A Tutorial on the Mechanisms, Dynamics, and Control of Atomic Force Microscopes

Daniel Y. Abramovitch, Sean B. Andersson, Lucy Y. Pao, and Georg Schitter

Abstract—The Atomic Force Microscope (AFM) is one of the most versatile tools in nanotechnology. For control engineers this instrument is particularly interesting, since its ability to image the surface of a sample is entirely dependent upon the use of a feedback loop. This paper will present a tutorial on the control of AFMs. We take the reader on a walk around the control loop and discuss each of the individual technology components. The major imaging modes are described from a controls perspective and recent advances geared at increasing the performance of these microscopes are highlighted.

I. INTRODUCTION

The Atomic Force Microscope (AFM), invented by Binnig, Quate, and Gerber [1], is one of the most versatile methods of imaging nanoscale structures (see Table I). An AFM is not restricted to imaging in a vacuum environment – as are the Transmission Electron Microscope (TEM) and the Scanning Electron Microscope (SEM). The sample preparation is far easier with an AFM than with a TEM. Furthermore the AFM is becoming a driving technology in nanomanipulation and nanoassembly [2] and is playing a burgeoning role in the field of molecular biology [3] - [9]. One of the interesting features of this tool is that imaging depends entirely on a feedback control loop. By and large, most AFMs use piezoelectric actuators, optical detection of cantilever deflection, and PI or PID control. AFMs can operate in a variety of modes, including contact mode (where the control loop tries to maintain constant contact force with the sample surface) and AC or intermittent contact mode (where the control loop tries to maintain a constant oscillation amplitude as the tip is bounced off of the sample surface).

Moreover, scientists and engineers interested in phenomena with nanometer-scale features are increasingly demanding better tools. Unfortunately for users, the joke about AFMs is that companies need to ship a Ph.D. with each system to keep the machines operating properly. The desire for faster and more repeatable results has led to considerable interest in advanced mechanics and controls for this problem.

A. AFM Basics

The standard layout of an Atomic Force Microscope is shown in Figure 1. The purpose of an AFM is to characterize a sample by bringing a sharp probe in close proximity to the sample surface. The probe tip is affected by the forces on the surface, some attractive and some repulsive [10], [11]. These forces cause a deflection of the cantilever on which the tip resides and this deflection is detected. While the original method of detection was via tunneling detection [1], the most common mode by far is the so called “optical lever,” in which a laser beam is bounced off of the cantilever and back onto a split photo detector [12]. Typical photo detectors for common AFMs now have four quadrants, allowing both the longitudinal bending modes and lateral torsional modes of the cantilever to be detected. The deflection in the cantilever results in a push pull signal on the detector which can be used to control the tip-sample interaction force. This is discussed in Section III-B.

The rest of this paper is organized as follows. Section II overviews a variety of the problems that are addressed by AFMs. Section III walks readers around the AFM control loop so that they understand the basic layout of the problem, and also delves into the individual technology components that comprise the loop. Section IV overviews the basics of AFM control, while Section V describes the operational modes of AFMs. Sections VI and VII discuss issues affecting AFM control and advanced AFM control topics, and concluding remarks are given in Section VIII.

II. A BRIEF SAMPLING OF AFM APPLICATIONS

The capability of AFM to image in vacuum, air, or in liquids with sub-nanometer resolution (see Table I), to manipulate objects with nanometer-scale features, and to measure forces with better than pico-Newton resolution makes it an extremely useful tool in a wide variety of disciplines. We highlight here a few applications. This is in no way a survey of the literature; such an undertaking is beyond the scope of this tutorial. Recent survey papers include [5], [9], [13]-[18].

A contrast between AFM methods and optical methods can be seen in Figure 2. While optical microscopy is a parallel and therefore faster measurement, the raster imaging of an AFM produces higher resolution. In addition, the optical image is a 2D image with the AFM image is a 3D surface map. The colors in an AFM image are computer
Fig. 1. An AFM Control Block Diagram. The diagram shows a scanned sample design, where the tip and cantilever are fixed and the sample is moved under the tip by the piezo actuator. In this mode, the controller attempts to maintain a constant level of deflection which corresponds to a constant level of contact force. The quantity to be measured, the surface profile, comes in as an unknown disturbance to the control loop. The deflection of the cantilever is sensed with optical detection.

TABLE I
COMPARISON OF AFM AND OTHER MICROSCOPE TECHNIQUES.

<table>
<thead>
<tr>
<th></th>
<th>AFM</th>
<th>TEM</th>
<th>SEM</th>
<th>Optical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Res.</td>
<td>Atomic</td>
<td>Atomic</td>
<td>1's nm</td>
<td>100's nm</td>
</tr>
<tr>
<td>Typical Cost (K$)</td>
<td>100–200</td>
<td>≥ 500</td>
<td>200–400</td>
<td>10–50</td>
</tr>
<tr>
<td>Imaging Environ.</td>
<td>air, fluid, vacuum, special gas</td>
<td>vacuum</td>
<td>vacuum</td>
<td>air, fluid</td>
</tr>
<tr>
<td>In-situ</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>In-fluid</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Sample prep.</td>
<td>Easy</td>
<td>Difficult</td>
<td>Easy</td>
<td>Easy</td>
</tr>
</tbody>
</table>

generated false color, to allow the user to discern height. Furthermore, the AFM can be used to characterize surface properties beyond the topology. An example of this is shown in Figure 3. In this case a magnetically sensitive coating is deposited on the tip so that beyond measuring the topology, the magnetic domains in a surface can be measured. Even though the topology has few obvious features, the map of the magnetic domains reveals a detailed structure.

The first applications were found in materials science and this field continues to make use of AFM today. Recent studies include experiments to understand the nanoscale-phenomena underlying improved photovoltaic cells [19], surface forces [20], thin films [21]-[23], crystallization [24], [25], and semiconductor properties [26]-[28].

Soon after its invention, it was recognized that the AFM can be used to image compliant samples, including biologically relevant materials. This capability has been taken advantage of to study mechanical properties and dynamics.
from the level of single cells down to single molecules. A few notable applications at the cellular level include the in situ study of drug-induced changes in cell structure, membrane stability, and receptor interaction forces [29] and the study of cell motility [30], [31]. Studies of single molecules have included the activity of RNA polymerase [32], [33], the motion of molecular motors such as proton-powered turbines [34] (see Figure 4) and myosin V [35], the transcription process [36]-[38] and the structure of a wide variety of viruses [39]-[42] (see Figure 5). AFM has also been used extensively as a force transducer to study the mechanical properties of biological structures and the forces of molecular interactions [4], [43]-[45].

The AFM is also capable of manipulating material and is a useful actuator for nanotechnology. The tip can apply a variety of forces, including contact, magnetic, thermal, and electrical using modified tips. It has been used in lithography [48], [49], in nanomanipulation [50]-[53], and in nanoassembly [2], [54], [55]. Another interesting application is the “millipede” project at IBM Research [56], [57]. This device consists of an array of cantilevers, operated in parallel (see Figure 6) and has the potential to achieve data storage densities of 1 Tb/in².

With the continued interest in understanding materials and biological systems at the nanoscale and with the promise of nanotechnology, AFM will continue to be an extremely important tool in the researcher’s toolbox.

III. A WALK AROUND THE AFM CONTROL LOOP

This description of the AFM loop will discuss contact (or constant force) mode, since it is the easiest to understand. Dynamic (or AC) mode will be described in Section V.

A schematic block diagram of a typical AFM control loop is shown in Figure 1. The AFM loop starts with a sample to image. The sample is typically on a surface which is scanned back and forth in a raster pattern. A sharp tip on the end of a cantilever is brought into close proximity to the surface where the interaction between the tip and surface (Figure 7) causes the cantilever to deflect. The nonlinear tip-sample interaction force can be represented by various models. Two popular models are the DMT (Derjagin, Muller, Toropov) model (e.g., [58]) and the Lennard-Jones potential combined with a modified Hertz model (e.g., [59]). The Lennard-Jones potential [10], [11] results in the interaction force,

\[ F(r) = k_1 \left( -\frac{\sigma}{r} \right)^2 + \frac{1}{30} \left( \frac{\sigma}{r} \right)^8, \]

where \( \sigma \) is an interaction parameter, \( r \) is the distance between the tip and the sample, \( k_1 \) is a constant which depends on the geometry and material of the tip and the sample, and \( F(r) \) represents the force between a spherical tip and a flat sample as a function of \( r \). The force of interaction between tip and sample in this model shows an attractive component (the first term) due to the Van der Waals’ forces and a repulsive component (the second term), which is attributed to the Pauli principle. More details about these models can be found in [10], [11], [58], [59], but the tip sample interaction has a general shape as shown in Figure 7, generated by (1). The nonlinearity of the interaction force clearly shows why...
Fig. 5. Isolated Human Immuno-deficiency Virus (HIV) particles. Individual viruses were fixed on glass cover slips and imaged using dynamic mode AFM under ethanol. (a to d) Groups of virus particles adhering to the glass substrate. The tendency to form clusters is likely due to the method used to isolate the viruses and prepare them for imaging. (e and f) Two isolated viruses imaged at high resolution showing the distinctive but arbitrary distribution of protein tufts covering their exterior surfaces. The roughly spherical particles have average heights of 120 nm, although some, as in panel f, are seen to be slightly compressed, probably due to contact with the substrate. The particles appear to be soft and easily deformed from a spherical shape. The images seen here are typical of many such particles found on the substrate. (Reprinted with permission from [47]. ©2003, the American Society for Microbiology.)

Fig. 6. The millipede project. 1024 cantilevers are operated in parallel for data storage applications with potential data densities at Tb/in². (Reprinted with permission from [56]. ©2003 IEEE.)

feedback operation for tracking the sample topography is crucial for obtaining reliable data about the sample surface. The forces acting on the tip cause the cantilever to deflect. The resulting deflection is measured by reflecting a laser spot off of the back of the cantilever onto an optical detector. The deflection signal from the optical detector is compared to a nominal deflection value, denoting the imaging force, and this difference is minimized using a feedback controller. The control signal itself is typically used as an estimate of the surface profile.

There are several variants on this. In a typical scanning sample design, the sample is moved below a stationary tip. The X, Y, and Z actuation are done by a single piezo tube actuator [60]. The Z actuation is done in closed-loop with the sample being moved vertically in response to the deflection of the cantilever. In a scanning tip design, the sample is stationary while the tip is moved in X, Y, and Z. In a third design, the X-Y motion is handled by a stage that moves the sample while the Z motion is handled by an actuator moving the cantilever up and down. The issues related to the choice of design are discussed in [61], [62].

The choice of designs depends greatly on the type of AFM measurement to be done. The single piezo tube actuator (used for scanning either the sample – as in the Veeco Multimode (www.veeco.com) – or the tip – as with the Agilent PicoPlus (www.molec.com) and the Veeco Dimension) – is the lowest cost. However because of the lower X-Y bandwidth and the bowing effect on the image, the typical maximum scan ranges are between 10 and 200 µm, depending on the actuator geometry and material. The scan ranges for separate commercial X-Y actuators quoted in the literature are between 0.4 and 400 µm [63]-[66]. While these ranges are comparable to the piezo tube scanners for most applications, the advantages of closed-loop operation and decoupling from the Z actuation often justify the extra hardware. The effects of mechanical cross coupling on the AFM control loops are discussed in [67]-[70]. Techniques for improving the mechanics of the system to achieve higher control bandwidths are discussed in [61].

A. The Cantilever and Tip

Drawings of typical cantilevers are shown in Figure 8. A SEM image of a NANOSENSORS bar cantilever is shown in Figure 9. The tip material can be chosen for specific properties of the surface it will interact with. There are two basic designs of cantilevers. The most common is the thin rectangular bar “diving board” shape, used in contact and AC mode operation (Figure 8A). With these cantilevers both the first bending mode and the torsional mode can be detected, using a quad photo detector. Triangular cantilevers
(Figure 8B) have been used in contact mode AFM with the idea that these would be more resistant to torsional twisting, but recent results have shown that these cantilevers are in fact more sensitive to torsion [71].

The most common materials for cantilevers are monocrystalline silicon (Si) and silicon nitride (Si₃N₄). Often cantilevers receive a coating of metal on the back side to improve their reflectivity for optical detection. For the rectangular bar cantilevers used in today’s commercial AFMs, typical parameter ranges are shown in Table II.

As can be seen from Table II, cantilever designs vary depending upon whether they are to be used for contact mode or AC mode. Contact mode cantilevers typically are more flexible but have lower resonant frequencies, while AC mode cantilevers are designed to have a specific resonant frequency. The resonance frequency will track with the spring constant of the cantilever. Going to lower resonant frequency cantilevers should, to a first approximation, lower the tip sample interaction force. Thus, for applications which benefit from a lower contact force in AC, the mid-frequency cantilevers can be used for “gentle AC.” Magnetic AC (MAC) uses cantilevers in this frequency range, but they have been coated with magnetic material on the back side to make them sensitive to the oscillations of a magnetic coil. Generally, different types of cantilevers are chosen depending on the sample and measurement mode.

The real-time capability of the AFM control system is a large part of what enables different measurement modes to be performed. As will be discussed in Section V-B, there are some advantages to higher frequency AC mode cantilevers, but the use of these is limited by the ability of the real-time system to extract the servo and surface information.

### B. Optical Position Detection

![Image of optical lever detection](image)

![Diagram of cantilever deflection detection](image)

A typical diagram of the “optical lever” method of detecting cantilever deflection is shown in Figure 10. A laser,
typically of wavelength around 690 nm, is reflected off of the back of the cantilever. When the cantilever is at a neutral deflection, the spot falls on the center of a photo-detector. As the cantilever is deflected more or less, the spot moves on the detector, providing a push-pull error signal. One of the key limiting factors of an AFM is therefore the noise in the optical detection system, since this is the sensor noise that will feed straight through to the closed-loop output (the tip position) [72]. Figure 10A shows the detection of the first bending mode of the cantilever, which is considered deflection or amplitude. The amplitude signal is the response of the cantilever in the vertical direction. This is detected by forming

\[ e = (A + D) - (B + C) \]  

in either an analog or digital circuit. In order to eliminate the influence of laser intensity fluctuations, the error signal can be normalized by dividing by the total optical intensity on the detector:

\[ e_{nor} = \frac{(A + D) - (B + C)}{A + B + C + D}. \]  

Figure 10B shows the detection of the first torsional mode of the cantilever, which is considered the friction signal. The friction signal is the response of the cantilever in parallel with the direction of the scan. This is detected by forming

\[ f = (A + B) - (C + D) \]  

in either an analog or digital circuit. As with the error signal the friction can be normalized,

\[ f_{nor} = \frac{(A + B) - (C + D)}{A + B + C + D}. \]  

C. Actuation

The most common form of actuation is a three degree-of-freedom piezo tube as developed by Binnig and Smith [60], shown in Figure 11. The structure is made up of a single tube of piezo material. The outside of the tube is actuated by 4 electrodes that move the tube in the X and Y directions. Another electrode actuates the piezo in the Z direction.

Fig. 11. Three degree-of-freedom piezo tube actuator. The structure is made up of a single tube of piezo material. The outside of the tube is actuated by 4 electrodes that move the tube in the X and Y directions. Another electrode actuates the piezo in the Z direction.

This type of actuator is compact and cost efficient, making it the basis for many commercial AFMs, including models from Veeco (www.veeco.com), Agilent Technologies (www.agilent.com), Quesant Instruments (www.quesant.com), NT-MDT (www.ntmdt.ru), and others.

There are two main drawbacks of such actuators. First of all, the X and Y directions are typically not provided with sensors, so the control of the X and Y directions is accomplished in open loop. Second, there can be coupling between the X-Y and Z directions [68], [70], [73]. A common method of decoupling an AFM system is to actuate the Z axis with a separate piezo from the X-Y actuator.

Fig. 12. A separate X-Y actuator. The sample is moved in-plane by the actuator, while the Z actuation is done separately. The frame within a frame approach decouples the motion of the X and Y stages, as do the flexures which are soft in the direction of applied motion and stiff in the orthogonal direction. Note that the sample size is not drawn to scale.

Fig. 13. The raster scan motion. The alternating shade areas represent the pixels of the image. Note that each pass of the X scan generates one line of pixels in an image. Passes in opposite directions are used for separate images.

1) X-Y Scanner: A separate X-Y actuator is shown schematically in Figure 12. A typical actuator includes an outer frame and an inner frame. This frame within a frame approach is designed to minimize the mechanical coupling
between X and Y motion. The inner frame is moved relative to the outer frame by means of actuators along a side. The flexures on the inner frame are flexible in the X direction and stiff in both the Y and Z directions. The flexures on the outside frame are flexible in the Y direction and stiff in the X and Z directions. The sample is placed in the center of the stage. The motion of the inner frame is measured in a direction by means of a sensor. Several options exist for these sensors, including capacitive, strain gauge, and LVDT sensors. Quite often, the stage is asymmetric, as one direction is moved much more rapidly than the other.

This approach decouples the X and Y directions from the Z direction. It eliminates the cross coupling and image artifacts associated with this coupling. An example of this is a bowing of the image that results from the piezo tube moving away from the surface as the X and Y position move away from the nominal point. Although these artifacts can be removed through image processing by accounting for the geometry of the problem, they are completely removed through the use of a separate X-Y actuator. Furthermore, these actuators allow for the implementation of sensors that enable closed-loop control of the X and Y positions. Typical X-Y actuators have resonances in the range of 200 Hz to 1.5 kHz, although these frequencies are often reduced in practice by the loading of the sample mass. The scanning ranges for commercial X-Y actuators are anywhere from 0.4 μm to 400 μm on a side [63]-[66]. A measured frequency response function for a nPoint (www.npoint.com) NPXY100A stage is shown in [67]. A more detailed discussion of the control of that stage is given there.

The most common form of scanning uses a triangle wave in the X-direction and a linear ramp in the Y-direction, as shown in Figure 13. In the scanner of Figure 12, the inner frame is chosen as the X-direction since it is less massive and thus can be moved more quickly for a given amount of energy. If the relative speeds of the two axes are properly set, each forward scan of the X-direction produces one line of pixels. The return scan in the X direction produces a second line of pixels. These two directions are not typically combined because the nonlinear coupling of the tip to the surface is different in each direction and combining them would distort the image.

The triangular scan pattern means that the tip spends the same amount of time over each pixel (except at the turnaround points). However, due to finite bandwidth of the scanner, this is an impossible curve to match. The distortion of the scan curve can be minimized through a combination of feedback and feedforward methods, which are discussed in [67].

2) Z Actuation : An example set of frequency response curves for a piezo tube is shown in Figure 14. The piezo tube resonances shown here are around 1 kHz, which is in the typical range of 500 Hz to 20 kHz. Some experimental systems have resonances above 40 kHz [75], [76].

In Figure 14 a series of five models of the piezo-cantilever system are plotted with the resonant frequency varying between 900 Hz and 1.1 kHz, and the quality (Q) factors varying between 10 and 30. The uncertainty results both from variation across multiple actuators and variation of the same actuator with varying signal amplitudes and environmental conditions. At higher frequency one sees a nominal 300 kHz cantilever resonance with a nominal Q of 100. Again, there is variation across different cantilevers in the same batch and different conditions for the same cantilever. All of this illustrates the need to do identification, whether as a preliminary calibration step or in an on-line form.

Note that these plots are idealized in that they neglect any extra dynamics – including non-minimum phase zeros – typically present in the actuator and cantilever. Furthermore, any dynamics of the electrical circuitry, such as low pass effects of the power amplifiers used to drive the piezos are neglected. Finally, these plots show no effects of transport or computational delay. However, even when using such an idealized model, the significant limitations and issues with AFM control are evident.

The effects of this structure on the feedback system can be immediately seen. If a feedback controller is to include a 300 kHz resonance in the model, then a typical rule of thumb sample rate of 10-20 times the highest dynamics of interest would imply a 3 – 6 MHz sample rate for the control system. Obviously, such a high sample rate puts severe constraints on the signal processing system, not just in accomplishing the needed processing between samples, but also in minimizing the latency of the computations, signal conditioning, and data conversion.

On the other end of the spectrum are control systems that will restrict bandwidth to be safely below the Z-piezo actuator’s resonance. For a 1 kHz resonance, this implies a sample rate of no less than 10 kHz. Thus, a typical sample rate for control on an industrial AFM is in the 50–100 kHz.
range [63], although newer controllers sample considerably faster – up to 500 kHz in the case of [77].

IV. FEEDBACK CONTROLLER

![Controller Responses (f_S = 50 kHz)](image1)

Fig. 15. PI and PII controllers for Z-piezo actuator and cantilevers of Figure 14.

Because the piezo actuator is modeled as a second-order resonance, the lack of integrators in the forward path necessitates the use of integral action for zero steady-state error to any steps in the surface height. The addition of a second integrator via PII control can provide zero steady-state error to surface slopes, which are common in many samples. Such controllers are necessarily low bandwidth, since the lack of phase lead means that the gain must be rolled off below the resonance of the actuator.

A look at the configuration of Figure 1 shows that the fundamental feature of this loop is that the control system only sees the deflection, not the surface. As such it is an output error loop, without direct access to any reference signal. This limits any attempt at feedforward in the Z direction to methods that use some prior Z measurement (such as the previous scan line). It also limits the bandwidth of any state-space controller that one may use, since the estimator error can go away no faster than the error in the overall control loop [78]. Furthermore, as one sees from the typical example shown in Figure 14, there is considerable variation in the response of the actuator at low frequency and overall control loop [78]. Furthermore, as one sees from the open-loop plots of Figures 17 and 19, where the PI controller has significantly less bandwidth, but also less ringing than the PII controller. The difficulty in finding a single robust controller for these varying plants which provides both reasonable bandwidth and acceptable gain and phase margins illustrates why there is so much hand tuning of AFM control loops by the end users of the instruments.

The typical industrial AFM control loop, whether done in contact or dynamic mode, is a low frequency PI or PII loop. A general form of an analog controller that admits PI, PD, PID, PII, and even PIID is:

\[
C(s) = \left( K_p + \frac{K_i}{s} + \frac{K_{ii}}{s^2} + K_ds \right) E(s)
\]

where \(E(s)\) is the Laplace transform of the error signal \(e(t)\). For a P, PI, PII, or PID controller, one or more of the \(K_i\),\( K_{ii}\) gains are set to zero. Note that as written the derivative term, \(K_ds\), is not practically implementable, but this is often rectified by having some low pass filter added to it. For digital implementation, the backward rectangular integration rule is most often used for PID controllers since this allows for direct translation from (6) [78], [79].

It is tempting to try to increase the bandwidth of the system by adding phase lead, such as with a PID controller. However, the use of this is limited by the uncertainty in the modeling of the piezo actuator. Furthermore, boosting the bandwidth with a PID requires lower noise in the optical measurement of deflection, otherwise this noise will be amplified by the effects of the derivative term.

For the model in Figure 14, a pair of controllers (PI and PII, \(K_d = 0\)) was synthesized as shown in Figure 15. The system was sampled at 50 kHz, and no attempt was made to add any extra computational or transport delay. Thus, the open-loop plots of Figures 16 and 18, representing the application of the PI and PII controller respectively, should be considered an idealized case. What is clear in these plots is that the open-loop crossover frequency must be substantially below the nominal resonant frequency for there to be any gain margin. Furthermore, the low frequency gain is quite limited in the case of the PI controller. The PII controller has more gain at low frequency, at the expense of decreased phase margin. The effects of these choices become clear in the closed-loop plots of Figures 17 and 19, where the PI controller has significantly less bandwidth, but also less ringing than the PII controller. The difficulty in finding a single robust controller for these varying plants which provides both reasonable bandwidth and acceptable gain and phase margins illustrates why there is so much hand tuning of AFM control loops by the end users of the instruments.

![Open-Loop Responses --- PI Controller (f_S = 50 kHz)](image2)

Fig. 16. Open-loop response for piezo/cantilevers of Figure 14 with digital PI controller of Figure 15.

Because tube scanners often lack sensors, much of the feedback control work is done only in the Z direction, leaving...
compensation of the X-Y directions to be done using open loop methods [74], [80]-[82]. Because scanning is most often a raster scan, with a fast axis (X) and a slow axis (Y), the compensation is often applied only to the fast axis.

Note also that the 50 kHz sample rate is only reasonable for actuators with their significant dynamics below about 5 kHz. For smaller actuators - such as those being proposed in higher bandwidth experiments - the control has to be done either with faster sampling or an analog controller [76], [83]-[85].

These issues are fundamental to the control of an AFM. The desire for a single robust, low-order controller is thwarted by the uncertainty in the system. The solution involves either an improved model and/or a higher-order robust controller. Because tube scanners often lack X-Y sensors, much of the original advanced feedback control work was done in the Z direction [59], [86], while feedforward controllers were developed for the X-Y motions [74], [81], [82]. The advent of sensored X-Y stages has led to feedback control methods being developed for X-Y motions as well [87], [88]. Combined feedforward and feedback controllers have also been investigated for both the Z and X motions [89], [90], [91].

In demonstrations of advanced control for nanopositioning, researchers have made careful models of a specific AFM under controlled conditions and then have been able to achieve significantly higher closed-loop bandwidths. While robust control methods may provide practical controllers in the presence of model uncertainty, development of adaptive control methods for AFMs remains an open area that may provide enhanced performance. Further discussion of the control problem from a multi-axis point of view is provided in [67]. An overview of the issues in AFM control and methods available to address these are given in Sections VI and VII, respectively.

V. MODES OF OPERATION

The two most common modes of operation for AFMs are known as contact mode and dynamic (or AC) mode. In either mode, it is important to recall that the servo system does not have access to the “reference” signal from the surface, making it an output error loop. The reference deflection (for contact mode) or the reference amplitude (for dynamic mode) are effectively constants. The surface is most commonly treated as an unknown disturbance input, so effectively this is an output error problem. As such, the surface estimate must come from the feedback loop itself. In commercial systems, the surface estimate comes from some function of the control signal. In some academic experiments, state-space methods are used to calculate the surface from an estimator [59], [89], [91]-[93].
Control Signal
Surface
Deflection
Fig. 20. Deflection of the AFM tip in contact mode. Note that the optical lever gives a signal proportional to an error signal. The control signal being sent to the actuator is a good, albeit band-limited, representation of the surface.

A. Contact Mode

In contact mode, the tip is dragged across the surface with the feedback loop minimizing the deflection of the cantilever away from its nominal position. By controlling the deflection of the cantilever, the force of the interaction with the surface is controlled, and thus this mode is also known as constant force mode. This mode is used when imaging materials that are not adversely affected by being in sheer with a sharp tip.

Figure 20 shows the key signals in a contact mode line scan across a surface with a step in height. As the tip is moved over this step, the interaction of the tip with the surface causes the cantilever to deflect, and this is detected on the optical detector (Figure 10). This change in deflection is seen as an error by the controller which moves the actuator away from the surface. The integral effect of the controller allows it to achieve zero steady state error and the control signal itself becomes a representation of the surface. As the surface drops away, the deflection once again moves away from the nominal value (but in the opposite direction). As before, the control loop responds to minimize this, and the control signal again represents the surface. The ability to control the cantilever deflection and image the surface is thereby limited by the bandwidth of the closed-loop system.

For contact mode, Table II shows that the cantilevers used have low bending mode frequencies (around 20 kHz). This puts them well above the actuator bandwidth, but far below that of AC mode cantilevers. The advantage of these contact mode cantilevers is that the force constant is typically much smaller than the AC mode cantilevers, allowing the tip to be dragged across a surface with less damage to either, as in the example of Figure 21.

B. Dynamic Mode

Figure 22 shows the general block diagram for dynamic mode AFM, which involves an oscillation of the cantilever in the proximity of the surface at a frequency close to the resonant frequency of the cantilever. In non-contact mode, the amplitude of the oscillation is slightly less than the nominal tip/surface distance so that while there is interaction between the tip and surface, this never enters into what would be considered contact. In the most common form of dynamic mode, also known as AC mode, intermittent contact mode, or by the trademarked name “tapping mode” [94], [95], the amplitude of the free oscillation is slightly larger than the nominal tip/surface distance. When the tip comes into proximity with the surface, the oscillation amplitude, phase, and frequency are modulated, as shown (for amplitude) in Figure 23. By detecting this modulation and closing a feedback loop on the amplitude of the oscillation, this amplitude can be maintained at a constant level (modulo the bandwidth of the system). Once again, the control signal represents the surface topography, as shown in Figure 24.

Dynamic mode imaging is done using cantilevers of various frequency ranges (Table II). Often as the cantilever resonant frequency goes up, they get stiffer and have a higher Q. The higher Q provides greater amplitude amplification of the drive signal and better frequency discrimination for small shifts due to surface interaction. However, the extra stiffness of the cantilever might damage some materials, so there is a tradeoff to be made on increasing the cantilever resonance. Because dynamic mode produces lower sheer forces on the sample than contact mode, the imaging of biological samples, such as the human chromosomes in Figure 25, is often done using this technique.

Because dynamic mode typically operates near the cantilever resonance [58], there is a relationship between the amplitude shift, phase shift, and frequency shift seen due to the surface/tip interaction. Thus, both the imaging and the Z-axis servo loop can be driven by one of several demodulated signals.

- **Amplitude Modulation (AM):** In this mode, the change in the amplitude of the oscillation is detected and used as the error signal for the feedback loop. The speed of AM-AFM is often limited by the high Q-factor of the cantilever, which slows the detection of surface features through the *Wile E. Coyote effect* seen...
in Figure 23, in which a tip goes off a cliff on the surface but doesn’t detect it for a while.

- **Phase Modulation (PM):** In this mode, the change in the phase difference between the cantilever drive and the returned deflection signal is detected. Applying feedback on the amplitude is easier to implement. However, simultaneously the phase signal can be used to measure other surface properties like energy dissipation [96].

- **Frequency Modulation (FM):** In this mode, the change in the oscillation frequency of the returned deflection signal is detected. FM-AFM typically requires extremely high-$Q$ cantilevers so that the frequency shift can be detected. This has meant that FM-AFM is most often done in a vacuum where the lack of air damping makes the cantilever $Q$ seem much larger. However, non-vacuum operation has been made possible by recent improvements in instrumentation.

1) **Actuating the Cantilever:** The $Z$ axis piezo actuators used in a typical AFM are relatively slow. Most piezo actuators are characterized with a simple second-order model, and for these, the resonance on the $Z$ actuator of a piezo tube is at a few kilohertz. This makes the standard piezo actuator unsuitable for providing stimulus for dynamic mode AFM. To compensate for this, several options are available, including a small piezo element to shake the cantilever (e.g., [84], [85], [94]) or vibrating the sample with acoustic means, directly actuated cantilevers [97], or magnetic actuators. This latter approach is used to generate Figure 25.

2) **Signal Demodulation:** Along with actuating the cantilever in AC mode, it is necessary to demodulate low frequency information from the optical deflection signal. Originally, the amplitude of the read-back signal was demodulated using non-synchronous demodulation via a RMS-to-
DC circuit. More recently, coherent demodulation, through the use of a lock-in amplifier has been used to extract both the magnitude and phase of the read-back signal.

AC mode usually has lower bandwidth than contact mode for several reasons:

- In AC mode, information about the surface is only available during the contact interval, which happens once every period of the oscillation. To have statistical significance, it is typical to average over multiple contact points, typically on the order of 10. Thus, the time constant of the vertical control loop is limited by the frequency of oscillation and the number of periods required.
- The Q factor of the cantilever affects the time response. The cantilever is usually oscillated near its resonant frequency to get reasonable deflection amplitudes with low levels of input signal. Due to nonlinear interactions with the surface, the tip oscillation amplitude responds almost instantaneously to a step up in the surface (see the left side of Figure 23). However, when there is a step down in the surface height, the response time of the cantilever oscillation will be proportional to $Q/\omega_o$, where $\omega_o$ is its resonant frequency (see the right side of Figure 23) [98]. The flywheel action (the Wile E. Coyote effect), also introduces a limitation on the imaging speed without imaging artifacts.
- The method used to demodulate the amplitude from the oscillatory deflection signal affects the time response.

VI. ISSUES IN AFM CONTROL

Despite their utility, there are substantial issues in the use of AFMs, and most of these lead back to control problems. A partial list includes:

- **Ease of use:** Unlike most instruments, an AFM usually requires an expert operator. This limits the utility and raises the expense of operation since measurements can only be made by a select few people. Furthermore, the exchange of cantilevers and tips leads to a need to readjust the system.
- **Repeatability/calibration:** Each measurement, each sample, and each new cantilever/tip combination requires the system to be adjusted again. Furthermore, images are rarely calibrated in an absolute sense. That is, the height measurements are truly estimates based on calibration samples and not individually known.
- **Speed of measurement:** AFM measurements tend to be slow. The features and size of the sample place spatial bandwidth requirements on the servo system. The actuators (Z and X-Y) have their own dynamic limitations. In combination, this means that depending upon the resolution of the desired image, AFM scans can take anywhere from under a minute to large fractions of an hour.

These issues largely arise from the characteristics of the piezo actuators used in the AFM. As mentioned in Section III-C.2, these actuators have a first resonance at frequencies ranging from 500 Hz to 20 kHz and their behavior is hysteretic. This makes reliable modeling of the actuator more difficult [74].

While piezo actuators enable high precision positioning, nonlinear hysteresis effects can significantly reduce the accuracy in long-range positioning, such as when imaging large (e.g., biological) samples. Another cause for loss in precision in piezos is drift due to creep effects, which become significant when positioning is required over extended periods of time (such as during slow operation of AFMs). Model-inversion based controllers [74] have been used to compensate for hysteresis and creep, and both integral and $H_\infty$ controllers [88] have been shown to mitigate these nonlinear effects in the closed-loop system behavior of AFMs.

Thermal noise of the cantilever is a fundamental limiting factor for AFM speed because the noise [99]-[101] feeds directly into the error signal. While the noise limits the eventual accuracy of the tip control, it also limits how much lead one can add to a control loop for a stable image. This accounts for the fact that most drive control loops have little or no lead. ($K_d = 0$ for PID controllers.)

VII. ADVANCED AFM CONTROL TOPICS

To deal with the issues presented in Section VI, there have been various thrusts. These include efforts geared at increasing the performance of conventional tube scanner-based microscopes through understanding and compensating for the nonlinear effects of hysteresis, creep, and varying voltage response [74], [102]. Many researchers have attacked the speed of measurement problem with approaches including a redesign of the actuator to achieve higher bandwidth [76], and applications of modern control theory to increase the scanning speed [74], [81], [82], [87], [88], [89]. These methods are discussed in more detail in [61]. Recent efforts include attempts to decrease the number of scan points through non-raster methods of generating images [103],
A survey of such non-raster methods is presented in [105]. In [67], an overview of several advanced control methods that have recently been developed and applied on AFMs is given. In particular, the use of $H_\infty$ in the feedback path only is described in [59], [87], [88]. The use of combined feedforward/feedback controllers is discussed using $H_\infty$ [89], [91], $\ell_1$ [90], and model-inverse based methods [68], [74]. Areas of future work are also outlined.

Speeding up dynamic (AC) mode AFM has been pursued by several researchers. One approach for faster AC mode imaging is to control the $Q$ of the cantilever. In [106], [107], a secondary piezo actuator has been built on the cantilever to allow an inner feedback loop to control the $Q$ of the cantilever. Another approach for faster AC mode imaging is to control the $Q$ of the Z piezo via an internal model as done in [84], [85]. These authors also try to speed up the detection of the amplitude of the oscillatory deflection signal by differencing the peaks of the oscillation at each half period [35]. However, this method is similar to peak detection – a method that can be quite susceptible to amplitude noise.

It is important to note that speeding up the AFM forces some very practical decisions about the implementation of the control law. While commercial AFMs have sample rates that are in the range of hundreds of kHz, this does not allow for demodulation of signals from 300 kHz cantilevers. Thus, a lot of higher speed experiments are done with analog electronics [83]-[85]. This is also the case with high speed experiments on faster actuators such as in [76]. Some commercial system controllers are moving to higher sample rates. For example, the Veeco NanoScope V has an output rate for the feedback loop of 500 kHz [77]. Sample rates this high force the control designer to give a lot of thought to how the control computations will be done. For example, these new controllers make use of the parallelism of FPGAs to speed up their operation. However, implementing signal processing on FPGAs can be much more tedious than floating point DSP programming. Furthermore, these sample rates are still too low to capture some of the higher harmonics of AC mode cantilevers. To enable digital demodulation of these higher harmonics, some manufacturers have moved to sampling the data at much faster rates, such as a single 5 MHz channel for the Asylum MFP-3D [63] or a single 50 MHz capture channel for the Veeco NanoScope V [77]. While control is not done at these frequencies, the data can be post-processed off line.

VIII. CONCLUSIONS

In this tutorial, we have examined the control of AFMs from a systems and control engineering point of view. The discussed spectrum of AFM applications demonstrates the versatility of this instrument. A walk around the AFM control loop discusses several components and points out potential bottlenecks in these kinds of instruments, which depend on the physics and technology behind each of the components. Efforts to improve the performance of AFMs typically involve an attempt to improve one or more of the loop components. This may be done either by re-designing the individual components or by implementing a better controller for the AFM, or both. However, it is the view of the authors that the performance of AFMs can truly be optimized only through a systems understanding of how adjustments to these components affect the overall feedback loop. In summary the AFM is an important instrument, which already has proven its huge potential for several applications on the nanometer scale. We believe that modern control engineering can significantly contribute to improve these systems even further and turn them from scientific instruments into well engineered machines for an even wider range of applications throughout various disciplines.

REFERENCES
