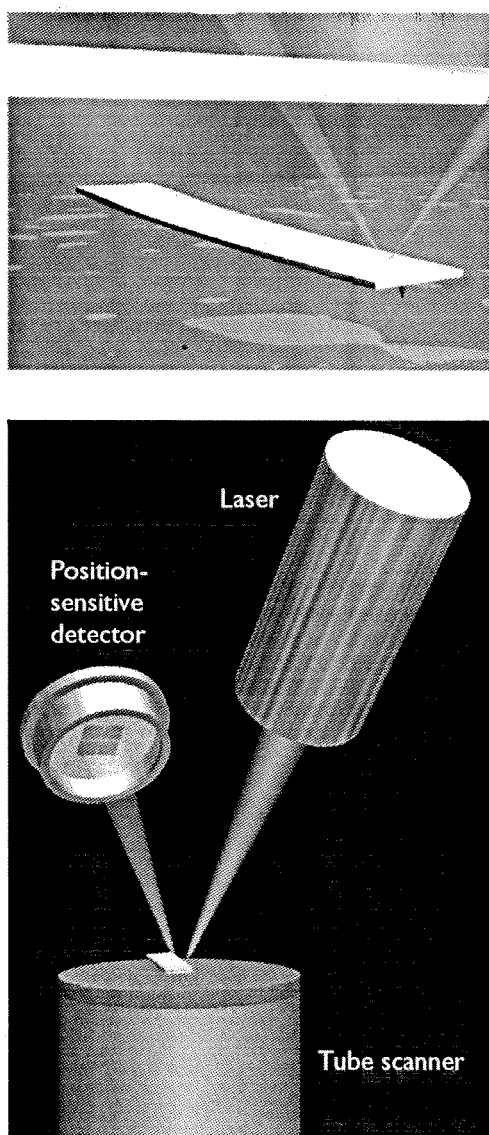


Figure 1. Concept of AFM and the optical lever: (top) a cantilever touching a sample; (bottom) the optical lever. Scale drawing; the tube scanner measures 24 mm in diameter, while the cantilever is 100  $\mu\text{m}$  long.

optical absorption, or magnetism — with a probe or “tip” placed very close to the sample. The small probe–sample separation (on the order of the instrument’s resolution) makes it possible to take measurements over a small area. To acquire an image the microscope raster-scans the probe over the sample while measuring the local property in question. The resulting image resembles an image on a television screen in that both consist of many rows or lines of information placed one above the other.

Unlike traditional microscopes, scanned-probe systems do not use lenses, so the size of the probe rather than diffraction effects generally limit their resolution.

The atomic force microscope measures topography with a force probe

AFM (figure 1) operates by measuring attractive or repulsive forces between a tip and the sample (Binnig et al., 1986). In its repulsive “contact” mode, the instrument lightly touches a tip at the end of a leaf spring or “cantilever” to the sample. As a raster-scan drags the tip over the sample, some sort of detection apparatus measures the vertical deflection of the cantilever, which indicates the local sample height.

Thus, in contact mode the AFM measures hard-sphere repulsion forces between the tip and sample.

In noncontact mode, the AFM derives topographic images from measurements of attractive forces; the tip does not touch the sample (Albrecht et al., 1991). Because it does not allow the imaging of samples under water, I have not used the attractive mode.

AFMs can achieve a resolution of 10 pm, and unlike electron microscopes, can image samples in air and under liquids.

In principle, AFM resembles the record player as well as the stylus profilometer. However, AFM incorporates a number of refinements that enable it to achieve atomic-scale resolution:

- Sensitive detection
- Flexible cantilevers
- Sharp tips
- High-resolution tip-sample positioning
- Force feedback

I describe these refinements below.

**Laser beam deflection offers a convenient and sensitive method of measuring cantilever deflection**

AFMs can generally measure the vertical deflection of the cantilever with picometer resolution. To achieve this most AFMs today use the optical lever, a device that achieves resolution

comparable to an interferometer while remaining inexpensive and easy to use (Meyer et al., 1988; Alexander et al., 1989).

The optical lever (figure 1) operates by reflecting a laser beam off the cantilever. Angular deflection of the cantilever causes a twofold larger angular deflection of the laser beam. The reflected laser beam strikes a position-sensitive photodetector consisting of two side-by-side photodiodes. The difference between the two photodiode signals indicates the position of the laser spot on the detector and thus the angular deflection of the cantilever.

Because the cantilever-to-detector distance generally measures thousands of times the length of the cantilever, the optical lever greatly magnifies motions of the tip. Because of this  $\sim 2000$ -fold magnification optical lever detection can theoretically obtain a noise level of  $10^{-14} \text{ m/Hz}^{1/2}$  (Putman et al., 1992). For measuring cantilever deflection, to date only the relatively cumbersome techniques of interferometry and tunneling detection have approached this value.

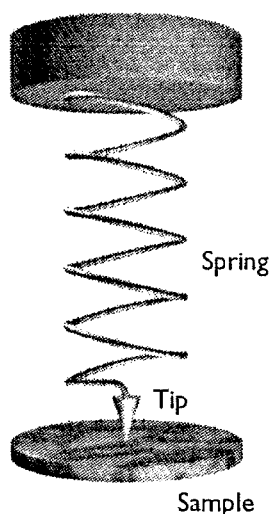


Figure 2. Schematic illustration of the meaning of “spring constant” as applied to cantilevers. Visualizing the cantilever as a coil spring, its spring constant  $k$  directly affects the downward force exerted on the sample.

### AFM cantilevers have high flexibility

A high flexibility stylus exerts lower downward forces on the sample, resulting in less distortion and damage while scanning. For this reason AFM cantilevers generally have spring constants of about  $0.1 \text{ N/m}$  (figure 2). For comparison, a Slinky is about ten times stiffer ( $1 \text{ N/m}$ ).

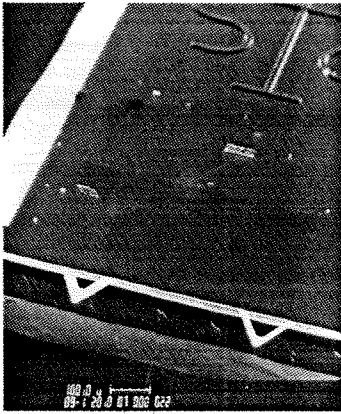


Figure 3. Electron micrograph of two 100 μm long V-shaped cantilevers (by Jean-Paul Revel, Caltech; cantilevers from Park Scientific Instruments, Sunnyvale, CA).

However, it would take a very long time to image a surface by dragging a Slinky over it (in the configuration of figure 2), because a Slinky cannot respond quickly as it passes over features. That is, a Slinky has a low resonant frequency, but an AFM cantilever should have a high resonant frequency.

The equation for the resonant frequency of a spring:

$$\text{resonant frequency} = \frac{1}{2\pi} \sqrt{\frac{\text{spring constant}}{\text{mass}}}$$

shows that a cantilever can have both low spring constant and high resonant frequency if it has a small mass. Therefore AFM cantilevers tend to be very small. Commercial vendors manufacture almost all AFM cantilevers by microlithography processes similar to those used to make computer chips. The Park Scientific Instruments cantilevers in figure 3 measure 100 μm in length and consist of silicon oxynitride with a thin coating of gold for reflectivity.

Micromachining techniques produce inexpensive, reasonably sharp tips

Most users purchase AFM cantilevers with their attached tips from commercial vendors, who manufacture the tips with a variety of microlithographic techniques.

A close enough inspection of any AFM tip reveals that it is rounded off. Therefore force microscopists generally evaluate tips

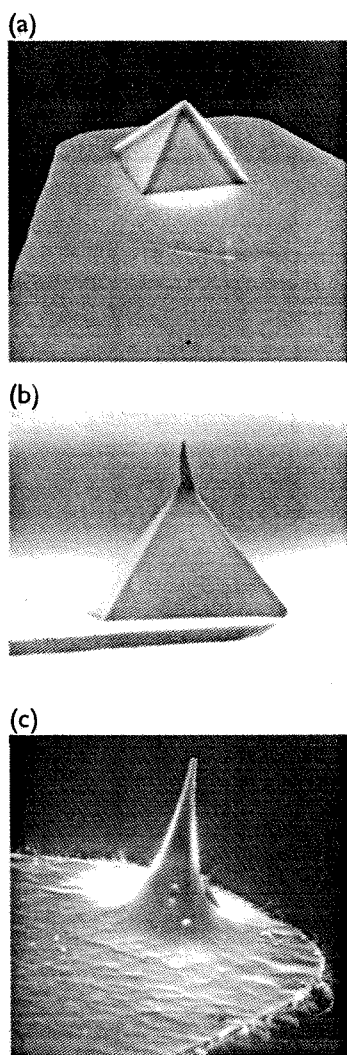


Figure 4. Three common types of AFM tip. (a) normal tip (3  $\mu\text{m}$  tall); (b) supertip; (c) Ultralever (also 3  $\mu\text{m}$  tall). Electron micrographs by Jean-Paul Revel, Caltech. Tips from Park Scientific Instruments; supertip made by Jean-Paul Revel.

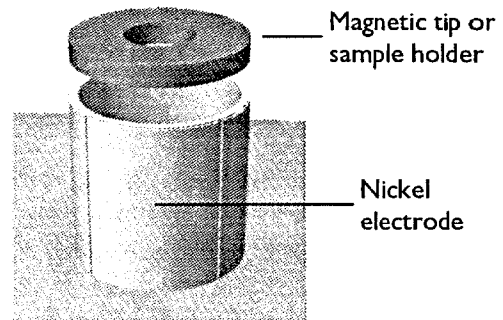
by determining their “end radius.” In combination with tip-sample interaction effects, this end radius generally limits the resolution of AFM. As such, the development of sharper tips is currently a major concern.

Force microscopists generally use one of three types of tip. The “normal tip” (figure 4a; Albrecht et al., 1990) is a 3  $\mu\text{m}$  tall pyramid with  $\sim 30$  nm end radius. The electron-beam-deposited (EBD) tip or “supertip” (figure 4b; Keller and Chih-Chung, 1992) improves on this with an electron-beam-induced deposit of carbonaceous material made by pointing a normal tip straight into the electron beam of a scanning electron microscope. Especially if the user first contaminates the cantilever with paraffin oil, a supertip will form upon stopping the raster of the electron beam at the apex of the tip for several minutes. The supertip offers a higher aspect ratio (it is long and thin, good for probing pits and crevices) and sometimes a better end radius than the normal tip. Finally, Park Scientific Instruments offers the “Ultralever” (figure 4c), based on an improved microlithography process. Ultralevers offers a moderately high aspect ratio and on occasion a  $\sim 10$  nm end radius.

Tube piezoceramics position the tip or sample with high resolution

Piezoelectric ceramics are a class of materials that expand or contract when in the presence of a voltage gradient or, conversely,

Figure 5. Exploded view of a tube scanner. Applying a voltage to one of the four outer quadrants causes that quadrant to expand and the scanner to tilt away from it (XY movement). A corresponding negative voltage applied to the opposite quadrant doubles the XY range while preventing vertical motion. Applying a voltage to the inner electrode causes the entire tube to expand or contract (Z movement).



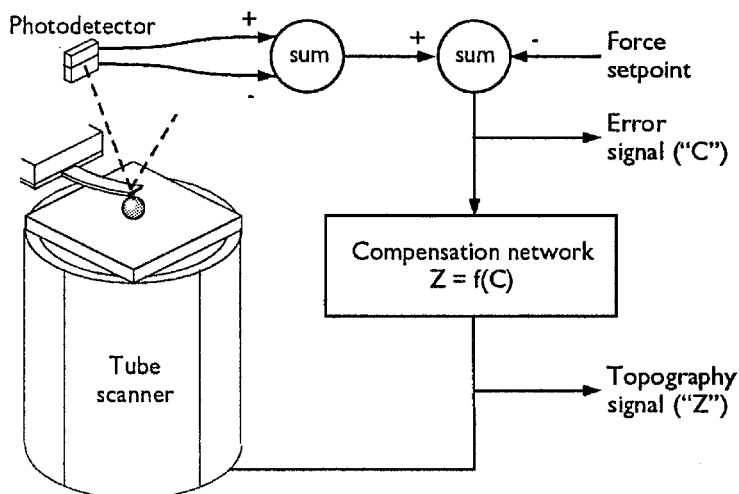
create a voltage gradient when forced to expand or contract (Gallego-Juárez, 1989). Piezoceramics make it possible to create three-dimensional positioning devices of arbitrarily high precision. Most scanned-probe microscopes use tube-shaped piezoceramics because they combine a simple one-piece construction with high stability and large scan range. Four electrodes cover the outer surface of the tube, while a single electrode covers the inner surface. Application of voltages to one or more of the electrodes causes the tube to bend or stretch, moving the sample in three dimensions (figure 5).

**AFMs use feedback to regulate the force on the sample**

The presence of a feedback loop is one of the subtler differences between AFMs and older stylus-based instruments such as record players and stylus profilometers. The AFM not only measures the force on the sample but also regulates it, allowing acquisition of images at very low forces.



Figure 6. The AFM feedback loop. A compensation network (which in my AFM is a computer program) monitors the cantilever deflection and keeps it constant by adjusting the height of the sample (or cantilever).



The feedback loop (figure 6) consists of the tube scanner that controls the height of the entire sample; the cantilever and optical lever, which measures the local height of the sample; and a feedback circuit that attempts to keep the cantilever deflection constant by adjusting the voltage applied to the scanner.

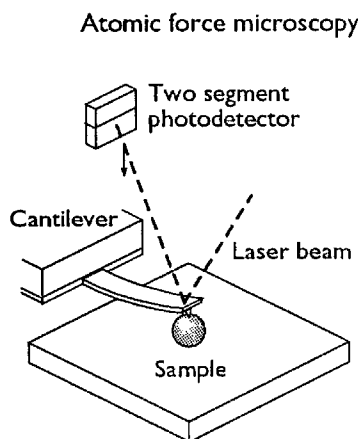
One point of interest: the faster the feedback loop can correct deviations of the cantilever deflection, the faster the AFM can acquire images; therefore, a well-constructed feedback loop is essential to microscope performance. AFM feedback loops tend to have a bandwidth of about 10 kHz, resulting in image acquisition times of about one minute.

## Alternative imaging modes

---

### AFMs have two standard imaging modes

Almost all AFMs can measure sample topography in two ways: by recording the feedback output (“Z”) or the cantilever deflection (“error”; see figure 6). The sum of these two signals always yields the actual topography, but given a well-adjusted feedback loop, the error signal should be negligible. As described below, AFMs may have alternative imaging modes in addition to these standard modes.



### Optical lever AFMs can measure the friction between tip and sample

If the scanner moves the sample perpendicular to the long axis of the cantilever (figure 7), friction between the tip and sample causes the cantilever to twist. A photodetector position-sensitive in two dimensions can distinguish the resulting left-and-right motion of the reflected laser beam from the up-and-down motion caused by topographic variations (Meyer and Amer, 1990).

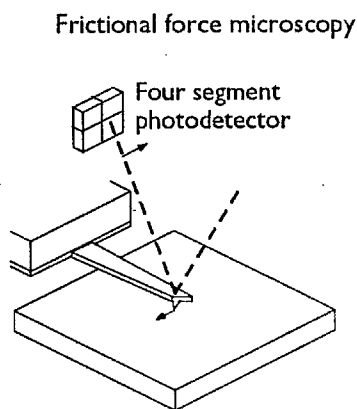


Figure 7. While topographic imaging uses the up-and-down deflection of the cantilever, friction imaging uses torsional deflection.

Therefore, AFMs can measure tip-sample friction while imaging sample topography. Besides serving as an indicator of sample properties, friction (or “lateral force,” or “lateral deflection”) measurements provide valuable information about the tip-sample interaction.

Figure 8.  $2.5 \times 2.5$  nm simultaneous topographic and friction image of highly oriented pyrolytic graphic (HOPG). The bumps represent the topographic atomic corrugation, while the shading reflects the lateral forces on the tip. The scan direction was right to left.

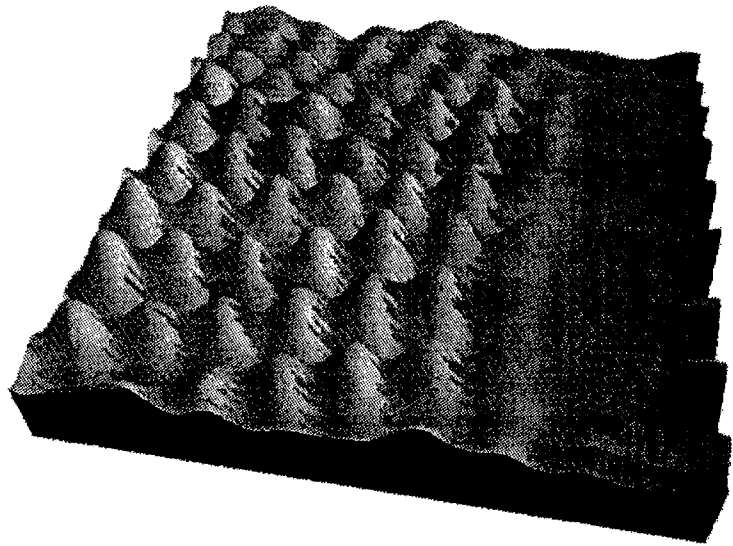


Figure 9. Cross-sectional profile of friction data from above image showing stick-slip behavior.

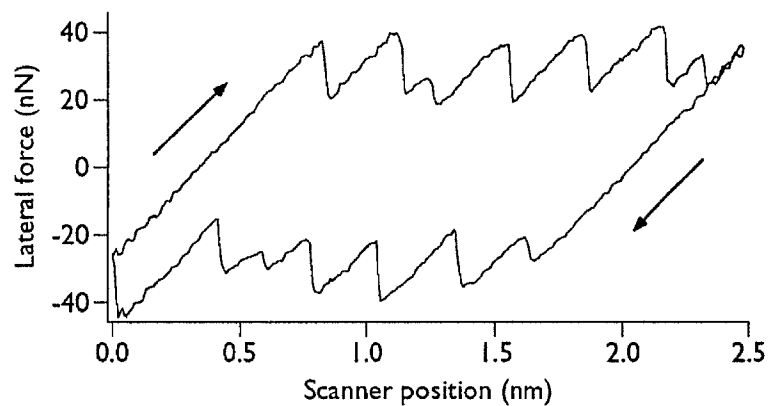


Figure 8 shows a simultaneous friction and topography image of graphite atoms in which I have plotted the topography image as a three-dimensional projection colored by the friction data. Each bump represents one carbon atom. As the tip moves from right to left, it bumps into an atom and gets stuck behind it. The scanner continues to move and lateral force builds up until the tip slips past the atom and sticks behind the next one. This “stick-slip” behavior creates a characteristic sawtooth waveform in the friction image (figure 9).

Figure 10. AFMs can image sample elasticity by pressing the tip into the sample and measuring the resulting cantilever deflection.

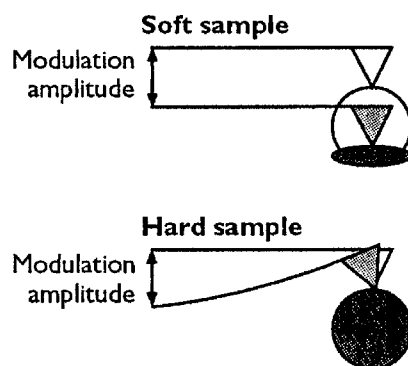
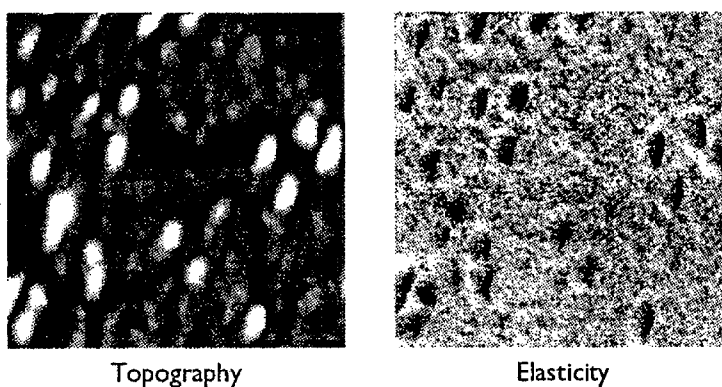


Figure 11.  $1 \times 1 \mu\text{m}$  simultaneous topography and elasticity images of bovine serum albumen on silicon (sample prepared by Sie-Ting Wong of Abbott Laboratories).



### AFMs can measure sample elasticity

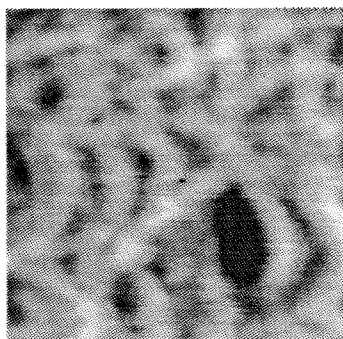
AFM can also image the softness of a sample by pressing the cantilever into it at each point in a scan. The scanner raises the sample or lowers the cantilever by a preset amount, the “modulation amplitude” (usually 1-10 nm). In response, the cantilever deflects an amount dependent on the softness of the sample: the harder the sample, the more the cantilever deflects (figure 10).

Figure 11 shows an image of bovine serum albumen (BSA) on silicon. A number of bumps appear in the topography image, each presumably corresponding to a single BSA molecule. The elasticity image reveals that each of the bumps is soft relative to the silicon substrate, a reasonable result for protein molecules.

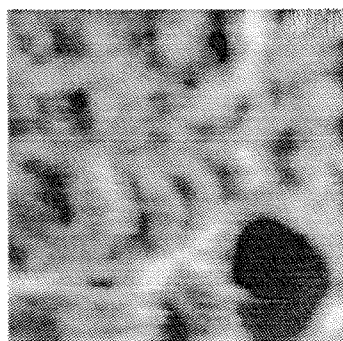
*AFM and biology*

---

Frame 1



Frame 50



Frame 100

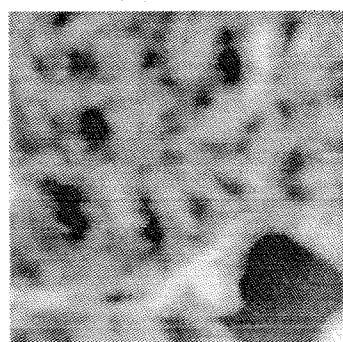


Figure 12. Three images of small collagen fibrils from a sequence of 100 images. Repetitive scanning of the same area progressively detaches the fibrils from the glass substrate, causing distortion in the direction of scanning, left-to-right and top-to-bottom.

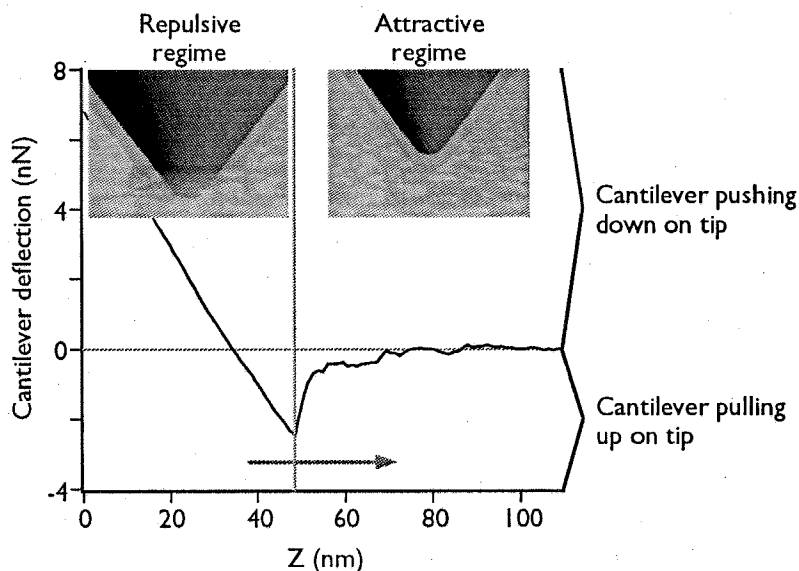
### Dull tips and tip-sample interaction forces prevent high-resolution imaging of biological structures

The ability of AFM to image at atomic resolution, combined with its ability to image a wide variety of samples under a wide variety of conditions, has created a great deal of interest in applying it to the study of biological structures. Images have appeared in the literature showing DNA, single proteins, structures such as gap junctions, and living cells (for a review see Hoh and Hansma, 1992).

Unfortunately, AFM cannot image all samples at atomic resolution. The end radii of available tips confines atomic resolution to flat, periodic samples such as graphite. In addition, because biological structures are soft, the tip-sample interaction tends to distort or destroy them. Figure 12, for example, shows how forces exerted on small collagen fibrils tend to detach them from the substrate over a period of time, resulting in progressively greater distortion.

A number of companies are attempting to develop sharper tips, primarily by improved microfabrication processes. I have concentrated on investigating the tip-sample interaction with alternative imaging modes.

Figure 13. Force curve taken in air. At  $Z=0$  nm the cantilever pushes down on the tip, and tip and sample are in contact. As  $Z$  increases, the cantilever exerts less force and then begins to pull up on the tip (negative force). Eventually the cantilever exerts enough force to pull the tip free of the meniscus (2 nN in this case, an unusually low figure). After this point, only attractive forces affect the cantilever deflection.



The meniscus force is the most important influence on the tip-sample interaction force when imaging in air

When imaging in air, a layer of water condensation and other contamination covers both the tip and sample, forming a meniscus that pulls the two together (Weisenhorn et al., 1989).

“Force curves” showing cantilever deflection as the scanner lowers the sample reveal the attractive meniscus force (figure 13): the cantilever has to exert an upward force to pull the tip free of the meniscus. This force equals the attractive force of the meniscus, usually 10–100 nN.

The great strength of the meniscus makes it the most important influence on the tip-sample interaction. Force microscopists often eliminate the meniscus by completely immersing both tip and sample in water.

## References

---

- Albrecht, T.R., Akamine, S., Carver, T.E., and Quate, C.F. (1990) Microfabrication of cantilever styli for the atomic force microscope. *J. Vac. Sci. Technol. A* 8(4), 3386–3396
- Albrecht, T.R., Grütter, P., Horne, D., and Rugar, D. (1991) Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity. *J. Appl. Phys.* 69(2), 668–673
- Alexander, S., Hellemans, L., Marti, O., Schneir, J., Elings, V., Hansma, P.K., Longmire, M., and Gurley, J. (1989) An atomic-resolution atomic-force microscope implemented using an optical lever. *J. Appl. Phys.* 65(1), 164–167
- Binnig, G., Quate, C.F., and Gerber, Ch. (1986) Atomic force microscope. *Phys. Rev. Lett.* 56(9), 930–933
- Gallego-Juárez, J.A. (1989) Piezoelectric ceramics and ultrasonic transducers. *J. Phys. E: Sci. Instrum.* 22, 804–816
- Hoh, J.H. and Hansma, P.K. (1992) Atomic force microscopy for high-resolution imaging in cell biology. *Trends Cell Bio.* 2, 208–213
- Keller, D.J. and Chih-Chung, C. (1992) Imaging steep, high structures by scanning force microscopy with electron beam deposited tips. *Surf. Sci.* 268, 333–339
- Meyer, G. and Amer, N.M. (1988) Novel optical approach to atomic force microscopy. *Appl. Phys. Lett.* 53(12), 1045–1047
- Meyer, G. and Amer, N.M. (1990) Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. *Appl. Phys. Lett.* 57(20), 2089–2091

Putman, C.A.J., De Grooth, B.G., Van Hulst, N.F., and Greve, J.

(1992) A detailed analysis of the optical beam deflection technique for use in atomic force microscopy. *J. App. Phys.* 72(1), 6–12

Weisenhorn, A.L., Hansma, P.K., Albrecht, T.R., and Quate, C.F.

(1989) Forces in atomic force microscopy in air and water. *Appl. Phys. Lett.* 54(26), 2651–2653



# Chapter 2. The stage

## *Introduction*

---

Chapter two describes aspects of AFM stage design learned during the construction of three stages

An AFM stage contains the sample, cantilever, tube scanner, and optical lever. It also contains an approach mechanism that can move the tip and sample close enough to lie within the range of the tube scanner.

I have constructed three AFM stages. The first two, AFM-1 and AFM-2, had similar design and used the standard scanned-sample architecture. AFM-2 had an improved optical lever alignment mechanism and accommodated larger samples than AFM-1.

I will not present specific details on these two stages, since I used them for none of the experiments in this thesis. The lessons learned from their design, however, constitute an essential part of this chapter. I will present specific details on the third stage, AFM-3, which uses a novel scanned-cantilever design. This stage

performed significantly better than the earlier ones, primarily because of better optics.

AFM-3 was also much easier to use. I have designed all parts of the AFM — the stage, electronics, and software — such that an expert user can perform experiments in a minimum of time. Noncommercial research instruments often neglect this quality, concentrating only on high performance. Nonetheless it is essential since it allows the user to perform more experiments, thereby perfecting techniques.

### *Vibrational and thermal stability*

---

Mechanical engineering concepts simplify the process of designing a stage stable to atomic dimensions

In order to obtain atomic resolution, an AFM stage must hold the tip and sample stationary relative to each other, resisting the effects of external vibrations and thermal expansion.

First-time designers of AFMs usually attempt to make the stage as stable as possible. This technique almost always backfires, yielding an unstable and cumbersome stage. I have found it much more effective to set a stability goal, then use standard mechanical engineering techniques to achieve that goal. The procedure I outline below has been developed by engineers to design structures such as bridges, electric shavers, and airplanes, and it works just as well for AFMs.

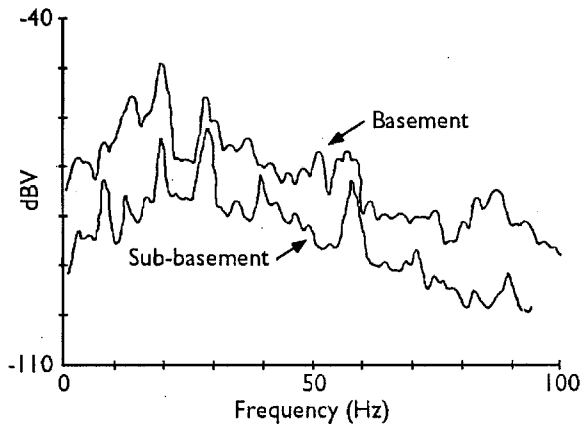


Figure 1. Seismograms taken in the basement and sub-basement of Noyes Laboratories. Taken by Robert Driscoll.

Given ambient vibrations of known magnitude, an AFM stage can be designed with a predetermined stability level

Specifying stage stability constitutes the first step of AFM design. For all of my instruments I have allowed 10 pm of noise from floor-driven stage vibrations and 10 pm from acoustically-driven stage vibrations. In combination with electronic noise (10 pm associated with measuring the cantilever

deflection and 10 pm from XYZ voltage noise) the instrument has 20 pm noise overall, sufficient for atomic scale imaging.

The next step involves measuring the floor-driven vibrations that the stage will have to resist. Figure 1 shows two seismograms taken in Noyes Laboratory. For each, the seismometer was located on the floor; ambient vibrations on a tabletop exceed floor vibrations by about an order of magnitude. Therefore I keep AFM stages on the floor. The ambient vibration spectrum of the sub-basement of Noyes laboratory peaks at about  $2\text{ }\mu\text{m}$  amplitude at 20 Hz.

The last step consists of proposing an AFM structure and determining how it will respond to this vibration. This involves determining the natural resonant frequency of each stage component. The higher the resonant frequency of a component, the greater its stability in the presence of low-frequency vibrations. To achieve a given stability, the resonant frequency must exceed:

$$f = F \sqrt{1 + \frac{A}{a}}$$

$f$  = resonant frequency of stage

$F$  = frequency of ambient vibration

$a$  = amplitude of stage vibration

$A$  = amplitude of ambient vibration

(derived from equations in Den Hartog, 1934). This equation applies if the object driving the vibrations (the floor) has a much greater mass than structure in question (the AFM stage).

Therefore, an AFM stage with 10 pm stability in the presence of a 2  $\mu$ m, 20 Hz ambient vibration must have a resonant frequency of at least 10 kHz. This criterion applies only to parts that support the tip or sample. Parts attached to the stage but not supporting the tip or sample do not degrade stability; in fact, they can improve stability by absorbing stage vibrations (Den Hartog, 1934).

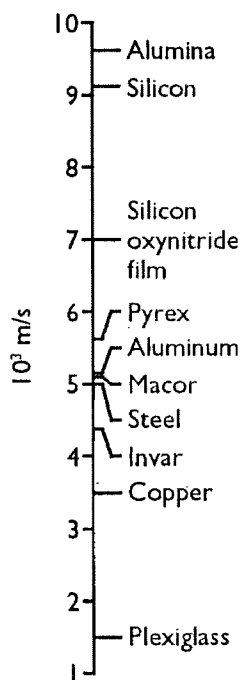


Figure 2. Square root of the elastic modulus: density ratio for several materials. From AFM design table I (appendix).

In general, small thick parts have a higher resonant frequency than large thin ones. Materials with a high stiffness-to-density ratio yield the best stability — the square root of this ratio is directly proportional to resonant frequency (figure 2).

A number of reference works provide formulas for the resonant frequency of simple geometries. To simplify matters, I have provided graphs that show how thick three common geometries — disks, rectangular plates, and rods — need to be, given their other dimensions (figure 3 and appendix). For example, the base of AFM-3 measures 3 inches in diameter; it therefore must

Figure 3. Thickness as a function of diameter for steel disks of 10 kHz resonant frequency. From AFM design table 7 (appendix).

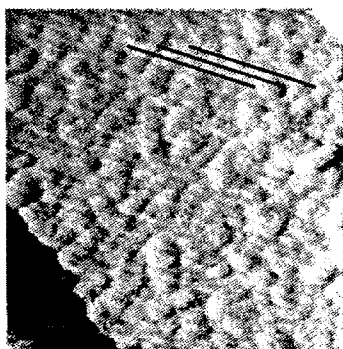
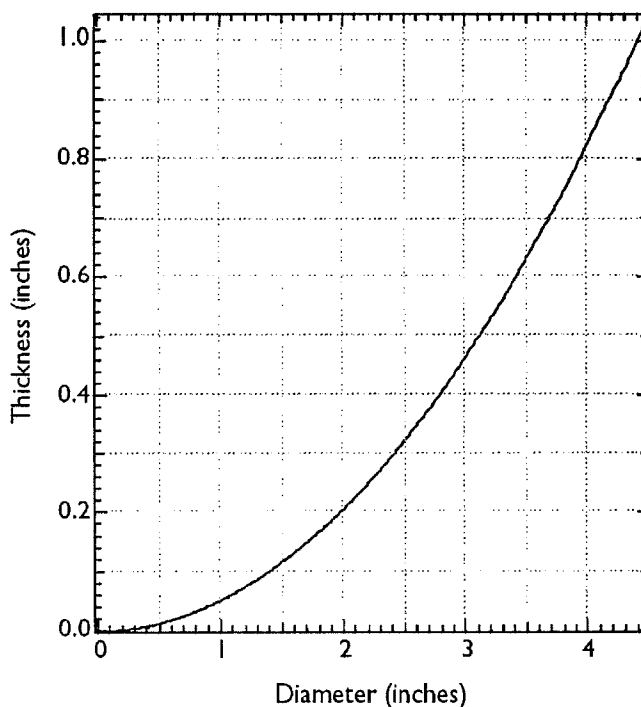


Figure 4. Typical "herring-bone" pattern (diagonal lines, highlighted) resulting from floor-driven vibrations.  $1\ \mu\text{m} \times 1\ \mu\text{m}$  portion of an image of collagen fibrils on glass. The bumps are part of the sample.

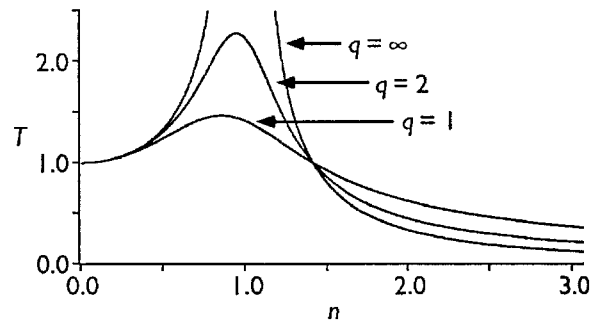
measure at least 0.5 inch in thickness. I have left a reasonable margin of error by doubling the thickness to 1 inch.

Images taken with low-stability stages exhibit evenly spaced diagonal lines with a frequency corresponding to the lowest resonant frequency of the stage (figure 4). Unstable stages also require longer feedback settling times (see chapter 3) to prevent the feedback from oscillating.

### Hanging the stage from rubber cords improves its stability

Although a well-designed stage can tolerate ambient vibrations, some sort of vibration isolation usually improves high-resolution images by ensuring an adequate margin for error and reducing the

Figure 5. Transmissibility of spring-and-mass vibration isolation as a function of normalized frequency. The vibration isolation is most effective at frequencies that are high compared to its own resonant frequency. Damping is beneficial at low frequencies but not at high frequencies.



effect of transient vibrations from doors slamming, people walking by, etc.

Hanging the stage from elastic cords helps to isolate it from ambient vibrations. Such spring-and-mass isolation acts most effectively if it has a low resonant frequency — generally about 1 Hz. Its transmissibility, defined as the ratio of transmitted to ambient force, is:

$$T = \frac{\sqrt{1 + \left(\frac{n}{q}\right)^2}}{\sqrt{(1 - n^2)^2 + \left(\frac{n}{q}\right)^2}}$$

$$n = \frac{\text{frequency of ambient vibration}}{\text{resonant frequency}}$$

$$q = \text{quality (damping) factor}$$

(figure 5)(den Hartog, 1934). Therefore, a 1 Hz resonant frequency vibration isolation system with large  $q$  transmits 0.25% of an ambient 20 Hz vibration.

Although better vibration isolation methods exist (for example, hanging a second spring-and-mass system from the first significantly improves results; see Park and Quate, 1987) the 1-Hz spring-and-mass system reduces floor-induced vibrations to negligible levels for 20 pm resolution AFM.

Since metal springs tend to transmit high-frequency vibrations, vibration isolation should use rubber cords. The mass of the stage should greatly exceed the mass of the cords to minimize transmission of higher-mode vibrations.

Note that the resonant frequency of a spring-and-mass system can be deduced by measuring the amount the spring stretches due to the mass:

$$f = \frac{1}{2\sqrt{\Delta l}}$$

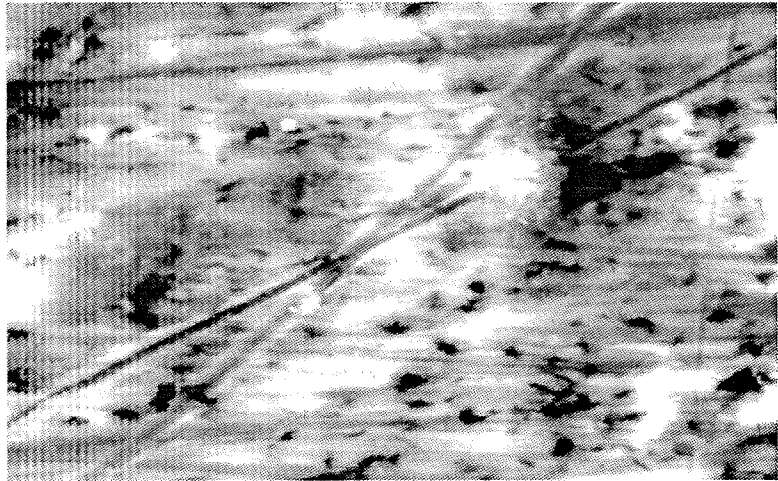
$\Delta l$  = length of spring with mass – length without mass  
( in meters)

Therefore, if the spring stretches 25 cm when suspending the mass from it, the system has a resonant frequency of 1 Hz.

### Smaller range scanners have more stability

Like other stage components, the scanner must have a minimum resonant frequency to achieve a specified vibrational stability. In

Figure 6.  $12.6 \times 7.9 \mu\text{m}$  image of an Al-Si alloy showing effects of scanner-driven vibrations in a large scan range ( $75 \mu\text{m}$ ) AFM. The vertical lines at left arise from vibrations generated when the scanner reverses its direction of movement. In most cases it is the scanner itself that is vibrating, since the scanner is usually the least stable part of large-range instruments. Vertical range 16 nm.



general, large range scanners have less stability than small range scanners and therefore yield lower resolution.

Scanner instabilities generally show up as evenly spaced vertical lines at one side of the image (figure 6). This pattern appears because the scanner “rings” after each change of scan direction.

Table 2 of the AFM design tables appendix presents the responsivity and resonant frequency of several commonly used scanners. AFM-3 uses a 0.5 in. diameter, 0.5 in. long EBL #1 tube ceramic with 0.020 in. thick walls. It achieves a scan range of  $1.8 \mu\text{m}$  with  $\pm 220 \text{ V}$  drive electronics and has a resonant frequency of 37 kHz. With a scanner this small, electronic rather than vibrational noise becomes the resolution-limiting factor. The drive electronics used with AFM-3 exhibit about 1 mV of noise, corresponding to 5 pm with the above tube ceramic.



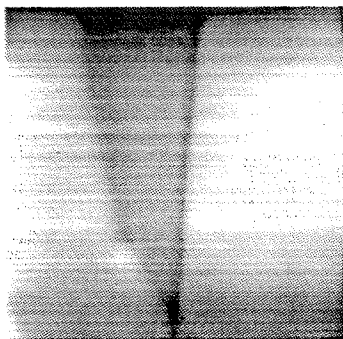


Figure 7. 340 × 340 nm image of boronated graphite showing low frequency noise (streaks) resulting from air current fluctuations. Vertical range 1 nm; scan rate 4 lines/second.

### Dense construction reduces air- and scanner- driven vibrations

So far I have considered vibrations driven by floor motions. Because of its relatively small mass, the stage does not affect these motions. It does, however, affect motions of the air (air currents and sound waves) as well as the scanner, two low-mass sources of vibration. Although high resonant frequency stages still resist these types of vibration, high density can also help. For example, a steel part resists sound waves better than an identical aluminum one, even though they have about the same resonant frequency. Because of its high surface area and high flexibility, the cantilever tends to exhibit particular sensitivity to air currents.

Air currents produce very low frequency noise ( $< 1$  Hz; see figure 7). This type of noise appearing to an unusual degree generally indicates some part of the instrument is loose. Most commonly, the user has not attached the sample firmly.

For high-resolution work I keep the AFM enclosed. Besides reducing air-driven vibrations, the enclosure shields the detector from noise-generating stray room light.

### Compact construction and special materials improve thermal stability

Thermal expansion and contraction causes the AFM to “drift” over time. A 10 mm long bar of stainless steel will expand by 150 nm when heated  $1^{\circ}$  C. Thermal drift distorts images and causes fluctuations in contact force.

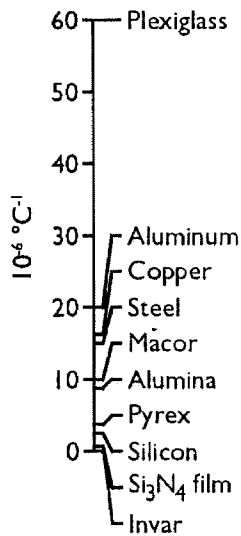


Figure 8. Coefficient of thermal expansion for several materials. From AFM design table I (appendix).

When heated an amount  $\Delta T$ , an object will expand by:

$$\Delta l = l\alpha\Delta T$$

$l$  = length of object

$\alpha$  = coefficient of thermal expansion

Therefore, using small components (small  $l$ ) results in low thermal drift. Since smaller elements also have higher resonant frequency, AFM stages generally benefit from compact design.

Special materials designed for thermal stability (low  $\alpha$ ) can also reduce thermal drift (figure 8). AFM designers favor Invar, a steel alloy. A 10 mm long bar of Invar expands only 1.3 nm when heated 1° C and thus affords a substantial advantage over steel. “Super-invar” alloys have even greater thermal stability but require re-annealing after machining.

Except for the top plate, made of nonmagnetic steel because of the nearby magnetic sample mount, AFM-3 consists of Invar.

## Scanner design

The position of a piezoceramic scanner depends on its recent history as well as on the voltage currently applied to it.

Piezoceramics exhibit three time-dependent nonlinearities: hysteresis, creep, and depoling.

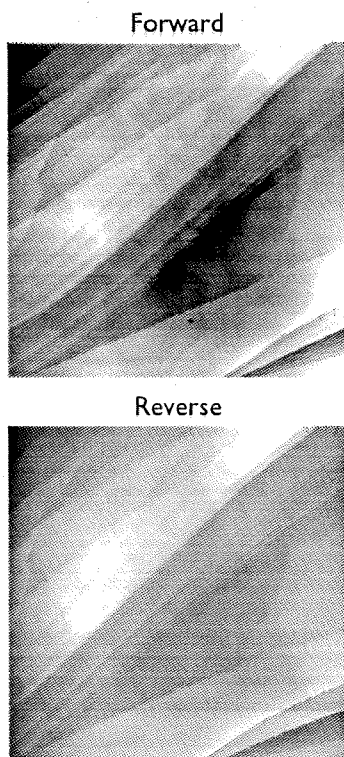


Figure 9. Simultaneous forward and reverse 1.62 x 1.62 nm scans of boronated pyrolytic graphite showing hysteresis effects. Note the curvature of the atomic steps, which should appear straight. Acquired with an EBL-2 (soft piezoceramic) scanner.

Hysteresis refers to the dependence of piezoceramic position on the most recent direction of movement. X hysteresis causes features to appear too far to the right in forward (left-to-right) scans and too far to the left in reverse scans (figure 9). The problem grows worse as image size increases.

“Creep” refers to the gradual settling that piezoceramics experience after the voltage applied to them changes. This phenomenon sometimes causes warping of the first few lines of an image while the scanner settles into its scanning pattern. It also means that the control system must keep the scan rate constant throughout the scan.

Finally, when exposed to high voltage of one polarity piezoceramics can become depoled, losing their ability to expand and contract. When this occurs the user generally has to repole the ceramic by applying high voltage of the opposite polarity (for AFM-3, positive voltage on the outside of the scanner).

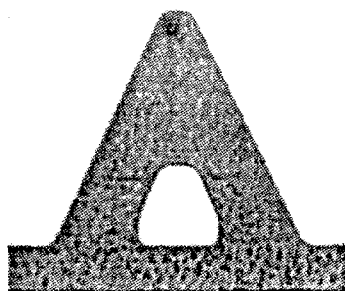
“Soft” piezoceramics exhibit all these nonlinearities in much more pronounced form than “hard” piezoceramics but also have a greater range. For AFM-3 I initially used a “soft” ceramic, EBL-2 from Stavely Sensors’ EBL division (East Hartford, CT), but later switched to the “hard” ceramic EBL-1. Both scanners measured 0.5” in diameter x 0.5” long x 0.020” thick. Over a  $\pm 220\text{V}$  voltage range the EBL-1 scanner has a 2.1  $\mu\text{m}$  XY range, while the EBL-2 had a 3.6  $\mu\text{m}$  range.

Barrett and Quate (1991) have developed a closed-loop XY positioning system to counter scanner nonlinearities. It measures scanner position and adjusts the applied voltage to yield a linear

scan. Ideally, this system should be extended to counter Z nonlinearities as well. I have not attempted to counter hysteresis in any way except by image processing (see chapter 4).

## Cantilever design

---



Spring constant	0.37 N/m
Res. frequency	66 kHz

Figure 10. An optical micrograph of a 22  $\mu\text{m}$  leg width, 100  $\mu\text{m}$  long, 0.6  $\mu\text{m}$  thick cantilever. This particular cantilever is an Ultralever, but normal cantilevers have the same shape and properties. Properties from AFM design table 6. (Cantilever from Park Scientific Instruments).

Engineering formulas describe relevant cantilever properties

As discussed in chapter 1, cantilevers should have low spring constant and high resonant frequency. Resonant frequency determines the maximum feedback bandwidth, which in turn limits the scan rate. Since stage stability generally limits feedback bandwidth to about 10 kHz, cantilever resonant frequency need not exceed this.

In AFM design tables 4 and 5 (appendix) I have provided formulas for resonant frequency and other relevant cantilever properties — angular deflection (see optical lever design, below) and the ratio of torsional and buckling to normal deflection per unit force (see chapter 5). Design table 6 presents results for the four Park Scientific Instruments cantilever models that I have used. Almost invariably I have used the 22  $\mu\text{m}$  wide, 100  $\mu\text{m}$  long model (figure 10).

Cantilevers usually have a V shape to provide lateral stability. Although V-shaped cantilevers provide less response to frictional forces, in my experience they yield significant improvement in image quality over I-shaped (“diving board”) cantilevers. I

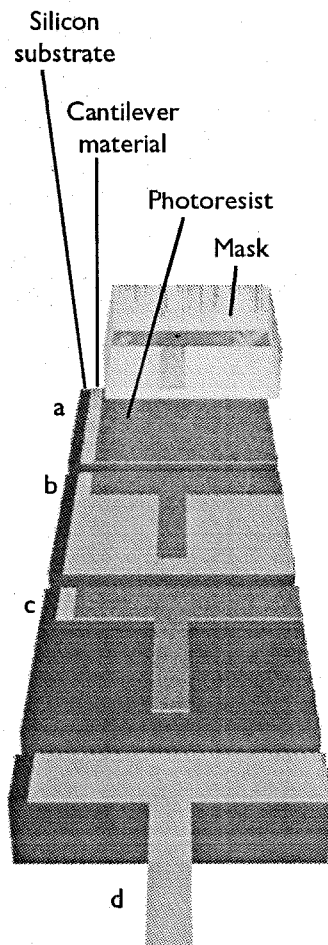


Figure 11. Steps of photolithography. (a) A layer of cantilever material and a layer of photoresist cover the silicon substrate. UV light passing through a photographic plate exposes certain parts of the photoresist. (b) After developing, only specific areas of the photoresist remain. (c) The photoresist protects underlying cantilever material during immersion in an etching solution. Uncovered areas etch away. (d) Another etching process removes excess silicon, leaving a free cantilever.

therefore prefer to use V-shaped cantilevers even for frictional force imaging.

### Most cantilevers are made by photolithography

Most manufacturers of cantilevers currently use photolithography (figure 11; Albrecht et al., 1990). This involves preparing a large-scale drawing of the cantilever and photographing it onto a fine-grain film plate to create a 1:1 scale “mask” having about 2  $\mu\text{m}$  resolution. Cantilevers with pyramidal tips consist of a 0.3–1.0  $\mu\text{m}$  thick layer of 70%  $\text{Si}_3\text{N}_4$  with 30%  $\text{SiO}_2$  (to reduce warping) prepared by chemical vapor deposition onto polished silicon {100}. Ultraviolet light passing through the mask exposes a photoresist layer deposited on top of this cantilever material. A developer dissolves exposed (or unexposed) portions of the photoresist, after which an acid bath can etch away the cantilever material under those portions. A second etching removes parts of the silicon substrate to create free cantilevers.

### Cantilever properties vary greatly due to imperfections

Micromachining techniques tend to produce inconsistent cantilever materials and dimensions, and even small imperfections in a cantilever alter its properties significantly. The spring con-

stant of a single model of cantilever, for example, can vary between 0.028 and 0.18 N/m (Cleveland et al., 1993).

Accurate force measurements require some way of calibrating the cantilever spring constant, such as that suggested by Cleveland et al. I have not done this for any of my experiments, which involve mostly qualitative or relative measurements.

### Thermal vibrations may limit the resolution of AFM

By equipartition, a rectangular cantilever at 25 °C and not in contact with the sample will experience random thermal movements with a root-mean-square amplitude of:

$$A = \frac{9.0}{\sqrt{k}}$$

where:

$A$  = vibration amplitude in pm

$k$  = spring constant of cantilever in N/m

A 0.1 N/m cantilever therefore undergoes thermal vibrations of 28 pm, comparable to AFM resolution. However, note that when the cantilever touches the sample the thermal vibrations become damped.

## Optical lever design

---

The theory of Gaussian beams predicts laser spot sizes

The optical lever design should focus the laser light to a spot having a diameter  $d_C$  small enough to fit on the cantilever. The reflected laser light should then form a spot at the detector having a diameter  $d_D$  about the same size as the detector. The theory of Gaussian beams provides formulas with which to calculate these spot sizes. Assuming a diode laser produces a collimated beam incident on a single focusing lens:

$$d_C = \frac{4\lambda}{\pi} \frac{I}{d_L}$$

$$d_D = d_L \frac{L}{I}$$

$I$  = lens-to-cantilever distance

$d_L$  = diameter of laser beam at lens

Therefore, a large beam diameter at the lens results in a small beam at the cantilever but a large beam at the detector. In AFM-3,  $d_L = 3$  mm,  $\lambda = 670$  nm,  $I = 4$  cm, and  $L = 4$  cm, so in theory  $d_C = 11$   $\mu$ m and  $d_D = 3$  mm.

Optical lever sensitivity depends on cantilever shape but not cantilever–detector distance

In optical lever detection, angular deflection  $\theta_l$  at the free end of the cantilever displaces the laser beam a distance  $V$  (at the detector) given by:

$$D = 2L\theta_l$$

$L$  = distance between cantilever and detector

AFM design table 3 provides formulas for  $\theta_l$  as a function of tip displacement  $v$ . Substituting into the above equation yields:

$$V = 3.5v \frac{L}{l} \text{ (rectangular cantilevers)}$$

$$V = 4v \frac{L}{l} \text{ (triangular cantilevers)}$$

In AFM-3  $L = 4 \text{ cm}$  and  $l = 100 \text{ }\mu\text{m}$ , and the V-shaped cantilevers used behave roughly like rectangular cantilevers, so  $V = 1400v$ . Shorter cantilevers and a triangular rather than rectangular shape would improve sensitivity.

Optical lever sensitivity does not depend on the cantilever–detector distance  $L$  because the laser beam expands with increasing distance from the cantilever. Inspecting the equation for detector output current (as discussed in chapter 1, this is the difference between the output of two adjacent photodetectors):



$$A - B = \frac{8}{\pi} \frac{V}{d_D} PR$$

$P$  = total wattage of beam incident on detector

$R$  = responsivity of detector

Substituting the equations for  $V$  and  $d_D$  yields the optical lever sensitivity:

$$\frac{A - B}{v} = \frac{28}{\pi} \frac{I}{ld_L} PR \text{ (rectangular cantilever)}$$

Therefore only the following changes can increase sensitivity:

- Longer lens-to-cantilever distance ( $l$ )
- Smaller beam diameter at lens ( $d_L$ ) or at detector (larger beam diameter at cantilever)
- Shorter cantilevers ( $l$ )
- More optical power at detector ( $P$ )
- A more responsive detector ( $R$ )

In practice the first measure may have little effect, since longer lens-to-cantilever distance also magnifies angular vibrations of the laser or focusing lens.

In AFM-3,  $P \sim 0.1$  mW (using a 1 mW diode laser) and  $R = 0.3$  A/W, yielding a theoretical optical lever sensitivity of 36 nA/nm. Typical experimental sensitivity is 50 nA/nm, with an

effective 10 pm noise level from the detector, amplifiers, and digitizing electronics combined.

## Alignment mechanisms

---

Tilt translators provide a reliable and stable alignment mechanism

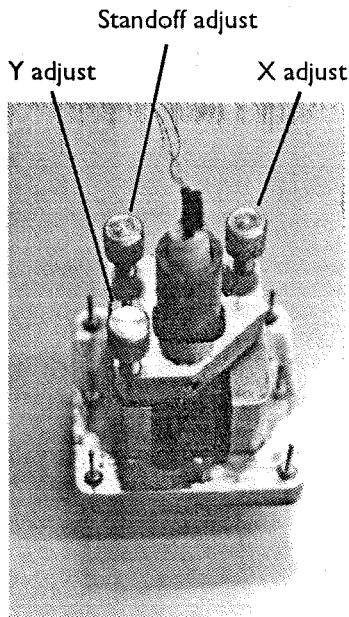


Figure 12. Bottom of AFM-3 head showing tilt translator with diode laser at center. The standoff adjust controls the overall height of the translator. Two springs hold the translator down with about 4 kg of force.

Optical lever AFMs require translators so the user can move the incident laser beam onto the cantilever and the reflected beam onto the detector. While commercially available translators have adequate precision, they lack the required stability, especially for the more demanding cantilever alignment.

Tilt translators, which work by rotating either a mirror or the source of the laser beam on a tripod made of three screws, provide an effective way to translate a laser beam. Their stability, compactness, and ease of construction has made them the most common translation device in AFM stages.

In AFM-3, a tilt translator holds the diode laser for cantilever alignment (figure 12). To prevent lateral play, a kinematic dimple-groove-flat system supports the tilting platform (figure 15). Each of the three support screws ends in a polished ball that rests on the body of the stage. One of the screws rests in a dimple, the second rests in a groove colinear with the dimple, and the third rests on a flat surface. This system tolerates small machining

Figure 13. The detector alignment system of AFM-3. A rotating mirror deflects the laser beam up and down, while the detector itself moves left and right.

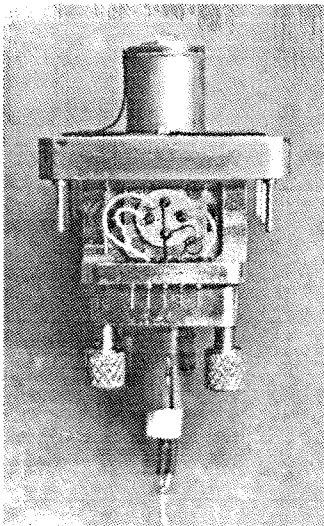
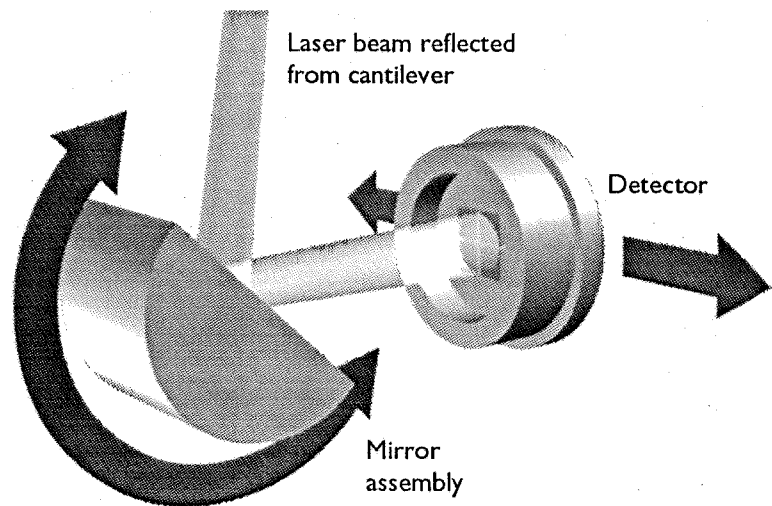


Figure 14. AFM-3 head showing the four-segment detector on its one-dimensional translator (center). A screw at left controls its position. A spring at right provides restoring force.

errors that might otherwise result in misalignment of the screws with the rests.

AFM-3 uses two one-dimensional translators to align the detector

Although a second tilt translator holding a mirror would work for aligning the laser beam to the detector, such an assembly would not fit into AFM-3. Therefore, I have used a rotating shaft holding a mirror for up-and-down alignment, and for left-and-right alignment placed the detector on a sliding puck (figures 13 and 14). Although this mechanism does not move as smoothly as a tilt translator, it suffices for detector alignment.

The optical lever should accommodate cantilevers warped up to  $\pm 5^\circ$ .

The optical lever should have the ability to accommodate laser beams as much as  $5^\circ$  off the theoretical path, first, because micromachined cantilevers exhibit varying amounts of warp, and second, because liquid cells can significantly alter the beam path. Through a combination of variable cantilever location, generous translator range, and a wide reflecting mirror, AFM-3 can accommodate cantilevers warped up to  $\pm 10^\circ$ . It would be acceptable to accommodate about half this range.

### *Tip-sample approach*

---

A motorized lever provides a reliable and practical way to bring the tip and sample into contact

The approach mechanism brings the tip and sample close enough for the scanner to take control of tip-sample separation. All scanned-probe microscopes have some type of approach mechanism, a device that usually dominates the stage design.

I have designed all of my AFM stages around a commercially available motorized screw (Series 860 motorizer, Newport Corporation, Fountain Valley, CA). This uses a DC motor — not a stepper motor. The screw moves a lever that serves to reduce the speed of the approach by a factor of about 10. Two

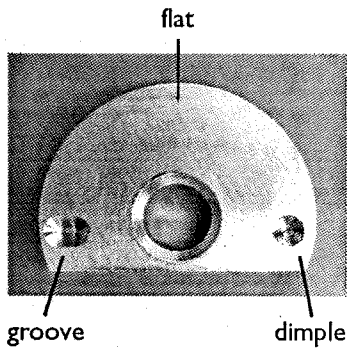


Figure 15. AFM-3 top plate showing dimple-groove-flat kinematic mount. The motorized screw contacts the flat, while the micrometer screws rest in the groove and dimple.

80 thread per inch (tpi) micrometer screws form the fulcrum of the lever and allow the user to adjust the sample tilt. The lever uses the same dimple-groove-flat support system used by the tilt translators (figure 15).

I have demonstrated the effectiveness of this mechanism by recording cantilever deflection during approaches as described in chapter 3. Tip and sample approach smoothly at a rate of about 1.5 nm/ms.

### The scanned-cantilever AFM

I have developed the scanned-cantilever architecture for its ease of use (Baselt and Baldeschwieler, 1993)

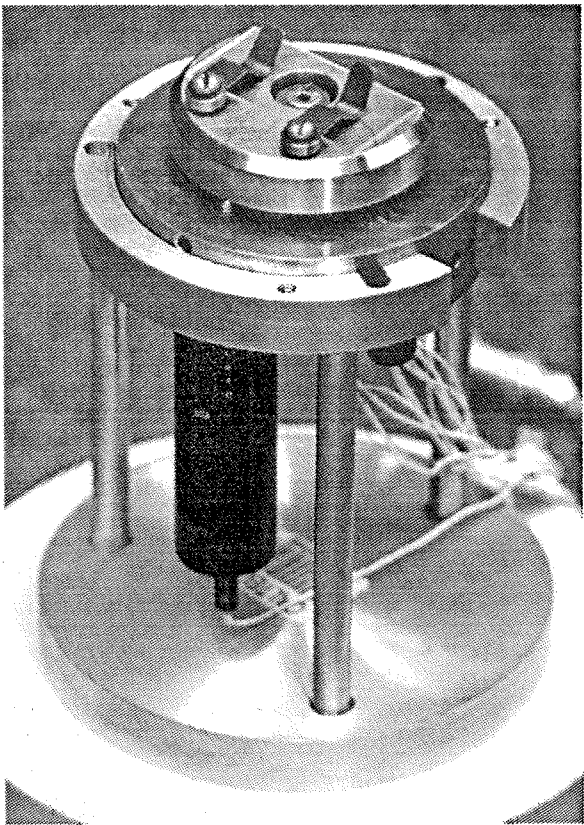
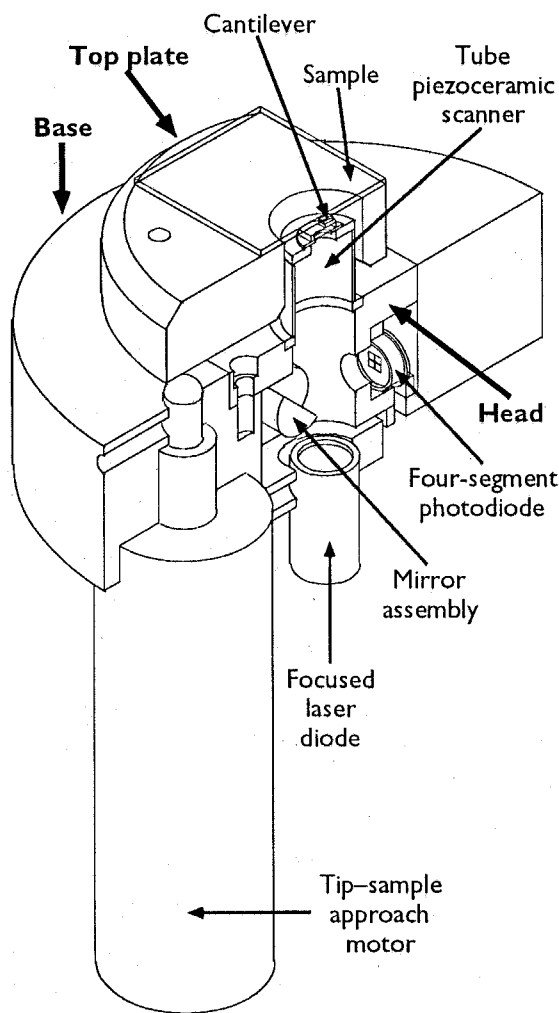
Most AFMs scan the sample in order to keep the cantilever stationary with respect to the detection apparatus. This configuration restricts the sample mounting area to the end of the scanner, a region at most half an inch in diameter, and massive samples degrade instrument performance by lowering the natural resonant frequency of the scanner. In addition, access to the sample for translation or replacement can be inconvenient.

I have developed a scanned-cantilever stage, AFM-3, that alleviates these problems. The instrument is easier to work with and more versatile than scanned-sample AFMs. Since the laser and detector remain stationary, the stage achieves high resolution but can only scan small areas.

For maximum accessibility, AFM-3 uses an inverted-sample configuration

AFM-3 (figure 16) positions the sample face-down at the top of the instrument to allow maximum accessibility. The cantilever mounts magnetically on the scanner. A focused diode laser (Hoetron, Inc., Sunnyvale, CA), mounted on the tiltable platform at the bottom of the instrument, emits a beam that shines up through the scanner to reflect off the cantilever. The laser beam

Figure 16. Schematic and photo of the scanned-cantilever AFM-3 stage. The stage is in the same orientation in both parts of the figure.



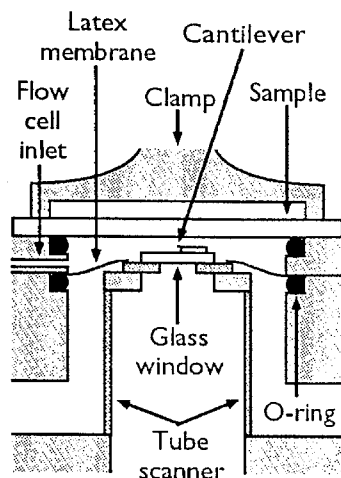


Figure 17. AFM-3 liquid flow cell. Liquid surrounds the cantilever and covers the sample; a glass window permits laser light to enter.

then reflects off a rotatable mirror assembly into a four-segment position-sensitive photodetector (SPOT-4D; United Detector Technology, Hawthorne, CA).

The sample mount accommodates a standard glass microscope slide. Since many of my samples are biomolecules deposited on glass, I can usually view the tip and sample (through the slide) with an optical microscope while scanning.

The inverted-sample configuration of this AFM complicates liquid flow cell design (note that the scanned-cantilever design does not require an inverted-sample configuration). After trying a number of flow cell designs I settled on one that uses two rubber O-rings and a latex glove rubber seal as illustrated in figure 17. The rubber seal has a hole in its center to admit laser light, and has sufficient elasticity to allow a tip-sample approach.

The scanned-cantilever design has the considerable advantage that the base contains all active parts of the instrument — scanner, optical lever, and tip-sample approach mechanism. The top plate, which holds the sample, has no wires or other attachments (unless the flow cell is in use) and thus can be removed and handled with ease.

The scanned-cantilever architecture gives rise to spurious signal variations

Scanning the cantilever while the optical-lever detection apparatus remains stationary gives rise to spurious variations of the

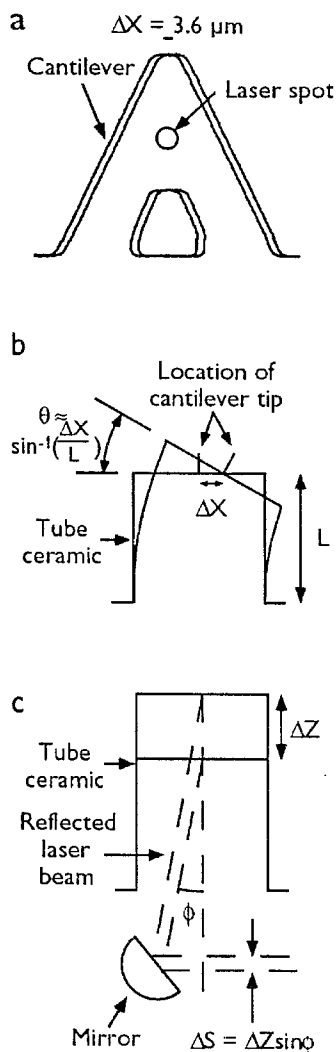


Figure 18. Sources of spurious signal variation. (a) irregularities in the mirror surface of the cantilever moving past the laser spot. (b) tilting of the tube ceramic during scanning. (c) Z motion of the tube ceramic.

cantilever deflection signal. If not corrected for, this could cause variations in the force exerted on the sample. I address here the theoretical effects of three possible sources of spurious signal variations.

As illustrated in figure 18a, the cantilever scans in X and Y (the directions parallel to the sample) while the beam reflected from it remains stationary; therefore, imperfections of the cantilever surface could cause fluctuations in the reflected laser beam as a scan progresses. These “imperfections” include the edge of the cantilever (that is, the cantilever could move out from under the beam), particles of dust, small surface imperfections (such as scratches), and overall warping of the cantilever. Since the cantilever measures about  $25 \mu\text{m}$  in width, the laser beam  $10 \mu\text{m}$  in diameter, and the maximum scan size  $2.1 \mu\text{m}$  on a side, with good alignment the cantilever should not move out from under the beam. The fact that the beam diameter exceeds the scan size also means that any spurious signal variations due to cantilever irregularities should have a first-order effect that can be removed with a first-order background subtraction, as described below.

Figure 18b diagrams the tilting of the scanner that accompanies XY motion. Since the optical lever cannot distinguish scanner tilt from cantilever deflection, a spurious signal variation results. Over a  $1 \mu\text{m}$  scan, the scanner tilts enough to produce a  $4 \text{ nm}$  signal variation. However, curvature of this magnitude appears even in scanned-sample AFMs, and thus I believe it is acceptable.

Finally, as illustrated in figure 18c, the laser beam reflected from the cantilever is displaced as the cantilever moves in Z (up



and down). However, the effect is very small, 1 Å of signal variation per  $\mu\text{m}$  of Z displacement.

In practice, spurious signal variations mostly arise from cantilever warp

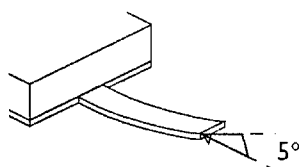


Figure 19. The most prominent type of cantilever warp bends the end of the cantilever up by about  $5^\circ$ .

In practice, spurious variations from cantilever warp dominate all other sources. By noting the location of the reflected laser beam in the AFM stage, I can estimate cantilever warp to about  $1^\circ$ . The commercial cantilevers I have used are generally warped  $3\text{--}5^\circ$  (figure 19).

In theory a cantilever warped by  $5^\circ$  should cause a spurious signal variation of 52 nm per  $\mu\text{m}$  of scan range in Y (the direction parallel to the long axis of the cantilever). The warp should not have any effect when moving the cantilever in the X or Z directions. In practice, the spurious signal variation in AFM-3 approximates this value: typically 60 nm per  $\mu\text{m}$  of Y displacement and 3.0 nm per  $\mu\text{m}$  of X displacement over a  $0.75\ \mu\text{m}$  displacement. The signal variation with Z displacement is unmeasurable. For comparison, a commercial scanned-sample instrument (Topometrix TMX 2000  $1\ \mu\text{m}$  AFM) demonstrated signal variations of 3.4 nm/ $\mu\text{m}$  in Y, 0.7 nm/ $\mu\text{m}$  in X, and 33 nm/ $\mu\text{m}$  in Z (possibly due to stray light reflecting from a metal backing on the sample). For these measurements the sample was a glass microscope slide.

## Background subtraction software can counteract spurious signal variations

I have implemented a first-order background-subtraction algorithm to counter the effect of spurious signal variations. Before approaching the tip and sample, the instrument-control software measures the deflection signal as a function of Y position and calculates a background, which it subsequently subtracts while scanning. This background subtraction reduces fluctuations in force resulting from spurious signal variations. The entire background measurement procedure takes only a second and does not detract from operation of the instrument. In fact, since all optical-lever instruments tend to exhibit spurious signal variations the background-subtraction procedure might be generally useful. Although I only calculate a first-order background, higher-order subtraction is certainly possible.

I have verified results from the scanned-cantilever AFM by repeating experiments performed on it with the Topometrix commercial scanned-sample instrument. Other than the above-mentioned signal variations, which are easily identified and removed, I have found no artifacts unique to the scanned-cantilever design.

## References

---

- Albrecht, T.R., Akamine, S., Carver, T.E., and Quate, C.F. (1990) Microfabrication of cantilever styli for the atomic force microscope. *J. Vac. Sci. Technol. A* 8(4), 3386–3396
- Barrett, R.C. and Quate, C.F. (1991) Optical scan-correction system applied to atomic force microscopy. *Rev. Sci. Instrum.* 62(6), 1393–1399
- Baselt, D.R. and Baldeschwieler, J.D. (1993) Scanned-cantilever atomic force microscope. *Rev. Sci. Instrum.* 64(4), 908–911
- Blevins, R.D. (1979) *Formulas for Natural Frequency and Mode Shape*. Robert E. Krieger Publishing Company, Malabar, Florida
- Cleveland, J.P., Manne, S., Bocck, D., and Hansma, P.K. (1993) A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Rev. Sci. Instrum.* 64(2), 403–405
- Den Hartog, J.P. (1934) *Mechanical Vibrations*. Dover Publications, Inc., New York
- Park, S. and Quate, C.F. (1987) Theories of the feedback and vibration isolation systems for the scanning tunneling microscope. *Rev. Sci. Instrum.* 58(11), 2004
- Rothbart, H.A. (1988) *Mechanical Engineering Essentials Reference Guide*. McGraw-Hill Book Company, New York
- Young, W.C. (1989) *Roark's Formulas for Stress and Strain*. McGraw-Hill Book Company, New York

## **AFM design tables**

revision 5  
David Baselt

Table 1. Properties of various engineering materials

Table 2. Properties of piezoceramics

Table 3. Calculated characteristics of nine scanning tubes

Table 4. Equations describing cantilever beams

Table 5. Responsivity of optical lever detection to cantilever bending modes

Table 6. Calculated characteristics of the standard PSI cantilevers

Stability criteria for three common component geometries

**Table 1.**  
**Properties of various engineering materials<sup>(1)</sup>**

mks units cgs units mm, mg, s	E 10 <sup>9</sup> Pa 10 <sup>10</sup> dyn/cm <sup>2</sup> 10 <sup>12</sup> mPa	ρ 10 <sup>3</sup> kg/m <sup>3</sup> g/cm <sup>3</sup> mg/mm <sup>3</sup>	$\sqrt{\frac{E}{\rho}}$ 10 <sup>3</sup> m/s 10 <sup>5</sup> cm/s 10 <sup>6</sup> mm/s	α <sup>(2)</sup> 10 <sup>-6</sup> °C <sup>-1</sup>	Q <sup>(3)</sup>
<b>Metals</b>					
Stainless steel type 316	193	7.92	4.94	14.7	
Invar <sup>(4)</sup>	150	8.00	4.33	0.13	
Aluminum alloy 2024-T4	72	2.71	5.15	21.6	metals: ~1000
Tungsten	407	19.25	4.60	4.5	
Copper	110	8.94	3.51	16.5	
<b>Nonmetals</b>					
Silicon	190	2.30	9.09	2.6	
SiO <sub>2</sub> bulk	76	2.20	5.90	0.49	
thin film	67	2.25	5.50		
Si <sub>3</sub> N <sub>4</sub> bulk	241	3.20	8.68	0.8	ceramics: ~10000
thin film	150	3.10	7.00		
Al <sub>2</sub> O <sub>3</sub>	366	3.97	9.60	8.8	
Macor <sup>(7)</sup>	65.4	2.52	5.09	9.4	
<b>Plastics</b>					
Teflon	4.8	2.17	1.50	100 <sup>(8)</sup>	
Nylon	2.8	1.14	1.60	80	plastics: ~10
Plexiglass	2.7	1.17	1.50	61	
<b>Other</b>					
Pyrex	69	2.20	5.60	3	500
Concrete <sup>(9)</sup>	28	2.60	3.30	0.3	100
Wood (fir)	9.7	0.304	5.60	6, 35 <sup>(10)</sup>	100
Wood (oak)	16	0.720	4.70	5, 55 <sup>(10)</sup>	100

1. E: Young's modulus;  $\rho$ : density;  $\alpha$ : coefficient of thermal expansion; Q: quality (damping) factor. Unless otherwise noted, data from Callister, W.D., *Materials Science and Engineering*, pp. 554-557, John Wiley and Sons Inc., New York, 1985 and/or J.H. Moore et al., *Building Scientific Apparatus*, p. 1.1, Addison-Wesley Publishing Company, London, 1983.
2. Thermal expansion coefficient at 20°C; values from Y.S. Touloukian et al., *Thermophysical Properties of Matter* volumes 12-13, IFI/ Plenum, New York, 1975.
3. B.J. Lazan, *Damping of Materials and Members in Structural Mechanics*, pp. 204-252, Pergamon Press, Oxford, 1968.
4. Except for  $\alpha$  and Q, values from ASM Handbook Committee, *Metals Handbook* volume 3, Ninth Edition, pp. 792-798, American Society for Metals, Metals Park, Ohio, 1980.
5. SiO<sub>2</sub> preparation, thermal, dry; Si<sub>3</sub>N<sub>4</sub>, CVD. K.E. Petersen, C.R. Guarnieri, "Young's modulus measurements of thin films using micromechanics," *J. Appl. Phys.* **50**(11), 6761 (1979).
6. C.T. Lynch, *Handbook of Materials Science* volume 2, CRC Press, Boca Raton, 1974.
7. Information and product from Corning Glass Works, Corning, New York; 1988.
8. Teflon undergoes a phase change at room temperature, resulting in widely varying values of the expansion coefficient. Value listed is from ref. 1; ref. 2 gives a value of 525.
9. W.S. La Londe, *Concrete Engineering Handbook*, pp. 1-49 – 1-50, McGraw-Hill Book Company, Inc., New York, 1961.
10. Parallel with and perpendicular to the grain, respectively.

**Table 2.**  
**Properties of piezoceramics<sup>(1)</sup>**

cgs units mm, mg, s	$E^{(2)}$ $10^{10}$ dyn/cm <sup>2</sup> $10^{12}$ mPa	$\rho$ g/cm <sup>3</sup> mg/mm <sup>3</sup>	$\sqrt{E/\rho}$ $10^5$ cm/s $10^6$ mm/s	$\alpha^{(3)}$ $10^{-6}$ °C <sup>-1</sup>	Q	d Å/V
PZT-5A, 3 <sup>(4)</sup>	53.2	7.75	2.62	2.0	75	3.74
PZT-5A, 1	61.0		2.80	1.5		-1.71
PZT-5H, 3	48.3	7.50	2.54	-3.9 <sup>(5)</sup>	65	5.93
PZT-5H, 1	60.6		2.84	7.8 <sup>(5)</sup>		-2.74

1. Unless otherwise noted, values are from "Piezoelectric Technology Data for Designers," Vernitron Piezoelectric division.
2. At constant electric field (as opposed to constant charge density).
3. At 25°C. These coefficients change if the piezoceramic is heated above 50°C; for PZT-5A the new values are 1, 2, 4.
4. Direction in which the property is measured; 3 refers to the axis of poling; 1 to any direction perpendicular to that axis.
5. A.M. Simpson, W. Wolfs, "Thermal expansion and piezoelectric response of PZT Channel 5800 for use in low-temperature scanning tunneling microscope designs," *Rev. Sci. Instrum.* **58**(11), 2193 (1987).

**Table 3.**  
**Calculated characteristics of nine scanning tubes**

diam. in.	length in.	thick. in.	m <sup>(1)</sup> mg	material <sup>(2)</sup>	dx/dV <sup>(3)</sup> Å/V	dz/dV <sup>(4)</sup> Å/V	f <sub>tr</sub> <sup>(5)</sup> kHz	f <sub>tor</sub> <sup>(6)</sup> kHz	f <sub>lon</sub> <sup>(5)</sup> kHz	t <sub>opt</sub> <sup>(7)</sup> mm
0.125	0.25	0.015	10	PZT-5A	23	29	34	67	104	0.333
				PZT-5H	36	46	32	68	105	0.338
0.25	0.5	0.020	40	PZT-5A	36	43	18	34	53	0.490
				PZT-5H	57 <sup>(8)</sup>	69 <sup>(8)</sup>	17 <sup>(8)</sup>	34 <sup>(8)</sup>	54 <sup>(8)</sup>	0.498
0.5	0.5	0.020	160	PZT-5A	24	43	37	33	51	0.979
				PZT-5H	38	69	38	33	51	0.994
0.25	1.0	0.020	40	PZT-5H	230	137	5.0	17	27	0.361
0.25	1.5	0.022	40	PZT-5H	531	187	2.3	12	18	0.298
0.5	1.5	0.030	160	PZT-5H	240	137	4.6	12	18	0.596

1. Mass of sample at end of tube; assumed disk of constant thickness. Corresponds to a 0.010 in. thick macor disk.
2. Piezoceramic properties as in table 1b.
3. V is the voltage difference between inner and outer electrodes. Assumes single-quadrant activation only. PZT-5H numbers calculated from the following equations in R.G. Carr, "Finite Element Analysis of PZT Tube Scanner Motion for Scanning Tunneling Microscopy," *J. Microsc.* **152**, 803 (1988):

$$\frac{dx}{dV} = 0.677 \frac{l^{2.066}}{d^{0.767} t^{1.086}}$$

$$\frac{dz}{dV} = 2.36 \frac{l^{1.005} d^{0.029}}{t^{1.01}}$$

where d = diameter, l = length, t = thickness. 0.25 x 0.5 in. tube from finite element analysis in same paper; PZT-5A numbers approximated by multiplying 5H numbers by ratio of 5A to 5H d<sub>31</sub>.

4. Calculated as d<sub>31</sub>(l/t).
5. Resonant frequencies calculated from the following formulas derived from R.D. Blevins, *Formulas for Natural Frequency and Mode Shape*, pp. 185 and 195, Robert E. Krieger Publishing Company, Malabar, Florida, 1979.

Transverse:

$$f_{tr} = 0.2461 \sqrt{\frac{E}{l^3} \frac{t(d-t) \left( \frac{d^2}{2} - dt + t^2 \right)}{m + 0.7125t(d-t)\rho}}$$

Longitudinal:

$$f_{\text{lon}} = \frac{\lambda}{2\pi l} \sqrt{\frac{E}{\rho}}$$

$$\lambda \tan \lambda - \frac{\pi \rho l t (d - t)}{b} = 0$$

Torsional:

$$f_{\text{tor}} = \frac{\lambda}{2\pi l} \sqrt{\frac{E}{2\rho(1 + \nu)}}$$

$$\lambda \tan \lambda - \frac{4\pi \rho l t (d - t) \left( \frac{d^2}{2} - dt + t^2 \right)}{md^2} = 0$$

d = diameter, l = length, t = thickness, m = end mass, E = Young's modulus of elasticity,  $\rho$  = density,  $\nu$  = Poisson's ratio.

6. Calculated as per footnote 5. The torsional frequency cannot be calculated accurately due to lack of published values of Poisson's ratio  $\nu$  for piezoceramics. Assumed value 0.3.
7. Thickness of tube wall that provides the greatest  $f_{\text{tr}}$ .
8. Experimental values are 50, 50, 8, -, 40; G. Binnig, D.P.E. Smith, *Rev. Sci. Instrum.* **57**(8), 1986.



**Table 4.**  
**Equations describing cantilever beams<sup>(1)</sup>**

	Trapezoid	Rectangle	Triangle
End deflection $V_L$	$\frac{6PL^3}{Eb_0h^3} \frac{2}{(1-n)^2} \left( \frac{n^2}{1-n} \ln\left(\frac{1}{n}\right) + \frac{1-3n}{2} \right)$	$\frac{4PL^3}{Eb h^3}$	$\frac{6PL^3}{Eb_0h^3}$
Resonant frequency $f_1$	$0.1634n^{-0.2689} \frac{h}{L^2} \sqrt{\frac{E}{\rho}}$	$0.1683 \frac{h}{L^2} \sqrt{\frac{E}{\rho}}^{(2)}$	$0.34 \frac{h}{L^2} \sqrt{\frac{E}{\rho}}$
End slope $\theta_L$	$\frac{12PL^2}{Eb_0h^3} \frac{1}{1-n} \left( \frac{n \cdot \ln n}{1-n} + 1 \right)$	$\frac{6PL^2}{Eb h^3}$	$\frac{12PL^2}{Eb_0h^3}$
Deflection equation $V_x$	See footnote 3	$\frac{6Px^2 \left( L - \frac{1}{3}x \right)}{Eb h^3}$	$\frac{6PLx^2}{Eb_0h^3}$
Slope equation $\theta_x$	$\frac{12PL^2}{Eb_0h^3} \frac{1}{1-n} \left( \frac{n \cdot \ln \left( 1 - \frac{x(1-n)}{L} \right)}{1-n} \right) + \frac{x}{L}$	$\frac{12Px \left( L - \frac{1}{2}x \right)}{Eb h^3}$	$\frac{12PLx}{Eb_0h^3}$

1.  $P$  = transverse force acting on end of cantilever;  $L$  = length;  $b_0$  = width at base;  $h$  = thickness;  $n$  = ratio of tip to base width (0 for triangle, 1 for rectangle);  $E$  = modulus of elasticity;  $x$  = longitudinal position (0 at base,  $L$  at tip). All calculations are for the transverse direction perpendicular to the plane of the cantilever.

Slopes and deflections calculated by integration of bending moment equation; see for example Gere and Timoshenko, *Mechanics of Materials*, pp. 351-361, PWS Engineering, Boston, 1984. Trapezoid frequency obtained by curve fitting calculated data from H.H. Mabie, C.B. Rogers, "Transverse Vibrations of Double-Tapered Cantilever Beams," *J. Acoust. Soc. Am.* **51**(5), 1771 (1972). Other frequencies obtained from R.D. Blevins, *Formulas for Natural Frequency and Mode Shape*, Robert E. Krieger Publishing Company, Malabar, Florida, 1979.

2. Using the energy method (see Gere and Timoshenko), it is possible to calculate the resonant frequency of a rectangular cantilever with a point mass at its end:

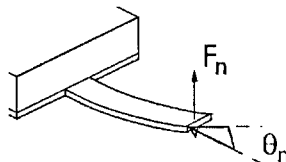
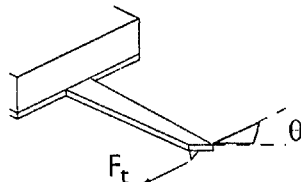
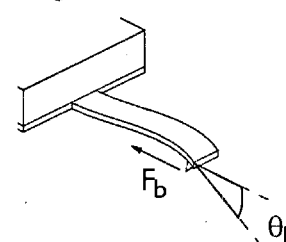
$$f = 0.08016 \sqrt{\frac{Eb h^3}{L^3} \frac{1}{m + 0.2268M}}$$

Where  $m$  = end mass and  $M$  = mass of cantilever.

- 3.

$$\frac{6PL^3}{Eb_0h^3} \frac{2n}{(1-n)^2} \left( \left( \frac{x}{L} - \frac{1}{1-n} \right) \ln \left( 1 - \frac{x}{L} (1-n) \right) - \frac{x}{L} + \left( \frac{x}{L} \right)^2 \frac{1-n}{2n} \right)$$

**Table 5.**  
**Responsivity of optical lever detection to cantilever bending modes**

Type of deflection	Responsivity
<p>Normal</p> 	$R_n = \frac{6l^2}{Ewt^3}$
<p>Torsional</p> 	$R_t = \frac{2al(l + \nu)}{KE}$ $K = wt^3 \left( 0.33 - \frac{0.2lt}{w} \left( 1 - \frac{t^4}{12w^4} \right) \right)$
<p>Buckling</p> 	$R_b = \frac{3al}{Ewt^3}$

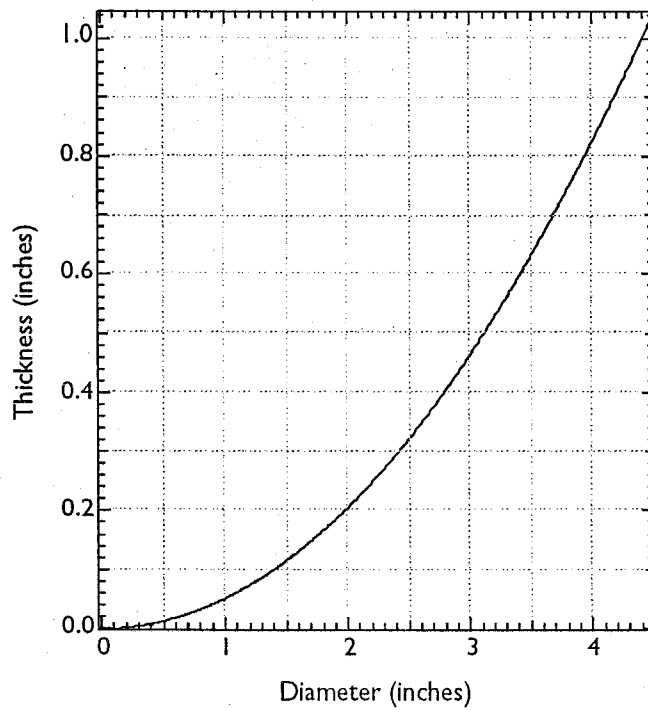
"Responsivity" defined as the rotational deflection  $q$  per unit force  $F$  acting as illustrated;  $a$ =tip height;  $l$ =cantilever length;  $w$ =width;  $t$ =thickness;  $E$ =modulus of elasticity;  $\nu$ =Poisson's ratio. Formulas derived from W.C. Young, *Roark's Formulas For Stress and Strain*, McGraw-Hill Book Company, New York. Normal: table 3-1a; torsional: table 20-4; buckling: table 3-3a.

**Table 6.**  
**Calculated characteristics of the standard PSI cantilevers**

	$k_n$ (N/m)	$R_n$ (rad/N)	$R_n/R_t$	$R_b/R_t$	$f_{tra}$ (kHz)
<b>Length 100 <math>\mu\text{m}</math></b>					
Wide V	0.37	30,900	30.7	0.463	66
Narrow V	0.21	52,900	36.0	0.540	66
Wide I	0.17	92,600	23.6	0.353	66
Narrow I	0.08	185,000	23.1	0.346	66
<b>Length 200 <math>\mu\text{m}</math></b>					
Wide V	0.064	93,400	70.9	0.498	17
Narrow V	0.032	168,000	81.8	0.574	17

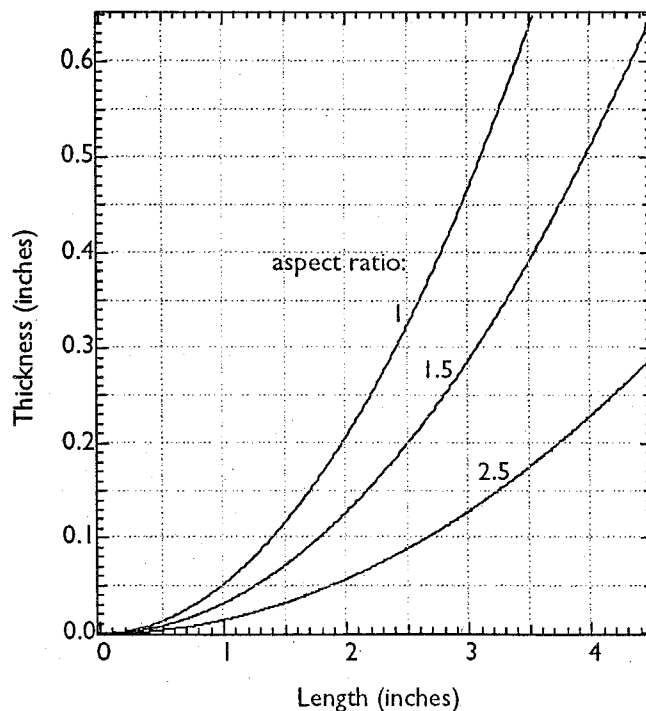
Calculated responsivity of the four standard silicon oxynitride V-shaped and the two obsolete I-shaped Park Scientific cantilevers. Assumed thickness 0.6  $\mu\text{m}$ , tip height 3  $\mu\text{m}$ , Poisson's ratio 0.33, and modulus of elasticity 150 GPa.  $k_n$  and  $f_{tra}$  from PSI Microlever sales literature (Park Scientific Instruments, Sunnyvale, CA). Other values calculated using the formulas in table 5. Torsional responsivity  $R_t$  estimated by averaging values calculated for a rectangle as wide as one leg of the cantilever and a rectangle as wide as the base of the cantilever.

### Thickness vs. diameter of steel disks having 10 kHz resonant frequency



Based on the formula  $h = 0.0517d^2$ , where  $h$  = thickness in inches and  $d$  = diameter in inches.  
Valid for steel disks simply supported at their circumference.  
For disks clamped at their circumference, multiply the value of  $h$  determined from this plot by 0.5.

### Thickness vs. length of rectangular plates having 10 kHz resonant frequency



Based on the formulas:

$$h = 0.0521a^2 \text{ (aspect ratio 1)}$$

$$h = 0.0321a^2 \text{ (aspect ratio 1.5)}$$

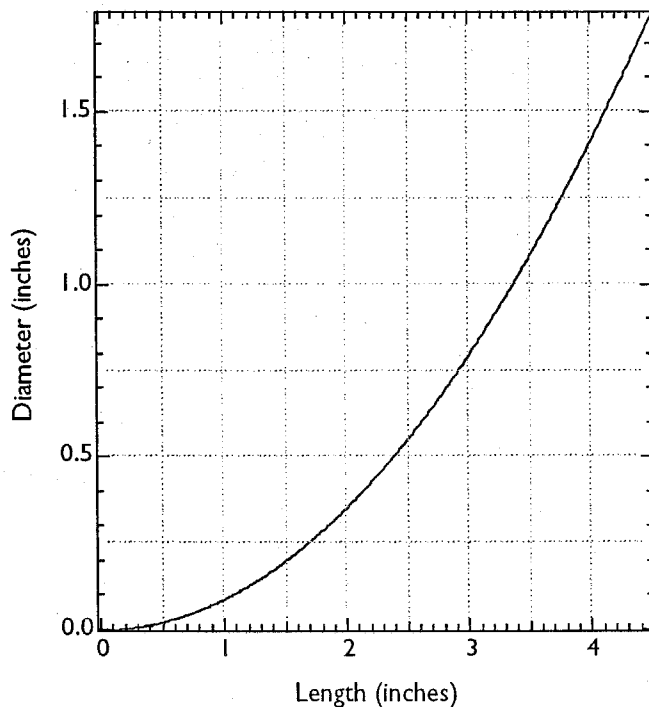
$$h = 0.0144a^2 \text{ (aspect ratio 2.5)}$$

where  $h$  = thickness in inches and  $a$  = length in inches.

Valid for rectangular plates simply supported at their periphery.

Aspect ratio refers to the ratio length/width.

### Diameter vs. length of steel rods having 10 kHz resonant frequency



Based on the formula  $d = 0.177l^2$ , where  $d$  = diameter in inches and  $l$  = length in inches.

Valid for cantilevered steel rods.

For rods free or clamped at both ends, multiply the value of  $d$  obtained from this plot by 0.157.

For rods hinged at both ends, multiply the value of  $d$  obtained from this plot by 0.441.

For materials other than steel, multiply the value obtained from the above plots by the following constant:

Alumina	0.57
Pyrex	0.88
Aluminum	0.96
Macor	0.97
Invar	1.1
Copper	1.4
Plexiglas	3.3

For resonant frequencies other than 10 kHz, multiply the thickness or diameter obtained from the above plots by the desired resonant frequency divided by 10 kHz.

Except for the materials information above, which is derived from table I, all information is from H.A. Rothbart, *Mechanical Engineering Essentials Reference Guide*, pages 5.65, 5.62, and 5.57, Mc-Graw-Hill Book Company, New York, 1988.

## Chapter 3. The control system

### *Introduction*

---

Chapter 3 discusses the digital signal processor-based control system of AFM-3 (Baselt et al., 1993)

The sensor that measures some local property of a sample, and the scanner that positions the sensor in three dimensions comprise a scanned-probe microscope (SPM) stage. The control system, or electronics, records the voltages that the sensor produces and applies the voltages that move the scanner.

AFM-1 and AFM-2 used commercial electronics that had a different circuit for each of its functions. The lack of versatility inherent in this architecture prompted me to construct the system described in this chapter.

I have used the digital signal processor (DSP)-based control system described here for atomic force microscopy and scanning tunneling microscopy; others have used it for nearfield scanning optical microscopy and low temperature tunneling spectroscopy. The system should also work for most other scanned-probe experiments.

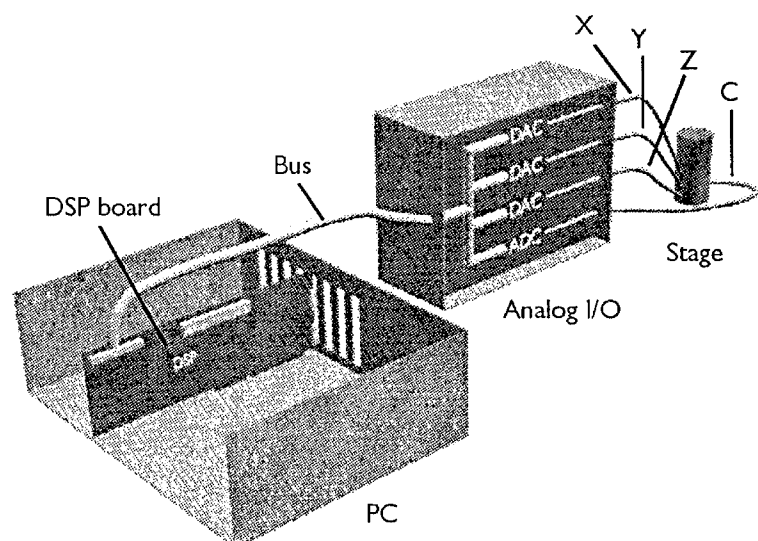
I have programmed the data acquisition package, SCAN. Topometrix, a commercial vendor of SPMs, has constructed the analog input/output (I/O; see below) based on my general specifications.

DSP-based control combines the versatility of digital systems with performance until recently available only in analog systems

In the DSP-based control system (figure 1), a microprocessor — the DSP — controls via digital-to-analog converters (DACs) the voltages applied to the SPM stage, and monitors via analog-to-digital converters (ADCs) the sensor output voltages. The software on the DSP carries out all control functions.

DSPs have two important advantages over other microprocessors: 1) they are designed for real-time processing of audio-

Figure 1. DSP-based SPM hardware. The DSP board plugs into a PC and communicates with an analog I/O box that resides outside the PC. For reasons of simplicity I have omitted the signal and high voltage amplifiers between the analog I/O and the stage. I have also omitted several ADCs and DACs; the electronics actually have three ADCs and eight DACs.



frequency signals similar to those used in SPM; and 2) they can be bought in convenient board-level products and programmed using high-level languages such as C. The DSP board plugs into a host personal computer (PC) and serves as a self-contained computer dedicated to SPM control.

As a result of the digital audio revolution, DSPs and associated analog I/O hardware have become inexpensive and widely available. Most commercial manufacturers of scanned-probe microscopes use DSP-based control systems to drive their instruments, since such control systems provide a combination of versatility and performance not readily available from other system architectures.

Using this technology I have developed a number of simple control algorithms that emulate analog circuits for feedback, raster generation, tip-sample approach, and modulation imaging. In addition, the control system has a programmable user interface. The speed of DSPs has eliminated the need for specially devised digital algorithms, while the straightforward nature of the algorithms, and their encoding in software rather than hardware, makes modification easy.



## Hardware components

---

Table 1. Spectrum Signal Processing, Inc. TMS320C30 processor board

CPU	32 bit floating point
Clock rate	33 MHz
Instruction cycle	60 ns
Memory (in 32-bit words):	
local (0 wait state)	0.5 Mw
dual port (3 wait state)	64 kw
Expansion bus (to analog I/O):	
data size	16-bit words
data rate	5 Mw/s

Table 2. Topometrix, Inc. TMX 2000 electronic control unit

Three 14-bit ADCs (feedback sensor, friction, auxiliary) with  $\pm 5$  V range and 1 mV peak-peak noise level  
 Five 16-bit DACs (+X, -X, +Y, -Y, Z) with  $\pm 220$  V amplifiers  
 One 12-bit motor voltage DAC with  $\pm 5$  V range and current amplifier  
 Two 16-bit DACs (STM tip voltage, auxiliary) with  $\pm 5$  V range

DSP-based control systems have two main components, the DSP board and its analog I/O

Commercial vendors now offer a wide range of DSP boards suitable for SPM control. AFM-3 uses the TMS320C30 processor board from Spectrum Signal Processing, Inc. (Table 1; Burnaby, BC, Canada), based on the Texas Instruments (Houston, TX) DSP of the same name. This board has three features that greatly simplify programming: its floating-point architecture simplifies development of control algorithms, while its dual-port memory simplifies PC-DSP communications and its on-board memory suffices to store several images if acquisition rates exceed the maximum PC-DSP data transfer rate.

Analog I/O consists of analog-to-digital converters (ADCs) and digital-to-analog converters (DACs), the latter usually having  $\pm 220$  V high-voltage amplifiers. Since many commercial I/O products are intended for prototyping purposes only and have noise levels inadequate for SPM, AFM-3 uses custom-built external I/O (Table 2; Topometrix, Inc., Santa Clara, CA).

### *PC-DSP communication*

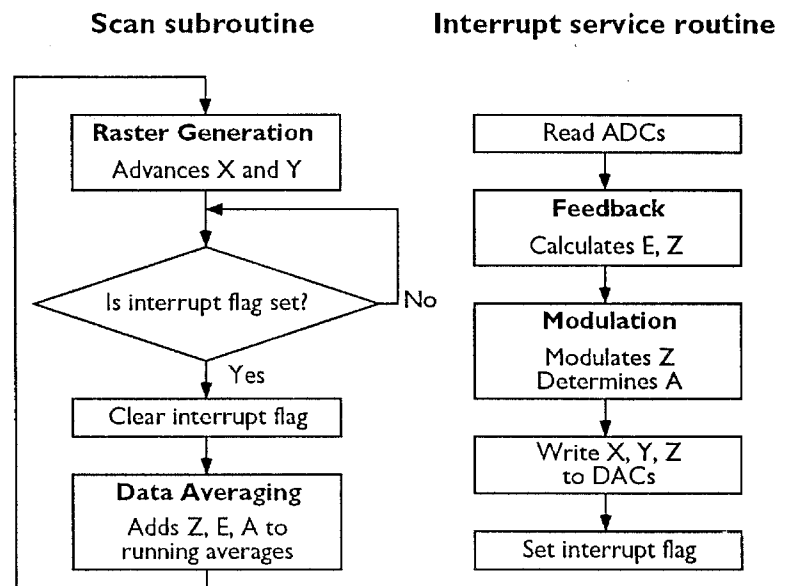
---

The PC and DSP communicate via the DSP's dual port memory

The DSP-based architecture lends itself to a division of tasks in which the DSP handles instrument control functions and the PC handles the user interface. The two microprocessors communicate via the DSP's dual port memory. The PC stores variables such as scan rate, feedback parameters, and scanner location in the dual port memory, while the DSP stores data such as the current scanner position, sensor output, and error messages. A shared memory location serves as an instruction mailbox. When the user launches the data acquisition program, the PC loads the DSP program onto the DSP. If the user then issues an instruction to take a scan, the PC places the "scan" instruction into the DSP's instruction mailbox. The DSP takes a scan and, when completed, places the "ready" instruction in the mailbox. Meanwhile, the PC displays the incoming image, using data placed in the dual port memory by the DSP.

A few notes on terminology: Z refers to the voltage that moves the probe toward or away from the sample, X to the fast raster voltage, and Y to the slow raster voltage. The "sensor output" C is an ADC reading proportional to the probe-sample separation (i.e., tunneling current in STM, cantilever deflection in AFM).

Figure 2. Software algorithms used during one timer period while scanning. Every  $10\ \mu\text{s}$  the DSP timer triggers an interrupt, causing an automatic jump to the interrupt service routine. This jump should occur during the waiting loop shown. After the DSP completes the interrupt service routine, it resumes the scan subroutine. Not shown: data storage (occurs less than once per timer period), tip-sample approach (not used while scanning), and user interface (independent of timer period). E: error signal. X, Y, Z: scanner voltages. A: modulation amplitude.



### Timing considerations

All functions of the software operate within a  $10\ \mu\text{s}$  interrupt cycle

Most SPM control algorithms require consistent and predictable timing to function properly. The characteristics of a digital feedback loop, for example, vary with its sampling rate. Likewise, an SPM needs a known and constant scan rate in order to predict and control effects of thermal drift, feedback characteristics, and scanner creep. In AFM-3, a timer built-in to the DSP ensures consistent timing of the control algorithms.

The timer period, usually  $10\ \mu\text{s}$ , serves as the fundamental clock of the SPM. Once per period, the timer triggers an interrupt that causes the DSP to advance the XY raster, read the

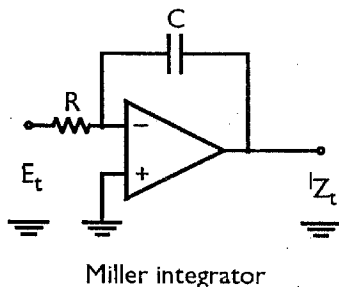
ADCs, perform the feedback loop calculations, and store image data (figure 2).

The timer period limits the number of features that the DSP can handle simultaneously. A 10  $\mu$ s period corresponds to 166 instruction cycles on our DSP and suffices for most experiments; complex experiments require more time. However, the shorter period allows faster scans and more oversampling on slow scans. Also, the optimum response rate of the feedback (its ability to follow features on the sample surface) depends in part on the timer period. Ultimately, the ADC conversion time of 10  $\mu$ s (for 16 bit ADCs) places a lower bound on the timer period.

Table 3. Proportional, trapezoidal-approximation integral, derivative, and PID feedback algorithms. See text for meaning of variables.

Proportional	$PZ_t = PE_t$
Integral	$Z_t = IZ_{t-1} + I(E_t + E_{t-1})$
Derivative	$DZ_t = D(E_t - E_{t-1})$
PID	$Z_t = PZ_t + IZ_t + DZ_t$

The integral feedback algorithm emulates a Miller integrator (below) with time constant  $RC = 0.5t/I$ , where  $t$  is the interrupt period and  $I$  is the integral feedback gain. For AFM-3  $I$  is typically 0.002, giving  $RC = 2.5$  ms. The actual feedback response rate also depends on various gain factors inherent in the AFM.



### Self-optimizing PID feedback

AFM-3 uses PID feedback

By continually adjusting  $Z$ , the feedback loop attempts to maintain the sensor output voltage at a constant level (the “setpoint”) specified by the user. The digital feedback I have implemented surpasses its analog prototype by supporting a wide range of feedback parameters, simplifying implementation of features such as spectroscopy and digital modulation, and permitting the self-regulating feedback described below.

AFM-3 uses a proportional-integral-derivative (PID) feedback loop, a commonly used formula that sums the outputs of three separate algorithms (table 3; Bucek, 1989). Proportional feedback

responds quickly to small features but cannot maintain a DC offset. Integral feedback maintains a DC offset but cannot respond to small features without oscillating. Derivative feedback reduces overshoot and oscillations, but amplifies high frequency (about half the sampling frequency) noise.

An operational formula for PID feedback follows:

$$Z_t = Z_{t-1} + aE_t + bE_{t-1} + cE_{t-2}$$

where:

$$E_t = C_t - S$$

and:

$t$  = time in timer periods

$Z_t$  =  $Z$  at time  $t$

$E_t$  = error signal at time  $t$

$C_t$  = ADC reading at time  $t$

$S$  = setpoint

$a = P+I+D$

$b = -P+I-2D$

$c = D$

$P$  = proportional feedback gain

$I$  = integral feedback gain

$D$  = derivative feedback gain

The user sets  $P$ ,  $I$ ,  $D$ , and  $S$ , from which SCAN determines  $a$ ,  $b$ , and  $c$ . These feedback gains control the response rate of the feedback: when too low the feedback responds slowly and features

wash out; when too high the feedback over-reacts and oscillates. In my experience with AFM and STM, setting P 10 to 100 times higher than I provides optimum feedback. Setting D about ten times greater than I does reduce overshoot, but the concurrent increase in high-frequency noise limits its usefulness. Negative values of D have the reverse and sometimes useful effect of decreasing high-frequency noise but increasing overshoot.

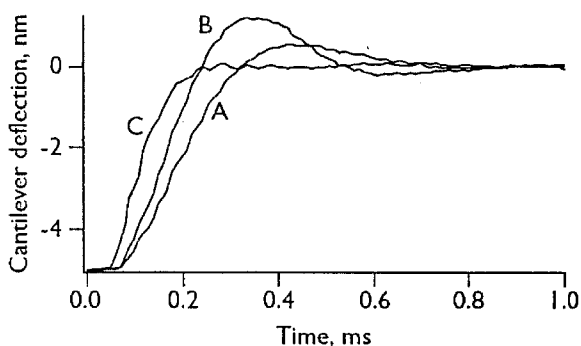
### Step response curves provide information about instrument performance

Adjusting P, I, and D in the usual manner, by inspecting image quality, generally fails to produce optimal instrument performance. Step response curves provide a better visual indicator of feedback rise time, noise level, and oscillations.

To generate a step response curve, the DSP changes the setpoint by an amount  $\Delta S$  and records the error signal for a few milliseconds afterward while the feedback adjusts to the new setpoint. The feedback behaves the same during a step response measurement as it would if the tip encountered a step on the sample and both the tip and step were perfectly sharp.

Using step response curves, the user can optimize feedback gains either manually or automatically. For manual optimization, the user adjusts P, I, and D based on the appearance of step response curves displayed after each iteration (figure 3). On AFM-3, optimized feedback usually settles within 20–50 feedback cycles, under one data point if the scan has 250 points/line

Figure 3. Manual optimization of AFM-3 feedback using step response curves.  $S = 5$  nm. A) Integral feedback only:  $P = 0$ ,  $I = 0.002$ ,  $D = 0$ . B) Raising  $I$  to 0.003 increases the response rate but also increases overshoot (or oscillations). C) Raising  $P$  to 0.03 increases response and decreases overshoot. The feedback now settles in about 0.25 ms — within one data point if a scan has 250 points/line acquired at 8 lines/second or slower.



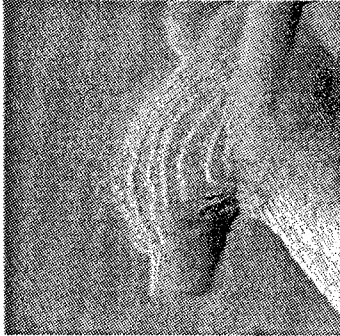
acquired slower than 4–10 lines/second. Thus, the feedback follows topography with a lag of no more than one data point. With the feedback optimized in this manner a slope-shaded topography (Z) image does not have any less detail than an error signal (E) image (figure 4).

I have developed a self-optimizing feedback mechanism

Control systems engineers have developed a number of optimal feedback algorithms that eliminate the need for manual adjustment of feedback parameters (Isermann et al., 1992). These algorithms use mathematical models of the stage and as such are instrument-specific. I have developed a general-purpose “evolutionary” algorithm that optimizes feedback parameters while making few assumptions about the nature of the stage. I use this algorithm after installing a new tip or sample or when oscillations or other signs of poor feedback performance appear. The soft-

Shaded topography ( $dZ/dX$ )

Error signal (E)



Correlation diagram

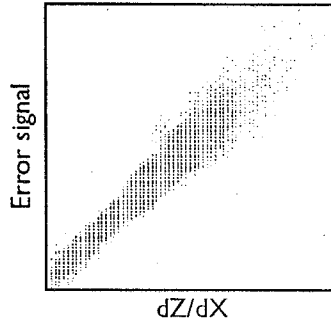


Figure 4. Comparison of slope-shaded Z with a simultaneous E image. Since the step response curve settles within one data point, the two images differ negligibly. The lines in the correlation diagram arise from the quantization of Z and C. The sample is probably a fragment of an electron-beam-deposited AFM tip. The substrate is float-polished quartz. Image size 385 nm.

ware generates nine “mutations” of the “parent” gain factors P, I, and D according to the following example:

$$iP_g = optP_{g-1} iR$$

where:

$iP_g$  = mutation  $i$  of proportional gain,  $i=1$  through 9

$optP_{g-1}$  = parent gain (optimal mutation of previous generation)

$g$  = generation

$iR$  = a random number between  $(1+V_g)$  and  $1/(1+V_g)$

$V_g$  (“variance”) = initially between 0.2 and 0.6

Note that the algorithm multiplies the three parameters  $optP$ ,  $optI$ , and  $optD$  by three *different* random numbers. The DSP generates five step response curves in which it increases the setpoint, and five in which it decreases the setpoint for the parent and each mutation, and calculates an optimization index for each curve:

$$O = oE_{max} + \frac{\sum_{t=0}^{t_{max}} |E_t|}{t_{max}} \quad (1)$$

where:

$O$  = optimization index

$E_{max}$  = overshoot

$o$  = user-specified overshoot factor (typically 0.3)

$t_{max}$  = number of points in step response curve (typically 100)

The feedback gains of the mutation having the lowest average optimization index replace the parent.



After each generation that a mutation replaces the parent, the software raises the variance  $V$  by 10%. Each generation that the software fails to find a mutation better than the parent, it lowers the variance 10%.  $V$  starts at 0.4; when it passes a preset threshold, usually 0.2, the optimization process terminates.

The optimizer usually cannot recover from oscillating feedback, so the user should initially set feedback parameters lower than optimal. To prevent the optimizer itself causing the feedback to go into oscillation, if the best optimization index increases by more than 150% over one generation, the PC turns the feedback off for 10 ms and reduces  $P$ ,  $I$ , and  $D$  by 50%.

Figure 5 demonstrates the optimization process using AFM-3. The graphs show standard deviation, overshoot, variance, and feedback gains over time. I set the feedback gains unusually low at the beginning to better demonstrate the optimization progress. Normally the algorithm would terminate when the variance falls below 0.2 (generation 40), but for the illustration I have disabled this feature. The entire 100-generation process required 10 seconds.

Figure 6 shows step response curves taken before and after the optimization process of figure 5. Note that after optimizing, the error signal first overshoots then undershoots zero. Increasing the overshoot factor  $\alpha$  reduces overshoot but tends to increase undershoot.

Figure 5. Evolutionary optimization of feedback parameters using step response curves. The optimization ran on a 2  $\mu\text{m}$  scan range AFM using a 50  $\text{\AA}$  step size and 100 points per step response curve. The software took six step response curves for each mutation. Total elapsed time, 10 seconds.

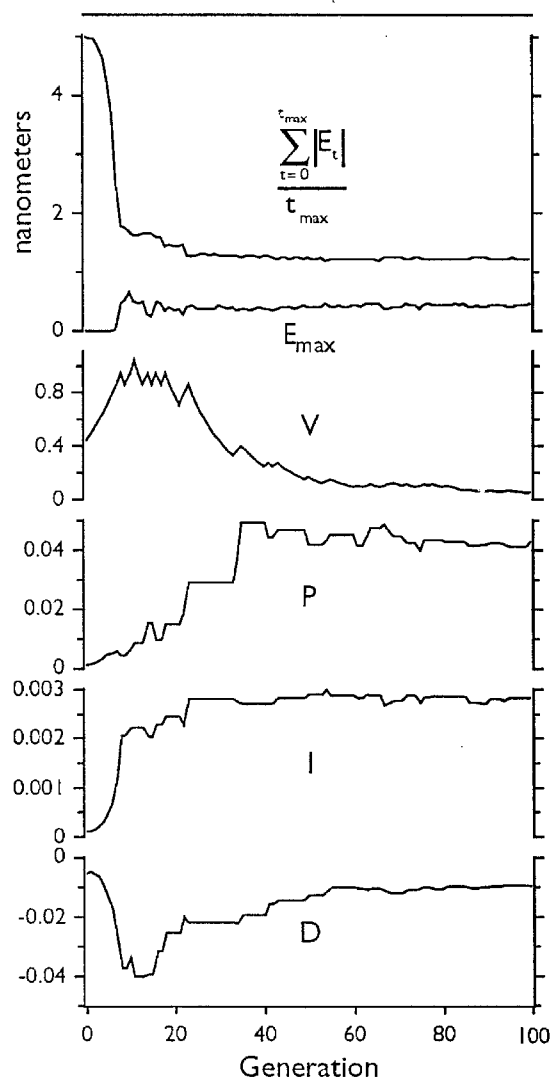
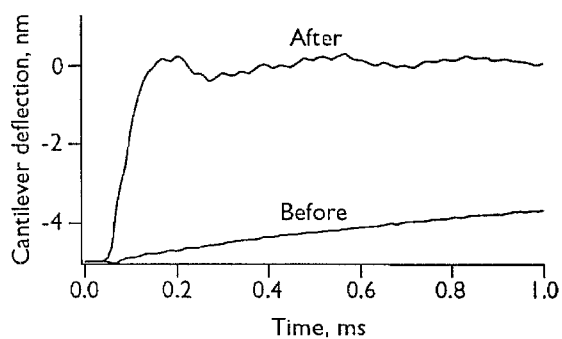


Figure 6. Step response curves taken before and after the automatic optimization shown in figure 5.



### Evolutionary optimization works with varying levels of effectiveness on different instruments

I have used the optimization algorithm on 1, 10, and 75  $\mu\text{m}$  scan range AFMs. It is most useful on long range instruments that tend to oscillate; it does an excellent job of maximizing response rate while minimizing oscillations. On more stable instruments an experienced user generally does a better job of optimizing the feedback, but requires much more time.

Evolutionary optimization works better on AFMs than on STMs, since STMs generally produce poor step response curves. The logarithmic dependence of tunneling current on tip-sample separation may be responsible, but I have not tried using a log amp to linearize the current.

AFMs with higher noise levels may benefit from more step response curves per mutation, a higher termination variance, or fewer mutations per generation. Greater noise translates into more uncertainty in the optimization index  $\mathcal{O}$ , so the optimizer more often finds mutations that appear to perform better than a fully optimized parent. The most direct solution is to average more step response curves together, thereby reducing the noise. However, the optimization process will then take longer. Alternatively, the user can lower the standard that defines “optimized feedback” either by increasing the termination variance or decreasing the number of mutations per generation (reducing the chance that the optimizer will find an apparently better mutation).

## Scanning

---

The scan subroutine emulates an analog raster by continually moving the scanner and time-averaging incoming data

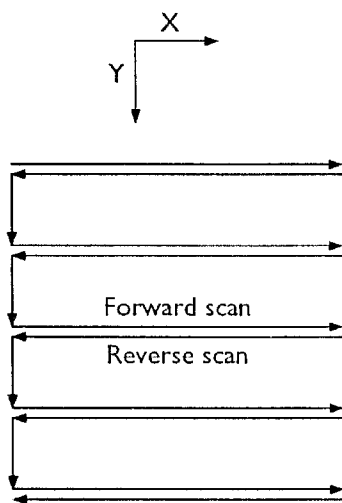


Figure 7. XY raster pattern. The coordinate system is left-handed, a common convention in graphics applications.

The scan subroutine generates the XY raster, ramping X repeatedly between selected minimum and maximum values and advancing Y after each cycle (figure 7). It also stores image data.

To achieve the user-specified scan rate, the raster algorithm adjusts the number of timer periods spent at each data point. To keep scanner motion as continuous as possible, the raster advances once per period rather than once per data point (one point usually spans 20–50 timer periods). Furthermore, the data recorded for each point represents an average value over the timer periods making up that point. The combination of continuous scanner motion and data averaging ensures that features somewhat smaller than the distance between data points still appear in the image. In addition, data averaging improves the signal-to-noise ratio.

The timer period determines the fastest possible scan rate. The DSP cannot scan faster than one timer period per data point, or 200 lines/second for an image with 250 points/line in each direction (left-to-right and right-to-left).

The scan subroutine uses a nonrecursive, floating-point algorithm

Two raster formulas encounter problems when used to generate continuous rasters: recursive algorithms tend to produce cumulative roundoff errors, and integer-based algorithms tend to produce scan size and rate limitations. Thus for programming simplicity SCAN uses the following formulas (all numbers are floating point):

$$X'_t = X'_0 + t\Delta X'$$

$$Y'_t = Y'_0 + t\Delta Y'$$

where:

$$X'_0 = (X_0 - X_c)\cos\theta + X_c - (Y - Y_c)\sin\theta - Y_c$$

$$Y'_0 = (X_0 - X_c)\sin\theta + X_c + (Y - Y_c)\cos\theta - Y_c$$

$$\Delta X' = \Delta X \cos\theta$$

$$\Delta Y' = \Delta X \sin\theta$$

and:

$(X, Y)$  = user coordinates

$(X', Y')$  = scanner coordinates

$(X_c, Y_c)$  = center of rotation

$\Delta X$  = amount to increment raster with each timer period

$X_0$  = starting point of scan

$t$  = time in timer periods

$X_t$  =  $X$  position at time  $t$

$\theta$  = scan rotation

The center of rotation generally coincides with the center of the image.

### A Z raster helps the feedback handle highly sloped samples

Adding a raster to Z during scanning tends to improve images of highly sloped surfaces by helping the feedback follow sample tilt. The following recursive algorithm takes about 5 instruction cycles:

$$Z'_t = Z_t + \Delta Z$$

The DSP inverts the slope  $\Delta Z$  with each change in scan direction. SCAN determines  $\Delta Z$  with a leveling routine that measures sample height at the corners of the current scan (the software also uses the resulting information to level plotted images). Tilt correction especially helps “fast scan” imaging, in which the user reduces the feedback gains and records the error signal E rather than Z.

Unlike the XY raster, finite word length does not cause problems with the Z raster because the feedback loop can correct for small errors.

### A tilted scan serves as an alternative to feedback for NSOM

Nearfield scanning optical microscopes (NSOMs) sometimes lack a feedback mechanism. With these instruments, sample tilt compensation in X and Y can help maintain reasonable tip-sample separation while scanning. Instead of the feedback loop, SCAN uses the following algorithm to determine Z (all numbers are floating point):

$$Z_t = iX_t + jY_t + k$$

The user determines the leveling factors  $i$ ,  $j$ , and  $k$  by inspecting image sharpness.

### Background subtraction counters the effect of spurious signal variations in scanned-cantilever AFMs

Optical-lever based AFMs, and scanned-cantilever instruments in particular, tend to exhibit spurious variations in cantilever deflection signal with the position of the scanner. As a result the downward force that the cantilever exerts on the sample varies with X and Y. With a scanned-cantilever stage, this signal variation arises mostly as a result of cantilever warp, and the variation with Y position greatly exceeds the variation with X.

I have implemented a background subtraction algorithm that counters the effect of spurious signal variation with  $Y$ . The algorithm adds a raster to the setpoint:

$$S' = S + Y \frac{dS}{dY}$$

A leveling routine determines the slope  $dS/dY$  by measuring  $S$  at the minimum and maximum  $Y$  when the tip is not touching the sample. I use this leveling routine once each time I replace the sample or cantilever.

### Tip-sample approach

---

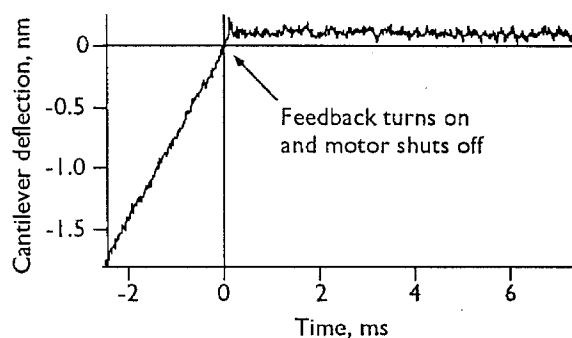
Tip-sample approach curves help to identify poor or false engages

To bring the tip within scanner range of the sample in AFM-3, a motor slowly turns a screw that determines tip-sample separation. A DAC voltage determines the motor speed. Generally, before starting the approach, the user has turned the feedback off and set  $Z$  at the middle of the scanner range. The engage routine sets the DAC to start the motor and then waits for the sensor signal to cross the setpoint. When it does, the DSP starts the feedback and shuts the motor off.

The software creates a digital oscilloscope display by recording the sensor output  $C$  to a circular buffer during the approach. It



Figure 8. AFM tip-sample approach curve taken in air. At -2 ms the tip and sample are already in contact and the motion of the approach motor is pushing the cantilever up. When the cantilever deflection reaches its setpoint (zero deflection), the DSP starts the feedback and shuts the motor off. Since the motor takes about 35 ms to come to a halt, the cantilever deflection does not immediately settle to the setpoint.



plots the output for a few milliseconds before and after achieving feedback as a check for false engages and tip crashes (figure 8).

## Modulation measurements

---

I have written DSP software to take modulation images by emulating a lock-in amplifier

In a modulation image, the DSP measures the dependence of an input on an output at each data point. In AFM, for example, the dependence of cantilever deflection on  $Z$  ( $dC/dZ$ ) yields a hardness image.

For lock-in detection of  $dC/dZ$ , the DSP adds to  $Z$  a sinusoidal modulation having a user-selected amplitude and frequency (for hardness images, amplitude usually about 1 nm and frequency about 1 kHz). The DSP modulates  $Z$  continually, interrupting neither the XY raster nor the feedback to take a modulation measurement. The user sets the frequency high enough and the feedback gains low enough so the modulation lies outside the bandwidth of the feedback.

$$Z'_t = Z_t + Hm_t$$

where:

$$m_t = \sin\left(\frac{2t}{p}\right)$$

$H$  = peak-to-peak modulation amplitude

$t$  = time in timer periods

$p$  = modulation period in timer periods

A circular buffer lookup table contains the values of  $m_t$ .

Assuming a linear sensor, the following formulas yield the magnitude of the error signal at the modulation frequency:

$$A_c = \frac{4}{np} \sum_{t=1}^{np} E_t m_t$$

$$A_{90} = \frac{4}{np} \sum_{t=1}^{np} E_t m_{t+\frac{p}{4}}$$

$$A = \sqrt{A_0^2 + A_{90}^2}$$

$A$  = peak-to-peak magnitude of error signal

$A_0$  = amplitude of  $E$   $0^\circ$  out of phase with  $Z$  modulation

$A_{90}$  = amplitude of  $E$   $90^\circ$  out of phase with  $Z$  modulation

$n$  = number of modulation periods per data point

The factor  $4/np$  provides a peak-to-peak normalization of the amplitudes.

Note that each data point must contain an integral number of modulation periods. For a 1 kHz modulation frequency, the fastest possible scan therefore has one 1 ms modulation period per data point, or 2 lines/second at 250 points/line. Slower rates improve signal-to-noise.

For the hardness images presented in this thesis, I have recorded only  $A_0$ , since the  $90^\circ$  component provided negligible information. Not recording the  $90^\circ$  component decreases the processor time required for modulation imaging from 25 to 18 instruction cycles and cuts the amount of data to store in half.

Michael Youngquist (1993) has found that for high-performance applications the differential nonlinearity of the Z DAC introduces significant noise. For such applications he suggests (though he has not tried) using separate Z and modulation DACs, summing their outputs with analog electronics.

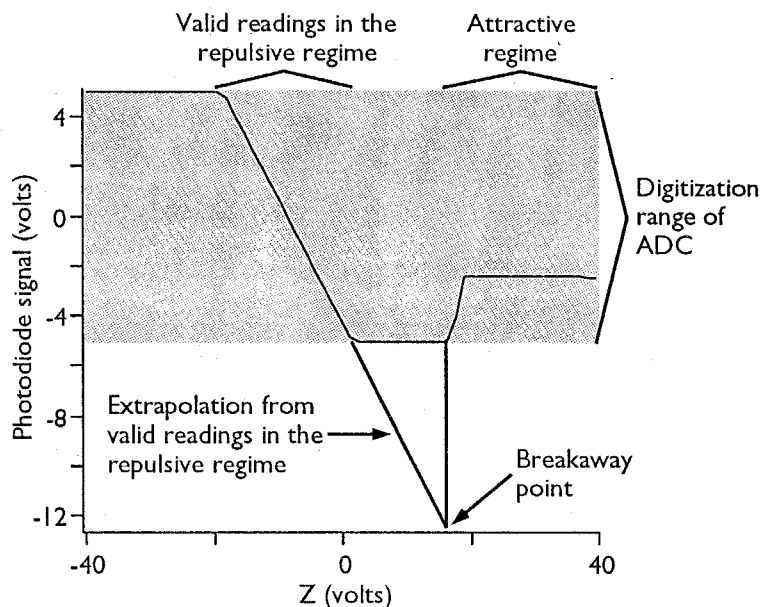
## Spectroscopy

---

SCAN can extract the meniscus force and optical lever response from force curves

When imaging in air, contamination films form a meniscus that pulls the tip and sample together and affects the tip-sample interaction. Force microscopists commonly use force curves,

Figure 9. Force curves often exceed the range of the ADC. In order to determine the meniscus force SCAN isolates the valid (within ADC range) repulsive-regime portion of the curve and fits a line to it. The slope of this line yields the sensitivity of the optical lever, while its Y intercept with the breakaway point yields the cantilever deflection at zero tip-sample contact force.



graphs of cantilever deflection vs.  $Z$ , to measure the meniscus attraction (Weisenhorn et al., 1989; see chapter 1).

SCAN determines both the meniscus force and optical lever sensitivity from force curves made by pulling the tip and sample apart. It isolates the repulsive-regime portion of the force curve and fits a line to it. The slope of the line indicates optical lever sensitivity. An extrapolation to the “breakaway” point where the tip comes out of contact with the sample indicates meniscus force (figure 9).

Imaging spectroscopy involves taking a spectrum at each point in a scan

By taking force curves at different locations, Mate et al. (1989) have measured one-dimensional cross-sections of liquid film

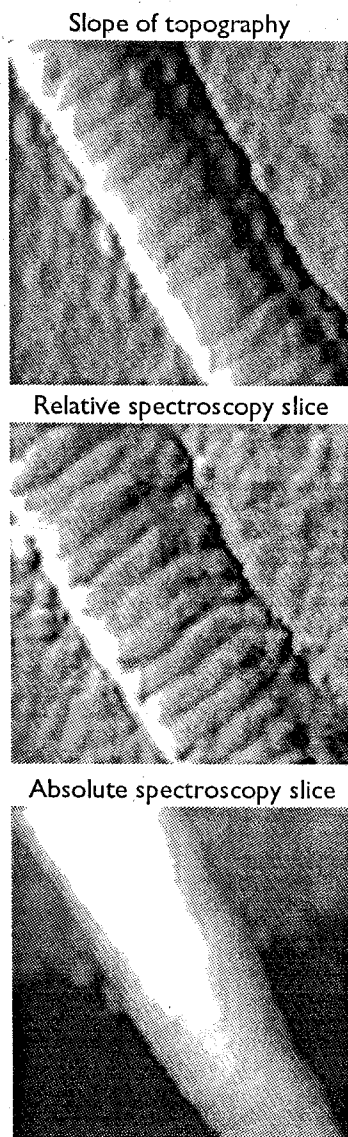


Figure 10. 700 × 700 nm slope-shaded topographic image of a collagen fibril with slices from relative and absolute force curve imaging spectroscopy stacks.

thickness. I have extended this technique to two dimensions, a development made practical largely by image processing techniques to handle the resulting data (see below). This “force curve imaging spectroscopy” resembles current imaging tunneling spectroscopy (CITS), a scanning tunneling microscope technique that involves acquiring a tunneling current vs. tip bias spectrum at each point in a scan (Hamers et al., 1986).

Imaging spectroscopy creates a three-dimensional “stack” of two-dimensional images or “slices.” Each slice is an image of cantilever deflection at a specific value of  $Z$ . The user can display individual slices and inspect the force curve for any given point.

SCAN can acquire force curve imaging spectra in “relative” and “absolute” modes

SCAN presents the user a choice between “relative” and “absolute” imaging spectroscopy modes. In either case, the user specifies minimum and maximum  $Z$  values for the spectra. In the relative mode, SCAN adds these values to the feedback-determined  $Z$  before taking each spectrum, while in the absolute mode, SCAN acquires all spectra between the same two values of  $Z$ . Relative-mode slices thus resemble error signal images, while absolute-mode slices resemble topography images (figure 10).

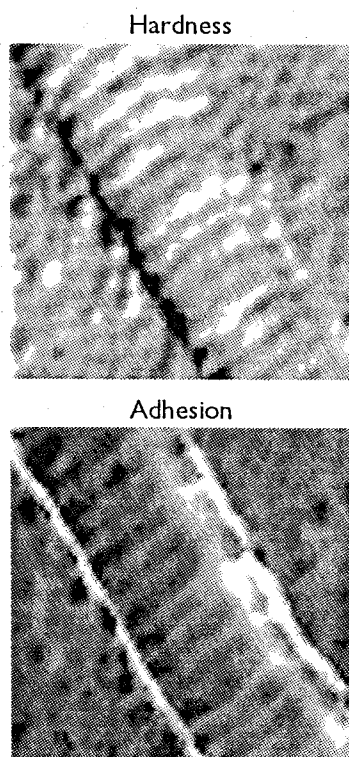


Figure 11. Hardness and adhesion images derived from the absolute-mode stack in figure 10. For interpretation of these images see chapters 7 and 6, respectively.

PROCESS can extract sample hardness and adhesion data from imaging spectroscopy stacks

Applied to force curve imaging spectroscopy stacks, the method described above for determining meniscus force and optical lever sensitivity results in the creation of adhesion and sample hardness images, respectively (figure 11). I have implemented this capability in my image processing program, PROCESS.

Hardness images thus derived resemble Z modulation images, while adhesion images resemble differential lateral deflection data (the difference between forward and reverse lateral deflection images; see chapter 6). Relative and absolute mode stacks yield similar hardness and adhesion images.

Line spectroscopy permits acquisition of several images at different forces

I have developed line spectroscopy primarily to acquire friction and modulation images at several different forces simultaneously. For this type of imaging, the software scans the same line several times at different forces before proceeding on to the next line. A set of several images results.

## *The user interface*

---

The programmable parser automates standard tasks while extending the flexibility of the software

I have automated the DSP-based control system using a programmable instruction parser.

An instruction parser accepts a typed command (like “scan”), carries it out, then prompts for the next command. SCAN has an instruction buffer so the user can issue hundreds of instructions at one time; it also has conditional and branch statements, the ability to load sets of instructions (macros) from text files, and the ability to display data and accept user input. For flexibility the instruction set consists of fundamental, low-level operations, effectively constituting an SPM programming language.

This programmable parser architecture has many uses. The availability of low-level instructions helps to perform the unusual operations often needed when testing new instruments. For completed instruments SCAN has a “shell” macro that carries out the routine functions of SPM and makes the programmable environment invisible to casual users. I have also developed macros that perform complete experiments under automatic control: for example, sequences of scans with varying conditions, strings of lithography operations, and area searches. Macros can recognize and rectify common problems such as false engages and use the automatic feedback optimizer to correct poor feedback settings.

The programmable parser offers flexibility but, with a shell macro, does not compromise ease of use. It also simplifies the architecture of the SPM control system by breaking the software into three distinct levels — DSP, PC, and macros.



## References

---

- Baselt, D.R., Clark, S.M., Youngquist, M.G., Spence, C.F., and Baldeschwieler, J.D. (1993) Digital signal processor control of scanned probe microscopes. To appear in *Rev. Sci. Instrum.*
- Bucek, V.J. (1989) *Control Systems*. Prentice Hall, Englewood Cliffs, NJ
- Hamers, R.J., Tromp, R.M., and Demuth, J.E. (1986) Surface electronic structure of Si(111)-(7×7) resolved in real space. *Phys. Rev. Lett.* 56(18), 1972–1975
- Isermann, R., Lachmann, K.-H., and Matko, D. (1992) *Adaptive Control Systems*. Prentice-Hall, New York 1992
- Mate, C.M., Lorentz, M.R., and Novotny, V.J. (1989) Atomic force microscopy of polymeric liquid films. *J. Chem. Phys.* 90(12), 7550–7555
- Weisenhorn, A.L., Hansma, P.K., Albrecht, T.R., and Quate, C.F. (1989) Forces in atomic force microscopy in air and water. *Appl. Phys. Lett.* 54(26), 2651–2653
- Youngquist, M.G.Y. (1993) *Atomic Scale Imaging and Spectroscopy Using Scanning Tunneling Microscopy*. Ph.D. dissertation, California Institute of Technology

# Chapter 4. Image processing

## *Introduction*

---

I have programmed a DSP-based image processing facility

Image processing accomplishes two tasks. First, it converts AFM data, which consists of lists of numbers, into images. Second, because AFM images have 65,536 levels, while the eye can only visualize 64 grey levels, in their unprocessed state most AFM images appear to contain no features whatsoever. The image processing software must discard less relevant information to create an informative image.

The image processing program I have programmed, PROCESS, contains many common algorithms available from digital image processing books and in commercial software packages. However, PROCESS handles 32-bit floating-point data, while most commercial software packages handle 8-bit integer data. Like SCAN, it runs on the DSP board with a PC-based user interface.

Chapter 4 describes the special algorithms I have developed to handle AFM data. It also contains a review of the standard

processing techniques used in later chapters, all of which I have programmed into PROCESS.

This thesis does not include code listings

A complete listing of the SCAN/PROCESS software runs over 500 pages and is not within the scope of this thesis. However, upon request I can provide explicit information on any of the algorithms in these programs.

### Review of standard SPM image processing algorithms

---

PROCESS normally displays height-keyed images

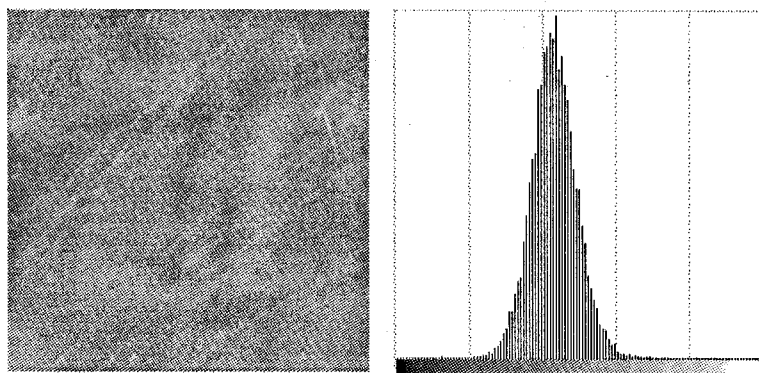
By default, PROCESS presents "height-keyed" images in which high areas are bright. The user can assign arbitrary colors to the image for aesthetic purposes; images have 256 color levels.

Histogram clipping serves to adjust contrast and brightness

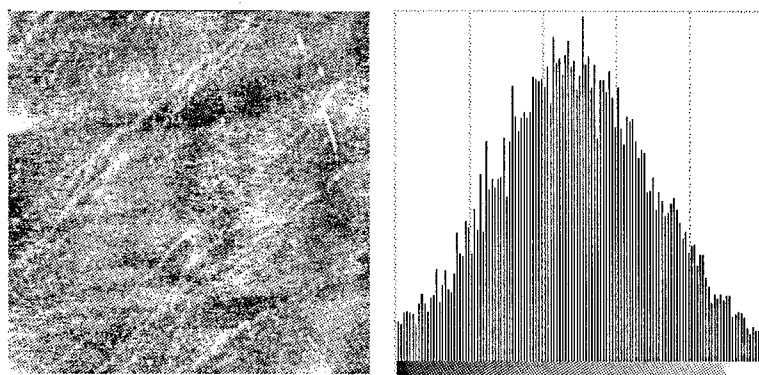
SPM images normally contain 16-bit data (65,536 levels), while video display devices can only display 256 or fewer color levels (8-bit data). Contrast and brightness adjustment converts 16-bit SPM data to 8-bit screen data.

Figure 1. Histogram clipping.

Top: a  $1.6 \times 1.6 \mu\text{m}$  AFM image of float-polished quartz after min/max scaling: the highest point in the image appears white and the lowest point appears black. Histogram at right.



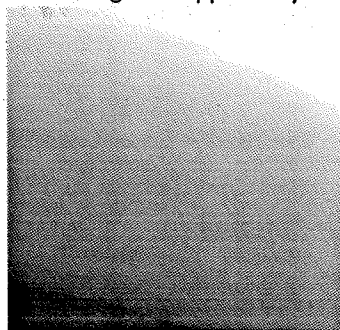
Bottom: the same image after 3% histogram clipping. The software ignores the highest and lowest 3% of data points. The highest remaining point and all higher points appear white; the lowest remaining point and all lower points appear black.



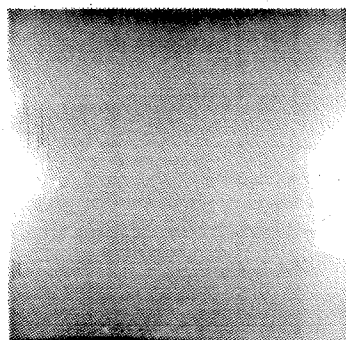
By default, PROCESS uses a histogram-clipping algorithm to adjust contrast and brightness (figure 1). The software creates a histogram of the data, a plot that shows the frequency with which specific values occur. It then discards the highest 5% and lowest 5% of the data. The software maps the lowest remaining point to video value 0 (black), the highest remaining point to value 255 (white), and all other points to intermediate values. PROCESS handles the “clipped” data by mapping it either to 0 (lowest 5%) or 255 (highest 5%).

PROCESS offers “contrast” and “brightness” options that modify the upper and lower clipping percentages. It also allows the user to manually clip the histogram.

Histogram clipped only



Plane subtracted



Paraboloid subtracted

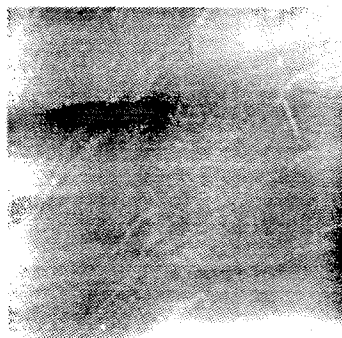


Figure 2.  $1.6 \times 1.6 \mu\text{m}$  AFM image of float-polished quartz demonstrating plane and subsequent paraboloid subtraction.

### Plane subtraction removes overall tilt

SPM images, topography images in particular, generally exhibit some overall tilt. Plane subtraction removes this tilt, improving the visibility of all features (figure 2). By default PROCESS performs a best-fit plane subtraction on all topography images, determining mathematically the degree of tilt to remove. PROCESS also offers a manual plane subtraction, in which the user specifies three points on the image that should all lie at the same height.

### Paraboloid subtraction removes spurious image curvature

The scanning process often leaves spurious image curvature. A best-fit paraboloid subtraction, similar in principle to plane subtraction, can remove much of this curvature. Like plane subtraction, paraboloid subtraction enhances the visibility of all features (figure 2).

### Light shading greatly improves the dynamic range of topography images

In many images, tall features obscure smaller ones; that is, the number of available color levels limits the range of feature sizes visible. The various forms of light shading make both small and

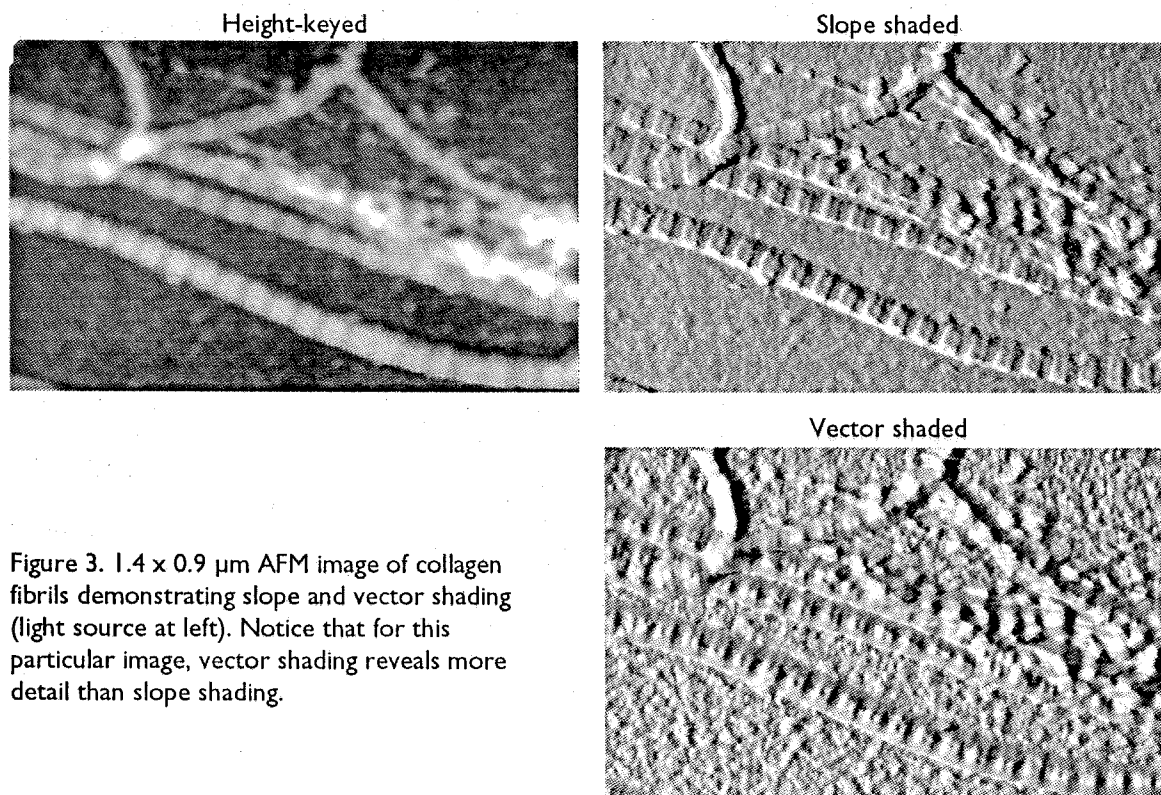


Figure 3.  $1.4 \times 0.9 \mu\text{m}$  AFM image of collagen fibrils demonstrating slope and vector shading (light source at left). Notice that for this particular image, vector shading reveals more detail than slope shading.

large features visible and also lend a realistic “solid-modeled” look to topography images.

In the simplest form of shading, which I refer to as “slope-keyed shading” (figure 3), the software replaces the value of each image pixel with the difference between that value and the value of a neighboring pixel. The colors of the resulting image reflect the slope, rather than the height, of the sample. PROCESS permits slope-keyed shading in nine directions, one for each neighboring pixel.

In the algorithm I have termed “vector shading” (figure 3), the software constructs a three-dimensional model from the image data and determines how this model would appear with a user-placed light source shining on it (Hill, 1990, p. 546). Vector

shading appears more realistic than slope shading but also takes about a hundred times longer. However, because the colors of the resulting image depend nonlinearly on the slope of the sample, vector shading often reveals features not visible with slope shading.

Because of the large number of points in SPM images, the vector shading algorithm in PROCESS (which calculates one intensity value for each data point) produces results similar to Phong shading (which interpolates between data points to provide smoother gradation of color; Hill, 1990, p. 559). For this reason I often refer to vector shading by the more conventional name of Phong shading.

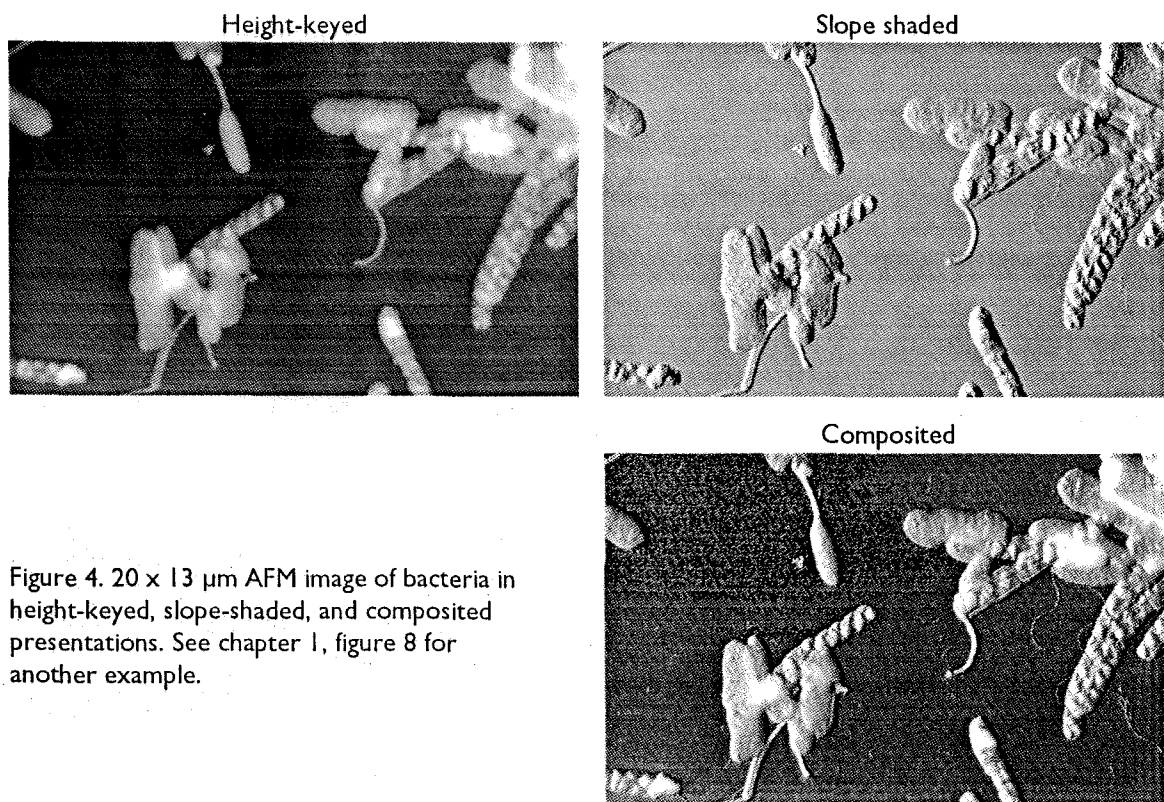


Figure 4. 20 x 13  $\mu\text{m}$  AFM image of bacteria in height-keyed, slope-shaded, and composited presentations. See chapter 1, figure 8 for another example.

### Compositing shows two data sets in one image

The process that I refer to as “compositing” combines two images. One image determines only the brightness of each pixel in the resulting image, while the second controls hue, saturation, and (to a limited extent) brightness. Thus, one image “colors” a second, black-and-white image (figure 4).

Coloring a light-shaded image with its height-keyed version can help to isolate features from a background. In addition, coloring a topography image with friction, hardness, or other simultaneously acquired data clarifies the spatial relationship between features in the two images.

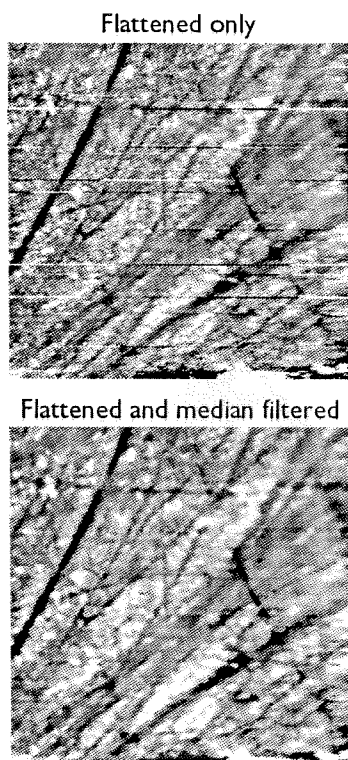


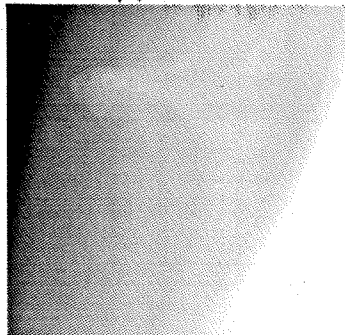
Figure 5.  $1.8 \times 1.8 \mu\text{m}$  AFM image of cast iron before and after median filtering. Median filtering removes the streaks caused by tip changes, but reduces the sharpness of the image.

### Smoothing and median filtering trade improved noise for reduced resolution

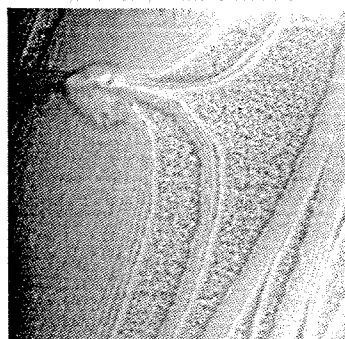
Smoothing and median filtering are sliding window convolutions that replace the value of each pixel with the average or median of the pixels in the immediate vicinity. Both algorithms presume that the user has acquired an “oversampled” image (higher resolution than necessary) and improve resolution by decreasing the effective sampling rate. In principle, the user can produce the same effect during acquisition by scanning at a slower rate or lowering the number of pixels in the image. Median filtering removes sharp, isolated noise spikes (figure 5).



Manually plane-subtracted



Statistical differenced



Slope shaded

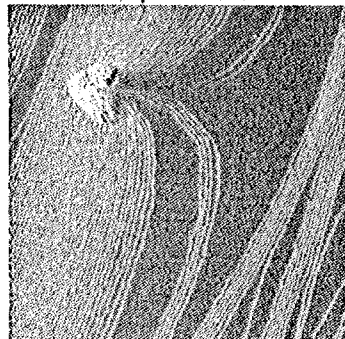


Figure 6. 730 x 730 nm STM image of atomic steps on gold comparing statistical differencing with shading. Sample prepared by Michael Youngquist. Image taken by Michael Youngquist and David Baselt.

### Statistical differencing improves dynamic range

Like light shading, statistical differencing makes minor features visible in the presence of tall ones. It is a sliding-window algorithm that brings the contrast (standard deviation) and brightness (mean) toward a global constant in the vicinity of each point (Niblack, 1986, p. 45). The user can adjust the strength of both effects — local standard deviation enforcement and local mean enforcement — independently. Statistical differencing can substitute for light shading (figure 6) and also provides additional dynamic range to already shaded images.

### Fourier filtering allows frequency-specific noise removal

The Fourier transform allows the user to remove features of specific frequencies — presumably high and low frequency noise. It is most useful for analysis and processing of images having regular, repeating features such as those that might appear in a crystalline material. Otherwise, the Fourier filter has limited use in SPM, since smoothing (to suppress high frequency noise) and flattening (for low frequency noise) algorithms produce similar results in less time.

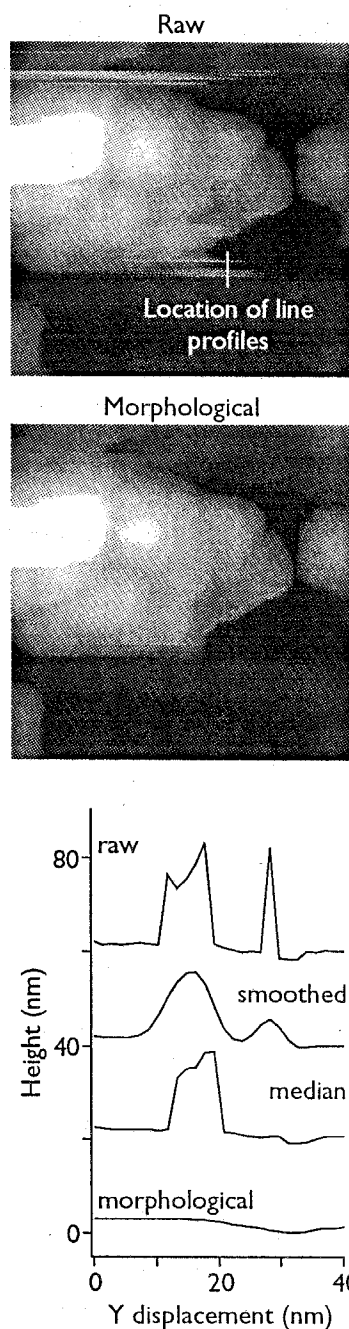


Figure 7. 360 x 360 nm scan of polystyrene beads in water demonstrating noise removal by opening. The line profiles compare the effects of three algorithms on noise spikes.

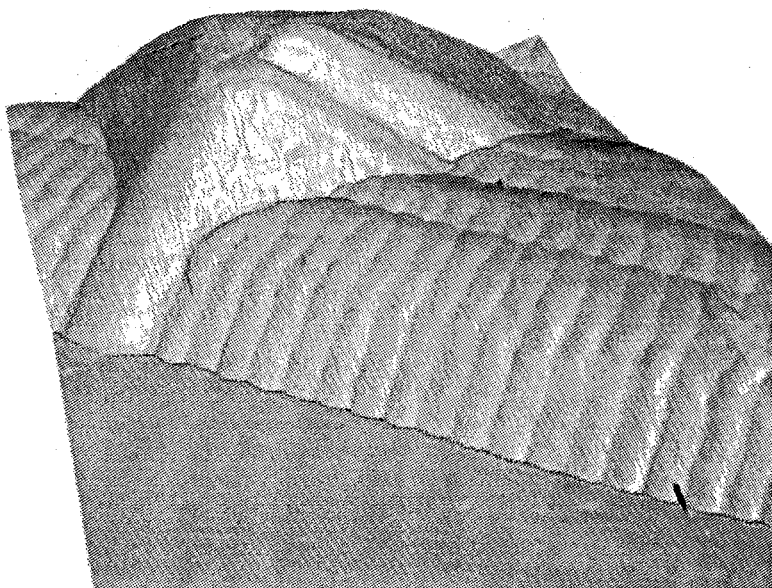
Morphological image processing removes or retains features based on their shape

Morphological image processing is a family of sliding-window processes (Dougherty, 1992). The fundamental process of “dilation” scans a three-dimensional model of an image with a user-defined “tip,” similar to the way an AFM scans a sample, and replaces the image with the path of the tip. Morphological image processing consists of combining one or more dilation steps with appropriate image arithmetic (multiplying the image by  $-1$ , subtracting two images, etc.).

Morphological image processing can eliminate small bumps and/or pits that the tip cannot fit into, or keep small features while eliminating larger ones. It differs from Fourier filtering and smoothing in that it does not blur large features when removing small ones or create “ringing” when removing large features.

Dilation can simulate the effects of dull or multiple SPM tips. The process referred to as “opening” (dilation followed by multiplication by  $-1$ , dilation, and multiplication by  $-1$  again) suppresses noise (figure 7). The process that I refer to as “morphological decurving” consists of opening an image with a very dull tip and subtracting the result from the original image, thereby keeping only features that the tip does not fit into and eliminating spurious image curvature.

Figure 8. AFM image of a collagen fibril in a three-dimensional projection. The pyramidal tip has convoluted a large, relatively sharp object on the fibrils, effectively imaging itself. The three-dimensional projection helps to visualize the size and shape of the tip in relation to the collagen fibrils. Top side of image measures 1.3  $\mu\text{m}$  in length. Tip image is 1  $\mu\text{m}$  tall. Vector shaded.

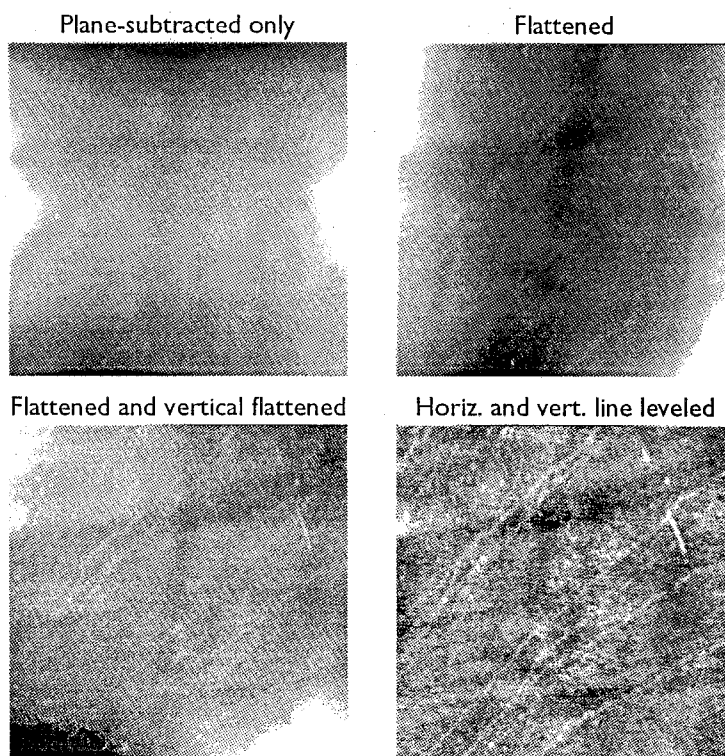


### Three-dimensional projections provide aesthetic appeal

Since SPMs quantitatively measure the topography of a surface, software algorithms can create three-dimensional models from SPM images and “project” them onto the video screen. In some cases, viewing SPM images as three-dimensional projections can help an audience to visualize the topography of a sample (figure 8). However, 3D projections tend to obscure the minor features critical to scientifically relevant imaging. In general, an audience will likely find normal “two-dimensional” presentations more easily interpreted if not as glamorous as three-dimensional presentations.

PROCESS offers a true three-point perspective with continuously variable user-selected viewing angle and distance (Ammerall, 1986, p. 74).

Figure 9.  $1.6 \times 1.6 \mu\text{m}$  AFM image of float-polished quartz demonstrating line leveling algorithms for decurving. Also compare with figure 2, which demonstrates decurving by paraboloid subtraction.



### Line profiles allow quantitative measurements

The line profile, a graph that represents a cross-section through an image (figure 7), serves as the customary tool for quantitative measurements of feature size.

### Newly developed SPM image processing algorithms

#### Line algorithms remove noise and spurious curvature

The Nanoscope SPM (Digital Instruments, Santa Barbara, CA) offers a “flattening” algorithm that removes low-frequency noise by subtracting the average of each horizontal line from that line. I

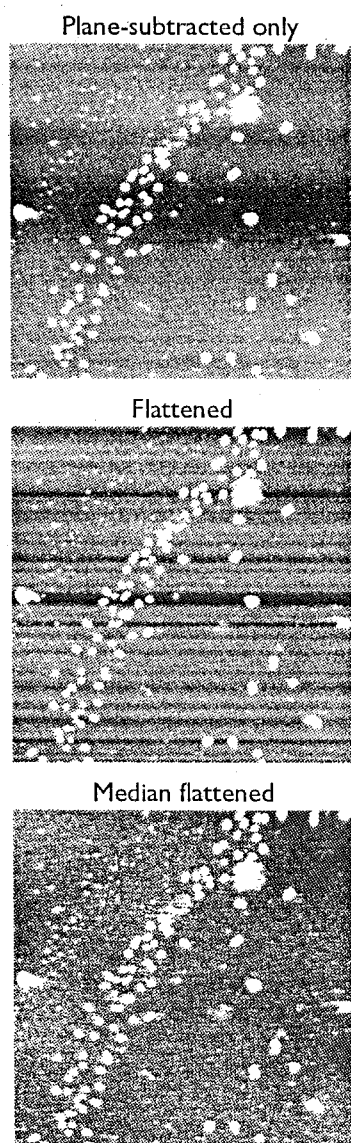


Figure 10.  $5 \times 5 \mu\text{m}$  image of a DNA preparation demonstrating line algorithms for low-frequency noise removal. The bumps in this image create trenches when flattening, but not when median flattening. DNA preparation by Solomon Ting.

have implemented this algorithm and created five more based on it.

Since flattening removes horizontal curvature, I added the complementary algorithm “vertical flattening” to remove vertical curvature. The two algorithms together remove much spurious image curvature (figure 9).

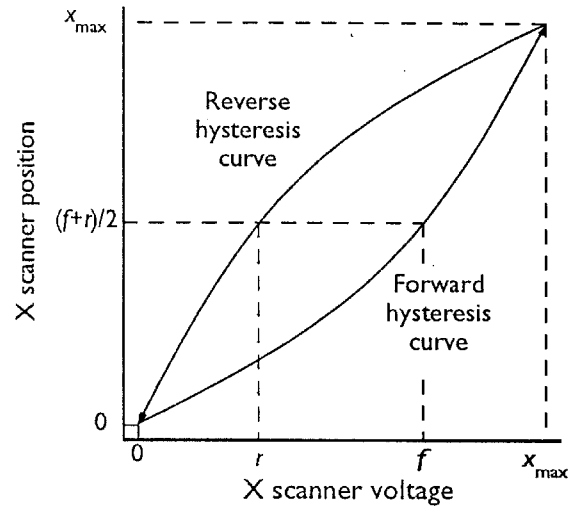
When it encounters isolated high features, flattening tends to produce trenches. Horizontal and vertical median flattening resists the average-skewing effects of a few high or low points by using the median instead of the average (figure 10).

To even more effectively remove image curvature, I have developed the horizontal and vertical “line leveling” algorithms. Instead of just subtracting the average from each line, these algorithms subtract a best-fit line (figure 9).

#### A second-order curve fit reduces hysteresis effects

Although most commercial instruments compensate for hysteresis by generating a nonlinear raster, SCAN/PROCESS uses a postacquisition algorithm to free up processor time during acquisition. The algorithm warps images by moving their data points according to a user-defined second-order curve (figure 11). It assumes that the user has taken both forward (scan direction left-to-right) and reverse (right-to-left) data. The user marks an easily recognized feature on both the forward and reverse images, generating two X coordinates respectively termed  $f$  and  $r$  (alterna-

Figure 11. SPM images show sample topography as a function of scanner voltage. Hysteresis correction transforms the images so they show topography as a function of scanner position. In PROCESS, the user selects two equivalent points  $f$  and  $r$  on the forward and reverse images, respectively. The software defines two second-order hysteresis curves to estimate the actual X position as a function of X voltage.



tively the software could determine these coordinates from a scanner calibration). The software warps the images such that  $f$  and  $r$  both transform to  $x = (f+r)/2$  and the edges ( $x = 0$  and  $x = x_{\max}$ ) remain unchanged. It does so by defining a second-order curve from the three points  $0, f$  or  $r$ , and  $x_{\max}$ , and making a lookup table according to the following equation (all X coordinates are measured in terms of data points):

$$x = \frac{1-b}{x_{\max}}(x')^2 + bx'$$

where:

$$b = \frac{x_{\max}(f+r) - 2d^2}{2d(x_{\max} - d)}$$

and:

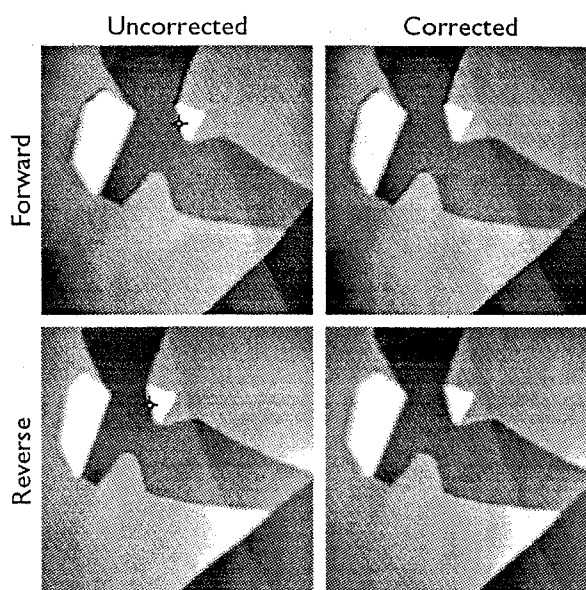
$x$  = old X coordinate

$x'$  = new X coordinate

$d = r$  to transform forward images

$d = f$  to transform reverse images

Figure 12. Example of hysteresis correction. The crosshairs show where I chose the points  $f$  and  $r$ . After hysteresis correction the forward and reverse images have a closer resemblance. The image is a 740 x 740 nm AFM scan of boronated pyrolytic graphite.



The algorithm creates a new image and scans through each point, determining its X coordinate  $x'$ . Using this coordinate the algorithm consults the lookup table and reads data from location  $x$  in the original image, interpolating when necessary. It then puts this data into the new image at the current location,  $x'$ . Figure 12 shows the effect of the algorithm.

If images scanned both bottom-to-top and top-to-bottom are available, the user can repeat the algorithm to remove hysteresis in Y.

Correlation diagrams show the relationship between two variables

PROCESS creates the scatterplots that I refer to as “correlation diagrams” by comparing two images. Any given pixel in the “X” image becomes the X coordinate of a point on the correlation

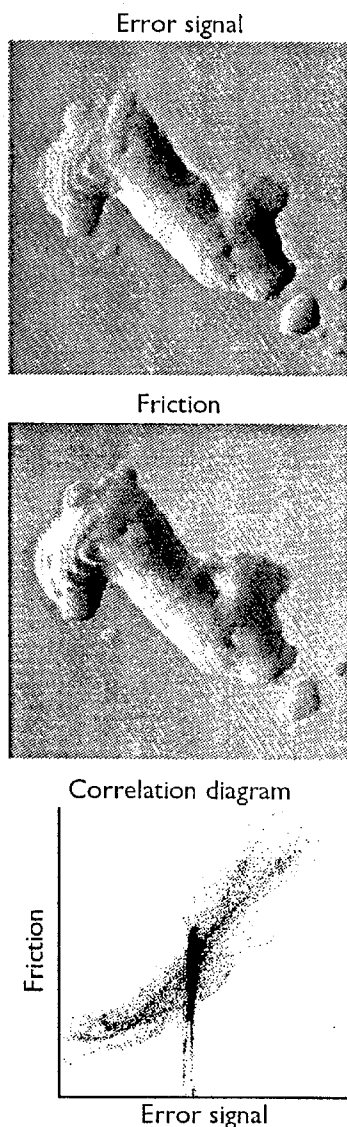


Figure 13. Simultaneous 765 nm x 765 nm AFM error signal and friction images of a deposit (possibly a broken-off supertip) on float-polished quartz, with a correlation diagram showing the dependence of friction signal on error signal. The nonlinear dependence may result from detector nonlinearity. The vertical streak results from the curvature of the friction image.

diagram. The corresponding pixel in the “Y” image (the pixel with the same location) becomes the Y coordinate. If the two images are identical, the correlation diagram shows a diagonal line. If the two images are completely uncorrelated Gaussian noise, the correlation diagram shows a diffuse ball of points.

Correlation diagrams can help the user analyze relationships between simultaneous imaging mechanisms such as topography and friction (figure 13).

PROCESS can derive adhesion force and sample hardness from AFM imaging spectroscopy data

PROCESS can extract adhesion and hardness images from force curve imaging spectroscopy stacks. I have placed the description of this capability along with the description of imaging spectroscopy in chapter 3.



## References

---

Ammeraal, L. (1986) *Programming Principles in Computer Graphics*.

John Wiley and Sons, Chichester

Dougherty, E.R. (1992) *An Introduction to Morphological Image*

*Processing*. SPIE Optical Engineering Press, Bellingham, WA

Hill, F.S. (1990) *Computer Graphics*. Macmillan Publishing Company, New York

Niblack, W. (1986) *Digital Image Processing*. Prentice/Hall International, Englewood Cliffs, N.J.

## Part 2

# Application

### *Introduction*

---

Part two describes progress toward understanding the tip-sample interaction in AFM

The presence of multiple imaging mechanisms and various artifacts complicates interpretation of friction and elasticity images. I have refined these tools, learning how to use them to study the tip-sample interaction.

I have also applied friction and elasticity imaging to the study of the tip-sample interaction. I have demonstrated that meniscus forces act laterally as well as normally, and that they vary with position. I have shown that, although imaging in water reduces interaction forces, water can also make biological samples more sensitive to force. Finally, I have demonstrated that, for robust biological samples imaged in air, tip shape presents a greater obstacle than tip-sample interaction forces to obtaining high-resolution images.

## Chapter 5. The lateral meniscus force

### *Introduction*

---

Chapter 5 describes aspects of the meniscus force revealed by lateral force images of atomic steps on graphite (Baselt and Baldeschwieler, 1992)

Using atomic steps on graphite as a generalized model for sample features, I have demonstrated that the meniscus force varies with sample curvature. Chapter 5 considers the relationship between contact force, cantilever deflection, and meniscus force, an important consideration when imaging biological molecules.

This chapter also presents other information obtained from lateral force studies of graphite surfaces: that boron impurities on boronated pyrolytic graphite may create local regions of high or low friction, and that lateral force images can reveal atomic structure more clearly than topographic images.

## Contact force and cantilever deflection

Contact force depends on both meniscus force and cantilever deflection

The fact that contact force (in combination with tip shape) limits AFM image quality makes it an important variable. Force microscopists often measure contact force by assuming that it is proportional to cantilever deflection. However, this method can produce erroneous results, since a number of forces affect the contact force without affecting deflection.

In air, the largest force of this type is the meniscus force (Weisenhorn et al., 1989; Burnham et al., 1991), which arises from the thin film of water vapor and other contamination adsorbed on the tip and sample. The meniscus force acts on the body of the tip and, discounting small sample elasticity effects, does not affect the deflection of a stationary cantilever.

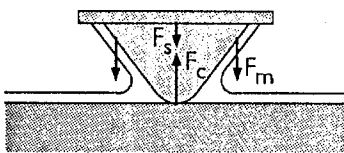


Figure 1. Forces acting on the tip.  $F_c$ : contact force.  $F_s$ : cantilever deflection force (spring constant times deflection).  $F_m$ : meniscus force. Positive forces act in the indicated directions. When the tip is not accelerating:

$$F_c = F_s + F_m$$

The feedback loop keeps  $F_s$  constant.

If a cantilever with zero deflection touches a sample in the absence of a meniscus, the contact force is zero. Now allow a meniscus to form, exerting a 10 nN force that pulls the tip and sample together. The contact force increases to 10 nN to balance the meniscus force, but deflection remains unchanged. Using deflection as a measure of contact force would, in this typical case, yield a figure that is much too low.

Contact force equals the sum of the cantilever deflection and meniscus forces (figure 1). Therefore, measuring contact force requires knowledge of both deflection and meniscus force.

Although force microscopists usually account for meniscus force by measuring it at one point in an image then subtracting it from all points in the image, if it varies from point to point, deflection will not accurately indicate contact force. Interpreting deflection images therefore requires care.

“Constant normal deflection” mode images do not actually exhibit perfectly constant normal deflection

I acquired all the images presented in this thesis using “constant normal deflection” mode. In this mode, a feedback loop monitors the normal deflection and tries to keep it constant by adjusting  $Z$  (see chapters 1 and 3). SCAN obtains a normal deflection image by recording deflection, and a topography image by recording  $Z$ . However, even with a perfect feedback loop normal deflection images still have features. Without fluctuations in normal deflection the feedback loop will not change  $Z$ , so if a normal deflection image shows no features, neither will the corresponding topography image. Faster feedback response and slower scan rates reduce the fluctuations in normal deflection, but these fluctuations never completely disappear.

Since the feedback loop algorithm generates  $Z$  by (among other elements; see chapter 3) integrating the normal deflection over time, the normal deflection at any given point in a scan is roughly proportional to the slope of topography at that point.

## Methods

---

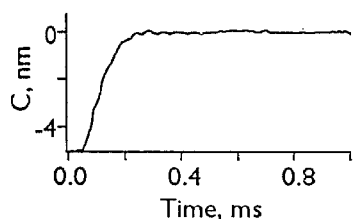


Figure 2. Typical step response curve showing settling within 0.25 ms. I acquired most of the scans in this thesis at a rate of 0.5 ms/data point.

I have adjusted scan parameters to make feedback loop errors negligible

The step response optimization described in chapter 3 is an essential aspect of the step shape measurements I present here. I have adjusted feedback parameters and scan rates such that, in step response measurements, normal deflection ( $C$ ) settled in the amount of time it took to acquire one data point in an image. Thus, step shape measurements are accurate to within one data point.

I have acquired both forward and reverse images

Each time the AFM scans one line, it moves the tip first from left to right, then right to left. Data acquired during the left-to-right trace becomes part of a “forward” image, while the right-to-left data forms the “reverse” image.

PROCESS automatically multiplies forward lateral deflection data by  $-1$  so that bright areas in both forward and reverse lateral deflection images represent regions of high tip-sample friction.

I have imaged both graphite and boronated graphite

For this study I have used freshly cleaved ZYA grade highly oriented pyrolytic graphite (HOPG) obtained from Union Carbide Coatings Service Corporation (Cleveland, OH) and boronated graphite (BPG) obtained from A.W. Moore, Union Carbide Coatings Service Corporation (Parma, OH).

BPG is a form of graphite in which boron atoms substitute for some (about 0.5% in the sample that I used) of the carbon atoms. The particular sample that I used has more steps and other defects than ZYA-grade HOPG.

Normal forces measured about 10 nN

Normal deflection forces measured about 10 nN, exerted downward on the sample. Normal contact forces measured about the same as normal meniscus forces, between 10 and 100 nN, as determined from force curves.

I have determined average lateral forces by subtracting uninverted forward lateral deflection images from the corresponding reverse lateral deflection images and finding the average value of the resulting “differential lateral deflection” images. I use this method so that I do not have to determine the lateral deflection zero point. Differential lateral deflection images also have other uses, as I will describe in chapter 6.

All force measurements rely on calculated values of cantilever spring constants, and are therefore only order of magnitude estimates.

I acquired some of the images in this chapter with “T” shaped cantilevers

I have used “T” shaped cantilevers from Park Scientific Instruments for some of the images below. No longer in production, these cantilevers had the same tips and consisted of the same materials as the “V” cantilevers described in chapter 2. They measured 100  $\mu\text{m}$  in length, 20  $\mu\text{m}$  in width, and 0.6  $\mu\text{m}$  in thickness. They had spring constants of 0.17 N/m for normal deflection and 53 N/m for lateral deflection of the end of the tip.

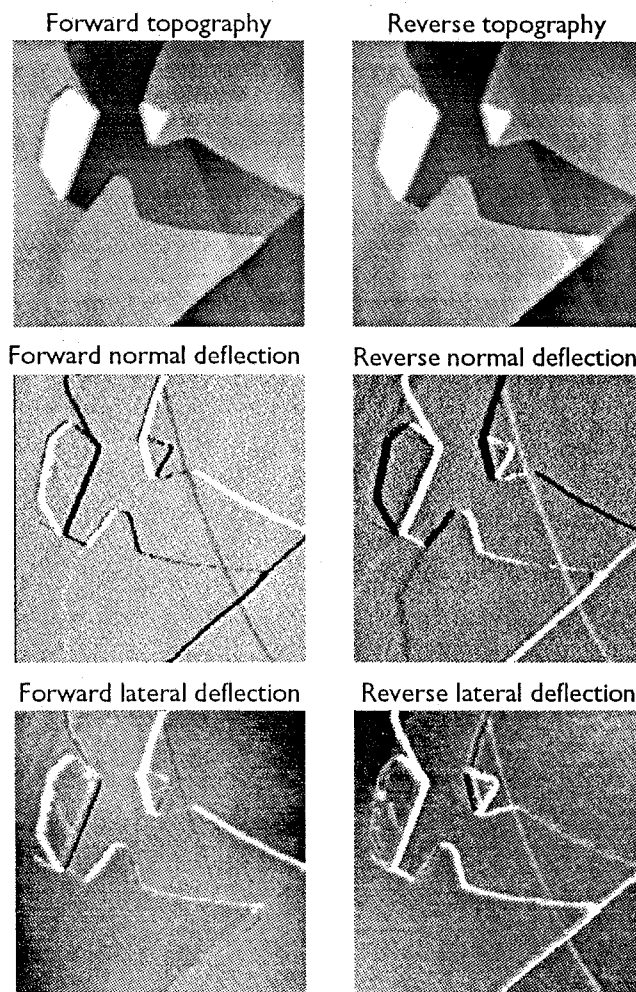
### Observations when scanning over steps

Several lateral force anomalies appear when scanning over steps

Figure 3 shows simultaneous topography, normal deflection, and lateral deflection images of an area in which peeled-back layers of graphite have formed steps 0.3-5.0 nm high. The normal meniscus force and the lateral deflection force both measure around 50 nN when scanning over flat areas. As expected, lateral



Figure 3. A 740 nm scan of BPG. I acquired all six images simultaneously. The diagonal line running through the entire scan is a single-atom-high step 0.4 nm tall. The larger steps measure 2–5 nm tall. Scan rate 4 lines/second. Corrected for X hysteresis; topography images paraboloid subtracted; lateral deflection images flattened.



deflection increases when scanning up steps. However, lateral deflection also increases when scanning down steps at a small angle to the scan direction. This effect grows stronger after exposing the graphite to air for several hours, so surface contamination probably causes it.

Figure 4 shows graphs of topography, normal deflection, and lateral deflection as the tip scans up and down a 5.0 nm step seen in figure 3. I have numbered the features of this graph in order of occurrence:

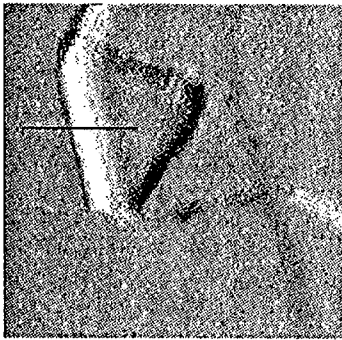
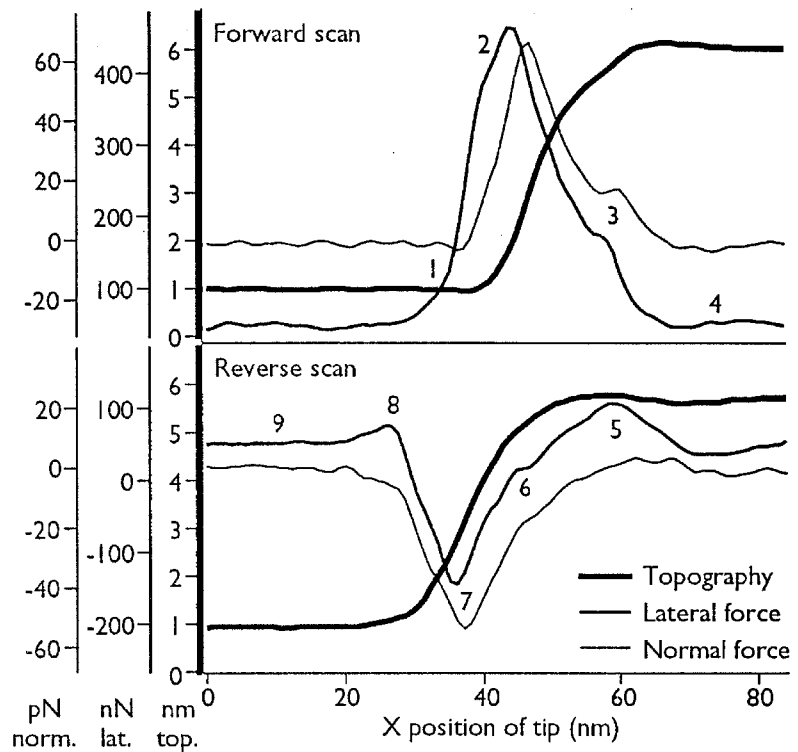


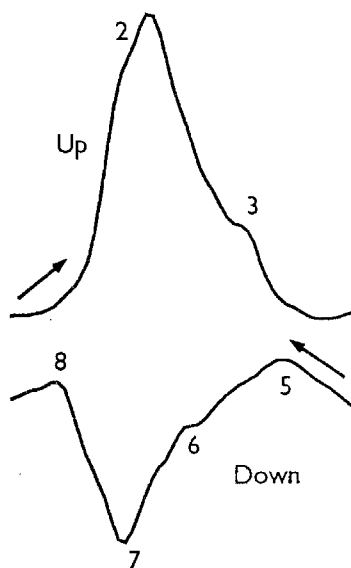
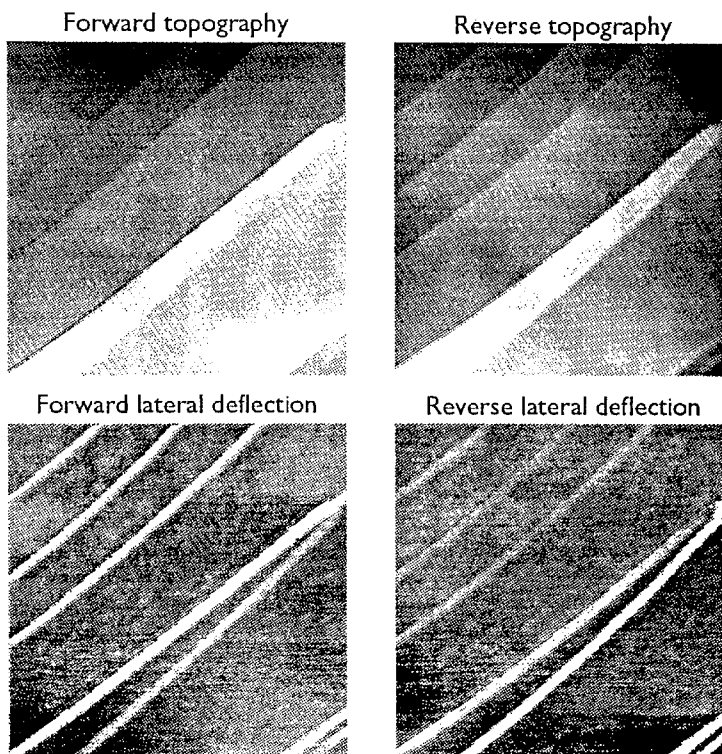
Figure 4. A 230 nm normal deflection scan of BPG, part of the area shown in figure 1, with graphs of topography, normal deflection force and lateral deflection force as the tip scans up and down a 5.0 nm tall step. The graphs come from the part of the image marked with a line, and each have 90 data points. Scan rate 4 lines/second. Topography and lateral deflection images leveled; all images smoothed 3x. The zero point on the topography scale is arbitrary.



*Scanning up a step:*

1. Lateral deflection increases.
2. Normal deflection increases and the feedback compensates by pulling the cantilever away from the sample, producing the rise in topography. Lateral deflection peaks, then normal deflection peaks.
3. Near the top of the step a shoulder appears, first in lateral deflection, then in normal deflection.
4. Normal and lateral deflection return to normal as the step comes to an end.

Figure 5. 820 x 820 nm image of BPG. Steps appear bright in both the forward and reverse lateral deflection images. X hysteresis corrected; horizontal and vertical line leveled.



The lateral deflection traces from figure 4, reproduced for reference.

*Scanning down a step:*

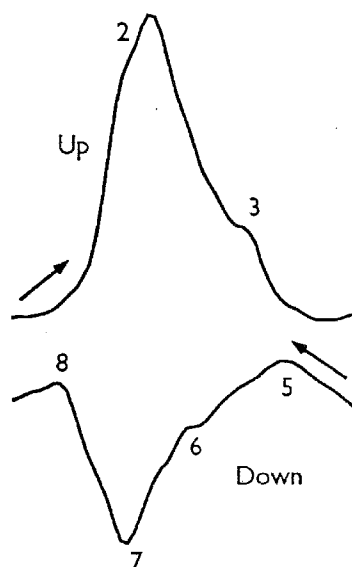
5. Lateral deflection increases before the tip encounters the step.
6. A shoulder appears in the lateral deflection trace soon after the start of the step.
7. First normal, then lateral deflection goes through a minimum near the bottom of the step.
8. Just after the step ends, lateral deflection increases momentarily.
9. Normal and lateral deflection return to normal.

Although some of the above effects (phases 3 and 6 in particular) look as if they might be the result of irregular tip or step shape, I have observed all of them with different tips and samples,

on both HOPG and BPG, and on both left-facing and right-facing steps.

On the other hand, not all of the above features appear in all images. Most notably, the phase 5 lateral deflection peak often obscures the phase 7 valley when scanning down a step. When this occurs, all steps appear bright in the lateral deflection image (figure 5). In addition, the phase 3 peak in lateral deflection sometimes appears quite high or broad. Its size has a general correlation with that of phase 5. As in the reverse scan, the phase 3 peak can eclipse phase 2, sometimes blending into it so only one peak appears. In these cases the only way to identify the type of the peak is from its position relative to the step.

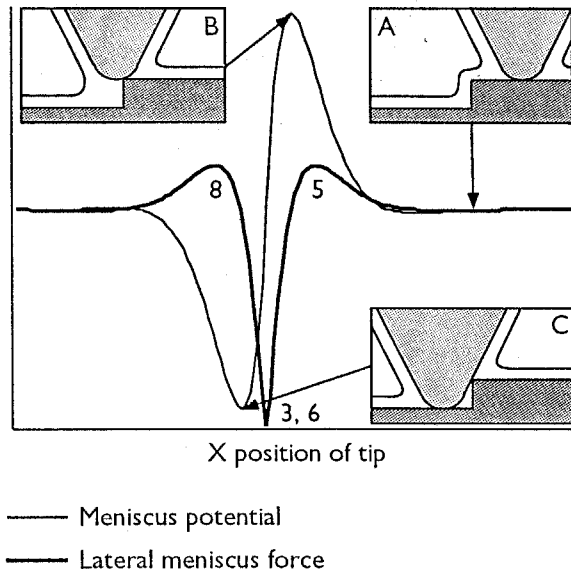
#### The meniscus can account for the observed lateral force behavior over steps



The lateral deflection traces from figure 4, reproduced for reference.

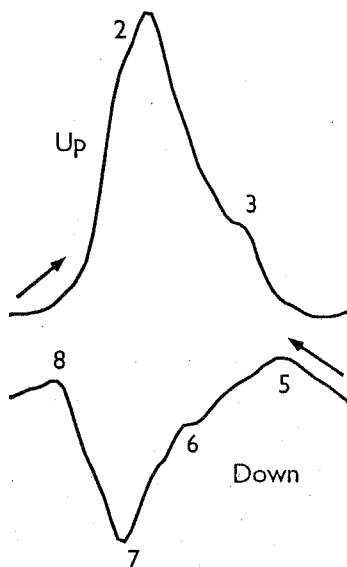
Because of the downward force it exerts on the sample, the tip should resist scanning up a step; lateral contact force should increase when the tip encounters a step up. Conversely, lateral contact force should decrease when scanning down a step. A prominent lateral deflection maximum does, in fact, appear when scanning up a step (phase 2), and a minimum appears when scanning down a step (phase 7), but in addition a number of other features appear.

Figure 6. A hypothesis to describe the observed lateral meniscus forces while scanning down a step. A) reference meniscus potential; B) highest meniscus potential; C) lowest meniscus potential. The lateral force curve for reverse scans, in bold, is the derivative of the potential taken right to left. For comparison with the forward lateral deflection curve in figure 4, invert the lateral force curve (i.e., take the derivative from left to right). Peaks 5 and 8 represent lateral force pulling the tip to the right, peak 3, 6 to the left, regardless of scan direction.



Lateral meniscus forces probably cause these features — phases 3, 5, 6, and 8 of figure 4 — since they have a similar magnitude (on the order of 10 nN) to the normal meniscus force.

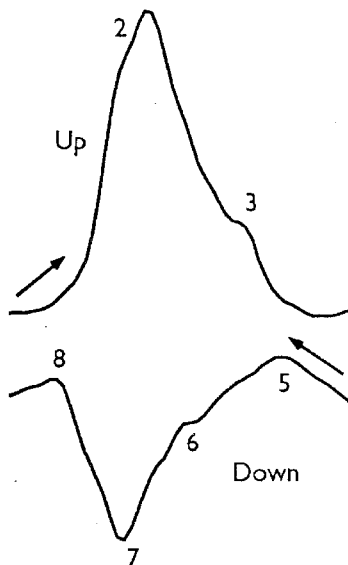
The model illustrated in figure 6 accounts for the observed pattern of lateral meniscus forces. Like any other attractive force, the meniscus force creates an interaction potential that increases with increasing tip-sample separation. Inset A of figure 6 shows the tip before it encounters the step. For the tip to move to position B, its left hand edge must pull further away from the sample, thereby increasing the meniscus potential. Conversely, in position C the right hand edge of the tip lies closer to the sample, lowering the potential. Thus, the thin line in figure 6 traces out the meniscus potential as a function of tip position,  $X$ . Differentiating this curve with respect to  $X$  yields the lateral meniscus force, shown with a bold line. Notice that  $X$  increases in the direction of scan, going right-to-left.



The lateral deflection traces from figure 4, reproduced for reference.

Figure 6 indicates that meniscus forces should cause two maxima and one minimum in lateral force when scanning down a step. This hypothesis therefore explains the high-low-high pattern of lateral deflection when scanning down a step, although the low probably corresponds to the small phase 6 shoulder rather than the larger phase 7 minimum. The pattern of lateral forces encountered when scanning up a step suggests that phase 7 probably results from contact forces.

Since PROCESS has inverted the forward lateral deflection curve in figure 4, analysis of the forward scan requires also multiplying the meniscus force curve in figure 6 by  $-1$ . Thus, the predicted curve has a low-high-low pattern, while the experimental curve has a double peak, a high-high pattern. Since the height of phase 3 tends to correlate with that of the meniscus-induced phase 5, phase 3 probably results from meniscus force, leaving phase 2 to result from contact force. Based on their positions relative to the step, phase 2 corresponds to phase 7 and phase 3 to phase 6. The two meniscus force minima corresponding to phases 5 and 8 do not show up in the forward scan, probably because the phase 2 maximum eclipses them.



The lateral deflection traces from figure 4, reproduced for reference.

Lateral meniscus forces can affect the imaging of biological samples

Lateral meniscus forces could play an important role when trying to image adsorbed molecules (DNA, polymers, proteins) in

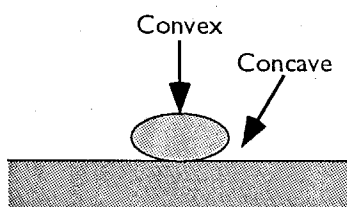


Figure 7. Schematic cross-section of a collagen fibril or other structure on a flat substrate with convex (hill) and concave (valley) regions indicated.



Figure 8. Differential lateral deflection image derived from the images in figure and showing increased meniscus forces at step edges. Note slight mottling.

air. Unlike normal and lateral contact forces, pulling up on the cantilever does not reduce lateral meniscus forces. Although they exert force over a wide area and so cannot cause much sample distortion, lateral meniscus forces could still pull molecules off the substrate. In fact, since lateral meniscus forces often exceed contact forces, the lateral meniscus force probably causes most tip-induced sample detachment.

Normal meniscus forces should also vary when scanning over a step: normal meniscus forces drop off with increasing tip-sample separation, and I have already described how tip-sample separation varies when scanning over a step. Referring back to figure 5, the normal meniscus force should be proportional to the meniscus potential. Since the feedback loop keeps the sum of normal meniscus and contact forces roughly zero, the contact force follows the same trend.

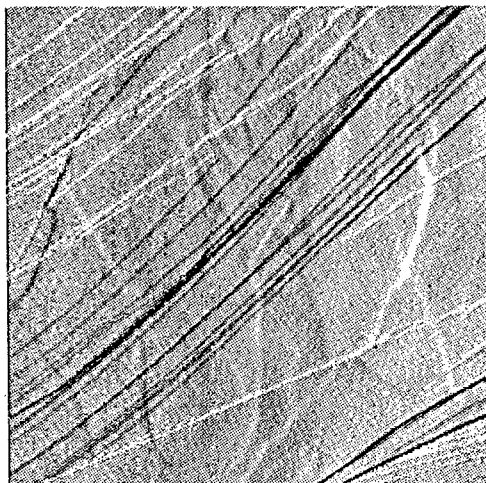
Generalizing, normal and lateral forces should increase whenever the tip is at concave features and decrease at convex features (figure 7). In chapter 6, I use a differential lateral deflection image and force spectroscopy data to reveal increased adhesion at the edges and between the bands of a collagen fibril. The maximum contact force in these areas exceeds the estimated value by a factor of two. Similarly, differential lateral deflection images of BPG indicate increased meniscus adhesion at steps (figure 8; see chapter 6 for an explanation of how differential lateral deflection images relate to tip-sample adhesion).

Figure 9. 200 x 200 nm STM image of BPG showing boron impurities. Image taken by Michael Youngquist.

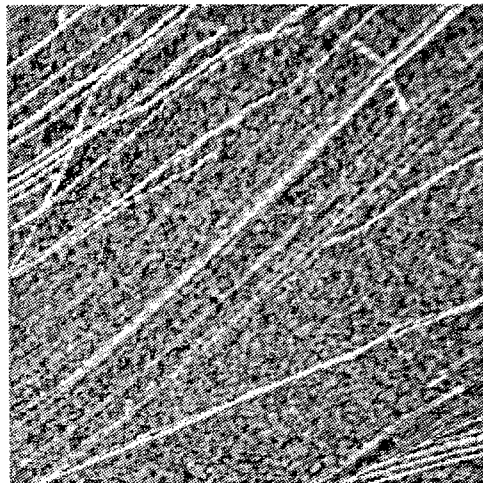


Figure 10. 1.6 x 1.6  $\mu\text{m}$  AFM image of BPG showing mottling. Taken with a supertip. X hysteresis corrected.

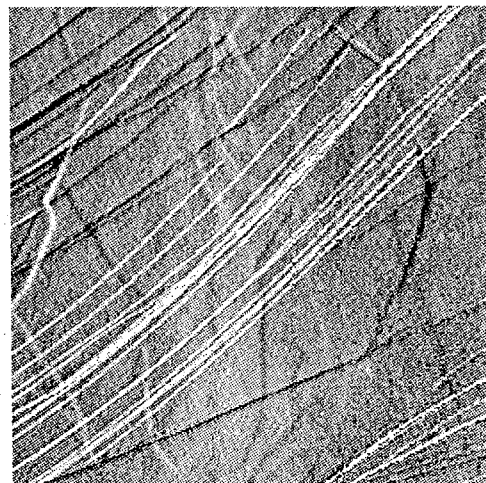
Forward normal deflection



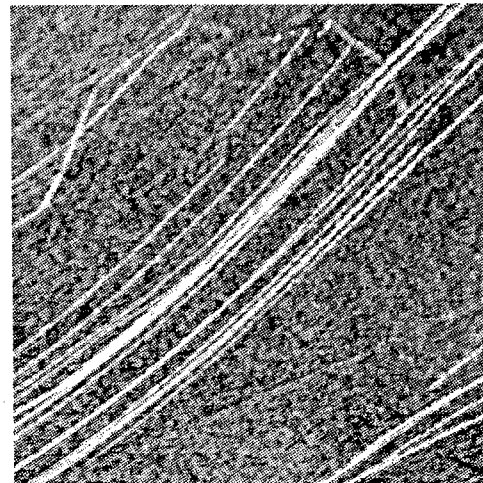
Forward lateral deflection



Reverse normal deflection



Reverse lateral deflection





### *Effect of boron impurities*

---

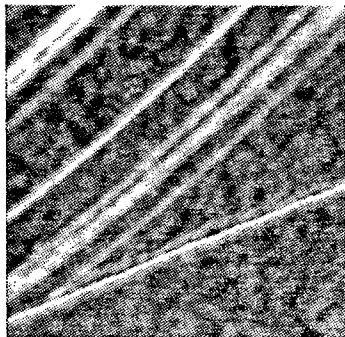


Figure 11. 570 x 570 nm differential lateral deflection image of BPG; same area as the lower-left corner of figure 8. Taken with a supertip.

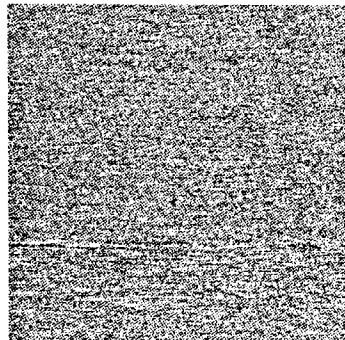


Figure 12. 700 x 700 differential lateral deflection of HOPG. The supertip used to acquire this image was able to reveal BPG mottling immediately before and after this image.

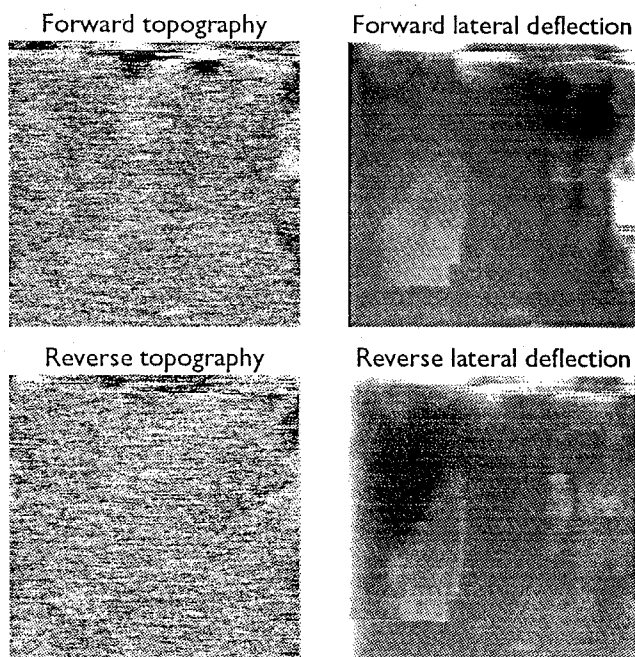
Boron impurities may have distinctive frictional characteristics

In STM images of BPG (figure 9), the reactive boron atoms (Hennig, 1965) show up as bright spots, and they might likewise create high-friction regions that show up in lateral deflection AFM images.

AFM scans of BPG surfaces sometimes exhibit a distinctive mottling pattern in the lateral force image (figure 10). This mottling does not correlate with any topographic feature, and only appears on BPG (figure 11, figure 12). Thus, as expected, the boron impurities may create regions of high (or possibly low) friction. The individual light areas probably represent clusters of boron atoms rather than single atoms. Alternatively, the high frequency of crystal imperfections in BPG may somehow cause the mottled appearance.

I have also observed larger, more well-defined regions of high lateral force (figure 13) that have no correlation to topography. These regions seem to have a different origin than BPG mottling, since they appear on both BPG and HOPG. Their polygonal shape indicates they may result from imperfections in the grain structure.

Figure 13. A 300 x 300 nm scan of HOPG. The lateral deflection images show "sticky spots" not visible in the topography images. Scan rate 4 lines/second, "I" cantilever. Corrected for X hysteresis. Topography images horizontal and vertical flattened; lateral deflection images leveled.

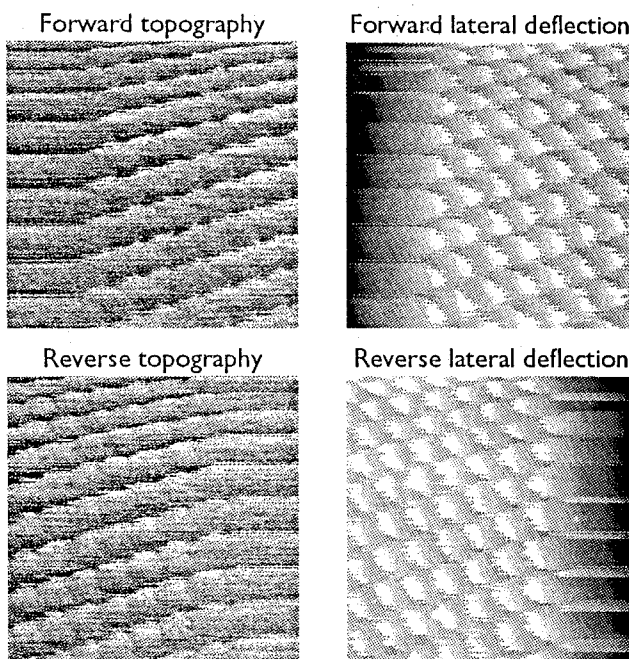


### Atomic resolution lateral deflection images

Lateral deflection images show atomic corrugation more clearly than topography images

I have acquired lateral deflection images showing the atomic corrugation of graphite (figure 14; see also chapter 1). These strongly resemble an image published by Mate et al. (1987), although the tip-sample interaction forces in the images from AFM-3 are approximately two orders of magnitude smaller. The atomic corrugation measures 250 pm in topography and 10 nN in lateral force. The average lateral force outside smeared regions

Figure 14. A 2.5 nm scan of HOPG. The smearing at the sides of the images is due to the cantilever twisting as lateral force builds up. Similar scans taken with "V" cantilevers do not show this smearing. Scan rate 10 lines/second, "I" cantilever. Topography images flattened.



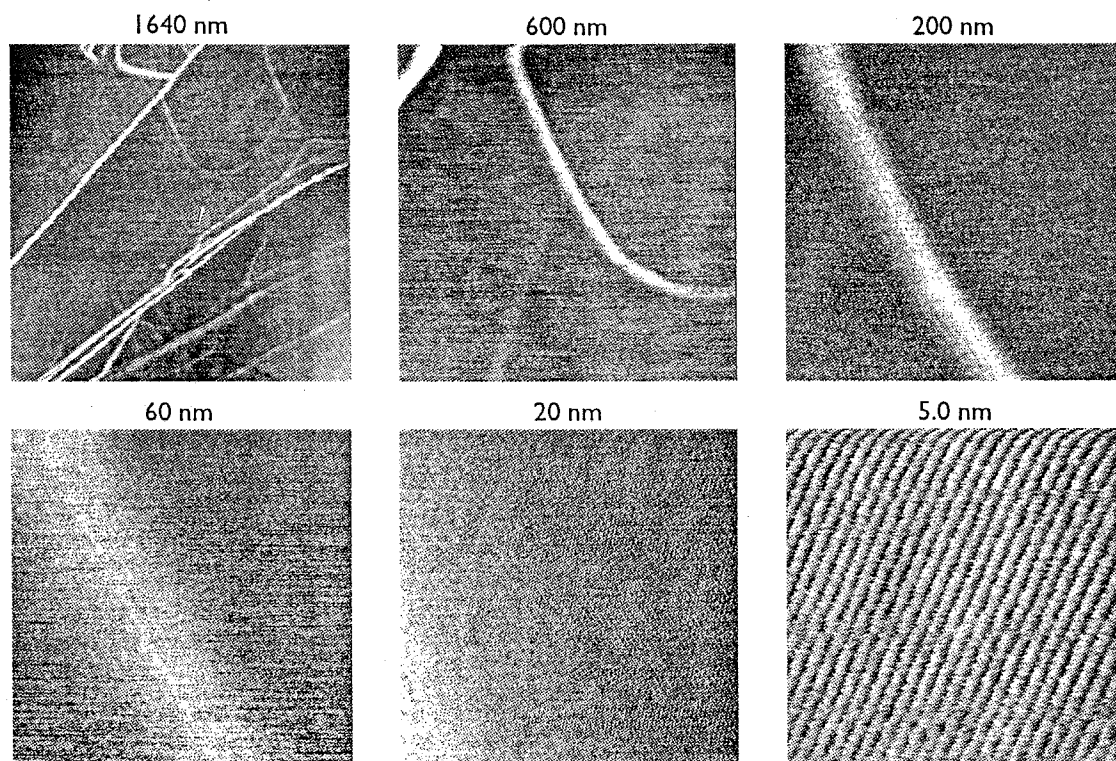
at the sides of the images is 30 nN. The slips in lateral force occur about 100 pm to the right of the peaks in topography in both the forward and reverse scans. Torsional cantilever warping may cause this anomaly, since atomic resolution images taken with "V" cantilevers show the lateral deflection and topography maxima in the same locations.

With both "I" and "V" cantilevers, atomic rows frequently appeared in lateral deflection images but not in normal deflection or topography images. Even in figure 14, where atomic corrugation appears in topography, it appears more clearly in the lateral deflection images. This phenomenon may arise from the high optical lever gain associated with lateral tip displacement (Meyer and Amer, 1990).

Tip-sample contact occurs over a broad region at the end of the tip

Atomic corrugation probably does not result from the interaction of a single atom on the tip with the atoms of the sample. Rather, a large region of the tip probably touches the sample. The apparent width of atomic steps on graphite indicates that most AFM tips have tip radii of over 30 nm (measured by dividing the square of the apparent step width by the step height), and frequently over 100 nm. Additionally, I have imaged graphite atoms within the region covered by a step (figure 15). Thus, the graphite and the tip flex enough that the tip touches both the upper and lower levels of the step, but the tip can still obtain atomic resolution over the entire image.

Figure 15. A series of lateral deflection images zooming in on a step. Atomic corrugation appears within the step. All images horizontal and vertical line leveled.



## References

---

- Baselt, D.R. and Baldeschwieler, J.D. (1992) Lateral forces during atomic force microscopy of graphite in air. *J. Vac. Sci. Technol. B* 10(5), 2316–2322
- Burnham, N.A., Colton, R.J., and Pollock, H.M. (1991) Interpretation issues in force microscopy. *J. Vac. Sci. Technol. A* 9(4), 2548–2556
- Flennig, G. (1965) Diffusion of boron in graphite. *J. Chem. Phys.* 42(1), 1167–1172
- Mate, C.M., McClelland, G.M., Erlandsson, R., and Chiang, S. (1987) Atomic-scale friction of a tungsten tip on a graphite surface. *Phys. Rev. Lett.* 59(17), 1942–1945
- Meyer, G. and Amer, N.M. (1990) Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. *Appl. Phys. Lett.* 57(20), 2089–2091
- Weisenhorn, A.L., Hansma, P.K., Albrecht, T.R., and Quate, C.F. (1989) Forces in atomic force microscopy in air and water. *Appl. Phys. Lett.* 54(26), 2651–2653

## Chapter 6. Contrast mechanisms in lateral force microscopy

### *Introduction*

---

Chapter 6 describes progress in characterizing lateral force microscopy

At the present time, only a few groups have presented simultaneous lateral force microscopy (LFM) and atomic force microscopy images (Meyer and Amer, 1990; Overney et al., 1992). Thus, because of its novel nature I have found it worthwhile to characterize the contrast mechanisms that contribute to optical lever-based lateral force microscopy. Chapter 5 discusses those contrast mechanisms that might be considered artifacts — that is, all mechanisms except adhesion between the tip and sample at the point where they touch.

Familiarity with the concepts of meniscus, contact, and deflection force presented in chapter 5 will greatly assist a reading of this chapter.

Assessment of lateral force microscopy images requires both forward and reverse scan acquisition

As this chapter will make apparent, the comparison of forward and reverse scans is essential to the analysis of LFM images. Although somewhat less vital to the interpretation of topographic images, in this task it can help with the correction of hysteresis and the identification of various artifacts.

## *Methods*

---

The imaging mechanisms described here appear both on scanned-cantilever and on scanned-sample instruments

All the contrast mechanisms that I discuss appeared both on AFM-3 and on a commercial scanned-sample AFM with interchangeable 1  $\mu\text{m}$  and 75  $\mu\text{m}$  scanners (Topometrix, Inc., Santa Clara, CA). Therefore, none of these mechanisms arise from the scanned-cantilever architecture of AFM-3. However, both instruments use optical lever detection, which very likely causes some of the artifacts I have observed.

I have imaged all samples in air unless otherwise indicated, at scan rates of 4–5 lines per second, with V-shaped, pyramidal-tipped silicon oxynitride cantilevers from Park Scientific Instruments (Sunnyvale, CA). PROCESS automatically multiplies forward lateral force images by  $-1$  so that areas of high friction appear bright in both the forward and reverse images.

I estimate lateral forces using calculations of cantilever  
responsivity

Table 1. Responsivity ratios  
for bending modes of the  
standard 0.6  $\mu\text{m}$  thick PSI  
cantilevers

type	$R_n/R_t$	$R_b/R_t$
<b>Length 100 <math>\mu\text{m}</math></b>		
wide	30.7	0.463
narrow	36.0	0.540
<b>Length 200 <math>\mu\text{m}</math></b>		
wide	70.9	0.498
narrow	81.8	0.574

$R_n$  = normal deflection  
responsivity (deflection per  
unit force);  $R_t$  = torsional;  $R_b$   
= buckling. From AFM design  
table 6, chapter 2 appendix.

I estimate lateral forces by combining force curve measurements  
of the AFM's sensitivity to normal cantilever deflection (see  
chapter 3) values of the AFM:LFM response ratio that I have  
calculated as described in the appendix to chapter 2 (table 1).

For example, when using the wide, 100  $\mu\text{m}$  long cantilever,  
force curves show that AFM-3 has a normal sensitivity  $R_n$  of  
about 270 mV (photodiode signal, indicating cantilever deflec-  
tion) per nN (normal cantilever deflection force). For this  
cantilever, the AFM:LFM response ratio  $R_n/R_t = 30$ , so the  
lateral force sensitivity  $R_t$  is about 9 mV/nN.

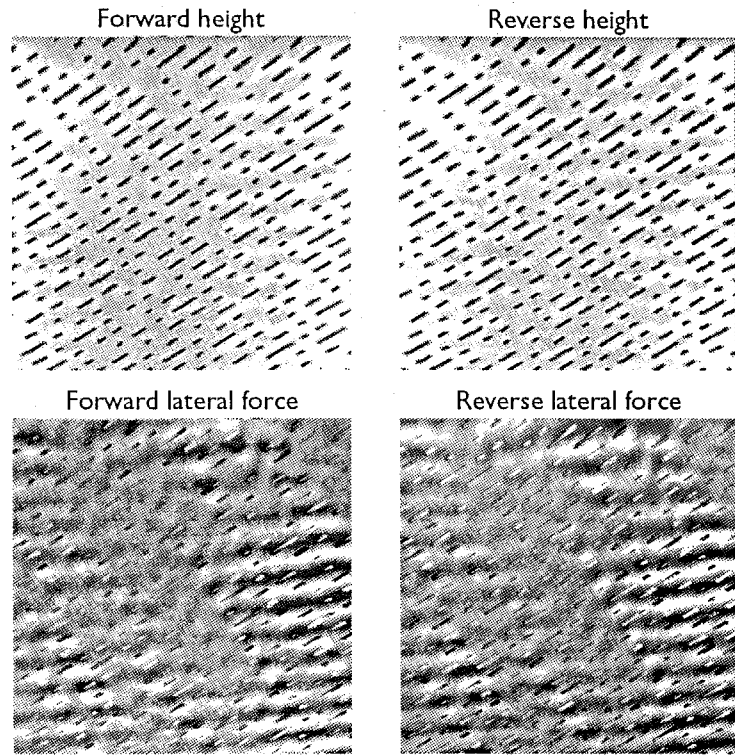
## Coupling with X, Y, and Z

### Apparent lateral force varies with scanner position

Large-scale lateral force images often show a periodic banding  
structure (figure 1). Closer inspection (figure 2) reveals that the  
banding depends not only on lateral position (XY) but also on  
sample height (Z). Thus, topographic features sometimes appear  
to have distinctive frictional characteristics: in the forward lateral  
force image of figure 2, the upper level of the compact disk  
appears to have a higher coefficient of friction than the lower.  
Comparison with the reverse image, in which the opposite is true,  
demonstrates that this frictional feature is an artifact. Since



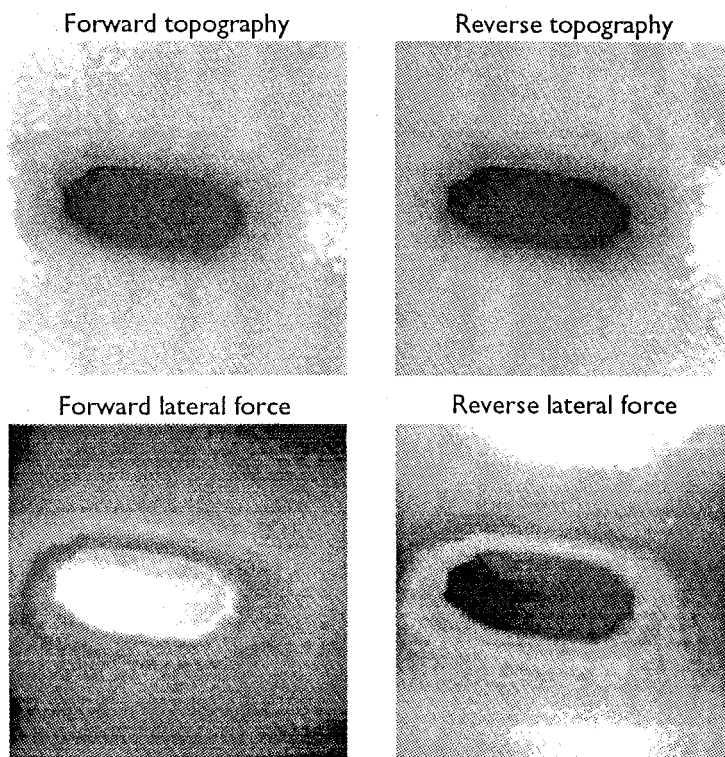
Figure 1. Example of XY coupling (ripples in lateral force images). Note that the topography image also shows some rippling.  $37.5 \times 37.5 \mu\text{m}$  image of a compact disk showing the pits that encode data. X hysteresis corrected. Taken with the Topometrix  $75 \mu\text{m}$  scanned-sample AFM.



PROCESS has multiplied the forward image by  $-1$ , the raw LFM signal at the bottom of the pit always exceeds that at the top; that is, the LFM signal depends on sample height. In both the forward and reverse portions of this particular image, the frictional force appears to differ by  $60 \text{ nN}$  over a height variation of  $160 \text{ nm}$ .

The  $2.1 \mu\text{m}$  scan range of AFM-3 is too small to observe significant XY coupling. I will therefore discuss only Z coupling, although my conclusions also apply to XY coupling.

Figure 2. Z coupling.  $1.0 \times 1.0 \mu\text{m}$  image of a compact disk taken under water with AFM-3. X hysteresis corrected.



XYZ coupling results from laser light reflected from the sample

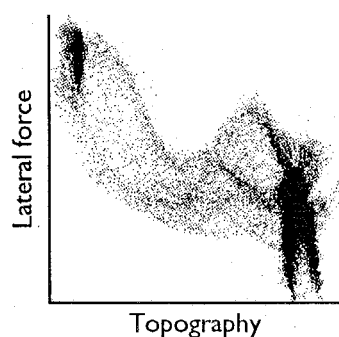
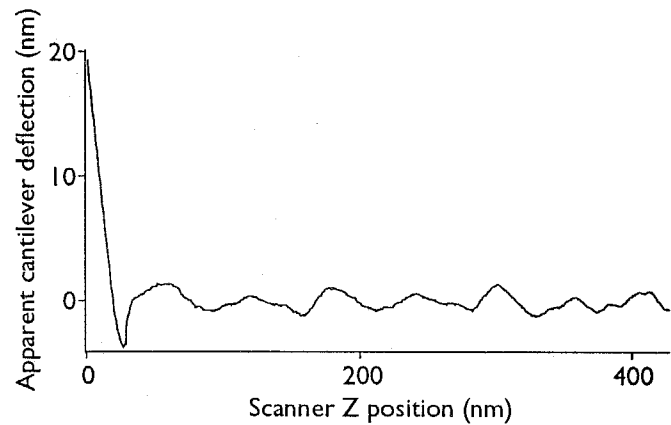


Figure 3. Correlation diagram derived from the forward images of figure 2 showing a roughly sinusoidal dependence of lateral force on sample height.

Since I acquired figure 2 in constant normal deflection mode, the height coupling observed cannot result from crosstalk between normal and lateral deflection. It seems likely that interference of light reflected from the cantilever with stray light reflected from the sample causes height coupling. A height vs. lateral force correlation diagram (figure 3) reinforces this hypothesis by revealing a sinusoidal dependence of apparent lateral force on height. If the hypothesis is true, large ( $>1 \mu\text{m}$ ) lateral deflection images might also exhibit coupling with local reflectance.

Figure 4. Force–distance curve showing 2 nm “rippling” of apparent normal cantilever deflection. AFM-3 with a wide 100  $\mu\text{m}$  cantilever under water. The sample was a compact disk.



Associated with height coupling is a ripple appearing in the attractive portion of normal deflection vs. Z force curves (figure 4). This ripple probably results from interference effects, since its frequency is too low to result from oscillations of the cantilever.

Users can detect and remove XYZ coupling by comparing forward and reverse images

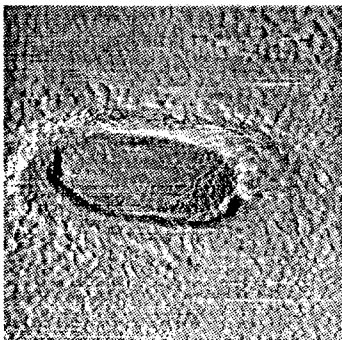
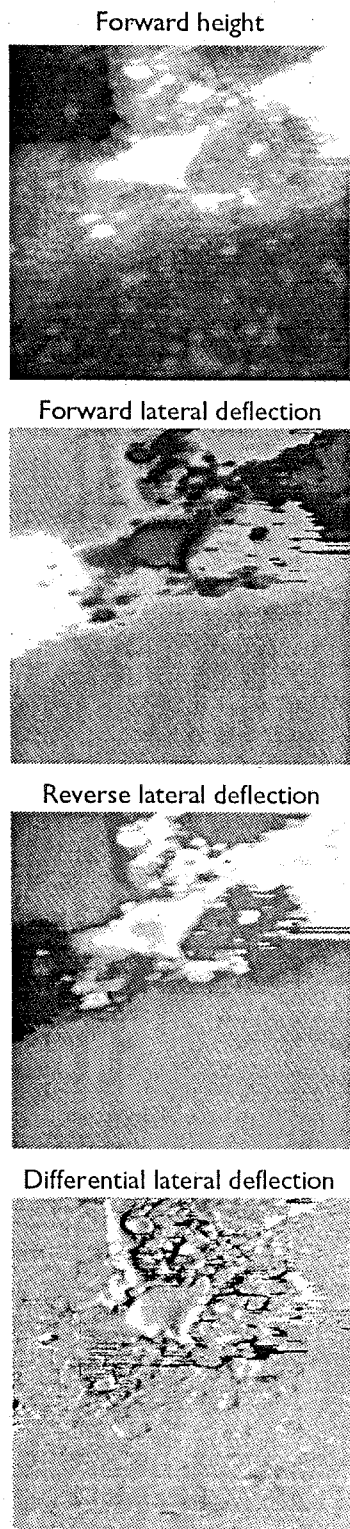


Figure 5. Differential lateral deflection image generated from the images in figure 2.

Force microscopists can detect and remove the effects of XYZ coupling by acquiring both the forward and reverse LFM images. Regions that actually do have high or low friction should appear the same in both images. Height-dependent features that invert between the two images are probably artifacts. Users can remove height coupling by adding the forward and reverse images, assuming the forward image has already been multiplied by  $-1$ , as demonstrated in figure 5. In this case, the resulting “differential lateral deflection” image shows an incomplete ring around the



CD pit, probably a result of imperfect hysteresis correction. The ring has a 5.2 nN magnitude, twice the background friction.

Figure 6, an image of cast iron, demonstrates the potentially deceptive nature of Z coupling. The high regions correspond to crystallites of the graphite phase that could have lower friction than the surrounding metallic surface. In fact, the forward LFM image appears to confirm this hypothesis: over a 4.5 nm rise the friction signal changes by 30 nN. However, comparison with the reverse image reveals that almost all the contrast results from height coupling. The differential lateral deflection image (figure 6) reveals some 7 nN features but little difference between the graphitic and metallic regions.

### Coupling with the error signal

Coupling of normal deflection into lateral deflection is insignificant

A number of mechanisms can cause normal deflection to couple into the lateral deflection signal. However, since the feedback keeps normal deflection constant while scanning, this coupling generally fails to produce significant artifacts.

Figure 6. Z coupling creates lateral deflection features that could appear real before comparison of forward and reverse data. 1.0 x 1.0  $\mu\text{m}$  image of cast iron from AFM-3.

Coupling of lateral deflection into the error signal produces artifacts in topography images

Conversely, lateral deflection can couple into the normal deflection signal. This effect warrants discussion since its impact on measured topography can influence the analysis, though not the appearance, of LFM images.

Lateral deflection coupling can have either optical or mechanical origins. Optical coupling occurs if the cantilever is rotated with respect to the photodetector; a  $6^\circ$  rotation causes 10% crosstalk. Mechanical coupling (den Boef, 1991) occurs when the cantilever is rotated with respect to the scan direction or when steps oblique to the scan direction pass under the tip. In such cases friction acts along the length of the cantilever, causing it to buckle (figure 7).

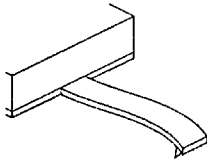
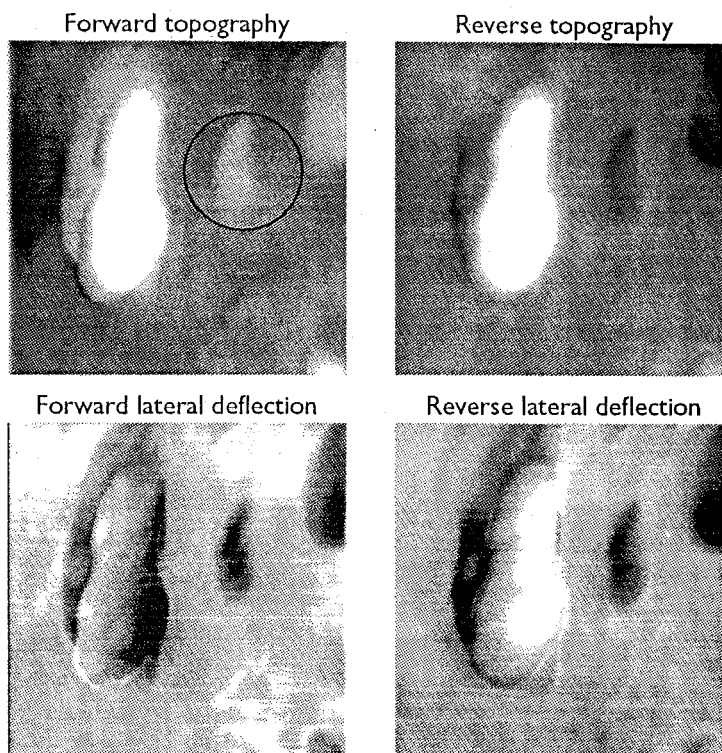


Figure 7. Buckling deflection of a cantilever.

As an example of friction coupling, figure 8 shows an image of a glass slide coated with a polymer solution. Several polymer aggregates with about 40% lower friction than the surrounding glass appear in the image. In the forward topography image one aggregate appears as a 3.0 nm high lump, but in the reverse image it appears as a 1.0 nm deep pit. It is reasonable to assume that the aggregate actually measures 1.0 nm high and, depending on scan direction, friction coupling either adds or subtracts 2.0 nm from its apparent height.

Two other groups (Thundat et al., 1992; Zenhausern et al., 1992) have also observed friction coupling, causing the height of

Figure 8. Coupling of lateral deflection into the error signal makes the apparent height of features dependent on their coefficient of friction. Note that the circled feature, which has low friction, appears high in the forward topography image and low in the reverse. 750 x 750 nm image of 1, 4-poly(cis-5,6-Bis(trimethylsiloxy)-1,3-cyclohexadiene) on glass. Taken with AFM-3. X hysteresis corrected. Polymer prepared by Douglas Gin.



DNA strands adsorbed on mica to vary with apparent humidity (i.e., meniscus force) or scan direction.

The calculated buckling:torsional response ratio fails to account for observed error signal coupling

Calculations based on rectangular cantilevers (table 1) show that the buckling mode is about half as sensitive to frictional force as the torsional mode. Analysis of figure 8, however, reveals 1:1 coupling of lateral deflection into the normal deflection signal; i.e., for every millivolt the lateral deflection signal changes, topography changes by an amount corresponding to one millivolt

of normal deflection signal. Thus if the mechanical coupling alone caused the observed coupling, the scan would have to be rotated about  $70^\circ$  from its correct orientation. Since I estimate that the rotation does not exceed  $10^\circ$ , cantilever defects or some other, unknown factors clearly play an important part.

The friction coupling in figure 8 is unusually but not uniquely high. The strength of friction coupling varies between 1:1 and unmeasurable.

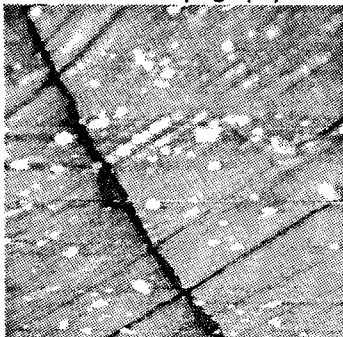
**Users can detect and remove lateral force coupling by comparing forward and reverse topography images**

The fact that the polymer aggregates appear too tall in one scan direction and too short in the other indicates that the normal force deviated from the setpoint in those areas. Averaging the forward and reverse topography images together eliminates the artifact. Subtracting the LFM data multiplied by a user-determined constant also removes friction coupling.

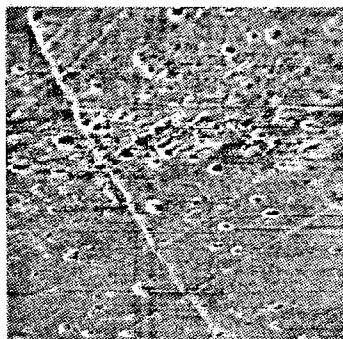
Users can also address friction coupling by eliminating its cause. Mechanically or electronically rotating the photodiode can eliminate optical coupling, while rotating the scan direction until the forward and reverse topography images match can eliminate mechanical coupling (an X-Y oscilloscope display helps).

## Meniscus effects

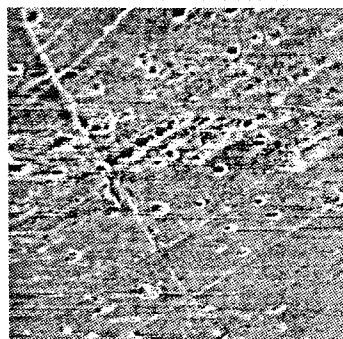
Forward topography



Forward lateral force



Reverse lateral force



When imaged in air, convex or concave features have distinct frictional characteristics

When imaging in air, meniscus forces can increase lateral forces when scanning both up and down steps (see chapter 5). The meniscus force tends to attract the tip into concave features and pull it away from convex features. It often makes all steps on the sample appear bright.

As an example of meniscus effects, figure 9 shows an image of cast iron in which many small lumps (presumably graphite crystallites) appear. An area of high lateral force appearing as a bright ring or horseshoe surrounds each crystallite. A user might interpret these areas, which have an apparent coefficient of friction about twice as high as the average, as a chemically distinct region at the interface of the graphitic and steel phases. However, they actually result from meniscus forces. These features tend to disappear at higher magnifications.

When imaging in air, any LFM features associated with convex or concave topography could result from meniscus forces. I know of no way to reliably remove these effects.

Figure 9. Regions around bumps exhibit high friction because of meniscus forces.  $1.8 \times 1.8 \mu\text{m}$  image of cast iron from AFM-3. X hysteresis corrected; topography horizontal line leveled and paraboloid subtracted; lateral force horizontal and vertical line leveled.



Meniscus-like effects may occur even in the absence of a meniscus

Meyer and Amer (1990) have shown that lateral force does not change significantly when scanning down steps of NaCl (001) in ultra-high vacuum, although they observe the expected increase when scanning up steps. This behavior probably results from attractive van der Waals forces acting in a manner similar to meniscus forces.

### *Slope-dependent contact force effects*

---

Lateral force images respond to slope

Although force microscopists already appreciate the slope dependence of lateral force, I describe it here for the sake of completeness.

Because the tip exerts a downward force on the sample (the “contact force”), lateral force increases when scanning up a slope. Conversely, lateral force decreases when scanning down a slope. A user can easily identify these effects by comparing friction and topography images, and can partially remove them by creating a differential lateral deflection image.

Figure 10 shows typical slope-dependent behavior. A correlation diagram (figure 11) shows that lateral force has a roughly first-order dependence on slope. Detector nonlinearities probably

Figure 10. It is well known that lateral force responds to topography. 765 x 765 nm image of debris, possibly part of a supertip, on polished quartz. From AFM-3; narrow 200  $\mu\text{m}$  cantilever; X hysteresis corrected.

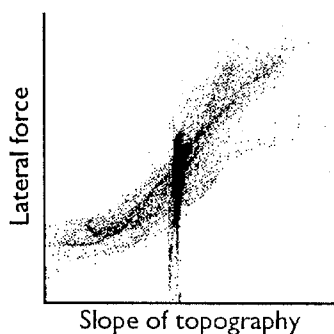
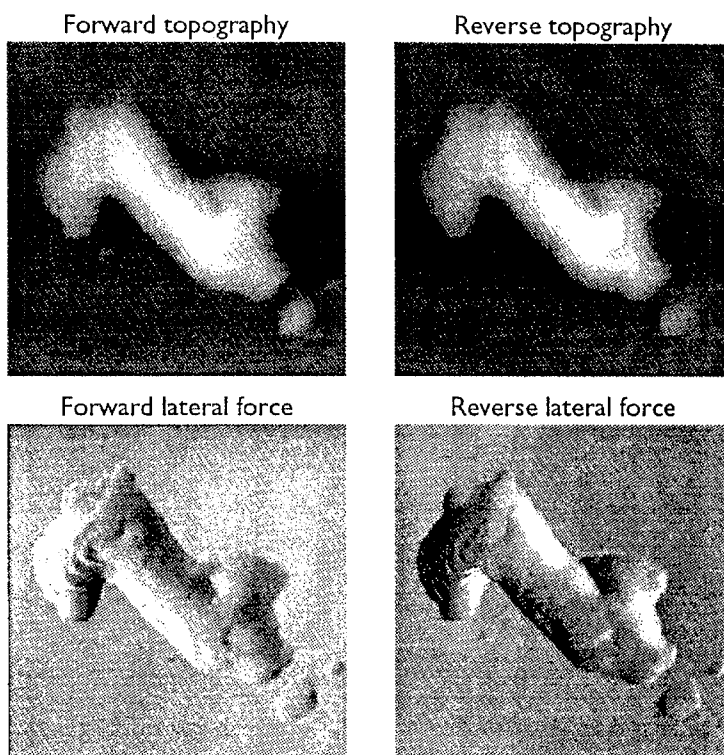


Figure 11. Correlation diagram derived from the images in figure 9 showing rough first-order dependence of lateral force on slope. The vertical smear results from image curvature.

cause the slight J-shape. Lateral force approximately doubles at a slope of one, while at a slope of  $-1$  it measures about zero.

Stick-slip and tip shape artifacts are less common manifestations of slope effects

Stick-slip is a special case of slope effect generally seen only at the atomic scale. The AFM tip “sticks” behind an atom and, as the scan progresses, lateral force builds up until the tip “slips” by the atom, at which point it “sticks” at the next atom (chapter 1).

Irregular tip shape might give rise to subtle slope-dependent variations in tip-sample interaction and thus in lateral force.

However, I have not observed this contrast mechanism.

### The differential lateral deflection image

Differential lateral deflection images show tip-sample adhesion

Adding the forward and reverse lateral force images to create a differential lateral deflection image eliminates two of the four LFM contrast mechanisms: XYZ coupling and slope effects. Since both of the remaining contrast mechanisms — meniscus effects and contact adhesion — result from attractive forces, the differential lateral deflection image essentially shows tip-sample adhesion. In fact, it bears a strong resemblance to adhesion images derived from force curve imaging spectroscopy (figure 12; see chapter 3).

The adhesion image in figure 12 shows on average 55 nN of adhesive force; the bright areas at the edges of the collagen fibril have about 120 nN of adhesive force. The differential lateral deflection image shows on average 13 nN of frictional force, and 37 nN at the edges of the fibril.

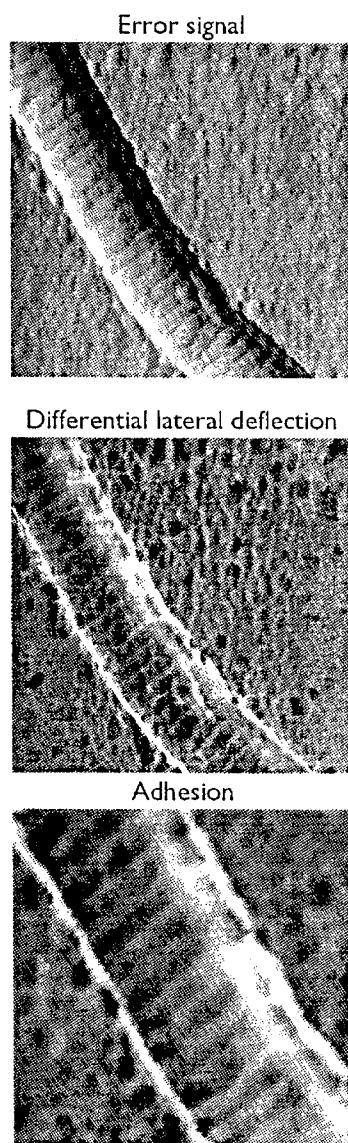


Figure 12. Error signal and simultaneous differential lateral deflection ( $1.3 \times 1.3 \mu\text{m}$ ) images of a collagen fibril compared with an adhesion ( $700 \times 700 \text{ nm}$ ) image. The latter two show the increased meniscus force at concave regions of the sample (bright areas; see chapter 6).

## References

---

- den Boef, A.J. (1991) The influence of lateral forces in scanning force microscopy. *Rev. Sci. Instrum.* 62(1), 88–92
- Meyer, G. and Amer, N.M. (1990) Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. *Appl. Phys. Lett.* 57(20), 2089–2091
- Overney, R.M., Meyer, E., Frommer, J., Brodbeck, D., Lüthi, R., Howald, L., Güntherodt, H.-J., Fujihira, M., Takano, H., and Gotoh, Y. (1992) Friction measurements on phase-separated thin films with a modified atomic force microscope. *Nature* 359, 133–135
- Thundat, T., Warmack, R.J., Allison, D.P., Bottomley, L.A., Lourenco, A.J., and Ferrell, T.L. (1992) Atomic force microscopy of deoxyribonucleic acid strands adsorbed on mica: the effect of humidity on apparent width and image contrast. *J. Vac. Sci. Technol. A* 10(4), 630–635
- Zenhausen, F., Adrian, M., ten Heggeler-Bordier, B., Eng, L.M., and Descouts, P. (1992) DNA and RNA polymerase/DNA complex imaged by scanning force microscopy: influence of molecular-scale friction. *Scanning* 14, 212–217

# Chapter 7. Hardness imaging

## *Introduction*

---

I have developed a method for imaging sample hardness with AFM

In the study of tip-sample interaction, hardness imaging complements lateral deflection imaging: while lateral deflection indicates how much force the tip exerts on the sample, hardness indicates how the sample responds to that force.

Simultaneously with other groups (Radmacher et al., 1992; Tao et al., 1992), I have developed methods for imaging sample hardness (see chapter 3). Inspired by analog techniques for barrier height imaging in scanning tunneling microscopy, the lock-in technique implemented in AFM-3 obtains high signal-to-noise ratio and minimal interference with other imaging modes.

Chapter 7 discusses the interpretation of hardness images. As with lateral force measurements, I have found that a number of artifacts appear in hardness measurements. These artifacts do not arise as a result of the lock-in technique but occur in all forms of hardness imaging.

### Measurement of sample compression

---

Hardness images show how much the sample has compressed under the normal pressure exerted by the tip

Referring back to chapter 3, hardness images present measurements of cantilever deflection  $A$  in response to a  $Z$  modulation of amplitude  $H$ . Force curves provide the tip-sample contact force  $F$ . This information suffices to determine the distortion  $D$  of the sample given the difference between the current cantilever deflection  $C$  and the cantilever deflection  $B$  at the tip-sample breakaway point (see chapters 1 and 3):

$$D = (C - B) \frac{H - A}{A}$$

Unfortunately, piezoceramic hysteresis and rolloff of scanner response with increasing frequency make measurements of  $H$  inaccurate. Since  $H \propto A$ , calculation of  $H - A$  yields a very high error. To alleviate this problem I generally replace  $H$  with  $A_{\max}$ , the value of  $A$  measured over the hardest surface in the image.

A more complex analysis requiring knowledge of tip shape, sample height, and Poisson's ratio of the sample could in principle yield Young's modulus (Burnham, 1989).

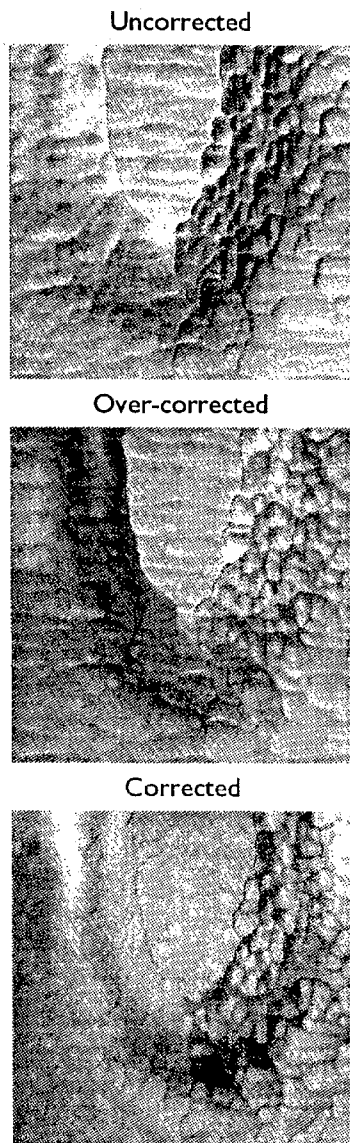


Figure 1. The slope artifact in a  $630 \times 630$  nm hardness image of a compact disk pit. The over-corrected image shows how orthogonality correction can change the direction of slope shading. In the properly corrected image, slope artifacts still appear, but with greatly reduced magnitude. The corrected image resembles a solid model with a light source directly above it.  $H = 10$  nm.

### *Slope coupling*

---

Unless modulation occurs normal to the sample surface, the tip slides over the sample during hardness measurements

Hardness images often resemble slope-shaded topography images, especially at larger modulation amplitudes (figure 1). The dependence of the hardness measurement on sample slope occurs because the tip slides over the sample during Z modulation.

Tip sliding occurs wherever the sample surface does not lie perpendicular to the direction of modulation. This effect reduces the cantilever deflection  $A$ , making the sample appear softer than it actually is. The magnitude of the effect depends both on the slope of the sample and on tip-sample friction; thus, features with distinctive frictional characteristics may exhibit spurious hardness characteristics.

Z nonorthogonality compounds slope coupling. For each 1 nm of motion in Z, the scanner in AFM-3 moves 0.2 nm parallel to the sample surface. When the instrument electronics modulate Z, the tip also moves in the XY plane, sliding over the sample surface. SCAN has a Z orthogonality correction that alleviates this problem (figure 1).

Since tip sliding always makes the sample appear softer, users should take the greatest hardness reading (the brightest area) of any given feature to determine sample compression.

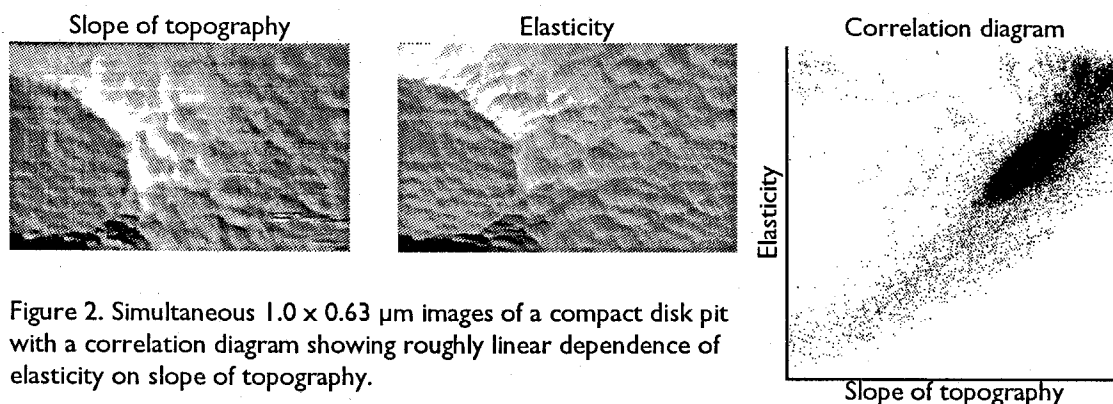
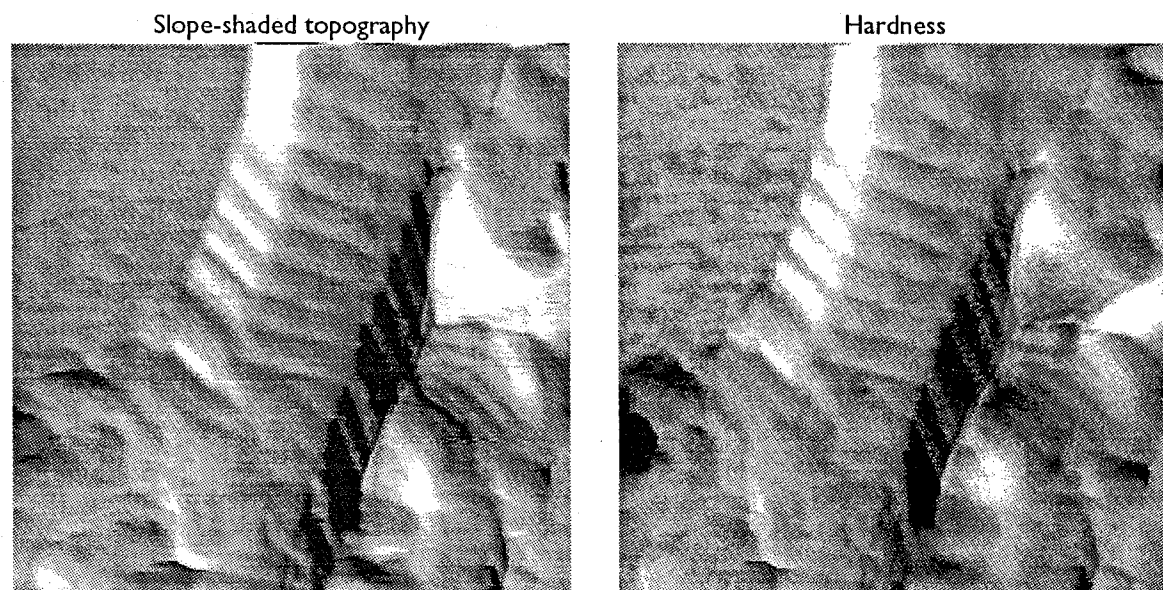


Figure 2. Simultaneous  $1.0 \times 0.63 \mu\text{m}$  images of a compact disk pit with a correlation diagram showing roughly linear dependence of elasticity on slope of topography.

Intentional use of the slope artifact produces slope-shaded images with low noise

Figure 3. Intentional use of the slope-shading artifact. Low-frequency noise can make it difficult to obtain good images of features that run parallel to the scan direction. In this example, the hardness image reveals the banding pattern of a collagen fibril without the low-frequency noise inherent in the simultaneous slope-shaded topography image. Light source at bottom left for both images.  $1.0 \times 1.0 \mu\text{m}$ .

On occasion, the slope artifact becomes a useful contrast mechanism (figure 3). Intentionally introducing nonorthogonality can produce slope images with any desired lighting direction. Unlike slope-shaded topography images, these images do not suffer from low-frequency noise.





## *Z coupling*

---

Optical lever artifacts sometimes create a Z dependence in hardness images

Hardness images sometimes show a dependence on sample height (figure 4). This artifact has the same cause as Z coupling in lateral force images: stray laser light reflected from the sample interferes with light reflected from the cantilever, causing optical lever sensitivity to vary with Z.

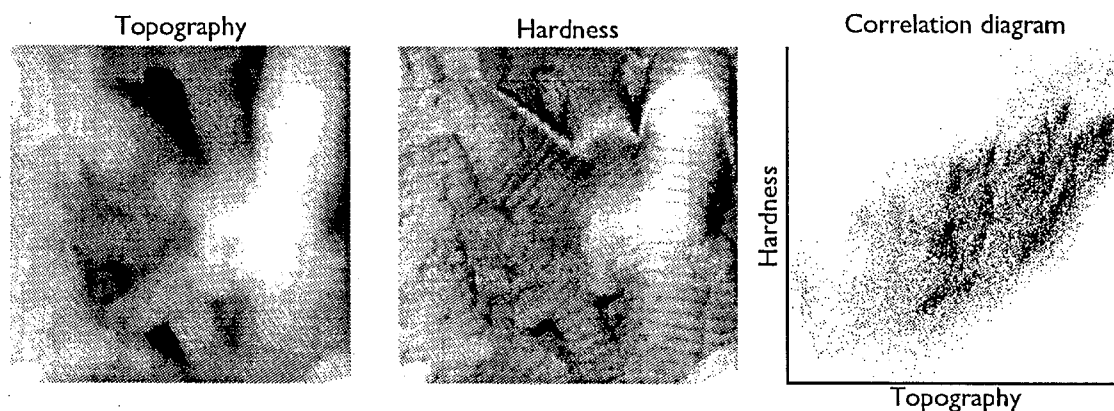
## *Nonlinear elastic behavior*

---

DNA can become incompressible under the pressure of the tip

Figure 4.  $1.7 \times 1.7 \mu\text{m}$  image of collagen fibrils showing Z coupling in the hardness measurement. The bright horizontal band near the top of the hardness image may result from a DAC zero-crossing error.  $H = 10 \text{ nm}$ .

Images of DNA deposited on mica often yield a counterintuitive result, showing that the hardness of the DNA exceeds that of the mica (figure 5). This effect cannot result from



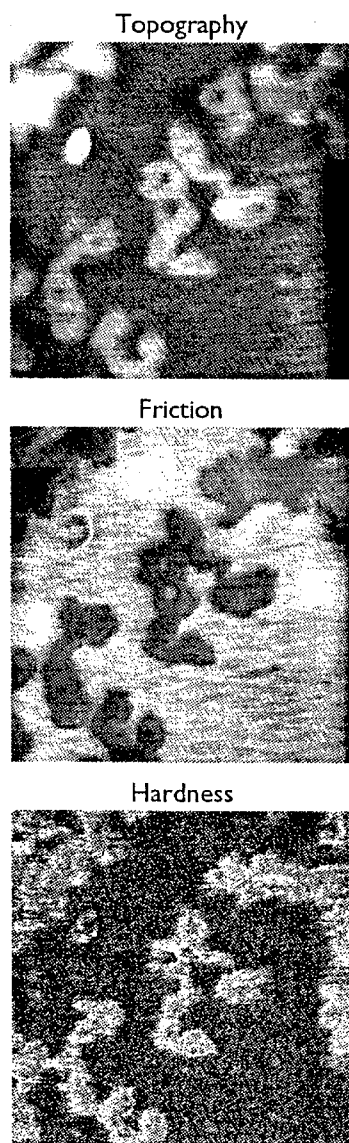


Figure 5. 830 x 830 nm images of pUC18 DNA showing 2 nm tall plasmids with low friction and high hardness.  $H = 10$  nm. Sample prepared by Solomon Ting.

tip sliding: since tip-sample friction is lower over the DNA, the tip should slide more and make the DNA appear softer than it actually is. The minuscule height of the DNA rules out Z coupling artifacts (in some preparations the DNA only measures 0.2 nm tall; see figure 6).

Preliminary results indicate that DNA becomes softer as contact force decreases. Thus, the tip probably compresses DNA to the point where hard-sphere repulsion effectively prevents further compression. The DNA has undergone a phase change, becoming harder under the pressure of the tip. The large and relatively stiff mica surface, on the other hand, exhibits relatively linear elastic behavior.

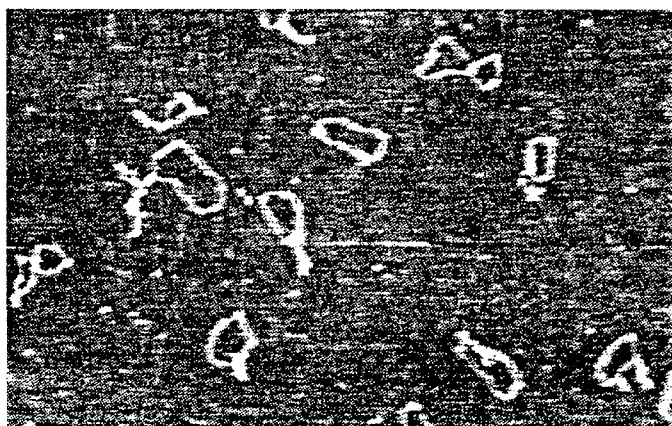


Figure 6. 2.1 x 1.3  $\mu\text{m}$  topography image of pUC18 plasmids on mica. The plasmids measure only 0.2 nm tall, probably an indication that they are single-stranded (since the height is the same in the forward and reverse scans, this is not a friction coupling artifact as described in chapter 5). The small height of these plasmids also makes them appear narrower in comparison with the 2 nm tall plasmids in figure 4. Morphologically decurved. Sample prepared by Solomon Ting.

## References

---

- Burnham, N.A. and Colton, R.J. (1989). Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *J. Vac. Sci. Technol. A* 7(4), 2906–2913
- Radmacher, M., Tillmann, R.W., Fritz, M., and Gaub, H.E. (1992) From molecules to cells: imaging soft samples with the atomic force microscope. *Science* 257, 1900–1905
- Tao, N.J., Lindsay, S.M., and Lees, S. (1992) Measuring the microelastic properties of biological material. *Biophys. J.* 63, 1165–1169

## Chapter 8. Float-polished quartz

### *Introduction*

---

Images of float-polished quartz have prompted the development of decurving algorithms and provided insight into the effects of tip shape (Baselt et al., submitted)

The finish quality of float-polished quartz has no immediate bearing on AFM tip–sample interactions, but the study described in this chapter does demonstrate how dull tips can create artifacts even on a very flat surface. It also prompted me to develop all of the curvature-removal processes described in chapter 4. The work described here thus has bearing on the interpretation of AFM images.

### *Float polishing*

---

Float polishing produces highly smooth surfaces in a damage-free process based on the controlled flow of a polishing slurry over the substrate

In float polishing (Namba and Tsuwa, 1978), a diamond-turned lap  $\sim 500$  mm in diameter rotates in a polishing slurry at about 55 rpm. The substrate to be polished, centered between the periphery and the axis of the lap, rotates at the same rate and in the same direction. In equilibrium the substrate floats over the lap, supported by pressure gradients at the lap ridges. As the slurry flows across the substrate, regions of high pressure form over and preferentially remove protrusions. On a microscopic level the slurry, consisting of a suspension of polishing powder in water, undergoes Couette flow. Its Reynolds number ideally approaches but does not exceed the critical value  $R_C$ , leading to rapid polishing and laminar rather than turbulent flow (Soares et al., 1993).

Evaluation by Nomarski microscopy, optical scatterometry, and photo-acoustic spectroscopy reveals that float polishing consistently yields samples with excellent surface and sub-surface characteristics, well suited for use in electronic, photonic, and acoustic devices (Soares et al., 1993).

I have imaged surfaces from which float-polishing has removed either 10 or 100  $\mu\text{m}$  of material

Schubert Soares prepared the float polished quartz. Natural quartz disks 25.4 mm in diameter and 0.9 mm thick were first lapped on a copper lap with a slurry of 2% by weight  $\text{TiO}_2$  in deionized water, removing either 10  $\mu\text{m}$  or 100  $\mu\text{m}$  of material at a rate of 1  $\mu\text{m}/\text{hr}$  (yielding either a “10  $\mu\text{m}$  face” or a “100  $\mu\text{m}$  face”). Two more float-polishing steps using tin laps followed, the first for four hours with 2% by weight  $\text{TiO}_2$ , the second for eight hours with 2% by weight  $\text{SiO}_2$ .

### Curvature removal

---

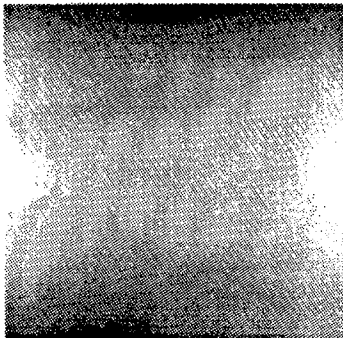


Figure 1. 1.8  $\mu\text{m}$  scan of polished quartz, 100  $\mu\text{m}$  face, raw data. This is the same image as figure 4b. Spurious image curvature dominates real features, complicating image analysis. The tonal range of the image (height difference between the brightest and darkest areas) is 14 nm.

Line leveling effectively removes spurious curvature from images of float-polished quartz

A typical 1.8  $\mu\text{m}$  image has 14 nm of curvature (figure 1). I have evaluated a number of algorithms (see chapter 4) to remove this curvature, among which the horizontal and vertical line leveling algorithms work best (as previously described, these subtract a best-fit line from each row and each column of data in the image). The procedures have the drawback of removing real features larger than about half the size of the image, as well as long features (such as scratches) aligned perfectly parallel or perpendicular to the scan direction. Nonetheless, I have found

line leveling a reliable method. Other curvature-removal algorithms such as best-fit paraboloid subtraction, line-by-line average subtraction, and Fourier filtering often fail to remove all the curvature or, worse, introduce spurious curvature of their own. Morphological decurving tends to remove larger features that, in the case of float-polished quartz, are often real.

### *Tip shape artifacts*

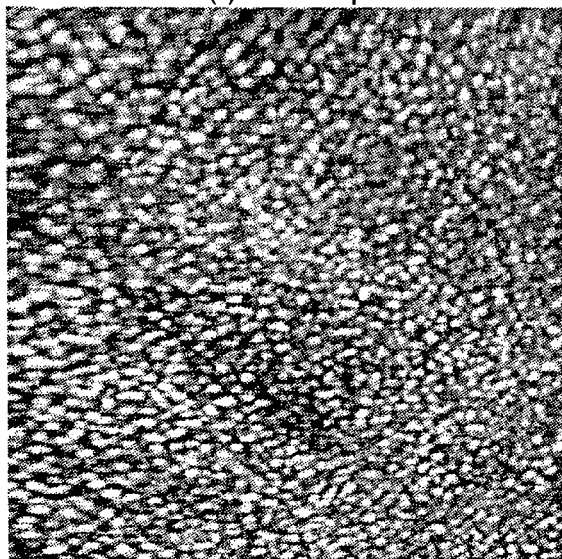
---

Dull tips create a “pebbled” appearance

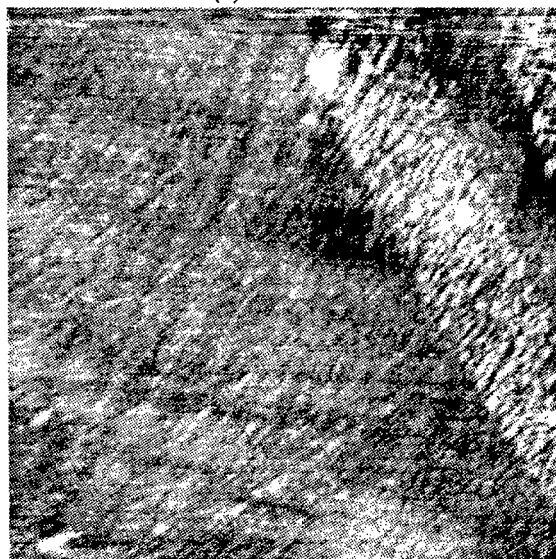
Figure 2. 1.8  $\mu\text{m}$  scans of a 10  $\mu\text{m}$  face taken with a normal tip and an Ultralever. Corrected for X hysteresis; horizontal and vertical line leveled.

Images taken with normal tips (figure 2) have a granular appearance, apparently revealing many similar particles  $\sim 50$  nm in diameter. Images taken with Ultralevers often have the same appearance, but sometimes show finer structure such as in figure 2. This particular image also shows a fairly common

(a) Normal tip



(b) Ultralever



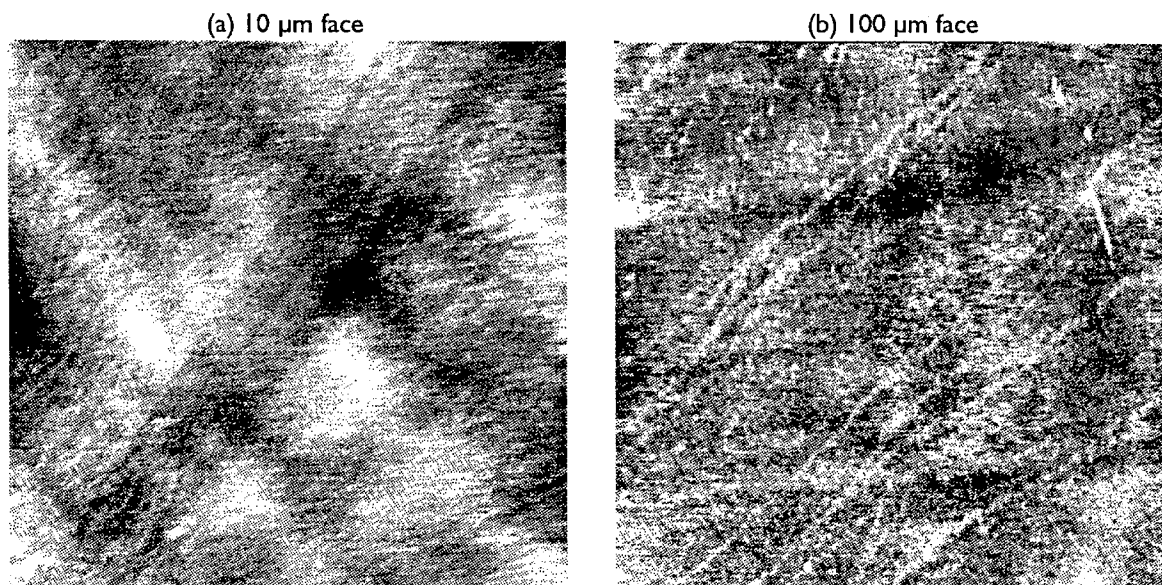


Figure 3. 1.8  $\mu\text{m}$  scans taken with electron-beam-deposited supertips. Corrected for X hysteresis; horizontal and vertical line leveled. 10  $\mu\text{m}$  face tonal range 2.0 nm; 100  $\mu\text{m}$  face tonal range 1.4 nm.

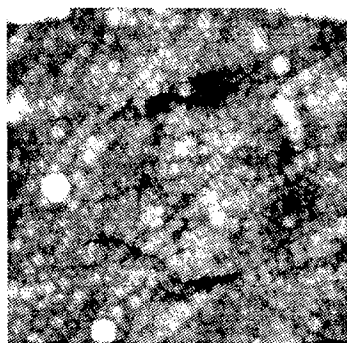


Figure 4. Same image as in figure 3b but processed to simulate a 750 nm diameter tip. The numerous round features thereby created bear a resemblance to those in figure 2a, suggesting that such features result from dull tips.

inhomogeneity: one region of the sample is significantly rougher than the rest. Finally, supertip images (figure 3) sometimes reveal still finer structure, including broad, shallow trenches that may be polishing marks, and narrow ridges that could be grain boundaries. Like Ultralever images, though, images from supertips often resemble those from normal tips.

Grütter et al. (1992) have reported extreme cases in which tip artifacts dominate AFM images. Dull tips can also cause the far more subtle granular appearance present in figure 2. I have demonstrated this by convolving figure 3b with an imaginary 750 nm radius tip, using the morphological image processing algorithm of dilation (figure 4). Assuming that the sharp features in figure 3b indicate a sharp tip, this process simulates how the image would have appeared had I acquired it with a 750 nm radius tip. Like figure 2a, the dilated image consists solely of round particles (their size reflecting the radius of the imaginary



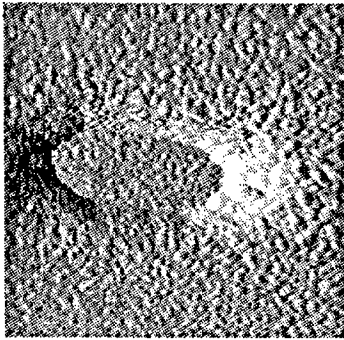


Figure 5.  $1.0 \times 1.0 \mu\text{m}$  error signal image of a compact disk showing a granular appearance similar to that in figure 2a.

tip), indicating that the different appearances of figures 2a, 2b, and 3 may result from different tip radii.

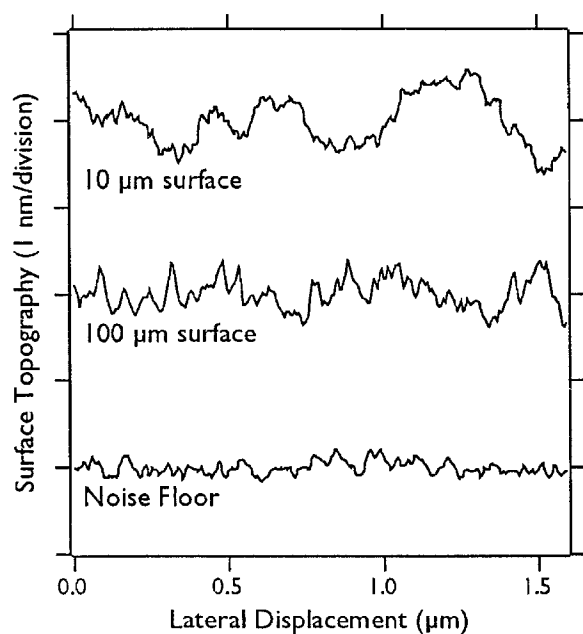
Images of flat surfaces such as compact disks (figure 5) often show this type of tip-shape artifact.

### Evaluation of surface finish

The  $100 \mu\text{m}$  face has a better finish than the  $10 \mu\text{m}$  face

Figures 3a and 3b compare surfaces that have undergone  $10 \mu\text{m}$  and  $100 \mu\text{m}$  removal. Figure 6 compares line profiles from these images. The surface becomes somewhat smoother after  $100 \mu\text{m}$  of removal, having about  $0.2 \text{ nm}$  roughness as opposed to about  $0.5 \text{ nm}$  after  $10 \mu\text{m}$  of removal. Thus, a longer polishing period

Figure 6. Line profiles of the images in figure 3.



results in a smoother finish, probably because it removes sub-surface damage that would remain after a short polishing period.

I noted no distinctive lateral force characteristics in any image of float-polished quartz.

## References

---

- Baselt, D.R., Soares, S.F., and Baldeschwieler, J.D. (submitted)  
Atomic force microscopy of float-polished quartz
- Grütter, P., Zimmermann-Edling, W., and Brodbeck, D. (1992)  
Tip artifacts of microfabricated force sensors for atomic force  
microscopy. *Appl. Phys. Lett.* 60(22), 2741–2743
- Namba, Y. and Tsuwa, H (1978) *Ann. CIRP* 27(1), 511
- Soares, S.F., Baselt, D.R., Black, J.P., Jungling, K.C., and Stowell,  
W.K. (1993) The float-polishing process and analysis of float-  
polished quartz. To appear in *Appl. Optics*

# Chapter 9. Type I collagen

## *Introduction*

---

I have observed the subfibrillar structure and mechanical properties of type I collagen (Baselt, Revel, and Baldeschwieler, in preparation)

Chapter 9 describes the application of the above-described tools to the study of a biological sample, type I collagen. Topography images show subfibrillar structure not previously seen by AFM, while elasticity images demonstrate that biological materials can become softer and thus more likely to undergo distortion when imaged underwater. Friction images reveal bands of high adhesion that may represent patterns of proteoglycan distribution. Finally, imaging this “dirty” sample has provided insight into the process of tip contamination.

## *Type I collagen*

---

### Collagen is a fibrous structural protein

Collagen is an extracellular protein responsible for most of the mass and strength of structural tissues such as bone, tendon, skin, and cartilage. Tropocollagen molecules form collagen fibrils by spontaneously aggregating into ordered arrays (for a review see Veis and Payne, 1988). In type I collagen fibrils, this order results in a prominent 60–70 nm transverse “D” periodicity. High-resolution studies reveal that these fibrils consist of numerous parallel microfibrils 4–5 nm in diameter, sometimes arranged helically (Ruggeri, 1979). X-ray diffraction suggests that each microfibril contains 5 tropocollagen molecules (Piez and Trus, 1981).

Dry rat tail tendon contains 90% by weight type I collagen (Brodsky and Eikenberry, 1982). The tendon itself consists mainly of large collagen fibrils with a mean diameter of 280 nm, oriented parallel to the axis of the tendon and having a straight or slightly helical microfibrillar structure. A loose, jellylike outer sheath (peritendineum and endotendineum; Strocchi et al., 1985) contains 40–50 nm mean diameter highly helical collagen fibrils wrapped around the tendon (Raspanti et al., 1990; Parry and Craig, 1977).

AFM has previously revealed the D periodicity of collagen fibrils (Chernoff and Chernoff, 1992).

## Sample preparation

---

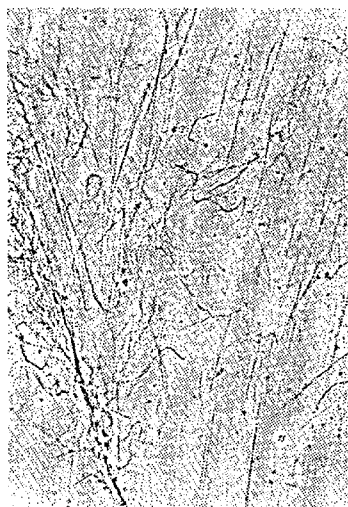


Figure 1. 110 x 160  $\mu\text{m}$  optical micrograph of a collagen spread. The spread contains both isolated, individual fibrils and bundles of fibrils.

I prepared native collagen by spreading rat tail tendon on glass

I prepared native rat tail collagen by peeling the skin from frozen rat tails and cutting a shaving off the underlying tendon. I then placed the shaving on a glass microscope slide and, using two razor blades, cut it several times while gently pulling apart (figure 1). By placing the outer surface of the shaving in contact with the slide, this technique yields mostly small outer sheath fibrils.

Conversely, placing the inner surface in contact with the slide yields mostly large collagen fibrils. I soaked the samples in 1:100 Triton X-100 for 10 minutes, then rinsed them with  $\sim 5$  mL 1:100 Triton X-100 followed by running distilled water. After soaking the samples in distilled water for 10 minutes, I allowed them to dry in air. I have imaged 34 such samples for this study.

I also prepared reconstituted collagen by heat gelation

I prepared reconstituted bovine dermal collagen by heat gelation (Brodsky and Eikenberry, 1982). I reconstituted 30 mL of a pH 7.4 100  $\mu\text{g}/\text{mL}$  solution of Vitrogen 100 (Collagen Corp., Palo Alto, CA) in Hank's Balanced Salts (Sigma Chemical Corp., St. Louis, MO) at  $37^{\circ}\text{C}$  for 12 hours. I filtered the solution through 0.4  $\mu\text{m}$  filter paper and washed with 1:100 Triton X-100 followed by distilled water. After recovering the mass of collagen I dissected it onto microscope slides as for rat tail collagen, but

with no further washing. I have imaged 31 such samples for this study.

## *Instrumentation*

---

I acquired most images with normal tips in air

As with all images in this thesis, unless otherwise noted I acquired the images in this chapter using normal pyramidal tips. Although I have tried ultralevs and electron-beam-deposited tips, these have proven less successful.

Unless otherwise specified, I acquired the images in air. Also note that the zero points of the height and lateral position scales on the cross sections are arbitrary.

## *Results*

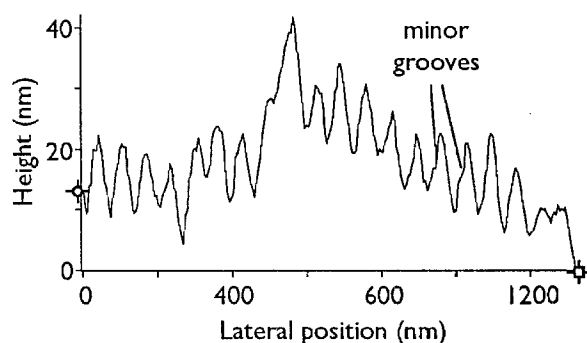
---

AFM reveals three types of subfibrillar structure: *D* bands, minor bands, and microfibrils

Reconstituted bovine dermal and native rat tail collagen yield similar results. The distinctive 60–70 nm *D* periodicity of type I collagen, consisting of alternating grooves and ridges, appears in almost all images (figure 2). The difference in height between the grooves and the ridges varies between ~5 nm on small-diameter fibrils and ~15 nm on large-diameter fibrils.



Figure 2.  $2.1 \times 1.3 \mu\text{m}$  image of native rat tail collagen showing D periodicity and evidence of microfibrils (ovals). The cross section, taken between the marked points, shows hints of intraperiod (minor) bands; the small fibril at top exhibits minor bands more clearly. Fibril height  $\sim 300 \text{ nm}$ . Imaging force  $63 \text{ nN}$ . Image processing: Phong shading, light source at left. Composited with unshaded plane subtracted image.  $443 \times 277$  points.



AFM occasionally reveals microfibrils on large-diameter fibrils. They are in evidence in figure 2, and more so in figure 3. They often have a slight right-handed helicity, with the microfibrils oriented at a  $\sim 5^\circ$  angle to the axis of the fibril. In addition, large fibrils tend to wind around each other in a right-handed helix.

The smaller-diameter sheath fibrils often exhibit two grooves per  $D$  period as in figure 4. The lower fibril in this image has a 22



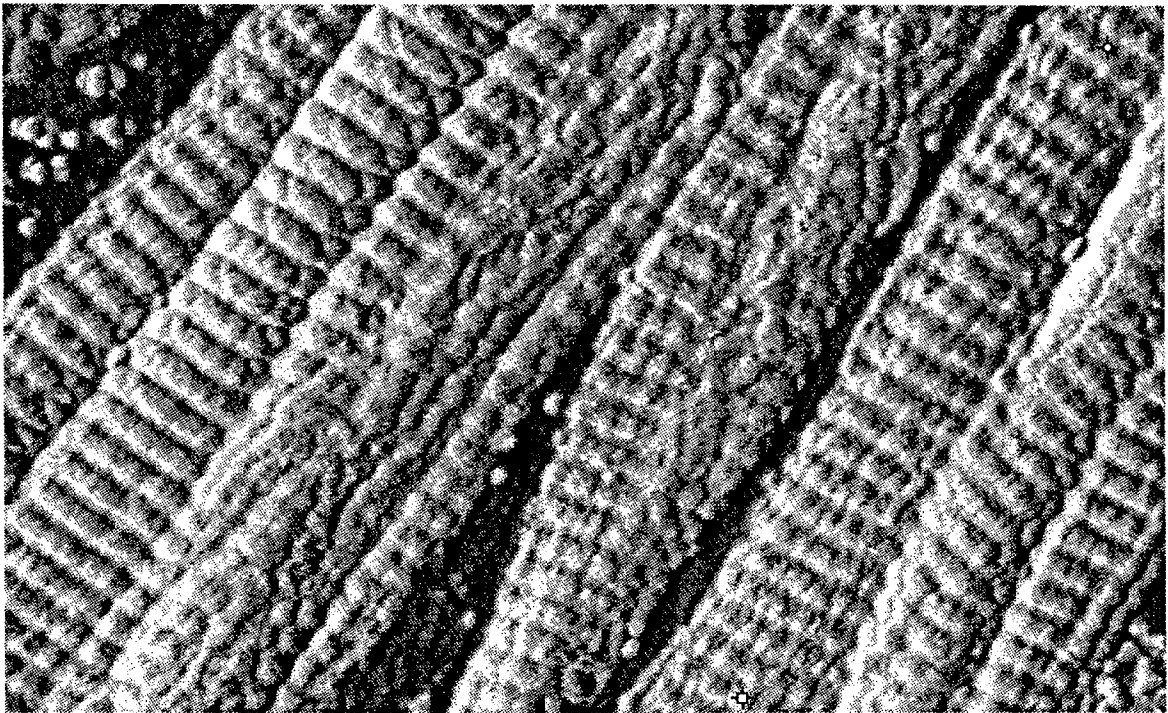
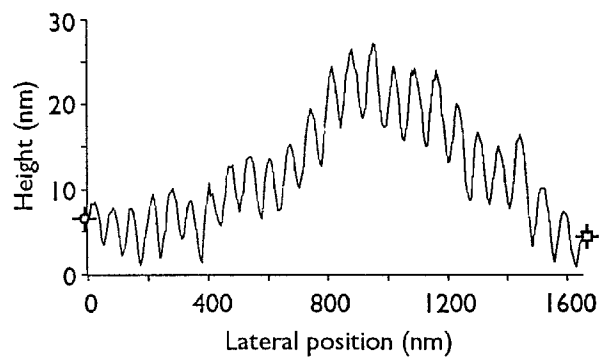


Figure 3.  $2.3 \times 1.4 \mu\text{m}$  image of reconstituted bovine dermal collagen showing microfibrils. This sample was fixed in 2.7% glutaraldehyde for 10 minutes. Fibril height  $\sim 50 \text{ nm}$ . Imaging force **35 nN**. Image processing: slope shading (light source at left) followed by statistical differencing. Composited with unshaded plane subtracted image.  $443 \times 277$  points.



nm height with large grooves measuring 3 nm deep and small grooves (which I will refer to as “minor bands”) 0.5 nm deep. Minor bands appear often enough that they probably do not arise from a double tip artifact. Although they appear most clearly on small fibrils, large fibrils can also show hints of minor bands’ (figure 2 graph).

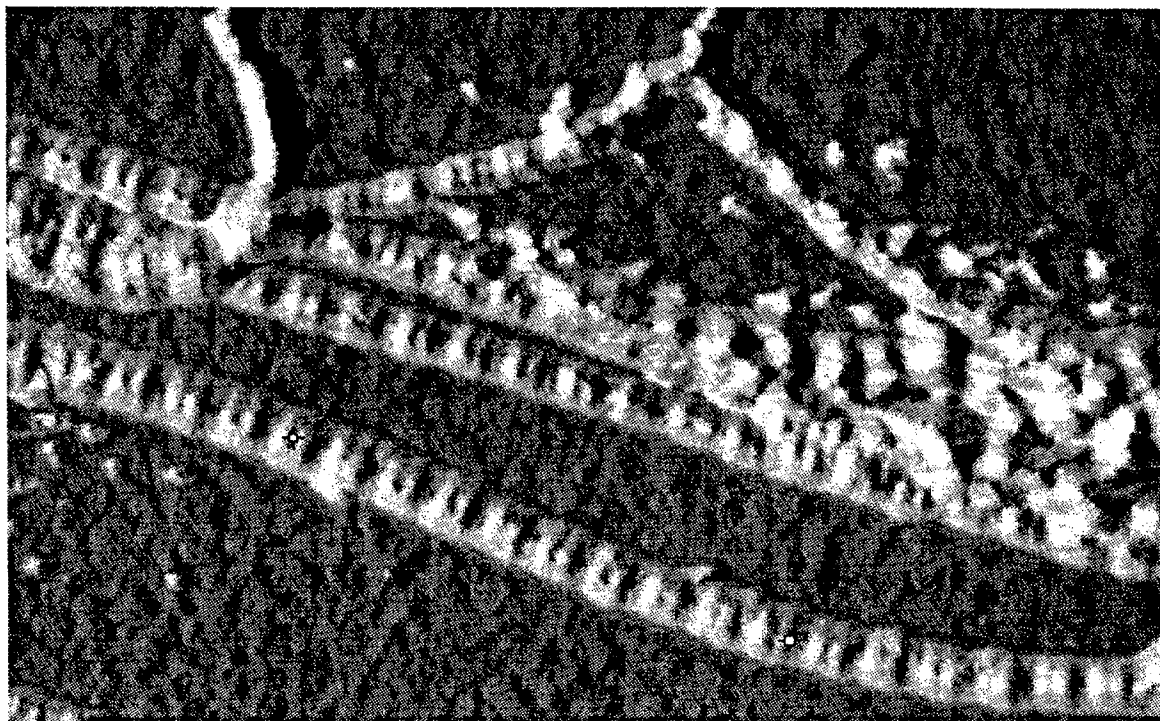
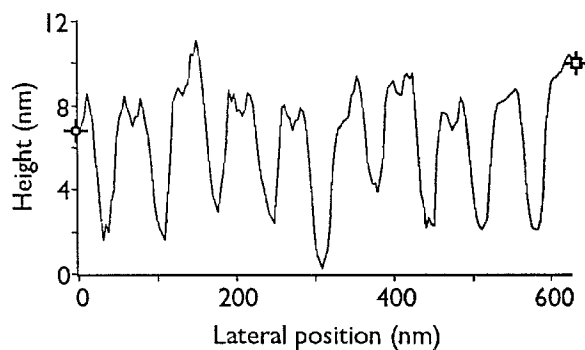


Figure 4.  $1.4 \times 0.89 \mu\text{m}$  image of native rat tail collagen showing banding substructure. Fibril height 22 nm. Imaging force 35 nN. Image processing: morphological opening followed by  $6 \times 6$  sliding window binomial averaging and Phong shading with light source at left. Composited with unshaded plane subtracted image.  $443 \times 277$  points.



We have fixed some samples with glutaraldehyde (figure 3), but this does not consistently improve the level of detail.

The pyramidal AFM tips that we have used produced images equal to or better than those acquired with electron-beam-deposited tips or Ultralevers. Ultralevers tend to rapidly become dull from contamination (see below).

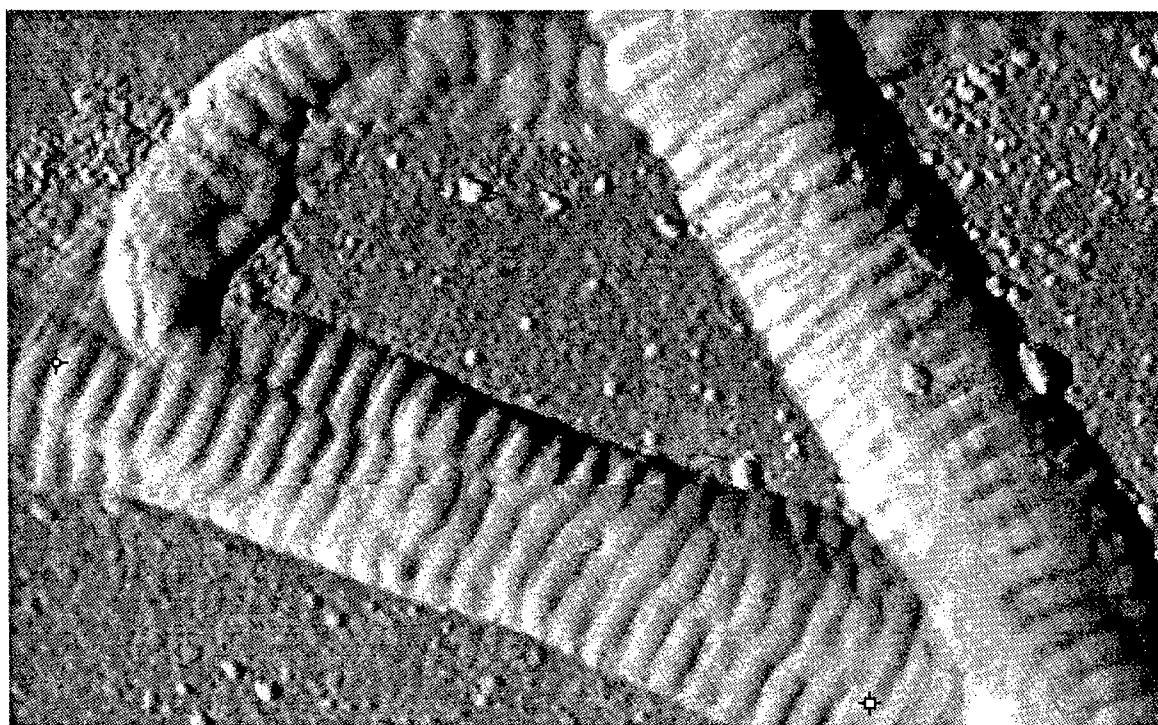
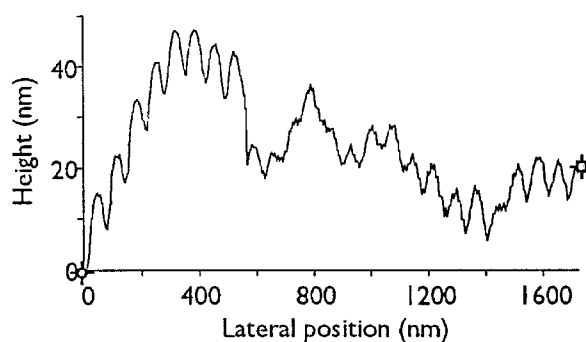


Figure 5.  $2.3 \times 1.4 \mu\text{m}$  image of reconstituted bovine dermal collagen showing tip changes. The scan progresses from top to bottom. Near the top of the image the definition of the banding pattern improves. At lateral position = 550 nm on the line profile, the definition abruptly becomes worse: the *D* periodicity grooves become shallower, and the noise level increases. Fibril height  $\sim 80$  nm. Imaging force 55 nN. Image processing: slope shading (light source at left). Composited with unshaded plane subtracted image.  $443 \times 277$  points.



### An unknown material sometimes obscures the *D* period

Frequently, a covering of some material partially or completely obscures the *D* period. Washing with Triton X-100 and water alleviates but does not completely eliminate this problem. The tip often disturbs material lying on or next to the fibrils while

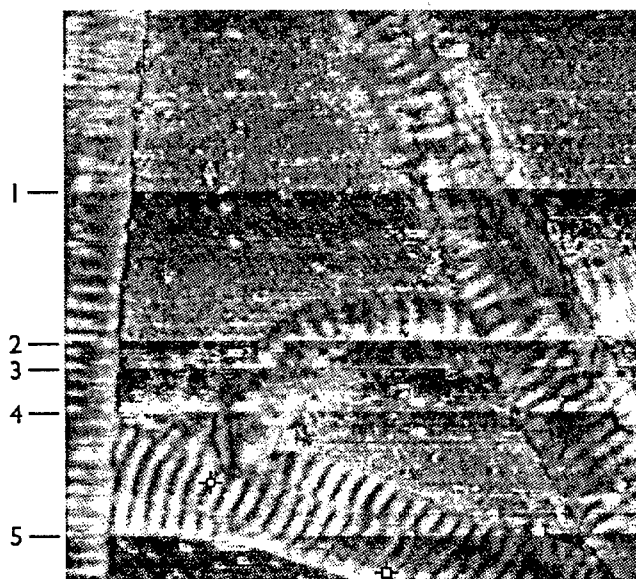
scanning the sample, and in many cases this material accumulates on the tip. Figure 5 shows an image of washed collagen that demonstrates this phenomenon. Within a sharply delineated horizontal band at its center, the image has good definition; at the top and bottom resolution appears poorer. The same pattern appeared in four different scans of this area.

Figure 6 shows a differential lateral deflection image acquired during an earlier, lower magnification scan of the same region. I acquired a height image (not shown) simultaneously with this image. I have calculated the average height (after plane subtraction) and tip-sample friction for each scan line and displayed the results as a graph (figure 6, graph B). This graph reveals that friction gradually increases, then suddenly decreases, in a repeating pattern. Ignoring the presence of the collagen fibrils and the plane-subtraction, the height follows the same pattern. Note that the discontinuities labeled 4 and 5 in figure 6 appear at the same positions as the two discontinuities visible in figure 5.

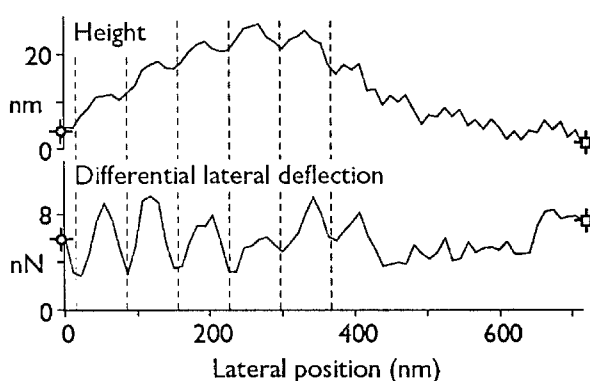
### Ridges exhibit more tip-sample adhesion than grooves

When topographic images exhibit good definition of the  $D$  periodicity, corresponding differential lateral deflection images show that the ridges of the periodicity have a stickier surface than the grooves (figure 6, graph A). Coupling of lateral deflection into error signal (see pages 127–129) could potentially cause features with high or low friction to appear more pronounced

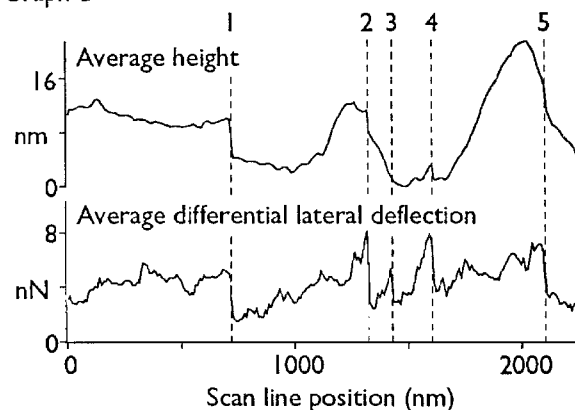
Figure 6.  $2.3 \times 2.3 \mu\text{m}$  differential lateral deflection image of the same region as shown in figure 4 (which corresponds to the lower half of this image). Graph A shows cross sections of the differential lateral deflection (adhesion) image and the average of the forward and reverse topography images (averaging the forward and reverse images together removes any lateral deflection crosstalk). Tip-sample adhesion is high at the ridges and low at the grooves of the collagen *D* periodicity. The banding pattern in the height cross section is not as well-defined as in the figure 4 graph, which has approximately twice the number of data points per nanometer. Graph B shows the average apparent height and tip-sample adhesion for each scan line. Friction and height gradually increase as material accumulates on the tip, then suddenly decrease when the material drops off the tip. I have labeled the five most prominent such “tip changes”; numbers 4 and 5 are also visible in figure 4. Image processing: plane subtraction of height image; subtraction of forward from reverse lateral deflection data to produce differential lateral deflection image.  $250 \times 250$  points.



Graph A



Graph B



than they should in topographic images. However, we can demonstrate that this artifact does not appear to a significant degree in figures 5–6. Crosstalk of lateral deflection into error signal is direction-dependent, but the topography contrast in figure 6 is direction-independent: a topography image taken with the tip scanning right-to-left (data not shown) appears identical to the presented image, taken with the tip scanning left-to-right.

Another possible source of the frictional contrast between ridges and grooves is the meniscus force, which tends to raise frictional forces in the vicinity of concave features and lower them near convex features (pages 113, 130, and 133). However, figure 6 reveals the opposite pattern, with the ridges exhibiting more friction than the grooves. We thus conclude that the ridges are actually stickier than the grooves.

Note that the image of collagen on page 133, taken with a relatively dull tip, does not show this banding pattern. Meniscus forces have the most prominent effect in this image.

### Fibrils become soft when immersed in water

Subfibrillar structure disappears when imaging under water (figure 7), although vague hints of the  $D$  periodicity sometimes appear. Hardness images reveal that the collagen becomes soft under these conditions. The fibril imaged under water in figure 7 measures 51 nm in height and has compressed 2.3 nm (5%) under the 11 nN force exerted by the tip. By comparison, the

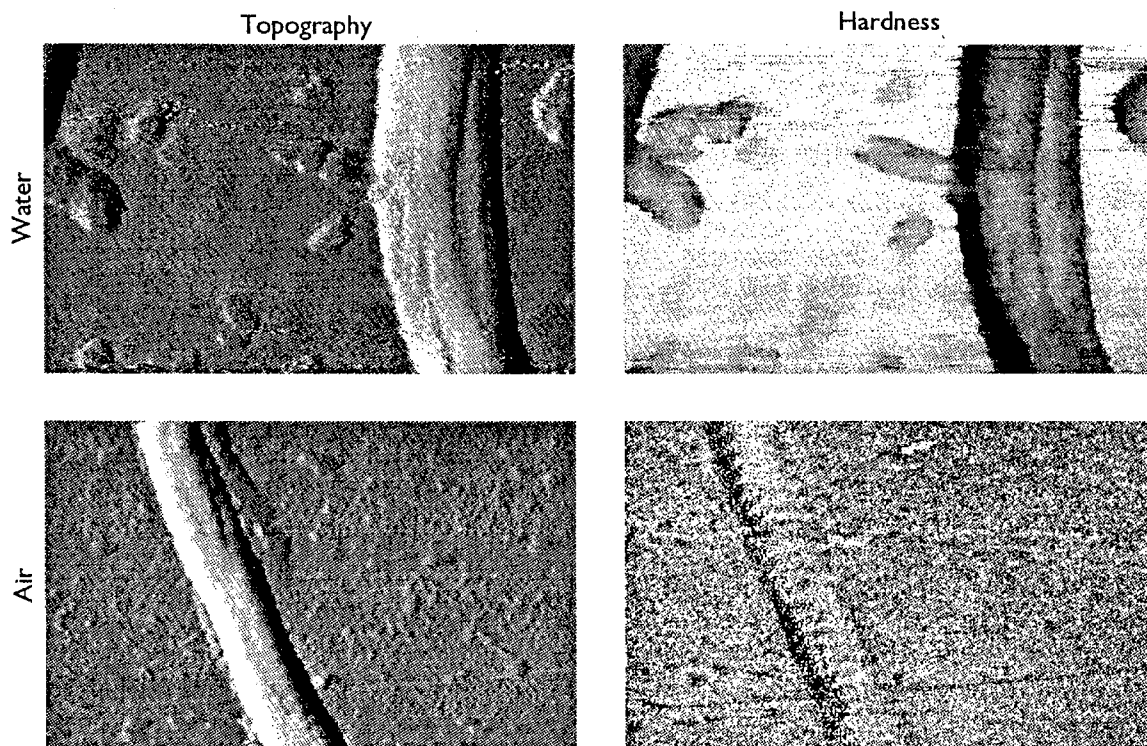
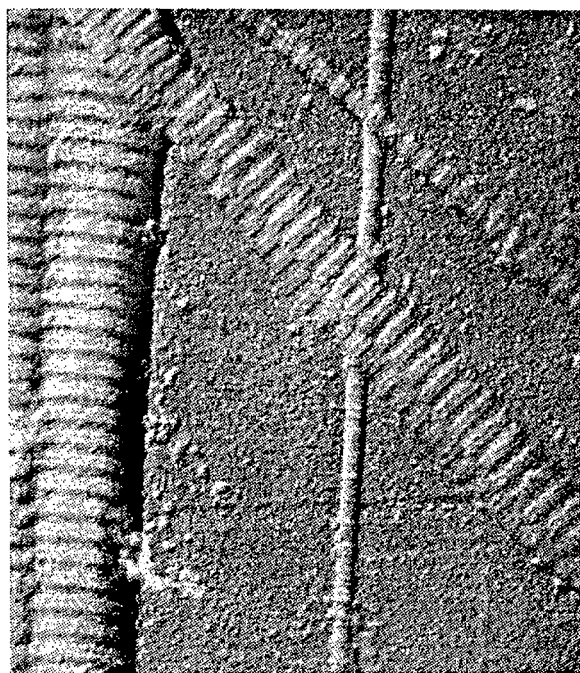


Figure 7.  $2.3 \times 1.4 \mu\text{m}$  images of reconstituted bovine dermal collagen showing effect of imaging in water. Hardness images taken simultaneously with topography images; brighter areas are harder. The dark bands to the left of the fibrils in the hardness images are an artifact caused by the tip slipping during the modulation process; the hardness measurement only yields a correct result where the surface is perpendicular to the direction of modulation. The hardness image taken in air shows negligible difference between the collagen fibril and the glass substrate, while the image of a different fibril taken in water shows that it has become soft. The absolute noise level is the same in both hardness images, but the image taken in air appears noisier because we have increased its contrast to reveal what features it does exhibit. Both fibrils measure 50 nm in height. Imaging force 11 nN under water, 44 nN in air. Modulation amplitude 10 nm. Image processing: topography images slope shaded with light source at left, composited with unshaded plane subtracted images. Hardness images unprocessed, except for contrast and brightness adjustment. 316 x 198 points.

fibril imaged in air (a different fibril in the same preparation) has not undergone measurable compression ( $<0.2 \text{ nm}$ ). Likewise, fibrils imaged in ethanol have negligible elasticity and show *D* banding, just as they do when imaged in air (data not shown).

Figure 8.  $1.6 \times 1.8 \mu\text{m}$  electron microscope image of platinum-shadowed collagen fibrils. Shadowed from top-left. Image processing: multiplied by  $-1$ . Image by Jean-Paul Revel.



## Discussion

---

Electron microscopy reveals structure similar to that seen by AFM

Electron microscopy has formerly revealed all of the subfibrillar features shown here, but the AFM images differ in that I have not coated, stained, freeze-dried, or (in most cases) fixed the samples. AFM images have a close resemblance to electron microscope images of shadowed collagen (figure 8). Electron microscopy of negatively-stained fibrils reveals a more extensive banding pattern (Olsen, 1963). Although it appears that regions of intense negative staining correspond to low regions in our images, the



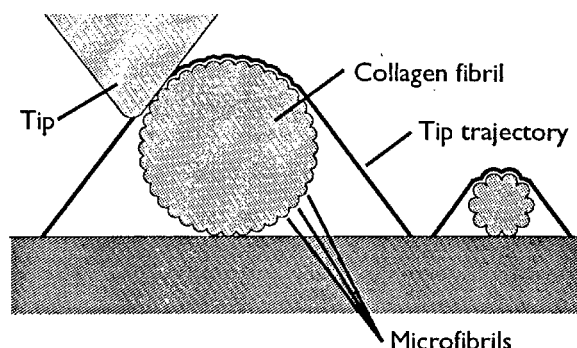
images do not rule out the opposite interpretation, that regions of intense negative staining correspond to high regions.

Freeze-fracture studies of rat tail tendon have revealed slightly helical microfibrils and the  $D$  periodicity (Raspanti et al., 1990; Ruggeri et al., 1990), while shadowing studies of skin tendon have revealed microfibrils and up to six intraperiod bands (Gross and Schmitt, 1948) (Note that these studies refer to the slightly helical appearance of the microfibrils as “nonhelical,” since other fibril types exhibit much more pronounced helicity).

#### Tip artifacts may allow AFM to image different subsets of features in small and large fibrils

AFM images only reveal microfibrils on large-diameter fibrils, while minor bands appear most prominently on small-diameter fibrils. Since electron microscopy has revealed both structures on fibrils of all diameters, this observation may represent an imaging limitation of the AFM. Specifically, the pyramidal shape of the tip limits resolution on highly-sloped regions of the sample (Grütter et al., 1992). Large-diameter fibrils have a wide unsloped region over which the tip can image microfibrils, while small-diameter fibrils have a much narrower topside, possibly too narrow to observe microfibrils (figure 9). Similarly, the greater corrugation of major bands in large-diameter fibrils may obscure minor bands.

Figure 9. Schematic cross-section of two collagen fibrils and the path the tip takes when scanning them. AFM cannot resolve highly sloped structures because the side rather than the end of the tip contacts them. Thus, AFM may reveal few if any microfibrils when imaging small-diameter fibrils.



Underwater images may lack detail because of tip-sample interactions or fibril swelling

Images of collagen under water may lack evidence of subfibrillar structure for one of three reasons. 1) Proteins become about eight times softer when hydrated from a dry state (Morozov et al., 1988). Therefore, forces exerted by the tip may distort fibrils immersed in water enough to obliterate the *D* periodicity and other subfibrillar structure. However, this seems unlikely since we estimate that the tip only distorts the fibril by 2.3 nm, while the grooves and ridges of the *D* periodicity differ in height by 10 nm. 2) Under water, the tip seems to pick up foreign material and permanently become dull within the first few scans. This may decrease the resolution enough to prevent observation of banding. 3) The banding may actually disappear under water because of fibril swelling. Electron micrographs of replicas of wet fibrils have previously suggested that type I collagen may not exhibit banding when fully hydrated (Gross and Schmitt, 1948).

Significantly, the elasticity measurements reveal that when imaging in air, distortion of the sample from tip-sample interaction forces does not limit resolution.

### Tip contamination can improve resolution

Figures 5 and 6 suggest that the tip gradually accumulates a sticky material from the sample, and the material suddenly drops off the tip at locations that can be consistent from scan to scan. In the figure 5 graph, two distinct changes occur when material drops off the tip. First, the depth of the D periodicity grooves decreases. The tip has apparently become duller and no longer fits into the grooves. Second, the noise level increases. Variations in noise level often accompany variations in the character of the tip. Changes in the amount or character of material adhering to the tip can change the speed with which the cantilever responds to variations in the feedback output  $Z$ , causing the feedback to overshoot or oscillate and resulting in increased noise.

Contrary to expectation, the accumulated material does not necessarily dull the tip; in some cases it actually seems to improve resolution. In the figure 5 graph resolution decreases when material drops off the tip; therefore, the contaminating material improved resolution. It is unlikely that part of the tip itself broke off, since the event occurred reproducibly from scan to scan. Perhaps small particles attached to the tip sometimes have a sharper profile than the tip itself.

In general, though, cleaner collagen samples yield higher-resolution images. Macroscopic pieces of both native and reconstituted collagen are very sticky, a property probably related to the high concentration of tip-contaminating material in our preparations. The stickiness presumably arises from proteoglycans or glycosaminoglycans associated with the collagen, removal of which may result in higher resolution images. However, treatment with glycosidases in a preliminary experiment only had a negative effect.

The frictional characteristics of ridges and grooves may arise from the distribution of proteoglycans

Since proteoglycans associate with the gap region of rat tail type I collagen (Scott and Orford, 1981), the frictional contrast that we observe between the ridges and grooves of the  $D$  periodicity may reflect the differing composition of these areas. For unknown reasons, proteoglycans apparently decrease tip-sample adhesion.

## Conclusion

---

AFM can bring useful new capabilities to the study of biology

This chapter demonstrates that on collagen, atomic force microscopy can achieve resolution approaching or equaling that

of electron microscopy, but with minimal sample preparation. Thus AFM can image fibril structure free from the effects of elements such as cryoprotectants, stains, embedments, or coatings. In addition, the ability of AFM to image under water and to measure elasticity and friction provides access to a number of new experimental possibilities.

## References

---

- Baselt, D.R., Revel, J.-P., and Baldeschwieler, J.D. (in preparation)  
Subfibrillar structure of type I collagen observed by atomic  
force microscopy
- Brodsky, B. and Eikenberry, E.F. (1982) Characterization of  
fibrous forms of collagen. *Meth. Enzymol.* 82, 127–174
- Chernoff, E.A.G. and Chernoff, D.A. (1992) Atomic force  
microscope images of collagen fibers. *J. Vac. Sci. Technol. A* 10(4),  
596–599
- Gross, J. and Schmitt, F.O. (1948) The structure of human skin  
collagen as studied with the electron microscope. *J. Exp. Medicine*  
88(5), 555–568
- Grütter, P., Zimmermann-Edling, W., and Brodbeck, D. (1992)  
Tip artifacts of microfabricated force sensors for atomic force  
microscopy. *Appl. Phys. Lett.* 60(22), 2741–2743
- Morozov, V.N., Morozova, T.Ya., Kachalova, G.S., and  
Myachin, E.T. (1988) Interpretation of water desorption  
isotherms of lysozyme. *Int J. Biol. Macromol.* 10, 329–336
- Olsen, B.R. (1963) Electron microscope studies on collagen: I.  
Native collagen fibrils. *Zeitschrift für Zellforschung* 59, 184–198
- Parry, D.A.D. and Craig, A.S. (1977) Quantitative electron  
microscope observations of the collagen fibrils in rat-tail  
tendon, *Biopolymers* 16, 1015–1031
- Piez, K.A. and Trus, B.L. (1981) A new model for the packing of  
type-I collagen molecules in the native fibril. *Bioscience Reports* 1,  
801–810

- Raspanti, M., Ottani, V., and Ruggeri, A. (1990) Subfibrillar architecture and functional properties of collagen: a comparative study in rat tendons. *J. Anat.* 172, 157–164
- Ruggeri, A., Benazzo, F., and Reale, E. (1990) Collagen fibrils with straight and helicoidal microfibrils: a freeze-fracture and thin-section study. *J. Ultrastructure Res.* 68, 101–108
- Scott, J.E. and Orford, C.R. (1981) Dermatan sulphate-rich proteoglycan associates with rat tail-tendon collagen at the d band in the gap region. *Biochem. J.* 197, 213–216
- Strocchi, R., Leonardi, L., Guizzardi, S., Marchini, M., and Ruggeri, A. (1985) Ultrastructural aspects of rat tail tendon sheaths. *J. Anat.* 140(1), 57–67
- Veis, A. and Payne, K. (1988) Collagen fibrillogenesis. In *Collagen* 1, 113–137 ed. Nimni, M.E. CRC, Boca Raton, FL

*Part 3*

# Nearfield Scanning Optical Microscopy

## *Introduction*

---

Part three discusses the design of a nearfield scanning optical microscope

Unlike atomic force microscopy, nearfield scanning optical microscopy (NSOM) can image samples under water without tip-sample contact. In addition, it permits access to the stains and labeling techniques developed for optical microscopy.

The similarity of all scanned-probe microscopes from the standpoint of control makes it simple to add new types of SPM to an existing system. Thus, a few minor modifications to the control system described in part one, plus the construction of the stage described here, permitted experimentation with a relatively new type of microscopy with potentially great relevance to biology.



# Chapter 10. Nearfield scanning optical microscopy

## *Introduction*

---

I have designed a nearfield scanning optical microscope

In 1992 Steve O'Connor of the Baldeschwieler group began a nearfield scanning optical microscope (NSOM) project. For this project I designed the NSOM stage described below.

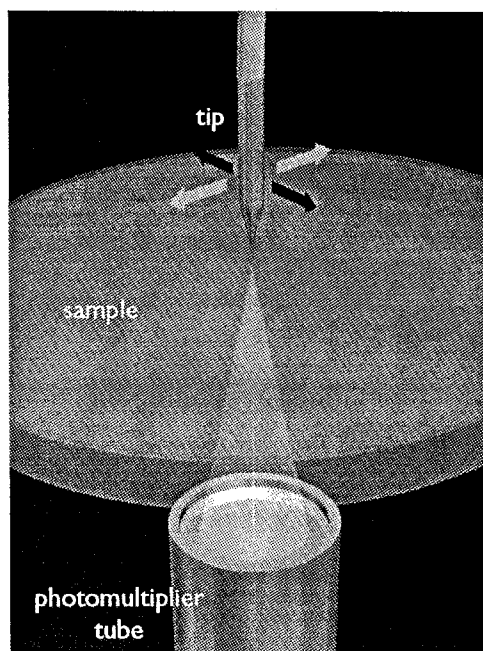
## *General principles*

---

NSOM uses the same contrast mechanisms as lens-based optical microscopy but attains 12 nm resolution

The nearfield scanning optical microscope (NSOM), one of the more recently developed scanned proximity probe microscopes, operates by scanning a point source of light over the surface of interest (figure 1; for a review see Betzig and Trautman, 1992). A photodetector measures the transmitted or reflected light, which

Figure 1. Concept of NSOM. A fiber optic pulled to a tip with a quartz micropipette puller serves as a point source of light. A photomultiplier tube collects the light transmitted through or reflected from the sample as the control system scans the tip.

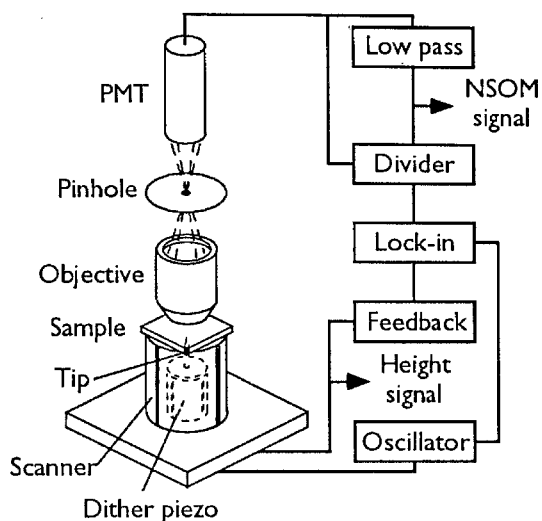


the system plots as a function of probe position to yield a high-resolution optical micrograph. Because NSOM does not require a focusing mechanism, diffraction does not limit its resolution.

Although several implementations of NSOM have been demonstrated, the most practical uses a fiber optic coupled to a laser at one end and having an aluminum-coated point at the other. In both theory and practice, this implementation can attain a resolution of 12 nm (Betzig et al., 1991).

NSOM can use any of the contrast mechanisms associated with lens-based optical microscopy — fluorescence, optical density, phase contrast, and so on (Betzig and Trautman, 1992); however, NSOM cannot image subsurface features.

Figure 2. Shear force feedback. A dither piezo oscillates the tip, and an objective images the light from the tip onto a pinhole. As the tip oscillates, the position of the light on the pinhole moves, resulting in a high-frequency modulation of light intensity at the detector. A low-pass filter distinguishes this modulation from the low-frequency signals that result from variations in sample transmissibility. After Betzig et al., 1992.



NSOM can operate without feedback or with shear force feedback

NSOM generally operates without feedback of any kind — the tip scans over the sample at a fixed height that the user determines on the basis of image sharpness. Recently, though, Betzig et al. (1991) have demonstrated feedback based on meniscus-induced shear forces between tip and sample (figure 2). In this scheme, a piezoceramic element dithers the tip near its resonant frequency, and a lock-in detector determines the amplitude and phase of tip resonance. Upon approaching within  $\sim 200$  nm of the sample, the phase begins to shift and the amplitude decreases. A feedback loop can thus keep the separation between tip and sample constant by monitoring the phase or amplitude of tip vibration. Shear force detection achieves sufficient resolution to warrant its use for

topographic imaging in atomic force microscopes (Yang et al., 1992).

Betzig et al. (1992) have also implemented shear-force feedback using separate feedback and imaging photodetectors, a more versatile arrangement than the one shown in figure 2.

### *The stage*

---

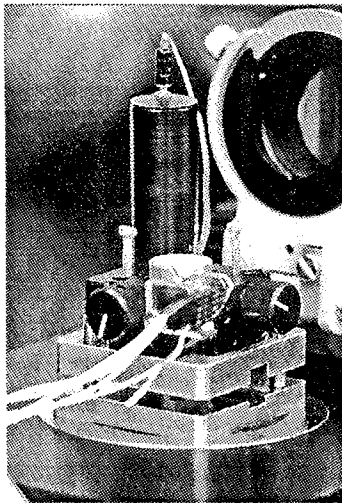


Figure 3. Photograph of the NSOM stage on an inverted microscope.

I have designed a general-purpose NSOM stage

I have designed an NSOM stage capable of imaging in air or under liquids and in transmission or reflection (figure 3, figure 4). Steve O'Connor constructed the stage.

The stage uses a 1.5" long, 0.5" diameter, and 0.022" thick EBL #1 piezoceramic tube (Staveley Sensors, EBL division, East Hartford, CT) having a 20  $\mu\text{m}$  XY range when driven with  $\pm 220\text{V}$  electronics. To stabilize this vibration-prone structure and prevent mechanical damage, I have placed the tube in an aluminum sheath 0.75" in diameter and filled the space between the tube and sheath with silicone rubber. When imaging under water, which could short out the scanner electrodes, the sheath also keeps the scanner dry.

For coarse XY positioning, a Spindler and Hoyer (Milford, MA) 06 5070 ball bearing-free X-Y displacement unit holds the scanner. For coarse Z positioning, a Newport (Fountain Valley, CA) 860A-05 motorized drive and two Newport AJS-02H

micrometer screws form a lever that rests in a hole-groove-flat support in the base (see chapter 3). The scanner, translator, and coarse Z mechanism is the "head."

Two spring-loaded latches (not shown in figure 4; springs from Century Spring Co., Inc., Los Angeles, CA) can, if desired, secure the head to the base with about 5 pounds of force.

For transmission imaging, a magnet (custom made by Magnet Sales and Manufacturing Co., Culver City, CA) at the end of the

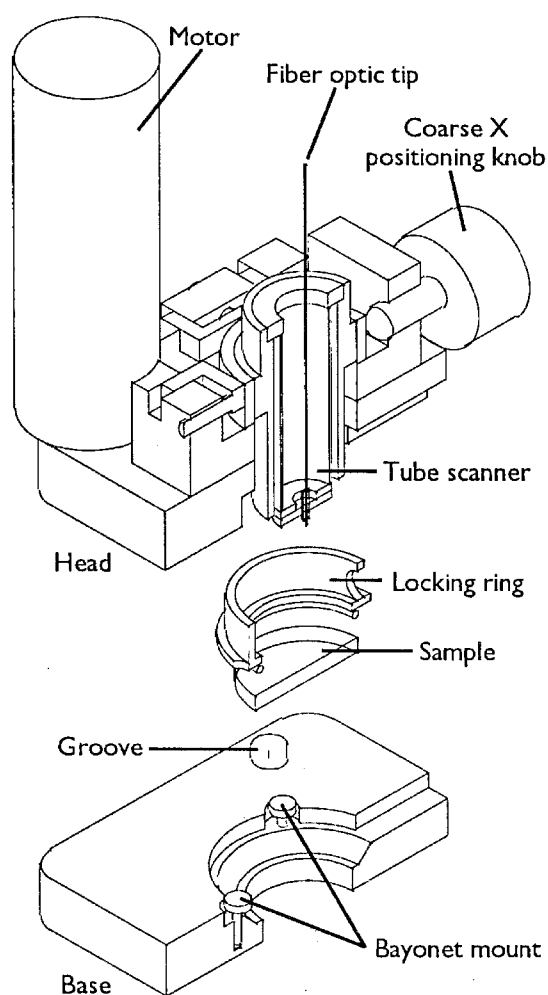
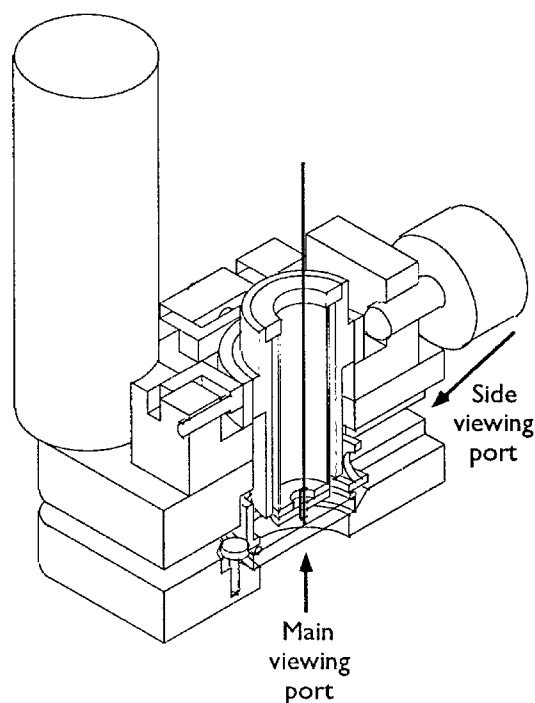


Figure 4. Cross-sections of the NSOM stage. The head consists of a custom-made steel plate holding a commercially-available translator, motor, and micrometer screws (not visible). The base consists of a bayonet mounting system to lock the sample in place and a hole-groove-flat kinematic rest for the head. When assembled, two viewing ports allow the user to inspect the sample and the tip-sample separation. Each side of the stage is 2.75 inches long.



scanner holds the tip, and a bayonet-mounted ring clamps the sample over a one-inch diameter hole in the "base." The ring also retains water for underwater imaging. For reflection imaging, the magnet holds the sample and the ring holds a one-inch diameter disk of polished quartz with a small hole through its center which holds the tip.

Although we have not attempted shear force feedback, the quartz tip holder can accommodate a tip dithering piezoceramic for feedback during reflection imaging. For transmission imaging it may be possible to use the scanner to modulate the tip; the scanner might, however, absorb most of the high-frequency vibrations.

The stage rests on an inverted microscope (Nikon Optiphot). A Hamamatsu (Bridgewater, NJ) HC124-01 current-mode photomultiplier tube placed in the camera mount detects the light emitted from the tip. A dissection microscope (Meiji EMZ-5 with U stand, DU adapter, and FS holder, Meiji Techno America, San Jose, CA) provides a side view to observe the tip-sample separation.

Since we do not have an optical table, I have constructed a laser coupling apparatus. A 30" long aluminum tube with 0.5" thick walls and a 1 kHz resonant frequency holds a Uniphase U-1320 1 mW He-Ne laser (Newport) and a Newport F-1015 high precision single-mode fiber coupler. We have used Newport F-SV single-mode 633 nm fiber optic.

## *Preliminary results*

---



Figure 5. 20 x 20  $\mu\text{m}$  NSOM image of red blood cells. Image taken by Steve O'Connor and David Baselt.

We have imaged red blood cells

We have used the NSOM without feedback to image red blood cells (figure 5). The contrast in this image results not from pigmentation in the cells but from refraction. The PMT signal averages 2 V, while the features measure 2 V p-p over a 20 mV noise level.

Larger scan range or greater scanner stability could be beneficial

Given the true 16-bit performance of our electronics (Clark et al., 1992), we could have a 200  $\mu\text{m}$  range scanner and still obtain 3 nm resolution, 4 $\times$  oversampling of the maximum 12 nm resolution. The larger scanner range would simplify imaging of cells and other large biological samples.

On the other hand, when used without feedback the NSOM can scan at high rates, limited only by signal-to-noise considerations and scanner stability. Unfortunately, stable scanners tend to have smaller range. The current 20  $\mu\text{m}$  range scanner is a compromise between the ideals of large range and high stability.

## References

---

- Betzig, E., Trautman, J.K., Harris, T.D., Weiner, J.S., and Kostelak, R.L. (1991) Breaking the diffraction barrier: optical microscopy on a nanometric scale. *Science* **251**, 1468–1470
- Betzig, E., Finn, P.L., and Weiner, J.S. (1992) Combined shear force and near-field scanning optical microscopy. *Appl. Phys. Lett.* **60** (20), 2484–2486
- Betzig, E. and Trautman, J.K. (1992) Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit. *Science* **257**, 189–195
- Clark, S.M., Baselt, D.R., Spence, C.F., Youngquist, M.G., and Baldeschwieler, J.D. (1992) Hardware for digitally controlled scanned probe microscopes. *Rev. Sci. Instrum.* **63**(10), 4296–4307
- Yang, P.C., Chen, Y., and Vaez-Iravini, M. (1992) Attractive-mode atomic force microscopy with optical detection in an orthogonal cantilever/sample configuration. *J. Appl. Phys.* **71**(6), 2499–2502