BEILSTEIN JOURNAL OF NANOTECHNOLOGY

Advanced atomic force microscopy techniques

Edited by Thilo Glatzel and Udo D. Schwarz

Imprint

Beilstein Journal of Nanotechnology
www.bjnano.org
ISSN 2190-4286
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Nanotechnology* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

Beilstein Journal of Nanotechnology

Advanced atomic force microscopy techniques

Thilo Glatzel*[1], Hendrik Hölscher[2], Thomas Schimmel[3], Mehmet Z. Baykara[4], Udo D. Schwarz*[5] and Ricardo Garcia[6]

Editorial

Open Access

Address:
[1]Department of Physics, University of Basel, Klingelbergstr. 82, 4056 Basel, Switzerland, [2]Institute of Microstructure Technology (IMT), Karlsruhe Institute of Technology (KIT), Campus North, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein Leopoldshafen, Germany, [3]Institute of Nanotechnology (INT), Karlsruhe Institute of Technology (KIT), 76021 Karlsruhe, Germany, [4]Department of Mechanical Engineering, Bilkent University, Ankara 06800, Turkey, [5]Department of Mechanical Engineering and Materials Science, Yale University, P.O. Box 208284, New Haven, CT 06520-8284, USA and [6]Instituto de Microelectronica de Madrid, CSIC Isaac Newton 8, 28760 Tres Cantos, Madrid, Spain

Email:
Thilo Glatzel* - thilo.glatzel@unibas.ch; Udo D. Schwarz* - udo.schwarz@yale.edu

* Corresponding author

Keywords:
atomic force microscopy

Beilstein J. Nanotechnol. **2012,** *3,* 893–894.
doi:10.3762/bjnano.3.99

Received: 04 December 2012
Accepted: 07 December 2012
Published: 21 December 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Editor-in-Chief: T. Schimmel

Although its conceptual approach is as simple as the technique used in record players already introduced in the 19th century, the invention of the atomic force microscope (AFM) in 1986 by Binnig, Quate, and Gerber was a milestone for nanotechnology. The scanning tunneling microscope (STM), introduced some years earlier, had already achieved atomic resolution, but is limited to conductive surfaces. Since its operational principle is based on the detection of the forces acting between tip and sample, this restriction does not exist for the AFM. Consequently, atomic force microscopy quickly became the standard tool for nanometer-scale imaging of all types of surfaces in all environments. True atomic resolution was first achieved in the 1990s. The most convincing results, however, were restricted to the so-called noncontact mode in vacuum for a long time, but recent technical developments overcame this limitation, and atomic-resolution imaging is now also a standard in liquids. Beyond pushing the resolution limit to the picometer range, the invention of the AFM triggered the development of a growing number of new scanning probe methods and approaches, ranging from an expansion of the properties that can be mapped to the active manipulation of surfaces and small particles. Practically every month, reports on the growing capabilities of AFMs appear. Nearly every physical effect that influences the tip–sample interaction has been used to improve existing modes and to develop new ones. For example, many recently presented techniques include the excitation of higher cantilever oscillation modes; it is amazing in how many ways the shaking of a simple cantilever can improve our knowledge about the tip–sample interaction. Another direction is high-speed atomic force microscopy, which is one of the eminent challenges that need to be solved in order to allow the in situ observation of biological processes. Data acquisition times have already

reached the millisecond range, enabling the visualization of the dynamic behavior of biological molecules and cells. Other recent accomplishments include imaging of organic molecules with unprecedented resolution, full three-dimensional mapping of surface force fields, and the imaging and discrimination of individual chemical bonds.

The development of advanced techniques is the focus of this Thematic Series, following the Thematic Series “Scanning probe microscopy and related techniques” edited by Ernst Meyer and the Thematic Series “Noncontact atomic force microscopy” edited by Udo Schwarz. The articles that are part of the series demonstrate that, despite its 25 years of history, the AFM is still far from reaching its limits, and today’s developments are far-reaching. As the number of research groups utilizing advanced atomic force microscopy techniques increases with each passing year, the technical improvements, data-acquisition approaches, analysis procedures, user friendliness, and application areas of the technique further diversify. With this Thematic Series, it is our intention to stimulate these improvements.

We thank all authors for contributing their excellent work to this series. Furthermore, we acknowledge all referees for their promptly provided reports keeping the publication times short and attractive for contributors. Finally, we are grateful to the open access policy of the *Beilstein Journal of Nanotechnology* providing the ground for unrestricted discussions on advanced atomic force microscopy techniques.

Thilo Glatzel, Hendrik Hölscher, Thomas Schimmel, Mehmet Z. Baykara, Udo D. Schwarz and Ricardo Garcia

December 2012

Beilstein Journal of Nanotechnology

Probing three-dimensional surface force fields with atomic resolution: Measurement strategies, limitations, and artifact reduction

Mehmet Z. Baykara[*1,2,3], Omur E. Dagdeviren[1,2], Todd C. Schwendemann[4], Harry Mönig[1,2,5], Eric I. Altman[2,6] and Udo D. Schwarz[*1,2,6]

Full Research Paper

Open Access

Address:
[1]Department of Mechanical Engineering and Materials Science, Yale University, New Haven, CT 06520, USA, [2]Center for Research on Interface Structures and Phenomena (CRISP), Yale University, New Haven, CT 06520, USA, [3]Department of Mechanical Engineering, Bilkent University, Ankara 06800, Turkey, [4]Physics Department, Southern Connecticut State University, New Haven, CT 06515, USA, [5]Physikalisches Institut at the Center for Nanotechnology (CeNTech), Westfälische Wilhelms-Universität, 48149 Münster, Germany and [6]Department of Chemical and Environmental Engineering, Yale University, New Haven, CT 06520, USA

Email:
Mehmet Z. Baykara* - mehmet.baykara@bilkent.edu.tr;
Udo D. Schwarz* - udo.schwarz@yale.edu

* Corresponding author

Keywords:
atomic force microscopy; force spectroscopy; NC-AFM; three-dimensional atomic force microscopy; tip asymmetry; tip elasticity

Beilstein J. Nanotechnol. **2012,** *3,* 637–650.
doi:10.3762/bjnano.3.73

Received: 13 June 2012
Accepted: 23 August 2012
Published: 11 September 2012

This article is part of the Thematic Series "Noncontact atomic force microscopy II".

Associate Editor: B. Bhushan

Abstract

Noncontact atomic force microscopy (NC-AFM) is being increasingly used to measure the interaction force between an atomically sharp probe tip and surfaces of interest, as a function of the three spatial dimensions, with picometer and piconewton accuracy. Since the results of such measurements may be affected by piezo nonlinearities, thermal and electronic drift, tip asymmetries, and elastic deformation of the tip apex, these effects need to be considered during image interpretation.

In this paper, we analyze their impact on the acquired data, compare different methods to record atomic-resolution surface force fields, and determine the approaches that suffer the least from the associated artifacts. The related discussion underscores the idea that since force fields recorded by using NC-AFM always reflect the properties of both the sample and the probe tip, efforts to reduce unwanted effects of the tip on recorded data are indispensable for the extraction of detailed information about the atomic-scale properties of the surface.

Introduction

Experimentally obtained information about atomic-scale interactions of specific surfaces with atoms, molecules, and other surfaces in their vicinity is crucial for a number of important scientific fields, including catalysis, thin-film growth, nanoscale device fabrication, and tribology, among others [1]. Shortly after the first atomic-resolution images of surfaces were obtained by noncontact atomic force microscopy (NC-AFM) [2,3], the method of dynamic force spectroscopy (DFS) was introduced, empowering experimentalists to characterize the tip–sample interaction in terms of normal forces F_n, potential energies E, and the distance z between the tip apex and the sample surface [4-7]. More recently, thanks to improvements in the design of atomic force microscopes [8,9] as well as the development of new data-acquisition strategies [10,11], DFS measurements have been extended to two and three spatial dimensions. As a result, tip–sample interaction forces and energies can be measured as a function of both the tip–sample distance z and the lateral position (x, y) of the tip apex above the sample surface. Force fields have now been recorded on NiO(001) [10,12,13], MgO/Ag(001) [14], NaCl(001) [15,16], Si(111)-(7×7) [17-19], HOPG [20,21], KBr(001) [9,22,23], Cu(111) [24], and $CaCO_3(10\bar{1}4)$ [25] surfaces, as well as single molecules of PTCDA [26,27], pentacene [28], CO [29], C_{60} [30], naphthalocyanine [31], and individual carbon nanotubes [32,33]. Moreover, differentiating the tip–sample interaction energy data in the lateral (x, y) directions has enabled the determination of atomic-scale lateral forces experienced by the probe tip [12]. From such data, the forces required to manipulate single atoms and molecules laterally on sample surfaces were quantified [34] and the lateral force field on graphite could be studied in detail [20]. Finally, three-dimensional force spectroscopy experiments performed in a liquid environment have revealed the spatial distribution of water molecules at a water–mica interface [35].

The methods most frequently reported in the literature to record two- and three-dimensional force fields above sample surfaces may be divided into two general categories (Figure 1):

1) The *curve-by-curve* method, in which individual curves of frequency shift versus tip–sample distance (Δf versus z) are recorded at a number of (x, y) locations on the sample surface and then combined to form full three-dimensional $\Delta f(x, y, z)$ arrays that are later converted to $F_n(x, y, z)$ and $E(x, y, z)$ data [12,15,16,19,22,23,25-28,30-32,36].

2) Alternatively, the three-dimensional $\Delta f(x, y, z)$ array may be recorded *layer-by-layer*, by combining a series of *topographical* or *constant-height* NC-AFM images that contain $\Delta f(x, y)$ information for certain tip–sample distances z [9,11,20,23,24]. A subset of this method involves recording the frequency shift along a single line as z is varied (*line-by-line* recording). This yields two-dimensional cuts of $\Delta f(x, z)$, which may be later converted to $F_n(x, z)$ and $E(x, z)$ maps [18,29,34].

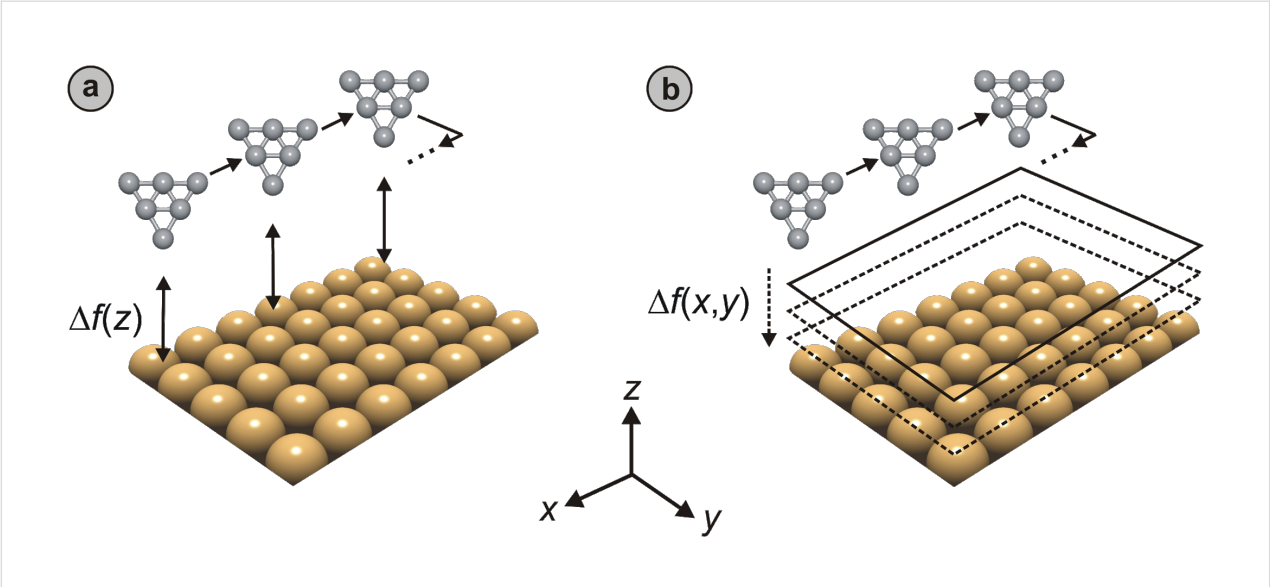

Figure 1: Schematic drawings illustrating data-acquisition procedures employed to record the atomic-scale surface force fields $F_n(x, y, z)$ experienced by a probe tip. While the *curve-by-curve* approach (a) relies on sequential recording of individual $\Delta f(z)$ curves at each (x, y) location on the surface that should be covered, the *layer-by-layer* approach (b) involves the consecutive recording of individual NC-AFM images at varying tip–sample distances z. In both cases, the resulting $\Delta f(x, y, z)$ array is converted to $F_n(x, y, z)$ data after data acquisition has been completed.

Regardless of the data-acquisition method, there are many reasons why data sets acquired on identical surfaces may vary significantly, both quantitatively and qualitatively, including:

- thermal and electronic drift during the measurement,
- nonlinearities and creep associated with piezoelectric scan elements used in the microscope,
- variability of tip-apex structure and chemistry between different experiments, and
- elastic deformations of the tip under the influence of the surface force field.

The intent of this paper is to provide a comprehensive discussion of all the major limitations intrinsic to three-dimensional force spectroscopy by scanned probes that have to be considered during data interpretation. To that end, the effect that each of the four items has on the recording of atomic-scale surface force fields is analyzed, and it will be shown that the four factors may be best alleviated by combining specialized data-recording schemes with post-acquisition correction procedures.

Results and Discussion

The goal of any microscopy technique is to obtain information on the investigated sample with as little ambiguity, uncertainty, and irreproducibility due to technique-inherent practical limitations as possible. In the case of NC-AFM, drift, piezo nonlinearities, and piezo creep result in an apparent spatial misalignment and distortion of characteristic image features compared to the true structure and location of the surface sites that induce them; elastic deformations of the probe tip can cause a lateral shift of features in data acquired at different heights; and tip asymmetry effects may further complicate the assignment of characteristic features observed in images, to actual sites on the sample surface. Finally, we need to consider that unavoidable variations in the tip-apex structure for independent measurements result in further irreproducibility. The first part of this section covers an in-depth analysis of the related issues, while the second part applies the findings to determine the optimum strategies for extracting reliable information on atomic-scale chemical and physical properties of sample surfaces.

Part I: Artifacts in force-field spectroscopy measurements

Drift

Virtually all atomic-scale scanning probe microscopy (SPM) experiments suffer from unwanted relative movement of sample and probe tip with respect to each other during imaging and force spectroscopy, as a result of thermal fluctuations and the difference in thermal expansion coefficients of the building blocks of scanning probe microscopes. Considering that the recording of dense data arrays of frequency shifts above sample surfaces in vacuum usually takes several hours [20,28], the associated imaging/spectroscopy artifacts are especially problematic for force-field measurements performed at room temperature [25], which often feature lateral drift rates of angstroms per minute. In contrast, performing the experiments at low temperatures can *suppress* thermal drift to as little as a few angstroms per day [8].

An elegant approach to *correct* the effects of thermal drift in lateral directions during SPM imaging involves the use of atom-tracking and feed-forward positioning methods. Atom tracking [37] comprises the determination of the drift vector by measuring the shift in the position of an individual maximum in subsequent SPM images followed by an appropriate correction of the tip location that compensates for this drift. In contrast, the feed-forward procedure [17] is based on the real-time correction of drift during data acquisition by applying appropriate voltages to the scan piezo, which are calculated based on the assumption that the drift vector can be adequately predicted based on prior measurements. The two approaches have been successfully implemented in the past to measure both two- and three-dimensional surface force fields at room temperature and low temperatures on various sample surfaces [16,18,19,23,25,30]. One drawback is that, typically, frequent updates of the drift vector (as much as one atom-tracking measurement before the recording of each Δf versus z curve in a *curve-by-curve* measurement [16,23]) are required due to the unpredictability of thermal drift and lack of control over temperature fluctuations. Thermal drifts leading to lateral displacements of the sample surface with respect to the probe tip, by more than one unit cell in the time required to collect an image, are also potentially problematic and in some cases limit *layer-by-layer* data acquisition to low temperatures [23].

An alternative approach to x–y drift correction involves manual post-data-acquisition shifting of images acquired by the *layer-by-layer* method [11]. In this approach, consecutive images that are part of the *layer-by-layer* dataset are laterally shifted against each other such that individual maxima in the images are aligned on top of one another. After all images in the dataset have been aligned accordingly, the (x, y) region common to all images is cut out and forms the basis for the $\Delta f(x, y, z)$ array that is later converted to interaction-force and energy data ($F_n(x, y, z)$ and $E(x, y, z)$ arrays, respectively). With a sufficiently dense dataset consisting of images separated by only a few picometers in the z direction, gradual lateral shifts between subsequent images due to thermal drift may be precisely monitored and corrected for, provided that lateral drifts between images are significantly lower than one unit cell.

Thermal drift is, however, not limited to lateral displacements, but may also affect the accuracy of z values. Fortunately, several easy-to-apply procedures can eliminate the consequences of z drift on the measurements. For arrays compiled using the *curve-by-curve* method, such drift-induced distortions can be corrected by using standard line- or plane-fit algorithms. When acquiring data *layer-by-layer*, the necessary adjustments can be carried out by comparison with site-specific high-density $\Delta f(z)$ calibration curves recorded directly before and/or after the individual layers needed to assemble the actual data array. Alternatively, curves of tunneling current versus distance can serve the same purpose if the tunneling current, which is recorded together with the frequency-shift data, does not decay too fast to provide accurate calibration at all distances covered by the 3-D set.

A completely different source of drift may originate from the use of analog electronics for oscillation detection and amplitude/phase-feedback control during NC-AFM operation. With such setups, the output voltage that is supposed to faithfully reflect the cantilever resonance frequency may shift over time even if the resonance frequency stays constant. If this happens, the data acquisition software interprets the shift as an apparent change in the tip–sample distance, which it counteracts by adjusting the z deflection of the scan piezo. As a consequence, unwanted variations in the tip–sample distance are induced [12]. While such electronic drifts may be manually corrected for by checking the atomic corrugation values at various points during force-spectroscopy experiments, the use of digital electronics for NC-AFM detection and control generally eliminates the effects of electronic drifts on measured data.

Piezo nonlinearities and piezo creep

Positioning devices that employ piezoelectric materials to realize voltage-controlled relative positioning of the tip and sample are widely used in SPM experiments (see, e.g., Figure 2) [38-40]. Despite subpicometer positioning accuracy, piezoelectric scanners display fundamental shortcomings. The most important limitation originates from the fact that the relationship between applied voltage and the amount of extension/contraction undergone by the piezoelectric material is *nonlinear* as well as *time*- and *history-dependent* [40,41]. As a result, piezo scanners extend or contract less at the beginning of a scan line than at the end, which leads to hysteresis loops [40,42,43]. Similarly, piezo *creep* manifests itself as an additional, logarithmically decaying deformation of piezo elements after the application of a change in voltage [44]. Both phenomena implicate an uncertainty regarding the assignment of the specific voltages applied to the electrodes to an actual location (x, y, z) over the sample surface, which ultimately manifests as distortion in the recorded image, as well as in a finite difference between the actual physical positions of the scanner at the same voltage, between forward and backward scan lines.

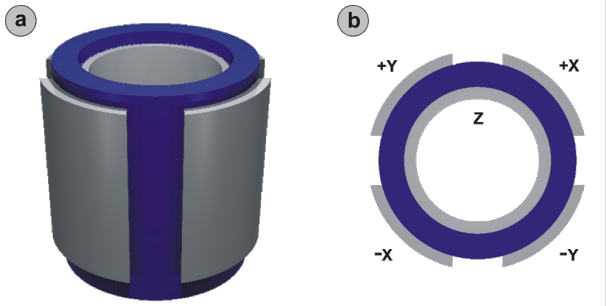

Figure 2: Isometric (a) and top (b) views of a piezoelectric scan tube employed as a precise positioning tool in scanning probe microscopes [38,39]. The tube itself (blue) is manufactured out of a piezoelectric material (usually lead zirconate titanate, PZT). Voltages applied to one of the four external metallic electrodes (grey; denoted with +X, −X, +Y, and −Y) are used to control the tube's deflection in lateral directions, whereas the potential experienced by the inner electrode (denoted with *Z*) governs the tube's vertical extension/contraction.

To correct for these effects, commercial SPM equipment often employs strategies such as closed-loop scan elements that track the actual (x, y) position with deflection sensors in real time or the application of voltages in the form of distorted waveforms so that the resulting motion is linear with respect to voltage [38,45]. Allowing the piezotube to settle down for a certain amount of time after the recording of each curve/image during data acquisition, helps to reduce the influence of creep on the measured data further. If atomic resolution is achieved, apparent lattice distortions may be corrected after acquisition has been completed by using the known size, symmetry, and orientation of surface unit cells as input [46]. Another strategy is to experiment at low temperatures, where the effects of piezo nonlinearities, creep, and hysteresis are suppressed. Combined with the benefits regarding thermal stability discussed in the previous section, low-temperature data recording is found to be ideally suited for the reliable long-term recording of atomic-scale surface force fields [8].

For completeness, it should be noted that the limitations addressed above do not represent a complete list. Additional distortions of atomic-scale force fields may be caused, e.g., by a cross-coupling between the x, y, and z channels of the scanner due to either structural imperfections of the piezo elements or design-inherent coupling issues, or both. For the tube scanner in Figure 2, such structural imperfections may be due to small variations in the thickness of the tube walls, or due to restrictions in the ability of the tube to flex, which are imposed by the soldering or glue spots that are used to contact the electrodes;

an example of a design-inherent issue is the tube sweeping out an arc when moved in lateral directions, which makes some distortion in z unavoidable. Another source of distortions may be due to small deviations of the scanner axes from the actual x, y, and z directions caused by incorrect alignment during microscope construction, where even differences of a few degrees may result in appreciable lateral shifts between tip and sample as the piezo material is deformed.

Tip-apex structure and chemical identity

Numerous theoretical and experimental studies have shown that the atomic-scale contrast in NC-AFM measurements is heavily dependent on the local structure of tip apices employed in the experiments, as well as on the chemical identity of the apex atoms (see, e.g., [28,47-55]). While, for specific cases, the contrast-formation mechanism may be explained by using a relatively simple picture of tip-apex polarity [47,54,55], generally more complicated tip-apex models and theoretical considerations need to be taken into account to understand the full effect of tip structure and chemistry on NC-AFM measurements [49-51]. Controlling the chemical identity of the probe tip employed in NC-AFM experiments down to the last few atoms of the tip apex has proven to be extremely difficult in the past due to oxidation issues associated with traditional Si cantilevers, as well as with metallic tips prepared by electrochemical etching, in addition to frequently observed tip changes that may lead to a modification of the tip apex on the atomic scale [56]. A notable exception presents itself in the form of metallic tip apices terminated by single molecules that have been deliberately picked up during SPM experimentation at low temperatures [28]. This approach, which has been previously employed in scanning tunneling microscopy imaging [57,58], has recently been applied to NC-AFM imaging and force spectroscopy experiments with great success [28,31]. As surface force fields recorded with such well-defined tips provide useful information about the interaction of the attached molecule with the probed surfaces, the application of molecule-terminated tips is expected to become more and more popular in force-spectroscopy experiments [1]. Lastly, let us note that a recently reported alternative method to control tip-apex chemistry involves the in situ deposition of metallic thin layers on commercial Si cantilever apices [56]. In combination with theoretical calculations, maxima in NC-AFM images provided by such well-characterized tips on the ionic surface of NaCl(001) have been shown to unambiguously coincide with surface anions, facilitating atomic-scale chemical identification.

Tip elasticity

In addition to drift and piezo effects, the accuracy with which a numerical value, obtained through two- or three-dimensional force field spectroscopy, can be straightforwardly assigned from its apparent position (x, y, z) in the data array to an actual location relative to the sample lattice, is further limited by elastic deformations of the tip apex under the influence of external forces as it is scanned over the sample surface. This is because these deformations cause the tip apex to be at a different location than we assume it to be, which results in a distortion of the recorded force field. The extent of this distortion depends on the local strength of the tip–sample interaction force as well as on the lateral and vertical stiffness of the specific tip.

The effect of tip-apex deformations on dense force-field spectroscopy experiments has been previously analyzed in the literature [9,16,23]. In particular, Such et al. listed several criteria that allow the identification of tip-apex relaxation effects in NC-AFM-based force spectroscopy experiments [9]: (a) Lines showing rapid changes in two-dimensional, horizontal topography/frequency shift maps of the surface; (b) significant shifts in contrast patterns observed in such maps as the tip moves closer to the surface; (c) force-versus-distance curves exhibiting extended plateaus, over several hundreds of picometers, of more or less constant force close to the surface, as opposed to the expected onset of repulsive force; and (d) two-dimensional vertical force cuts where the force on several lattice sites becomes maximum at a certain height above the surface and stays constant until the plane of closest approach is reached, which is a direct consequence of (c). Additionally, Kawai et al. [16] and Fremy et al. [23] performed drift/creep-corrected three-dimensional force-field-spectroscopy experiments using atom-tracking, on NaCl(001) and KBr(001), respectively, in which shifts of characteristic maxima in atomic-scale images and significant distortions of the observed contrast patterns attributed to tip apex elasticity were observed as a function of tip–sample distance.

On our path towards finding strategies that reduce the impact of tip deformation on the recorded data, we start by noting that atom-tracking and feed-forward techniques, which have been successful in correcting for the effects of thermal drift and, at least partially, piezo effects, do not offer viable solutions. Next, we recognize that caution has to be exercised when analyzing data sets where relaxations lead to significant distortions in contrast patterns, because the associated data may be so heavily influenced by the properties of the probing tip that little useful information can be gained about the sample surface. To avoid misinterpretations, such measurements should be discarded. For the analysis of data acquired with tips that display only minor gradual contrast changes with distance, lateral tip apex bending may arise from two sources: (i) lateral forces inducing atomic-scale relaxations in the last few atomic layers of the tip apex, based on the local position of the tip above the surface, and (ii) normal (i.e., vertical) forces that will cause a bending and an

effective overall lateral movement of the tip apex in a particular direction, which occurs if the tip used to probe the sample surface is asymmetric (Figure 3). Deformations due to (i) could be accounted for if the lateral stiffness of the apex were known, which it is usually not. However, if measurements are restricted to distances at which low site-specific lateral forces manifest, rough estimations of the expected deformations using typical values [59,60] suggest that they may be small enough to be ignored for all practical purposes. Deformations due to (ii), on the other hand, may be largely compensated by employing the post-data-acquisition correction procedures described earlier for the layer-by-layer approach [11].

Tip asymmetry

Even though imaging artifacts observed in atomic-scale scanning probe experiments are often associated with the use of asymmetric tips [16,61], a comprehensive understanding of the link between asymmetric tip geometries and the imaging artifacts they cause is still not complete. In this part of the paper, we present the highlights from a systematic study of the fundamental effects that asymmetric tips have on the measurement of atomic-scale surface force fields. The corresponding simulations, which use Matlab-based code [62], feature basic model geometries for tip and sample consisting of rigid atoms that interact through analytical potentials (both Lennard-Jones (L-J) and ionic). Even though these assumptions represent an oversimplification, as tip–sample contacts will relax upon tip approach and short-range interactions may differ substantially from those predicted by the potentials employed here, we still expect such simulations to provide valuable insights into the general trends that describe how tip asymmetry manifests in 3-D data sets.

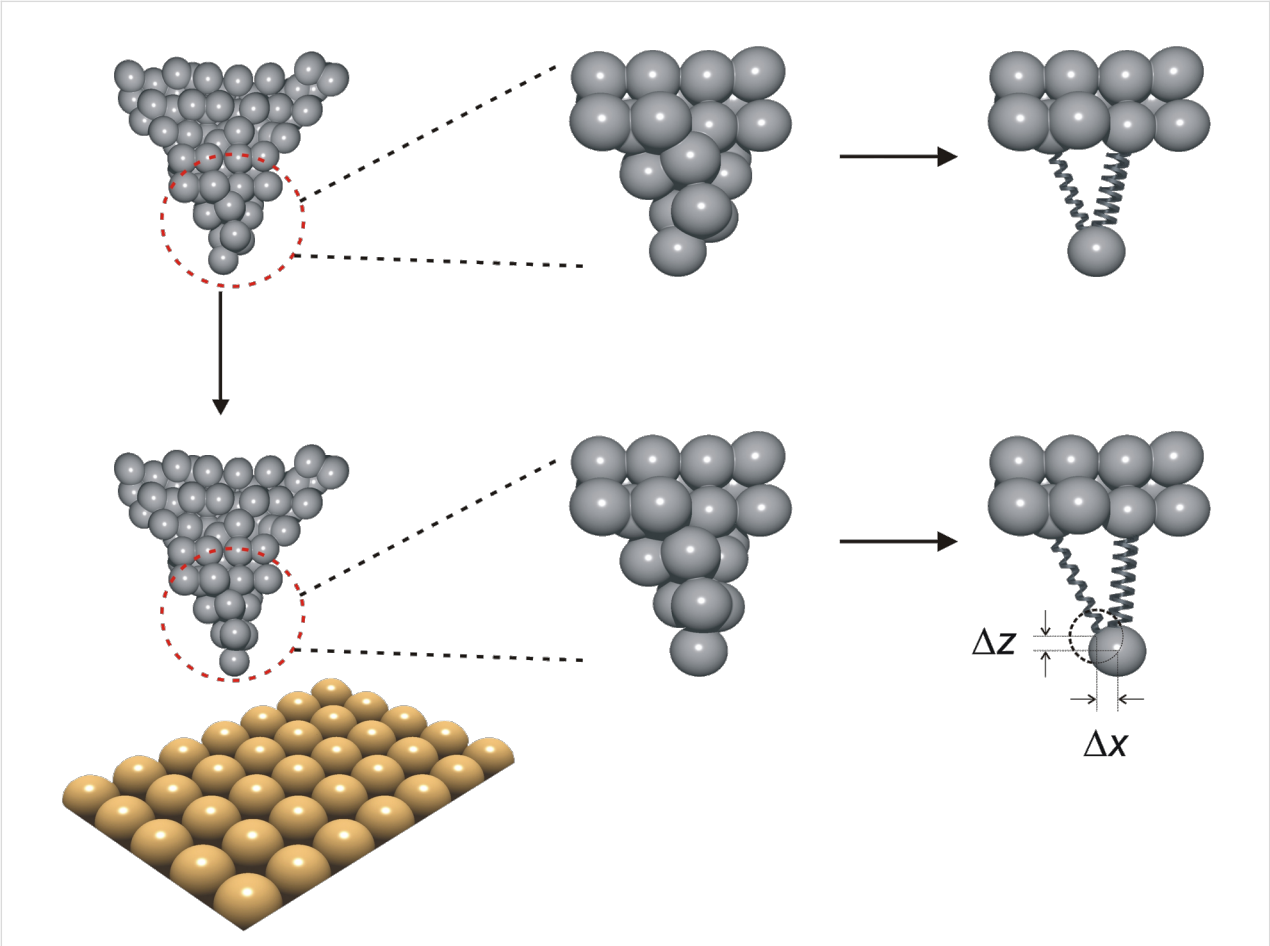

Figure 3: Schematic drawing describing the elastic deformation/bending of an asymmetric tip apex as the tip–sample distance is reduced. Top row, left: an asymmetric tip located away from any surface; middle: magnification of the tip apex; and right: a mechanical spring model that mirrors the elastic properties of the last tip atom. Note that due to the asymmetric character of the tip apex, the spring on the right is stiffer than the spring on the left. Bottom row, left: The tip is now located close to a surface, where it feels an attractive interaction in the *z* direction; middle: zoom onto the tip end; and right: equivalent mechanical model. The surface forces pull the last tip atom a distance Δ*z* towards the surface when compared to the location of the atom relative to the tip base in the top row (dashed circle). Since the right spring in the equivalent mechanical model is stiffer than the spring on the left, this motion also results in a net lateral displacement Δ*x*.

For the computations, the Lennard-Jones potential between two atoms $V_{L\text{-}J}$ was calculated by using

$$V_{L-J} = 4\varepsilon\left[(\sigma/r)^{12} - (\sigma/r)^{6}\right]. \quad (1)$$

Here r denotes the distance between the centers of the two atoms, ε the depth of the potential well, and σ the finite distance at which the potential vanishes (note that $\sigma = d/1.12$, where d is the hard-sphere diameter of the atom). To obtain appropriate σ and ε values, the Lorentz–Berthelot mixing rules ($\sigma_{12} = (\sigma_1 + \sigma_2)/2$ and $\varepsilon_{12} = (\varepsilon_1 \times \varepsilon_2)^{0.5}$ [63]) were employed. In cases where the interacting atoms were ionized, a Coulomb potential V_C

$$V_C = k_e q_1 q_2 / r, \quad (2)$$

was added to the Lennard-Jones interaction with k_e reflecting the Coulomb constant, q_1 and q_2 the ionic charges, and r the distance between the ions. Total interaction potentials are obtained by summing up the individual potentials between each tip and substrate atom. The normal force F_n is then calculated by taking the derivative of the total interaction potential in the vertical direction.

The model tip apices used in the study are constructed from six close-packed atoms arranged in a three-layer planar configuration with the structural characteristics of a single Pt(111) plane (Figure 4a). The desired asymmetry in the tip-apex structure is obtained by rotating the model structures by an angle θ around the front-most atom (Figure 4b and Figure 4c). Even though we focus solely on this particular type of tip for the present discussion, we have calculated all cases for closed-packed planar tips featuring anything between a single atom and up to 15 atoms (five-layer 2-D tips), as well as for a full set of 3-D closed-packed tips with up to 25 atoms (five-layer 3-D tips). As a general rule, all trends displayed in the results below are more emphasized the more atoms are included in the tip, in particular for ionic interactions; but comparison also confirmed that the observed effects are representative for the overall behavior of asymmetric tips within the limited range of validity of this simple conceptual approach.

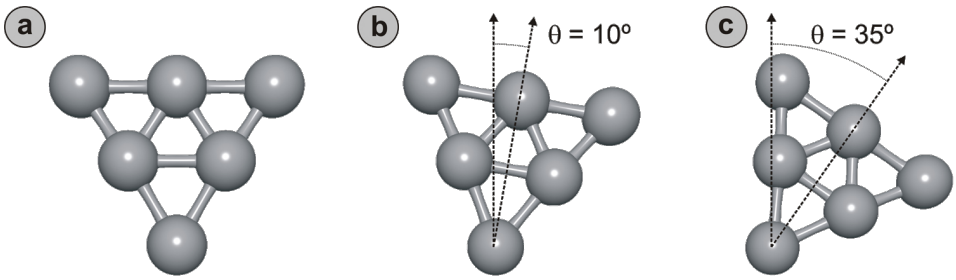

Figure 4: Planar three-layer tip apex models used in the analytical simulations, featuring close-packed atoms arranged with 2.70 Å nearest-neighbor distance (2.20 Å for ions with +1*e* charge). ε is 694 meV (680 meV if ionized) [64]. (a) Symmetric tip apex. (b) and (c) successively more asymmetric tip apices are obtained by rotating the tip model of (a) around the front-most atom with increasing angles θ.

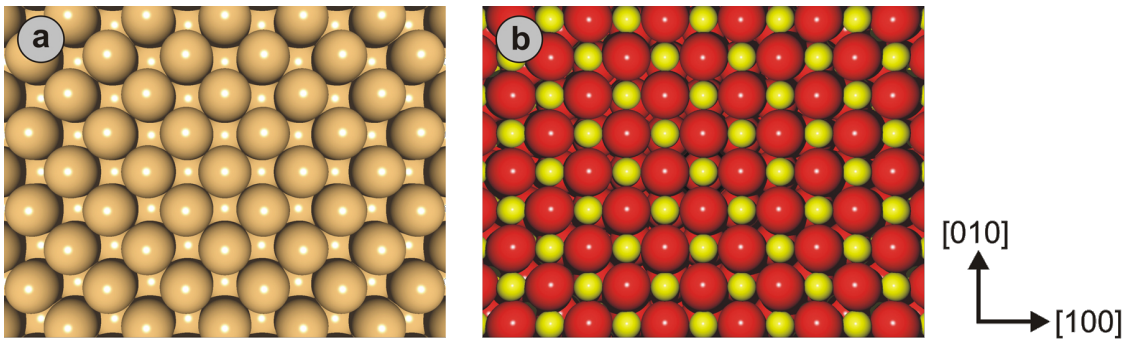

Figure 5: Illustrations of the model surfaces investigated in this section. (a) The Cu(001) surface in FCC configuration. Beige spheres represent individual copper atoms with a diameter of d_{Cu} = 2.60 Å and ε_{Cu} = 415 meV [64]. (b) The ionic NaCl(001) surface in which the large, red spheres are Cl^- ions with d_{Cl-} = 3.34 Å and ε_{Cl-} = 30.0 meV, and the small, yellow spheres are Na^+ ions with d_{Na+} = 2.32 Å and ε_{Na+} = 15.4 meV. Crystallographic directions are indicated for both surfaces on the right.

Two different surfaces were investigated as part of the simulations described here (Figure 5): a surface that features all the structural characteristics of Cu(001), but interacts purely through L-J forces with the tip (i.e., we do not reproduce "true" metallic interactions, to keep calculations simple); and a surface with the structure of NaCl(001), where ionized atoms feel a Coulomb potential in addition to the ubiquitous L-J contribution. As is the case for the tips, substrate atoms were assumed to be immobile hard spheres, i.e., material relaxation effects are excluded. Surface cells were chosen to comprise 33 × 33 atoms with a thickness of five layers, as this cell size could be calculated quickly but was found to be large enough to avoid boundary effects.

To start, we focused on the Cu(001) surface using two-dimensional vertical cuts representing normal L-J forces along the [100] direction. When a symmetric tip is used (i.e., θ = 0°), it is observed that the individual force fields associated with surface atoms are *symmetric*, i.e., evolve in a *straight* configuration while moving away from the sample surface. Thereby, the force maxima are situated directly above the surface copper atoms (referred to as *atomic sites* in the following) for most tip–sample distances (Figure 6a). At very close separations, however, the atomically sharp tip apex employed in the simulations experiences a larger attractive force on the site of the minima of the surface potential (the *hollow sites*). This force contrast flip causes a crossing of the $\Delta f(z)$ curves recorded above the atomic and hollow sites similar to the one previously observed in simulations carried out for a Xe(111) surface [65], which in turn limits the closest distance at which constant-frequency-shift images can be recorded. Therefore, high-resolution 3-D force field measurements using the *layer-by-layer* approach with active feedback are likely restricted to tip–sample distances that are unaffected by contrast flips.

This situation changes if *asymmetric* tips are employed. For θ values of 45° and 58° (Figures 6b and Figure 6c, respectively), the simulated atomic-scale force fields become lopsided, especially in close proximity to the sample surface, where an increasing lateral offset between the locations of perceived force maxima and the (x, y) positions of the surface copper atoms develops with decreasing tip–sample distance. Similarly shaped force spectra have previously been reported for two-dimensional force-section measurements on the surface of graphite [21]. While an elastic deformation of the tip apex in the lateral directions during the measurements could be responsible for such effects, our simulations suggest that strongly asymmetric tips provide an alternate explanation for the observed patterns. We note that for θ = 58°, one side of the tip is almost parallel to the surface, which thus may represent an illustrative but extreme case for tip asymmetry under practical conditions.

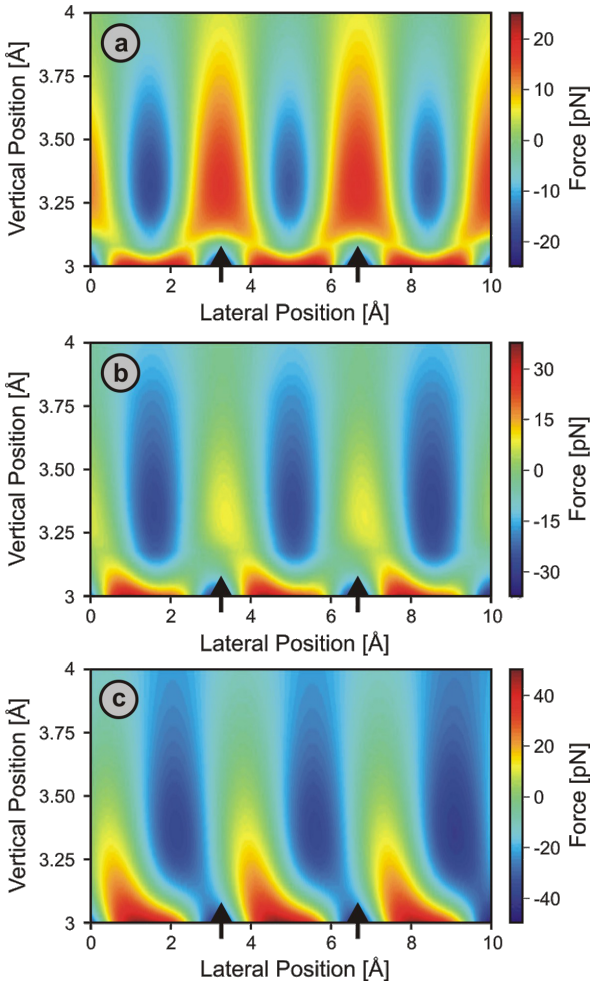

Figure 6: Simulated 2-D force fields over a Cu(001) surface along the [100] direction probed by symmetric (a: θ = 0°) as well as asymmetric (b: θ = 45°; c: θ = 58°) model tip apices that feature geometries as illustrated in Figure 4 while being assumed to interact solely through Lennard-Jones forces with the sample. The difference between the interaction force at each point and the mean interaction force at that height is displayed for better contrast visibility. The vertical position axis indicates the distance in the z direction between the center of the front-most tip atom and the x–y plane defined by the centers of the surface copper atoms. Black arrows mark the lateral positions of the surface Cu atoms.

The fact that the location of the maximum attractive force smoothly moves from the atomic to the hollow site upon decreasing the tip–sample distance is of particular concern when applying the *layer-by-layer* approach for data acquisition in combination with post-data-acquisition correction procedures. As the symmetry of the surface unit cells of atomic and hollow sites of the Cu(001) sample is identical, the gradual shifting of the force maxima from one to the other becomes indistinguishable from the effects of thermal drift or overall elastic bending of the probe tip. Considering that most tips employed in NC-AFM measurements are asymmetric to a

certain extent, this experimental approach should thus be avoided for surfaces with such characteristics, or, at the very least, limited to regions sufficiently away from the sample surface. However, even though such complications can in principle be avoided by applying atom-tracking or feed-forward techniques, we note that their inherent inability to correct for elastic deformations of the tip apex jeopardizes the reliability of such real-time approaches to drift correction to the extent that it is unclear a priori which approach to drift correction, i.e., real-time or post-acquisition, is superior.

The effects of tip asymmetry on simulated force spectra acquired on ionic surfaces are inherently different from the simulation results obtained for Cu(001), as we will see below with the case of the NaCl(001) surface. Towards this end, we first outfit our planar six-atom tip as displayed in Figure 4 with a trapped positive, unity charge localized at the center of the front-most atom, which reflects a reasonable way of representing a charged tip [48]. The findings are then compared to the case where each of the atoms features a single trapped unity charge. Though this situation would be unrealistic for a real platinum tip, it allows us to get a feel for the trends that charge distributions in tips may impose on image contrasts. Figure 7 shows the results obtained along the [110] direction, which includes Cl^- ions only, while results along the [100] direction, which features both Cl^- and Na^+ ions, are displayed in Figure 8.

We find that the locations of the force maxima coincide with the actual positions of the chlorine surface atoms in all cases except for Figure 7d, in which we observe an absolute lateral shift. But even in this instance, the individual force-field spectra associated with surface atoms evolve in a straight configuration, as the type of lattice site responsible for maximum forces (the Cl^- ions) does not change throughout the simulated height regime, in contrast to the simulations with L-J force only, performed on Cu(001). The reason for this difference lies in the distance dependence of the underlying atomic potentials. Since the ionic potential dominant in simulations on NaCl(001) has a long-range character as opposed to the L-J potential, the force field spectra are not as sensitive to the local changes in tip-apex structure induced by tip asymmetry for the majority of simulated heights. Perhaps more interestingly, the effect of asymmetric tip apices on simulated force spectra is nearly absent in the [100] direction for all investigated tips (Figure 8). The underlying reason is that the [100] direction on the sample

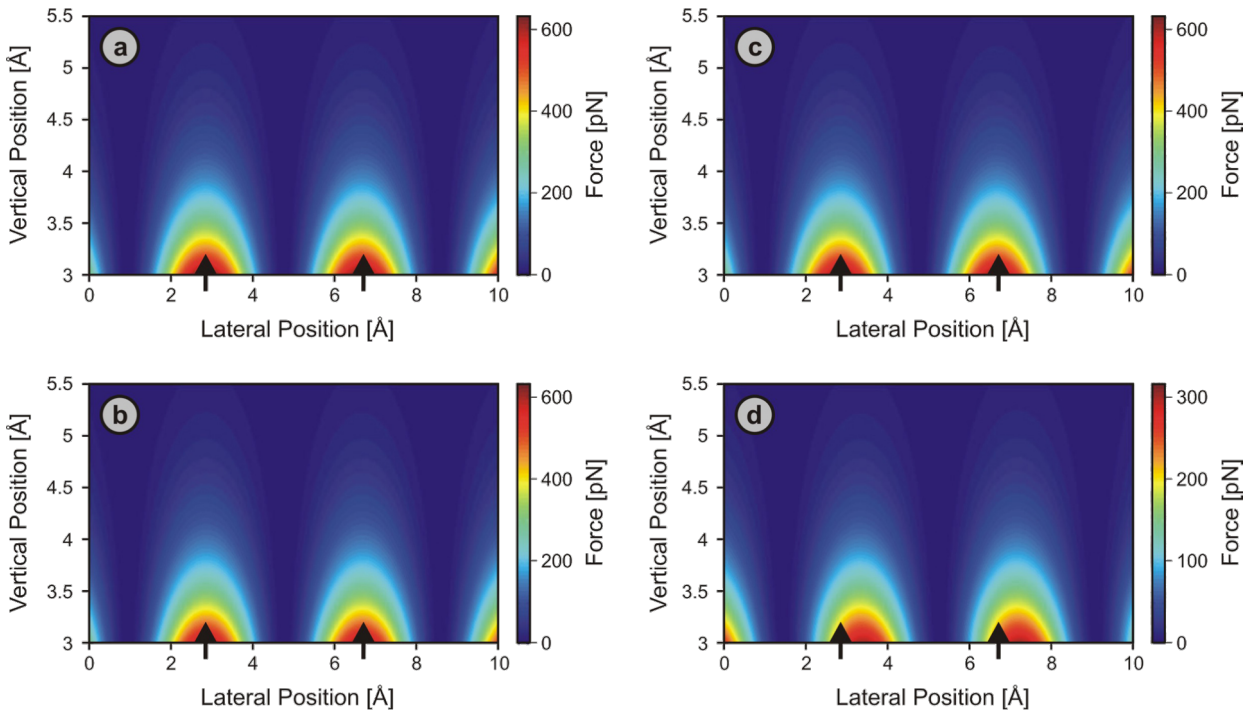

Figure 7: Simulated 2-D force fields over the NaCl(001) surface along the [110] direction probed by a symmetric (θ = 0°; a, c) and an asymmetric (θ = 56°; b, d) tip apex. In (a) and (b), only the front-most tip atom carried a trapped positive charge, while in (c) and (d) all atoms had been artificially ionized with +1 to explore the effect of charge distributions in asymmetric tips. In all panels, the difference between the absolute interaction force at each point and the minimum calculated force is displayed for better contrast. Black arrows mark the positions of the chlorine surface ions, which coincide with the force maxima in all cases but (d), in which the force maxima and the atomic positions are shifted by about 0.7 Å.

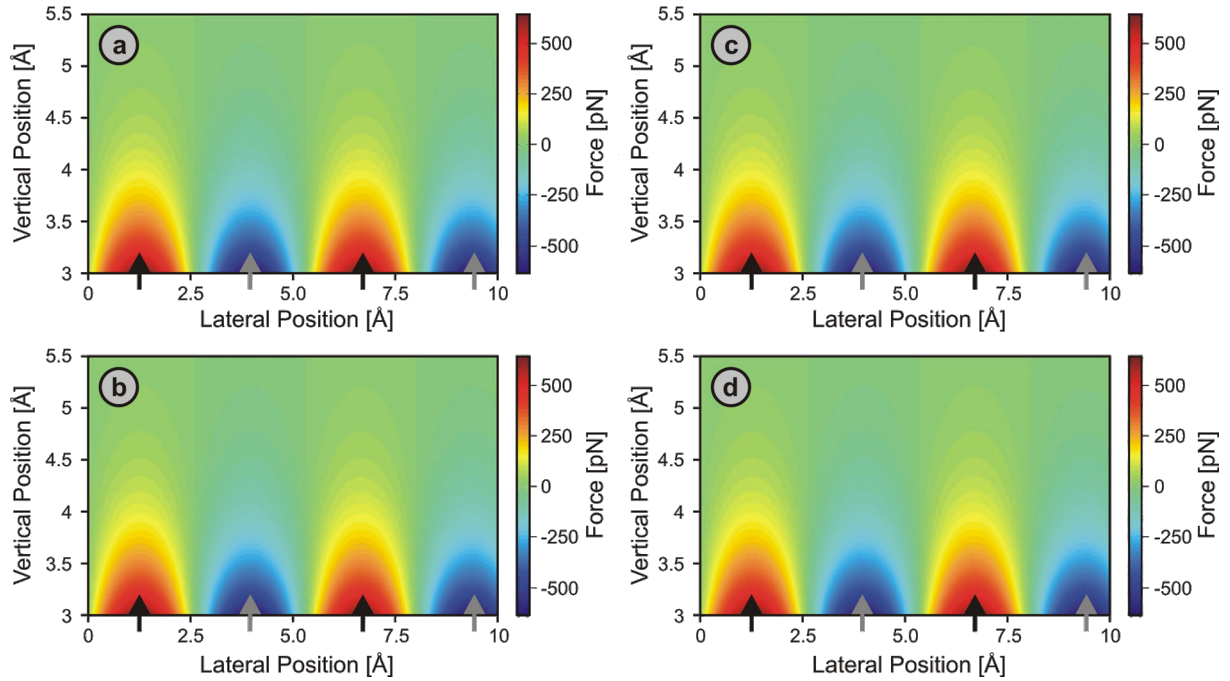

Figure 8: 2-D force fields over the NaCl(001) surface simulated using parameters identical to the corresponding panels in Figure 7, but this time taken along the [100] direction. The difference between the absolute interaction force at each point and the mean calculated force is displayed for better contrast. In all cases, force maxima (red) coincide with the positions of the chlorine ions of the surface (indicated by the black arrows), while force minima (blue) concur with the sodium-ion lattice sites.

surface includes both Cl^- and Na^+ ions, and the additional attractive forces experienced by a *right-leaning* tip due to Coulombic interactions with the Cl^- ions are cancelled out by an equal increase in the amount of repulsive interactions with the Na^+ ions of the surface, indicating that asymmetry effects in force spectra over ionic surfaces are strongly direction-dependent.

To summarize this section, we have found that:

1. For long-range interactions, force fields evolve straight into the space above the surface, even with asymmetric tips, because there is not much change in the relative contribution of individual tip atoms to the total tip–sample interaction with distance. Therefore, force spectroscopy experiments may be reliably performed and interpreted at all distance regimes, and post-data-acquisition correction procedures in conjunction with *layer-by-layer* data acquisition may be employed. With the example of ionic surfaces and Coulomb-force-dominated tip–sample interactions, it was found that the actual lattice sites responsible for the force maxima in the attractive tip–sample interaction regime (i.e., Cl^- ions in the case of NaCl(001) and a positively charged tip) coincide with the perceived force-maxima locations, except for in the case of very asymmetric tips featuring charge distributions (multiple localized charges of the same sign) at the apex, in which a lateral shift between force maxima and the locations of the surface ions occurs.
2. In contrast, when interactions between individual tip and sample atoms are sufficiently short-ranged, the lattice site exhibiting the most attraction on the front-most tip atom may move from an atomic to a hollow site upon tip approach. In such a case, tip asymmetry may lead to distortions in force fields, with the force maximum experienced by the tip as a whole smoothly moving from being near the atomic site towards being near the hollow site with decreasing tip–sample distance, as illustrated in Figure 6. Considering that most tips employed in NC-AFM measurements are asymmetric, 3-D force-mapping experiments performed on such samples may be instructive only for large enough tip–sample distances. If smaller distances are included in the analysis, post-data-acquisition correction procedures should be generally avoided if the effects of tip asymmetry cannot readily be distinguished from thermal drift or lateral shifting due to overall elastic bending of the probe tip.

3. For both short-range and long-range interactions, force-field characteristics associated with individual surface atoms on defect-free surfaces exhibit a *straight* and *symmetric* nature when probed with *symmetric tips* consisting of *immobile, hard, sphere-like atoms*. Thus, any experimentally observed deviations from this straight, symmetric character are necessarily due to either tip asymmetry or the elastic properties of the tip, or a combination of both.

Part II: Comparison of data-acquisition procedures for atomic-scale force field spectroscopy

To help facilitate a reliable interpretation of 3-D force field data, we will summarize in this section the key points that have to be considered for selected experimental approaches. Thereby, we will focus on the following three methods, as they have been the main methods reported in the literature so far:

- *Layer-by-layer* data acquisition at low temperatures with post-data-acquisition correction procedures, referred to in the following as "method A" [11,20];
- *Curve-by-curve* data acquisition, mostly at room temperature, with atom-tracking/feed-forward procedures ("method B") [16,19,23,25,30];
- *Curve-by-curve* or *line-by-line* data acquisition at low temperatures, involving the use of a *reference image* for drift correction at set time intervals ("method C") [28,31].

Table 1 summarizes to what extent the methods listed above address the specific experimental issues discussed earlier. Note, however, that other combinations of experimental and data analysis procedures are possible, and experimentalists can tailor the exact approach to represent the best combination of experimental capabilities, post-acquisition processing, and artifact avoidance.

From Table 1, we see that all three methods are able to satisfactorily account for the effects of thermal drift and piezo nonlinearities/creep. The variability of the tip-apex structure and chemistry between different experiments remains an inherent problem associated with NC-AFM, and efforts to obtain well-defined tips (such as the deliberate adsorption of a CO molecule on the tip apex; see [28]) are expected to be utilized in an increasing number of experiments, regardless of the specific methods used to perform the force spectroscopy. The effect of tip asymmetry on recorded force spectra is also an inherent tip-related problem that is especially critical for surfaces where the tip–sample interactions responsible for atomic resolution are predominantly short-range. Even though none of the methods above can eliminate the effect of tip asymmetry on the recorded force spectra, the application of post-data-acquisition shifting of subsequent images in method A could lead to misleading results on surfaces where the force maxima shift from one lattice site to another with indistinguishable symmetries. However, as long as tip–sample interactions are predominantly electrostatic in nature (e.g., on ionic crystals such as NaCl and KBr, as well as most metal oxides), method A holds a notable advantage over the

Table 1: Overall comparison of methods for atomic-scale force field spectroscopy. Items marked with ✓ are satisfactorily resolved, while entries labeled with ✗ remain to be addressed.

	method A	method B	method C
thermal drift	✓ low thermal drift corrected post-acquisition	✓ data acquisition has to be interrupted frequently to correct drift	✓ data acquisition has to be interrupted to correct drift
piezo nonlinearities and creep (*x*–*y* plane)	✓ reduced piezo effects can be corrected post-acquisition	✓ remaining piezo effects can be corrected post-acquisition	✓ reduced piezo effects can be corrected post-acquisition
variability of tip structure/chemistry	✗ tip changes easily detectable during data acquisition	✗ tip changes may not be visible during data acquisition	✗ tip changes may not be visible during data acquisition
tip elasticity	~ ✓ minor tip elasticity effects can be eliminated post-acquisition	✗	✗
tip asymmetry	✗ may cause problems for post-acquisition drift correction on certain surfaces	✗	✗
additional notes	contrast distortions readily detectable during data acquisition	irreversible tip changes much more likely at room temperature	–

other methods, as it allows the correction of lateral shifts in force maxima due to overall elastic deformations/bending of an asymmetric tip apex with increasing external normal forces at decreasing tip–sample distance.

While the post-data-acquisition correction procedures employed in method A do not allow researchers to distinguish between the exact source of such lateral shifts (which may be either caused by drift, creep, or elastic bending of asymmetric tips due to normal forces) [23], they nevertheless correct for such effects to a certain extent. Post-data-acquisition drift correction is best suited for low-temperature experiments, since (i) the drift rates are low enough that lateral shifts between consecutive images are small (significantly less than the lattice constant of the sample surface in question); (ii) thermal drift is often not random, but gradual, as the microscope temperature asymptotically approaches the equilibrium temperature of the cryogen [11]; and (iii) gradual changes in contrast (such as the appearance of stripe-like features between force maxima [20]) are clearly observable using sufficiently dense datasets. On the other hand, it is unsuitable for datasets in which contrast patterns alter so much with changing tip–sample distance that the alignment of characteristic maxima in subsequent images becomes problematic. However, no current method allows one to account for the effects of local, site-specific lateral forces causing such atomic-scale elastic relaxations in the tip apex. Lastly, it should be pointed out that one aspect where *layer-by-layer* acquisition of frequency-shift data proves especially advantageous, as opposed to *curve-by-curve* recording, is that distance-dependent distortions in the measured tip–sample interaction due to tip asymmetry and/or elastic relaxations manifest clearly during data acquisition as contrast changes, whereas in the *curve-by-curve* data-acquisition strategy such effects generally become only observable after data processing.

Conclusion

Various methods and procedures employed to measure atomic-resolution surface force fields by NC-AFM have been reported and compared with respect to the extent to which they address issues such as drift and piezo nonlinearities, as well as tip-related problems of asymmetry and elasticity. While drift and piezo creep may be addressed in a number of ways, including the use of atom-tracking and feed-forward methodologies, a combination of *layer-by-layer* data acquisition with post-data-acquisition correction procedures allows the additional correction of minor gradual lateral shifts due to an overall elastic bending of the tip apex, in the case of datasets in which the type of contrast remains largely undistorted. Simulations based on simplified pairwise potentials acting between model nonionic surfaces suggest that tip asymmetries may lead to appreciable distortions of atomic-scale force spectra that are absent when the surface force field is probed with symmetric tip apices. As such, it can be argued that the distortions in force-field maps that have been reported in the literature are due to irregularities associated with the probing tip apex, such as inherent asymmetry and/or extensive elastic response due to interaction forces, and are thus unlikely to represent intrinsic properties of the investigated sample surfaces. Since the general aim of the measurement of atomic-scale force fields is to study the physical and chemical properties of the sample surfaces in question rather than to probe artifacts, datasets exhibiting unusual changes in contrast patterns should ideally be disregarded. The use of post-data-acquisition correction procedures in suitable datasets, on the other hand, proves beneficial when addressing problems associated with drift and piezo creep, as well as the gradual lateral bending of asymmetric tip apices with decreasing tip–sample distance during the measurement of atomic-scale surface force fields.

Acknowledgements

U. D. S. thanks Alexander Schwarz for helpful discussions on the subject of tip elasticity. This work was financially supported by the US Department of Energy (Basic Energy Sciences grant No. DE-FG02-06ER15834) and the National Science Foundation through the Yale Materials Research Science and Engineering Center (grant Nos. MRSEC DMR-1119826 and DMR-0520495) and the Materials World Network program (grant No. MWN DMR-0806893).

References

1. Baykara, M. Z.; Schwendemann, T. C.; Altman, E. I.; Schwarz, U. D. *Adv. Mater.* **2010,** *22,* 2838–2853. doi:10.1002/adma.200903909
2. Sugawara, Y.; Ohta, M.; Ueyama, H.; Morita, S. *Science* **1995,** *270,* 1646–1648. doi:10.1126/science.270.5242.1646
3. Giessibl, F. J. *Science* **1995,** *267,* 68–71. doi:10.1126/science.267.5194.68
4. Hölscher, H.; Schwarz, A.; Allers, W.; Schwarz, U. D.; Wiesendanger, R. *Phys. Rev. B* **2000,** *61,* 12678–12681. doi:10.1103/PhysRevB.61.12678
5. Lantz, M. A.; Hug, H. J.; Hoffmann, R.; van Schendel, P. J. A.; Kappenberger, P.; Martin, S.; Baratoff, A.; Güntherodt, H.-J. *Science* **2001,** *291,* 2580–2583. doi:10.1126/science.1057824
6. Hoffmann, R.; Kantorovich, L. N.; Baratoff, A.; Hug, H. J.; Güntherodt, H.-J. *Phys. Rev. Lett.* **2004,** *92,* 146103. doi:10.1103/PhysRevLett.92.146103
7. Sugimoto, Y.; Pou, P.; Abe, M.; Jelinek, P.; Pérez, R.; Morita, S.; Custance, Ó. *Nature* **2007,** *446,* 64–67. doi:10.1038/nature05530
8. Albers, B. J.; Liebmann, M.; Schwendemann, T. C.; Baykara, M. Z.; Heyde, M.; Salmeron, M.; Altman, E. I.; Schwarz, U. D. *Rev. Sci. Instrum.* **2008,** *79,* 033704. doi:10.1063/1.2842631
9. Such, B.; Glatzel, T.; Kawai, S.; Koch, S.; Meyer, E. *J. Vac. Sci. Technol. B* **2010,** *28,* C4B1. doi:10.1116/1.3382230
10. Hölscher, H.; Langkat, S. M.; Schwarz, A.; Wiesendanger, R. *Appl. Phys. Lett.* **2002,** *81,* 4428–4430. doi:10.1063/1.1525056

11. Albers, B. J.; Schwendemann, T. C.; Baykara, M. Z.; Pilet, N.; Liebmann, M.; Altman, E. I.; Schwarz, U. D. *Nanotechnology* **2009,** *20,* 264002. doi:10.1088/0957-4484/20/26/264002
12. Schwarz, A.; Hölscher, H.; Langkat, S. M.; Wiesendanger, R. *AIP Conf. Proc.* **2003,** *696,* 68–78. doi:10.1063/1.1639679
13. Langkat, S. M.; Hölscher, H.; Schwarz, A.; Wiesendanger, R. *Surf. Sci.* **2003,** *527,* 12–20. doi:10.1016/S0039-6028(03)00076-1
14. Heyde, M.; Simon, G. H.; Rust, H.-P.; Freund, H.-J. *Appl. Phys. Lett.* **2006,** *89,* 263107. doi:10.1063/1.2424432
15. Schirmeisen, A.; Weiner, D.; Fuchs, H. *Phys. Rev. Lett.* **2006,** *97,* 136101. doi:10.1103/PhysRevLett.97.136101
16. Kawai, S.; Glatzel, T.; Koch, S.; Baratoff, A.; Meyer, E. *Phys. Rev. B* **2011,** *83,* 035421. doi:10.1103/PhysRevB.83.035421
17. Abe, M.; Sugimoto, Y.; Namikawa, T.; Morita, K.; Oyabu, N.; Morita, S. *Appl. Phys. Lett.* **2007,** *90,* 203103. doi:10.1063/1.2739410
18. Sugimoto, Y.; Namikawa, T.; Miki, K.; Abe, M.; Morita, S. *Phys. Rev. B* **2008,** *77,* 195424. doi:10.1103/PhysRevB.77.195424
19. Sugimoto, Y.; Ueda, K.; Abe, M.; Morita, S. *J. Phys.: Condens. Matter* **2012,** *24,* 084008. doi:10.1088/0953-8984/24/8/084008
20. Albers, B. J.; Schwendemann, T. C.; Baykara, M. Z.; Pilet, N.; Liebmann, M.; Altman, E. I.; Schwarz, U. D. *Nat. Nanotechnol.* **2009,** *4,* 307–310. doi:10.1038/nnano.2009.57
21. Kawai, S.; Glatzel, T.; Koch, S.; Such, B.; Baratoff, A.; Meyer, E. *Phys. Rev. B* **2010,** *81,* 085420. doi:10.1103/PhysRevB.81.085420
22. Ruschmeier, K.; Schirmeisen, A.; Hoffmann, R. *Phys. Rev. Lett.* **2008,** *101,* 156102. doi:10.1103/PhysRevLett.101.156102
23. Fremy, S.; Kawai, S.; Pawlak, R.; Glatzel, T.; Baratoff, A.; Meyer, E. *Nanotechnology* **2012,** *23,* 055401. doi:10.1088/0957-4484/23/5/055401
24. Such, B.; Glatzel, T.; Kawai, S.; Meyer, E.; Turansky, R.; Brndiar, J.; Štich, I. *Nanotechnology* **2012,** *23,* 045705. doi:10.1088/0957-4484/23/4/045705
25. Rahe, P.; Schütte, J.; Schniederberend, W.; Reichling, M.; Abe, M.; Sugimoto, Y.; Kühnle, A. *Rev. Sci. Instrum.* **2011,** *82,* 063704. doi:10.1063/1.3600453
26. Braun, D.-A.; Weiner, D.; Such, B.; Fuchs, H.; Schirmeisen, A. *Nanotechnology* **2009,** *20,* 264004. doi:10.1088/0957-4484/20/26/264004
27. Braun, D.-A.; Langewisch, G.; Fuchs, H.; Schirmeisen, A. *J. Vac. Sci. Technol. B* **2010,** *28,* C4B6. doi:10.1116/1.3322735
28. Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. *Science* **2009,** *325,* 1110–1114. doi:10.1126/science.1176210
29. Sun, Z. X.; Boneschanscher, M. P.; Swart, I.; Vanmaekelbergh, D.; Liljeroth, P. *Phys. Rev. Lett.* **2011,** *106,* 46104. doi:10.1103/PhysRevLett.106.046104
30. Pawlak, R.; Kawai, S.; Fremy, S.; Glatzel, T.; Meyer, E. *ACS Nano* **2011,** *5,* 6349–6354. doi:10.1021/nn201462g
31. Mohn, F.; Gross, L.; Meyer, G. *Appl. Phys. Lett.* **2011,** *99,* 053106. doi:10.1063/1.3619829
32. Ashino, M.; Obergfell, D.; Haluška, M.; Yang, S. H.; Khlobystov, A. N.; Roth, S.; Wiesendanger, R. *Nat. Nanotechnol.* **2008,** *3,* 337–341. doi:10.1038/nnano.2008.126
33. Ashino, M.; Obergfell, D.; Haluška, M.; Yang, S. H.; Khlobystov, A. N.; Roth, S.; Wiesendanger, R. *Nanotechnology* **2009,** *20,* 264001. doi:10.1088/0957-4484/20/26/264001
34. Ternes, M.; Lutz, C. P.; Hirjibehedin, C. F.; Giessibl, F. J.; Heinrich, A. J. *Science* **2008,** *319,* 1066–1069. doi:10.1126/science.1150288
35. Fukuma, T.; Ueda, Y.; Yoshioka, S.; Asakawa, H. *Phys. Rev. Lett.* **2010,** *104,* 016101. doi:10.1103/PhysRevLett.104.016101
36. Schmidt, R.; Lazo, C.; Kaiser, U.; Schwarz, A.; Heinze, S.; Wiesendanger, R. *Phys. Rev. Lett.* **2011,** *106,* 257202. doi:10.1103/PhysRevLett.106.257202
37. Abe, M.; Sugimoto, Y.; Custance, O.; Morita, S. *Appl. Phys. Lett.* **2005,** *87,* 173503. doi:10.1063/1.2108112
38. Moheimani, S. O. R. *Rev. Sci. Instrum.* **2008,** *79,* 071101. doi:10.1063/1.2957649
39. Binnig, G.; Smith, D. P. E. *Rev. Sci. Instrum.* **1986,** *57,* 1688–1689. doi:10.1063/1.1139196
40. Eaton, P.; West, P. *Atomic Force Microscopy;* Oxford University Press: Oxford, U.K., 2010. doi:10.1093/acprof:oso/9780199570454.001.0001
41. Akila, J.; Wadhwa, S. S. *Rev. Sci. Instrum.* **1995,** *66,* 2517–2519. doi:10.1063/1.1145651
42. Riis, E.; Simonsen, H.; Worm, T.; Nielsen, U.; Besenbacher, F. *Appl. Phys. Lett.* **1989,** *54,* 2530–2531. doi:10.1063/1.101064
43. Jørgensen, J. F.; Carneiro, K.; Madsen, L. L.; Conradsen, K. *J. Vac. Sci. Technol. B* **1994,** *12,* 1702–1704. doi:10.1116/1.587267
44. Stoll, E. P. *Ultramicroscopy* **1992,** *42,* 1585–1589. doi:10.1016/0304-3991(92)90488-6
45. Hug, H. J.; Stiefel, B.; van Schendel, P. J. A.; Moser, A.; Martin, S.; Güntherodt, H.-J. *Rev. Sci. Instrum.* **1999,** *70,* 3625–3640. doi:10.1063/1.1149970
46. Cai, C. Z.; Chen, X. Y.; Shu, Q. Q.; Zheng, X. L. *Rev. Sci. Instrum.* **1992,** *63,* 5649. doi:10.1063/1.1143395
47. Hofer, W. A.; Foster, A. S.; Shluger, A. L. *Rev. Mod. Phys.* **2003,** *75,* 1287–1331. doi:10.1103/RevModPhys.75.1287
48. Enevoldsen, G. H.; Pinto, H. P.; Foster, A. S.; Jensen, M. C. R.; Kühnle, A.; Reichling, M.; Hofer, W. A.; Lauritsen, J. V.; Besenbacher, F. *Phys. Rev. B* **2008,** *78,* 045416. doi:10.1103/PhysRevB.78.045416
49. Pou, P.; Ghasemi, S. A.; Jelinek, P.; Lenosky, T.; Goedecker, S.; Perez, R. *Nanotechnology* **2009,** *20,* 264015. doi:10.1088/0957-4484/20/26/264015
50. Oyabu, N.; Pou, P.; Sugimoto, Y.; Jelinek, P.; Abe, M.; Morita, S.; Pérez, R.; Custance, Ó. *Phys. Rev. Lett.* **2006,** *96,* 106101. doi:10.1103/PhysRevLett.96.106101
51. Bechstein, R.; González, C.; Schütte, J.; Jelínek, P.; Pérez, R.; Kühnle, A. *Nanotechnology* **2009,** *20,* 505703. doi:10.1088/0957-4484/20/50/505703
52. Arai, T.; Gritschneder, S.; Tröger, L.; Reichling, M. *J. Vac. Sci. Technol. B* **2010,** *28,* 1279–1283. doi:10.1116/1.3511505
53. Schwarz, A.; Allers, W.; Schwarz, U. D.; Wiesendanger, R. *Phys. Rev. B* **2000,** *61,* 2837–2845. doi:10.1103/PhysRevB.61.2837
54. Lauritsen, J. V.; Foster, A. S.; Olesen, G. H.; Christensen, M. C.; Kühnle, A.; Helveg, S.; Rostrup-Nielsen, J. R.; Clausen, B. S.; Reichling, M.; Besenbacher, F. *Nanotechnology* **2006,** *17,* 3436–3441. doi:10.1088/0957-4484/17/14/015
55. Enevoldsen, G. H.; Foster, A. S.; Christensen, M. C.; Lauritsen, J. V.; Besenbacher, F. *Phys. Rev. B* **2007,** *76,* 205415. doi:10.1103/PhysRevB.76.205415
56. Teobaldi, G.; Lämmle, K.; Trevethan, T.; Watkins, M.; Schwarz, A.; Wiesendanger, R.; Shluger, A. L. *Phys. Rev. Lett.* **2011,** *106,* 216102. doi:10.1103/PhysRevLett.106.216102
57. Bartels, L.; Meyer, G.; Rieder, K.-H. *Appl. Phys. Lett.* **1997,** *71,* 213. doi:10.1063/1.119503
58. Repp, J.; Meyer, G.; Stojković, S. M.; Gourdon, A.; Joachim, C. *Phys. Rev. Lett.* **2005,** *94,* 026803. doi:10.1103/PhysRevLett.94.026803
59. Hölscher, H.; Schwarz, U. D.; Wiesendanger, R. *EPL* **1996,** *36,* 19. doi:10.1209/epl/i1996-00181-4

60. Hölscher, H.; Schwarz, U. D.; Zwörner, O.; Wiesendanger, R. *Phys. Rev. B* **1998,** *57,* 2477. doi:10.1103/PhysRevB.57.2477
61. Oja, R.; Foster, A. S. *Nanotechnology* **2005,** *16,* S7. doi:10.1088/0957-4484/16/3/002
62. Matlab computing environment and programming language developed by MathWorks, Natick, MA, USA.
63. Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids;* Oxford University Press: Oxford, U.K., 1989.
64. Agrawal, P. M.; Rice, B. M.; Thompson, D. L. *Surf. Sci.* **2002,** *515,* 21–35. doi:10.1016/S0039-6028(02)01916-7
65. Hölscher, H.; Allers, W.; Schwarz, U. D.; Schwarz, A.; Wiesendanger, R. *Appl. Phys. A* **2001,** *72,* S35. doi:10.1007/s003390100724

License and Terms

Beilstein Journal of Nanotechnology

The memory effect of nanoscale memristors investigated by conducting scanning probe microscopy methods

César Moreno[*1,2], Carmen Munuera[2,§], Xavier Obradors[2] and Carmen Ocal[*2]

Full Research Paper

Open Access

Address:
[1]International Center for Young Scientists, National Institute for Materials Science, Tsukuba, Ibaraki 305-0047, Japan and [2]Institut de Ciència de Materials de Barcelona, CSIC, Campus de la UAB, 08193 Bellaterra, Spain

Email:
César Moreno[*] - MORENOSIERRA.Cesar@nims.go.jp;
Carmen Ocal[*] - cocal@icmab.es

* Corresponding author
§ Current address: Instituto de Ciencia de Materiales de Madrid, CSIC, Cantoblanco 28049 Madrid, Spain

Keywords:
conductive scanning probe micoscopy; memristor; 3-D modes; resistive switching; scanning probe microscopy

Beilstein J. Nanotechnol. **2012,** *3,* 722–730.
doi:10.3762/bjnano.3.82

Received: 31 July 2012
Accepted: 28 September 2012
Published: 06 November 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We report on the use of scanning force microscopy as a versatile tool for the electrical characterization of nanoscale memristors fabricated on ultrathin $La_{0.7}Sr_{0.3}MnO_3$ (LSMO) films. Combining conventional conductive imaging and nanoscale lithography, reversible switching between low-resistive (ON) and high-resistive (OFF) states was locally achieved by applying voltages within the range of a few volts. Retention times of several months were tested for both ON and OFF states. Spectroscopy modes were used to investigate the *I–V* characteristics of the different resistive states. This permitted the correlation of device rectification (reset) with the voltage employed to induce each particular state. Analytical simulations by using a nonlinear dopant drift within a memristor device explain the experimental *I–V* bipolar cycles.

Introduction

The current knowledge-based society requires a new, more-powerful memory technology for the development of any field concerning human activity, such as biomedicine, space research, meteorological predictions, simulation in basic research science, and entertainment. Nowadays computers use two types of memory, i.e., dynamic random-access memory

(DRAM) and static random-access memory (SRAM). The handicap of these fast memories is that they are volatile, i.e., data are lost when the power supply is removed. For this reason computers also use a nonvolatile memory, i.e., hard-disk drives (HDDs). HDDs are nonvolatile but also slow, thus increasing the computer start-up time. Another kind of nonvolatile memory commonly employed for data storage in hard disks, digital cameras, USB memory sticks, or cellular disks is Flash memory. Its main drawbacks are a slow writing speed and the limited number of write/erase cycles that can be endured.

At this point, the need for new materials with higher performance has encouraged extensive basic research in this field. Resistive random-access memory (RRAM) based on a memristive behaviour has become an exciting and strongly developing field, because when combined with transistors in a hybrid chip, memristors could radically improve the performance of digital circuits without the need for further reduction of transistor dimensions [1]. These memory resistors are similar to nonlinear resistors exhibiting memory effects and are essentially two-terminal devices whose resistance depends on the polarity of the applied voltage. The simplest memristor consists of a thin oxide or semiconductor doped film (of thickness w) between top and bottom metallic electrodes. The slope of the functional relationship between the time integral of the current and the time integral of the voltage across the device element is the so-called memristance, $M(w)$, similar to a variable resistance, which provides a pinched effect to the characteristics of the devices.

Memristance becomes more important as the critical dimensions shrink to the nanometre scale [2] and therefore the characterization of the local electrical properties becomes more and more important. In this work we have combined conductive scanning force microscopy imaging and single-point current–voltage spectroscopy, with more advanced spectroscopy measurements (3-D modes) to characterize the nanoscale electrical response of thin films presenting memristive behaviour. The systems of choice for benchmarking are ultrathin films of conducting $La_{0.7}Sr_{0.3}MnO_3$ (LSMO) with a thickness in the range of 10 to 24 nm, grown by chemical solution deposition (CSD). The films are transferred through air to the SFM equipment, and therefore, a thin top layer likely with a slightly different oxygen content is expected to cover the conducting LSMO film, which otherwise acts as the bottom electrode. The present work particularly aims at demonstrating and evaluating the electrical responses, including the memory effect, of nanoscale memristors, by a combination of contact C-SFM modes, which reveal the technique as an ideal tool for the research and development of memristive systems on a nanometre scale.

Results and Discussion

Resistive-switching procedure

Conductive scanning force microscopy was used either to modify (writing/erasing) or to characterize (reading) the electrical properties of the sample under study. Current images were acquired in a noninvasive manner (no sample indentation) by using the contact operation mode at the lowest possible applied load while still obtaining stable signals. The conducting tip (top electrode) was placed in direct contact with the sample surface, under controlled load, i.e., by using a normal force feedback, and the current between the tip and sample was measured. Simultaneous topographic images $z(x,y)$ and current maps $I(x,y)$ over a given region at a fixed voltage permitted observation of the complete writing (positives V_{tip}), reading (small negative V_{tip}) and erasing (negatives V_{tip}) process. In our setup, the sample was always grounded and the voltage was applied to the tip.

Given the insulating character of the substrates used ($SrTiO_3$), the direct electrical contact to ground was established through a metallic clamp (counter electrode) firmly attached to the film at the sample border (millimetres apart from the tip–surface contact). For topographic images a colour code is commonly used in which bright colours indicate high values and dark means low; however, interpretation of colour-coded current maps depends on the voltage sign as well as the absolute current magnitude. Thus higher currents appear darker in C-SFM images taken at negative V_{tip}, while brighter for positive V_{tip}. The current–voltage (I–V) characteristics of the contact were measured as a function of the bias voltage applied between tip and sample, starting from negative tip voltages.

Once the surface location to be modified had been selected by large-scale topographic imaging, typical writing–reading C-SFM experiments, as presented in Figure 1, were based on the following protocol:

1. The scan range is set to the reading area. The tip is biased at V_{rd}, and topography and current images are simultaneously acquired to characterize the conductance state of the surface.
2. The scan range is reduced to the writing area. The V_{tip} is set to a given positive value, and topography and current images are simultaneously acquired.
3. The V_{tip} is brought back to the reading value (V_{rd}) and any possible modification is checked by repeating (i). The process in (ii) is repeated at increased V_{tip} until no current is detected within the writing region. The corresponding voltage, which will be seen to depend slightly on film thickness, is considered as the onset for writing, V_{wr}.

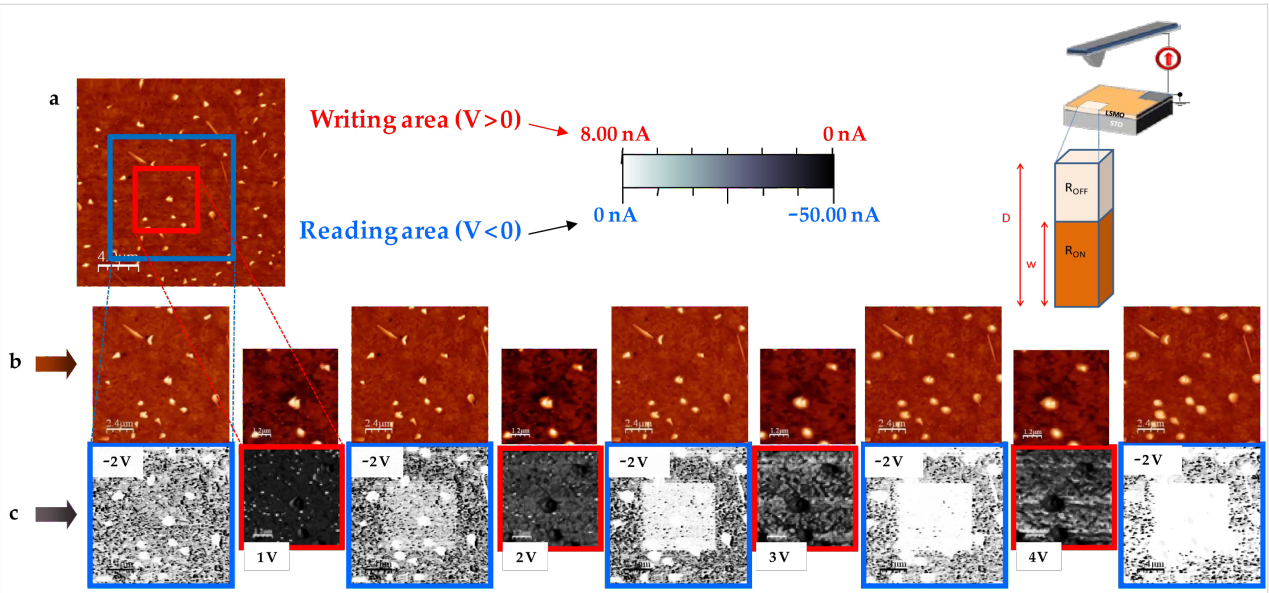

Figure 1: Resistance-switching sequence: Writing and reading of local conductance modifications made on the LSMO surface. The transition between low-resistive (LR) and high-resistive (HR) states is performed by writing at an appropriate V_{tip} with the tip of the conducting scanning force microscope (C-SFM). (a) Topographic image (20 μm × 20 μm) of a 24 nm thick LSMO thin film. A certain concentration of outcropped insulating $(La,Sr)_xO_y$ (LSO) nanodots (50–80 nm high) are used as an in situ reference for topographic and conductivity measurements. Concentric squares on the image correspond to the writing (inner square) and reading (outer square) regions. Top-right inset: Schematics of the measuring setup indicating relevant parameters. (b), (c) Series of topographic and current-map images displaying the reading area (12 μm × 12 μm) taken at V_{rd} = −2 V, before (virgin state) and after writing a HR area (6 μm × 6 μm) in several steps V_{wr} = +1, +2, +3, +4 V. Note: the colour code in the current maps depends on the voltage sign, and therefore a dark colour means higher current for V_{tip} < 0 but lower current for V_{tip} > 0 (see scale bar).

The retention time of the written high-resistive state of Figure 1 is further investigated in the following sections. Nonetheless, as illustrated in Figure 2, if desired, a total or partial switch back to the conductive state can be accomplished by scanning the modified region (or a portion of it) at a negative V_{tip} for which the conducting character is recovered. This is called the erasing voltage, V_{er}, and has been found to be at least $V_{er} = -V_{wr}$ [3]. The reversible transition between differently resistive states of the LSMO thin film are performed (writing and erasing) and followed (reading) in a noninvasive way.

Analytical simulation of the memristor *I–V* behaviour

Typical *I–V* characteristics conducted by application of a voltage cycle at a fixed location on the virgin (preswitched) 10 nm thick LSMO thin film surface presented a bipolar switching behaviour, such as that shown in Figure 3a. The clearly asymmetric curve indicates that diverse microscopic processes govern charge transport under each polarity; the different branches are capable of being described by *I–V*s of electrical circuits in which both resistors and rectifiers are

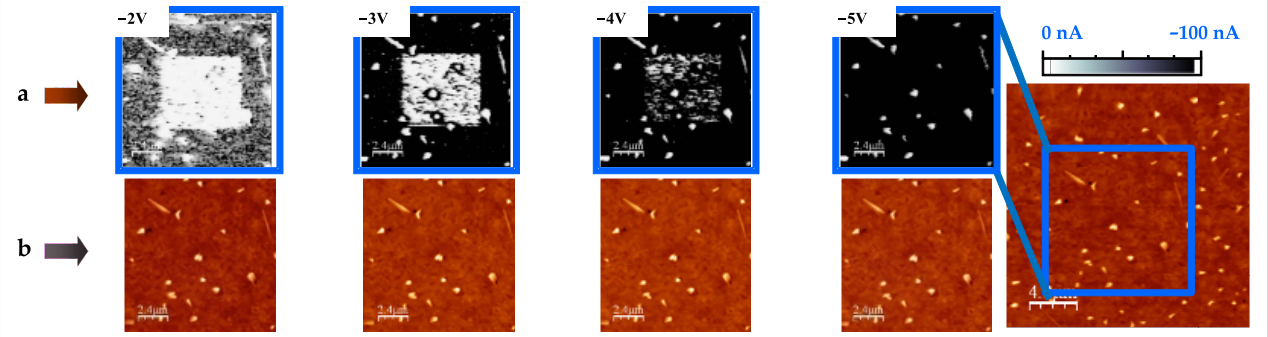

Figure 2: Resistance-switching sequence: Erasing back to a LR state starting from the local HR (nonconducting) modification, performed on the LSMO surface at V_{wr} = +4 V. The nonconducting character of the written region (white square) is reversed by changing the voltage polarity: (a) Reading process at V_{rd} = −2V and three steps of the erasing process V_{er} = −3, −4 and −5 V. The conducting character of the initial LSMO is recovered after applying $V_{er} \approx -V_{wr}$. (b) Series of topographic images of the erased region show that the process does not result in morphological changes.

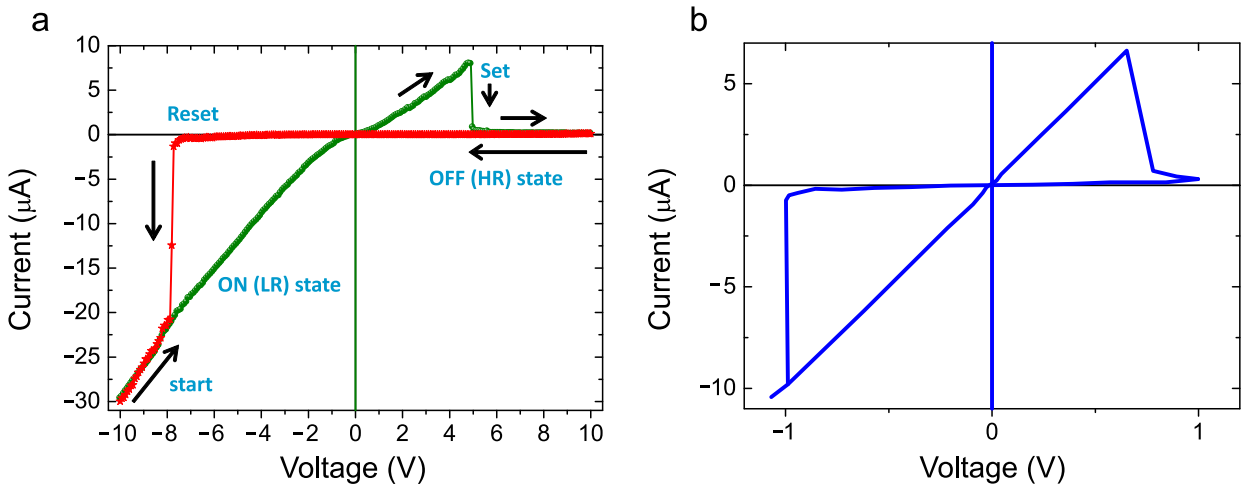

Figure 3: Experimental (a) and simulated (b) *I–V* curves. (a) Starting from a conducting (LR) state, a voltage ramp is applied to the tip, from negative to positive values and back, to complete the cycles as indicated by the arrows. *Set* indicates the change from the initial conductive state to a more resistive one and *Reset* corresponds to the restoration of the ohmic behaviour. LR and HR states are identified as ON and OFF, indicating the pass or blockage of electric current through the device. To have access to the complete range of currents (1 pA to 1 mA) an external *I–V* converter was used (see Experimental section). (b) Calculated *I–V* curve for a memristor with a nonlinear dopant drift under an external sinusoidal bias with a voltage amplitude v_0 = 10 V, a resistance ratio R_{OFF}/R_{ON} = 70 and a memristor length D = 10 nm. The units of the axes are expressed in units of voltage amplitude v_0 = 10 V and the unit of current is $I_0 = v_0/R_{ON}$ = 10 μA. I_0 is the maximum current reached through the memristor.

involved, as already observed in other thin-oxide-based memristor systems [4].

The fundamental memristive system theory states that the observation of a pinched current-versus-voltage hysteresis loop measured from an experimental two-terminal device, when driven by a dc and/or sinusoidal signal of any frequency, implies that the device is a memristor [5]. A basic mathematical definition of a memristor, assuming an ohmic electron conduction and linear ionic drift in a uniform field, is given by the following equations [2]:

$$M(w) = \frac{v(t)}{I(t)} = \frac{w}{D}R_{ON} + R_{OFF}\left(1 - \frac{w}{D}\right) \quad (1)$$

$$\frac{dw}{dt} = \mu E = \frac{\mu R_{ON}}{D} I(t) \quad (2)$$

where M is the memristance, w is the size of the doped region, D is the thickness of the film, R_{ON} and R_{OFF} correspond to the resistance of the memristor for completely uniformly doped or undoped cases, respectively, and μ is the drift velocity of the dopants under an electric field E across the doped region in the presence of a current $I(t)$.

The switching characteristic observed for a particular memristive system helps to classify the nature of the dopant drift inside the memristor. If the generated electric field can be assumed to be small, the linear dopant-drift model can approximate the dynamics of a memristor. However, this model is not valid in our case. During the electrical measurement of the *I–V* characteristic in contact mode, an enormous electric field can be generated considering the small tip–sample contact area and the Schottky-like tip–film contact. Numerical simulation for 60 nm thick films shows that the electric field within a few nanometres from the tip–sample contact is a factor of ~20 larger than that in region of the film closer to the substrate interface [6].

The influence of a nonuniform electric field significantly suppresses the drift of the dopants. Nonlinear dopant drift can be taken into account in the memristor simulation *I–V* curve by introducing an appropriate window function in Equation 2. Thus, by using the window function proposed by Joglekar [7], i.e.,

$$f(x) = 1 - (2x - 1)^p \quad (3)$$

with p being a positive exponent , we obtain the simulated *I–V* curve showed in Figure 3b, which corresponds to a memristor in a nonlinear model in which the dopant drift is heavily suppressed only near the boundaries (p = 10). Experimental and simulated curves are in excellent agreement, pointing out that the measured hysteretic *I–V* curve is consistent with a hard-switching system governed by nonlinear ion drift in a high electric field. These nonlinearities have been related to long retention times and fast switching processes [8] in accordance with the endurance test displayed in Figure 4.

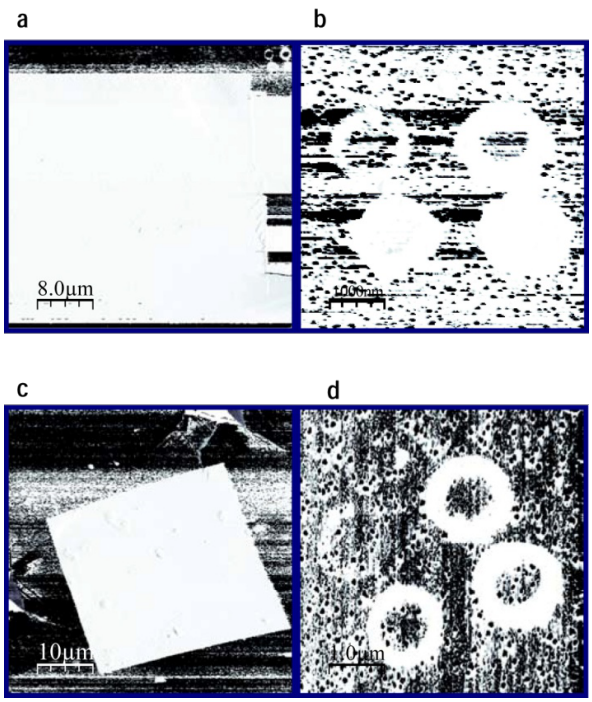

Figure 4: Current-map reading images of different high-resistive modifications. As-fabricated nonconducting square region (a) and a series of HR rings written at different increasing V_{wr} from left to right and top to bottom (b). The bottom images (c and d) were taken after four months by relocating the samples in the C-SFM system to observe the same surface regions (slightly rotated). The samples were annealed at up to 400 K and submitted to magnetic fields (up to 3 T) as part of endurance tests between sets of images. The images were taken at V_{rd} = −2 V and the total scale goes from 0 to −100 nA (white to black).

In order to gain further insight into the nature of the process, the durability of the tip-voltage-induced modifications was tested as a function of time, temperature and magnetic field. As can be seen in the current images presented in Figure 4 for different geometries of the modifications (square and rings) and different V_{wr}, the initial high-resistive states (a and b) were found to persist for at least up to four months, as well as upon annealing up to 400 K, and show no important degradation under magnetic fields up to 3 T (c and d). The magnetic-field independence is what is to be expected from the memristor equations. Note that, as well as the square-shape modification (a, c), the HR rings persist in all cases after this long period of time, excluding important degradation. The small differences in the measured current between b and d are ascribed to slightly different reading conditions. However, a thin contamination layer adsorbed during exposure to air and the endurance treatments, connecting the inner ring with the outer conducting region, cannot be excluded. A certain contribution arising from these effects would be similar for all modifications, independently of the initial writing parameters; therefore, a comparative spectroscopic investigation is presented in the following to obtain direct comparison between electric-current behaviour for different HR states.

Complementary memory tests: 3-D modes and *I–V* spectroscopy

Figure 5a depicts a reading current map, $I(x,y)$, performed over a surface region containing a series of squared HR regions obtained by writing at V_{tip} voltages ranging from V_{wr} = +1 to +6 V. As seen in the line profile $I(x)$ of Figure 5b, taken along the green dotted line in Figure 5a, at the employed V_{rd} = −2 V the current approaches zero for all modifications performed with V_{wr} > +1 V. The maximum current measured (±100 nA) corresponds to the saturation of the current amplifier in the experimental setup used.

To analyse the memory effect of the fabricated memristors further in terms of the electronic current passing through different devices and the restoring voltages, a complementary C-SFM method is further proposed. This method is based on the so-called 3-D modes [9,10] and is used here to measure the current (I) as a function of the applied voltage ($V = V_{tip}$) along a single surface line (x), thus obtaining a $I(x,V)$ image in which the colour scale is the current measured at each pixel, the fast scan continuously sweeps the same surface line, and the voltage is changed at every line of the slow scan. This was the procedure used to obtain the 3-D image of Figure 5c: the tip was positioned to cross the whole series of modifications at the same green dotted line of Figure 5a and V_{tip} was changed from −4 to +4 V (fast scan). In such a way horizontal profiles in the image of Figure 5c provide the current values along the surface line at a given voltage (Figure 5d) and vertical profiles correspond to the characteristic *I–V* curve at a given point (x,y) (Figure 5e). As a result, the whole set of *I–V* curves were simultaneously obtained for an ensemble of surface locations lying along the selected line.

The first noticeable piece of information is extracted from the current profile taken at V_{tip} = −4 V and depicted in Figure 5d. Note that the $I(x)$ measured on modifications made at V_{wr} = +2 and +3 V has considerably increased from that shown in Figure 5b (V_{tip} = −2 V). It is worth noting that as commented above the resistive switching is reversed for $V_{er} > -V_{wr}$ and therefore, this observation is the result of a reversible switch (erasing process) from each HR to a more conducting LR state. Conversely, as expected HR states written at +4, +5 and +6 V remained unaltered. The reading image of Figure 5f recorded after the 3-D imaging further illustrates the reversibility of the process. On the one hand, as a consequence of the scanning at −4 V a line has been erased (HR to LR) on top of the squares written at +3 and +2 V. On the other hand, note that the V_{tip} reached during the 3-D imaging goes up to +4 V, modifying the

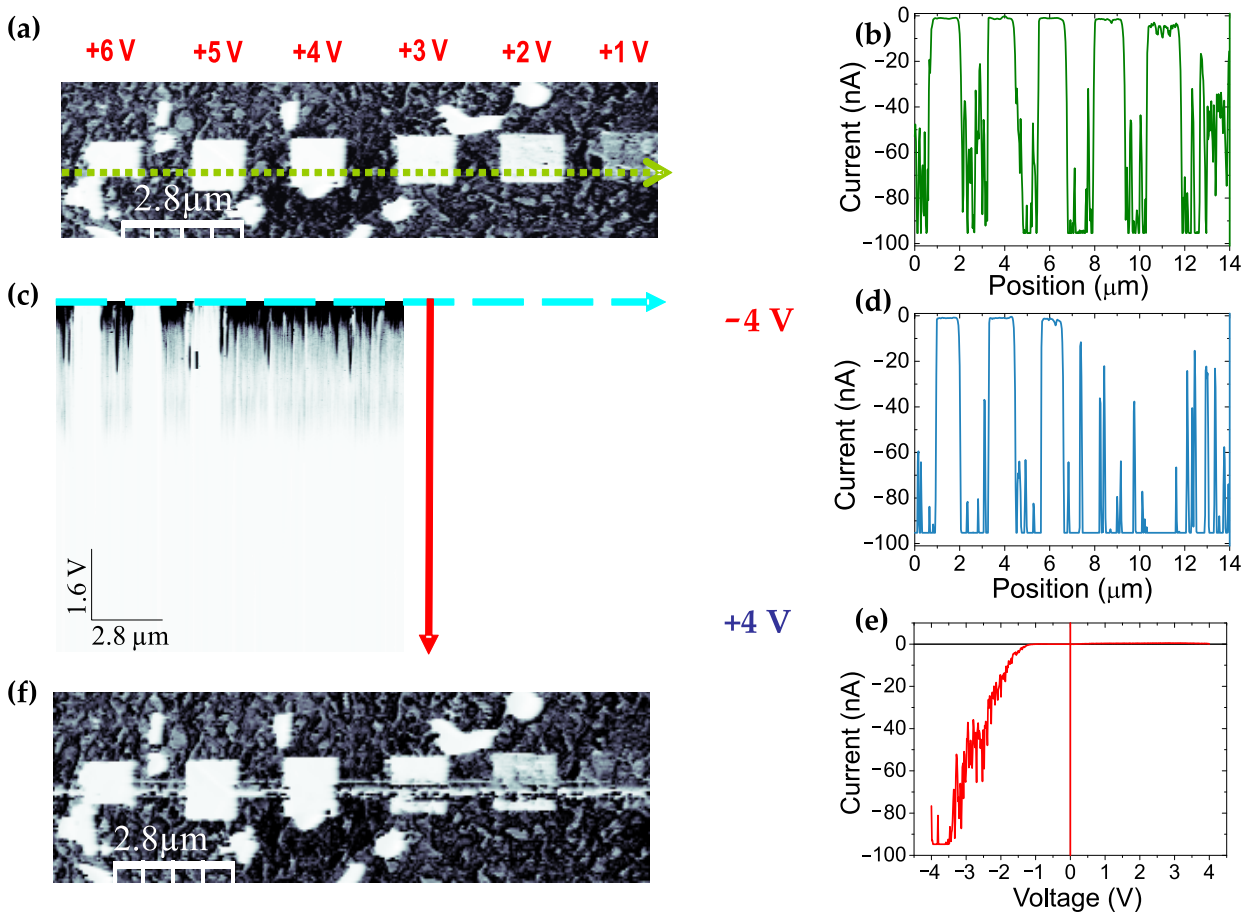

Figure 5: 3-D mode spectroscopy. (a) Current-map reading image (V_{rd} = −2 V) of different HR states written at V_{wr} from +1 to +6 V. (b) Current profile along the dotted green line marked in (a) (V_{tip} = −2 V). (c) $I(x,V)$ image obtained at the same surface line: the colour scale corresponds to the current measured as a function of the lateral position (horizontal axis) and the voltage (slow scan) ranging from V_{tip} = −4 V (top) to +4 V (bottom). (d) Current profile along the horizontal blue dashed line in (c) (V_{tip} = −4 V) and (e) I–V curve corresponding to the vertical profile (red line in c) centred in the modification made at V_{wr} = +3 V. (f) Current-map reading image (V_{rd} = −2 V) after 3-D image. Because the voltage range (±4 V) of the 3-D image is high enough to modify the surface, a line has been written (LR to HR) between modifications at the tip passage. For a similar reason a line has been erased (HR to LR) at those modifications where the corresponding V_{er} has been exceeded. Note: the maximum currents measured (±100 nA) correspond to the saturation of the amplifier setup.

virgin film surface (LR to HR) between the pre-existing marks. From the above results one can think about the capability of the proposed process to fabricate wire-like modifications, which are promising for nanodevice applications, made by the sweeping action of the tip at different voltages. The longitudinal dimension of such wires is only limited by the total piezo scan range (up to hundreds of microns in common SFM systems) and the short dimension is limited by the contact area of the tip and, therefore, dependent on the tip radius. At present, lines as thin as about 50–100 nm can be performed, demonstrating the scalability of the resistive-switching effect to nanoscale devices. It is worth mentioning here that, by comparing Figure 5a and Figure 5f we can rule out any tip effects in I–V measurements due to tip contamination or coating wear. The nonmodified regions of the LSMO films serve as an in situ quality test of the tip performance. No measurable changes were observed in the current measured on these nonmodified regions after the whole series of experiments in Figure 5, ensuring the good electrical performance of the tip and the reliability of the measurements.

As will be confirmed below and has been previously demonstrated by means of Kelvin probe force microscopy [3], the results presented here support the memristor memory effect capable of memorizing the amount of charge that has passed through it. In order to gain insight to this issue, analysis of the I–V characteristics is needed. As commented above, vertical profiles in the $I(x,V)$ would provide, in principle, such curves at each point of the selected line. An example of such I–V is

shown in Figure 5e for a location within the modification performed at +3 V (vertical red line in Figure 5a). Though the curve is relatively noisy, it shows a clear rectifying behaviour. It is worth remarking that in the 3-D modes the current is measured during tip scanning, i.e., the point at which the tip contacts the surface is continuously changing, and therefore, some averaging and deviation of the true *I*–*V* behaviour can exist. In addition, as has been described above, the resistive state of a given modification is altered if the total range exceeds the absolute value of the corresponding V_{wr}. Therefore, in order to explore the true *I*–*V* curve by improving the signal-to-noise ratio and using an adequate voltage range for each HR state, we also measured individual *I*–*V* curves, such as those presented in Figure 6.

I–*V* curves (Figure 6e and Figure 6f) obtained on the modifications made at the lower voltages (V_{wr} = +1 and +2 V) are nearly symmetric, with a pseudo-gap around $V = 0$. Interestingly, all *I*–*V* curves for larger V_{wr} (Figure 6a to Figure 6d) show a clear rectifying behaviour such that charge injection occurs exactly at $V_{Reset} = -V_{wr}$ indicating that to change the memory state of the device, carriers must overcome a barrier height equal to the switch-on voltage employed in each case. At this point, the particular HR state is reversed to the LR state. In other words, $V_{Reset} = V_{er}$. As already proposed for other memristive devices showing similar behaviour [4], a combination of a non-ohmic contact interface (maybe Schottky-like) and quasi-ohmic contact at the opposite interface could be responsible for the observed features in the electrical responses. This interpretation is in agreement with the nonuniform-dopant-drift analysis of the hysteretic characteristic cycles. Due to air exposure, the oxygen content of the topmost surface region of the LSMO may be different to the stoichiometric composition constituting the active doped element of the memristor.

As an alternative to other investigations made on memristor devices with two planar electrodes in which the active regions are buried under the metal contact, here we have been able to analyse the memory effect of nanoscale resistive switching with the very same probe that was used for the fabrication, providing direct evidence for the restoring phenomena as an activated process involving carrier injection. Integration of the presented nanoscale memristor devices opens new possibilities for the low cost and scalable development of ultrahigh-density resistive and nonvolatile memory cells.

Conclusion

In this work we have shown the full capabilities of electrical scanning probe microscopy modes to modify and characterize memristive LSMO thin films on the nanoscale. Hard switching

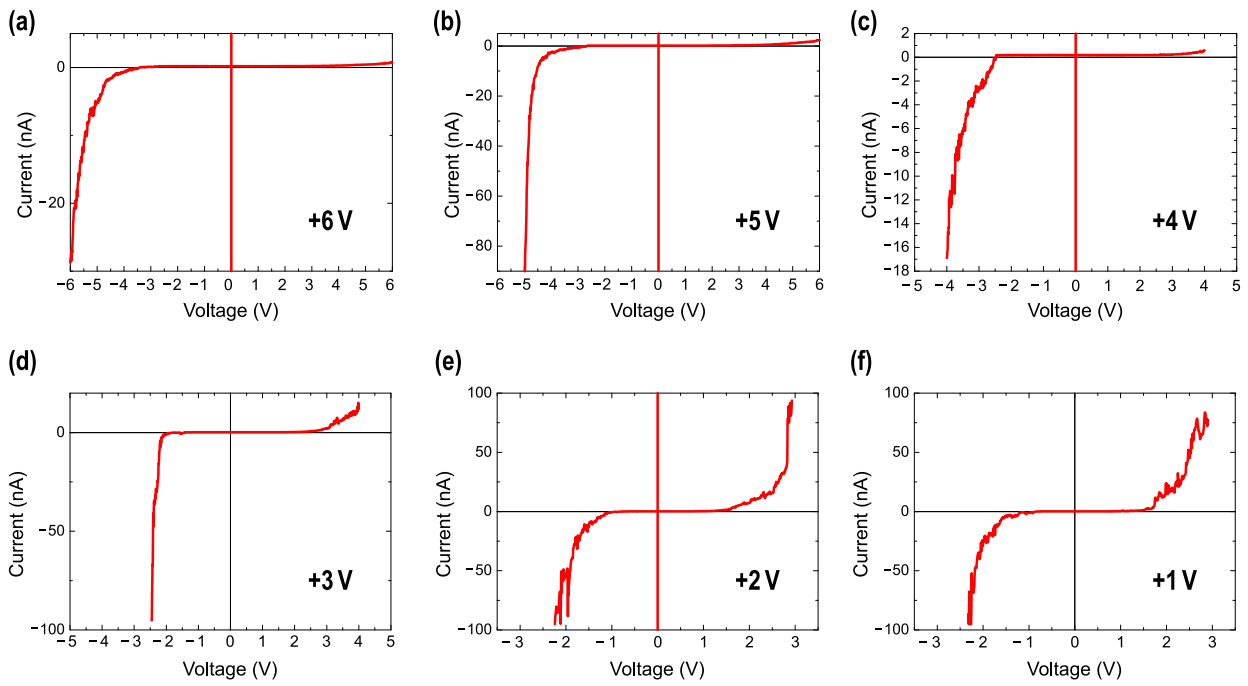

Figure 6: *I*–*V* curves obtained at specific surface locations within the different resistive modifications of Figure 5. The corresponding V_{wr} is indicated in each panel. The voltage range has been expanded to different values in order to visualize that rectification (reset) occurs at $V_{tip} \approx -V_{wr}$.

with long retention times and fast switching process are consistent with the nonuniform-drift model used here. We have proposed a nanoscale methodology to fabricate and characterize durable memristors using the tip of the C-SFM as a movable top electrode, as an alternative to standard two-plate devices.

The superb control of the top electrode and perfect interface between the native oxide and the bottom conducting LSMO electrode make the electrical characteristics of the memristor highly reproducible.

Experimental

Sample preparation

$La_{0.7}Sr_{0.3}MnO_3$ (LSMO) films with a thickness in the range of 10–24 nm were grown by chemical solution deposition (CSD) on (100)$SrTiO_3$ (STO) substrates [11]. The metal–organic solution was deposited by spin coating, at 6000 rpm for 2 min, and annealed at temperatures in the range of 900–1000 °C under an oxygen gas flux, for different times up to 12 h. Heteroepitaxial growth of LSMO films was confirmed by θ–2θ XRD patterns and XRD *q*-plot measurements showed that the films are fully strained. The values of resistivity, magnetoresistance and Curie temperature are very similar to those observed in LSMO films grown by physical vapour-deposition techniques, such as sputtering or pulsed-laser deposition, leading us to conclude that a similar epitaxial quality is achieved with films grown by our CSD approach. More experimental details of the growth process, as well as structural and magnetic characterization, have been reported elsewhere [12,13].

Between the retention and endurance tests, the samples were always stored at ambient conditions in a plastic box supplied with silica powders.

Scanning force microscopy

Scanning force microscopy (SFM) measurements were performed using a commercial head and software from Nanotec [14] under a N_2 environment (RH < 5%) to diminish any possible humidity effects. For the conductivity measurements (C-SFM), we used either commercial conductive B-doped diamond-coated tips with k = 2.8 N/m (Nanosensors) or Cr–Pt coated Si tips with k = 3 N/m, resonance frequency of 75 kHz and contact resistance of 300 Ω on a Pt thin film surface (Budgetsensors in-factory specifications). The same tip was used in all the C-SFM experiments of at least one series. To check tip–sample conditions, the adhesion force was systematically determined from force-versus-distance curves prior to and after each conductivity experiment. In addition, comparison of the conducting properties of nonmodified regions prior to and after the experiments was used as an in situ quality test to assess proper tip performance. Applied forces were commonly in the 50–100 nN range. For all measurements in this study, electrical testing was performed by applying the bias voltage to the top electrode (tip) keeping the bottom electrode (LSMO film) grounded. When needed, an external *I–V* converter (Stanford Systems) was used to gain access to a wide range of compliance currents (1 pA to 1 mA).

Acknowledgements

This work has been supported by the Spanish Government through grants MAT2010-20020 and NANOSELECT CSD2007-00041. C. Moreno is presently supported by the Japanese Ministry for Education, Culture, Sports, Science and Technology (MEXT) through the ICYS program. C. Munuera thanks financial support from the Spanish "Juan de la Cierva" postdoctoral program JCI-2011-08815.

References

1. Williams, R. *IEEE Spectrum* **2008,** *45,* 28–35. doi:10.1109/MSPEC.2008.4687366
2. Strukov, D. B.; Snider, G. S.; Stewart, D. R.; Williams, R. S. *Nature* **2008,** *453,* 80–83. doi:10.1038/nature06932
3. Moreno, C.; Munuera, C.; Valencia, S.; Kronast, F.; Obradors, X.; Ocal, C. *Nano Lett.* **2010,** *10,* 3828–3835. doi:10.1021/nl1008162
4. Yang, J. J.; Pickett, M. D.; Li, X.; Ohlberg, D. A. A.; Stewart, D. R.; Williams, R. S. *Nat. Nanotechnol.* **2008,** *3,* 429–433. doi:10.1038/nnano.2008.160
5. Chua, L. *Appl. Phys. A* **2011,** *102,* 765–783. doi:10.1007/s00339-011-6264-9
6. Lee, M. H.; Kim, K. M.; Song, S. J.; Rha, S. H.; Seok, J. Y.; Jung, J. S.; Kim, G. H.; Yoon, J. H.; Hwang, C. S. *Appl. Phys. A* **2011,** *102,* 827–834. doi:10.1007/s00339-011-6266-7
7. Joglekar, Y. N.; Wolf, S. J. *Eur. J. Phys.* **2009,** *30,* 661. doi:10.1088/0143-0807/30/4/001
8. Strukov, D. B.; Williams, R. S. *Appl. Phys. A* **2009,** *94,* 515–519. doi:10.1007/s00339-008-4975-3
9. Munuera, C.; Barrena, E.; Ocal, C. *Nanotechnology* **2007,** *18,* 125505. doi:10.1088/0957-4484/18/12/125505
10. Gómez-Navarro, C.; Gil, A.; Álvarez, M.; De Pablo, P. J.; Moreno-Herrero, F.; Horcas, I.; Fernández-Sánchez, R.; Colchero, J.; Gómez-Herrero, J.; Baró, A. M. *Nanotechnology* **2002,** *13,* 314. doi:10.1088/0957-4484/13/3/315
11. Hassini, A.; Pomar, A.; Gutiérrez, J.; Coll, M.; Romà, N.; Moreno, C.; Ruyter, A.; Puig, T.; Obradors, X. *Supercond. Sci. Technol.* **2007,** *20,* S230–S238. doi:10.1088/0953-2048/20/9/S18
12. Moreno, C.; Abellán, P.; Hassini, A.; Ruyter, A.; Pérez del Pino, A.; Sandiumenge, F.; Casanove, M.-J.; Santiso, J.; Puig, T.; Obradors, X. *Adv. Fun. Mat.* **2009,** *19,* 2139–2146. doi:10.1002/adfm.200900095
13. Moreno, C.; Munuera, C.; Pérez del Pino, A.; Gutiérrez, J.; Puig, T.; Ocal, C.; Obradors, X.; Ruyter, A. *Phys. Rev. B* **2009,** *80,* 094412. doi:10.1103/PhysRevB.80.094412
14. Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. *Rev. Sci. Instrum.* **2007,** *78,* 013705. doi:10.1063/1.2432410

Beilstein Journal of Nanotechnology

Friction and durability of virgin and damaged skin with and without skin cream treatment using atomic force microscopy

Bharat Bhushan*[1], Si Chen[1,2] and Shirong Ge[2]

Full Research Paper

Open Access

Address:
[1]Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics, The Ohio State University, 201 W 19th Ave., Columbus, OH 43210 USA and [2]Institute of Tribology and Reliability Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116 China

Email:
Bharat Bhushan* - Bhushan.2@osu.edu

* Corresponding author

Keywords:
atomic force microscopy; damaged skin; pig skin; rat skin; skin cream

Beilstein J. Nanotechnol. **2012,** *3,* 731–746.
doi:10.3762/bjnano.3.83

Received: 20 July 2012
Accepted: 17 October 2012
Published: 08 November 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

Skin can be damaged by the environment easily. Skin cream is an effective and rapid way to moisten the skin by changing the skin surface properties. Rat skin and pig skin are common animal models for studies and were used as skin samples in this study. The nano- and macroscale friction and durability of damaged skin were measured and compared with those of virgin (intact/undamaged) skin. The effect of skin cream on friction and durability of damaged and virgin skin samples is discussed. The effects of velocity, normal load, relative humidity and number of cycles were studied. The nanoscale studies were performed by using atomic force microscope (AFM), and macroscale studies were performed by using a pin-on-disk (POD) reciprocating tribometer. It was found that damaged skin has different mechanical properties, surface roughness, contact angle, friction and durability compared to that of virgin skin. But similar changes occur after skin cream treatment. Rat and pig skin show similar trends in friction and durability.

Introduction

Skin is the largest outer organ. The skin structure of mammals is mainly composed of three distinct layers: subcutis, dermis, and epidermis [1-6]. Rat skin and pig skin are common models used for skin in health and cosmetics studies. Figure 1 shows the epidermis and dermis of pig and rat skin [6,7]. The epidermis is the outer layer of skin. It contains four distinct cellular layers: basal layer, spinous layer, granular layer, and stratum corneum (keratin layer) [6]. Figure 1 also shows the chemical structure of the major components of the stratum corneum. Corneocytes with lipid bilayers exist in stratum corneum [6]. Intact (i.e., undamaged; hereafter referred to as "virgin") skin is covered with a thin hydrophobic lipid layer in its outer layer, containing triglycerides, diglycerides, fatty acids, wax esters, squalene, cholesterol, and cholesterol esters [8]. The

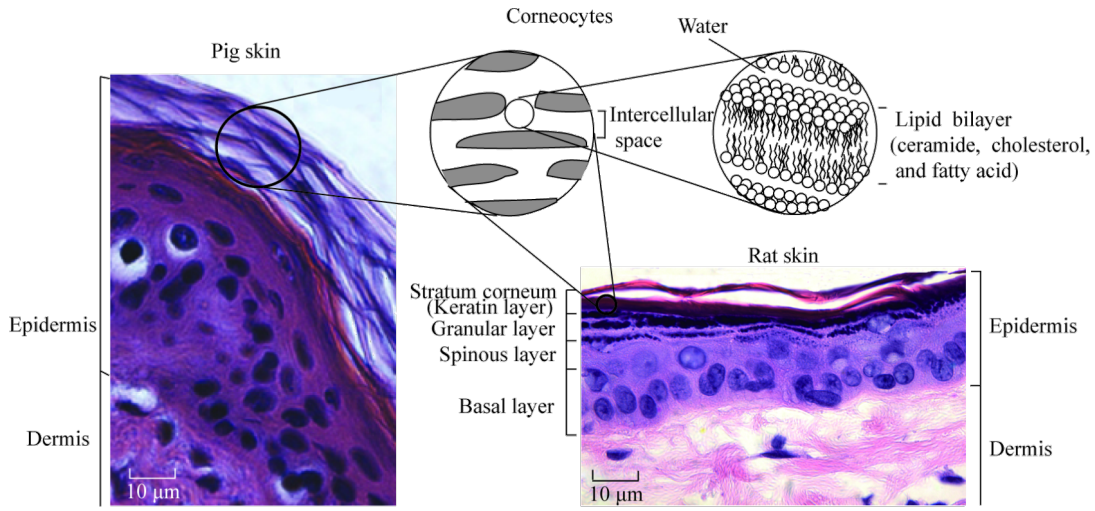

Figure 1: Histology of pig skin (photography reprinted from [7]) and rat skin (photography reprinted from [6]), and chemical structure of major lipid components of the stratum corneum.

intercellular lipids play an essential role in the establishment or maintenance of water-retention capacity in the stratum corneum [9]. Skin supports the body with protection, sensation, heat regulation, water resistance and so on, but environmental conditions such as dry and cold weather can reduce the moisture content of skin and induce epidermal hyperplasia, mast cell degranulation, cytokine secretion, increased skin roughness and physical discomfort [6,10,11].

Skin care products are developed to moisturize skin surface by humectants, which are a component of skin creams that can increase the moisture retention of stratum corneum and reduce the incidence of dry and flaky skin in vivo [12]. Humectants in skin cream attract and hold water in the skin, acting on the inside (i.e., moisture from the dermis to the epidermis/stratum corneum) and on the outside (i.e., moisture from the environment to the skin) [13]. Glycerin, lactic acid, potassium lactate, urea, sodium PCA, and propylene glycol are the humectants in common skin creams [11]. In general, polyols are the most effective humectants, especially the trihydroxylated molecule glycerol. Moisturizers containing glycerol provide enduring moisturization by binding and retaining water or by minimization of water loss. Glycerol can also hinder crystal-phase transitions induced by humidity in stratum corneum lipids and thus enhance the function of the skin as a barrier. In healthy skin, as corneocytes migrate to the skin surface they mature from a fragile to a resilient phenotype. Envelopes of fragile corneocytes can be seen on exposed body parts such as the face, especially in winter [13]. When moisturizers are used, however, the corneocytes may still mature to the resilient phenotype. It has been demonstrated through in vivo studies that moisturizers containing glycerin promote the maturation of these corneocytes, probably by activating the residual transglutaminase activity retained within the stratum corneum [13]. Skin cream can smooth the skin surface and increase the hydrophilic properties by helping the stratum corneum restore the lost moisture and the regular packing of the lipid lamellae, as well as repair the function and improve the feel of the skin [14].

For the studies of aesthetic repair and percutaneous absorption of cosmetics and drugs, pig skin has been used [15-17]. In studies of the mechanical [18-21] and tribological properties [6,22] of skin, and the percutaneous absorption of cosmetics and drugs [23,24], rat skin has also been used as an animal model. Pig skin and rat skin have been compared for percutaneous absorption [25,26], epidermal barrier layer lipids and morphology [27,28]. Table 1 shows some surface characteristics of virgin rat and pig skin. The stratum corneum and epidermis of pig skin are thicker than those of rat skin. The pig skin has fewer hair follicles than rat skin. Pig skin has been reported to be the most suitable model for human skin because of its similar surface properties, such as body mass and skin-to-body surface-area ratio, sparse hair, thick epidermis, hair-follicle density, epidermal turnover kinetics, lipid composition and the biophysical properties of the lipids [27,28], and similar permeability, i.e., the fluxes through the skin and concentrations in the skin are of the same order of magnitude for both tissues [25,26].

Understanding the differences between friction and durability of normal and damaged skin and the role of cream treatment of the damaged skin on friction and durability is of importance. Many

Table 1: Comparison of selected surface characteristics between virgin rat and pig skin samples.

	virgin rat skin	virgin pig skin
RMS (nm)	148 ± 6	274 ± 10
nanohardness (MPa)	7 ± 1	19 ± 3
elastic modulus (MPa)	70 ± 7	91 ± 28
stratum corneum thickness (μm)	5.0 ± 0.8	12.3 ± 0.7
epidermis thickness (μm) [27]	21.7 ± 2.2	51.9 ± 1.5
number of hair follicles (per cm^2) [27]	289 ± 21	11 ± 1

studies have focused on the macroscale friction and durability of normal and damaged skin, but the skin properties are related to the nanoscale structures, and therefore an understanding of the nanoscale friction and durability is necessary. In this research, friction and durability of virgin and damaged skin were measured on the nanoscale to study the differences between them. The skin treated with skin cream was compared to untreated skin to study the effect of skin cream. The effect of velocity, normal load, relative humidity and number of cycles on friction was studied. Experiments were also carried out on the macroscale to study the scale effects. Both rat skin and pig skin were used as skin models in this study as they are common models for human skin. Surface roughness and mechanical properties are known to affect friction and these have been measured. A low contact angle is desirable for the adsorption of skin cream and has also been measured.

Experimental

Sample preparation

There were four categories of skin samples used in the tests: virgin skin, treated virgin skin, damaged skin, and treated damaged skin, for both rat and pig skin. The method to prepare virgin skin samples was described by Tang and Bhushan [6]. After the animal was sacrificed, the dorsal skin was immediately excised, subcutaneous tissues were scraped off with scissors, and the hair was shaved carefully with a razor. Then, the skin was gently cleaned with a 10% (v/v) soap solution (liquid hand soap, Kroger Co., Cincinnati, OH), rinsed with tap water for 30 s, and leveled on the table to dry under ambient conditions (22 °C, RH 25–35%). After that, the skin was rinsed with a commercial facial cleanser treatment, Clean & Clear Shine Control facial cleanser (Johnson & Johnson, Skillman, NJ). The facial cleanser was applied evenly on the skin surface with a cotton swab. Skin was lathered for 30 s and rinsed with tap water for 60 s. Then, the skin was leveled on the table and dabbed with filter papers (Whatman International Ltd., Maidstone, England) to remove excess water. After that the skin was cut into 10 mm × 10 mm samples and attached to the AFM sample pucks with rapidly drying glue. This virgin skin was considered to be the reference group.

A dry (damaged) skin can be realized by repeated skin wash with harsh soaps/detergents containing sodium lauryl sulfate (SLS) or by sodium dodecyl sulfate (SDS) surfactant [9,29,30], or by a 20 minutes treatment of the skin with acetone/ether (1:1), which causes removal of skin lipids and induces a chapped and scaly appearance. Scanning electron microscope studies of SDS-treated stratum corneum revealed selective depletion of the lipids from the intercellular spaces accompanied by marked disruption of multiple lamellae structures, and lipid analysis showed a considerable and selective loss of intercellular lipids such as cholesterol, cholesterol ester, free fatty acid, and sphingolipids [9,29]. Another approach is to use sticky cellophane tape to remove upper layers of skin, which also results in skin damage and scaly appearance after one day [31,32]. Tape stripping has been reported to produce results similar to treatment with a surfactant of a 5% aqueous solution of SDS under an occlusive dressing for 4 h [31]. In this study, SDS was chosen to prepare damaged skin without any inflammatory reaction accompanied by a significant decrease in its water-retention function. To produce a controllably damaged skin sample, a 5% weight aqueous solution of SDS, prepared by adding 5 g SDS (Bio-Rad Laboratories, Hercules, CA) into 100 g demineralized distilled water, was applied to the virgin skin surface by rubbing with a cotton swab for 30 seconds, and the skin was allowed to dry for 10 minutes, and the process was repeated.

Table 2 shows the composition of the skin cream (Vaseline Intensive Care Lotion, Unilever, Trumbull, CT) used in this study. Virgin skin and damaged skin were treated with common skin cream, which was rubbed over the entire skin surface for 30 s with a cotton swab. For the nanoscale tests 0.2 mg of commercial skin cream was applied to obtain a 150 nm film thickness. On the macroscale, 2 mg was applied forming a film of 1.8 μm thickness [6]. The same methodology was used both on rat and pig skin.

Surface roughness and coefficient of friction measurements

Nanoscale surface roughness and coefficient of friction were measured by using a commercial AFM system (Dimension

Table 2: Composition of common skin cream used in the study (based on manufacturer information).

skin cream	composition
common skin cream	water, glycerin, stearic acid, *helianthus annuus* seed oil, glycine soja, lecithin, tocopheryl acetate, retinyl palmitate, urea, collagen amino acids, sodium stearoyl lactylate, sodium isostearoyl lactate, mineral oil, sodium PCA, potassium lactate, lactic acid, petrolatum, dimethicone, avena sativa, keratin, glyceryl stearate, cetyl alcohol, methyl palmitate, magnesium aluminum silicate, fragrance, carbomer, stearamide AMP, triethanol amine, corn oil, methylparaben, DMDM hydantoin, disodium EDTA, BHT, propylene glycol, titanium dioxide

Nanoscope IIIa, Veeco, Santa Barbara, CA) under ambient conditions. Fort A-20 tips (Si, N-type, 10 nm radius, spring constant of 3 N/m) (Appnano, Santa Clara, CA) were used. The coefficient of friction was calibrated by the method described by Bhushan [33]. The friction force measurements were made over a scan length of 10 μm and at a scan rate of 1 Hz at various increments of normal load ranging from 25 nN to 250 nN. By plotting the friction force as a function of the normal load, an average value of the coefficient of friction was obtained from the slope of the fitted line of the data.

The macroscale coefficient of friction was measured by using a POD reciprocating tribometer, with measurement techniques described in detail by Bhushan [34,35]. The tests were carried out in an ambient environment over a stroke length of 10 mm and at a velocity of 0.4 mm/s and at a normal load ranging from 20 mN to 60 mN, unless otherwise noted. A sapphire ball with a 1.5 mm radius and surface roughness of about 2 nm RMS was fixed in a stationary holder. The normal load and friction force were measured with the semiconductor strain gages mounted on a crossed-I-beam structure. By plotting the friction as a function of normal load, an average coefficient of friction was obtained from the slope of the fitted line of the data. For each property, a minimum of six measurements was made. The $\pm 1\sigma$ values were presented in the data plots.

Effect of velocity and normal-load measurements

To study the effect of velocity on nanoscale friction, experiments were carried out using the AFM by changing the scan frequency from 0.1 to 50 Hz while the scan size was maintained at 10 μm, which allowed a range of velocity from 2 to 1000 μm/s. For the effect of normal load experiments, the normal load was varied from 50 to 750 nN at a 10 μm scan length and a scanning velocity of 20 μm/s. The macroscale experiments were performed by using a POD reciprocating tribometer. To study the effect of velocity, the velocity was varied from 0.4 to 4 mm/s over a stroke length of 2.5 mm at the normal load of 50 mN. For the effect of normal load experiments, the normal load was varied from 10 to 50 mN over a stroke length of 2.5 mm at a velocity of 0.4 mm/s.

Effect of relative humidity measurements

A homemade humidity-control chamber system [36] was used to study the effect of relative humidity on the friction and durability of skin samples. A humidity detector was used to monitor the humidity inside AFM chamber. The relative humidity ranged from 4 to 95%. The skin sample was placed at each humidity value for about 1 h prior to the test.

Durability measurements

The durability measurements were carried out by repeated cycling tests. The nanoscale durability tests were carried out by using the AFM at a velocity 20 μm/s and at a normal load of 250 nN for 3600 cycles. The macroscale durability tests were performed by POD reciprocating tribometer with a velocity 1 mm/s and at a normal load of 50 mN for 3000 cycles.

Contact-angle measurements

The apparent contact angles were measured for various samples. Measurements were made with a Rame-Hart automated goniometer model 290-F4, where 5 μL water droplets were deposited onto the sample surface and the contact angle was measured.

Nanoindentation measurements

The nanoindentation measurements were made by using a Hysitron Triboscope (Hysitron Inc., Minneapolis, MN) in the constant displacement rate loading mode with a three-sided pyramidal diamond (Berkovich) tip. In this study, the maximum indentation displacement was controlled to be 1000 nm [6]. The method for the hardness (*H)* and the elastic modulus (*E)* determination was based on established methods [37,38]. Briefly, H was calculated from

$$H = \frac{P_{\max}}{A} \tag{1}$$

where $P_{\max}$ is the maximum imposed load, and A is the projected contact area. The relationship between the contact area and the contact depth was obtained from calibrating the tip with a standard material of known mechanical properties such that A is readily obtained from the load–displacement data.

E was analyzed according to the following equations:

$$E = \frac{\left(1-\nu^2\right)}{\frac{1}{E_r} - \frac{1-\nu_t{}^2}{E_t}} \quad (2)$$

where E_t and ν_t are the elastic modulus and the Poisson's ratio of the indenter tip respectively; ν is the Poisson's ratio of skin assumed to be 0.5 [6]; E_r is the reduced modulus given as follows:

$$E_r = \frac{\sqrt{\pi}}{2}\frac{S}{\sqrt{A}} \quad (3)$$

where S is the contact stiffness obtained from the slope of the unloading curve.

Results and Discussion

The nanoindentation properties are presented in the first section. Then the surface roughness, contact angle and nano- and macroscale friction data of rat skin are presented. Finally, data of pig skin are presented.

Nanoindentation properties of rat and pig skin

Mechanical properties of rat and pig skin were measured by using a nanoindenter. The load–displacement curves for rat and pig skin are presented in Figure 2a. Under the same displacement control, the load required for a given displacement for pig skin is larger than that for the rat skin, which means the pig skin is harder than rat skin. The nanohardness and elastic modulus data are presented in Figure 2b. Table 1 summarizes the mechanical properties data for virgin pig and rat skin. Both the nanohardness and elastic modulus of pig skin samples are higher than those of rat skin samples, and those of the damaged skin are higher than virgin skin for both rat and pig skin. The differences between the damaged skin and virgin skin for pig skin are greater than those for rat skin.

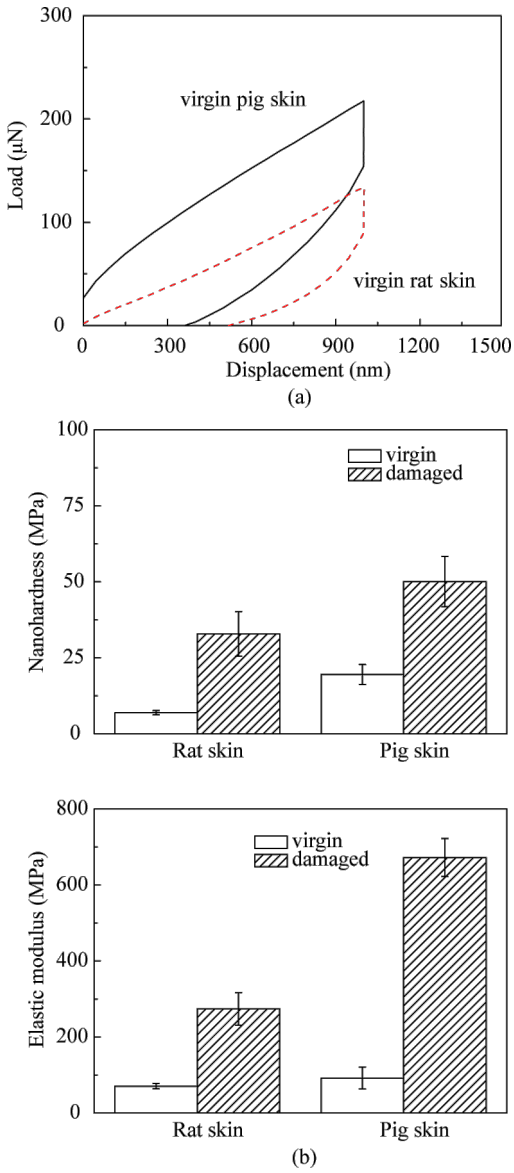

Figure 2: (a) Load–displacement curves and (b) nanohardness and elastic modulus data for rat and pig skin.

Surface roughness, contact angle and friction properties of rat skin

Surface roughness, contact angle and nanoscale friction

Figure 3 shows topography maps and corresponding height profiles of the cross section indicated by the arrows on a 20 μm × 20 μm scan size for virgin skin, damaged skin, cream-treated virgin skin and cream-treated damaged skin. The height profiles appear smoother for virgin skin compared with damaged skin, and for cream-treated skin compared with untreated skin. The RMS roughness data, which serve as quantified expressions of the surface characteristics, are shown in Figure 4a. The damaged skin has a higher roughness than virgin skin. After treatment with skin cream, the roughness of virgin skin and damaged skin decreased. The reasonable explanation is that the skin cream can fill the gap between the cells of stratum corneum.

The contact angle data for virgin skin and damaged skin are shown in Figure 4b. The contact angle of virgin skin is lower than damaged skin due to physical and chemical changes to the skin surface. An increasing surface roughness may be partially responsible for an increase in contact angle of the damaged skin

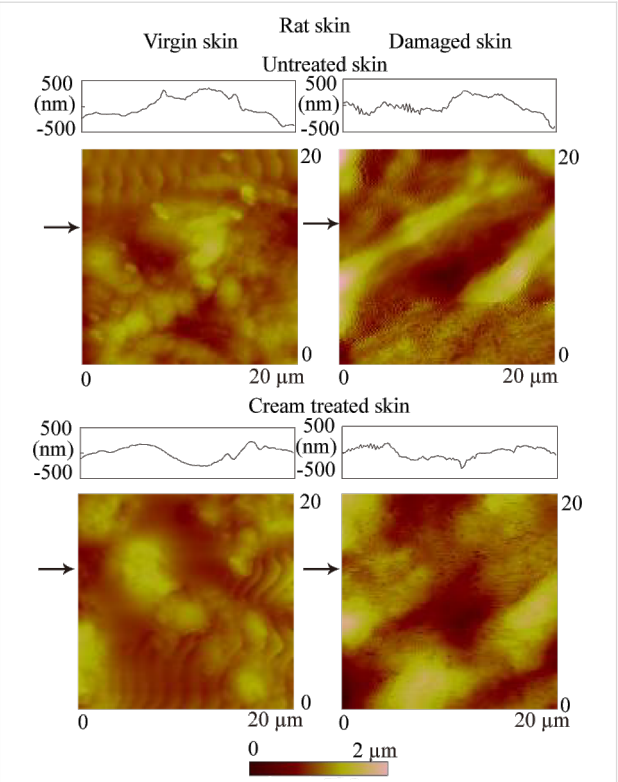

Figure 3: AFM topography maps and roughness profiles taken at arrows indicated for virgin rat skin, damaged skin, treated virgin skin and treated damaged skin.

[39]. After treatment with skin cream, the contact angles of virgin and damaged skin decreased. The hydrophilic groups in skin cream, such as hydroxyl group, amines group, and carboxyl group in the humectants, increase the surface hydrophilicity and lead to a lower contact angle.

Figure 5 shows curves of friction force as a function of normal load for virgin rat and pig skin. An average value of coefficient of friction was obtained from the slope of the fitted line of the data. The intercept on the horizontal axis of normal load is the adhesive force, which is dominated by the meniscus contribution. Figure 4c presents the coefficients of friction of various skin samples. The coefficient of friction of damaged skin is higher than that of virgin skin, and increases for both virgin and damaged skin after treatment. Schematics show various rat skin interfaces. Damage to skin results in greater surface roughness and shrinking of the stratum corneum cells due to water loss, which hence increases the number of asperities on the surface [35]. The natural lipids present also deplete. Cream treatment for both skin types increases friction. Liquid films (lipid and condensed water vapor) present on the skin surface reduce the interfacial shear strength leading to lower friction; however, a thicker film forms meniscus bridges at asperity contacts leading to higher friction [34,35,40]. Cream treatment moistens and

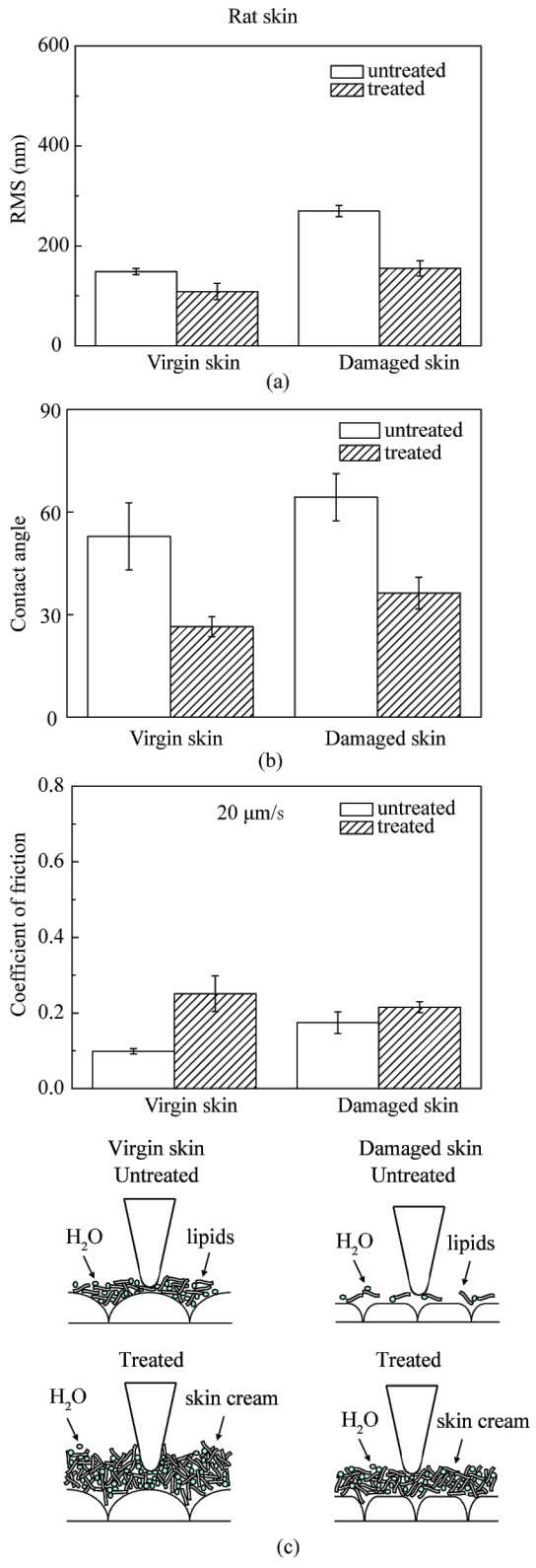

Figure 4: (a) RMS roughness, (b) contact angle and (c) coefficient of friction on nanoscale and schematic cartoons of the tip–skin-cream interaction of virgin rat skin, damaged skin, treated virgin skin and treated damaged skin.

softens the skin, which leads to a greater ductility and larger real area of contact. Larger contact area and formation of meniscus bridges are responsible for higher friction in cream-treated skin [6].

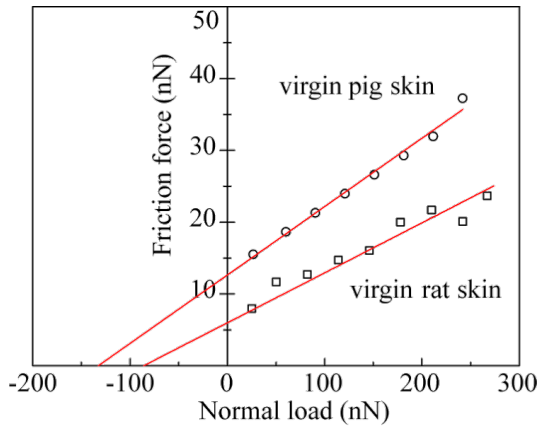

Figure 5: Friction force as a function of normal load curves for virgin rat and pig skin.

Effect of velocity, normal load, relative humidity and number of cycles on nanoscale friction

Figure 6a shows the coefficient of friction as a function of velocity for various skin samples. The data shows that friction decreases with an increase of velocity for all skin samples. At low velocity, the friction is dominated by meniscus force as proposed by Tang and Bhushan [6]. The tip sliding results in shearing and reformation of meniscus bridges. As the velocity increases, the meniscus bridges cannot be fully reformed, resulting in a drop in adhesive force and coefficient of friction. In the case of cream-treated skin, the skin cream is typically a shear-thinning fluid, and the viscosity decreases with the increasing shear rate leading to a decrease in the coefficient of friction [6,36].

Figure 6b shows the coefficient of friction as a function of normal load. The data shows that the friction for untreated skin samples first decreases then levels off, whereas, for the treated skin samples, it first decreased then increases above a certain load. As the tip moves towards the sample, a sudden mechanical instability occurs, and the tip jumps into contact with the film and a meniscus bridge is formed. But the tip does not slide in a steady manner on the surface at low normal load, and it may get rid of the meniscus bridges and bounce leading to a high deflection of the tip resulting in high friction data at the beginning. At higher load, the tip penetrates into film and slides in a steady manner, and the meniscus force dominates the friction. The coefficient of friction of treated skin samples increases above a certain load. It is believed that at larger load, the tip penetrates into the thick film and the formation of large meniscus bridges provides additional resistance responsible for the increasing friction [41,42].

Figure 6c shows the coefficient of friction as function of relative humidity. As the relative humidity increases, the coefficient of friction of all untreated and treated skin samples increases. As discussed earlier, the hydrophilic groups in the humectants of skin cream tend to form hydrogen bonds with water molecules, such that the humectants help the skin surface to attract water molecules in the environment, which increases the adhesive force leading to increasing coefficient of friction especially at high humidity [6]. Due to the hydrophobic lipid layer of virgin skin and some still present on the damaged skin, water hardly absorbs or penetrates into the skin surface, and humidity has less effect on it.

For durability studies, the friction experiments were performed by cycling the tip over the samples. Figure 7 shows the effect of the number of cycles on various skin samples. For untreated virgin and damaged skin, the coefficient of friction in the initial cycles is related to the removal of the thin lipid film on the skin surface, and then remains constant because the interaction between skin cream, skin surface, and environment reaches equilibrium. For cream-treated skin, the coefficient of friction decreases with the increasing number of cycles. This is believed to be caused by the change of cream film thickness. When cream is first applied to the skin surface, the cream cannot be absorbed immediately by the skin, and the cream liquid accumulates at the contact interface, resulting in a larger liquid height and greater viscous drag to motion. However, after several scans, because of the absorption of the skin cream and the evaporation of the water content, the cream film thickness decreases, which is responsible for the decrease in adhesive force and coefficient of friction. The skin cream finally covers the skin surface as a stable gel network (surfactant, fatty amphiphile, and water) and friction remains constant [6].

Macroscale friction and the effect of velocity, normal load and number of cycles

Macroscale data for coefficient of friction for various skin samples is shown in Figure 8. The coefficient of friction of damaged skin is comparable to virgin skin. For damaged skin, as discussed earlier, the levels of the fragile corneocytes generally increase, so the stratum corneum of damaged skin is torn rapidly at high loads in macroscale experiments forming a lubricant layer between the tip and the skin surface, which is more easily sheared, and may compensate the loss of the lipid layer. After the application of skin cream, the skin surface properties change, and the skin is moistened and softened by the skin

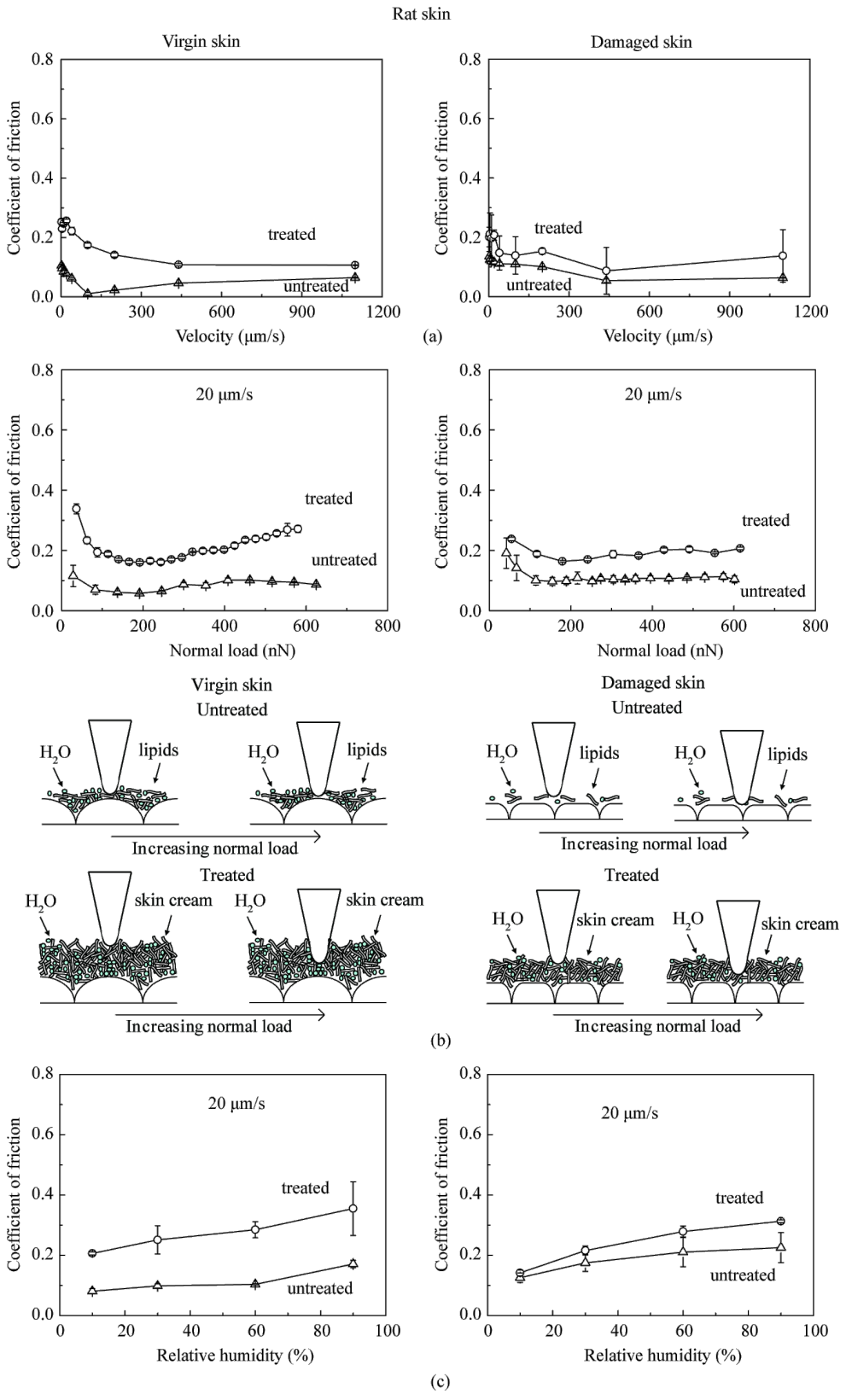

Figure 6: Effect of (a) velocity, (b) normal load, and schematic cartoons of tip–skin interaction, and (c) effect of relative humidity on the coefficient of friction on the nanoscale for virgin rat skin, damaged skin, treated virgin skin and treated damaged skin.

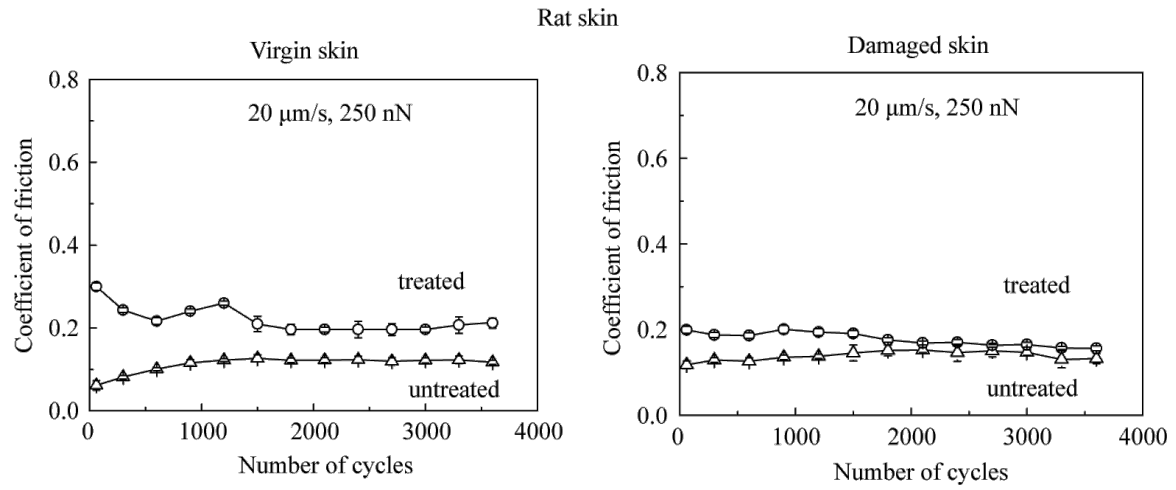

Figure 7: Effect of the number of cycles on the coefficient of friction on the nanoscale for virgin rat skin, damaged skin, treated virgin skin and treated damaged skin.

cream, which leads to a greater ductility and a larger real area of contact resulting in stronger adhesion, such that the coefficient of friction of cream-treated skin is higher than that of virgin skin [6].

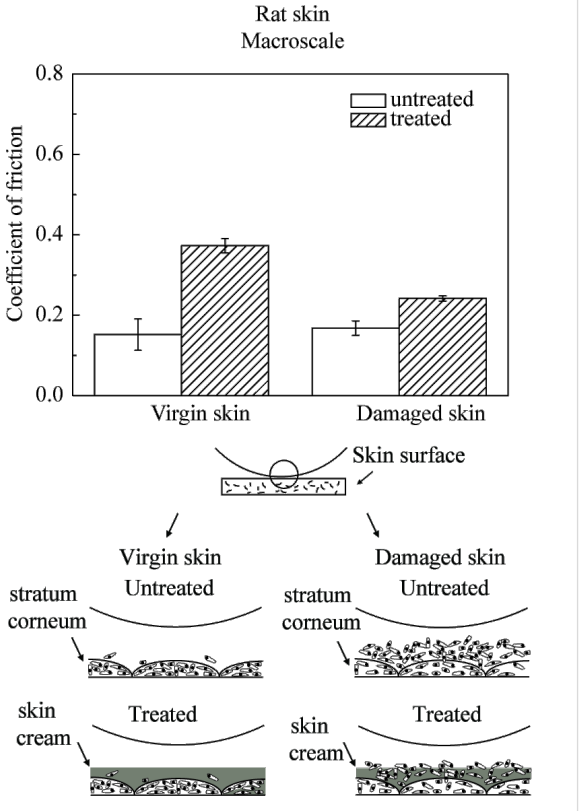

Figure 8: Coefficient of friction and schematic cartoons of the tip–skin interaction on macroscale for virgin rat skin, damaged skin, treated virgin skin and treated damaged skin.

Figure 9 presents the effect of velocity, normal load and number of cycles on the macroscale friction. Figure 9a shows that the coefficient of friction decreases as the velocity increases. The treated skin samples show a greater change than untreated skin samples. The reduction is similar to that on the nanoscale, since skin cream is a shear-thinning fluid as mentioned earlier.

Figure 9b shows that the coefficient of friction decreases as the normal load increases. Increased surface roughening and a large quantity of wear debris are believed to be responsible for the decrease of friction with an increase of normal load [40]. Asperity deformation of skin is primarily elastic, and as the normal load increases, elastic deformation at the asperities is large, such that the individual asperities on the contacting surface are totally deformed, and the contact region approximates to the contact of a large single asperity [35]. In this case, $\mu \propto W^{-1/3}$, and the coefficient of friction μ decreases with the increase of normal load W [6]. Figure 9c shows that the coefficient of friction remains almost constant on the macroscale for the four skin samples with the number of cycles, which suggests little damage during the cycling test.

Surface roughness, contact angle and friction properties of pig skin

Surface roughness, contact angle and nanoscale friction

Figure 10 shows AFM topography on a 20 μm × 20 μm scan size for virgin pig skin, cream-treated virgin skin, damaged skin and cream-treated damaged skin. Figure 11a shows that the damaged skin has a higher surface roughness than virgin skin, i.e., the same trend as for rat skin, but the difference between virgin and damaged pig skin is more distinct than that for rat

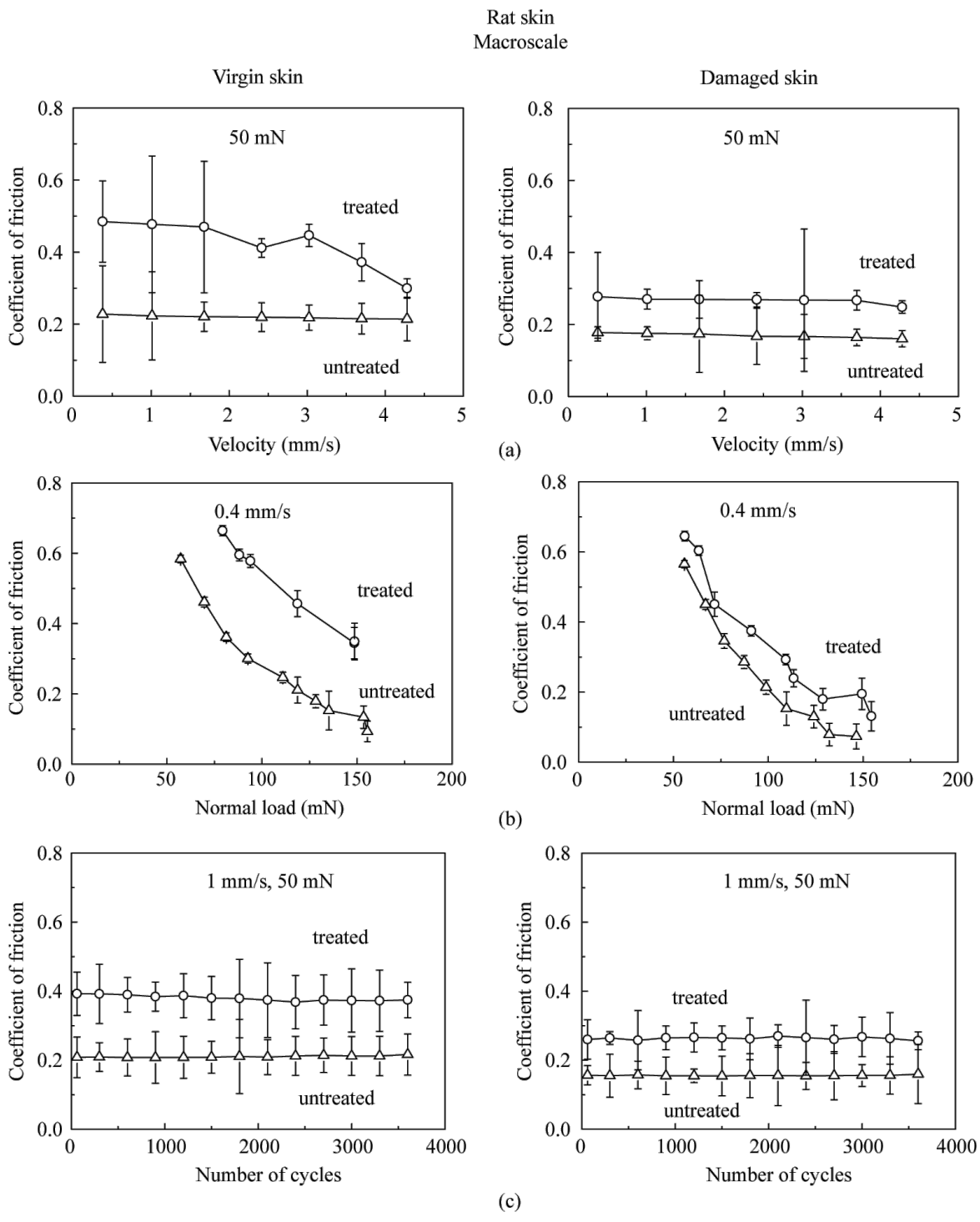

Figure 9: Effect of (a) velocity, (b) normal load, and (c) number of cycles on the coefficient of friction on macroscale for virgin rat skin, damaged skin, treated virgin skin and treated damaged skin.

skin. After treatment, the roughness of both virgin and damaged skin decreased. As shown in Figure 11b, the contact angle of damaged skin is higher and decreases after treatment with skin cream, as observed earlier for rat skin.

Figure 11c shows the coefficient of friction of various skin samples. The coefficient of friction of damaged skin is higher than virgin skin. After treatment, the coefficient of friction of virgin and damaged skin increases. The coefficient of friction of

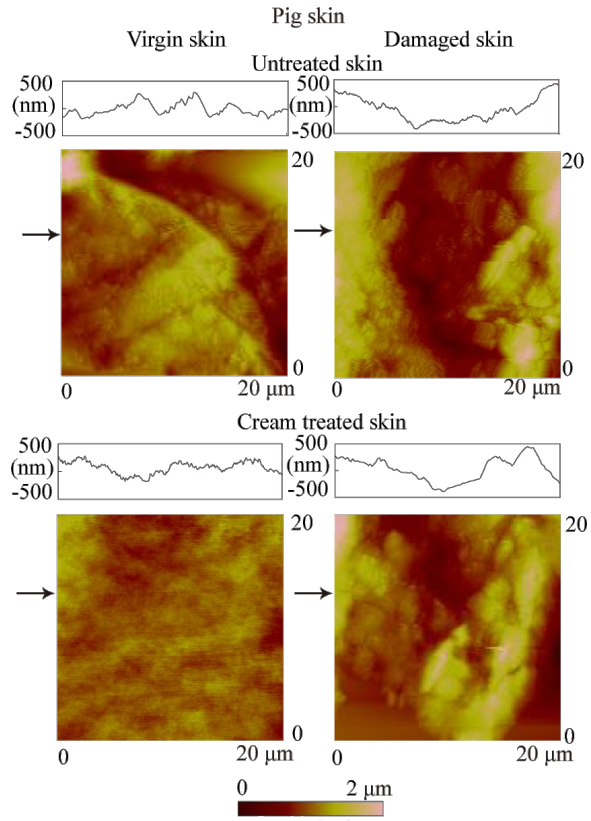

Figure 10: AFM topography maps and roughness profiles taken at arrows indicated for virgin pig skin, damaged skin, treated virgin skin and treated damaged skin.

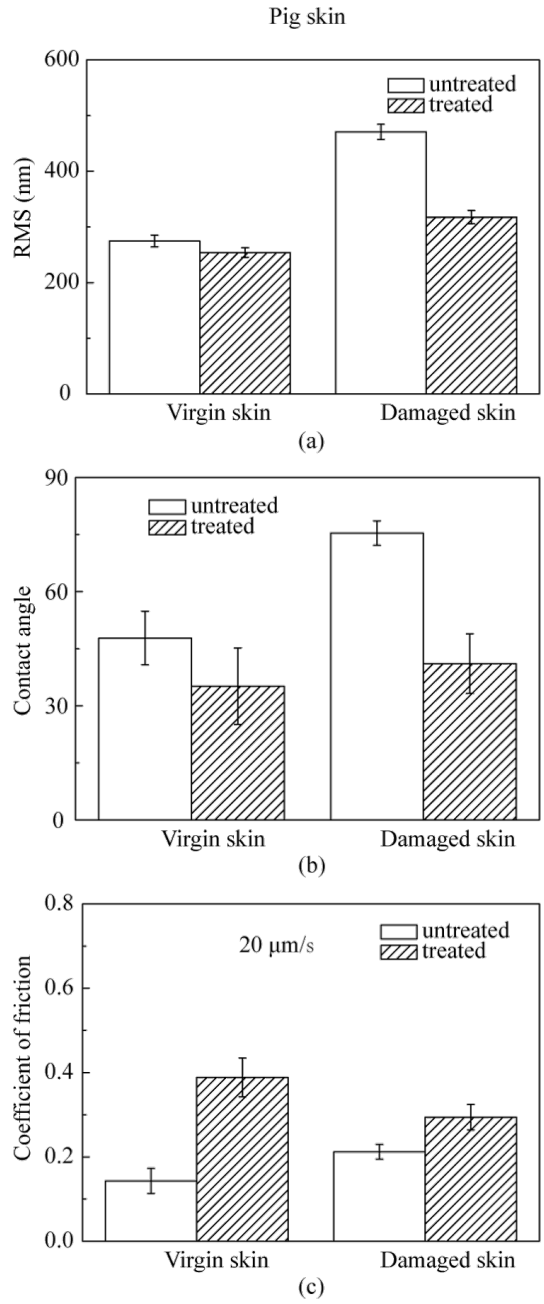

Figure 11: (a) RMS roughness, (b) contact angle, and (c) coefficient of friction on the nanoscale of virgin pig skin, damaged skin, treated virgin skin and treated damaged skin.

pig skin is higher than that of rat skin because of the different surface characteristics discussed earlier.

Effect of velocity, normal load, relative humidity and number of cycles on nanoscale friction

Figure 12 shows the effect of velocity, normal load and relative humidity on various skin samples. The coefficient of friction slightly decreases initially with an increase of velocity; the decrease is significant with an increase in normal load. It increases as the relative humidity increases. The trends are the same as those for rat skin.

Figure 13 shows the effect of the number of cycles on various skin samples. The coefficient of friction of treated pig skin samples shows a greater decrease than untreated pig skin samples. The reason is as discussed for rat skin.

Macroscale friction and effect of velocity, normal load effect and number of cycles

Figure 14 shows the coefficient of friction of the four pig-skin samples on the macroscale. Trends and the values of the coefficient of friction are similar to those of the rat skin. Figure 15 shows the effect of velocity, normal load, and number of cycles on the macroscale. The coefficient of friction does not show a significance change with the increasing velocity. The coefficient of friction decreases as the normal load increases. The coefficient of friction remains constant with an increase in the number of cycles. Again trends are similar to those for the rat skin.

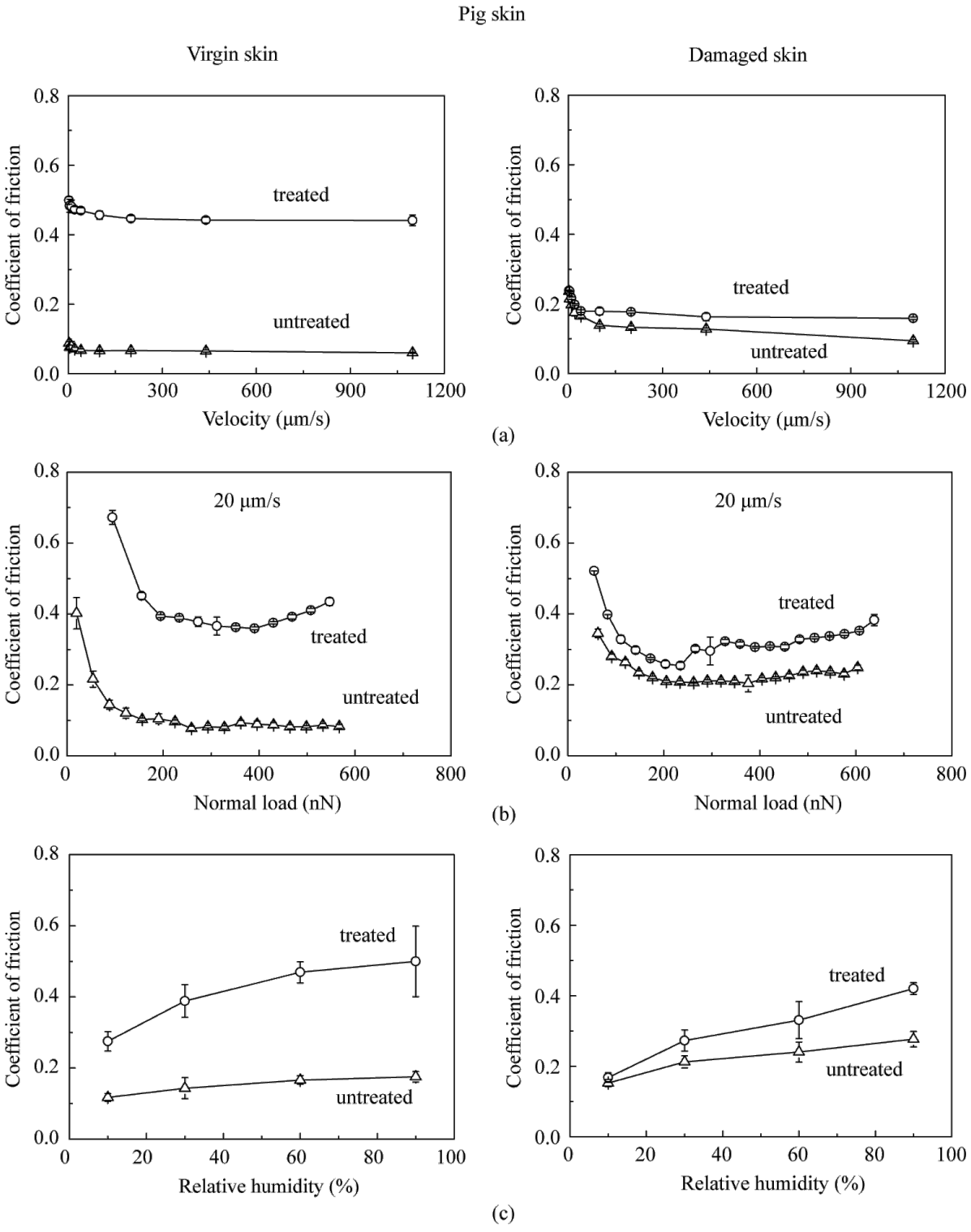

Figure 12: Effect of (a) velocity, (b) normal load and (c) relative humidity on the coefficient of friction on the nanoscale for virgin pig skin, damaged skin, treated virgin skin and treated damaged skin.

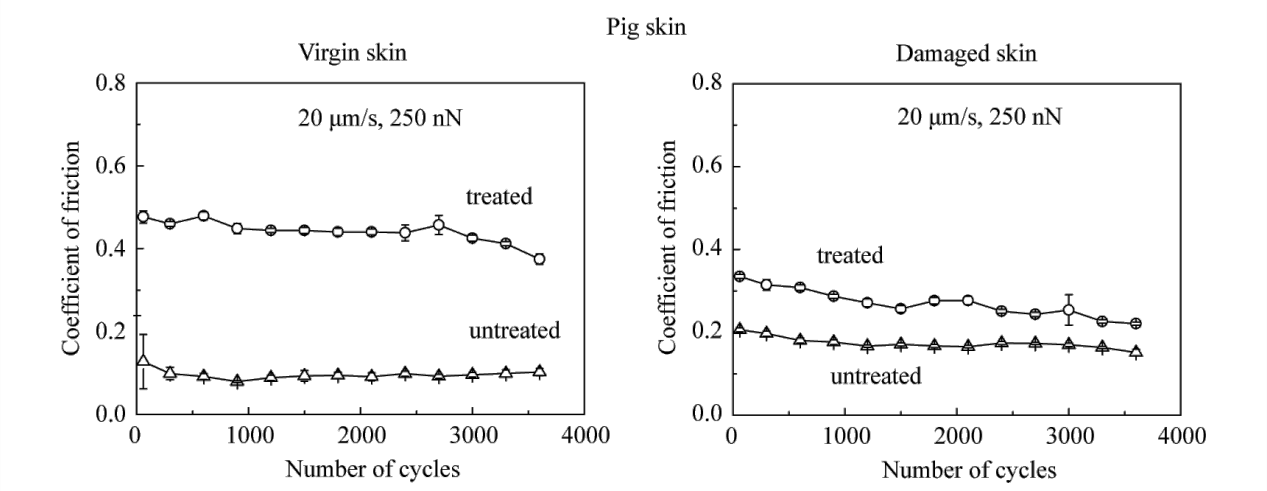

Figure 13: Effect of the number of cycles on the coefficient of friction on the nanoscale for virgin pig skin, damaged skin, treated virgin skin and treated damaged skin.

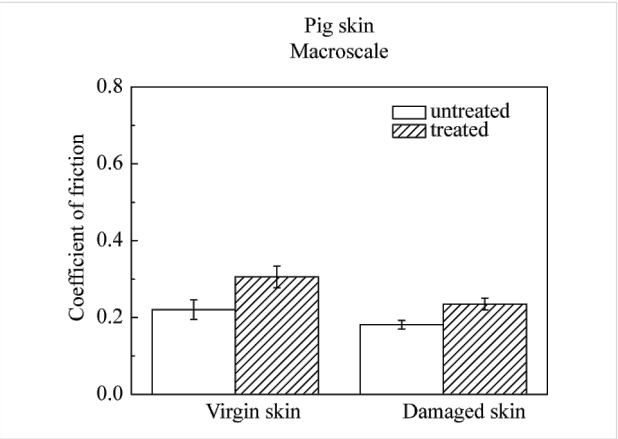

Figure 14: Coefficient of friction on macroscale for virgin pig skin, damaged skin, treated virgin skin and treated damaged skin.

Conclusion

In this study, rat skin and pig skin were used as experimental samples to study their friction properties and durability. The effect of velocity, normal load, and relative humidity on the coefficient of friction for virgin skin, cream-treated virgin skin, damaged skin and cream-treated damaged skin were studied on the nano- and macroscale. The durability of skin samples was also studied by repeated cycling tests.

The virgin rat skin has lower nanohardness and elastic modulus than that of virgin pig skin, and that of the damaged skin is higher than that of the virgin skin for both rat and pig skin. For rat skin, damaged skin has a larger roughness than virgin skin. After treatment with skin cream, the roughness decreases. The contact angle value of virgin skin is lower than damaged skin. The contact angle decreases after treatment due to an increasing hydrophilicity. Skin cream increases the hydrophilic properties of the skin. For pig skin, the roughness shows a similar trend to that of rat skin, but the contact angle of damaged skin shows a significant increase as compared to virgin skin because of the reduction in lipids present on the surface.

On the nanoscale, for both rat and pig skin samples, the coefficient of friction of damaged skin is larger than that of virgin skin. After treatment, the coefficient of friction increases because of meniscus formation. The effect of velocity on the nanoscale coefficient of friction of rat and pig skin samples shows the same trend. When the velocity increases, the coefficient of friction decreases. For untreated skin it is because the meniscus force cannot be fully reformed during sliding, but for treated skin it is because the viscosity decreases with the increasing shear rate. At the beginning as the normal load increases, the coefficient of friction decreases for rat and pig skin on the nanoscale, because the tip does not slide in a steady manner on the surface at low normal load, leading to a high deflection of the tip. After a certain value of the normal load, when the normal load further increases, the coefficient of friction of untreated skin samples remains constant, while that of treated skin samples shows a slightly increase due to an increasing meniscus force. As the relative humidity increases, the coefficient of friction of untreated rat and pig skin on the nanoscale does not increase much, because the water hardly absorbs or penetrates into skin surface because of a thin hydrophobic lipid layer. For the treated skin samples, the humectants help the skin surface to attract water molecules, which increase the adhesive force and coefficient of friction.

The coefficient of friction does not show a significant change in the durability tests. The coefficient of friction on the macroscale is larger than on the nanoscale. On macroscale, the coefficient

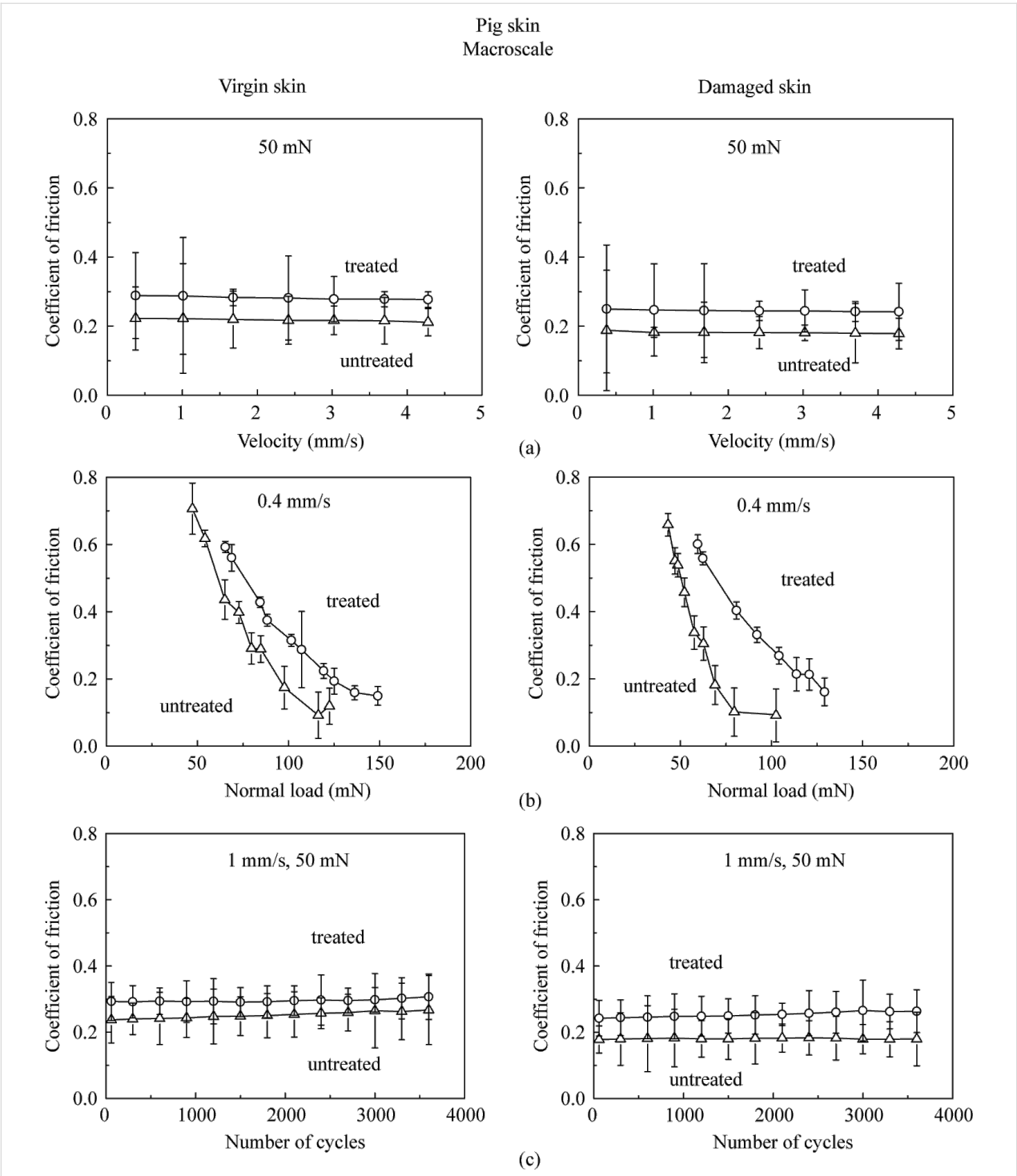

Figure 15: Effect of (a) velocity, (b) normal load and (c) number of cycles on the coefficient of friction on the macroscale for virgin pig skin, damaged skin, treated virgin skin and treated damaged skin.

of friction of damaged skin is comparable to that of virgin skin. After treatment, the coefficient of friction increases compared with untreated skin samples because of the formation of meniscus bridges. On macroscale, the velocity and number of cycles do not have an obvious effect on the coefficient of friction. When the normal load increases, the coefficient of friction decreases due to an increased surface roughening and a large quantity of wear debris. The coefficient of friction of pig skin is larger than that of rat skin on the nanoscale. The effect of velocity, normal load, and relative humidity on pig skin has the

same trend as that for rat skin both on the nano- and macroscale, as does the durability. The differences of friction properties between the four skin samples on pig skin are more distinct than those of the rat skin samples.

This research demonstrates that skin cream can smooth the skin surface and increase the hydrophilic properties of skin. The damaged skin surface condition can be improved by skin cream. The coefficient of friction of skin depends on the velocity, normal load, relative humidity and number of cycles.

Acknowledgements

S.C. acknowledges financial support from the Chinese Scholarship Council. The authors would like to thank Dr. Carrie Freed and Dr. McBride Feed for providing the skin samples. Suggestions from Dave Maharaj, Dalei Jing and Yunlu Pan of NLBB, and Dr. Wei Tang of CUMT were very helpful during the research.

References

1. Sutton, R. L. *The Skin: A handbook;* Doubleday: Garden City, NY, 1962.
2. Elsner, P.; Berardesca, E.; Maibach, H. I., Eds. *Bioengineering of the Skin: Water and the Stratum Corneum;* CRC Press: Boca Raton, 1994.
3. Pugliese, P. T., Ed. *Physiology of the Skin;* Allured Publishing: Carol Stream, IL, 1996.
4. Wilhelm, K. P.; Elsner, P.; Berardesca, E.; Maibach, H. I., Eds. *Bioengineering of the Skin: Skin Surface Imaging and Analysis;* CRC Press: Boca Raton, 1997.
5. Fuchs, E.; Raghavan, S. *Nat. Rev. Genet.* **2002,** *3,* 199–209. doi:10.1038/nrg758
6. Tang, W.; Bhushan, B. *Colloids Surf., B* **2010,** *76,* 1–15. doi:10.1016/j.colsurfb.2009.09.039
7. Lawson, L. B.; Freytag, L. C.; Clements, J. D. *Clin. Pharmacol. Ther.* **2007,** *82,* 641–643. doi:10.1038/sj.clpt.6100425
8. Downing, D. T.; Strauss, J. S.; Pochi, P. E. *J. Invest. Dermatol.* **1969,** *53,* 322–327. doi:10.1038/jid.1969.157
9. Imokawa, G.; Akasaki, S.; Minematsu, Y.; Kawai, M. *Arch. Dermatol. Res.* **1989,** *281,* 45–51. doi:10.1007/BF00424272
10. Rogers, J.; Harding, C.; Mayo, A.; Banks, J.; Rawlings, A. *Arch. Dermatol. Res.* **1996,** *288,* 765–770. doi:10.1007/BF02505294
11. Leyden, J. J.; Rawlings, A. V., Eds. *Skin Moisturization;* Marcel Dekker: New York, 2002.
12. Middleton, J. D.; Roberts, M. E. *J. Soc. Cosmet. Chem.* **1978,** *29,* 201–205.
13. Rawlings, A. V.; Canestrari, D. A.; Dobkowski, B. *Dermatol. Ther.* **2004,** *17,* 49–56. doi:10.1111/j.1396-0296.2004.04S1006.x
14. Bhushan, B. *J. Colloid Interface Sci.* **2012,** *367,* 1–33. doi:10.1016/j.jcis.2011.10.019
15. Pflucker, F.; Hohenberg, H. *Int. J. Cosmet. Sci.* **1999,** *21,* 399–411. doi:10.1023/A:1005450606035
16. Braye, F. M.; Stefani, A.; Venet, E.; Pieptu, D.; Tissot, E.; Damour, O. *B. J. Plast. Surg.* **2001,** *54,* 532–538. doi:10.1054/bjps.2001.3620
17. Yuan, Y.; Verma, R. *Colloids Surf., B* **2006,** *48,* 6–12. doi:10.1016/j.colsurfb.2005.12.013
18. Vogel, H. G.; Denkel, K. *Arch. Dermatol. Res.* **1985,** *277,* 484–488. doi:10.1007/BF00510067
19. Belkoff, S. M.; Haut, R. C. *J. Biomech.* **1991,** *24,* 711–720. doi:10.1016/0021-9290(91)90335-K
20. Kendall, M. A. F.; Chong, Y.-F.; Cock, A. *Biomaterials* **2007,** *28,* 4968–4977. doi:10.1016/j.biomaterials.2007.08.006
21. Bhushan, B.; Tang, W.; Ge, S. *J. Microsc. (Oxford, U. K.)* **2010,** *240,* 135–144. doi:10.1111/j.1365-2818.2010.03407.x
22. Tang, W.; Bhushan, B.; Ge, S. *J. Vac. Sci. Technol., A* **2010,** *28,* 1018–1028. doi:10.1116/1.3374341
23. Yourick, J. J.; Bronaugh, R. L. *J. Appl. Toxicol.* **1997,** *17,* 153–158. doi:10.1002/(SICI)1099-1263(199705)17:3<153::AID-JAT426>3.0.CO;2-E
24. Yourick, J. J.; Jung, C. T.; Bronaugh, R. L. *Toxicol. Appl. Pharmacol.* **2008,** *231,* 117–121. doi:10.1016/j.taap.2008.04.006
25. Bartek, M. J.; Labudde, J. A.; Maibach, H. I. *J. Invest. Dermatol.* **1972,** *58,* 114–123. doi:10.1111/1523-1747.ep12538909
26. Schmook, F. P.; Meingassner, J. G.; Billich, A. *Int. J. Pharm.* **2001,** *215,* 51–56. doi:10.1016/S0378-5173(00)00665-7
27. Monterio-Riviere, N. A. In *Marzulli and Maibach's dermatotoxicology,* 7th ed.; Zhai, H.; Wilhelm, K.-P.; Maibach, H. I., Eds.; CRC Press: Boca Raton, Florida, 2007; pp 39–50.
28. Stahl, J.; Niedorf, F.; Kietzmann, M. *Eur. J. Pharm. Biopharm.* **2009,** *72,* 310–316. doi:10.1016/j.ejpb.2008.09.013
29. Fulmer, A. W.; Kramer, G. J. *J. Invest. Dermatol.* **1986,** *86,* 598–602. doi:10.1111/1523-1747.ep12355351
30. Egawa, M.; Oguri, M.; Hirao, T.; Takahashi, M.; Miyakawa, M. *Skin Res. Technol.* **2002,** *8,* 41–51.
31. Denda, M.; Hori, J.; Koyama, J.; Yoshida, S.; Nanba, R.; Takahashi, M.; Horii, I.; Yamamoto, A. *Arch. Dermatol. Res.* **1992,** *284,* 363–367. doi:10.1007/BF00372041
32. Alikhan, S.; Maibach, H. I. Biology of Stratum Corneum: Tape Stripping and Protein Quantification. In *Textbook of Aging Skin;* Farage, M. A.; Miller, K. W.; Howard, I., Eds.; Springer: Berlin, Heidelberg, 2010; pp 401–408. doi:10.1007/978-3-540-89656-2_40
33. Bhushan, B., Ed. *Nanotribology and Nanomechanics,* 3rd ed.; Springer: Heidelberg, 2011.
34. Bhushan, B. *Principles and Applications of Tribology;* Wiley: New York, 1999.
35. Bhushan, B. *Introduction to Tribology;* Wiley: New York, 2002.
36. Liu, H.; Bhushan, B. *Ultramicroscopy* **2003,** *97,* 321–340. doi:10.1016/S0304-3991(03)00058-5
37. Oliver, W. C.; Pharr, G. M. *J. Mater. Res.* **1992,** *7,* 1564–1583. doi:10.1557/JMR.1992.1564
38. Bhushan, B.; Li, X. *Int. Mater. Rev.* **2003,** *48,* 125–164. doi:10.1179/095066003225010227
39. Bhushan, B. *Biomimetics: Bioinspired Hierarchical-Structured Surfaces for Green Science and Technology;* Springer-Verlag: Berlin, Heidelberg, 2012.
40. Bhushan, B. *Tribology and Mechanics of Magnetic Storage Devices,* 2nd ed.; Springer: New York, 1996.
41. Chen, N.; Bhushan, B. J. *J. Microsc. (Oxford, U. K.)* **2006,** *221,* 203–215. doi:10.1111/j.1365-2818.2006.01553.x
42. Lodge, R. A.; Bhushan, B. *J. Vac. Sci. Technol., A* **2006,** *24,* 1258–1269. doi:10.1116/1.2180275

Beilstein Journal of Nanotechnology

Large-scale analysis of high-speed atomic force microscopy data sets using adaptive image processing

Blake W. Erickson[1], Séverine Coquoz[1], Jonathan D. Adams[1], Daniel J. Burns[2] and Georg E. Fantner[*1]

Full Research Paper

Open Access

Address:
[1]Laboratory for Bio- and Nano-Instrumentation, École Polytechnique Fédérale de Lausanne, Batiment BM 3109 Station 17, 1015 Lausanne, Switzerland and [2]Mechatronics Laboratory, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States of America

Email:
Blake W. Erickson - blake.erickson@epfl.ch; Jonathan D. Adams - jonathan.adams@epfl.ch; Daniel J. Burns - danburns@mit.edu; Georg E. Fantner* - georg.fantner@epfl.ch

* Corresponding author

Keywords:
adaptive algorithm; artifact correction; atomic force microscopy; high-speed atomic force microscope; image processing

Beilstein J. Nanotechnol. **2012,** *3,* 747–758.
doi:10.3762/bjnano.3.84

Received: 26 April 2012
Accepted: 08 October 2012
Published: 13 November 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

Modern high-speed atomic force microscopes generate significant quantities of data in a short amount of time. Each image in the sequence has to be processed quickly and accurately in order to obtain a true representation of the sample and its changes over time. This paper presents an automated, adaptive algorithm for the required processing of AFM images. The algorithm adaptively corrects for both common one-dimensional distortions as well as the most common two-dimensional distortions. This method uses an iterative thresholded processing algorithm for rapid and accurate separation of background and surface topography. This separation prevents artificial bias from topographic features and ensures the best possible coherence between the different images in a sequence. This method is equally applicable to all channels of AFM data, and can process images in seconds.

Introduction

Atomic force microscopes (AFMs) are a useful tool for investigating nanoscale surfaces. They have applications in physics, materials science, chemistry, biology and nanotechnology. AFMs generate detailed three-dimensional images of surfaces with nanometer and subnanometer resolution [1-12]. The raw imaging data has to be post-processed to eliminate artifacts arising from distortions inherent in the technique or a specific instrument. Sample tilt, scanner bow and other artifacts corrupt the true topography. Removal of these artifacts is requisite to obtain the true topography of a sample [13-17]. In most AFM studies, the goal is to generate a single image of a sample surface or surfaces of multiple samples. The resulting data sets

are relatively small and easy to correct by hand. In an emerging part of the field, the goal is no longer to see just the detailed surface topography; but rather, to also see how this topography changes as a function of time or treatment [18-24]. High-speed AFM (HS-AFM) focuses on reducing imaging time to track faster dynamics than can be observed with existing instruments [25-29]. Where conventional AFM generates only a few images, HS-AFM generates hundreds of images. This amount of data creates a significant increase in the amount of image processing needed to extract the true sample topography, which can no longer be performed by the user on an image-by-image basis [15,17]. In HS-AFM, there is a real need for an automated processing routine that optimizes the image processing for each image without losing essential topographic information. One of the most important considerations in automating processing is that it must not introduce processing artifacts, and it must maintain coherence between images in a dataset to allow for rapid comparisons between frames [30,31]. Automated, AFM quality-control check in semiconductors and other nanotechnologies [32-34] also creates large data sets, which would also benefit from reliable automated image processing. In order to automate the image processing, it is important to define a metric by which the success of the image processing can be judged, and for which the processing is optimized. In the next sections we will evaluate the source and effect of typical AFM image distortions and derive a suitable optimization metric.

Sources of inherent distortions

All AFM images will contain artifacts in the *x*–*y*-direction as well as the *z*-direction. This work will address only the distortions found in the *z*-direction. There are two fundamental classes of image distortions in the *z*-direction, i.e., 2-D distortions and 1-D distortions. Both of these distortions are well-known and have been commented on since the earliest prototypes. 2-D distortions manifest over the whole image and deform 2-D surfaces into 3-D. The most common example of this is sample tilt. Tilting the sample normal relative to the scanner normal gives the entire image an apparent tilt. Piezo tube scanners (which are often used in AFM instruments) also generate inherent distortions in the image, creating an additional bow in the apparent topography [13]. Some of the early work in the field addresses the need to correct these background distortions appropriately by fitting only the background regions to a polynomial of the correct order [35]. 1-D distortions cause relative offsets from scan line to scan line. The sources of these distortions include laser-mode hopping and changes in the tip–sample interaction. These 1-D distortions cause apparent discontinuities in the topography of the sample, but do not represent an actual topographic feature. Most importantly, 1-D distortions generally appear as offsets in the *z*-data. Potential corrections for these types of artifacts have been addressed well in a direct fashion from starting principals by Starink and Jovin [35], and again in a less direct fashion that has some potential pitfalls with an emphasis on biological samples [36]. Figure 1 shows how these types of artifacts distort an artificial image along with corresponding histograms. In Figure 1A, half-spheres of arbitrary size were placed randomly within the image. A histogram of the *z*-heights in the image shows a spike at the background level, and a distribution of heights from the sphere (along the axis *x*-axis). Figure 1B shows the addition of a small amount of random noise added in the *z*-direction. This noise leads to broadening of the histogram peak. The peak now appears Gaussian with a small distribution.

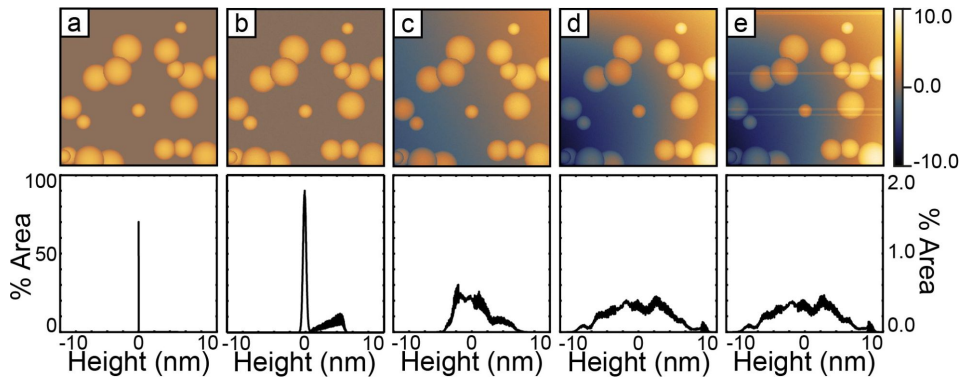

Figure 1: This figure shows the cumulative effects of typical distortions on model AFM data. Panels A though E show a sample with arbitrary topographic features (randomly placed and sized half spheres) and the corresponding histograms below. The first column shows the pure topography. The second column shows the addition of random noise. The third column shows the effects of an arbitrary sample tilt. The fourth column shows the effects of an additional second-order polynomial representing scanner bow. The fifth column adds arbitrary 1-D offsets in the fast-scan direction. Each artifact successively broadens the peaks in the histogram. The histogram in column A uses the vertical scale on the left. The histograms columns B through E share a common vertical scale shown on the right. All images share a common color scale shown on the right. All the histograms are plotted over this same range.

Figure 1C imparts a small tilt to the sample (less than 10 nm over 3 μm in either direction, which is very good for most samples). The background peak has been broadened so far as to be nearly indistinguishable. Figure 1D adds a second-order polynomial distortion to simulate scanner bow (less than 10 nm over the scan range, which is well within reason for most AFMs). The peak is no longer visible in the histogram. Finally, Figure 1E adds random vertical offsets to some of the scan lines. These offsets simulate the sorts of line skips commonly observed while imaging. As can be seen from the histograms, each successive distortion broadens the peak in the histogram. We will, therefore, test the suitability of calculating the distribution of the histogram as a metric for judging the success of automated-image-processing steps.

Basic requirements for successful image correction

The goal of the AFM image-processing algorithm described in this paper is to remove imaging artifacts and generate the best possible representation of the true sample topography. The achievement of this goal has one basic prerequisite, an area that can be used as a reference geometry. In most cases, this will be an area that can be considered as flat in the true topography of the sample. The image-processing algorithm then attempts to extract the global distortions in the image from observations of the flat reference region. These global distortions are subtracted from the complete image. In practice, this criterion is often met in AFM images, since the sample preparation generally utilizes flat substrates, such as mica or silicon wafers. In the case of semiconductor processing, the reference geometry could also be a specific area of the processed device that is flat or has a known shape. For the purpose of this work, we use the assumption that there is some part in the image that corresponds to a flat background, since this is the most common case. The algorithm automatically determines this background region and performs the necessary background correction. The algorithm uses minimization of the standard deviation of the image histogram to judge the success of the background detection and the image processing.

Image term description

Throughout the rest of this paper we will use the following conventions when referring to images for the sake of clarity. First, an image is defined as an array of scan lines along the slow scan axis,

$$\overleftrightarrow{\mathbf{I}} = \left(\mathbf{L}_1, \mathbf{L}_2, ..., \mathbf{L}_{\text{lines/image}}\right) \tag{1}$$

where each scan line is an array of pixels in the image along the fast scan axis.

$$\mathbf{L}_n = \left(\text{sample}_1, \text{sample}_2, ..., \text{sample}_{\text{samples/line}}\right) \tag{2}$$

Standard deviation as a metric for image flatness

In order to show that the standard deviation of an image is a suitable metric for monitoring the progress of image flattening we can use uncertainty propagation. We can start by describing the AFM image, $\overleftrightarrow{\mathbf{I}}$, as the sum of two height fields, the sample topography, $\overleftrightarrow{\mathbf{H}}$, and the inherent geometric distortions, $\overleftrightarrow{\mathbf{G}}$:

$$\overleftrightarrow{\mathbf{I}} = \overleftrightarrow{\mathbf{H}} + \overleftrightarrow{\mathbf{G}} \tag{3}$$

The square variance of the image can be written as:

$$\begin{aligned} \sigma^2_{\overleftrightarrow{\mathbf{I}}} &= \sigma^2_{\overleftrightarrow{\mathbf{H}}} + \sigma^2_{\overleftrightarrow{\mathbf{G}}} + 2\cdot\text{cov}\left(\overleftrightarrow{\mathbf{H}}, \overleftrightarrow{\mathbf{G}}\right) \\ \sigma^2_{\overleftrightarrow{\mathbf{I}}} &= \sigma^2_{\overleftrightarrow{\mathbf{H}}} + \sigma^2_{\overleftrightarrow{\mathbf{G}}} \end{aligned} \tag{4}$$

The equation simplifies to the final form because the covariance of the two fields is zero (they are uncorrelated variables for the types of distortions described in this paper). This means that $\sigma_{\overleftrightarrow{\mathbf{I}}} \geq \sigma_{\overleftrightarrow{\mathbf{H}}}$; thus, the geometric distortions will always broaden the distribution of observed heights relative to the true sample topography. This also means that as long as the processing steps only affect $\overleftrightarrow{\mathbf{G}}$ (flattening is only performed on truly flat regions, and not real topography), the minimization of $\sigma_{\overleftrightarrow{\mathbf{I}}}$ will improve the image.

Commonly used methods for correction of inherent distortions

The goal of AFM image processing is to correct the inherent distortions mentioned above and recover an accurate representation of the undistorted surface topography. A simple, often used, method is 1-D line fitting, which is an effective way to get a rough representation of the surface by removing much of the scanner tilt and bow at the expense of some induced artifacts. Methods for the subtraction of 2-D distortions, which add less artifacts than does 1-D line fitting, are 2-D polynomial fits. 2-D polynomial fits can be performed by using either the whole figure for the fit, or only certain regions of the image determined by using thresholds. Figure 2 shows a comparison of different processing methods on a standard calibration grating as well as on a lipid bilayer of mixed composition. Figure 2A and Figure 2E show the starting data. If present, the 1-D errors must be corrected first because any attempts to correct for the 2-D distortions will be biased by these 1-D offsets. To do this, the offset caused by the 1-D distortions is removed from each line. Higher order 1-D operations, such as line fits or higher order polynomial subtraction, should not be used to correct the

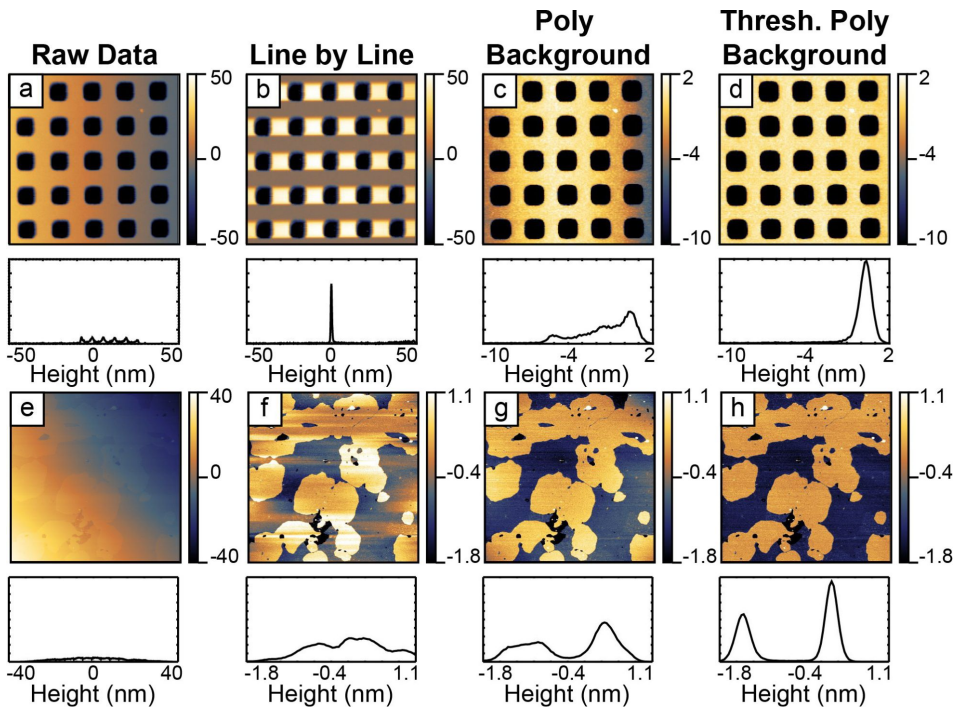

Figure 2: Panel A shows the raw data of a calibration grating. Panel B shows the result of line-by-line polynomial subtraction on the grating. Panel C shows global 2-D polynomial subtraction. Panel D shows a thresholded 2-D polynomial subtraction excluding the pits from the polynomial fit before subtraction. The histograms of each image are shown below. The vertical axis is the same for each histogram. In both the images, and the histograms, the scale has been focused on the upper portion of the grid in order to highlight the distortions in the image. Panel E shows the raw data of a mixed lipid bilayer on mica. Panel F shows the results of line-by-line second-order polynomial subtraction. Panel G shows 1-D artifact correction followed by 2-D second-order polynomial subtraction. Panel H shows the results of 1-D artifact correction followed by 2-D second-order polynomial subtraction by using the highest lipid domain to calculate the background for removal. The histograms of each image are shown below. The vertical axis is the same for each histogram. The median of each panel has been set to zero to allow visual comparisons. Panels D and H both show improvements in the representation of the surface in comparison to the other results. These improvements are shown clearly in the histograms. The grid was imaged under tapping mode in air. The lipid sample is the height data from QNM mode in fluid. The vertical scales of each image are in nanometers.

inherent 2-D distortions because they will destroy interline relationships in the data, and can generate false artifacts between lines [13,31,37], see Figure 2B and Figure 2F which show the results of 1-D second-order polynomial removal. This line-by-line polynomial subtraction generates many artifacts in the data, for example the surface surrounding the pits appears raised along the fast scan axis in Figure 2B, and continuous levels in Figure 2F have offsets from line to line and are not perpendicular to the image plane. Figure 2C and Figure 2G show the removal of 1-D offsets followed by a second-order background polynomial removal. For the grid, there is a significant improvement in the representation of the surface. In Figure 2F, the sample topography is of the same order as the 2-D distortions. While 2-D operations are less prone to induce artifacts, performing a global 2-D polynomial fit and background subtraction leaves significant residual distortions (Figure 2G). These distortions can be avoided by using a thresholded flattening, instead of a global flattening Figure 2D and Figure 2H (details discussed in Section Results and Discussion, "Algorithm description"). We conclude that a suitable way to process the images is to first determine and subtract the line-by-line offsets. Second, fit only the part of the image that is *the flat background* with a 2-D polynomial. Finally, subtract the calculated 2-D polynomial from the entire image. For an automated algorithm, the problem reduces to accurately distinguishing *the flat background* from the sample topography. In the rest of this paper, we describe a method for iteratively determining both *the flat background*, and the line-by-line offsets. Once these quantities are known, subtracting the line offsets and correcting the 2-D distortions can be performed with only two image-processing steps on the raw data.

Results and Discussion

1 Algorithm description

Figure 3 shows a general flow diagram of the processing algorithm from data import to final output. The algorithm has three major blocks: (a) **identify the background** region, **generate a mask** of the background region and **estimate the polynomial background**; (b) **determine 1-D background offsets** from the raw data within the mask; and (c) **subtract 1-D offsets** from the raw data followed by a single **masked background flattening**. This final step ensures that a minimal number of modifications

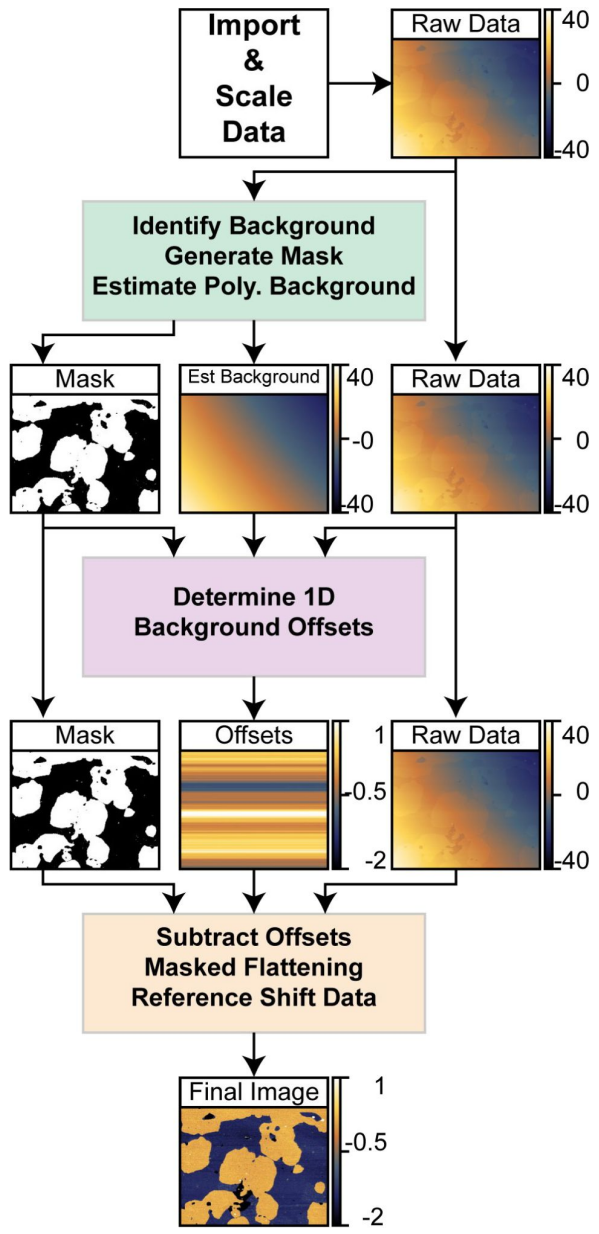

Figure 3: This figure outlines how an image is corrected automatically. There are three main operational blocks used in the process. Each block starts with the raw data as input and generates unique outputs. The first major block applies a series of 1-D offset corrections and thresholded polynomial background subtractions to generate a masked image of the region to be flattened as well as an estimate of the polynomial background. The second block subtracts this estimated background from the raw data and runs thorough a series of median corrections and flattening steps only on the data within the masks. This sequence determines the unique 1-D offset for each line. The third block subtracts these 1-D offsets from the raw data and performs a single polynomial fitting on the data within the mask. This background is subtracted from the entire image (raw-offsets). Finally, the height of the background is shifted to a defined reference value to ensure continuity within the image sequence. All vertical scales in the figure are in nanometers.

are performed on the raw data to generate the final output. The final output is offset to consistent values and exported for later use. The results of the automated processing routine on an example lipid bilayer of mixed composition are shown as the inputs and outputs of each major block. Each block will be discussed separately in detail.

1.1 Identify the background, generate a mask, estimate the polynomial background

The purpose of this section is to identify the background region in the image and estimate the initial polynomial background. In order to accomplish this, the algorithm does the following: (a) roughly corrects 1-D offsets present in the image (scars and offsets); (b) monitors the distribution of height values in the image to track the progress of each step, and rejects steps that broaden the distribution; (c) intelligently removes global tilt and scanner bow (often greater than sample topography); (d) adaptively identifies the background; (e) fits 2-D polynomials to the identified background; and (f) removes these fitted polynomials from the image.

1.1.1 Scar identification and median correction: The first process is to import the data from the binary format and convert it into relevant units. The standard deviation of the image is calculated as a reference point. Figure 4A shows the data as captured. Sample tilt dominates the image.

A basic scar identification and correction step is used to minimize the influence of the scars on the correction of 1-D artifacts. Scars are identified by detecting sections within a line $\mathbf{L}_n$ that are greater than surrounding lines by more than two times the standard deviation of the difference in medians. (Equation 5). Then each data point identified within a scar is replaced with the median of five neighboring data points in the last line before the scar:

$$\text{scar threshold} = 2 \cdot stdev\left(diff\left(median\left(\mathbf{L}\right)\right)\right) \qquad (5)$$

diff($\mathbf{X}$) calculates differences between adjacent elements of $\mathbf{X}$. If $\mathbf{X}$ is an array, then *diff*($\mathbf{X}$) returns an array, one element shorter than $\mathbf{X}$, of differences between adjacent elements: $[\mathbf{X}_2 - \mathbf{X}_1, \mathbf{X}_3 - \mathbf{X}_2, \ldots, \mathbf{X}_n - \mathbf{X}_{n-1}]$ of all lines in the image.

1-D artifacts, which introduce relative offsets, are corrected for by using *median corrections*. First, a line-by-line median-offset correction is applied to all lines. For each scan line, $\mathbf{L}_n$, the median of the scan line is subtracted from all the values in that scan line (Equation 6).

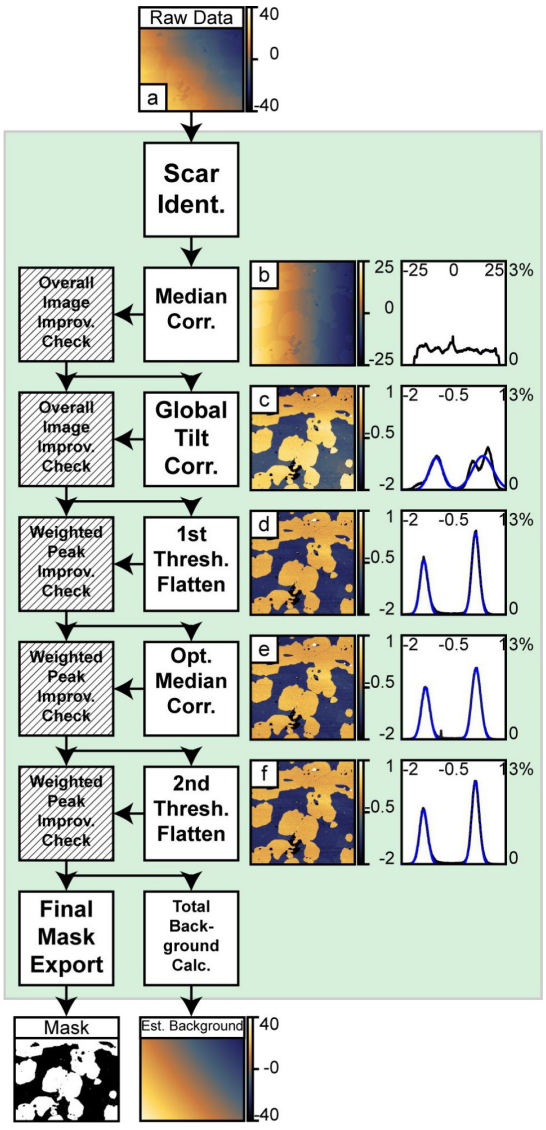

Figure 4: This figure shows the structure of the processing algorithms for identifying the region to be flattened and estimating the background. An initial scar detection followed by median correction and a median-difference correction is applied. This corrects for line skips in the source image. Following the 1-D artifact correction, a global correction is applied. The primary function is to remove overall sample tilt and some second-order distortions. The image is then adaptively thresholded and a 2-D polynomial is subtracted. After this step there is an optional second median difference correction to correct discontinuities missed in the previous median-difference correction. A second round of thresholded processing corrects for residual image distortions. After each step that manipulates data, i.e., the first median correction through to the final thresholded processing, a weighted improvement check ensures that the step had improved the image. In the case when a step degrades the image, the algorithm passes the previous result. The final mask generated after the last successful step as well as the sum background of all successful flattening steps is passed on to the next block of the algorithm. The right half of the figure shows the corresponding images and histograms for each step. Gaussian fits to the two distinct levels are shown with dotted blue curves. The AFM data in this figure was collected in QNM mode in fluid. The vertical scale in all images and the horizontal scale in all histograms is in nanometers.

$$\mathbf{L}_n = \mathbf{L}_n - median\left(\mathbf{L}_n\right) \tag{6}$$

Second, a weighted median-difference correction corrects artifacts from the median correction caused by differences in topography in the image. This process is only performed on lines not previously identified as scars. The median difference correction calculates the median of the difference between two consecutive scan lines excluding outliers greater than two times the standard deviation of the differences between scan lines. This outlier exclusion reduces influence from edges within a scan line. The median difference value is subtracted from all the values within the second scan line, see Equation 7.

$$\begin{aligned}\mathbf{C}_n &= \mathbf{L}_n - \mathbf{L}_{n-1}\\ \mathbf{C}'_n &= \left\{\forall E \in \mathbf{C}_n,\right.\\ &\quad median\left(\mathbf{C}_n\right) - 2\cdot stdev\left(\mathbf{C}_n\right) \le E\\ &\quad \left.\le median\left(\mathbf{C}_n\right) + 2\cdot stdev\left(\mathbf{C}_n\right)\right\}\\ \mathbf{L}_n &= \mathbf{L}_n - median\left(\mathbf{C}'_n\right)\end{aligned} \tag{7}$$

Third, an additional iterative refinement minimizes residual offsets generated from changes in topography. This subtracts a running average of the median difference between the scan line and up to the preceding k scan lines within a defined threshold. This does two things: one, it further excludes edge effects along topographic contours; and, two, it maintains continuity over long distances in the image and prevents the build up of high-frequency errors from scan line to scan line. This process runs iteratively in alternating directions. Currently, the algorithm uses four iterations with a maximum span, k, equal to 15 lines.

$$\begin{aligned}\mathbf{D}_n &= \left(\frac{1}{k}\right)\sum_{i=n-k,i\ge 1}^{n-1} \mathbf{L}_n - \mathbf{L}_i\\ \mathbf{D}'_n &= \left\{\forall E \in \mathbf{D}_n,\right.\\ &\quad median\left(\mathbf{D}_n\right) - 2\cdot stdev\left(\mathbf{D}_n\right) \le E\\ &\quad \left.\le median\left(\mathbf{D}_n\right) + 2\cdot stdev\left(\mathbf{D}_n\right)\right\}\\ \mathbf{L}_n &= \mathbf{L}_n - median\left(\mathbf{D}'_n\right)\end{aligned} \tag{8}$$

These steps handle discontinuities in the original image that do not represent actual topography. In cases where there is significant topographic variation parallel to the fast scan axis, i.e., a calibration structure, these corrections can actually introduce significant artifacts into the data and should not be used. In these cases, these 1-D corrections can be turned off in the algorithm. In Figure 4B, the initial median correction improves the

image by removing both the 1-D artifacts as well as most of the tilt in the vertical direction. It is important to note, that in all these operations, there is always only a constant offset subtracted from all points in a line, and not a linear fit or higher polynomial. Such higher-order 1-D corrections can create artifacts that cannot be undone at a later stage.

1.1.2 2-D Tilt removal: The third process is a basic polynomial surface fit of either first or second order and subsequent removal to facilitate the thresholding process by removing enough distortion to make it possible to separate topographic features from the background. Primarily, this removes sample tilt, which has a significant contribution to the variation in the observed heights. In cases of very flat samples where the changes of the topography are on the same order or smaller than the sample tilt or scanner bow, removal of a second-order surface generally improves the separation of the topography from the background. The algorithm does a separate first-order and second-order removal process and calculates the standard deviation of both results. The standard deviations are compared to the standard deviation after the median correction. The image with the lowest standard deviation progresses to the next step. Figure 4C shows the results of the first, unthresholded background correction. A second-order polynomial removal improved the image the most. After this step, the two distinct levels in the system manifest in both the image and the corresponding histogram. Some residual tilt remains in the image, as the bottom is somewhat higher than the top.

1.1.3 Thresholded flattening: The fourth process is an adaptive image thresholding algorithm and background removal. In order to calculate the appropriate threshold for the 2-D background removal, the algorithm identifies discrete peaks in the histogram corresponding to different topographic levels in the image. In the absence of discrete levels, only one general peak is used. The histogram is fit to a summation of $m \leq 8$ Gaussian functions centered at each peak, z_i, where i ranges from 1 to m. In the case that the number of levels is greater than m, only m peaks with the greatest area will be fit. From the Gaussian fit, the standard deviations, σ_i, and areas covered, A_i, are calculated. Nominally, the threshold for the 2-D surface fitting is $z_j \pm n\sigma_j$, where n is a user-defined range and j is the peak used for the 2-D surface fit. (The algorithm default is to use the highest peak found in the histogram. Based on prior knowledge of the sample, the user can override this default and specify which peak to fit.) Since this span can include contributions from other peaks, the actual span can be narrowed to the intersection of the jth peak with the next closest peak. All the data within the threshold is fit to a polynomial surface of a given order (up to fifth). The extrapolated background is subtracted from the whole image.

1.1.4 Improvement check: In order to test if a particular step has improved the image, the peak detection and Gaussian fit is repeated on a histogram of the corrected data. While the thresholded flattening only operates on a single level, the improvement check monitors the progress of all the fit peaks. This allows the algorithm to follow the progress of a sample with many levels. The criteria for this conditional check are:

1. The total area of the fits must not decrease by more than a small tolerance with respect to the image before the processing step.
2. A weighted average, $\sum_i A_i\sigma_i / \sum_i A_i$, of the standard deviations is computed both before and after. This term must not increase by more than a small threshold.

If either of these conditions fails, the algorithm rejects the process and passes the previous image to the next step. In the case of the initial background removal the whole image is used. After the initial background removal, only regions identified as independent levels are considered. The user defines the number of levels present in the system. In the case of the lipids or grids, both independent levels are considered; whereas, in the case of a random sample on a substrate, only the substrate would be considered. Figure 4D shows the results of the first thresholded processing. This results in a significant improvement of the overall image flatness, shown in both the image and histogram. The algorithm based the thresholding on the higher level since it had the greatest area.

1.2 Determine 1-D background offsets

The purpose of this section is to accurately determine the 1-D offsets that exist from line to line. While the operations used above in Section 1.1 do a reasonable job, they are still prone to influence from sample topography. In order to minimize the influence of topography the algorithm performs its corrections only on the masked background region determined in Section 1.1. First, the estimated background from Section 1.1 is subtracted from the raw data in order to generate a reasonable starting point for the first masked median correction (Figure 5A). The masked median correction subtracts the median of all the data within the mask on a given line, from that line (Figure 5B).

$$\mathbf{L}_n = \mathbf{L}_n - median\left(\mathbf{L}_n \cap \mathbf{mask}_n\right) \tag{9}$$

The second step is an additional polynomial flattening of the masked data in order to ensure a flat background (Figure 5C). This step is followed by an additional masked median correction (Figure 5D). Following each step, a check is performed to ensure that the step decreased the standard deviation of the data

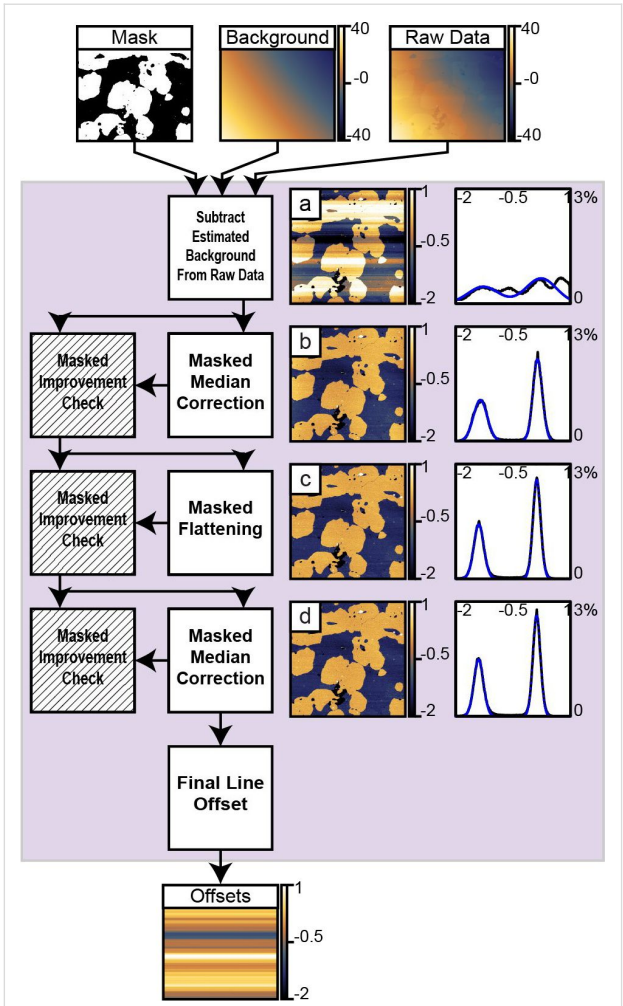

Figure 5: This figure shows the structure used to determine the 1-D offsets present in the image. First, the estimated background from the prior block is subtracted from the raw image. Second, a median correction is performed on the data within the mask. Third, an additional polynomial flattening on the masked region reduces the influence from the topography. Fourth, a second masked median correction refines the calculated offsets. After each step, except the initial background subtraction, an improvement check on the masked region verifies that the step improved the image. The final offsets are determined and passed through to the next section. The vertical scale of all images and horizontal scales of all histograms are in nanometers.

within the mask. Finally, the cumulative offsets are calculated for each line. In the case where a line has no pixels contained in the mask, the algorithm calculates the offset from the final output of Section 1.1 instead. These final outputs are passed on to the next section.

1.3 Subtract offsets, masked flattening, reference shift data

This final section of the algorithm subtracts the offsets calculated in Section 1.2 from the raw data (Figure 6A). Next, the data within the mask calculated in Section 1.1 is fit to a single 2-D polynomial. The resulting polynomial is subtracted from all

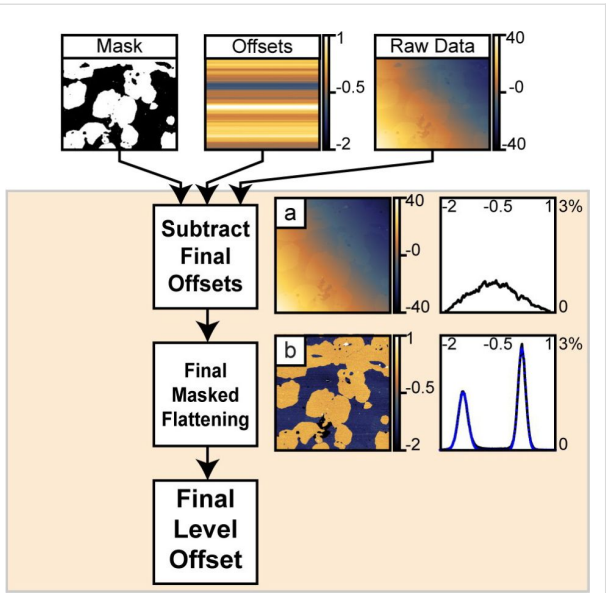

Figure 6: This figure shows the steps used for the final image correction. First, the 1-D offsets are subtracted from the raw image. Second, a final masked polynomial flattening is performed. Third, a reference offset is added to all data in the image prior to export. The vertical scale of all images and horizontal scales of all histograms are in nanometers.

the data (Figure 6B). A final histogram is computed and fit with Gaussians. The difference between the defined background reference height and the fitted center of the background peak in the final histogram is added to all the data. This final data is exported for future use.

2 Processing high-speed dynamic images in biological systems

2.1 HS-specific artifacts

Operating AFMs at high speeds has the potential to generate new distortions not normally observed at conventional speeds. Such artifacts include structural scanner resonances coupling into the sample topography. These sorts of spurious resonances can be observed in piezo tube scanners as a turn-around ripple in the images (damped oscillations of a given frequency). This stems from excitations in the lateral direction coupling into the vertical motion. Because the scanner used in these experiments has a relatively low structural resonance, this turn-around can be excited to a substantial extent at even moderate scan speeds. Given that the amplitude of the turnaround ripple in tube scanners like the one used in these experiments can easily be greater than the topography of the system, the turn-around ripple must be dealt with in some fashion. The best way is to avoid it altogether through either input shaping [38-41] of the drive signals or through electrical damping of the resonances [42]. For our experiments, we use a self optimizing method that determines the scanner resonances and compensates them with an input

shaper [38]. Using this resonance-compensator system, the amplitude of the turn-around ripple in the image can be nearly eliminated in a fashion that is sample independent and completely transparent to the rest of the experiment. While this method works very well, some residual, subnanometer distortions can remain.

2.2 Vertical median correction

While model-based filters do a very good job of reducing the amplitude of the turn-around ripple, they may not completely eliminate it. This is especially problematic on very flat samples with nanometer scale topography, where a subnanometer distortion is a significant part of the overall topography, such as in the images of lipid membranes shown in Figure 7. This residual turn-around ripple can be seen on the right-hand side of the images in the middle row. The turn-around ripple appears on the right side of the image because retrace images are being shown. This residual error can be corrected by using an additional vertical median correction on the masked data after the final background subtraction in Section 1.3 and before the final offset. This vertical median correction is exactly the same as Equation 9, but it runs in the perpendicular direction. This correction is only applicable to HS-AFM data that exhibits some turn-around ripple and should not be used otherwise.

Figure 7 demonstrates the utility of the automated processing algorithm on a dynamic model biological system. The sample consists of two phase-segregated lipids, DLPC (blue) and DPPC (gold). The image sequence shows the rapid degradation of the DPPC domains following rapid heating from interactions with the AFM tip. The sequence shows every third image. The upper sequence shows the results of standard unthresholded flattening and the lower sequence shows the results of the fully automated processing. Both sequences use the same polynomial background orders. The fully automated processing does a much better job of identifying the background and flattening it correctly. This makes it much easier to follow the changes in

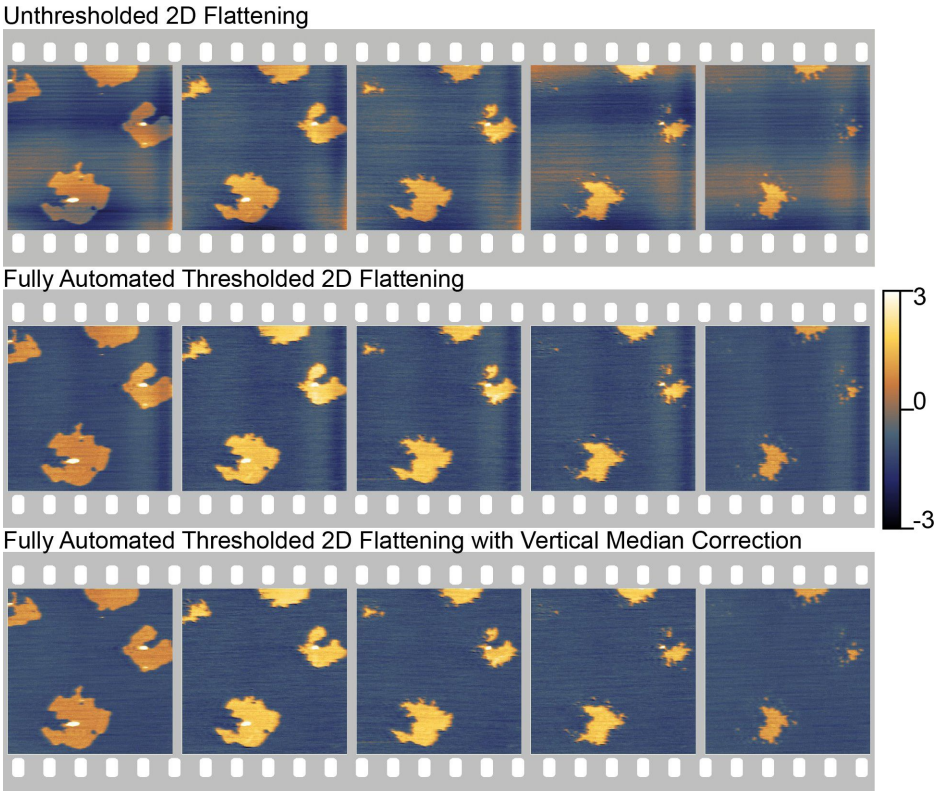

Figure 7: Comparison of unthresholded 2-D background removal to the full, iterative thresholded background removal for a sequence of HS-AFM data. The upper row shows the unthresholded result. The fully automated processing using the lower level for the background calculation is shown in the middle row. In the bottom row, an additional vertical median correction has been used. This corrects some of the residual artifacts from the turn-around ripple of the scanner, which are not corrected with the resonance compensator described by Burns et al. [38]. Each image is a 4 μm image with 256 lines captured at 78 lines/second while tapping in fluid. These imaging parameters give an approximate acquisition of 3 seconds per image. Each frame in this figure represents every third image in the overall sequence, approximately 9 seconds apart. (Full image sequence as Figure S1 in Supporting Information File 1 or in movie format as Supporting Information File 2). In the unthresholded sequence, many artifacts remain, making it difficult to track the progression. In the fully automated sequence, the automated algorithm does a much better job of correcting the data and outputting consistent images for comparison.

the system. Moreover, this automated flattening process allowed for a significant improvement in the speed of processing without sacrificing the quality of the results. The images shown in Figure 7 and Figure S1 in Supporting Information File 1 are only a few sample images from a much longer image sequence many hundreds of images long. The entire sequence was processed in only a few hours, representing a significant time saving compared to hand-correcting each image.

Conclusion

We have developed an automated, adaptive image-processing algorithm for high-speed AFM image sequences. This is achieved by identifying the background region and determining line-by-line offsets using an iterative process. The output of this iterative process is used to perform a single line-by-line offset correction, followed by a single 2-D polynomial-background removal step. This order of operations ensures that a minimal amount of manipulation is performed on the raw data to generate the final output. During the iterative background-identification process, the algorithm uses the standard deviation of one or more peaks in the height histogram of the image as a metric for determining the accuracy of the thresholded background identification. A similar metric is used to determine the accuracy of the line-by-line offset correction. The algorithm is specifically tailored to image sequences, ensuring consistent processing, offsets and contrast settings between frames in AFM movies in a short time period. On a standard PC, each channel takes about 10 seconds to process. This algorithm can be used on AFM data from most modes. We have demonstrated its applicability to data acquired in tapping mode in air, tapping mode in fluid, and quantitative nanomechanical mapping (QNM) in fluid. Finally, this algorithm can also be applied to strongly stepped samples, such as atomic layers. Figure S1 in Supporting Information File 1 shows the correction of samples with four and eight distinct levels. In both cases, the error is less than the noise in the image. With the release of modern HS-AFM systems by commercial manufacturers, this sort of automated processing will provide a significant benefit to this emerging research area in surface science and related fields. The algorithm and an associated graphical user interface are available at http://lbni.epfl.ch under the software section.

Experimental

Lipid preparation

Small unilamelar vesicle mixtures of 1,2-dilauroyl-*sn*-glycero-3-phosphocholine (DLPC) and 1,2-dipalmitoyl-*sn*-glycero-3-phosphocholine (DPPC) were prepared by sonication. Both lipid types were purchased from Avanti Polar Lipids Incorporated (Alabaster, AL, USA). Lipid powders were mixed prior to vesicle formation at a nominal molar ratio of 1:2, DLPC/DPPC. Vesicle solutions at 1 mg/mL were formed by transferring an appropriate mass of lipid into glass vials and dissolved with chloroform. The chloroform was evaporated off with dry nitrogen gas, leaving a thin film on the glass vial. The film was hydrated with Milli-Q water (Millipore, Billerica, MA, USA), generating large multilaminar vesicles (LMVs). The LMVs were then sonicated with a probe sonicator (BioLogics Inc, Manassas, VA, USA) to generate small unilaminar vesicles (SUVs). The SUVs were centrifuged to remove metal particles left from the probe sonicator. A 35 µL amount of the lipid preparation was warmed to 37 °C and deposited onto freshly cleaved mica surfaces, forming bilayers through vesicle fusion. Surfaces were allowed to incubate for at least a half hour in a humid environment at room temperature.

AFM imaging in air

Images were captured on a Multimode system with an E-scanner (Bruker Nano: Santa Barbara, CA, USA). Standard TESPA (Bruker AFM Probes: Camarillo, CA, USA) tapping cantilevers were used. The cantilever was driven at 348.44 kHz with an amplitude of 20.14 mV. Images at a size of 14.25 µm were captured at 512 × 512 pixels with a line rate of 1.5 lines/second. The data shown in the paper are crops from the center of the acquired image.

Quantitative Nanomechanical Mapping (QNM) – AFM imaging in fluid

Images were captured on a Multimode system with an E-scanner (Bruker Nano: Santa Barbara, CA, USA). A standard DNP-A (Bruker AFM Probes: Camarillo, CA, USA) cantilever was used with a spring constant of 0.40 N/m. Images at 5 µm were captured at 512 × 512 pixels with a line rate of 1 Hz. Manual control of the QNM parameters was used to minimize the applied force on the sample and the QNM drive amplitude. Each channel's acquired limits were minimized to limit bit quantization in the DSP.

HS-AFM imaging in fluid

Images were captured on a modified Multimode system with an E-scanner (Bruker Nano: Santa Barbara, CA, USA). A customized small-lever head allowed for the use of small cantilevers (SCL-Sensor.Tech., Vienna, Austria). The cantilever had a resonance frequency in fluid of 266.49 kHz, a spring constant of 0.54 N/m and a Q value of 2.68. Square areas of 4 µm were scanned at 78 lines/second at 256 × 256 pixels giving an approximate image acquisition time of 3 seconds/image. The fast-scan drive signal was passed through a custom filter designed to minimize the excitation of the tube scanner resonance [38]. The internal PID feedback of the Nanoscope 5 controller was bypassed with a Labview controlled FPGA based PID with a loop rate of 575 kHz.

Supporting Information

Supporting Information File 1

Further details on imaging and image processing
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-3-84-S1.pdf]

Supporting Information File 2

Image sequence movie
The movie shows the entire image sequence of the fully corrected data, with vertical median correction, from Figure 7. Each frame is approximately three seconds apart. The vertical scale of all images is in nanometers.
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-3-84-S2.mov]

Supporting Information File 3

User Manual
The user manual presented here contains a brief description of how to use the program and the parameters available for each channel.
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-3-84-S3.pdf]

Acknowledgements

The authors would like to thank Jaco de Groot (j.groot@ucl.ac.uk) for writing the initial Matlab import script for Nanoscope files. The authors would also like to thank the Atelier de l'Institut de microtechnique (ATPR) at EPFL for their on-going support and fabrication of research equipment. This work has been funded by the following sources: Swiss National Science Foundation (205321_134786), National Institutes of Health under Award No. RO1 GM65354, and the European Union's Seventh Framework Programme FP7/2007-2011 under grant agreement n° 286146.

References

1. Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986,** *56,* 930–933. doi:10.1103/PhysRevLett.56.930
2. Czajkowsky, D. M.; Sheng, S.; Shao, Z. *J. Mol. Biol.* **1998,** *276,* 325–330. doi:10.1006/jmbi.1997.1535
3. Scheuring, S.; Müller, D. J.; Ringler, P.; Heymann, J. B.; Engel, A. *J. Microsc. (Oxford, U. K.)* **1999,** *193,* 28–35. doi:10.1046/j.1365-2818.1999.00434.x
4. Scheuring, S.; Ringler, P.; Borgnia, M.; Stahlberg, H.; Müller, D. J.; Agre, P.; Engel, A. *EMBO J.* **1999,** *18,* 4981–4987. doi:10.1093/emboj/18.18.4981
5. Müller, D. J.; Fotiadis, D.; Scheuring, S.; Müller, S. A.; Engel, A. *Biophys. J.* **1999,** *76,* 1101–1111. doi:10.1016/S0006-3495(99)77275-9
6. Müller, D. J.; Heymann, J. B.; Oesterhelt, F.; Möller, C.; Gaub, H.; Büldt, G.; Engel, A. *Biochim. Biophys. Acta, Bioenerg.* **2000,** *1460,* 27–38. doi:10.1016/S0005-2728(00)00127-4
7. Hassenkam, T.; Fantner, G. E.; Cutroni, J. A.; Weaver, J. C.; Morse, D. E.; Hansma, P. K. *Bone* **2004,** *35,* 4–10. doi:10.1016/j.bone.2004.02.024
8. Czajkowsky, D. M.; Hotze, E. M.; Shao, Z.; Tweten, R. K. *EMBO J.* **2004,** *23,* 3206–3215. doi:10.1038/sj.emboj.7600350
9. Lodowski, D. T.; Salom, D.; Le Trong, I.; Teller, D. C.; Ballesteros, J. A.; Palczewski, K.; Stenkamp, R. E. *J. Struct. Biol.* **2007,** *158,* 455–462. doi:10.1016/j.jsb.2007.01.017
10. Casuso, I.; Kodera, N.; Le Grimellec, C.; Ando, T.; Scheuring, S. *Biophys. J.* **2009,** *97,* 1354–1361. doi:10.1016/j.bpj.2009.06.019
11. Picas, L.; Montero, M. T.; Morros, A.; Cabañas, M. E.; Seantier, B.; Milhiet, P.-E.; Hernández-Borrell, J. *J. Phys. Chem. B* **2009,** *113,* 4648–4655. doi:10.1021/jp8102468
12. Czajkowsky, D. M.; Shao, Z. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106,* 14960–14965. doi:10.1073/pnas.0903805106
13. Schönherr, H.; Vancso, G. J. Atomic Force Microscopy in Practice. *Scanning Force Microscopy of Polymers;* Springer: Berlin Heidelberg, 2010; pp 25–75.
14. Burns, A. R. Atomic force microscopy of lipid domains in supported model membranes. In *Lipid Rafts;* McIntosh, T. J., Ed.; Methods in Molecular Biology, Vol. 398; Humana Press: Totowa, N.J., 2007; pp 263–282.
15. Venkataraman, S. S.; Allison, D. P.; Qi, H.; Morrell-Falvey, J. L.; Kallewaard, N. L.; Crowe, J. E., Jr.; Doktycz, M. J. *Ultramicroscopy* **2006,** *106,* 829–837. doi:10.1016/j.ultramic.2006.01.014
16. Bickmore, B. R.; Rufe, E.; Barrett, S.; Hochella, M. F., Jr. *Geol. Mater. Res.* **1999,** *1,* No. 5.
17. Tsaftaris, S. A.; Zujovic, J.; Katsaggelos, A. K. Automated line flattening of Atomic Force Microscopy images. In *Proceedings of the 15th IEEE International Conference on Image Processing,* IEEE Press: Piscataway, NJ, 2008; pp 2968–2971.
18. Giocondi, M.-C.; Milhiet, P. E.; Dosset, P.; Le Grimellec, C. *Biophys. J.* **2004,** *86,* 861–869. doi:10.1016/S0006-3495(04)74161-2
19. Richter, R. P.; Brisson, A. R. *Biophys. J.* **2005,** *88,* 3422–3433. doi:10.1529/biophysj.104.053728
20. Watari, F. *J. Electron Microsc.* **2005,** *54,* 299–308. doi:10.1093/jmicro/dfi056
21. Kindt, J. H.; Thurner, P. J.; Lauer, M. E.; Bosma, B. L.; Schitter, G.; Fantner, G. E.; Izumi, M.; Weaver, J. C.; Morse, D. E.; Hansma, P. K. *Nanotechnology* **2007,** *18,* 135102. doi:10.1088/0957-4484/18/13/135102
22. Blanchette, C. D.; Lin, W.-C.; Orme, C. A.; Ratto, T. V.; Longo, M. L. *Langmuir* **2007,** *23,* 5875–5877. doi:10.1021/la7004584
23. Blanchette, C. D.; Orme, C. A.; Ratto, T. V.; Longo, M. L. *Langmuir* **2008,** *24,* 1219–1224. doi:10.1021/La702364g
24. Erickson, B.; DiMaggio, S. C.; Mullen, D. G.; Kelly, C. V.; Leroueil, P. R.; Berry, S. A.; Baker, J. R., Jr.; Orr, B. G.; Banaszak Holl, M. M. *Langmuir* **2008,** *24,* 11003–11008. doi:10.1021/La801497d
25. Ando, T.; Kodera, N.; Takai, E.; Maruyama, D.; Saito, K.; Toda, A. *Proc. Natl. Acad. Sci. U. S. A.* **2001,** *98,* 12468–12472. doi:10.1073/pnas.211400898
26. Hansma, P. K.; Schitter, G.; Fantner, G. E.; Prater, C. B. *Science* **2006,** *314,* 601–602. doi:10.1126/science.1133497
27. Milhiet, P.-E.; Yamamoto, D.; Berthoumieu, O.; Dosset, P.; Le Grimellec, C.; Verdier, J.-M.; Marchal, S.; Ando, T. *PLoS One* **2010,** *5,* e13240. doi:10.1371/journal.pone.0013240
28. Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. *Nature* **2010,** *468,* 72–76. doi:10.1038/nature09450

29. Fantner, G. E.; Barbero, R. J.; Gray, D. S.; Belcher, A. M. *Nat. Nanotechnol.* **2010,** *5,* 280–285. doi:10.1038/nnano.2010.29
30. Högström, R.; Korpelainen, V.; Riski, K.; Heinonen, M. *Metrologia* **2010,** *47,* 670–676. doi:10.1088/0026-1394/47/6/006
31. Kienberger, F.; Pastushenko, V. P.; Kada, G.; Puntheeranurak, T.; Chtcheglova, L.; Riethmueller, C.; Rankl, C.; Ebner, A.; Hinterdorfer, P. *Ultramicroscopy* **2006,** *106,* 822–828. doi:10.1016/j.ultramic.2005.11.013
32. Pfeifer, T.; Freudenberg, R.; Dussler, G.; Bröcher, B. *Measurement* **2001,** *30,* 1–18. doi:10.1016/S0263-2241(00)00050-6
33. Vorburger, T. V.; Dagata, J. A.; Wilkening, G.; Lizuka, K.; Thwaite, E. G.; Lonardo, P. *CIRP Ann.* **1997,** *46,* 597–620. doi:10.1016/S0007-8506(07)60881-6
34. Girard, P. *Nanotechnology* **2001,** *12,* 485–490. doi:10.1088/0957-4484/12/4/321
35. Starink, J. P. P.; Jovin, T. M. *Surf. Sci.* **1996,** *359,* 291–305. doi:10.1016/0039-6028(96)00367-6
36. Chen, S.-W. W.; Pellequer, J.-L. *BMC Struct. Biol.* **2011,** *11,* 7. doi:10.1186/1472-6807-11-7
37. Lenihan, T. G.; Malshe, A. P.; Brown, W. D.; Schaper, L. W. *Thin Solid Films* **1995,** *270,* 356–361. doi:10.1016/0040-6090(95)06747-7
38. Burns, D. J.; Youcef-Toumi, K.; Fantner, G. E. *Nanotechnology* **2011,** *22,* 315701. doi:10.1088/0957-4484/22/31/315701
39. Schitter, G.; Stemmer, A. *Microelectron. Eng.* **2003,** *67–68,* 938–944. doi:10.1016/S0167-9317(03)00157-6
40. Schitter, G.; Stark, R. W.; Stemmer, A. *Ultramicroscopy* **2004,** *100,* 253–257. doi:10.1016/j.ultramic.2003.11.008
41. Schitter, G.; Stemmer, A. *IEEE Trans. Control Syst. Technol.* **2004,** *12,* 449–454. doi:10.1109/TCST.2004.824290
42. Aphale, S.; Fleming, A. J.; Moheimani, S. O. R. *Micro Nano Lett.* **2007,** *2,* 9–12. doi:10.1049/mnl:20065075

License and Terms

Beilstein Journal of Nanotechnology

Effect of spherical Au nanoparticles on nanofriction and wear reduction in dry and liquid environments

Dave Maharaj and Bharat Bhushan*

Full Research Paper

Open Access

Address:
Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics (NLBB), The Ohio State University, 201 W. 19th Avenue, Columbus, Ohio 43210-1142, USA

Email:
Bharat Bhushan* - Bhushan.2@osu.edu

* Corresponding author

Keywords:
AFM; drug delivery; friction; gold nanoparticles; MEMS/NEMS; nanomanipulation

Beilstein J. Nanotechnol. **2012,** *3,* 759–772.
doi:10.3762/bjnano.3.85

Received: 18 July 2012
Accepted: 18 October 2012
Published: 15 November 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

Nano-object additives are used in tribological applications as well as in various applications in liquids requiring controlled manipulation and targeting. On the macroscale, nanoparticles in solids and liquids have been shown to reduce friction and wear. On the nanoscale, atomic force microscopy (AFM) studies have been performed in single- and multiple-nanoparticle contact, in dry environments, to characterize friction forces and wear. However, limited studies in submerged liquid environments have been performed and further studies are needed. In this paper, spherical Au nanoparticles were studied for their effect on friction and wear under dry conditions and submerged in water. In single-nanoparticle contact, individual nanoparticles, deposited on silicon, were manipulated with a sharp tip and the friction force was determined. Multiple-nanoparticle contact sliding experiments were performed on nanoparticle-coated silicon with a glass sphere. Wear tests were performed on the nanoscale with AFM as well as on the macroscale by using a ball-on-flat tribometer to relate friction and wear reduction on the nanoscale and macroscale. Results indicate that the addition of Au nanoparticles reduces friction and wear.

Introduction

Nano-objects are continually studied in tribological applications and increasingly in other applications that require controlled manipulation and targeting in liquid environments. The need for suitable forms of lubrication for micro/nano-electromechanical systems (MEMS/NEMS) and the ability to control and transport nano-objects in liquids, requires an understanding of nano-object behavior, with regards to friction, adhesion and wear, which is essential to their successful and continued application.

Increasing the lifetime and efficiency of individual components of systems is crucial to the commercialization of MEMS/

NEMS [1]. As one moves from the macroscale to the micro/nanoscale, surface to volume ratio increases. Therefore, adhesive and friction forces, which are dependent on surface area, become more significant. With MEMS/NEMS devices, the initial start-up forces and torques needed become high, which can hinder device operation and reliability [2]. The choice of a suitable lubricant on these scales becomes crucial.

Nano-objects are also used for applications that require controlled manipulation and targeting mechanisms in biomedicine and the oil industry. Applications include, but are not limited to, their use in targeted drug delivery and chemical sensors in the identification of oil, removal of contaminants and enhanced oil recovery (EOR). Au, iron oxide, polymer and silica nanoparticles have been studied in targeted drug delivery [3-8]. In cancer treatment, nanoparticles are either functionalized with biomolecules that recognize and attach to the cancer cells, [6,7] or in the case of iron-oxide nanoparticles, the nanoparticles are directed by an external magnetic field [9]. The cells are destroyed by drugs that coat the nanoparticles or by increasing the temperature of the nanoparticles to which cancer cells are susceptible. Figure 1a shows a nanoparticle loaded with a therapeutic drug and functionalized with a biomolecule (ligand), which selectively attaches to receptors in the cancer cell. The drug is then released as the nanoparticle diffuses into the diseased cell resulting in cell death.

Several factors need to be considered for the successful use of nanoparticles in targeted drug delivery. Biological barriers, including physical surfaces and the reticulo-endothelial system (RES), which detects and sequesters blood-borne particles, can prevent nanoparticles from reaching their intended target [7]. Smaller nanoparticles can diffuse through surfaces and avoid detection by the RES. Studies have shown that forces such as hydrodynamic and van der Waals forces along with the nanoparticle size influence lateral drift (margination) and adhesion to cell walls [5,11], which are important factors for effective drug delivery.

In oil-detection studies, as in the example shown in Figure 1b, oxidized carbon-black nanoparticles with a polyvinyl alcohol shell are coated with an oil-detecting agent (2,2',5,5'-tetrachlorobiphenyl (PCB)). The release of this agent on contact with hydrocarbons is used as an indication of the presence of oil on recovery of the nanoparticles [10]. In contaminant removal, nanocomposites composed of collagen and superparamagnetic iron-oxide nanoparticles (SPIONs) have been investigated. The collagen selectively absorbs the oil by motion of the nanoparticles towards the oil in a magnetic field [12]. Magnetic nanoparticles are also of interest in enhanced oil recovery (EOR) since they can be dispersed in fluid and manipulated and monitored by an external magnetic field [13,14]. In both oil detection and EOR, agglomeration of nanoparticles can prevent flow through

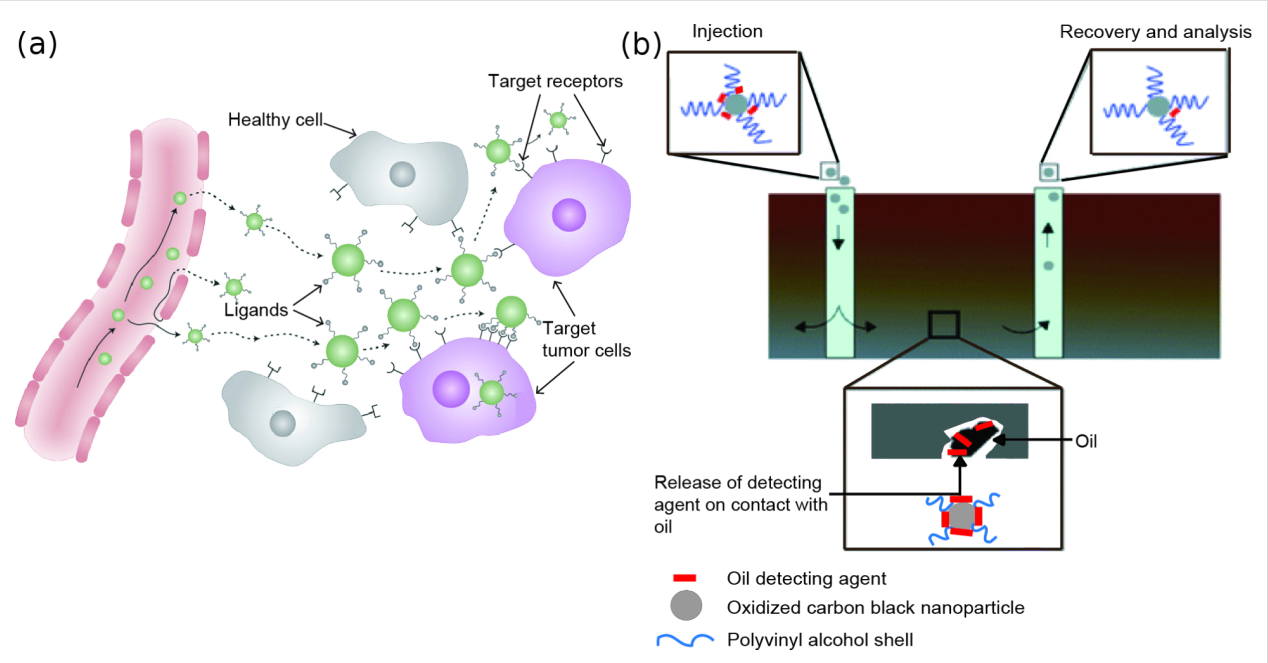

Figure 1: (a) Schematic of drug-carrying nanoparticles targeting cancer cells and releasing their therapeutic payload resulting in death of the cancer cell. Reprinted by permission from Macmillan Publishers Ltd [8], copyright 2011. (b) Showing the process of oil detection with nanoparticles. The carbon-black nanoparticles are coated with an oil-detecting agent. After injection into the ground, the agent is released on contact with hydrocarbons and this is used as an indication of the presence of oil on recovery and analysis of the nanoparticle [10]. Reproduced by permission of the Royal Society of Chemistry.

porous media. Nanoparticles can adhere to the surface over which they flow, which results in losses and prevents their eventual recovery [12,15]. Studies have shown that surface charge can cause nanoparticles in liquids to adhere to sites in porous media and hinder mobility [15]. Functionalizing nanoparticles with a hydrophilic polymer has been shown to reduce aggregation and improve flow [12].

For many of these applications, control of the friction of nanoparticles moving in the fluids, as well as the friction and adhesion as nano-objects come into contact with each other and surfaces present in their working environment, is necessary.

Nano-object additives have proven to be successful in macroscale studies in reducing friction and wear when added to solid materials and base-liquid lubricants and are expected to provide similar benefits on the micro/nanoscale. Some examples of nano-objects in liquids and their reported sizes, for friction and wear reduction, with studies carried out on the macroscale, are as follows: WS_2 platelets (0.5 μm) in commercial mineral oil [16], ferric oxide nanoparticles (20–50 nm) in 500 solvent neutral (SN) mineral oil [17], spherical MoS_2 (15–60 nm) in poly-alpha-olefin (PAO) and 150 SN [18], spherical WS_2 nanoparticles (50–350 nm) in SN 150 and SN 190 [19], spheroidal carbon-nano-onion nanoparticles (<10 nm) in PAO [20], WS_2 nanoparticles (120 nm) in paraffin oil [21], MoS_2 spheres (0.5–3 μm) in 500 SN oil [22] and carbon spheres (420 nm) in water [23]. Mechanisms for friction and wear reduction have been reported as tribofilm formation, rolling, sliding, and reduced contact area. It is expected that the reduced contact area and mobility offered by nano-objects observed on the macroscale will also lead to friction reduction and wear protection on the micro/nanoscale. These micro/nanoscale contacts are relevant for MEMS/NEMS devices.

In MEMS/NEMS devices, commercial lubricant oils are unacceptable as base liquids on machine components running in liquid. This is due to energy losses associated with the large viscous drag. In experiments where electrostatic micromotors are operated in a liquid environment, there have been problems of excessive drag and damping, which limited operating speeds, due to the use of high viscosity (20–60 cSt) oils [24]. However, studies have also demonstrated that friction and wear can be reduced with liquids of low viscosities [25]. Liquids such as glycerol and dodecane have been shown to reduce friction and wear. Glycerol has a dynamic viscosity (934 mPa·s) that is significantly higher than water (0.89 mPa·s) and studies were performed on the macroscale by using pin-on-disk testers [26]. In these studies, glycerol was also combined with water to lower the viscosity, which may be feasible for micro/nanoscale applications. Dodecane has been used as a base fluid with ZnS nanorod additives [27], which also resulted in a reduction in the coefficient of friction and wear. Tests were performed by using a surface force apparatus (SFA) with crossed-mica geometry with a 0–1600 μm^2 contact area.

To characterize friction forces associated with controlled manipulation and to understand the nature of the mechanism of friction and wear reduction of nanoparticles in MEMS/NEMS devices, studies have been carried out in both single-nanoparticle contact and multiple-nanoparticle contact with the aid of an AFM. Both mechanisms are described in detail in the following section.

Single-nanoparticle contact

In single-nanoparticle contact, a sharp AFM tip, as shown in Figure 2a as an example, is used to push the nanoparticle laterally (lateral manipulation). Manipulation studies of nanoparticles, with the aid of an AFM have shown that there is a contact-area dependence of the friction force. Several types of nanoparticles with reported diameters, such as latex spheres (80–100 nm) [28], Sb nanoparticles (120–400 nm) [29], (50–500 nm) [30], spherical SiO_2 nanoparticles (30 nm) [31] and spherical Au nanoparticles (25 nm) [32], (30–50 nm) [31] and (80 nm) [33] have been studied in both contact and intermittent-contact modes in dry environments. In liquid environments, Au nanoparticles (20–30 nm) have also been manipulated in water and ethanol with an AFM operated in intermittent-contact mode [34]. In addition to the contact-area dependence of friction observed in these studies, the relative-humidity (RH) dependence of friction was investigated by Mougin

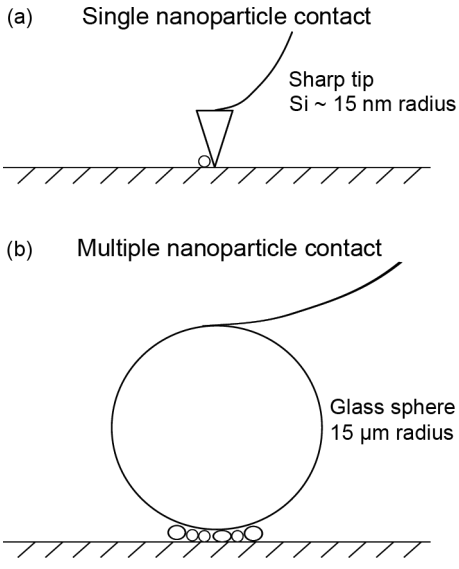

Figure 2: Schematics of (a) a sharp tip pushing a particle in single-particle contact and (b) a glass sphere sliding over several particles in multiple-particle contact.

et al. [32] and Palacio and Bhushan [31]. In the study by Mougin and co-workers [32], it was found that Au nanoparticles could not be moved in an ultrahigh vacuum (UHV) as compared to an ambient environment under otherwise identical manipulating conditions. Palacio and Bhushan [31] found that for larger nanoparticles, the friction force was lower at lower RH (10%) compared to higher RH (40%) for both Au and SiO_2 particles. Both studies were performed on silicon substrates. This would suggest that some adsorbed moisture between the nanoparticle and substrate is necessary for enhanced lubricity.

Manipulation studies of nanoparticles submerged in liquid environments, to simulate nanoscale contacts and characterize friction forces, are limited. Such studies are necessary for simulating the kinds of environments that involve controlled-manipulation and targeting-mechanism applications of nanoparticles. In addition, these studies provide insights into the interactions of single nanoparticles with a surface, in dry and submerged-in-liquid environments.

Multiple-nanoparticle contact

In addition to determining the friction force due to lateral manipulation, the effect of the normal load on the friction force has also been investigated. In multiple-nanoparticle contact, a glass sphere attached to an AFM cantilever, as shown in Figure 2b as an example, was used to slide over several nanoparticles. This type of study simulates the contacts experienced by MEMS/NEMS devices when nanoparticles are introduced for the purpose of friction and wear reduction.

Previous studies have been performed using a colloidal glass sphere attached to an AFM cantilever on bare silicon surfaces [35] and in multiple-nanoparticle contact with both immobile asperities on polymer surfaces [36] and mobile nanoparticles, such as spherical Au and SiO_2 nanoparticles on silicon surfaces [31]. In these studies, friction forces were reduced due to the reduced contact area provided and, in the case of Au and SiO_2, the possible sliding and possible rolling of individual nanoparticles. Similar to single-nanoparticle contact studies, AFM studies of multiple-nanoparticle contacts submerged in a liquid environment are also lacking. These studies are crucial to determine the added advantage of dispersing nanoparticles in liquids, in cases where the entire MEMS/NEMS system is submerged in a liquid environment. This has the ability to eliminate the adhesive effects of meniscus forces associated with the formation of capillary bridges due to adsorbed moisture on a surface.

Objective of this research

In this study, spherical Au nanoparticles are investigated to determine their effect on friction and wear under dry conditions and submerged in water. Lateral manipulation of single nanoparticles with a sharp tip is used to determine the friction force between the nanoparticle and the silicon substrate by AFM. The coefficient of friction is also investigated, with the aid of a glass sphere attached to an AFM cantilever sliding over multiple nanoparticles. Wear tests were performed on the nanoscale by using AFM and on the macroscale by using a ball-on-flat tribometer. This helps to link the nanoscale friction and wear to that observed on the macroscale and to fully understand the mechanisms involve.

Experimental

Choice of nanoparticle and operating liquid

As mentioned previously, Au is attractive for use in biomedicine since it is a noble metal, does not oxidize readily and has low to no toxic effects [6]. This also makes it suitable for use as a solid lubricant, and studies in liquids on the nanoscale to determine friction reduction and wear protection have not been reported. Spherical Au nanoparticles have a well-characterized shape and this thus eliminates orientation effects of tubular or cylindrical nanoparticles on the observed friction forces. Additionally, its small contact area and mobility is expected to contribute to friction force reduction and, when added to a base liquid, to further reduce the coefficient of friction and wear. Au nanoparticles are also suitable for manipulation studies since they are found in applications requiring controlled manipulation and targeting.

Liquids such as glycerol and dodecane have been shown to reduce friction and wear, as previously mentioned. However, our attempts to combine glycerol with Au nanoparticles suspended in deionized (DI) water lead to agglomeration of the nanoparticles. Evidence of agglomeration was observed by a change in color of the solution, from red to purple. In the case of dodecane, its immiscibility with water prevented its use with Au nanoparticles suspended in DI water.

In addition to the fact that Au nanoparticles are provided already suspended in DI water, the low viscosity of water and its ability to provide a surface of low shear strength [35] for sliding, makes it a good candidate as an operating fluid. Water is also attractive due to its environmentally friendly nature.

Materials and sample preparation

Si (100) silicon wafers with a native oxide layer (University Wafers, Boston, MA) were ultrasonically cleaned in DI water, followed by isopropyl alcohol (IPA) and finally acetone for 15 min each. For experiments involving nanoparticle-coated surfaces under dry conditions, several droplets of Au nanoparticles suspended in DI water (Nanopartz, Inc., Loveland, CO) were deposited onto the clean Si (100) substrate by using a syringe. A 25% concentration of an initial 0.05 mg/mL solution

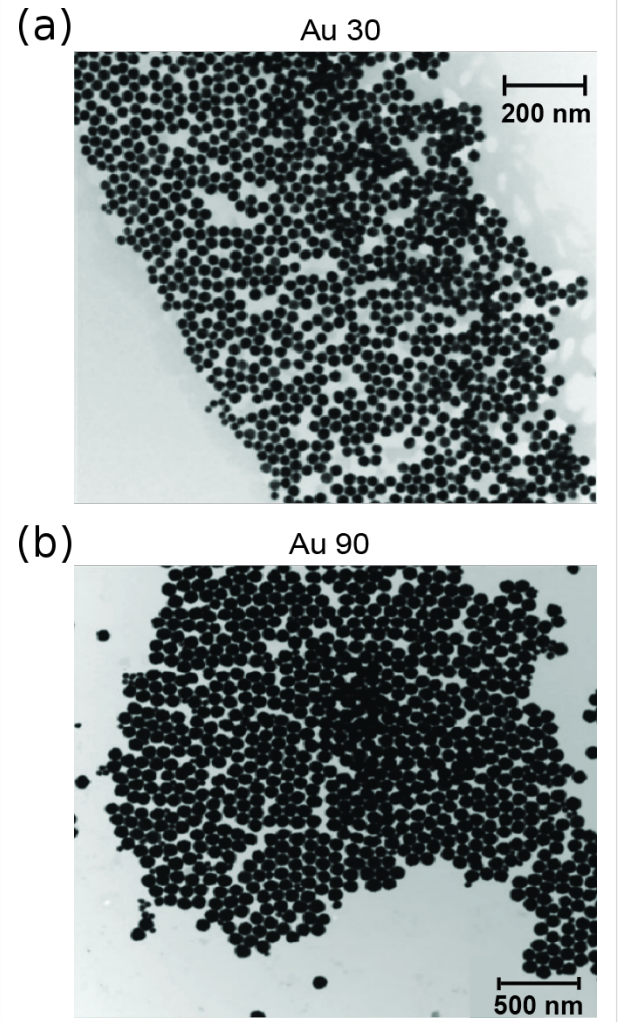

Figure 3: TEM images of spherical Au nanoparticles approximately (a) 30 nm in diameter and (b) 90 nm in diameter [37].

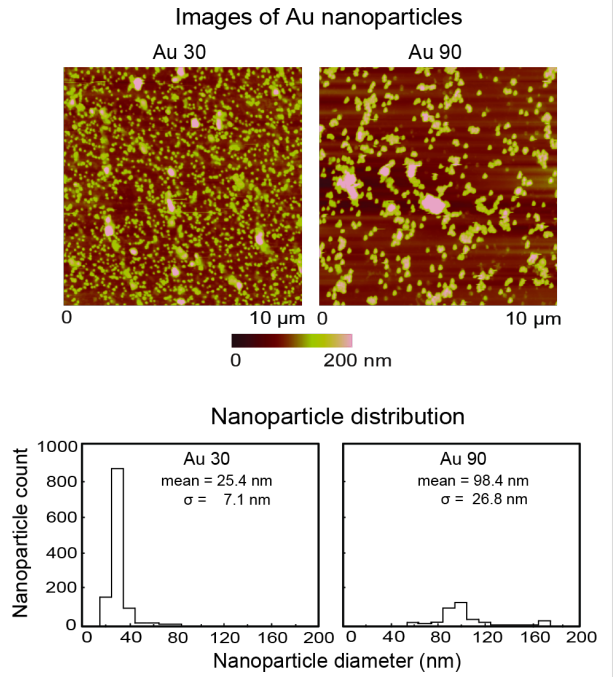

Figure 4: Topography map of Au 30 and Au 90 with corresponding histograms depicting the nanoparticle size distribution. The average nanoparticle diameter for Au 30 is 25.4 ± 7.1 nm and for Au 90 is 98.4 ± 27 nm from each histogram.

was used for all sliding and wear experiments unless otherwise stated. The substrate was then placed on a hot plate and heated to a temperature of about 70–80 °C and left until the water evaporated. The nominal diameters for the Au nanoparticles as reported by the manufacturer were about 30 and 90 nm, to be henceforth referred to as Au 30 and Au 90, respectively. Figure 3 shows transmission electron microscopy (TEM) images of spherical Au 30 and Au 90 nanoparticles [37].

Typical nanoparticle distributions obtained from topography images by using a commercial AFM (Multimode, Bruker, Santa Barbara, CA) are shown in Figure 4, for a 10 µm × 10 µm scan size. The average nanoparticle diameter for Au 30 is 25.4 ± 7.1 nm and for Au 90 is 98.4 ± 27 nm. The nanoparticle coverage analysis was performed by using SPIP 5.1.11 (Image Metrology A/S, Horshølm, Denmark). Imaging was performed at a normal load of 10 nN.

For experiments on nanoparticles submerged in water, a fluid cell consisting of a standard multimode cantilever holder (Bruker, Santa Barbara, CA) with a glass plate glued just above the cantilever was placed over the silicon substrate. The resulting meniscus bridge formed between the glass plate and substrate completely encloses the cantilever, which eliminates any viscous effects and adhesion due to meniscus forces.

Nanoscale friction force

Friction force data for nanoscale experiments were obtained by using the previously mentioned AFM for both single- and multiple-nanoparticle contact. The friction signals obtained in both cases were converted to forces by using an established calibration method [2,38]. Normal loads were determined by multiplying the cantilever vertical deflection by the cantilever stiffness [2]. The vertical deflection in turn was obtained by operating the cantilever in force-calibration mode, in which the deflection sensitivity obtained from the force curve was multiplied by the change in setpoint voltage.

For single-nanoparticle contact, a sharp silicon tip (FORT series, Applied NanoStructures, Inc., Santa Clara, CA,) with a spring constant k = 3 N/m and nominal radius of 15 nm was used for manipulation of a single Au nanoparticle under dry conditions. For submerged-in-water conditions, a silicon nitride tip of lower force constant was used (Orc8 series, Bruker,

Camarillo, CA) with k = 0.05 N/m and a nominal radius of 15 nm. For dry conditions, a 10% concentration of Au nanoparticles was used to ensure that isolated single nanoparticles could be found and would not be hindered by other nanoparticles during a manipulation attempt. The average value of the friction force presented is the result of five manipulations.

In multiple-nanoparticle contact, to determine the coefficient of friction, a soda lime glass sphere (Duke Scientific Corporation, Palo Alto, CA) of nominal radius 15 μm attached to a silicon probe (FORT series, Applied NanoStructures, Inc., Santa Clara, CA,) with a spring constant k = 3 N/m was used. Coefficient of friction data were obtained by plotting the friction force as a function of normal load from five random spots on the test samples.

Nanoscale wear

Wear tests on the nanoscale were performed by using the glass sphere attached to a silicide coated cantilever (NANOSENSORS, Neuchatel, Switzerland) with k = 40 N/m. In this case a cantilever of higher stiffness was used to obtain a normal load of 20 μN, which is not possible with the cantilevers mentioned previously. This was carried out for 1, 10 and 100 cycles at 10 Hz, over a 5 μm × 5 μm scan size. A larger scan (10 μm × 10 μm) was then taken of the area enclosing the wear region for comparison. Tests were not performed under submerged-in-water conditions, since at lower cycles, as the test is completed, nanoparticles suspended in solution will continue to be deposited on the surface as the water evaporates, and hence cover the wear area. Representative data for 5–6 tests are summarized in the results section. All experiments were performed at room temperature (23 °C) and 50–55% relative humidity.

Macroscale friction and wear

For comparison to the nanoscale, macroscale friction tests were conducted by using a ball-on-flat tribometer to determine if similar effects would be observed on both scales. For these tests, Au 90 was chosen as a representative nanoparticle. A sapphire ball of 1.5 mm radius was fixed to a stationary holder. A normal load of 200 mN was applied to the surface of the substrate and the tribometer was operated in a reciprocating manner [25]. The stroke length was 10 mm with an average speed of 3.5 mm/s. Friction forces were measured with semiconductor strain gages for 500 cycles. In liquid environments, 2–3 droplets of DI water with and without Au 90 nanoparticles were deposited onto the silicon substrate with a syringe. The sapphire ball was then slid over the substrate. The coefficient of friction was obtained as a function of the number of cycles. Wear was characterized by using an optical microscope by taking micrographs of the wear scars created during the test.

Results and Discussion

In this section, results for experiments in single- and multiple-nanoparticle contact are given for dry conditions and submerged-in-water conditions. In single nanoparticle contact, the manipulation technique is first described for each condition, and then the friction forces for both Au 30 and Au 90 nanoparticles are compared and discussed for both dry conditions and submerged in water. For multiple-nanoparticle contact, friction forces and corresponding coefficient of friction data are given and explained in detail. In addition, wear data for both nanoscale wear in dry conditions and macroscale wear in dry and submerged-in-water conditions, with and without the addition of Au nanoparticles are also presented. AFM wear images are shown for nanoscale wear. For macroscale wear, optical and SEM micrographs and corresponding data for the coefficient of friction are given and discussed.

Single-nanoparticle contact: Lateral manipulation of nanoparticles over a silicon substrate

Manipulation technique in dry and liquid environments

For single-nanoparticle contact under dry conditions, a sharp tip is used to push Au nanoparticles in the lateral direction. Figure 5 shows examples of topography images of nanoparticles, highlighted by the squares before and after manipulation within the same scan area, for both Au 30 and Au 90. A normal load of 10 nN was used during imaging. This takes place on the scan line illustrated by the black arrows. A 2 μm × 2 μm scan area is used to ensure that the nanoparticle can still be seen in the same image after manipulation. This clearly shows that the nanoparticles are being moved by the AFM tip. The smaller Au 30 nanoparticles indicated by 1 and 2 are pushed a shorter distance as the tip stays in contact for a shorter time compared to a larger Au 90 nanoparticle (1). This occurs since the Au 90 nanoparticles have a larger radius and take longer to roll or slide out of contact with the tip.

Figure 6 shows examples of topography maps and 2-D friction force profiles for Au 30 and Au 90 nanoparticles. A 1 μm × 1 μm area is imaged before manipulation to identify nanoparticles of interest. During manipulation in the same scan area, as shown in Figure 6a, scanning proceeds in the slow scan direction (bottom to top), and the normal load is increased from 10 nN to 300 nN at the approximate center of the nanoparticle. This corresponds to an increase in the friction signal (A–B) on a single scan line, as illustrated by the black horizontal arrows. Increasing the normal load prevents the nanoparticle from being imaged, as the cantilever tip remains in contact with the substrate and does not slide over the nanoparticle to track its height. As the sharp tip continues to slide along the scan line from left

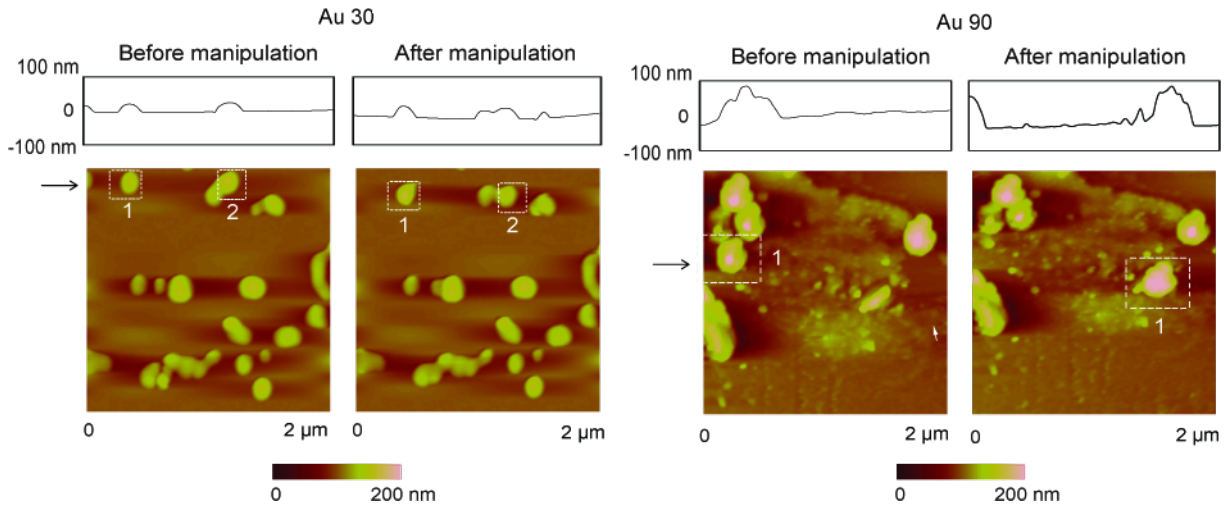

Figure 5: Two examples of topography maps and height profiles, at sections shown by the arrows, of Au 30 and Au 90 nanoparticles highlighted by the squares, manipulated within the same scan area. For Au 30, nanoparticles 1 and 2 are moved, and for Au 90, nanoparticle 1 is moved.

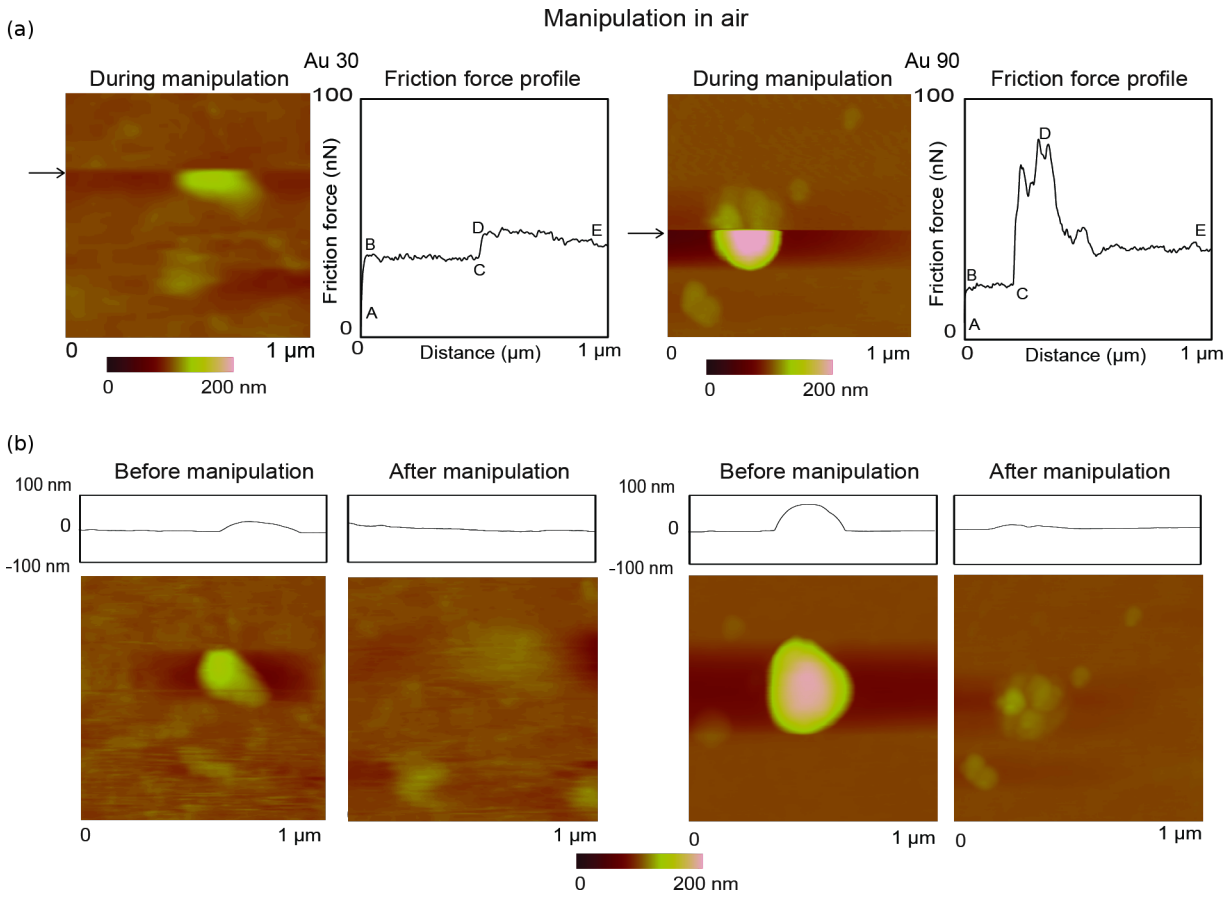

Figure 6: (a) Topography maps and 2-D friction force profiles of Au (30 nm in diameter) and Au (90 nm in diameter) during manipulation of single nanoparticles in dry conditions. During manipulation, the normal load is increased (A–B) at the approximate center of the particle image (indicated by the black arrows) from 10 nN to 300 nN on a single scan line. On this scan line, the sharp tip slides laterally (B–C) until contact is made with the particle. The lateral friction force (C–D) increases, as the sharp tip pushes against the particle, until it begins to slide. E shows the end of the scan line. (b) Two examples of topography images and height profiles showing the same scan area before and after manipulation.

to right (fast-scan direction), there is a further increase in the friction signal (C–D), which directly correlates to the twisting of the cantilever as it pushes against the nanoparticle, until friction is overcome and sliding begins. Point E represents the end of the scan line on which the manipulation takes place.

Corresponding profile and topography images of the same scan region before and after nanoparticle manipulation are shown in Figure 6b, which provide proof that the nanoparticle is pushed out of the scan area. Imaging of the nanoparticle was done in contact mode for both pre- and postmanipulation. The change in lateral force (C–D) is used to quantify the friction force between the nanoparticle and silicon surface as sliding is initiated.

For submerged-in-water conditions as shown in Figure 7, a 10 µm × 10 µm area is imaged at a normal load of 1 nN. This allows for multiple manipulations within a single topography scan. As an example, a single partial image of a nanoparticle, highlighted by the white squares, is used to demonstrate the manipulation technique. The associated topography and friction-force scan lines, before manipulation (Figure 7, top), during manipulation (Figure 7, middle) and after manipulation (Figure 7, bottom), for the above-mentioned nanoparticle are also shown. In this case the normal load does not have to be increased since the nanoparticles can be pushed due to the low adhesion between them and the substrate during scanning. As the nanoparticle is imaged, there is a rise in friction force associated with twisting of the cantilever (Figure 7, top) corresponding to the nanoparticle profile. In the middle set of scan lines, the topography is flat since the tip no longer follows the nanoparticle profile as it is being pushed, which corresponds to an increase in friction force. This increase (A–B) corresponds to the friction force between the Au nanoparticle and the silicon substrate at the initiation of sliding. In the bottom set of scan lines, both the topography and friction-force scan lines remain flat, which proves that the nanoparticle has been pushed out of the area. The friction-force results for nanoparticle manipulation under dry conditions and submerged in liquid is presented and discussed in the following section.

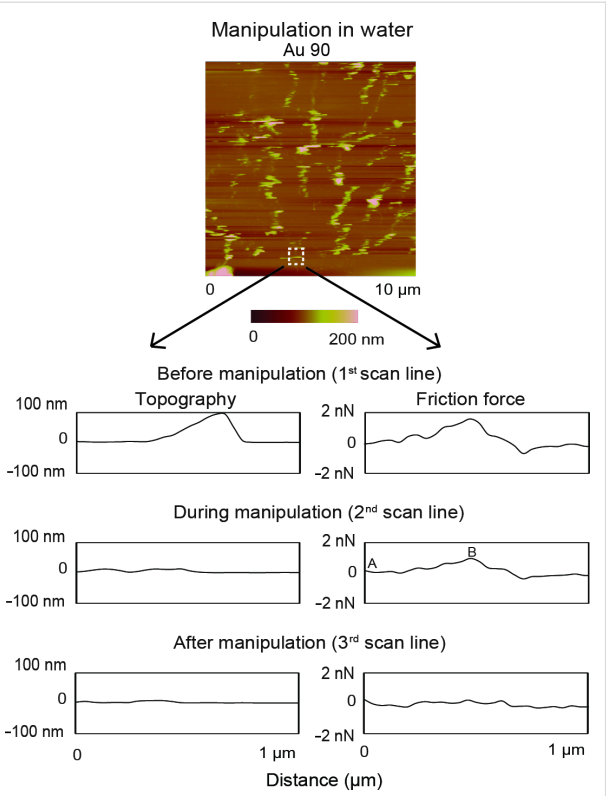

Figure 7: Topography map of Au (90 nm in diameter) nanoparticle submerged in water. Imaging is performed by using a cantilever with a low normal load of 1 nN. The highlighted particle is examined and the associated topography and friction force scan lines before manipulation, during manipulation and after manipulation are show. Before manipulation there is a rise in friction force corresponding to the particle profile as the cantilever deflects during imaging. During manipulation the topography profile is flat as the nanoparticle is being pushed, which corresponds to a rise in friction force (A–B). After manipulation, both the topography and friction force profiles remain flat as the nanoparticle is no longer there.

Comparison of friction forces obtained during lateral manipulation under dry and submerged- in-water conditions

Figure 8 shows the friction forces during nanoparticle manipulation of Au 30 and Au 90 nanoparticles under dry conditions and submerged in water. The data shows that Au 90 exhibit higher friction forces compared to Au 30. The normal load acting on the nanoparticle is due only to the mass of the nanoparticle since it is pushed from the side and the friction force is the result of adhesion between the nanoparticle and the silicon substrate. The adhesive force can include van der Waals forces under both dry and submerged-in-water conditions and meniscus forces under dry conditions. In this regime the friction force is not proportional to the normal load since it is dependent on the contact area. The friction force in this case, for single-nanoparticle contact of spherical shapes is proportional to $(\text{normal load})^{2/3}$ [1,2,31,39]. The normal load comprises the external normal load and the adhesive force. Since the adhesive force is dependent on surface area, it is expected that the larger Au 90 nanoparticles will display higher changes in friction force compared to the smaller Au 30 nanoparticles, and this is confirmed from the results shown in Figure 8a–b for both dry and submerged-in-water conditions.

In addition to comparing the friction force for Au 30 and Au 90 in dry conditions, Figure 8a also compares the friction force when the normal loads are 150 nN and 300 nN. The friction forces are comparable as is expected, since the normal load should only influence the interaction between the tip and the

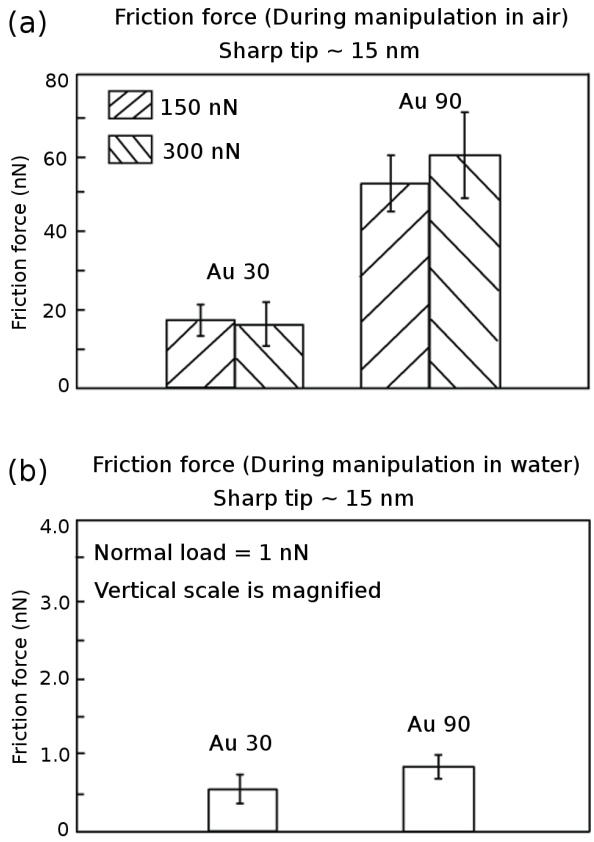

Figure 8: Friction force for Au 30 and Au 90 nanoparticles on the silicon substrate during manipulation, at normal loads of (a) 150 nN and 300 nN in air and (b) 1 nN in water. The vertical scale has been magnified for data in water.

substrate. The friction force between the Au nanoparticle and the substrate provides the additional offset C–D in the friction signal shown in Figure 6a during manipulation in air, independent of the tip–substrate friction force [30]. If the tip–substrate friction force approaches that of the nanoparticle–substrate force, it is expected that the friction signal during manipulation would be masked. This would necessitate lowering of the normal load as in the case of nanoparticles submerged in water.

Figure 8b presents the results of measurements of the average friction force for the Au 30 and Au 90 nanoparticles submerged in water. Since the high normal loads used during manipulation under dry conditions would mask the friction-force signal, a lower normal load of 1 nN is used for nanoparticle manipulation. This is sufficient since the nanoparticles are weakly adhered to the substrate and can be easily moved during the manipulation process shown in Figure 7. The adhesive forces are due to van der Waals interactions since there are no meniscus bridges formed under the submerged-in-water conditions. The lower friction forces observed under the submerged-in-water conditions compared to the dry conditions can thus be attributed to the nanoparticles sliding on an easily sheared surface and the elimination of meniscus forces.

Multiple-nanoparticle contact-sliding of a glass sphere over several nanoparticles under dry and submerged-in-water conditions

Nanoscale friction

In multiple-nanoparticle contact, the effect of the normal load acting on the Au nanoparticles between two surfaces is studied to determine the effects on the friction force. Figure 9 summarizes the friction forces and coefficient of friction under dry and

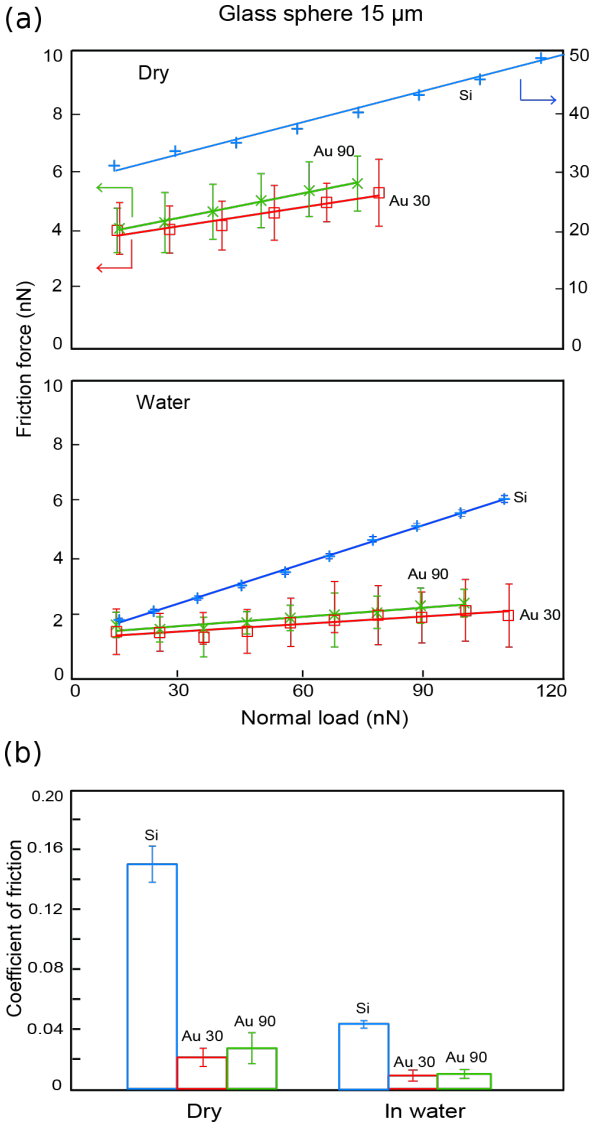

Figure 9: (a) Friction force as a function of normal load and (b) coefficients of friction for both dry and in-water conditions, with and without Au nanoparticles.

submerged-in-water conditions. In general, the friction forces were lower for sliding in water as compared to sliding under dry conditions, as shown in Figure 9a, for both nanoparticle-coated and uncoated surfaces. The same trend is observed in the data for the coefficient of friction (Figure 9b). Sliding in multiple-nanoparticle contact results in lower coefficients of friction under dry and submerged-in-water conditions, as compared to sliding on the silicon substrate. The coefficient of friction is also lower for sliding on Au 30 nanoparticles compared to Au 90. This is expected since the lateral manipulation of the nanoparticles resulted in lower friction forces for Au 30 nanoparticles compared to Au 90. The difference is more pronounced under the dry conditions compared to sliding in water. One reason for this could be that, under the submerged conditions, since the nanoparticles and cantilever are completely covered by water, the meniscus force contribution to the friction force is eliminated. One must also consider that, since the glass sphere is glued to the cantilever, the addition of the epoxy could contribute to an increased stiffness k of the cantilever, making it less sensitive to detecting changes in the lateral friction-force signal, especially for sliding in water where friction-force signals are lower. In the case of multiple-nanoparticle contact with an applied external load, the friction force shows a linear relationship as evidenced by the results for the coefficient of friction in Figure 9b.

It has also been demonstrated that sliding on multiple asperities on nanopatterned surfaces [36] results in the reduction of friction. In this particular case, the asperities are immobile, and reduction occurs as a result of the reduced contact area. For sliding on Au nanoparticles, friction-force reduction can be attributed to the mobility of the nanoparticles in addition to the reduced contact area. It is expected that as the glass sphere comes into contact with the Au nanoparticles, some of them will be deformed, since the larger nanoparticles will be encountered first and experience the highest contact pressures, due to fewer particles supporting the normal load. The resulting friction-reduction mechanism can thus be attributed to the reduced contact area, the sliding over deformed nanoparticles, and individual nanoparticles sliding with the glass sphere. In addition, it is also possible for some rolling to take place as the sphere encounters a greater number of nanoparticles and the contact pressure is reduced, leading to undeformed nanoparticles, which may roll between the surfaces.

In water, the presence of a liquid film between the glass sphere and the silicon substrate provides an interface of low shear strength resulting in a lower coefficient of friction [35]. In addition, since the glass sphere, cantilever and Au nanoparticles are completely covered in water, meniscus forces are eliminated, which also contributes to the reduction in the friction force.

Nanoscale wear

For a potential lubricant to be considered effective, it must not only be able to reduce the coefficient of friction, but also protect the underlying surface. Figure 10 summarizes the wear data for sliding on Si, and Si coated with Au 30 and Au 90 for 1, 10, and 100 cycles under dry conditions. As sliding progresses, a greater degree of wear is observed for the uncoated silicon substrate for 10 cycles compared to 1 cycle, and 100 cycles compared to 10 cycles, as seen within the first column. As we move from 1 to 10 cycles there is some roughening of the surface evidenced by the height profile. After 100 cycles a small amount of material has been removed, with a wear depth close to 0.5 nm. The very small amount of material removed at a load of 20 μN after 100 cycles would indicate that the wear mechanism is most likely due to breaking and removal of sharp asperities, as seen in adhesive wear [25], and eventual polishing of the surface as evidenced by the smoother height profile as the number of cycles progresses from 10 to 100.

For the surface coated with Au 30, as the number of cycles progresses from 1 to 10, it can be observed that nanoparticles still remain in the wear area, with evidence from the height profiles suggesting that they become compressed and deformed. At 100 cycles the nanoparticles are completely removed from the wear area and show agglomeration on the edges. For the surface coated with Au 90, after 1 cycle the nanoparticles are just beginning to be pushed out of the area and are completely removed after 10 cycles, in contrast to the surface coated with Au 30. This can be explained by the number of nanoparticles typically found on the surface. From the coverage data displayed in Figure 4, it can be seen that there is a much higher nanoparticle count for the Au 30 nanoparticles compared to Au 90 nanoparticles. It is expected that it would therefore take a longer time (more cycles) to completely remove the Au 30 nanoparticles from the surface.

It is also expected, since the softer Au nanoparticles remain in the wear area after 1 cycle for Au 90 and 10 cycles for Au 30, that the damage to the silicon surface should be less than that of an initially uncoated substrate, since the glass sphere is not directly sliding on the underlying surface and the contact load being exerted contributes towards deformation of the Au nanoparticles. Additionally, since Au 30 nanoparticles remain in the wear area longer than Au 90 nanoparticles, less wear of the surface is expected. Adhesive and abrasive wear of the silicon substrate is thus minimized since the asperities of the softer Au nanoparticles are more likely to deform and fracture during sliding than are those of the substrate or the glass sphere. At 100 cycles it is therefore expected that there would be less wear for the Au 30 surface compared to Au 90, with the greatest wear occurring on the bare silicon substrate.

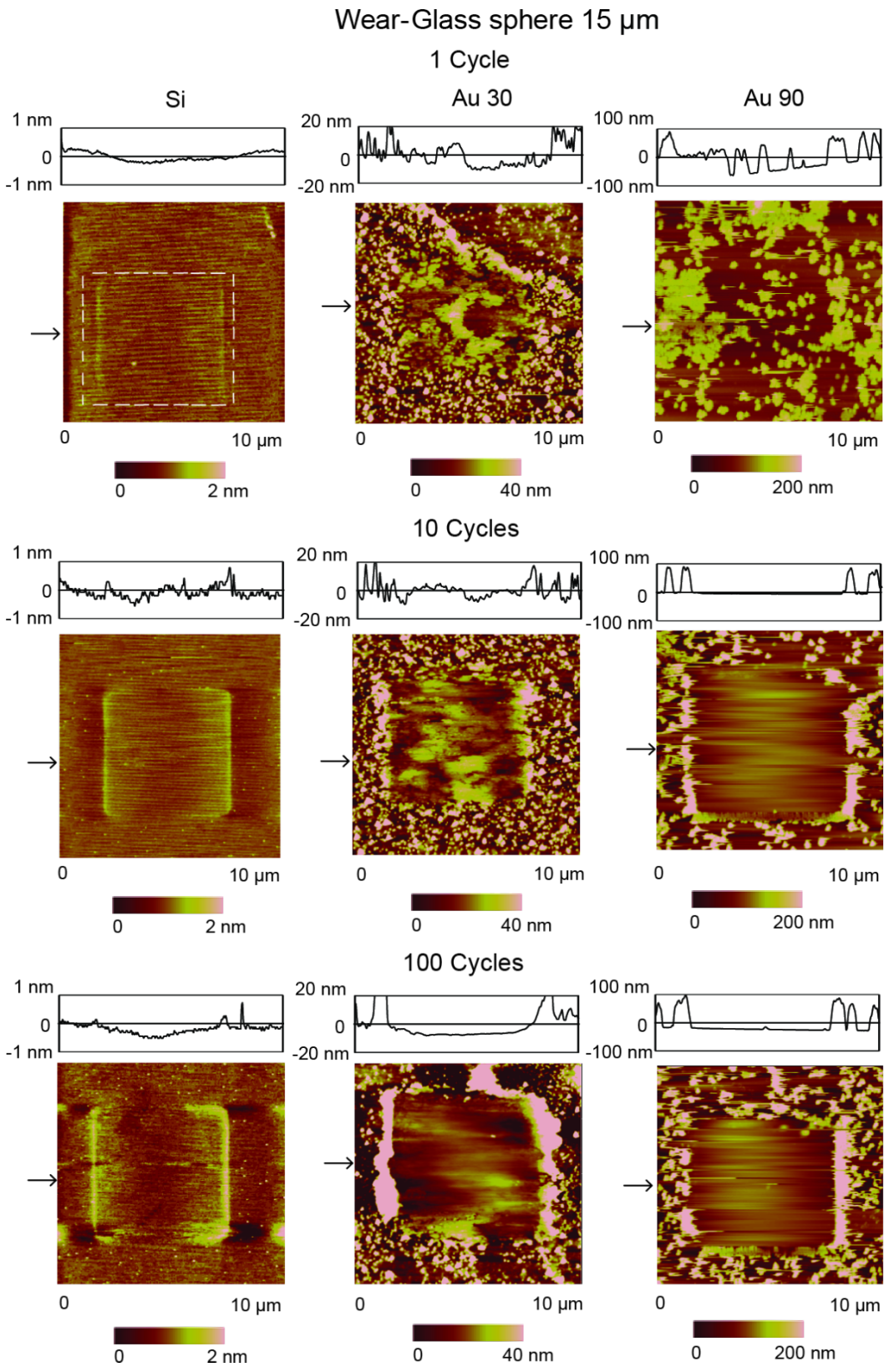

Figure 10: Topography maps and 2-D profiles, at sections shown by the arrows, after sliding for 1, 10 and 100 cycles with a normal load of 20 µN on Si and Si coated with Au 30 and Au 90.

Macroscale friction and wear

The results of the ball-on-flat wear tests are shown in Figure 11. Optical micrograph images of the wear scars for dry and water conditions with and without Au 90 nanoparticles are displayed in Figure 11a. In general, the widths of the wear scars shown are larger for sliding under dry conditions compared to sliding under water conditions, as the amount of wear is greater. Under dry conditions, the addition of Au 90 nanoparticles reduces the amount of wear compared to the uncoated silicon substrate. Under water conditions, the widths of the wear scars are comparable with or without the addition of nanoparticles.

Figure 11b shows a magnified scanning electron microscope (SEM) micrograph of the Au 90 wear scar under dry conditions, where agglomerations of Au 90 nanoparticles can be seen, highlighted by the squares. This is in contrast to Figure 3b, which shows TEM images of single unagglomerated nanoparticles. Agglomeration occurs during the wear process as nanoparticles are pressed together. The presence of the Au 90 nanoparticles within the wear scar contributes to the reduction in the coefficient of friction and wear by reducing the contact area, sliding and rolling of the nanoparticles.

Figure 11c shows data of the coefficient of friction for the four wear cases over 500 cycles. Higher coefficients of friction occur under dry conditions compared to water conditions. The lowest coefficients of friction coincide with the cases of least wear observed in Figure 11a. The results are similar to those on the nanoscale, where the submerged-in-water conditions generally produce lower coefficients of friction than those under the dry conditions, with the lowest coefficients of friction being observed with Au nanoparticles as shown in Figure 9b.

The addition of Au nanoparticles creates a barrier between the two surfaces and reduces the contact area. Similar to the nanoscale friction, it is expected that the observed reduction in the coefficient of friction is due to sliding on deformed nanoparticles, where deformation can occur when the initial contact is made with larger nanoparticles (fewer in number) resulting in higher contact pressure. It is also expected that some nanoparticles slide along with the ball as the number of cycles increases. As more nanoparticles support the load the contact pressure is reduced. This increases the chances of rolling for the undeformed nanoparticles as part of the friction-reduction mechanism.

In the case of DI water without Au nanoparticles, a surface of low shear strength is obtained, which makes it easier for the sapphire ball to slide [35]. Eventually, as the number of cycles continues, the coefficient of friction increases as wear of the substrate begins to take place and progresses. With the addition of Au 90 nanoparticles in DI water, the coefficient of friction is initially high and becomes lower after an initial settling-in period, which coincides with the initial deformation of the larger nanoparticles and eventual formation of a surface of low shear strength for sliding. In addition, as sliding progresses, the Au nanoparticles are continually being deposited on the surface,

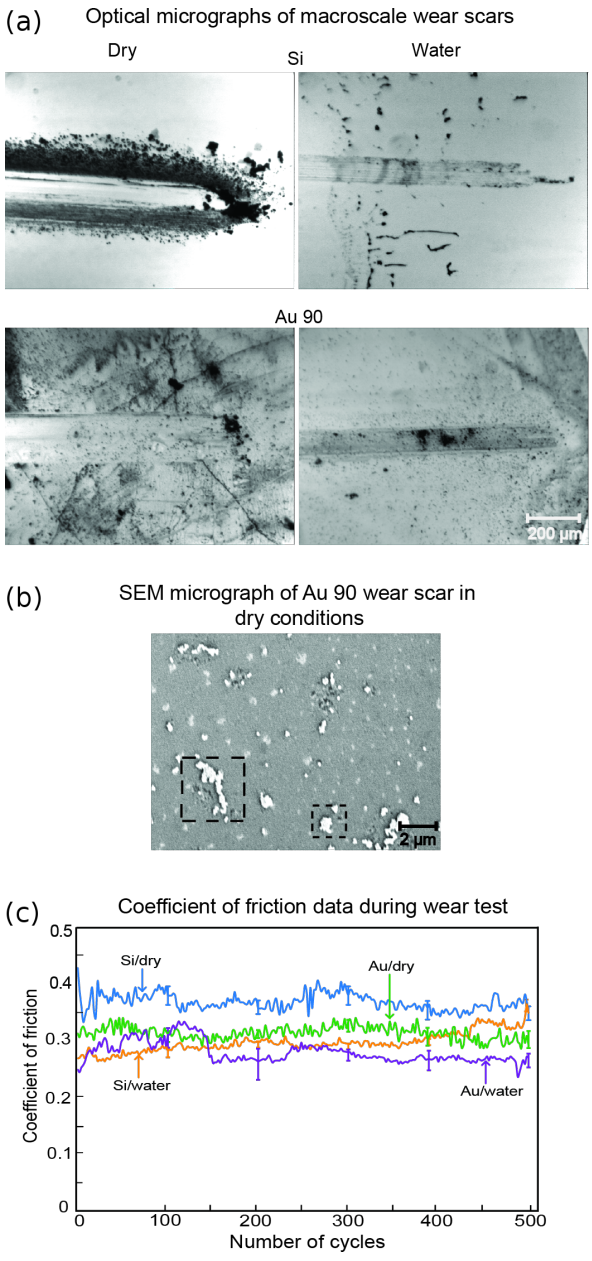

Figure 11: (a) Optical micrographs of the wear scars taken after 500 cycles. (b) SEM micrograph of the wear scar for agglomerated Au 90 nanoparticles, within the squares, under dry conditions. (c) Coefficients of friction from ball-on-flat tests, for both dry and water conditions, with and without Au 90 nanoparticles.

which replenishes the supply of nanoparticles for the sapphire ball to slide on. This combined with the low shear strength of the water contributes to the lowest observed coefficient of friction.

Conclusion

An investigation of the effects of spherical Au nanoparticles on friction and wear reduction was carried out. Both single- and multiple-nanoparticle contact cases were studied by using an AFM for nanoscale studies. For macroscale studies, a ball-on-flat tribometer was used.

For single-nanoparticle contact, there is a friction-force dependence on the size of the nanoparticle with lower forces observed under submerged-in-water conditions. For multiple-nanoparticle contact, sliding over Au nanoparticles in general reduced the coefficient of friction as compared to sliding on the bare silicon substrate. Coefficients of friction were also lower under submerged-in-water conditions compared to dry conditions for all surfaces due to the low shear strength of the surface provided.

In nanoscale wear experiments, addition of the Au nanoparticles provides protection from wear of the underlying substrate by preventing the glass sphere from coming directly into contact with the surface. Evidence of nanoparticle deformation was found in the case of the Au 30 nanoparticles. A larger nanoparticle count was also responsible for the better wear protection afforded by the Au 30 nanoparticles compared to the Au 90 nanoparticles.

Macroscale studies using a ball-on-flat tribometer showed similar trends to those on the nanoscale. The addition of Au 90 nanoparticles under dry conditions and suspended in water resulted in lower coefficients of friction. The addition of Au 90 nanoparticles also resulted in better wear resistance in both cases, with the best wear protection and lowest coefficients of friction being observed in water.

From the results obtained, Au nanoparticles prove to be a good potential lubricant as it lowers the coefficient of friction and minimizes wear. Further studies with other nano-objects under dry conditions and as an additive to water or other low-viscosity liquids could open up the possibilities for new types of hybrid lubricants. Such lubricants are expected to contribute to the increased lifetime and efficiency of MEMS/NEMS devices, which will lead to their successful commercialization. In addition, the study of manipulation of new types of nanoparticles in different liquids will lead to an understanding of their suitability for various applications in which friction forces are of concern in controlled manipulation and targeting mechanisms.

Acknowledgements

The financial support of this research was provided by a grant from the National Science Foundation, Arlington, VA (Grant # CMMI-1000108). We are especially grateful to Dr. Manuel Palacio for providing support in some of the data interpretation, for scientific discussion, and for editing the manuscript.

References

1. Bhushan, B. *Springer Handbook of Nanotechnology,* 3rd ed.; Springer: Heidelberg, Germany, 2010. doi:10.1007/978-3-642-02525-9
2. Bhushan, B. *Nanotribology and Nanomechanics, I and II,* 3rd ed.; Springer: Heidelberg, Germany, 2011.
3. Duncan, R. *Nat. Rev. Drug Discovery* **2003,** *2,* 347–360. doi:10.1038/nrd1088
4. Ferrari, M. *Nat. Rev. Cancer* **2005,** *5,* 161–171. doi:10.1038/nrc1566
5. Decuzzi, P.; Pasqualini, R.; Arap, W.; Ferrari, M. *Pharm. Res.* **2009,** *26,* 235–243. doi:10.1007/s11095-008-9697-x
6. Panyala, N. R.; Pena-Mendez, E. M.; Havel, J. *J. Appl. Biomed.* **2009,** *7,* 75–91.
7. Ferrari, M. *Trends Biotechnol.* **2010,** *28,* 181–188. doi:10.1016/j.tibtech.2009.12.007
8. Irvine, D. J. *Nat. Mater.* **2011,** *10,* 342–343. doi:10.1038/nmat3014
9. Tran, N.; Webster, T. J. *J. Mater. Chem.* **2010,** *20,* 8760–8767. doi:10.1039/c0jm00994f
10. Berlin, J. M.; Yu, J.; Lu, W.; Walsh, E. E.; Zhang, L.; Zhang, P.; Chen, W.; Kan, A. T.; Wong, M. S.; Tomson, M. B.; Tour, J. M. *Energy Environ. Sci.* **2011,** *4,* 505–509. doi:10.1039/c0ee00237b
11. Decuzzi, P.; Lee, S.; Bhushan, B.; Ferrari, M. *Ann Biomed Eng* **2005,** *33,* 179–190. doi:10.1007/s10439-005-8976-5
12. Thanikaivelan, P.; Narayanan, N. T.; Pradhan, B. K.; Ajayan, P. M. *Sci. Rep.* **2012,** *2,* 230. doi:10.1038/srep00230
13. Matteo, C.; Candido, P.; Vera, R.; Francesca, V. *Am. J. Appl. Sci.* **2012,** *9,* 784–793.
14. Ryoo, S.; Rahmani, A. R.; Yoon, K. Y.; Prodanovi, M.; Kotsmar, C.; Milner, T. E.; Johnson, K. P.; Bryant, S. L.; Huh, C. *J. Pet. Sci. Eng.* **2012,** *81,* 129–144. doi:10.1016/j.petrol.2011.11.008
15. Johnson, P. R.; Sun, N.; Elimelech, M. *Environ. Sci. Technol.* **1996,** *30,* 3284–3293. doi:10.1021/es960053+
16. Rapoport, L.; Bilik, Y.; Feldman, Y.; Homyonfer, M.; Cohen, S. R.; Tenne, R. *Nature* **1997,** *387,* 791–793. doi:10.1038/42910
17. Hu, H. Z.; Dong, J. X.; Chen, G. X. *Tribol. Int.* **1998,** *31,* 355–360. doi:10.1016/S0301-679X(98)00042-5
18. Cizaire, L.; Vacher, B.; Le Mogne, T.; Martin, J. M.; Rapoport, L.; Margolin, A.; Tenne, R. *Surf. Coat. Technol.* **2002,** *160,* 282–287. doi:10.1016/S0257-8972(02)00420-6
19. Greenberg, R.; Halperin, G.; Etsion, I.; Tenne, R. *Tribol. Lett.* **2004,** *17,* 179–186. doi:10.1023/B:TRIL.0000032443.95697.1d
20. Joly-Pottuz, L.; Dassenoy, F.; Belin, M.; Vacher, B.; Martin, J. M.; Fleischer, N. *Tribol. Lett.* **2005,** *18,* 477–485. doi:10.1007/s11249-005-3607-8
21. Rapoport, L.; Nepomnyashchy, O.; Lapsker, I.; Verdyan, A.; Soifer, Y.; Popovitz-Biro, R.; Tenne, R. *Tribol. Lett.* **2005,** *19,* 143–149. doi:10.1007/s11249-005-5095-2
22. Fu, X.; Zhou, X.; Shi, H.; Wu, D.; Hu, Z. J. *Tribol.* **2007,** *129,* 913–919. doi:10.1115/1.2768612
23. St. Dennis, J. E.; Jin, K.; John, V. T.; Pesika, N. S. *ACS Appl. Mater. Interfaces* **2011,** *3,* 2215–2218. doi:10.1021/am200581q

24. Reddyhoff, T.; Ku, I. S. Y.; Holmes, A. S.; Spikes, H. A. *Tribol. Lett.* **2011,** *41,* 239–246. doi:10.1007/s11249-010-9704-3
25. Bhushan, B. *Introduction to Tribology;* Wiley: New York, USA, 2002.
26. Nalam, P. C.; Clasohm, J. N.; Mashaghi, A.; Spencer, N. D. *Tribol. Lett.* **2010,** *37,* 541–552. doi:10.1007/s11249-009-9549-9
27. Akbulut, M.; Belman, N.; Golan, Y.; Israelachvili, J. *Adv. Mater.* **2006,** *18,* 2589–2592. doi:10.1002/adma.200600794
28. Ritter, C.; Heyde, M.; Stegemann, B.; Schwarz, U. D.; Rademann, K. *Langmuir* **2002,** *18,* 7798–7803. doi:10.1021/la011318n
29. Ritter, C.; Heyde, M.; Stegemann, B.; Rademann, K.; Schwarz, U. D. *Phys. Rev. B* **2005,** *71,* 085405. doi:10.1103/PhysRevB.71.085405
30. Dietzel, D.; Monninghoff, G. S.; Jansen, L.; Fuchs, H.; Ritter, C.; Scharwz, U. D.; Schirmeisen, A. *J. Appl. Phys.* **2007,** *102,* 084306. doi:10.1063/1.2798628
31. Palacio, M.; Bhushan, B. *Nanotechnology* **2008,** *19,* 315710. doi:10.1088/0957-4484/19/31/315710
32. Mougin, K.; Gnecco, E.; Rao, A.; Cuberes, M. T.; Jayaraman, S.; McFarland, E. W.; Haidara, H.; Meyer, E. *Langmuir* **2008,** *24,* 1577–1581. doi:10.1021/la702921v
33. Hui, X.; Regnier, S. *IEEE Trans. Nanotech.* **2012,** *11,* 21–33. doi:10.1109/TNANO.2010.2041359
34. Resch, R.; Lewis, D.; Meltzer, S.; Montoya, N.; Koel, B. E.; Madhukar, A.; Requicha, A. A. G.; Will, P. *Ultramicroscopy* **2000,** *82,* 135–139. doi:10.1016/S0304-3991(99)00152-7
35. Bhushan, B.; Sundararajan, S. *Acta Mater.* **1998,** *46,* 3793–3804. doi:10.1016/S1359-6454(98)00062-7
36. Burton, Z.; Bhushan, B. *Nano Lett.* **2005,** *5,* 1607–1613. doi:10.1021/nl050861b
37. Nanopartz.com. Accurate spherical gold nanoparticles. http://www.nanopartz.com/bare_spherical_gold_nanoparticles.asp (accessed Sept 20, 2012).
38. Ruan, J.-A.; Bhushan, B. *ASME. J. Tribol.* **1994,** *116,* 378–388. doi:10.1115/1.2927240
39. Schwarz, U. D.; Zwörner, O.; Köster, P.; Wiesendanger, R. *Phys. Rev. B* **1997,** *56,* 6987–6996. doi:10.1103/PhysRevB.56.6987

License and Terms

Beilstein Journal of Nanotechnology

Growth behaviour and mechanical properties of PLL/HA multilayer films studied by AFM

Cagri Üzüm[1], Johannes Hellwig[1], Narayanan Madaboosi[2,3], Dmitry Volodkin[1,3] and Regine von Klitzing[*1]

Full Research Paper

Open Access

Address:
[1]Stranski-Laboratorium, Department of Chemistry, TU Berlin, Strasse des 17. Juni 124, D-10623 Berlin, Germany, [2]Fraunhofer Institute for Biomedical Engineering, Am Mühlenberg 13, 14476 Potsdam-Golm, Germany and [3]Max Planck Institute of Colloids and Interfaces, Am Mühlenberg 1, 14476 Potsdam-Golm, Germany

Email:
Regine von Klitzing[*] - klitzing@mailbox.tu-berlin.de

* Corresponding author

Keywords:
atomic force microscopy; polyelectrolyte multilayers; stress relaxation; viscoelasticity; Young's modulus

Beilstein J. Nanotechnol. **2012,** *3,* 778–788.
doi:10.3762/bjnano.3.87

Received: 15 August 2012
Accepted: 21 October 2012
Published: 21 November 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

Scanning- and colloidal-probe atomic force microscopy were used to study the mechanical properties of poly(L-lysine)/hyaluronan $(PLL/HA)_n$ films as a function of indentation velocity and the number of polymer deposition steps n. The film thickness was determined by two independent AFM-based methods: scratch-and-scan and newly developed full-indentation. The advantages and disadvantages of both methods are highlighted, and error minimization techniques in elasticity measurements are addressed. It was found that the film thickness increases linearly with the bilayer number n, ranging between 400 and 7500 nm for n = 12 and 96, respectively. The apparent Young's modulus E ranges between 15 and 40 kPa and does not depend on the indenter size or the film bilayer number n. Stress relaxation measurements show that PLL/HA films have a viscoelastic behaviour, regardless of their thickness. If indentation is performed several times at the same lateral position on the film, a viscous/plastic deformation takes place.

Introduction

Polyelectrolyte multilayers (PEMs) have been studied intensely for the past two decades [1,2]. Despite their complex structure and wide range of applicability, PEMs can be prepared simply by alternating deposition of polycations and polyanions by dipping/spraying a substrate into/with the corresponding polyelectrolyte solutions. Potential applications of PEMs (e.g., filtration, paper making and biomaterials) require control of their adhesive behaviour by tuning the elastic/viscoelastic properties [3-5]. For example, the adsorption behaviour of proteins and cells, which is highly sensitive to the elasticity of the substrate, must be known and controlled for the development of contact lenses and antifouling materials [3,5,6]. Nevertheless, thin films have to be studied in the form in which they are available, making the use of macroscopic methods unsuitable. There-

fore, scanning- or colloidal-probe atomic force microscopy have been widely used for studying the topography and the mechanical properties of PEMs [3-5,7,8].

One of the first measurements of elastic modulus with atomic force microscopy (AFM) on biological films was performed on lung-cancer cells, back in 1993 [9]. Further measurements include different strains of *E. coli* with a colloidal probe [10], elastic modulus of human platelet cells [11], human bone cell or skeletal muscle cells [12], breast cancer cells [13,14], hydrogel films [15-17], or nanoribbons [18], as well as single hydrogel particles [19-22]. Recent advances in the area have been summarized by Picart and co-workers [23,24].

Several studies on the bio-applicability of polymer-based films showed that if cells are deposited on a surface with an elasticity gradient, they move from the softer region to a relatively harder one [3,6,25]. Richert et al. showed that chondrosarcoma cells adhere much more strongly on chemically cross-linked poly(L-lysine)/hyaluronic acid PLL/HA films than on native PLL/HA films due to the elasticity difference between the two structures, the cross-linked surface being harder [6]. Engler et al. reported a similar spreading behaviour for smooth-muscle cells [3]. An enhanced cell growth was observed also for cross-linked chitosan/hyaluronan multilayer films, as compared to the native ones [8]. These reports suggest that polyelectrolyte multilayer films are ideal matrices for bio-applications as their elasticity can be tuned in a wide range simply by changing the cross-linker content in the structure. In that manner, PLL/HA films gained more attention than their analogues, and a wide range of Young's moduli between 3 and 400 kPa was accessed by cross-linking (mostly with 1-ethyl-3-(3-(dimethylamino)propyl)carbodiimide, EDC) [3-5,7].

Although native (non-cross-linked) PLL/HA films were previously produced and well characterised, it is still a requirement to precisely measure not only the elasticity but also other rheological properties of these matrices on time scales suitable for biological processes [3]. In this work, scanning- and colloidal-probe AFM were used to perform nanoindentation on poly (L-lysine)/hyaluronan $(PLL/HA)_n$ films with n = 12–96, in order to better understand their growth behaviour, apparent Young's modulus, and viscoelastic properties.

Results and Discussion

Bilayer number *n* versus film thickness *h*

The thickness h of $(PLL/HA)_n$ films with n = 12, 24, 36, 48, 60, 72, 84, 96 was measured both to determine the growth regime and to be able to study the mechanical properties. Two methods were used to determine the thickness. The first one is the scratch-and-scan method and comprises the removal of a small part of the film and subsequent imaging of the surface with a scanning-probe AFM. An optical microscope image of the scratched area on a PLL/HA film is shown in Figure 1 together with its AFM micrograph. The scratched area can be clearly seen due to its smoothness and lower height. In order to ensure that the film removal was successful, force measurements were made on this area and no indentation was observed. The thickness was extracted from the cross-section profile by calculating the height difference of the higher and the lower areas (excluding the rim) as shown in Figure 1c.

The scratch-and-scan method has some advantages such as the possibility to observe any possible damage done to the film during indentation, by scanning a large area that includes the originally studied position on the film. Another advantage is that this method gives the prerequisite image as well as the sample thickness in force-mapping studies [13,14,21,22,26,27]. The disadvantages mainly lie in the difficulty of obtaining a detailed image on soft surfaces, especially in liquid media. A magnetically driven cantilever as used in this study enhances the image quality since the surrounding medium is not excited

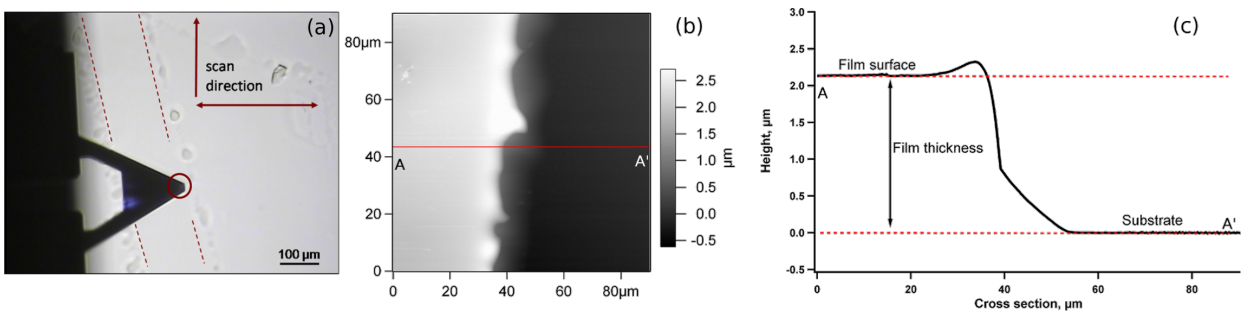

Figure 1: (a) Top camera view of the cantilever and the scratch on the film (delimited with the dashed lines). The circle points out the position of the tip. (b) 2-D AFM image of the scratch and the cross-section line (A–A′). The grey scale bar shows the height. (c) Height profile through the cross-section line (A–A′) for a measurement made on $(PLL/HA)_{36}$.

to vibrations. On the other hand, this type of cantilever driving is still not available in most commercial systems and the appropriate cantilevers are still not common [28].

The second method to determine the film thickness is the full-indentation method, introduced in the experimental section. The total penetration depth in a force measurement has already been used to determine the thickness of nanometre-scale coatings, e.g., lipid bilayers [29], but, to our knowledge, the total thickness of micron-scale polymeric films has not yet been extracted in this way. The full-indentation method requires a stiff, calibrated cantilever equipped with a tip that is significantly longer than the film is thick. The measurements are reasonably fast and reproducible, and the film damage caused is limited to a small area ($r \approx 50$ nm). The spring constant of the cantilever or the exact shape and size of the indenter do not have to be determined; however, an optical sensitivity calibration is required. A typical F versus δ curve for $(PLL/HA)_{72}$ and determination of its thickness are presented in Figure 2.

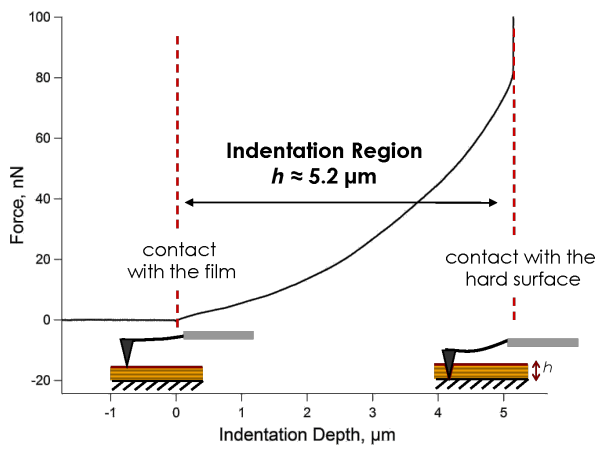

Figure 2: Calculation of the film thickness by the full-indentation method. The dashed lines show the tip-film and tip-substrate contact points. The measurement was made on $(PLL/HA)_{72}$.

A comparison of the film thickness obtained by scratch-and-scan and full-indentation is presented in Figure 3. The results of the two completely independent methods coincide very well within the experimental errors. Film thickness increases linearly with increasing bilayer number n (with an exception of $n = 48$) and ranges from about 0.4 to 7 µm for $n = 12$ and $n = 96$, respectively. Linear thickness growth with bilayer number indicates stacked polyelectrolyte layers that interpenetrate only into the neighbouring layers, forming periodic structures [5,24]. Although PLL/HA films are known to be highly hydrated and less ordered, it has been previously suggested that they grow exponentially [24,30,31] up to a certain bilayer number and thereafter the growth regime switches to a linear one [4,6,24,32,33]. For the dry state, a transition from exponential to linear growth was observed for $n = 12$–18 depending on the polyelectrolyte molecular weight [34,35]. Hence, our observation of a linear growth coincides well with the previous reports [4,6,24,32,33].

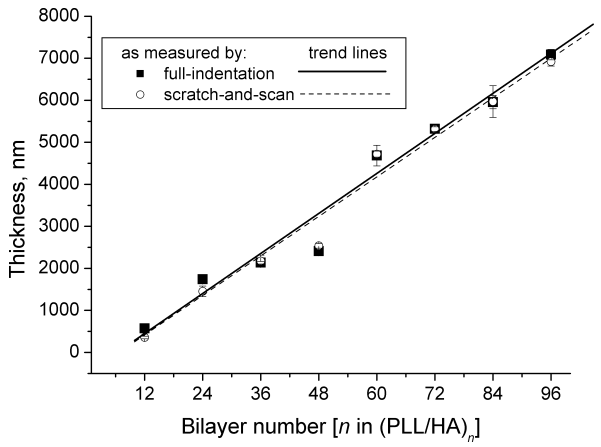

Figure 3: Thickness of $(PLL/HA)_n$ as a function of n as measured by scratch-and-scan and full-indentation methods. There is a good agreement between the two methods, and the thickness grows linearly with n. Error bars indicate the standard deviations.

The contact point and film thickness issues

The Hertzian model [36] calculates the Young's modulus E for each exact indentation depth, and thus a wrong indentation-depth determination may give rise to dramatic errors [37]. This problem is illustrated in Figure 4b.

Figure 4b shows the calculated Young's modulus E versus the indentation depth δ for $(PLL/HA)_{24}$ with two different contact points chosen on the raw F–δ curve as pointed out in Figure 4a. The first contact point is where the probe "jumps" to the surface, which can be recognized by an abrupt decrease in the measured force. The second contact point is chosen randomly at a further 35 nm, which is a clear shift considering the small film thickness. Although at deeper indentation depths the difference in the calculated E can be ignored, the deviation is dramatic in the region of small indentation depth, as previously reported [37-39]. In this study, the position of the minimum force was taken as the contact point, but in the case that there is no such minimum in the force curve, more effort is needed to determine an approximate contact point [21,37,38,40].

Another source of error in mechanical measurements of thin films is the substrate effect [37]. Although film stiffening due to the hard substrate is a real effect, it should be eliminated in the calculations since it does not reflect the material properties of

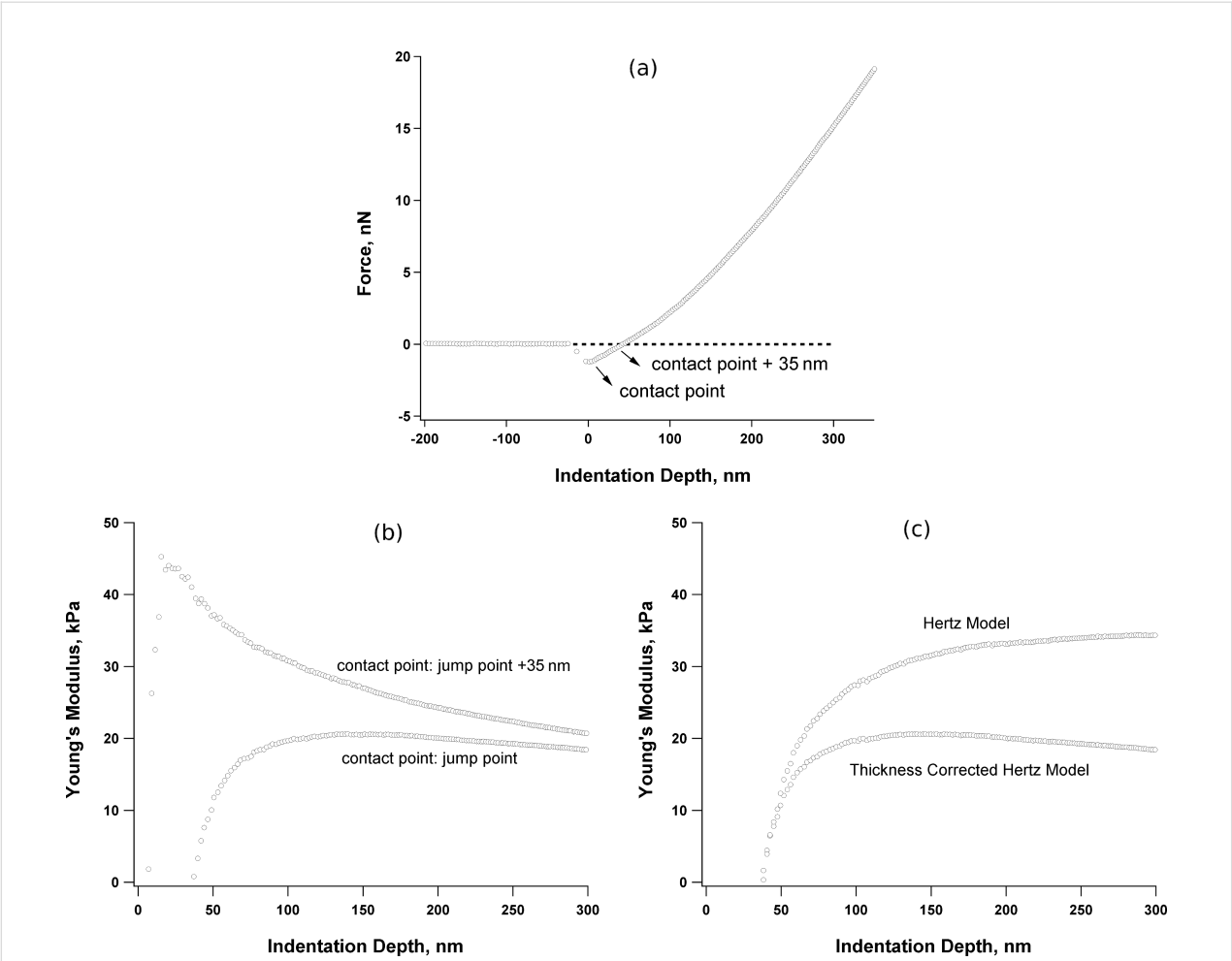

Figure 4: (a) A typical force–indentation depth curve as measured by CP-AFM on $(PLL/HA)_{24}$. The jump point of the colloidal-probe to the film surface was taken as the contact point. (b) Young's modulus versus indentation depth for two contact-point choices and (c) Young's modulus calculated with the classical [36] and thickness-corrected [37] Hertzian models.

the coating and may not be detected by the practical sensors, e.g., cells. Figure 4c shows the difference between E calculated by the classical Hertzian model [36] and the thickness-corrected one with Equation 2 [37], for $(PLL/HA)_{24}$ (thickness ≈ 1500 nm). E calculated with the classical Hertzian model is nearly two fold larger than that from the thickness-corrected one. It was observed that the substrate effect is less stressed for thicker films, but it cannot be neglected for a film thinner than ≈5000 nm.

Repetitive indentation measurements at one lateral position

In order to test the reversibility of film deformation in the indentation studies, a series of force measurements were performed repetitively at one fixed lateral position. Two different indentation velocities, i.e., 400 and 2000 nm/s were used, each at a different lateral position. Enough time was left between any two sequential measurements to allow a complete relaxation of the film. E of $(PLL/HA)_{72}$ as a function of the number of repeat measurements is shown in Figure 5 (E relative to the first measured one).

For both indentation velocities, repetitive measurements result in a continuous decrease in the apparent E, with the exception of some strong fluctuations, which are most likely measurement errors. The decrease is more stressed for slower indentation: E decreases by nearly 80% at 400 nm/s and by nearly 30% at 2000 nm/s. Presumably, the film is damaged once the applied load exceeds a certain limit, regardless of the indentation velocity, and it undergoes an irreversible viscous [41] or plastic deformation [42], leading to a thinner film and a softer structure. In order to avoid any irreversible deformation effects, in the following sections the lateral position of the measurement was changed each time after an indentation process.

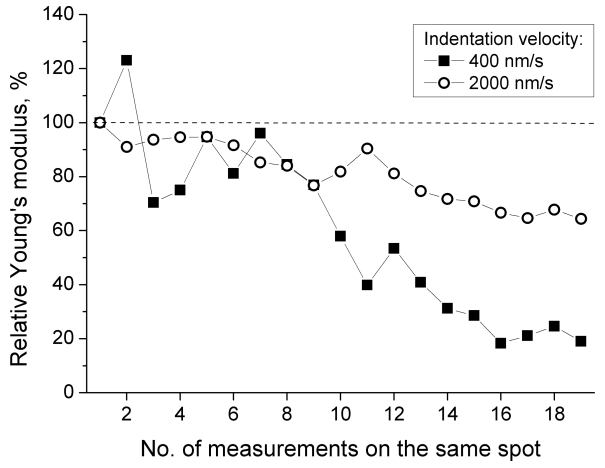

Figure 5: Relative percentage Young's modulus *E*% for $(PLL/HA)_{72}$ as measured on one fixed spot. All the values are normalised to the *E* calculated from the first indentation. Indentation velocities are 400 nm/s (squares) and 2000 nm/s (circles).

Indentation velocity versus Young's modulus *E*

Since the measurements take sometimes up to an hour, it was tested whether the mechanical properties of the film varied with time due to, e.g., instrument-based heating or contamination. The indentation velocity in the successive measurement sets was changed randomly in the following order: 400 nm/s, 400 nm/s, 50 nm/s, 100 nm/s, 500 nm/s, 1000 nm/s, 2000 nm/s, 1500 nm/s, 800 nm/s, 300 nm/s, 400 nm/s. The measurements were performed in a row without changing any further parameters except for the lateral position on the film. The dependence of the apparent *E* on the indentation velocity is shown in Figure 6a for $(PLL/HA)_{72}$. In this graph, the indentation velocity values in the *x*-axis are sorted according to the order in which they were applied. The first two measurement sets with 400 nm/s gave the same *E* value as the other 400 nm/s set, which was made nearly two hours later, indicating that the measurement duration does not affect *E*, at least within a period of several hours.

On the other hand, *E* clearly depends on the indentation velocity. Figure 6b shows the same data sorted according to indentation velocity. The Young's modulus *E* increases continuously from ≈10 kPa for 50 nm/s up to ≈60 kPa for 2000 nm/s. Previously, Francius et al. also showed that *E* of PLL/HA films was influenced by the indentation velocity [5]. They reported a nearly constant *E* below 500 nm/s, except for some fluctuations, and above that an increasing *E* with increasing indentation velocity. In the current study, *E* increases continuously even when the indentation velocity is below 500 nm/s. The difference between the two studies can be explained by differences in the nature of the films (due to the molecular weight of the polymers, preparation conditions, etc.) as well as in the measurement and data-handling procedures.

The dependence of *E* on the indentation velocity clearly indicates a viscoelastic film character [23,41] as will be further discussed below in terms of stress-relaxation measurements. In fact, due to the viscoelastic liquid character of the PLL/HA films, their equilibrium *E* is expected to be very low if not zero [23]. The Young's modulus reported in this study refers to the *apparent* Young's modulus, which arises as a reaction to a sudden load. Despite this fact, determining an apparent *E* on the time scale of cellular processes (<500 nm/s) gives a first insight into the film's mechanical properties and makes a comparison of different surfaces possible [23]. Therefore, a fixed indentation velocity of 400 nm/s was used for the elasticity measurements in the following section.

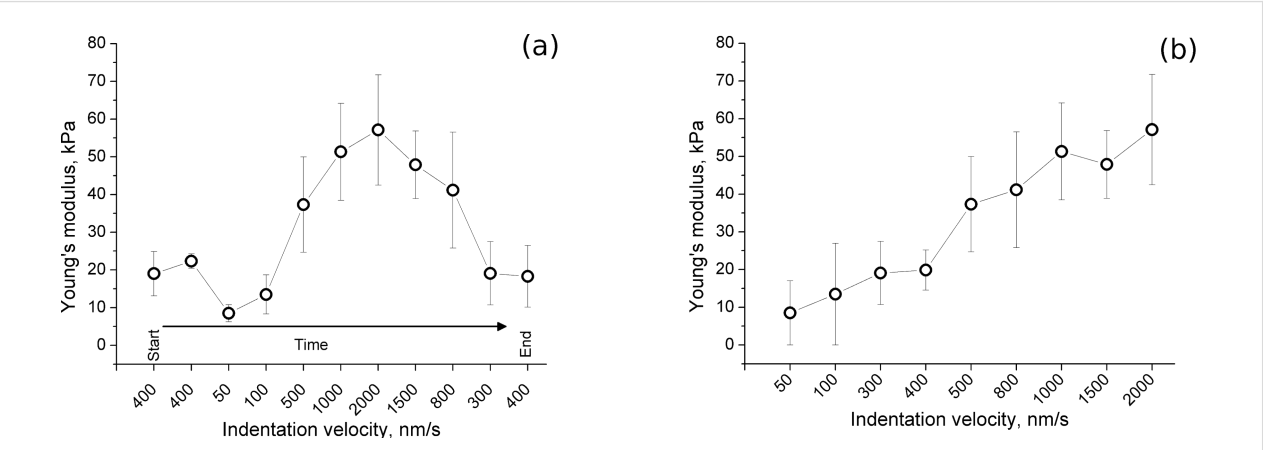

Figure 6: Calculated Young's modulus *E* for $(PLL/HA)_{72}$ with different indentation velocities sorted by (a) the measurement order (b) indentation velocity. Error bars show the standard deviation from 15 measurements.

Effect of film thickness and indenter size on Young's modulus

Small indenters, e.g., AFM tips, prevent errors due to insufficient indentation forces [15] and allow all scanning and force measurements to be done without changing any parameters [13,17,21]. On the other hand, shape and size determination of AFM tips is not straightforward, tips are more vulnerable to deformation during the measurements, and they can apply high loads on the film, invalidating the Hertzian assumptions [4,37,43]. A larger indenter, on the other hand, is advantageous due to the ease of size determination and attachment to the cantilever. In this study, tips were used for determination of the film thickness while colloidal probes were preferred in mechanical measurements. One way or another, the size of the indenter should not change the calculated Young's modulus E since the indenter radius R is already included in Equation 2. Two colloidal probes that are 2.37 and 3.35 μm in radius R were compared in order to ensure the applicability of Equation 2 for the studied systems. The apparent E as a function of film bilayer number n and probe radius R is presented in Figure 7 showing indeed no systematic dependence of E on the indenter radius R.

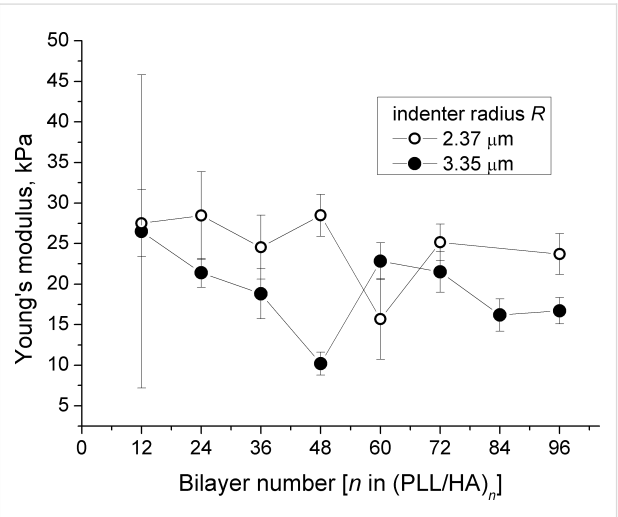

Figure 7: Young's modulus E as a function of $(PLL/HA)_n$ bilayer number n for an indenter radius of 2.37 μm (open circles) and 3.35 μm (closed circles). Neither n nor R seem to affect the calculated E in a systematic way at an indentation velocity of 400 nm/s.

Figure 7 suggests also that E does not change in a systematic way with the film bilayer number n. This result contradicts the results of Richert et al. [4], who reported a decrease in E from 90 kPa for n = 20 to 40 kPa for n = 60. The authors suggested a possibility of film softening due to greater hydration of the upper layers rather than due to a change in surface structure, heterogeneity or roughness. The film thickness range reported in the mentioned study [4] is nearly threefold larger than ours, which may result in significant structural differences.

Despite some fluctuations, E in our study ranges between 10 and 40 kPa, being in the same order of magnitude with the previous reports [3-5,7,44,45]. Only the film with 12 bilayers shows a very high variation in Young's modulus, from one measurement position to the other. This can be explained by inhomogeneities on the film surface, which will affect the local film thickness dramatically compared to the thicker films. The determination of E is very sensitive to the thickness h for thinner films, as Equation 2 suggests: For larger h values, the correction terms on the right side approach 1, making Equation 2 independent of h. Another problem with thinner films is the difficulty of extracting the F versus δ data for the indentation depth at 5–10% of the total thickness. For very thin films, this range was extended up to 20% of the total thickness, restricted to cases for which the calculated E does not change abruptly.

Relaxation time measurements

Measuring the viscoelastic properties of films in the thin-film limit can be challenging due to the requirement of precise information on the surface charge as well as on the film thickness, indentation depth and the indenter shape/size [41]. In the current set of measurements, a colloidal probe indenter was driven into the PLL/HA film with a constant velocity, but unlike a standard elasticity measurement, the indenter was not retracted instantly after the initial indentation. It was left to dwell in the film while the z-piezo drive was paused. The AFM detector continued collecting the cantilever deflection data F as a function of time t. The dwell time was set to 40 s. A summary of this process is shown in Figure 8 and Figure 9: The colloidal probe indents into the film first due to the fast, short-time z-piezo driving ("Indentation" in Figure 9, mostly elastic response) and then due to a slow, long-time relaxation of the cantilever stress ("Relaxation" in Figure 9, viscous response). After 40 s ("Dwell Time" in Figure 9) the cantilever was retracted from the film, as in the case of a standard elasticity measurement ("Retraction" in Figure 9). Data points were collected every 0.1 s.

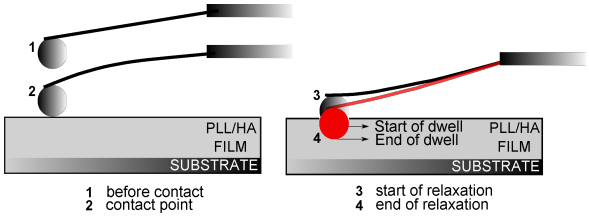

Figure 8: A summary of the stress-relaxation measurements. "Start of dwell" is the indentation depth at which the driving of the z-piezo is stopped. "End of dwell" shows the maximum indentation depth caused by the relaxation on the cantilever.

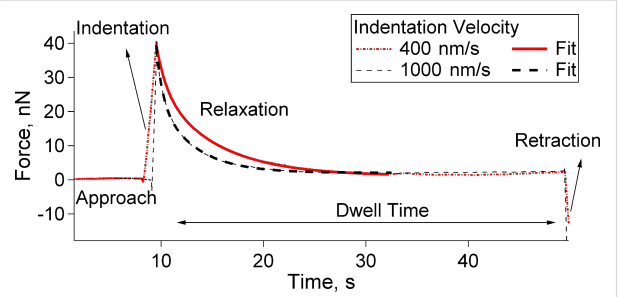

Figure 9: Stress relaxation curve with a dwell time of 40 s. The biexponential fit is represented by the thick lines on the decaying part of the force curve, in the dwell regime.

Typical F versus t relaxation curves on $(PLL/HA)_{72}$ are presented in Figure 9 for two initial indentation velocities of 400 nm/s and 1000 nm/s. Regardless of the velocity, the cantilever stress relaxes totally to zero force. A total relaxation means that the film behaves as a viscoelastic liquid having either zero or very small equilibrium elasticity, in agreement with the results of reflection-interference-contrast microscopy from Picart and co-workers [23].

Figure 9 shows that for a fixed maximum load of 40 nN, the force decays slower for 400 nm/s than for 1000 nm/s. This behaviour was observed for other indentation velocities as well, indicating that the PLL/HA films act less viscously in the case of a faster initial indentation. In contrast to that, the instantaneous elastic response of the films is stronger for a faster indentation as shown in Figure 6.

For a quantitative comparison of the stress-relaxation behaviour for different initial indentation velocities, a multiexponential decay fit can be used as the model details can be found elsewhere [13,14]. Although not shown, an exponential decay with one relaxation component was unable to fit the curves. Therefore, the force relaxation was described with a biexponential decay with the assumption that the studied multilayer films have one short and one long relaxation scale, one corresponding to the dynamics of individual segments and one to the collective dynamics of many chains. For a spherical indenter of 3.35 μm in radius, two relaxation times τ_1 and τ_2 were calculated by fitting the F versus t curves by the biexponential decay below [13,14]:

$$F = F_0 + A_1 \exp\left\{\frac{-(t-t_0)}{\tau_1}\right\} + A_2 \exp\left\{\frac{-(t-t_0)}{\tau_2}\right\}, \quad (1)$$

where t_0 is the initial time and F is the force exerted on the cantilever at a time t. F_0 corresponds to the elastic component of the relaxation. The fits are shown as thick lines in Figure 9. It should be noted that this fit is much simpler compared to the creep-compliance function [41,46], it resembles the stress relaxation fit used for heterogeneous materials [13,14], and allows for a qualitative comparison of the cantilever relaxation time rather than giving the actual material relaxation time.

Before discussing the outcome of the fits, two possible corrections should be mentioned. The first is the elimination of the effect of limited film thickness in the elastic component [13,14] F_0, as discussed above. In an attempt, the effect of the limited film thickness was included in the fit function by using Equation 2; however, probably due to the small equilibrium value of the elastic component [23] F_0, this correction did not result in a significant change in the fit, aside from making the fit function extremely complicated. The second possible correction can be performed to normalize the contact area of the colloidal probe during the stress relaxation, since the biexponential formula given above assumes a constant contact area during the process [13,14]. On the contrary, the contact area in our stress relaxation measurements changes as the colloidal probe moves deeper into the film. The change in the area can be calculated by using the (indentation-depth)–(time) relation, and the detected force at a certain time can be divided by the instantaneous contact area. Although the contact area may change by a maximum of 40% during the stress relaxation, this change results in an error in τ_1 and τ_2 that is much smaller than the uncertainty of the measurements, and thus the contact area was assumed to be constant for the sake of simplicity.

The effect of the initial indentation velocity on τ_1 and τ_2 was studied with velocities of 100, 400, 1000, 2000, 4000 and 6000 nm/s at a maximum initial load of 40 nN. Ten measurements on different lateral positions were performed for each indentation velocity. The calculated τ_1 and τ_2 are presented in Figure 10.

Figure 10 shows a very clear dependence of τ_1 and τ_2 on the indentation velocity, indicating that PLL/HA films have a viscoelastic character, as has been previously suggested for similar multilayers [4,5,23,32], crosslinked PDMS films [47], human platelets [48], agar gel [49] and cancer cells [13,14]. From 100 nm/s to 6000 nm/s, τ_1 continuously decreases from 2.17 s to 0.18 s and τ_2 decreases from 9.03 s to 1.20 s. This decrease means that the film is less viscous when it is exposed to a faster initial load and gives a hint about a non-Newtonian, shear-thinning behaviour [50].

Conclusion

Mechanical properties of layer-by-layer assembled PLL/HA films with varied bilayer number were studied by scanning- and colloidal-probe atomic force microscopy. Detailed measure-

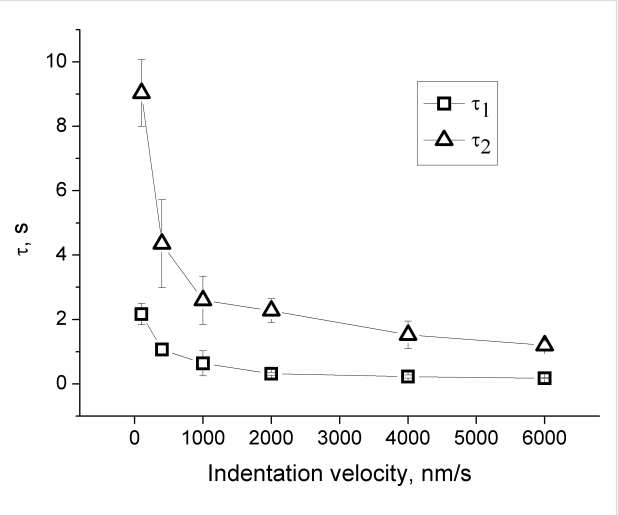

Figure 10: The cantilever's stress relaxation time τ_1 and τ_2 as a function of the initial indentation velocity. Error bars indicate the standard deviation from 10 measurements.

ment and data analysis techniques were addressed. Two independent AFM-based methods were used to measure the film thickness: Scratch-and-scan and full-indentation. Film thickness depends linearly on the bilayer number and ranges from ~400 nm to ~7500 nm for 12 and 96 bilayers, respectively. The apparent Young's modulus of the films ranges between 15 to 40 kPa and the thinner films present larger error bars presumably due to the inhomogeneity of the surface. Film thickness and indenter size have no significant effect on the apparent Young's modulus providing that the film-thickness-corrected Hertzian model is used to analyse the AFM force data. Multiple indentations at a fixed lateral film position can trigger a viscous or plastic deformation, continuously softening the structure.

Regardless of their thickness, PLL/HA films show a viscoelastic liquid behaviour. This is evidenced by the fact that the apparent Young's modulus increases with indentation velocity and the cantilever stress relaxes to zero force after a while. Stress relaxation measurements show a biexponential decay indicating two relaxation processes, one due to the individual multilayer segments and the other to the collective film dynamics. Both relaxation times decrease with increasing initial indentation velocity, suggesting a non-Newtonian, shear-thinning fluid character. Frequency-dependent AFM force [41] and quartz crystal microbalance measurements are planned for a better understanding of shear and friction effects on the mechanical response of polymeric films.

Experimental

Preparation of polyelectrolyte films

The polyelectrolyte films PEI–(HA/PLL)$_n$–HA, where n represents the number of deposited polymer pairs, were prepared by the layer-by-layer (LbL) technique [1] using a dipping robot (Riegler & Kirstein GmbH, Germany). The films were deposited on microscopy cover glasses (14 mm in diameter, Marienfeld GmbH, Germany). Before deposition, the glass slides were cleaned by consecutive incubation in hot solutions (60 °C) of 2% (w/v) Hellmanex (Hellma GmbH, Germany), 0.01 M sodium dodecyl sulphate, and 0.1 M HCl during 15 min for each solution followed by multiple rinsing with pure water. The film build-up was pursued at 25 °C by alternating dipping of the glass slides into PLL and HA solutions (0.5 mg/mL in 10 mM Tris-buffer containing 15 mM NaCl, pH 7.4) over 10 min with an intermediate washing step with the buffer (10 min). As with the precursor layer, PEI was adsorbed under the same conditions as other polymers. Before use, polyelectrolyte solutions were filtered through a 0.22 μm filter. The films were stored in Tris-buffer containing 0.15 M NaCl at 4 °C and never allowed to dry during the measurements.

Microsphere attachment to the cantilever

Silica particles with a radius of 3.35 or 2.37 μm (Bangs Laboratories, Inc., USA) were used as the indenting probes. The probes were glued on CSC12 cantilevers (μMasch, USA). Before the attachment process, the tipless cantilevers were cleaned in a plasma chamber for 20 min in order to get rid of any organic contaminants on their surface so that the glue spreads and adheres better. A two-component epoxy adhesive (UHU plus endfest 300, UHU GmbH, Bühl, Germany) was prepared and stored for 20 min under ambient conditions so that it is less fluid and easier to handle. The silica particles and the adhesive were placed on a glass slide. The adhesive was apportioned in fine stitches with a needle of a syringe or a very thin metal wire so that it had small separate droplets, ideally the size of the silica spheres. The cantilever was then moved by using a micro-manipulator. The far end of the cantilever was brought into contact with a drop of adhesive and finally was brought to a soft contact with an individual silica particle, leading to the attachment. After a successful attachment, the cantilever was stored for 24 h in ambient conditions. Finally, before each use, cantilevers were cleaned in an air plasma chamber for another 20 min.

Cantilever calibration

Before each set of force measurements, the cantilever was calibrated in the medium (air, water, buffer, etc...) where the measurements were to be performed. This was done by bringing the cantilever into contact with a rigid surface and driving it further down by the piezo unit for a known distance. As there could not be any indentation on a hard surface, the driving distance was equal to the deflection of the cantilever. This step is crucial both for determination of the spring constant and for indentation measurements. Although the spring constant k_c was

given by the manufacturer as 0.05 N/m (unless stated otherwise), its exact value was determined before each measurement by the thermal noise method, which is a built-in procedure in the MFP-3D instrument (Asylum Research, CA, USA).

Thickness measurements

Two independent methods were performed with an MFP-3D AFM instrument (Asylum Research, CA, USA) in order to determine the thickness of the polyelectrolyte films. All measurements were performed as close to the centre of the substrate as possible to avoid any effect of film inhomogeneity.

Scratch-and-scan method

A part of the PEM film was removed by scratching it with a sharp needle or tweezers. Scratching the film in one direction with the correct angle led to a very small rim on one side of the scratch, as shown in Figure 1. The scratched area was imaged in the buffer medium with a magnetically driven *iDrive*™ setup by using AR-iDrive-N01 (Asylum Research, USA) cantilevers. Scan rate was fixed to 0.1 Hz on a 90 μm × 90 μm area. The thickness of the film was calculated by using a cross-section profile on the AFM micrograph. Three different regions were scanned and up to five cross sections per image were used to obtain an average thickness. The three regions on the scratch were selected so that the distance between them was around 2 mm.

Full-indentation method

An AFM force measurement setup was used for indentation in the *z*-direction. The optical lever sensitivity was determined on a hard surface before any measurements on the soft films. μMasch CSC37 cantilevers with a high spring constant $k_c \approx$ 0.3 N/m and a pyramidal tip with a length of 20–25 μm were used. The aim was not to measure any meaningful force but to penetrate through the film down to the hard substrate. When the tip reached the hard surface, the slope in the F versus δ curve changed drastically as indentation was no longer possible, as Figure 2 shows. The distance between the tip–film contact point and the tip–substrate contact point simply gave the thickness of the film. An average thickness was calculated over 10 to 20 measurements on different lateral positions.

Elasticity measurements

Force measurements were performed with a commercial MFP-3D instrument (Asylum Research, CA, USA) by using the calibrated cantilevers with an attached silica microsphere. All measurements on $(PLL/HA)_n$ films were performed in a tris-buffer environment containing 0.1% NaN_3 in order to prevent a bacterial contamination and a possible damage through drying. Force measurements were made with an indentation velocity of 50–6000 nm/s. The applied load on the films was controlled by a trigger point, which was set to a relative deflection corresponding to a 1.50 V detector signal after the contact. Unless stated otherwise, the lateral position of the measurement was changed after each indentation and the next measurement was taken on a spot around 50 to 100 μm away from the first one. Fifteen to thirty different spots were chosen in the central area of the substrate to calculate an average Young's modulus E.

E was calculated from the raw F versus δ data by a thickness-corrected Hertzian model presented by Dimitriadis et al. [37]. The Young's modulus of a film bonded to a hard substrate is given as

$$E = \frac{9F}{16}\frac{1}{R^{1/2}\delta^{3/2}} \times \frac{1}{\left(1+1.133\chi+1.283\chi^2+0.769\chi^3+0.0975\chi^4\right)} \quad (2)$$

where F is the force exerted on the surface at an indentation depth δ, R is the radius of the spherical indenter and $\chi = \sqrt{R\cdot\delta/h}$, h being the film thickness. The Poisson's ratio was set to 0.5 for the PEM films due of their high water content, meaning that they are incompressible [37].

The above relation requires a precise determination of h and R. The thickness of the films h was determined by AFM as discussed above. The radius of the indenter curvature R was extracted from scanning electron microscope (SEM) images taken in TU Berlin ZELMI, with a high-resolution field emission microscope (S4000, Hitachi, acceleration field of 10 kV and 20 kV, no gold-coating).

Raw F versus δ data measured by CP-AFM were analysed by home-written procedures using Igor Pro software package (Wavemetrics Inc., USA). Average Young's modulus was calculated in the indentation region $0.05h \leq \delta \leq 0.2h$. The probe–sample contact point was taken as the "jump-to-contact" of the probe to the surface. This point is shown in Figure 4a.

Viscoelasticity measurements were performed with some principle differences and the details are given in the Discussion section.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (VO1716-2/1, KL 1165-11/1, KL 1165-12/1) and Alexander von Humboldt Foundation for the financial support.

References

1. Decher, G.; Hong, J. D.; Schmitt, J. *Thin Solid Films* **1992,** *210-211,* 831–835. doi:10.1016/0040-6090(92)90417-A
2. Decher, G. *Science* **1997,** *277,* 1232–1237. doi:10.1126/science.277.5330.1232
3. Engler, A. J.; Richert, L.; Wong, J. Y.; Picart, C.; Discher, D. E. *Surf. Sci.* **2004,** *570,* 142–154. doi:10.1016/j.susc.2004.06.179
4. Richert, L.; Engler, A. J.; Discher, D. E.; Picart, C. *Biomacromolecules* **2004,** *5,* 1908–1916. doi:10.1021/bm0498023
5. Francius, G.; Hemmerlé, J.; Ohayon, J.; Schaaf, P.; Voegel, J.-C.; Picart, C.; Senger, B. *Microsc. Res. Tech.* **2006,** *69,* 84–92. doi:10.1002/jemt.20275
6. Richert, L.; Boulmedais, F.; Lavalle, P.; Mutterer, J.; Ferreux, E.; Decher, G.; Schaaf, P.; Voegel, J.-C.; Picart, C. *Biomacromolecules* **2004,** *5,* 284–294. doi:10.1021/bm0342281
7. Schneider, A.; Francius, G.; Obeid, R.; Schwinté, P.; Hemmerlé, J.; Frisch, B.; Schaaf, P.; Voegel, J.-C.; Senger, B.; Picart, C. *Langmuir* **2006,** *22,* 1193–1200. doi:10.1021/la0521802
8. Schneider, A.; Richert, L.; Francius, G.; Voegel, J.-C.; Picart, C. *Biomed. Mater.* **2007,** *2,* S45–S51. doi:10.1088/1748-6041/2/1/S07
9. Weisenhorn, A. L.; Khorsandi, M.; Kasas, S.; Gotzos, V.; Butt, H.-J. *Nanotechnology* **1993,** *4,* 106. doi:10.1088/0957-4484/4/2/006
10. Li, X.; Logan, B. E. *Langmuir* **2004,** *20,* 8817–8822. doi:10.1021/la0488203
11. Considine, R. F.; Dixon, D. R.; Drummond, C. J. *Water Res.* **2002,** *36,* 3421–3428. doi:10.1016/S0043-1354(02)00082-9
12. Mathur, A. B.; Collinsworth, A. M.; Reichert, W. M.; Kraus, W. E.; Truskey, G. A. *J. Biomech.* **2001,** *34,* 1545–1553. doi:10.1016/S0021-9290(01)00149-X
13. Moreno-Flores, S.; Benitez, R.; Vivanco, M. d.; Toca-Herrera, J. L. *Nanotechnology* **2010,** *21,* 445101. doi:10.1088/0957-4484/21/44/445101
14. Moreno-Flores, S.; Benitez, R.; Vivanco, M. d.; Toca-Herrera, J. L. *J. Biomech.* **2010,** *43,* 349–354. doi:10.1016/j.jbiomech.2009.07.037
15. Matzelle, T. R.; Ivanov, D. A.; Landwehr, D.; Heinrich, L. A.; Herkt-Bruns, C.; Reichelt, R.; Kruse, N. *J. Phys. Chem. B* **2002,** *106,* 2861–2866. doi:10.1021/jp0128426
16. Matzelle, T. R.; Geuskens, G.; Kruse, N. *Macromolecules* **2003,** *36,* 2926–2931. doi:10.1021/ma021719p
17. Junk, M. J. N.; Berger, R.; Jonas, U. *Langmuir* **2010,** *26,* 7262–7269. doi:10.1021/la903396v
18. Radji, S.; Alem, H.; Demoustier-Champagne, S.; Jonas, A. M.; Cuenot, S. *J. Phys. Chem. B* **2010,** *114,* 4939–4944. doi:10.1021/jp909819h
19. Tagit, O.; Tomczak, N.; Vancso, G. J. *Small* **2008,** *4,* 119–126. doi:10.1002/smll.200700260
20. Hashmi, S. M.; Dufresne, E. R. *Soft Matter* **2009,** *5,* 3682–3688. doi:10.1039/b906051k
21. Burmistrova, A.; Richter, M.; Uzum, C.; von Klitzing, R. *Colloid Polym. Sci.* **2011,** *289,* 613–624. doi:10.1007/s00396-011-2383-2
22. Burmistrova, A.; Richter, M.; Eisele, M.; Üzüm, C.; von Klitzing, R. *Polymers* **2011,** *3,* 1575–1590. doi:10.3390/polym3041575
23. Picart, C.; Senger, B.; Sengupta, K.; Dubreuil, F.; Fery, A. *Colloids Surf., A* **2007,** *303,* 30–36. doi:10.1016/j.colsurfa.2007.02.015
24. Boudou, T.; Crouzier, T.; Ren, K.; Blin, G.; Picart, C. *Adv. Mater.* **2010,** *22,* 441–467. doi:10.1002/adma.200901327
25. Lo, C.-M.; Wang, H.-B.; Dembo, M.; Wang, Y. *Biophys. J.* **2000,** *79,* 144–152. doi:10.1016/S0006-3495(00)76279-5
26. Touhami, A.; Nysten, B.; Dufrêne, Y. F. *Langmuir* **2003,** *19,* 4539–4543. doi:10.1021/la034136x
27. Hurley, D. C.; Kopycinska-Müller, M.; Kos, A. B. *JOM* **2007,** *59,* 23–29. doi:10.1007/s11837-007-0005-8
28. Xu, X.; Raman, A. *J. Appl. Phys.* **2007,** *102,* 034303. doi:10.1063/1.2767202
29. Li, J. K.; Sullan, R. M. A.; Zou, S. *Langmuir* **2011,** *27,* 1308–1313. doi:10.1021/la103927a
30. Lavalle, P.; Picart, C.; Mutterer, J.; Gergely, C.; Reiss, H.; Voegel, J.-C.; Senger, B.; Schaaf, P. *J. Phys. Chem. B* **2004,** *108,* 635–648. doi:10.1021/jp035740j
31. von Klitzing, R. *Phys. Chem. Chem. Phys.* **2006,** *8,* 5012–5033. doi:10.1039/b607760a
32. Picart, C.; Lavalle, P.; Hubert, P.; Cuisinier, F. J. G.; Decher, G.; Schaaf, P.; Voegel, J.-C. *Langmuir* **2001,** *17,* 7414–7424. doi:10.1021/la010848g
33. Picart, C.; Mutterer, J.; Richert, L.; Luo, Y.; Prestwich, G. D.; Schaaf, P.; Voegel, J.-C.; Lavalle, P. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 12531–12535. doi:10.1073/pnas.202486099
34. Porcel, C.; Lavalle, P.; Ball, V.; Decher, G.; Senger, B.; Voegel, J.-C.; Schaaf, P. *Langmuir* **2006,** *22,* 4376–4383. doi:10.1021/la053218d
35. Porcel, C.; Lavalle, P.; Decher, G.; Senger, B.; Voegel, J.-C.; Schaaf, P. *Langmuir* **2007,** *23,* 1898–1904. doi:10.1021/la062728k
36. Hertz, H. *Journal für die reine und angewandte Mathematik* **1882,** *92,* 156–171. doi:10.1515/crll.1882.92.156
37. Dimitriadis, E. K.; Horkay, F.; Maresca, J.; Kachar, B.; Chadwick, R. S. *Biophys. J.* **2002,** *82,* 2798–2810. doi:10.1016/S0006-3495(02)75620-8
38. Lin, D. C.; Dimitriadis, E. K.; Horkay, F. *J. Biomech. Eng.* **2007,** *129,* 430–440. doi:10.1115/1.2720924
39. Lin, D. C.; Horkay, F. *Soft Matter* **2008,** *4,* 669–682. doi:10.1039/b714637j
40. Lin, D. C.; Dimitriadis, E. K.; Horkay, F. *J. Biomech. Eng.* **2007,** *129,* 904–912. doi:10.1115/1.2800826
41. Attard, P. *J. Phys.: Condens. Matter* **2007,** *19,* 473201. doi:10.1088/0953-8984/19/47/473201
42. Butt, H.-J.; Cappella, B.; Kappl, M. *Surf. Sci. Rep.* **2005,** *59,* 1–152. doi:10.1016/j.surfrep.2005.08.003
43. Kopycinska-Müller, M.; Geiss, R. H.; Hurley, D. C. *Ultramicroscopy* **2006,** *106,* 466–474. doi:10.1016/j.ultramic.2005.12.006
44. Collin, D.; Lavalle, P.; Garza, J. M.; Voegel, J.-C.; Schaaf, P.; Martinoty, P. *Macromolecules* **2004,** *37,* 10195–10198. doi:10.1021/ma048683g
45. Picart, C.; Sengupta, K.; Schilling, J.; Maurstad, G.; Ladam, G.; Bausch, A. R.; Sackmann, E. *J. Phys. Chem. B* **2004,** *108,* 7196–7205. doi:10.1021/jp037971w
46. Attard, P. *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.* **2001,** *63,* 061604. doi:10.1103/PhysRevE.63.061604
47. Gillies, G.; Prestidge, C. A.; Attard, P. *Langmuir* **2002,** *18,* 1674–1679. doi:10.1021/la011461g
48. Radmacher, M.; Fritz, M.; Kacher, C. M.; Cleveland, J. P.; Hansma, P. K. *Biophys. J.* **1996,** *70,* 556–567. doi:10.1016/S0006-3495(96)79602-9
49. Tyrrell, J. W. G.; Attard, P. *Langmuir* **2003,** *19,* 5254–5260. doi:10.1021/la0207163
50. Rubinstein, M.; Colby, R. *Polymer Physics;* Oxford University Press: Oxford, 2003.

Beilstein Journal of Nanotechnology

Spring constant of a tuning-fork sensor for dynamic force microscopy

Dennis van Vörden*, Manfred Lange, Merlin Schmuck, Nico Schmidt and Rolf Möller

Full Research Paper

Open Access

Address:
Faculty of Physics, University of Duisburg-Essen, Lotharstr. 1–21 47048 Duisburg, Germany

Email:
Dennis van Vörden* - dennis.vanvoerden@uni-due.de

* Corresponding author

Keywords:
atomic force microscopy; finite element method; spring constant; thermal fluctuation; tuning fork

Beilstein J. Nanotechnol. **2012,** *3,* 809–816.
doi:10.3762/bjnano.3.90

Received: 16 August 2012
Accepted: 05 November 2012
Published: 29 November 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We present an overview of experimental and numerical methods to determine the spring constant of a quartz tuning fork in qPlus configuration. The simple calculation for a rectangular cantilever is compared to the values obtained by the analysis of the thermal excitation and by the direct mechanical measurement of the force versus displacement. To elucidate the difference, numerical simulations were performed taking account of the real geometry including the glue that is used to mount the tuning fork.

Introduction

Quartz tuning forks provide excellent self-sensing probes in scanning probe microscopy, offering several advantages compared to the standard microfabricated silicon-based cantilevers [1,2]. Frequency-modulation atomic force microscopy (FM-AFM) with a tuning-fork sensor has had a major impact on fundamental and scientific research, e.g., by resolving the structure of a molecule [3] or even determining the structure of an unknown organic molecule [4].

In FM-AFM, the motion of the sensor is given in very good approximation by a harmonic oscillator. For the limit of small amplitudes the measurement of the frequency shift provides the average force gradient caused by the interaction between the tip and sample surface, according to

$$\left\langle k_{TS}(z)\right\rangle = 2k\frac{\Delta f}{f_0}, \qquad (1)$$

where $\left\langle k_{TS}(z)\right\rangle$ is the average force gradient between tip and sample, Δf is the frequency shift, k is the spring constant of the sensor and f_0 is the resonance frequency of the sensor without interaction with the sample.

While the resonance frequency may be measured accurately in the experiment, it is more difficult to evaluate the spring constant k. However, the latter is required to evaluate the force gradient and other physical quantities, e.g., the energy dissipated due to the interaction between tip and sample.

For microfabricated cantilevers several methods to evaluate k have been demonstrated. The most obvious, although technically difficult, method is to measure the static deflection as a function of the applied force [5-11]. If an additional mass is attached to the cantilever the spring constant can be determined by the change of the resonance frequency [12-14]. By measuring the amplitude of the thermal noise, k can be evaluated in situ, e.g., in a vacuum system prior to the measurement, without any modification of the experimental arrangement [15-20]. However, this requires a good signal-to-noise ratio for the measurement of the beam deflection.

For the analysis of data obtained by tuning-fork sensors k is often simply calculated based on the formula for a cantilever beam [21]. In principal the experimental techniques mentioned above can be applied for tuning forks sensors as well [22-24].

In the present paper we compare the results for the determination of the spring constant of tuning fork sensors in the qPlus configuration [1,2] based on the following methods: a simple calculation for a cantilever beam; the measured deflection as a function of the applied force; the thermal noise; and a numerical simulation by the finite-element method.

Result and Discussion

Calculation for a rectangular beam

The formula for the spring constant of a beam that is clamped on one side is

$$k = \frac{E \cdot w \cdot \tau^3}{4 \cdot L^3}, \qquad (2)$$

where E is the Young's modulus (for quartz), τ is the thickness, w the width, and L the length of a prong. For the cantilevers used in our experiments, E = 78.7 GPa, t = 0.41 ± 0.003 mm, w = 0.24 ± 0.003 mm, and L = 2.94 ± 0.003 mm, yielding a value of

$$k = 12840 \pm 410\,\mathrm{N/m}. \qquad (3)$$

Experimental evaluation of the spring constant

Beam deflection as function of applied force

To a good approximation, the force F exerted by the tuning fork (TF) in the qPlus configuration is given by Hooke's law $F = -kz$, with the spring constant k and the deflection z.

For the experiment, the TFs are glued to a holder in exactly the same way as for the low-temperature noncontact AFM developed in our group [25]. To apply a force on the TF a loop is formed by a thin wire, which is hooked as far as possible to the end of the free prong (Figure 1). Two weights with masses of m_1 = 14.5 g and m_2 = 19.3 g are used to apply the force. The measurements are performed at room temperature. The runs of loading and unloading are repeated three times for each TF. The deflection is monitored by a CCD-chip (Sony ICD098BQ Color: 640 × 480, 4.5 mm diagonal, 5.6 µm × 5.6 µm) and an objective lens (Cosmicar Pentax TV lens, f = 16 mm, 1:1.4) providing an optical resolution of about 7 µm. The images are calibrated by using a scale. By comparing the images for the loaded and unloaded TF, the deflection at the end of the prong is determined, using WSxM [26] and Corel Draw X5 (Corel Corporation, Ottawa, Canada).

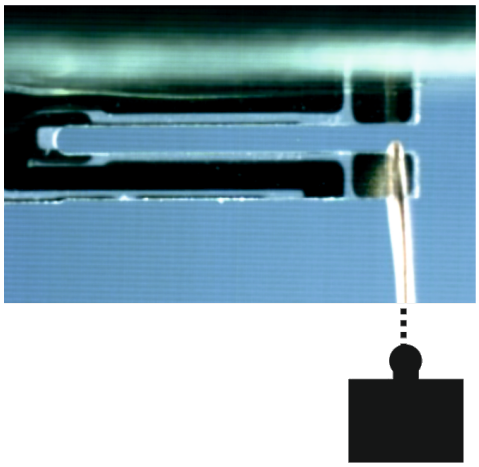

Figure 1: Setup for evaluating the spring constant by Hooke's law. The deflection is measured by analyzing differences between images of the loaded and unloaded TF using WSxM [26].

Two different kind of glues were used. The results for a set of 24 TFs glued to the holder by Torr Seal® (ThorLabs GmbH, Dachau, Germany) are plotted in Figure 2. Similarly a set of 16 TFs glued by UHU plus endfest 300® (UHU GmbH, Bühl, Germany) was measured. Both glues were prepared accurately, according to the corresponding recipe. Care was taken to use approximately the same amount of glue. The results of an analysis of both data sets are given in Table 1.

The results are valid for a point of application of the force that is about 0.3 mm from the end of the prong. According to the numerical calculations by the method of finite elements presented in the following, these values can be extrapolated for the end of the prong by a reduction of 1090 and 910 N/m for Torr Seal and UHU endfest, respectively. The numbers are given in the right column. The values of the spring constants for both glues are significantly smaller than the one calculated for

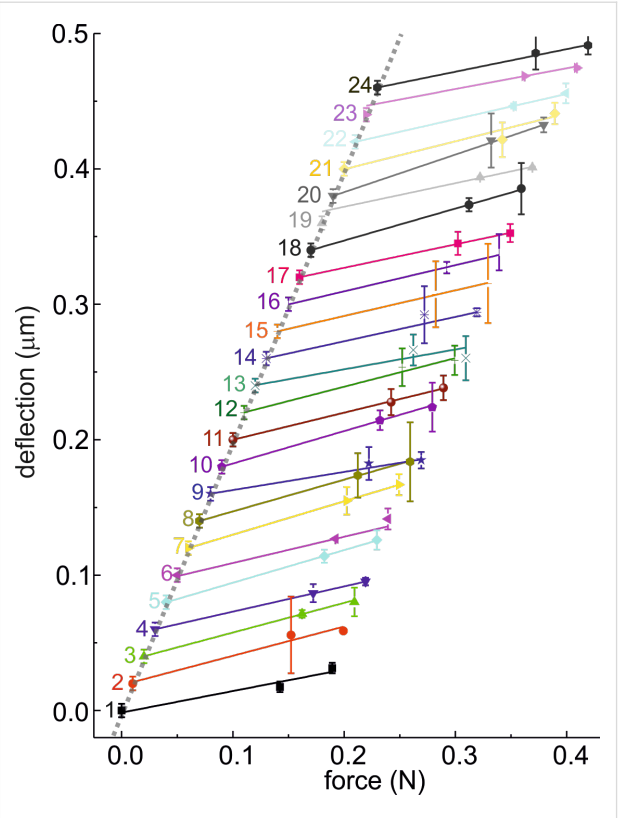

Figure 2: Deflection versus force for 24 TFs using Torr Seal. The individual lines are displaced for clarity.

Table 1: Spring constants measured by Hooke's law. k refers to the position at which the load is applied, while k_{end} is obtained by extrapolation for the end of the free prong.

	k (N/m)	k_{end} (N/m)
Torr Seal	9280 ± 960	8190 ± 960
UHU endfest	7500 ± 1520	6590 ± 1520

the rectangular beam. The rather large difference between the two types of glues clearly demonstrates that, at least at room temperature, the glue has a major impact on the effective spring constant, although the thickness of the layer of glue is only about 0.04–0.06 mm.

Amplitude of the thermal fluctuations

According to the equipartition theorem a thermal energy of $\frac{1}{2}k_BT$ will be dissipated in every energetically accessible degree of freedom of a system [27]. The tuning fork in the qPlus configuration has one degree of freedom:

$$\frac{1}{2}k\left\langle A_{\text{therm}}^2\right\rangle = \frac{1}{2}k_BT \Leftrightarrow k = \frac{k_BT}{\left\langle A_{\text{therm}}^2\right\rangle}, \qquad (4)$$

where k is the spring constant, k_B is the Boltzmann's constant, T is the temperature in Kelvin and A_{therm} the amplitude of thermal deflection of one prong.

The experiments were performed in situ by using a home-built low-temperature tuning-fork AFM (LT-TF-AFM) [25]. Figure 3a shows the carrier onto which the TFs are mounted in the qPlus configuration, i.e., the top side of one prong is glued to a relatively massive part. Based upon initial results for all further experiments Torr Seal was used as the glue. The evaluation of the spring constant was performed for a TF with a tip (e.g., a 25–100 µm tungsten wire) attached to the front face of the free prong of the tuning fork. It is connected separately through a metallic wire (e.g., a 25 µm gold wire) to collect the tunneling current, avoiding crosstalk with the frequency-shift signal.

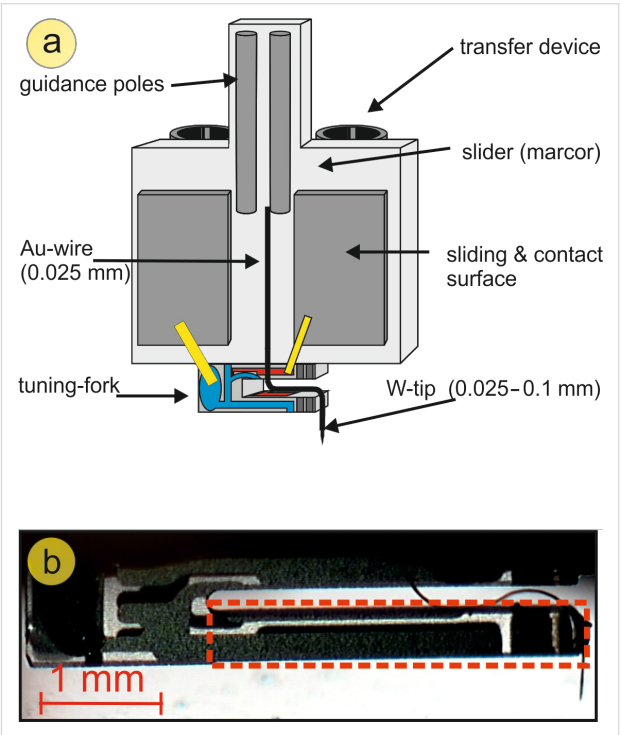

Figure 3: (a) Sketch of the qPlus configuration in our setup. (b) Picture of a mounted tuning fork. The area marked in red indicates the part of the tuning fork that is used to estimate the spring constant based on the formula for a rectangular beam.

Prior to the measurements, the sensitivity of the TF in millivolts per nanometer (mV/nm) was calibrated, including the electronics for detection. This is done in several steps. First, the z-piezo of the scanning unit is calibrated by measuring the topography of a surface (Cu(111), Ag(111) and Si(111)) with atomic steps of well-known height. Next, the amplitude of the TFs oscillation is varied, while simultaneous monitoring the change in height at constant tunneling current, and the electric

signal is detected by the TF. The inset in Figure 4 displays the corresponding graph. This was performed separately at room temperature and at 80 K. Figure 4 shows the power spectral density $s(f)$ of the thermal fluctuations in picometers squared per hertz (pm^2/Hz) of a TF at room temperature and at 80 K measured by a FFT-Analyzer (SR 760). A close inspection reveals that not only is the amplitude of the fluctuation lower at 80 K, but also the width is reduced, and the resonance frequency is slightly shifted. To evaluate the amplitude of the thermal fluctuations a Lorentzian was fitted to the spectra:

$$s(f) = \frac{\left\langle A_{\text{therm}}^2 \right\rangle}{2\pi} \frac{\Gamma}{(f - f_0)^2 + (\frac{\Gamma}{2})^2} + C\,, \qquad (5)$$

where Γ is the width, f is the frequency, f_0 is the resonance frequency and C is the offset.

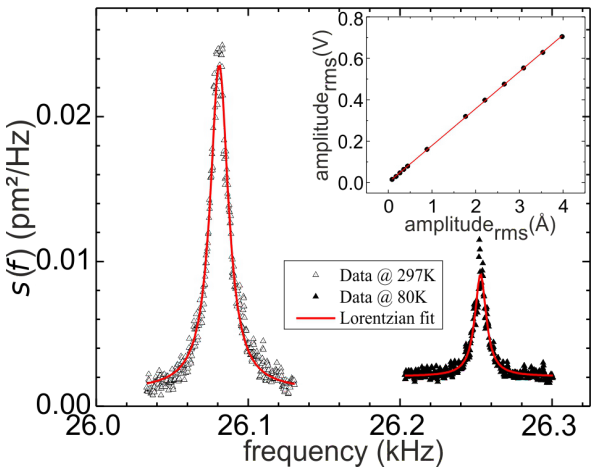

Figure 4: Power spectral density of the thermal fluctuations of a tuning fork at 300 and 80 K. The inset shows the data used to calibrate the sensitivity of the tuning fork.

According to Equation 4 the spring constants listed in Table 2 are obtained.

Table 2: Spring constants measured by thermal fluctuations. All measurements were performed with Torr Seal as the glue.

T (K)	k (N/m)
300	8000 ± 500
80	11200 ± 500

The higher spring constant at 80 K is most probably due to the increase in Young's modulus of the glue. The value at room temperature is rather close to the value obtained by the measurements using Hooke's law.

Numerical simulation using the finite element method

To get more insight into the relevant details of the qPlus configuration, numerical calculations by using the method of finite elements were performed. In contrast to the experiment this enables analysis of the influence of one specific parameter, e.g., the thickness of the glue, keeping all the others exactly the same. The simulations were performed with the commercial FEM-software COMSOL Multiphysics 3.5 (COMSOL Multiphysics GmbH, Berlin, Germany) and the additional module "Structural Mechanics". The model introduced below is simulated by the 3-D model "Solid, Stress-Strain" and statically analyzed by the module "Structural Mechanics".

A grid model is built for the TF (see Figure 5a) including the glue that is used to attach it to the rigid support. To simulate the experimental configuration as well as possible, it is assumed that the prong of the TF is partially embedded in the glue. Several microscopic images were taken to analyze the experimental geometry in detail. Figure 5b sketches the configuration and the parameters. The thickness is t = 0.04 mm, the overlap o = 0.15 mm, and the additional width at both sides w = 0.1 mm. The given data are valid for the average bonding.

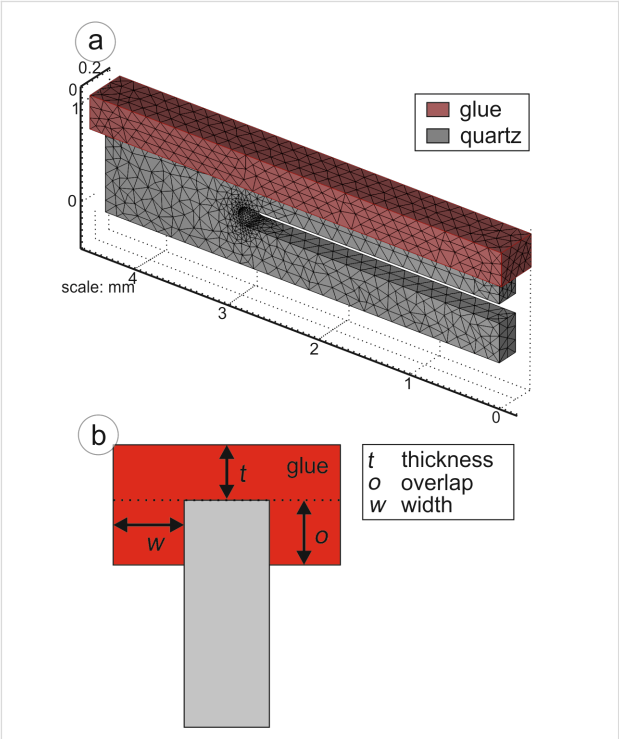

Figure 5: (a) Mesh-grid model of the tuning fork used for FEM simulations. (b) Sketch of the glue enclosing the tuning fork.

In the simulation, both materials, the glue and the tuning fork (quartz), are considered isotropic. COMSOL Multiphysics needs three different material-specific parameters: Young's modulus, the Poisson ratio and mass density. The values and the origin of these parameters can be seen in Table 3. As a boundary condition, the interface between the support and the glue is fixed in all directions. The deformation of the TF is calculated for a force applied in the *z*-direction at the end of the free prong. In an iterative procedure, a closer mesh is generated in areas of high deformation after each step, until changes between two sequential steps are marginal. In detail, adaptive mesh refinement was performed by using quadratic Lagrange elements and h-refinement. Figure 6 shows a typical result of the FEM simulation. The mesh is refined in regions of higher stress, e.g., at the link between both prongs.

At first a calculation without glue (zero thickness, overlap and width) was performed leading to a spring constant of $k = 10100$ N/m. As to be expected, this value is lower by 19% than for the rectangular beam, because the latter neglects the contribution of the area of the TF linking the two prongs.

To study the significance of the different parameters, they were varied one by one while keeping the standard values for the remaining ones. Figure 7a displays the calculated effect of a layer of glue between the prong of the TF and the support. As expected, the spring constant decreases with increasing thickness of the glue. Due to the higher Young's modulus of Torr Seal the reduction is less than for UHU endfest. For the thickness of about 0.04 mm estimated for the experimental setup by microscopic inspection, values for the spring constant for both

Table 3: Material constants used for FEM simulations. The quartz parameters are taken from COMSOL's material library, except the Young's modulus, which was taken from [21]. The glue parameters are taken from [28-30], while the Young's modulus thereof was determined by converting the Shore D strength using [31].

Material constant	SiO_2	Torr Seal	UHU endfest
Shore D hardness	—	80	70
Young's modulus (GPa)	78.7	9.39	5.54
Poisson ratio	0.17	0.45	0.35
mass density (kg/m^3)	2200	1600	1054

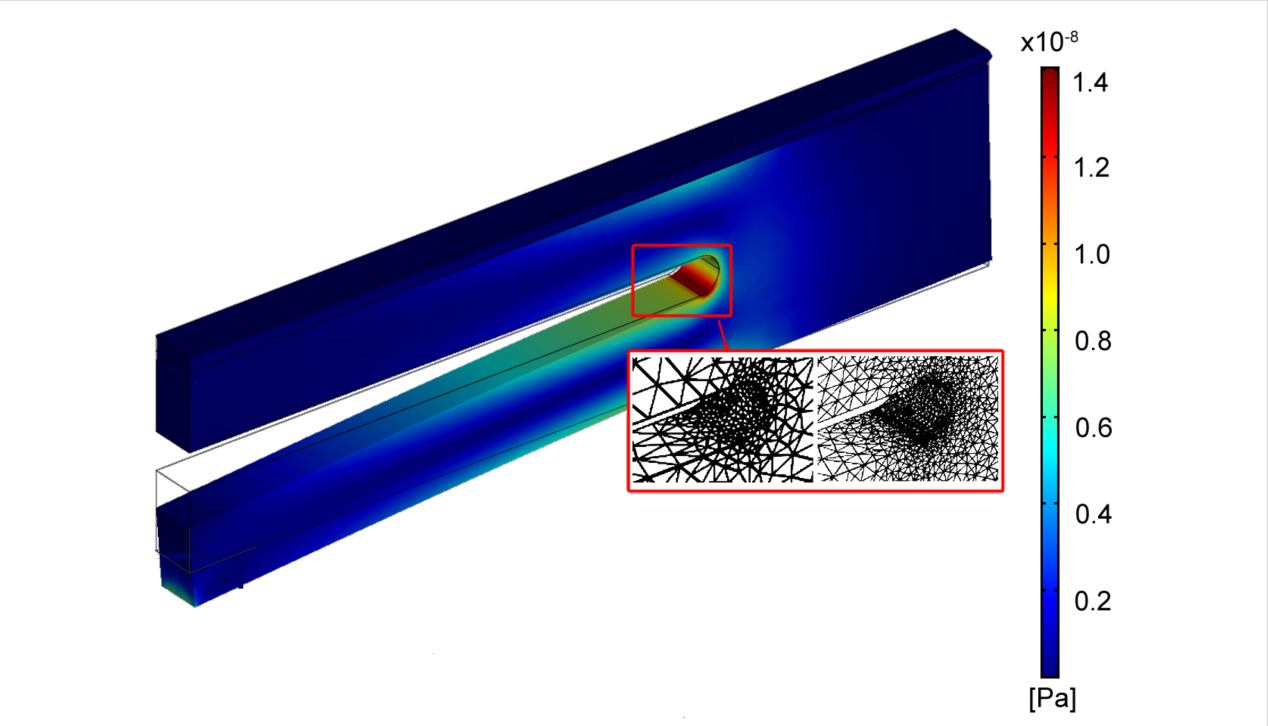

Figure 6: FEM simulation of a deflected TF. The stress is given by the color scale. The inset (red box) shows the refinement of the mesh in the stressed region.

glues were obtained that are lower than the experimental ones. The influence of the overlap and the width of the overlapping glue is displayed in Figure 7c. The spring constant increases with the width of the overlapping layer of glue up to about 0.05 mm; a further increase of the width has no significant influence. Similarly, the spring constant increases with increasing overlap.

In practice, it is rather difficult to perfectly align the TF with the support. The effect of a tilt is shown in Figure 7b. Tilting around an axis parallel to the long side of the prong results in a decrease of the spring constant that is symmetric with angle. Tilting around an axis that is perpendicular to the TF, as indicated in the figure, leads to an increase of the spring constant that is not symmetric, since either the joint or the end of the TF approaches the support.

For the experimental evaluation of the spring constant using Hooke's law, the force could not by applied at the very end of the prong. As can be seen in Figure 1, the wire is about 0.2 mm from the end. This leads to an increase of the observed spring constant since the relevant part of the prong is shorter. The resulting difference has been calculated and is displayed in Figure 7d. If the point of application of the force is at a distance of about 0.3 mm from the end, as marked by the hatched area in the figure, the measured spring constant should be about 1000 N/m higher. Based on the FEM calculation (see Figure 7d), the k value at the end of the prong can be estimated by extrapolating. This leads to a reduction of the spring constant of about 1090 N/m. The calculated values agree rather well with the ones observed in the experiment for both types of glue.

Conclusion

The comparison of the different methods to evaluate the spring constant of a TF in the qPlus configuration reveals the importance of the details of the way in which the TF is mounted. Table 4 summarizes the results. An estimate for the spring constant may be obtained by the formula for a rectangular beam

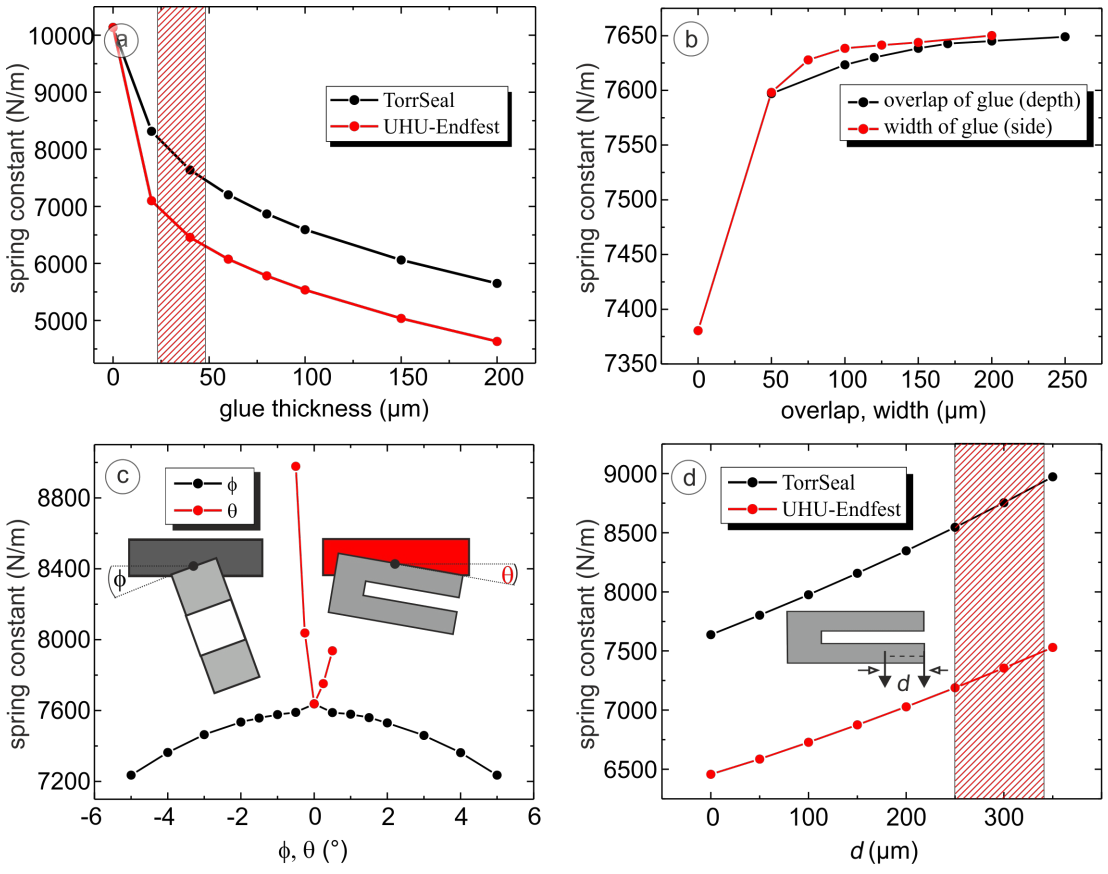

Figure 7: Results of the FEM simulations, when one parameter is varied. The standard parameters are thickness t = 0.04 mm, overlap o = 0.15 mm, width w = 0.1 mm, Φ = 0°, and θ = 0°. The parameters that are not varied are set to the standard values. (a) Spring constant with Torr Seal (black) or UHU endfest (red) as a function of the glue thickness. The hatched area marks the experimental range. (b) Spring constant as a function of the overlap and width of the glue enclosing the upper prong of the tuning fork. (c) Spring constant versus Φ (red) and versus θ (black). While the spring constant decreases symmetrically as a function of Φ, it is asymmetric as a function of θ. (d) Spring constant as a function of the point at which the force was applied for Torr Seal (black) and UHU endfest (red). The shaded area displays the experimental range.

Table 4: Comparison of the different evaluation methods.

Method	T (K)	k (N/m)	
		UHU endfest	Torr Seal
rectangular beam (no glue)	300	12840 ± 410	12840 ± 410
Hooke's law	300	6590 ± 1520	8190 ± 960
thermal fluctuation	300		8080 ± 300
thermal fluctuation	80		10950 ± 500
numerical calculation	300	6490	7650
numerical calculation with t = 0.0 mm	300	10136	10136

with the dimensions of the free prong of the TF. For the TFs used in the present experiments this yields k = 12840 N/m. However, this assumes that the one end of the TF is ideally clamped. Since the influence of the part of the TF between the prongs and the attachment to a rigid support is neglected, this value represents an upper limit. When evaluating the spring constant experimentally by measuring the deflection as a function of the applied force, the point of application has to be considered carefully. In the present experiment this was at a distance of 0.2–0.3 mm from the free end, which, according to the numerical simulations presented, leads to an increase of the observed spring constant by about 1000 N/m. The same is true if the tip of the AFM is not mounted at the very end of the prong of the TF. The values listed in Table 4 are calculated by extrapolating the measured values to the very end of the prong.

While this method could only be applied ex situ, the spring constant may be evaluated in situ by the amplitude of the thermal fluctuations. Use of a low-temperature setup enables us to analyze the influence of the temperature on the spring constant. In Table 4 the values for 300 and 80 K are given.

The values of the spring constant for both experimental techniques agree well within the experimental errors. As expected these values are lower than for the ideal beam. To analyze the cause of the difference in detail, the method of finite elements (FEM) was applied. It reveals that the value of the rectangular beam of 12450 N/m is reduced to 10136 N/m, if one prong is rigidly attached to the support. This is due to the area connecting the two prongs, which is deformed during the oscillation. This is not only important for the spring constant but also for the dissipation of the TF. Dynamic measurements have shown that applying glue to that area will reduce the quality factor of the TF by a factor of about 2. As expected the contribution of the layer of glue used to attach the TF cannot be neglected either. Not only the type of glue, but also the thickness and the embedding is important. FEM was used to study the influence of each parameter separately, which is experimentally not possible. As expected the layer of glue should be thin and the prong should be embedded a little bit to obtain a high spring constant. These results explain also the huge error in k measured during the ex situ experiments. Some tuning forks are probably glued with some small tilt or a small variation in the amount of the used glue.

Simon et al. [21] already performed FEM simulations for a tuning-fork sensor that is rather different from ours, because both prongs are free. Their calculation as well as ours shows that the area linking the two prongs substantially reduces the spring constant of the tuning fork. The qPlus configuration presented here is different from the conventional qPlus configuration used by Omicron. The results cannot be compared directly to the latter configuration. Nevertheless, it is to be expected that the effect of the glue on the spring constant will not be negligible either.

In summary, the combination of experimental techniques and numerical simulation provides insight into the contributions of various parameters to the spring constant of tuning fork sensors used for dynamical force microscopy. This is of major importance whenever quantitative values for the force gradient, the force, or the dissipated power are to be evaluated.

Acknowledgements

Financial support from the Deutsche Forschungsgemeinschaft through SFB616 "'Energy dissipation at surfaces"' is gratefully acknowledged. The authors would like to thank D. Utzat for improving the electronics.

References

1. Giessibl, F. J. *Phys. Rev. B* **1997,** *56,* 16010–16015. doi:10.1103/PhysRevB.56.16010
2. Giessibl, F. J. *Appl. Phys. Lett.* **1998,** *73,* 3956–3958. doi:10.1063/1.122948
3. Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. *Science* **2009,** *325,* 1110–1114. doi:10.1126/science.1176210
4. Gross, L.; Mohn, F.; Moll, N.; Meyer, G.; Ebel, R.; Abdel-Mageed, W. M.; Jaspars, M. *Nat. Chem.* **2010,** *2,* 821–825. doi:10.1038/nchem.765

5. Morel, N.; Ramonda, M.; Tordjeman, P. *Appl. Phys. Lett.* **2005,** *86,* 163103. doi:10.1063/1.1905803
6. Cumpson, P. J.; Hedley, J.; Clifford, C. A.; Chen, X.; Allen, S. *J. Vac. Sci. Technol., A* **2004,** *22,* 1444–1449. doi:10.1116/1.1763898
7. Cumpson, P. J.; Hedley, J.; Zhdan, P. *Nanotechnology* **2003,** *14,* 918–924. doi:10.1088/0957-4484/14/8/314
8. Jericho, K. S.; Jericho, H. M. *Rev. Sci. Instrum.* **2002,** *73,* 2483–2485. doi:10.1063/1.1471351
9. Gibson, C. T.; Watson, G. S.; Myhra, A. *Nanotechnology* **1996,** *7,* 259–262. doi:10.1088/0957-4484/7/3/014
10. Torii, A.; Sasaki, M.; Hane, K.; Okuma, S. *Meas. Sci. Technol.* **1996,** *7,* 179–184. doi:10.1088/0957-0233/7/2/010
11. Senden, J. T.; Ducker, A. W. *Langmuir* **1994,** *10,* 1003–1004. doi:10.1021/la00016a600
12. Cleveland, J. P.; Manne, S.; Bocek, D.; Hansma, P. K. *Rev. Sci. Instrum.* **1993,** *64,* 403–405. doi:10.1063/1.1144209
13. Sader, J. E. *Rev. Sci. Instrum.* **1995,** *66,* 4583–4587. doi:10.1063/1.1145292
14. Gibson, C. T.; Weeks, B. L.; Lee, J. R. I.; Abell, C.; Rayment, T. *Rev. Sci. Instrum.* **2001,** *72,* 2340–2343. doi:10.1063/1.1361080
15. Hutter, L. J.; Bechhoefer, J. *Rev. Sci. Instrum.* **1993,** *64,* 1868–1873. doi:10.1063/1.1143970
16. Sader, J. E.; Larson, I.; Mulvaney, P.; White, R. L. *Rev. Sci. Instrum.* **1995,** *66,* 3789–3798. doi:10.1063/1.1145439
17. Sader, J. E.; Chon, J. W. M.; Mulvaney, P. *Rev. Sci. Instrum.* **1999,** *70,* 3967–3969. doi:10.1063/1.1150021
18. Lévy, R.; Maaloum, M. *Nanotechnology* **2002,** *13,* 33–37. doi:10.1088/0957-4484/13/1/307
19. Green, C. P.; Lioe, H.; Cleveland, J. P.; Proksch, R.; Mulvaney, P.; Sader, J. E. *Rev. Sci. Instrum.* **2004,** *75,* 1988–1996. doi:10.1063/1.1753100
20. McFarland, A. W.; Poggi, M. A.; Bottomley, L. A.; Colton, S. J. *J. Micromech. Microeng.* **2005,** *15,* 785–791. doi:10.1088/0960-1317/15/4/016
21. Simon, G. H.; Heyde, M.; Rust, H.-P. *Nanotechnology* **2007,** *18,* 255503. doi:10.1088/0957-4484/18/25/255503
22. Castellanos-Gomez, A.; Agraït, N.; Rubio-Bollinger, G. *Nanotechnology* **2009,** *20,* 215502. doi:10.1088/0957-4484/20/21/215502
23. Shaw, G. A.; Pratt, J. R.; Jabbour, Z. J. Small mass measurements for tuning fork-based atomic force microscope cantilever spring constant calibration. In *MEMS and Nanotechnology, Vol. 2 – Proceedings of the 2010 Annual Conference on Experimental and Applied Mechanics;* Proulx, T., Ed.; Springer: New York, 2011; pp 4956 ff.
24. Rychen, J.; Ihn, T.; Studerus, P.; Herrmann, A.; Ensslin, K.; Hug, H. J.; van Schendel, P. J. A.; Güntherodt, H. J. *Rev. Sci. Instrum.* **2000,** *71,* 1695–1697. doi:10.1063/1.1150521
25. Wintjes, N.; Lange, M.; van Vörden, D.; Karacuban, H.; Utzat, D.; Möller, R. *J. Vac. Sci. Technol., B* **2010,** *28,* C4E21–C4E23. doi:10.1116/1.3374720
26. Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. *Rev. Sci. Instrum.* **2007,** *78,* 013705. doi:10.1063/1.2432410
27. Welker, J.; de Faria Elsner, F.; Giessibl, F. J. *Appl. Phys. Lett.* **2011,** *99,* 084102. doi:10.1063/1.3627184
28. Technical Specifications for Torr Seal (Ladd Research). http://www.laddresearch.com/tds/30022tds.pdf (accessed Aug 16, 2012).
29. UHU plus endfest 300, Technisches Datenblatt (APM Technica AG). http://www.apm-technica.com/media/downloadcenter/pdf/tdb/TDB%20APM%20431595.pdf (accessed Aug 16, 2012).
30. UHU plus endfest 300, UHU Technisches Merkblatt. http://www.uhu-profi.de/uploads/tx_ihtdatasheets/tds_plus_endfest300.pdf (accessed Aug 16, 2012).
31. Gent, A. N. *Rubber Chem. Technol.* **1958,** *31,* 896–906. doi:10.5254/1.3542351

License and Terms

Beilstein Journal of Nanotechnology

Pinch-off mechanism in double-lateral-gate junctionless transistors fabricated by scanning probe microscope based lithography

Farhad Larki*[1], Arash Dehzangi[2], Alam Abedini[1], Ahmad Makarimi Abdullah[3], Elias Saion[1], Sabar D. Hutagalung[3], Mohd N. Hamidon[4] and Jumiah Hassan[1]

Full Research Paper

Open Access

Address:
[1]Department of Physics, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia, [2]Institute of Microengineering and Nanoelectronics (IMEN), Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia, [3]School of Materials and Mineral Resources Engineering, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia, and [4]Functional Devices Laboratory, Institute of Advanced Technology, Universiti Putra Malaysia, 43400 Serdang, Selangor, Malaysia

Email:
Farhad Larki* - farhad.larki@gmail.com

* Corresponding author

Keywords:
AFM nanolithography; junctionless transistors; pinch-off; scanning probe microscope; simulation

Beilstein J. Nanotechnol. **2012,** *3,* 817–823.
doi:10.3762/bjnano.3.91

Received: 21 September 2012
Accepted: 09 November 2012
Published: 03 December 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

A double-lateral-gate p-type junctionless transistor is fabricated on a low-doped (10^{15}) silicon-on-insulator wafer by a lithography technique based on scanning probe microscopy and two steps of wet chemical etching. The experimental transfer characteristics are obtained and compared with the numerical characteristics of the device. The simulation results are used to investigate the pinch-off mechanism, from the flat band to the *off* state. The study is based on the variation of the carrier density and the electric-field components. The device is a pinch-off transistor, which is normally in the *on* state and is driven into the *off* state by the application of a positive gate voltage. We demonstrate that the depletion starts from the bottom corner of the channel facing the gates and expands toward the center and top of the channel. Redistribution of the carriers due to the electric field emanating from the gates creates an electric field perpendicular to the current, toward the bottom of the channel, which provides the electrostatic squeezing of the current.

Introduction

The fabrication of transistors without junctions and a doping concentration gradient has been introduced recently as a potential way to overcome the major obstacles in ultrascaled transistors [1,2]. Accordingly, based on simulation studies, performance estimates of junctionless transistors (JLTs), quantum ballistic transport, and novel structures such as bulk planar junc-

tionless transistors (BPJLTs) have also been investigated [3-5]. The idea behind the JLTs, or pinch-off transistors [6], is to simplify the source/drain engineering by removing the conventional junctions, and at the same time, facilitating the scaling of the transistors. The structures of proposed JLTs utilize a thin channel with homogeneous doping polarity and high doping concentration across the source/drain and the channel. High doping concentration can provide a higher value of current for such a thin channel, but at the same time causes an unavoidable scattering effect and subthreshold swing (SS) fluctuation. The latter case can justify new experiments with low doping concentration for JLTs.

The fabrication of low-doped single-lateral-gate (SG) and double-lateral-gate junctionless transistors (DGJLT) by scanning probe microscope based lithography (SPL) via local anodic oxidation (LAO) was reported previously [7-9]. The experimental characteristics were also investigated and single-gate and double-gate structures were compared [10]. The principle of SPL on silicon-on-insulator (SOI) was described for the first time by Snow et al. [11]. Subsequent results by this technique are presented in the references [12,13]. In fact, fabrication of nanostructures by SPL and particularly by using atomic force microscope (AFM) nanolithography has been developed with prominent results, and similar structures have been fabricated and experimentally characterized [14,15]; however, the lack of sufficient explanation for the behavior of these structures is still an interesting issue. Moreover, similar structures have never been investigated numerically, or ever used as pinch-off devices.

In this work, we investigate some of the influential factors on SPL by AFM nanolithography, to obtain the optimized parameters for fabrication of the DGJLT. We also used 3-D TCAD simulations to investigate the principles of the DGJLT in the *off* state. We investigate the electron/hole density distribution and electric-field components along the channel and the source/drain extension in order to obtain a better understanding of the device performance through the pinch-off mechanism.

Methodology

The DGJLT structure was physically fabricated by using the local anodic oxidation (LAO) process and two wet-chemical-etching processes. Pre-oxidation sample preparation steps, fabrication method, and parameters were elaborately mentioned in our previous works [8,10]. LAO by AFM nanolithography was carried out on a lightly doped (10^{15} cm^{-3}) p-type (100) SOI wafer with top silicon thickness of 100 nm and a 145 nm buried oxide (BOX) thickness with a resistivity of 13.5–22.5 Ω cm [16], by using scanning probe microscope (SPM) (SPI3800N/4000). The buried oxide layer in the SOI wafer was used as an insulator between the device and the handle silicon wafer, and also as the etch-stop in the wet-etching process. All predesigned oxide masks were fabricated in contact mode employing a Cr/Pt coated tip with a force constant of 0.2 N m^{-1} and a resonance frequency of 13 kHz. The room humidity (RH) was controllable from 50% to 80% with an accuracy of 1%. At a constant RH and contact force, the AFM tip voltage and writing speed (exposing time) are important factors to determine the size of the oxide patterns. Figure 1a and Figure 1b show that oxide protrusions were produced on the hydrogen passivated silicon surface with various voltages from 6 to 9 V at a constant writing speed of 1 μm/s, at 65% RH, and in contact mode. The effect of different writing speeds at constant voltage of 9 V is shown in Figure 1c. The optimized oxidation parameters correspond to a voltage of 9 V on the tip, with a speed of 1 μm/s, with the RH in the range of 65–67%.

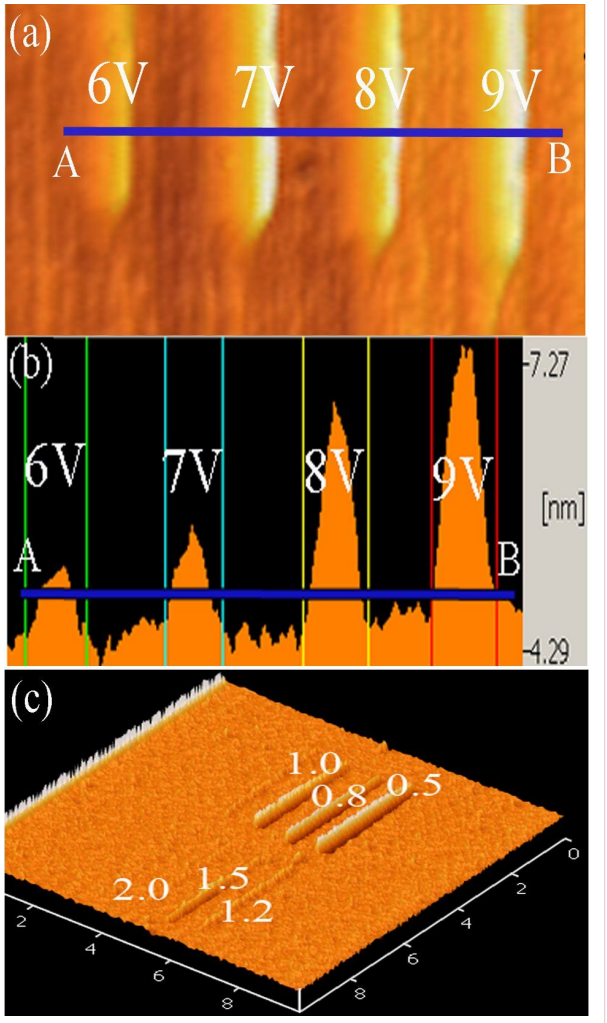

Figure 1: (a) AFM topographic images depicting a series of oxide protrusions produced by applying various voltages in the range 6–9 V. (b) Cross-section profiles along AB lines indicated in (a). (c) Effect of writing speed at a constant voltage of 9 V.

After patterning of the oxide mask, the first step of wet chemical etching was carried out with a solution of 30 wt % potassium hydroxide (KOH) saturated with 10 vol % isopropyl alcohol (IPA) at 63 °C for 20 s, in order to remove the unmasked Si layer. IPA was used as an initiator to improve the cleaning process: it reduces the etch rate, improves the surface roughness and makes the etching process more controllable [17]. This step of device fabrication is very significant and high accuracy and precision is required in order to obtain a smooth and uniform surface. The final step of the device fabrication is the etching of the silicon oxide, which allows the removal of the oxide mask. The silicon oxide etch was performed with diluted hydrofluoric acid (H_2O/HF 100:1). The immersion time of the sample in the HF solution was about 16 to 18 s. Figure 2 shows the AFM topography image of the DGJLT after two steps of successive etching.

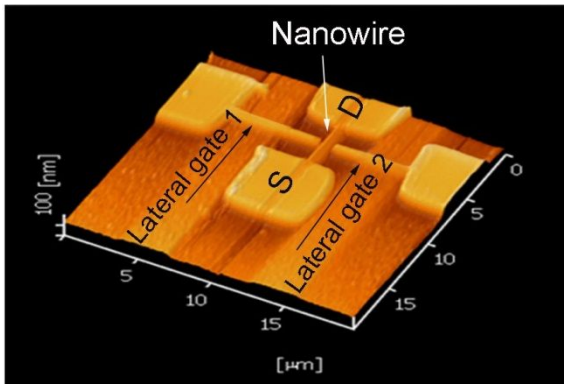

Figure 2: AFM topography image of the final structure.

Final parameters of the fabricated structure, with the best gate symmetry and reproducibility of dimensions that we achieved, are given in Table 1. According to the fabrication method, the whole structure has the same thickness (100 nm) and doping concentration of 10^{15} cm^{-3} throughout the channel, source/drain, and lateral gates. The lateral gates are located 100 nm away from the channel region.

Table 1: Fabricated/simulated device parameters.

	width	thickness	length
nanowire	100 nm	100 nm	4.2 µm
S/D/Gates pad	5 µm	100 nm	5 µm
lateral gates	200 nm	100 nm	4.5 µm

In this paper, 3-D simulations of the DGJLT were carried out by the Sentaurus 3-D device simulator [18]. All parameters in the simulation process are analogous to the fabricated device as mentioned in Table 1. The isometric view of the simulated device structure is schematically presented in Figure 3a. In accordance with the fabricated structure, the simulated structure also consists of a nanowire, two wires as lateral gates, and four square pads as source, drain, and gate contacts. The complete structure sits on a 145 nm ideal oxide. The origin in all simulation results is considered to be in the center at the bottom of the channel with the BOX interface. According to the material of the tip used in the experimental measurements (Tungsten), the work function of the contacts is taken to be 5.12 eV in all simulation steps. The simulations were carried out by using the hydrodynamic model. The doping-dependent

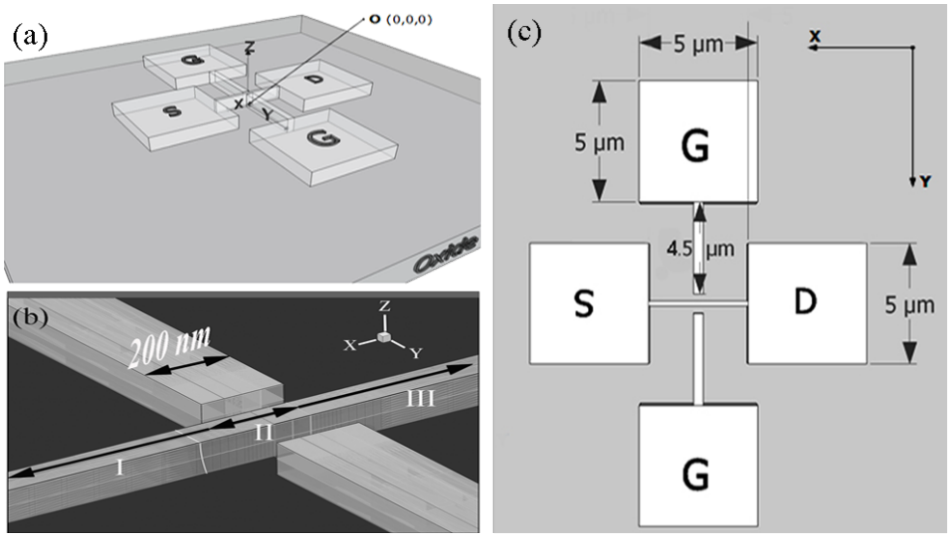

Figure 3: (a) Isometric view of the simulated device. (b) Blow up view of the gated area and different nanowire zones. (c) Top view of the simulated device.

Masetti mobility model, which incorporates the high-field-saturation Canali model, is used in order to examine the high-electric-field effect. Doping-dependent Shockley–Read–Hall recombination-generation [19] was applied in order to consider the leakage current and recombination through deep defect levels in the gap. As is shown in Figure 3b, for ease of explanation in future references, the nanowire between the source and the drain contacts is divided into three different zones, labeled as X_I, X_{II}, and X_{III}, with lengths of 2 µm, 200 nm, and 2 µm, respectively. In Figure 3c, the top view of the simulated structure is shown schematically. Here, the dimensions of the source, drain, and gate pads, as well as lateral gate length are demonstrated.

Results and Discussion

A scanning electron microscope (SEM) image of the gated area of the device with the best gate symmetry is shown in Figure 4a. Transfer characteristics (I_D–V_G) from experimental measurements and 3-D TCAD simulation results for DGJLT are depicted in Figure 4b. The electrical characteristics of the device were measured by the HP4156C semiconductor parameter analyzer (SPA, Agilent) at room temperature. It should be noted that the measurement setup has four highly accurate source/monitor units (SMUs) and is designed for Kelvin connections. Four lakeshore Tungsten tips with 3 µm radius were used with the SMUs to measure the electrical characteristics of the device. An overall good agreement is found between measurement and simulation results. The on-currents of simulated results and experimental measurement are of the same order of magnitude. However, the experimental *off* state current agreement cannot be determined, due to the limitations of our measurement instrument. A small variation in threshold voltages (V_{th}) between the simulation and experimental curves can be illustrated by the presence of fixed interface charge, work function differences, or both [20].

It should be mentioned that, during the KOH anisotropic etching process, very complicated three-dimensional structures based on the etchant concentration, temperature, and angle between the silicon surface and the mask [21] are likely to be created. However, in all simulation steps the ideal initial cubic shape is considered, for simplicity. We believe that, according to the device dimensions in the simulation study, considering the exact shape of the device edges can only slightly affect the critical characteristics of the devices.

Figure 4b indicates that, at zero gate voltage, the device is in its *on* state when it is biased by a nonzero drain–source voltage. A sufficient positive bias applied to the lateral gates depletes the region under the gates and cause an *off* state. The transfer characteristic curves also show that the device has on–off ratio of 10^6 and 10^7 for the fabricated and simulated device between $V_G = 0$ V and $V_G = 2$ V, respectively. In the case of accumulation MOSFETs (AMOSFETs) and JLTs (gated resistors), the depletion of the channel region is in the *off* state at $V_G = 0$ V, caused by the work-function difference between the gate material and the highly doped channel [2,20]. As a result, in all fabricated devices a gate bias voltage equal to the work-function difference between channel and the gate is required to achieve a flat-band condition. It is worth noting that the JLT is principally a gated resistor that is normally an *on* device at $V_G = 0$ V [22]. When zero gate bias is applied to DGJLT, the entire channel region is neutral (i.e., not depleted), and the device is in a flat-band condition. The similar and low doping concentration of the channel and the gates eliminates the effect of a work-function difference between the channel and the

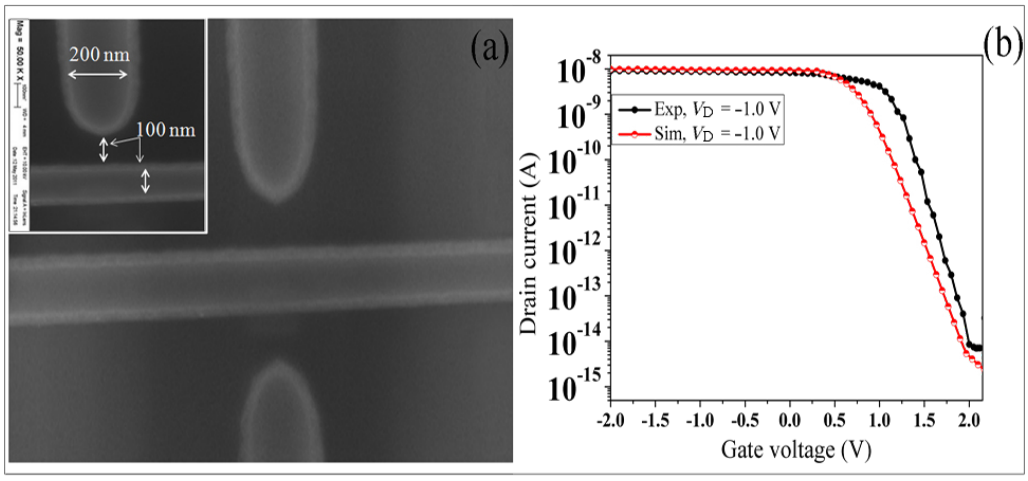

Figure 4: (a) SEM image of the gated area. (b) Experimental measurement and simulated transfer characteristics of DGJLT at $V_{DS} = -1.0$ V.

gates, which provides the flat-band condition at zero gate voltage.

According to the equation proposed in [10] the saturation drain current of DGJLTs is given by,

$$I_{\mathrm{Dsat}} \approx \frac{1}{2}\frac{q\mu N_{\mathrm{A}} T_{\mathrm{Si}} W_{\mathrm{Si}}}{L} (V_{\mathrm{Ch}}^{\mathrm{Sat}})^2, \qquad (1)$$

where W_{Si} is the width of the silicon, T_{Si} is the thickness of the silicon, N_{A} is the doping concentration, L is of the order of L_{G}, and $V_{\mathrm{Ch}}^{\mathrm{Sat}}$ is the effective channel voltage, which obeys

$$\left|V_{\mathrm{Ch}}^{\mathrm{Sat}}\right| < \left|V_{\mathrm{DSat}}\right|.$$

In JLTs, the operation is designed to start from the *off* state at zero gate voltage, and be driven into the *on* state by means of a proper gate voltage [5,23], but in a pinch-off device such as a DGJLT, the operation is implemented in the reverse direction, in order to force the device into the *off* state from the *on* state.

The mechanism of depletion due to the lateral gate voltage can be demonstrated by the hole distribution in the channel. Figures 5a–d show the simulated results for the hole density distribution in the channel profile (*X*-cut) for different gate voltages from V_{G} = +0.5 to +2.0 V.

When the device is in the *on* state and a positive gate voltage is applying to the lateral gates, the channel starts to deplete, and at a sufficient positive gate voltage the channel is fully depleted and the device is in the pinch-off condition. For the DGJLT, due to the specific shape of the device and having only one interface with the BOX at the bottom, the undepleted (neutral) area is located in the center of the channel [10]. As is shown in Figure 5, the depletion starts at the bottom corner of the channel, facing the gates, and expands toward the center and the top of the channel.

Figure 6 shows the hole and the electron density distribution along the channel axis at four different positive gate voltages, from near flat-band to pinch-off state. As the graphs of the hole density distribution show, the main depletion starts from the drain side of the gate ($X_{\mathrm{II/III}}$ interface).

The holes in the channel are repelled by the gate voltages and swept up by the drain contact; therefore, the higher depletion rate occurs at the higher positive gate voltage (Figure 6a).

When the positive gate voltage increases, the electrons start to accumulate under the gated area, and the higher voltage provides more accumulated electrons in the channel. Since the

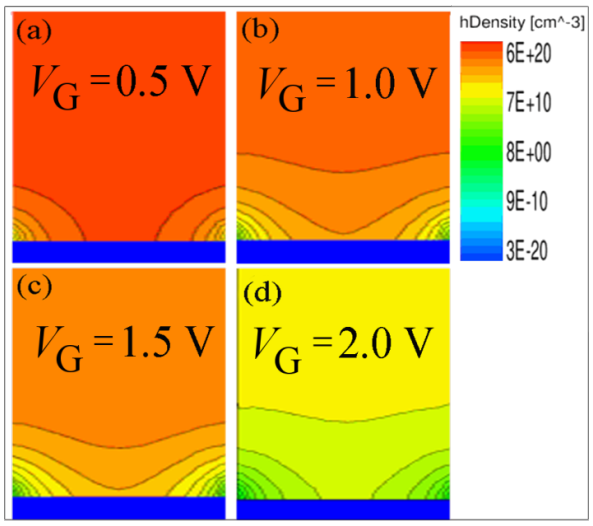

Figure 5: Hole concentration for a vertical cut of DGJLT at (*X* = −100 nm) for four different gate voltages (a) V_{G} = +0.5, (b) V_{G} = +1.0., (c) V_{G} = +1.5, and (d) V_{G} = +2.0 V. The contours present the depletion of carriers due to the gate effect.

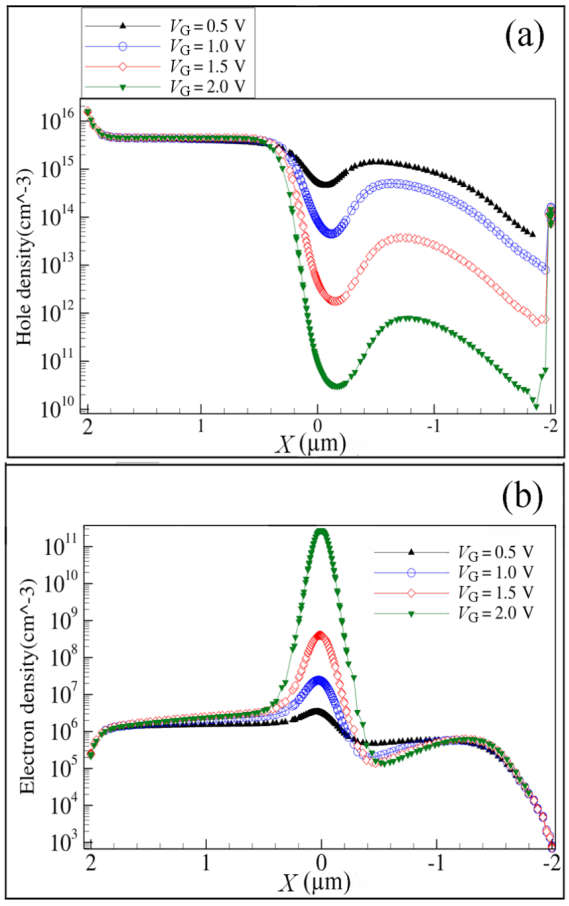

Figure 6: (a) Hole density and (b) electron density distribution as a function of position along a horizontal cross section, at four different gate voltages (0.5, 1.0, 1.5, 2.0 V), V_{DS} = −1.0 V.

electrons in the gated region are mostly absorbed by the positive gate voltage and simultaneously repelled toward the channel by the negative drain voltage, lower electron density occurs in the drain extension (Figure 6b).

The complete depletion mechanism of the device relies on the ability of the electric field of the lateral gates to deplete the channel of holes, and the incapability of the drain to supply an influential electron population into this area. In fact, the introduced electrons from the drain contact are really negligible compare to the hole density (Figure 6b). The accumulated electrons in the channel create an area of higher electron potential energy. This potential difference establishes an electric field toward the zone X_I and provides a barrier against the holes passing through the channel from the source to the drain, which facilitates the pinch-off effect in the channel. Figure 7a and Figure 7b show the electric field parallel and perpendicular to the current flow at four different gate voltages from the flat-band to the pinch-off state. Wherever the accumulated electrons are higher (Figure 6b), due to the increased positive gate voltage, the components of the electric field are stronger. As the normal component of electric field indicates, and according to the mechanism of depletion (from the bottom to the center and top of the channel), a redistribution of the accumulated electrons and depleted holes provokes an electric field toward the bottom of the channel (Si/BOX).

Figure 7 indicates that the peak of the electric field is located at the area of lowest hole concentration, which confirms the normal behavior of JLTs [24]. It is worth noting that the region of high electric field is located in the drain extension of the DGJLT outside of the gated region, since current blocking is caused by pure electrostatic pinch-off effect over the channel. This is in contrast to conventional junction transistors in which the peak of the electric field is located in the channel region, right next to the metallurgical junction [23]. As a result, the influence of the drain electric field on the channel region of DGJLT is much smaller than in conventional transistors (inversion mode). The configuration of the electric field, in a direction perpendicular to the current, indicates that the minimum value of the normal electric field is below the threshold voltage, due to the formation of the electrically neutral conducting wire below the threshold voltage between the source and the drain. By increasing the gate voltage above the threshold voltage, the carriers are affected by a large electric field normal to the current flow. This is completely opposite to the case of accumulation-mode (AM) devices in which the highest electric field appears in the channel when the device is turned on [20]. By increasing the drain voltage, a high electric field in the drain extension creates a depletion area in zone X_{III}. At sufficiently high electric field, full depletion of the nanowire in the drain

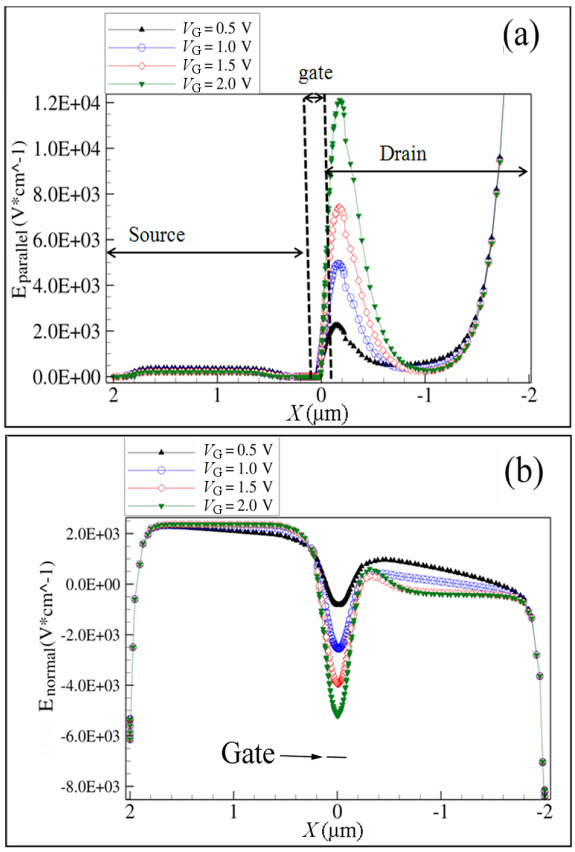

Figure 7: Simulated (a) parallel and (b) normal electric field along a horizontal cut line at the center of the channel (source side to drain side), V_G = 0.5, 1, 1.5, and 2.0 V, V_{ds} = −1.0 V.

extension acts as a buffer and prevents the drain electric field from propagating into the channel. This mechanism reduces the channel modulation effects and at high drain voltage, this barrier causes the saturation of current.

Conclusion

We have presented fabrication of p-type DGJLTs using an unconventional method of scanning probe lithography and a numerical study of the same structures using 3-D TCAD simulation results. The analysis presented shows that the device can be considered as a simple FET device that includes no doping concentration gradient and no junction. Unlike the accumulation metal–oxide–semiconductor field-effect transistors (AMOSFETs) and junctionless nanowire transistors (JLNWT), in which the channel is depleted when zero gate voltage is applied to the device, the channel of the DGJLT is fully *on* at this gate voltage and the entire channel region is neutral. These simple devices are normally *on* transistors and the device can be forced to the pinch-off state as a result of full depletion of channel by the electric field created by the lateral gates.

Acknowledgements

The authors gratefully acknowledge that this work was financially supported by the Science Fund from the Ministry of Science, Technology and Innovation (MOSTI), Malaysia, under project no. 03-01-05-SF0384, the USM Short Term Grant under project number 304/PBAHAN/6039035, and UPM FRGS number 5524051. The authors also thank the anonymous referees for valuable comments and suggestions.

References

1. Colinge, J.-P.; Lee, C.-W.; Afzalian, A.; Akhavan, N. D.; Yan, R.; Ferain, I.; Razavi, P.; O'Neill, B.; Blake, A.; White, M.; Kelleher, A.-M.; McCarthy, B.; Murphy, R. *Nat. Nanotechnol.* **2010,** *5,* 225–229. doi:10.1038/nnano.2010.15
2. Colinge, J. P.; Kranti, A.; Yan, R.; Lee, C. W.; Ferain, I.; Yu, R.; Akhavan, N. D.; Razavi, P. *Solid-State Electron.* **2011,** *65–66,* 33–37. doi:10.1016/j.sse.2011.06.004
3. Lee, C.-W.; Ferain, I.; Afzalian, A.; Yan, R.; Akhavan, N. D.; Razavi, P.; Colinge, J.-P. *Solid-State Electron.* **2010,** *54,* 97–103. doi:10.1016/j.sse.2009.12.003
4. Sels, D.; Sorée, B.; Groeseneken, G. *J. Comput. Electron.* **2011,** *10,* 216–221. doi:10.1007/s10825-011-0350-2
5. Gundapaneni, S.; Ganguly, S.; Kottantharayil, A. *IEEE Electron Device Lett.* **2011,** *32,* 261–263. doi:10.1109/LED.2010.2099204
6. Sorée, B.; Magnus, W.; Vandenberghe, W. *Appl. Phys. Lett.* **2011,** *99,* 233509. doi:10.1063/1.3669509
7. Dehzangi, A.; Larki, F.; Saion, E. B.; Hatagalung, S. D.; Abdullah, M.; Hamidon, M. N.; Hassan, J. *Am. J. Appl. Sci.* **2011,** *8,* 872–877. doi:10.3844/ajassp.2011.872.877
8. Larki, F.; Hutagalung, S. D.; Dehzangi, A.; Saion, E. B.; Abedini, A.; Abdullah, A. M.; Hamidon, M. N.; Hassan, J. *Microelectron. Solid State Electron.* **2012,** *1,* 15–20.
9. Dehzangi, A.; Larki, F.; Saion, E. B.; Hutagalung, S. D.; Hamidon, M. N.; Hassan, J. In *IEEE, Regional Symposium on Micro and Nanoelectronics (RSM),* Kota Kinabalu, Sept 28–30, 2011; pp 104–107. doi:10.1109/RSM.2011.6088302
10. Dehzangi, A.; Abdullah, A. M.; Larki, F.; Hutagalung, S. D.; Saion, E. B.; Hamidon, M. N.; Hassan, J.; Gharayebi, Y. *Nanoscale Res. Lett.* **2012,** *7,* 381. doi:10.1186/1556-276X-7-381
11. Snow, E. S.; Campbell, P. M. *Appl. Phys. Lett.* **1994,** *64,* 1932–1934. doi:10.1063/1.111746
12. Vijaykumar, T.; Kulkarni, G. U. *Solid State Commun.* **2007,** *142,* 89–93. doi:10.1016/j.ssc.2007.01.027
13. Malshe, A. P.; Rajurkar, K. P.; Virwani, K. R.; Taylor, C. R.; Bourell, D. L.; Levy, G.; Sundaram, M. M.; McGeough, J. A.; Kalyanasundaram, V.; Samant, A. N. *CIRP Ann. - Manuf. Technol.* **2010,** *59,* 628–651. doi:10.1016/j.cirp.2010.05.006
14. Ionica, I.; Montès, L.; Ferraton, S.; Zimmermann, J.; Saminadayar, L.; Bouchiat, V. *Solid-State Electron.* **2005,** *49,* 1497–1503. doi:10.1016/j.sse.2005.07.012
15. Pennelli, G. *Microelectron. Eng.* **2009,** *86,* 2139–2143. doi:10.1016/j.mee.2009.02.032
16. SOITEC: PTF, 38190 Bernin, France; http://www.soitec.com/en/index.php
17. Pennelli, G.; Piotto, M.; Barillaro, G. *Microelectron. Eng.* **2006,** *83,* 1710–1713. doi:10.1016/j.mee.2006.01.144
18. *TCAD-Sentaurus User Manual,* Version D 2010, 03; Synopsys, Inc.: Mountain View, CA, 2010.
19. Sze, S. M.; Ng, K. K. *Physics of semiconductor devices;* Wiley-Blackwell, 2007.
20. Iqbal, M. M.-H.; Hong, Y.; Garg, P.; Udrea, F.; Migliorato, P.; Fonash, S. J. *IEEE Trans. Electron Devices* **2008,** *55,* 2946–2959. doi:10.1109/TED.2008.2005174
21. Rađenović, B.; Radmilović-Rađenović, M. *Hem. Ind.* **2010,** *64,* 93–97. doi:10.2298/HEMIND100205008R
22. Ansari, L.; Feldman, B.; Fagas, G.; Colinge, J.-P.; Greer, J. C. *Appl. Phys. Lett.* **2010,** *97,* 062105. doi:10.1063/1.3478012
23. Colinge, J.-P.; Ferain, I.; Kranti, A.; Lee, C.-W.; Akhavan, N. D.; Razavi, P.; Yan, R.; Yu, R. *Sci. Adv. Mater.* **2011,** *3,* 477–482. doi:10.1166/sam.2011.1163
24. Colinge, J.-P.; Lee, C.-W.; Ferain, I.; Akhavan, N. D.; Yan, R.; Razavi, P.; Yu, R.; Nazarov, A. N.; Doria, R. T. *Appl. Phys. Lett.* **2010,** *96,* 073510. doi:10.1063/1.3299014

Beilstein Journal of Nanotechnology

Reversible mechano-electrochemical writing of metallic nanostructures with the tip of an atomic force microscope

Christian Obermair*[1], Marina Kress[1], Andreas Wagner[1] and Thomas Schimmel[1,2]

Full Research Paper

Open Access

Address:
[1]Institute of Applied Physics and Center for Functional Nanostructures (CFN), Karlsruhe Institute of Technology (KIT), 76128 Karlsruhe, Germany, and [2]Institute of Nanotechnology (INT), Karlsruhe Institute of Technology (KIT), 76021 Karlsruhe, Germany

Email:
Christian Obermair* - christian.obermair@physik.uni-karlsruhe.de

* Corresponding author

Keywords:
atomic force microscopy; electrochemical deposition; electrochemistry; nanoelectronics; nanofabrication; nanolithography; nanotechnology; MEMS and NEMS; reversible processes; scanning probe microscopy and lithography

Beilstein J. Nanotechnol. **2012,** *3,* 824–830.
doi:10.3762/bjnano.3.92

Received: 14 August 2012
Accepted: 14 November 2012
Published: 05 December 2012

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We recently introduced a method that allows the controlled deposition of nanoscale metallic patterns at defined locations using the tip of an atomic force microscope (AFM) as a “mechano-electrochemical pen”, locally activating a passivated substrate surface for site-selective electrochemical deposition. Here, we demonstrate the reversibility of this process and study the long-term stability of the resulting metallic structures. The remarkable stability for more than 1.5 years under ambient air without any observable changes can be attributed to self-passivation. After AFM-activated electrochemical deposition of copper nanostructures on a polycrystalline gold film and subsequent AFM imaging, the copper nanostructures could be dissolved by reversing the electrochemical potential. Subsequent AFM-tip-activated deposition of different copper nanostructures at the same location where the previous structures were deleted, shows that there is no observable memory effect, i.e., no effect of the previous writing process on the subsequent writing process. Thus, the four processes required for reversible information storage, “write”, “read”, “delete” and “re-write”, were successfully demonstrated on the nanometer scale.

Introduction

The electrochemical deposition of metallic contacts, wires and patterns on the nanometer scale is of tremendous technological relevance for contacting micro- and nanoelectronic integrated circuits in the semiconductor industry [1]. With the shrinking dimensions of semiconductor structures, there is a need to understand and control electrochemical deposition processes on

the nanometer scale. This also applies to the field of micro- and nano-electromechanical systems (MEMS and NEMS). At the same time, much progress was achieved in recent years in understanding the mechanisms and developing new methods for the deposition and control of metallic structures and contacts on the nanometer scale and even down to the atomic scale [2-13]. Nanowires showing a dramatic increase of the yield strength by more than one order of magnitude have been fabricated by electrochemical deposition [14]; transport through single molecules between metallic nanocontacts was studied intensively [15-24]. Even a single-atom transistor [25-29] was demonstrated, i.e., a device that allows the controlled switching "on" and "off" of an electrical current by the controlled and reversible movement of one individual atom.

While electrochemical processes on the nanometer scale are of increasing technological relevance, there is also a need to visualize, study and control these processes locally on this length scale. Atomic force microscopy (AFM) and related techniques are valuable tools for imaging surfaces and surface processes on the nanometer scale. Examples of recent work can be found in the literature [30-40]. At the same time, the scanning tips of the atomic force microscope, the scanning tunneling microscope (STM) and related instruments can be used as tools for surface modification and nanolithography [9-12,41-48], even down to the atomic scale [41-45]. The great advantage of these instruments for nanoelectrochemistry is the fact that they also allow the in situ and real-time observation of the processes within the electrochemical cell [49-53], making them a valuable tool for studying electrochemical processes on the nanometer scale.

Recently, we demonstrated the local electrochemical deposition of metallic nanostructures and nanowires mechanically induced with the tip of an AFM [12]. In the semiconductor industry, mechano-electrochemical processes are frequently used in the form of mechano-electrochemical polishing. While this is only a macroscopic method with very poor lateral resolution, the tip of an AFM allows the site-selective control of electrochemical deposition on the nanometer scale, e.g., by mechanically depassivating a surface. Here, we study this process in more detail. We demonstrate that the technique of AFM-tip-induced mechano-electrochemical writing is fully reversible, and we investigate the long-term stability of the resulting structures. We also discuss the implications of these results on the understanding of the deposition mechanism.

Results and Discussion

Reversible writing, deleting and re-writing

Writing of metallic structures with the tip of an atomic force microscope as a "nano-electrochemical pen" was performed as described in our previous work [12]. A glass substrate covered with a polycrystalline gold film is used as a substrate. When placed into an electrolyte solution containing copper sulfate and H_2SO_4, a thin copper oxide containing film is formed on the gold surface, its thickness and properties depending on the pH value of the solution. This film can act as a passivation layer preventing electrochemical deposition of copper when an electrochemical deposition potential is applied to the gold electrode. Our tip-activated deposition method now works in the following way: while an electrochemical deposition voltage is applied to the gold electrode, the AFM tip repeatedly scans along a certain pathway at force loads on the order of 10 nN, in this way locally destroying or damaging the oxide-based passivation layer on the gold electrode. As a consequence, copper is deposited on the gold electrode at exactly the locations where the tip has destroyed the passivation layer, i.e., along the lines scanned by the tip when the deposition voltage is applied. After this site-selective, tip-activated deposition, the deposition potential is switched off and an electrochemical potential close to the electrochemical equilibrium is applied, i.e., no further deposition and no further dissolution occurs. Subsequently, AFM contact-mode imaging of the resulting structures is performed, i.e., nanopatterning and AFM imaging are performed with one and the same tip in situ within the electrolyte. This process has the advantage that an entire sequence of deposition and dissolution processes can be performed, and after each step, the resulting structures and their changes can be visualized and quantitatively characterized by AFM. A schematic representation of the experimental setup is given in Figure 1.

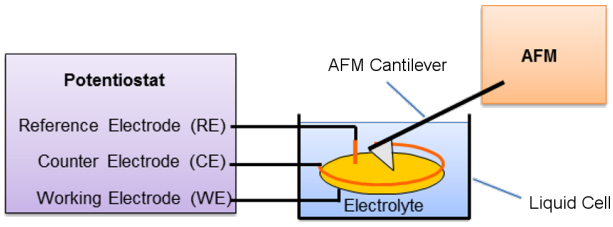

Figure 1: Schematic diagram of an electrochemical atomic force microscope, comprising an atomic force microscope with a laser-beam detection system (not shown), an electrochemical liquid cell with reference electrode (RE), counter electrode (CE) and working electrode (WE), and a computer-controlled potentiostat. The cantilever is dipped into the electrolyte but not directly connected to the electrodes.

In order to gain further insights into the mechanism of tip-induced electrochemical deposition, the erasability of the deposited structures and the reversibility of the deposition process were investigated in more detail. Therefore, a copper structure was deposited electrochemically by local mechano-electrochemical AFM-tip activation. Subsequently the copper structure was dissolved by applying an electrochemical dissolution potential. In a following step a new and differently shaped

copper nanostructure was deposited at the same location on the sample.

In Figure 2 the result of such a deposition experiment is given. Figure 2 shows the AFM topography of one and the same 1.6 µm × 1.6 µm scanning area of a polycrystalline gold surface in an electrolyte containing Cu^{2+} ions. The electrolyte consisted of an aqueous solution of 50 mM H_2SO_4 (suprapure, Merck) with 1 mM $CuSO_4$ (p.a., Merck). All potentials were measured against a Cu/Cu^{2+} reference electrode, using a Cu counter electrode and the polycrystalline gold surface as the working electrode.

For the experiment, it was important to make sure that after the first Cu deposition, after dissolution of this Cu nanostructure and deposition of a new Cu structure on the gold surface, always exactly the same area was imaged with the AFM. For this reason, an area was selected for the experiment which contained an easily identifiable label within the scanning area in the form of a gold-coated hillock on the otherwise smooth gold film (see arrow in the AFM images of Figure 2a–c). This gold-coated hillock-like defect was not dissolved, even at a high underpotential of +1 V, which guaranteed that it remained unchanged at the potentials applied during the copper deposition and dissolution experiments.

In the first step, a deposition potential of −60 mV was applied and the gold surface was scanned with the AFM tip following the contour of the shape of the digit “6” with 13 repetitions within 5 s. The scanning rate of the AFM tip on the sample surface was about 8 µm/s. The force load of the AFM tip was approx. 10 nN. In this way a line structure of largely interconnected copper islands with the shape of the digit “6” was deposited. Subsequently a potential of −30 mV (the “holding potential” at which no further Cu deposition took place) was applied to the gold working electrode and the surface area was then scanned with the AFM.

The contact-mode AFM image of Figure 2a shows that the deposition process is highly selective, i.e., the deposition occurred at exactly the sites scanned by the AFM tip, whereas absolutely no copper islands were deposited at other places. The metal deposition starts in the form of small islands, most probably nucleating at atomic-scale defects and grain boundaries. The deposited metal in the beginning of the deposition process thus forms a chain of islands along the path of the AFM tip. These islands finally touch and overgrow each other, in this way forming a continuous wire in the shape of the digit “6”. The line width of the structure is approx. 50 nm. The force load of the AFM tip during imaging was the same as during the deposition process (approx. 10 nN).

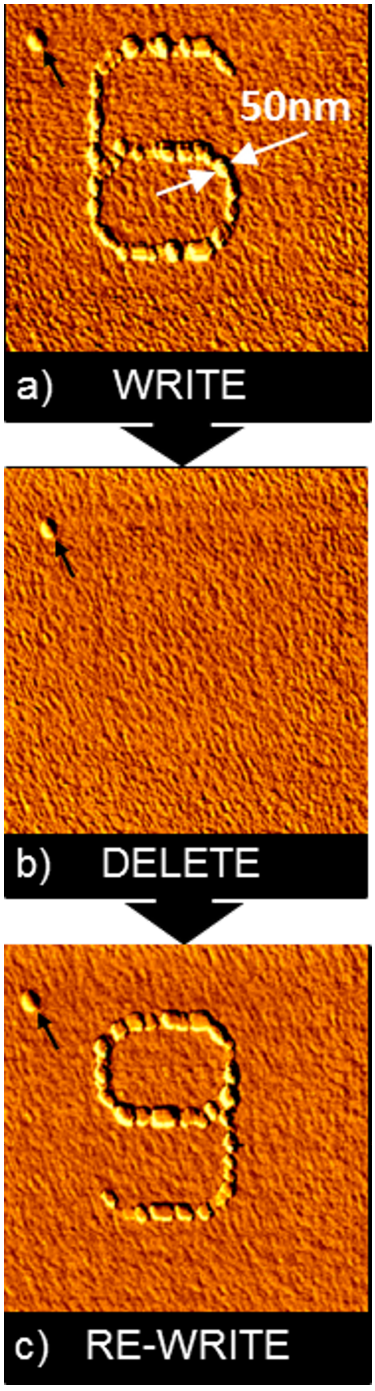

Figure 2: Reversible mechano-electrochemical writing of Cu nanostructures. (a) WRITE: In situ AFM image of an initial Cu island structure (“6”). The “6”-structure was deposited by tip-induced mechano-electrochemical writing on a gold substrate, applying a deposition potential of −60 mV versus Cu/Cu^{2+} for 5 s. (b) DELETE: In situ AFM image of the gold substrate after deletion. For deletion a potential of +0.2 mV versus Cu/Cu^{2+} was applied for 180 s. (c) RE-WRITE: In situ AFM image of a further Cu island structure (“9”). The “9”-structure was deposited on the same area where previously the “6”-structure had been written and deleted. For orientation see the point-shaped defect in the upper-left corner of the images, marked by an arrow. Typical height of the islands: 15 nm. Scan size: 1.6 µm × 1.6 µm.

In the second step a dissolution potential of +0.2 V was applied for 180 s to the Au working electrode. Figure 2b shows an AFM image of the same scanning area as shown in Figure 2a, but after application of the dissolution potential for 180 s. All copper islands are dissolved. During the dissolution process, the AFM tip was still in contact with the sample surface, but it was not scanning, i.e., it was not moved at all in relation to the sample surface. In Figure 2b no tip-induced surface defects are found on the sample surface.

Following the same procedure described above for the deposition of the copper structure with the shape of the digit "6", a new structure with the shape of the digit "9" was deposited after dissolution of the "6", at exactly the same position of the sample. During the deposition time of approx. 5 s the deposition potential of −60 mV was applied. During this time, the AFM tip was continuously scanning along the trace of the figure "9". In total, this scan was repeated 13 times within this deposition time. This means that the repetition rate with which the tip reactivated the same position of the sample during deposition was once every 0.4 seconds. The result of subsequent AFM imaging, again at a "holding" potential of −30 mV, is given in Figure 2c. In the same scanning area, where previously copper islands with the shape of the digit "6" were deposited and subsequently dissolved, new cooper islands with the shape of the digit "9" were deposited. Again, the deposition process was highly selective, i.e., no copper islands were found in the scanned area, except those arranged in the shape of the digit "9".

Neither in Figure 2b nor in Figure 2c can any indications of the previous deposition processes be found. This leads to the conclusion that during AFM-tip-induced electrochemical deposition with force loads of about 10 nN, as used in the previous experiment, the gold surface itself was not modified. Obviously, only the passivation layer on top of the gold surface was locally removed by means of the AFM tip scanning along the path of the digits "6" and "9". During none of our electrochemical deposition experiments was any observable tip-induced surface damage or defect found.

This allows performing sequential electrochemical deposition on one and the same area of the sample, because no modification of the sample area remains after the dissolution of the previously deposited structures. The tip-induced electrochemical deposition of copper islands onto gold surfaces is fully reversible.

Self-passivation and long-time stability

A key feature for the technological relevance and application perspectives of electrochemically deposited nanostructures is their long-term stability. The nanostructures have to remain stable under ambient air after being taken out of the electrolyte.

To study the long-term stability of electrochemically deposited nanostructures written with the AFM tip as a mechano-electrochemical pen, a copper structure was deposited following the procedure described above. To make the finding of the same structure after longer periods of time as well as its study with scanning electron microscopy (SEM) easier, a larger structure was deposited by AFM with a total size of 6 µm × 6 µm. The copper structure with the shape of a square frame was again deposited by AFM tip-induced electrochemical deposition on a polycrystalline gold substrate. During deposition, the structure written in the shape of a square-shaped spiral with an inner-edge length of 2 µm and an outer-edge length of 6 µm was scanned 45 times with the tip of the AFM. The AFM path during the structuring process consisted of 15 concentric squares becoming subsequently smaller in size starting at a length of 6 µm and ending at a length of 2 µm. The structures were written by using a tip speed relative to the sample surface of approx. 50 µm/s, resulting in a repetition rate of 5 seconds with which the tip subsequently retraced the same profile, i.e., the same position along the scanning path of the tip was hit by the tip every 5 seconds. The whole writing cycle was repeated 45 times, resulting in a total writing time of 3.7 minutes. The force load of the AFM tip again was approx. 10 nN. During deposition, a deposition potential of −70 mV versus Cu/Cu^{2+} was applied to the gold electrode. An aqueous solution of 100 mM H_2SO_4 (suprapure, Merck) with 1 mM $CuSO_4$ (p.a., Merck) was used as electrolyte. After deposition, a holding potential of −20 mV versus Cu/Cu^{2+} was applied to the gold electrode.

After deposition, the AFM tip was taken out of the liquid cell and the electrolyte was poured out of the cell as quickly as possible to avoid further deposition and dissolution during this very short transitional phase between the sample being immersed in the electrolyte and the sample being dry in ambient air, i.e., a transitional phase with an undefined electrochemical potential. During the pouring out of the electrolyte, the holding potential was kept applied to the electrodes in order to maintain defined electrochemical conditions as long as possible. Subsequently, the gold substrate with the copper structure was rinsed with bi-distilled water and then dried at room temperature.

Figure 3a–c shows SEM images of the as-prepared copper structure one day after deposition. The SEM images were taken with a Zeiss SEM equipped with a field-emission electron source. In Figure 3a a 32 µm × 32 µm scanning area of the sample is given. In the middle of the scanning area, the square structure of copper islands with an open area in the center is

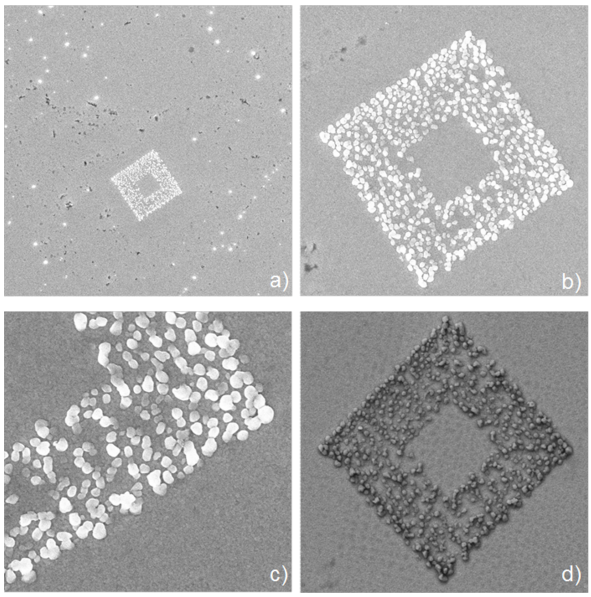

Figure 3: SEM image demonstrating the ex situ stability of the tip-induced electrochemically deposited copper islands on a gold substrate. (a–c) SEM image immediately taken after deposition, (d) SEM image of the same sample area after 1.5 years. Scan size: (a) 32 µm × 32 µm, (b) 9 µm × 9 µm, (c) 4 µm × 4 µm, (d) 9 µm × 9 µm.

visible. Figure 3b and Figure 3c are zoomed SEM images of the same structure.

To demonstrate the long-term stability of the electrochemically deposited copper structure, the same sample area of Figure 3b was investigated again after more than one year. Figure 3d shows an SEM image of the same 9 µm × 9 µm scanning area as in Figure 3b 1.5 years after the electrochemical deposition of the copper structure, thus showing the long-term stability of the resulting structures under ambient conditions. The typical humidity of the laboratory air over this time period of 1.5 years varied between 30% and 90% relative humidity, for most of the time being between 40% and 60%.

The different imaging contrasts between Figure 3a–c and Figure 3d are due to the use of different SEM detectors. For Figure 3a–c an in-lens SEM detector was used. For the image of Figure 3d the signal from a secondary electron detector was used leading to dominant topographic information.

The copper structure in Figure 3d remains unchanged, compared to Figure 3b although the sample was stored for 1.5 years under ambient conditions, i.e., ambient air and ambient humidity. If two different detectors for the SEM imaging had not been used, the structures would not have shown any discernible change at all during this 1.5-year period of time.

Nucleation density, pattern homogeneity and island size

When comparing the structures of Figure 2 with the structures of Figure 3, striking differences are observed:

- In both cases, the structures consist of a sequence of copper islands. In the case of Figure 2, these islands are significantly smaller (approximate diameter: 50 nm) than in Figure 3 (typically 80–200 nm).
- In the case of Figure 2 the islands are largely interconnected to form continuous lines (“nanowires”), whereas in Figure 3 they are mostly found in the form of single, isolated islands or in the form of smaller aggregates of 2–5 islands interconnected with each other, but separated from the rest of the copper islands.
- In the case of Figure 2, the islands form a very regular sequence and have a much narrower size distribution. In Figure 3, the positions of the islands as well as their local areal density strongly fluctuate, which is also true for the island sizes.

These differences are remarkable, considering that the structures of Figure 2 and Figure 3 have been deposited with the same technique and by applying similar electrochemical parameters such as concentrations and pH of the electrolyte, deposition potential, and holding potential. The major difference was in the very different experimental parameters for the repetition rate of the AFM scanning process during structure growth, which was by a factor of more than 12 lower for the pattern of Figure 3 than for the structure of Figure 2. This, in turn, allows conclusions to be drawn concerning the structuring mechanism.

The results described above can be explained in the following way: obviously, in both cases, there is a growth of individual copper islands along the lines where the passivating layer is damaged or removed with the tip of the AFM. The results of previously reported [12] already showed that once an area or line along the surface is activated with the AFM tip, the passivation layer starts to regrow within a certain time, which was typically on the order of seconds (depending on the chemical parameters of the system such as concentrations, pH value), requiring a continuous reactivation of the active area of the surface with the tip of the AFM to allow for the nucleation of further islands. Assuming this, a repeated scanning of the same line or trace with the AFM tip is needed during the structuring process to prevent repassivation of the tip-activated areas before the deposition process is completed. For this purpose, obviously the repetition rate of tip-induced reactivation has to be higher than the inverse time constant of the self-passivation process of the surface.

Thus the low density of islands in the case of Figure 3 is easily explained by the much lower rate of reactivation of the surface along the scan trace with the AFM tip (only every 5 seconds), allowing a repassivation between consecutive AFM scans, resulting in the growth of fewer, larger islands. In the case of Figure 2, the reactivation took place every 0.4 seconds, allowing the formation of a continuous chain of much smaller islands and showing no effects of repassivation in the form of parts of the line in which no islands grew. This indicates a repassivation time in between the repetition times of these two experiments, in agreement with values in the range of seconds estimated previously [12].

Conclusion

The controlled and reversible electrochemical writing of copper nanostructures on a polycrystalline gold surface was achieved with the tip of an atomic force microscope used as a mechano-electrochemical pen. After dissolution of a certain structure by reversing the electrochemical potential, new, different structures were written at the same position with no observable memory effect of the previous writing process. This demonstrates that the tip-induced modification is limited to the passivation layer and does not change the topography or the properties of the gold substrate, and it supports the proposed mechanism of site-selective, local AFM-tip-induced depassivation of the surface. Finally, it is shown that a sufficiently high repetition rate of the AFM scan during the growth process is needed to prevent repassivation of the once-activated surface area during the growth process, and an estimate of the repassivation time constant is derived from the experimental results.

To conclude, the four basic processes required for reversible information storage, i.e., “Write”, “Read”, “Delete” and “Rewrite”, were successfully demonstrated on the nanometer scale. Long-term stability of the resulting structures under ambient air for more than 1.5 years was demonstrated, which can be explained by self-passivation of the surface.

Apart from the aspect of reversible nanolithography, the results also help in understanding microscopic mechanisms of mechanically activated or mechanically assisted electrochemical processes on metallic surfaces, e.g., during electropolishing or in combined mechanical wear and corrosion processes.

Experimental

The experimental setup was described previously [12] in more detail. An AFM, constructed at our institute, that was specially equipped for electrochemical nanolithography was used. An electrochemical cell (diameter: approx. 20 mm), equipped with Cu reference and counter electrodes (consisting of copper wires of 0.5 mm diameter from Goodfellow) were used. Glass substrates with evaporated gold films were used as working electrodes, on which the deposition experiments were performed. Rinsing, sonication and gold coating of the glass substrates was performed following the procedure described previously [12]. Deposition and dissolution was directed by a computer-controlled bi-potentiostat.

All AFM images were taken in situ within the electrolyte in contact mode and represent original data that are shown without filtering. V-shaped silicon nitride cantilevers were used with force constants between 0.03 and 0.1 N/m. Cyclic voltammograms were measured both before and after each experiment.

Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft within the DFG-Center for Functional Nanostructures (CFN) and by the Baden-Wuerttemberg Foundation within the Research Network of Excellence “Functional Nanostructures”.

References

1. Schlesinger, M.; Paunovic, M., Eds. *Modern Electroplating (The ECS Series of Texts and Monographs, 52),* 5th ed.; Wiley-VCH: Weinheim, 2010.
2. Kolb, D. M.; Simeone, F. C. Characterization and modification of electrode surfaces by in-situ STM. In *Scanning Tunneling Microscopy in Surface Science, Nanoscience and Catalysis;* Bowker, M.; Davies, P., Eds.; Wiley-VCH: Weinheim, 2010; pp 119–146.
3. Heß, C.; Borgwarth, K.; Ricken, C.; Ebling, D. G.; Heinze, J. *Electrochim. Acta* **1997,** *42,* 3065–3073. doi:10.1016/S0013-4686(97)00154-0
4. El-Giar, E. M.; Said, R. A.; Bridges, G. E.; Thomson, D. J. *J. Electrochem. Soc.* **2000,** *147,* 586–591. doi:10.1149/1.1393237
5. Li, W.; Hsiao, G. S.; Harris, D.; Nyffenegger, R. M.; Virtanen, J. A.; Penner, R. M. *J. Phys. Chem.* **1996,** *100,* 20103–20113. doi:10.1021/jp962328d
6. Ziegler, J. C.; Engelmann, G. E.; Kolb, D. M. *Z. Phys. Chem.* **1999,** *208,* 151–166. doi:10.1524/zpch.1999.208.Part_1_2.151
7. Pötzschke, R. T.; Staikov, G.; Lorenz, W. J.; Wiesbeck, W. *J. Electrochem. Soc.* **1999,** *146,* 141–149. doi:10.1149/1.1391577
8. Schindler, W.; Hofmann, D.; Kirschner, J. *J. Electrochem. Soc.* **2001,** *148,* C124–C130. doi:10.1149/1.1343107
9. LaGraff, J. R.; Gewirth, A. A. *J. Phys. Chem.* **1994,** *98,* 11246–11250. doi:10.1021/j100095a003
10. LaGraff, J. R.; Gewirth, A. A. *J. Phys. Chem.* **1995,** *99,* 10009–10018. doi:10.1021/j100024a050
11. Obermair, C.; Müller, M.; Klinke, C.; Schimmel, T. *Chin. Phys. (Beijing, China)* **2001,** *10,* S151–S156.
12. Obermair, C.; Wagner, A.; Schimmel, T. *Beilstein J. Nanotechnol.* **2011,** *2,* 659–664. doi:10.3762/bjnano.2.70
13. Obermair, C.; Kuhn, H.; Schimmel, T. *Beilstein J. Nanotechnol.* **2011,** *2,* 740–745. doi:10.3762/bjnano.2.81
14. Zhong, S.; Koch, T.; Wang, M.; Scherer, T.; Walheim, S.; Hahn, H.; Schimmel, T. *Small* **2009,** *5,* 2265–2270. doi:10.1002/smll.200900746
15. May, F.; Wegewijs, M. R.; Hofstetter, W. *Beilstein J. Nanotechnol.* **2011,** *2,* 693–698. doi:10.3762/bjnano.2.75

16. Hong, W.; Valkenier, H.; Mészáros, G.; Zsolt Manrique, D.; Mishchenko, A.; Putz, A.; Moreno García, P.; Lambert, C. J.; Hummelen, J. C.; Wandlowski, T. *Beilstein J. Nanotechnol.* **2011,** *2,* 699–713. doi:10.3762/bjnano.2.76
17. Perrin, M. L.; Martin, C. A.; Prins, F.; Shaikh, A. J.; Eelkema, R.; van Esch, J. H.; van Ruitenbeek, J. M.; van der Zant, H. S. J.; Dulić, D. *Beilstein J. Nanotechnol.* **2011,** *2,* 714–719. doi:10.3762/bjnano.2.77
18. Todorov, T. N.; Dundas, D.; Paxton, A. T.; Horsfield, A. P. *Beilstein J. Nanotechnol.* **2011,** *2,* 727–733. doi:10.3762/bjnano.2.79
19. Strange, M.; Thygesen, K. S. *Beilstein J. Nanotechnol.* **2011,** *2,* 746–754. doi:10.3762/bjnano.2.82
20. Nakashima, S.; Takahashi, Y.; Kiguchi, M. *Beilstein J. Nanotechnol.* **2011,** *2,* 755–759. doi:10.3762/bjnano.2.83
21. Lü, J.-T.; Gunst, T.; Hedegård, P.; Brandbyge, M. *Beilstein J. Nanotechnol.* **2011,** *2,* 814–823. doi:10.3762/bjnano.2.90
22. Solomon, G. C.; Bergfield, J. P.; Stafford, C. A.; Ratner, M. A. *Beilstein J. Nanotechnol.* **2011,** *2,* 862–871. doi:10.3762/bjnano.2.95
23. Bergfield, J. P.; Barr, J. D.; Stafford, C. A. *Beilstein J. Nanotechnol.* **2012,** *3,* 40–51. doi:10.3762/bjnano.3.5
24. Bode, N.; Kusminskiy, S. V.; Egger, R.; von Oppen, F. *Beilstein J. Nanotechnol.* **2012,** *3,* 144–162. doi:10.3762/bjnano.3.15
25. Xie, F.-Q.; Nittler, L.; Obermair, C.; Schimmel, T. *Phys. Rev. Lett.* **2004,** *93,* 128303. doi:10.1103/PhysRevLett.93.128303
26. Xie, F.-Q.; Maul, R.; Augenstein, A.; Obermair, C.; Starikov, E. B.; Schön, G.; Schimmel, T.; Wenzel, W. *Nano Lett.* **2008,** *8,* 4493–4497. doi:10.1021/nl802438c
27. Xie, F.-Q.; Maul, R.; Obermair, C.; Wenzel, W.; Schön, G.; Schimmel, T. *Adv. Mater.* **2010,** *22,* 2033–2036. doi:10.1002/adma.200902953
28. Obermair, C.; Xie, F.-Q.; Schimmel, T. *Europhys. News* **2010,** *41,* 25–28. doi:10.1051/epn/2010403
29. Maul, R.; Xie, F.-Q.; Obermair, C.; Schön, G.; Schimmel, T.; Wenzel, W. *Appl. Phys. Lett.* **2012,** *100,* 203511. doi:10.1063/1.4719207
30. Gnecco, E. *Beilstein J. Nanotechnol.* **2010,** *1,* 158–162. doi:10.3762/bjnano.1.19
31. Merlijn van Spengen, W.; Turq, V.; Frenken, J. W. M. *Beilstein J. Nanotechnol.* **2010,** *1,* 163–171. doi:10.3762/bjnano.1.20
32. Malegori, G.; Ferrini, G. *Beilstein J. Nanotechnol.* **2010,** *1,* 172–181. doi:10.3762/bjnano.1.21
33. König, T.; Simon, G. H.; Heinke, L.; Lichtenstein, L.; Heyde, M. *Beilstein J. Nanotechnol.* **2011,** *2,* 1–14. doi:10.3762/bjnano.2.1
34. Magonov, S.; Alexander, J. *Beilstein J. Nanotechnol.* **2011,** *2,* 15–27. doi:10.3762/bjnano.2.2
35. Glatzel, T.; Zimmerli, L.; Kawai, S.; Meyer, E.; Fendt, L.-A.; Diederich, F. *Beilstein J. Nanotechnol.* **2011,** *2,* 34–39. doi:10.3762/bjnano.2.4
36. Stegemann, B.; Klemm, M.; Horn, S.; Woydt, M. *Beilstein J. Nanotechnol.* **2011,** *2,* 59–65. doi:10.3762/bjnano.2.8
37. Darwich, S.; Mougin, K.; Rao, A.; Gnecco, E.; Jayaraman, S.; Haidara, H. *Beilstein J. Nanotechnol.* **2011,** *2,* 85–98. doi:10.3762/bjnano.2.10
38. Elias, G.; Glatzel, T.; Meyer, E.; Schwarzman, A.; Boag, A.; Rosenwaks, Y. *Beilstein J. Nanotechnol.* **2011,** *2,* 252–260. doi:10.3762/bjnano.2.29
39. Jaafar, M.; Iglesias-Freire, O.; Serrano-Ramón, L.; Ricardo Ibarra, M.; de Teresa, J. M.; Asenjo, A. *Beilstein J. Nanotechnol.* **2011,** *2,* 552–560. doi:10.3762/bjnano.2.59
40. Tian, C.; Liu, Z.; Jin, J.; Lebedkin, S.; Huang, C.; You, H.; Liu, R.; Wang, L.; Song, X.; Ding, B.; Walheim, S.; Schimmel, T.; Fang, J. *Nanotechnology* **2012,** *23,* 165604. doi:10.1088/0957-4484/23/16/165604
41. Crommie, M. F.; Lutz, C. P.; Eigler, D. M. *Science* **1993,** *262,* 218–220. doi:10.1126/science.262.5131.218
42. Eigler, D. M.; Schweizer, E. K. *Nature* **1990,** *344,* 524–526. doi:10.1038/344524a0
43. Meyer, G.; Repp, J.; Zöphel, S.; Braun, K.-F.; Hla, S. W.; Fölsch, S.; Bartels, L.; Moresco, F.; Rieder, K. H. *Single Mol.* **2000,** *1,* 79–86. doi:10.1002/(SICI)1438-5171(200004)1:1<79::AID-SIMO79>3.0.CO;2-R
44. Moresco, F.; Meyer, G.; Rieder, K.-H.; Tang, H.; Gourdon, A.; Joachim, C. *Appl. Phys. Lett.* **2001,** *78,* 306–308. doi:10.1063/1.1339251
45. Fuchs, H.; Schimmel, T. *Adv. Mater.* **1991,** *3,* 112–113. doi:10.1002/adma.19910030212
46. Irmer, B.; Blick, R. H.; Simmel, F.; Gödel, W.; Lorenz, H.; Kotthaus, J. P. *Appl. Phys. Lett.* **1998,** *73,* 2051–2053. doi:10.1063/1.122364
47. Schumacher, H. W.; Keyser, U. F.; Zeitler, U.; Haug, R. J.; Eberl, K. *Appl. Phys. Lett.* **1999,** *75,* 1107–1112. doi:10.1063/1.124611
48. Lenhert, S.; Brinkmann, F.; Laue, T.; Walheim, S.; Vannahme, C.; Klinkhammer, S.; Xu, M.; Sekula, S.; Mappes, T.; Schimmel, T.; Fuchs, H. *Nat. Nanotechnol.* **2010,** *5,* 275–279. doi:10.1038/nnano.2010.17
49. Nakamura, M.; Endo, O.; Ohta, T.; Ito, M.; Yoda, Y. *Surf. Sci.* **2002,** *514,* 227–233. doi:10.1016/S0039-6028(02)01634-5
50. Toney, M. F.; Howard, J. N.; Richer, J.; Borges, G. L.; Gordon, J. G.; Melroy, O. R.; Yee, D.; Sorensen, L. B. *Phys. Rev. Lett.* **1995,** *75,* 4472–4475. doi:10.1103/PhysRevLett.75.4472
51. Lucas, C. A.; Marković, N. M.; Ross, P. N. *Phys. Rev. B* **1997,** *56,* 3651–3654. doi:10.1103/PhysRevB.56.3651
52. Wilms, M.; Broekmann, P.; Stuhlmann, C.; Wandelt, K. *Surf. Sci.* **1998,** *416,* 121–140. doi:10.1016/S0039-6028(98)00550-0
53. Spaenig, A.; Broekmann, P.; Wandelt, K. *Z. Phys. Chem.* **2003,** *217,* 459–477. doi:10.1524/zpch.217.5.459.20454

License and Terms

Beilstein Journal of Nanotechnology

Characterization of the mechanical properties of qPlus sensors

Jan Berger*, Martin Švec, Martin Müller, Martin Ledinský, Antonín Fejfar, Pavel Jelínek and Zsolt Majzik

Full Research Paper

Open Access

Address:
Institute of Physics, Academy of Sciences of the Czech Republic, Cukrovarnicka 10, 162 53, Prague, Czech Republic

Email:
Jan Berger* - bergerj@fzu.cz

* Corresponding author

Keywords:
AFM; Cleveland's method; cross talk; force; qPlus; stiffness; STM; thermal noise; tuning fork

Beilstein J. Nanotechnol. **2013,** *4,* 1–9.
doi:10.3762/bjnano.4.1

Received: 14 September 2012
Accepted: 28 November 2012
Published: 02 January 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

In this paper we present a comparison of three different methods that can be used for estimating the stiffness of qPlus sensors. The first method is based on continuum theory of elasticity. The second (Cleveland's method) uses the change in the eigenfrequency that is induced by the loading of small masses. Finally, the stiffness is obtained by analysis of the thermal noise spectrum. We show that all three methods give very similar results. Surprisingly, neither the gold wire nor the gluing give rise to significant changes of the stiffness in the case of our home-built sensors. Furthermore we describe a fast and cost-effective way to perform Cleveland's method. This method is based on gluing small pieces of a tungsten wire; the mass is obtained from the volume of the wire, which is measured by optical microscopy. To facilitate detection of oscillation eigenfrequencies under ambient conditions, we designed and built a device for testing qPlus sensors.

Introduction

The invention of scanning tunneling microscopy [1] and atomic force microscopy (AFM) [2] opened new horizons in characterization and modification of surfaces and nanostructures. STM is routinely used nowadays as a standard technique to characterize and modify objects at the atomic scale. However, its application is limited only to conductive samples as the tunneling current flowing between a probe and a sample is employed as the feedback signal.

This limitation was surpassed by AFM techniques, which are essentially based on force detection acting between the tip and the sample. In principle, this approach can be adapted to any arbitrary sample, independently of its conductivity properties. Consequently AFM techniques have found many applications across different scientific fields, including biology, chemistry and physics. In particular, noncontact atomic force microscopy [3] (nc-AFM) has developed into a powerful technique for

imaging with true atomic resolution [4,5], chemical sensitivity [6-8] or for performing single atom manipulation [9-11] on all types of surfaces, including insulators.

Shortly after the invention of AFM and STM, the first attempt to combine static AFM and STM measurements was made by Dürig et al. in 1986 [12]. A few years later, combined nc-AFM/STM using a Si cantilever was reported for the first time (see [13]), showing the capability to record simultaneous STM and AFM signals with atomic resolution on a metal surface [14]. At the same time F. J. Giessibl introduced so-called qPlus sensors [15], which allow simultaneous acquisition of the tunneling current and the forces with a small oscillation amplitude. This method increases substantially the sensitivity to the tunneling current signal comparing to traditional Si cantilevers. This new approach opens new possibilities in the characterization of surfaces and nanostructures on the atomic scale (see, e.g., [16]). Not surprisingly, qPlus sensors have become frequently used for noncontact measurements nowadays (see, e.g., [7,9,17-21]).

The heart of the qPlus-based AFM/STM microscope is a tuning fork with one prong fixed and the other, with a metallic tip positioned at the very end, freely oscillating. The variation of the resonant frequency of the prong reflects directly the interaction of the sensor with the sample. At the same time, the presence of a metallic tip allows simultaneous acquisition of the average tunneling current flowing between tip and sample when the bias voltage is applied. In principle, the force acting between the tip and sample might be accessible from the detected frequency modulation. However, a reliable estimation of the measured force depends on several factors, among them the proper calibration of the mechanical properties (stiffness) of the sensor.

Originally all mass-produced tuning forks are tuned to the same frequency by laser trimming. However, it has been shown that the sensor fabrication process may strongly alter the stiffness [22,23]. The value of the stiffness can also be influenced by fixing the tip on the prong and by the gluing process. Furthermore, it has been shown that a slight shortening of the oscillating prong may be beneficial for improving electrical sensitivity of the sensor [20]. What is more, when the tunneling current is collected on the tip, an additional wire, usually gold, needs to be attached for STM measurements in order to avoid any interference between the tunneling current and deflection channels [20]. This wiring may have a certain impact on the mechanical properties of the sensor. Therefore to ensure proper estimation of detected forces the mechanical properties of each sensor should be calibrated carefully.

In the past, several methods have been developed to estimate the stiffness of sensors [24]. The first class of methods uses the variation of the resonant frequency before and after the addition of some small mass to the end of the prong. One example of this method is the added-mass (Cleveland's) method [25]. The second method estimates stiffness from the bending of the sensor as a function of the applied force [26]. The third method uses the thermal motion of the sensor to estimate the stiffness (for details see [27]). Additionally, the stiffness can be calculated directly from the elastic properties of the sensor [28]. All methods have some advantages and limitations, as we will discuss later.

The aim of this paper is to compare and critically discuss the following methods for estimating the stiffness of qPlus sensors: (i) the added-mass method; (ii) thermal excitation; and (iii) a method based on the continuum theory of elasticity. In particular we will estimate and compare the stiffness of several home-built and commercial qPlus sensors using different methods. We will also briefly describe a simple testing device for reliable detection of the mechanical properties of sensors under ambient conditions.

Experimental

Sensors

In this work, we used two types of sensors: commercial (Omicron–Oxford instruments) (com) and home-built (hb) sensors. The commercial sensors have a length of ~2.4 mm and the tip is placed on the side of the prong usually around 0.05 mm from the end. No additional wire is used in this design for the tip connection.

The home-built qPlus sensors were built from commercially available tuning forks from Micro Crystal, in the SMD package MS1V-T1K. The original tuning forks were shortened in order to reach higher sensitivity (charge produced by deflection), which allows us to operate with lower amplitudes [20,29]. The schematic figure of the sensor and a detail of a tip mounting is shown in Figure 1. Ceramic plates were used as a support for tuning forks. One side of these plates was covered by a very thin copper layer for shielding [20]. The tip itself is connected to one of the pins of the sensor by a combination of gold (ø = 25 μm) and copper (ø = 500 μm) wire. In order to minimize unnecessary vibrations the copper wire is fixed to the copper shield by nonconductive epoxy (Torr seal). As a result, only the thin gold wire is allowed to vibrate with the motion of the prong. The gold wire is attached to the end of the prong by the same nonconductive epoxy as used before and, as the last step during the sensor construction, the tip is carefully mounted directly to the gold wire by a conductive epoxy (EPO-TEK H21D) in such a way that there is no additional electrical connection to the rest of the prong. Tips are etched from 0.125 μm tungsten wire in 2 M solution of NaOH. We use the

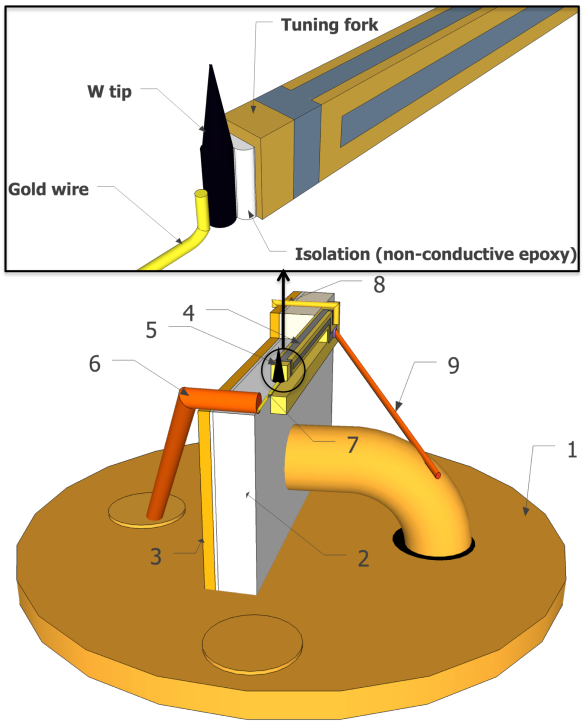

Figure 1: Home-made qPlus sensor consists of the following parts: (1) base stage (from Omicron–Oxford instruments), (2) ceramic plate, (3) copper shield that is glued to the ceramic plate (4) (shortened) tuning fork, (5) tip, (6) copper wire for supporting the gold wire, (7) gold wire used for making connection to the tip, (8) gold wire providing connection between copper shield and the shielding electrode, (9) gold (or copper) wire for collecting the deflection signal [20]. Detail shows the fixing of the tip to the prong and gold wire for collecting the tunneling current.

drop-off method for etching our tips, thus the final tips have the shape of a droplet and they are between 200 and 300 µm long.

Testing device for measurements in ambient conditions

Cleveland's method requires detection of an eigenfrequency. To facilitate these measurements, we constructed a simple device for testing the mechanical properties of qPlus sensors under ambient conditions. The tester has two amplification stages as shown in Figure 2. The first stage contains two current-to-voltage converters (IVC) that are used to generate voltage signals from the small currents produced by two electrodes of the fork during its sinusoidal motion. OP 111 operational amplifiers (OPA) in a TO 99 package were used as IVCs with 100 MΩ feedback resistors. In order to gain the maximum performance of the device, the inputs of amplifiers were brought as close as possible to the outputs of the sensor: the lengths of the connecting wires were about 1 cm only. Furthermore, the SMD-packed feedback resistor was mounted directly between the input and the output legs minimizing the length of wiring to a few millimetres. By this construction the input and the parasitic capacitance of the feedback resistor could be minimized. Note that nowadays the old OP 111 can be replaced by a faster, less noisy OPA (e.g., OP 637, etc.). However in the frequency range of our interest, OP 111 performed satisfactorily.

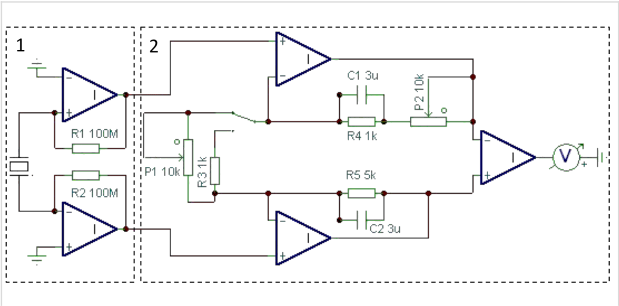

Figure 2: Two-stage amplifier used for testing qPlus sensors. The first stage works as a charge amplifier in the frequency range of our tuning forks. The second stage is a differential amplifier, which allows us to improve the signal-to-noise ratio [30] and can eliminate the driving signal, which is coupled by stray capacitances, from the deflection channel.

The second stage is a differential stage, in which the output voltage of the first stage can be optionally enhanced up to 103 times. Due to the electrode configuration of the prong, the signals from the first stage are in inverse phase with respect to each other. Therefore the application of a differential amplifier leads to a better signal-to-noise ratio [30]. Furthermore, the two OPAs in the first stage were mounted in a way that the stray capacitances are tuned to be nearly the same. By this design, the driving signal is coupled by stray capacitances almost identically (phase and amplitude) to each input of the OPAs. Consequently this unwanted component becomes almost completely nulled out by the differential stage. This procedure can be applied to minimize coupling from other sources, e.g., bias modulation used for Kelvin probe measurements.

A piezo tube is used for mechanical excitation. Apart from the rubber legs of this small instrument there is no additional vibration isolation. Finally, SPECS Nanonis OC-4 PLL was used for data acquisition.

UHV measurements

The measurements requiring precise vibration and sound isolation were carried out in UHV on an Omicron VT XA qPlus AFM/STM at room temperature (RT) at a base pressure below 1×10^{-10} mbar. The resonant frequency v was determined by using the same PLL as for the ambient measurement. A high-quality PC sound card ASUS Xonar Essence ST in combination with a free FFT software Spectrum Lab (Audio Signal Analyser) was used to record thermal-noise density spectra.

Results and Discussion

We carried out a series of measurements of sensors with prong lengths varying from 1.8 to 2.4 mm. Due to the variation of the length and of the mass of the attached tips, the resonant frequency altered between 25 and 60 kHz. For commercial sensors, the tip was attached on the side of the prong, effectively shortening the tuning fork. Consequently two values of stiffness can be considered; one taking into account the whole length of the prong, and another that considers the length defined by the tip mounting position (effective stiffness). Compared to the case of real physical shortening (cutting), the remaining part of the prong adds extra mass, and therefore it has certainly a negative effect on the resonant frequency. The interesting question is whether this kind of tip attachment has a significant effect on the stiffness measurement. In the following part of the text, the effective stiffness will be shown in brackets right after the value corresponding to the total length.

Stiffness calculation from continuum theory of elasticity

The tuning fork can be considered as a pure prism with a rectangular cross section. Therefore the stiffness of the tuning fork can be expressed by using the continuum theory of elasticity [31] as follows:

$$k = \frac{Et^3 w}{4l^3}, \tag{1}$$

where E is Young's elastic modulus (for quartz E = 78.7 GPa), t is the thickness (0.214 mm), w is the width (0.126 mm), l is the length of the prong, and k is the stiffness of the tuning fork. In order to determine the stiffness in this work, the length, which is the only variable, needs to be known. All remaining parts can be considered as constants because (i) we assume the Young's modulus to be constant for quartz tuning forks; and (ii) we found from repeated measurements that variations in t and w are negligible (less than ±2 μm) for our purposes.

The dimensions of the prong were determined by an optical microscope Leica DM LM prior to insertion into UHV. The horizontal resolution of this microscope with objective Leica N PLAN L 50× (NA 0.5) is at least 2 μm. The measured lengths and calculated stiffness values are listed in Table 1. By using the minimal resolution for the microscope, the error of measuring the length of the prong is below 0.1% and errors of the thickness and the width are below 1.6%. The final precision of this method is in the range of 5%.

Added-mass method

Cleveland et al. presented a very intuitive way for precise determination of the stiffness, which is based on the measurement of the change in the resonant frequency. This change is induced by loading an extra mass to the lever, M, which is relatively small compared to the total mass of the oscillator, m^*. The effect of M on the resonant frequency ν can be estimated as follows:

$$\nu = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{M + m^*}}, \tag{2}$$

where ω is the angular velocity. The Equation 2 can be rewritten in a more usable form:

$$M = k\left(2\pi\nu\right)^{-2} - m^*. \tag{3}$$

From Equation 3 we directly see that the loaded mass M is linearly proportional to $(2\pi\nu)^{-2}$. Therefore the dependence of M versus $(2\pi\nu)^{-2}$ should be linear, with a slope corresponding to the stiffness k.

Consequently the stiffness k can be directly calculated from the values of resonant frequencies before (ν_0) and after (ν_1) loading the lever with the mass M:

$$k = \left(2\pi\right)^2 \frac{M}{\left(\frac{1}{\nu_1^2}\right) - \left(\frac{1}{\nu_0^2}\right)}. \tag{4}$$

Because the frequency can be measured with very high precision, the precision of the method relies mainly on the mass measurement. For silicon cantilevers, several ideas have been proposed for how to load them with known masses (see

Table 1: Measured lengths of tested sensors and calculated stiffnesses. In the case of the commercial sensor, the length defined by the tip mounting position and the resulting effective stiffness is shown in brackets. Estimated error is in the range of 5%.

Sensor	hb_1	hb_2	hb_3	hb_4	com
l [mm]	1.853	1.842	1.816	2.314	2.4 (2.35)
k [N/m]	3818 ± 83	3887 ± 78	4056 ± 101	1943 ± 56	1757 ± 43 (1872 ± 43)

[25,32,33]). In our case, the large tuning forks make the task somewhat easier because they can support heavier objects. For this reason, we introduce a method that does not require any expensive microbalance scale or other types of high precision instrument. In this method, we fix a small piece of tungsten wire as close as possible to the tip. The mass of the load was estimated from its dimensions measured with the optical microscope and afterwards calculated using the bulk density of tungsten $\rho = 19.3$ g/cm^3.

The amount of added mass must be chosen carefully as overloading the tuning fork may lead to anharmonicity, and consequently Equation 3 will be not valid anymore. In order to be sure that overloading does not occur, we added several loads to a single sensor and the resonant frequency was measured for each added mass. As it is clearly visible in Figure 3 the linearity is still kept and thus Equation 3 and Equation 4 remain valid in the range of the given masses.

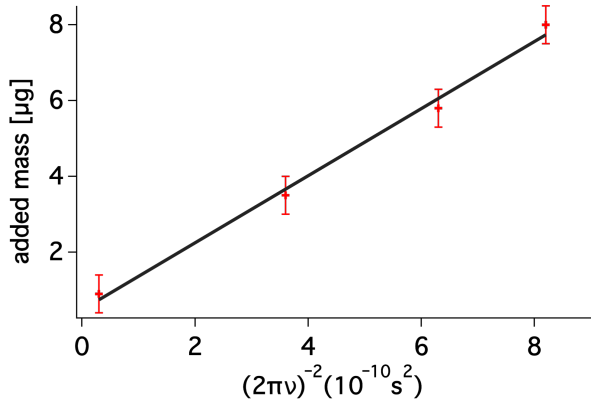

Figure 3: A plot of added mass versus $(2\pi\nu)^{-2}(10^{-10}s^2)$ for a single commercial sensor. A simple linear regression of the measured data gives a stiffness, in this case the slope of linear fit, of 1753 ± 135 N/m.

Naturally, the tungsten wire needs to be fixed in a rigid way. For this purpose, we used the Torr seal® epoxy, which was used for sensor construction as well. On the one hand, experiments showed that the adhesion force between tuning fork and glue was strong enough to hold the tungsten wire during the measurement. On the other hand, the surface of the tuning fork is smooth. Therefore the glue can be easily removed after measuring, without damaging the sensor and voiding its subsequent application. Unquestionably, the glue adds some extra mass that we do not take into account in our method. However, the amount of the glue used to fix the extra wire causes a change in frequency that is negligible compared to the large frequency shift caused by the tungsten wire loading. Figure 4 gives a typical example of the change in the resonant frequencies after adding an extra load.

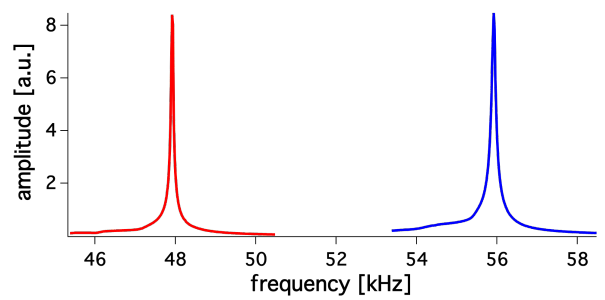

Figure 4: Graph shows two resonant curves. The blue curve corresponds to an unloaded tuning fork with a resonance frequency of 55872 Hz. The red one refers to the same fork end-loaded with 11.43 µg. The resonance frequency shifts to 47787 Hz.

Results obtained by the added mass method for different sensors are summarized in the Table 2. Lengths and diameter of wires used as added masses were about 300 µm and 50 µm, respectively. The optical microscope utilized in this work has a resolution of about 2 µm, from this we estimate an error of the length measurement less than 1% and for the diameter of ≈8%. Consequently, the estimation of the added mass has a maximal error of about 11%. Because the spring constant depends on the added mass linearly and the error arising from the determination of frequencies is negligible, the estimated maximal error of the stiffness from Cleveland's method is also around 11%.

Table 2: Eigenfrequencies of tuning forks before (ν_0) and after (ν_1) loading of an extra mass M to the prong. In the case of the commercial sensor, the length defined by the tip mounting position and the resulting effective stiffness is shown in brackets. Sensor hb_3 was measured with and without gold wire (with the same added mass). These values show that the influence of the gold wire is minor for shortened tuning forks.

Sensor	hb_1	hb_2	hb_3	hb_3–w/o wire	hb_4	com
ν_0 [Hz]	55872	48596	57030	58200	25929	25573
ν_1 [Hz]	47787	43203	46283	46997	23728	23581 (23286)
k [N/m]	3835 ± 211	3846 ± 218	3945 ± 217	3991 ± 220	1857 ± 105	1782 ± 101 (1829 ±103)
M [µg]	11.43	10.96	15.94	15.94	13.60	12.17 (14.62)

Thermal excitation

The harmonic oscillator in equilibrium with its environment fluctuates in response to the thermal noise. The prong of the tuning fork is assumed to behave like a simple harmonic oscillator. The normal stiffness of the qPlus sensor can be related by the equipartition theorem to its thermal energy during vibration, leading to the relationship

$$k = \frac{k_B T}{2A_{th}^2}, \tag{5}$$

where A_{th} is the amplitude of the thermal motion of the free prong, k_B is the Boltzman constant ($k_B = 1.38 \times 10^{-23}$ J·K^{-1}), T is the temperature of the system (in our case it is RT) and k is the stiffness of the tuning fork. This method can be used for stiffness measurement if the oscillation is only thermally excited without any additional excitation (for instance mechanical vibration). This method is the only one mentioned in this paper that is able to estimate the spring constant during the course of an UHV AFM experiment. Generally, UHV AFM/STM instruments are able to reach a high level of vibrational isolation, which is needed to minimize extra mechanical excitation from outside of the system.

The square of the output voltage of the thermally exited tuning fork can be expressed as

$$V_{th}^2 = \int_{\nu_0 - 1/2B}^{\nu_0 + 1/2B} n_{th}^2(\nu)\,d\nu, \tag{6}$$

where n_{th} is the thermal noise density and B is the bandwidth [34]. In Equation 6 the range of the measurement determined by B has to contain all the resonant features. In principle, as long as this rule holds, the bandwidth should not play any role in the measurement. If the peak width is much larger than the bandwidth, inaccurate results are expected.

There is another factor that has to be considered: In Equation 6 the thermal noise density can be obtained from the power noise density, $n_V^2(\nu)$, by subtracting the electrical noise density $n_{el}^2(\nu)$ from it, thus $n_{th}^2(\nu) = n_V^2(\nu) - n_{el}^2$. The electrical noise density, n_{el}^2, should have white-noise character around the resonance, as a consequence, n_{el}^2 does not depend on the frequency. This requirement is usually well fulfilled around the resonant frequency of our homemade sensors.

In order to be able to evaluate the physical magnitude of the oscillation amplitude, the conversion factor ($C = V_{th}/A_{th}$) needs to be determined. C was estimated during dynamic STM measurement on a clean Si(111)-(7 × 7) surface with a precision of ≈10%. It is worth noting that, as the amplitude A_{th} is represented in Equation 5 as a square value, the error in the stiffness measurement will be two times higher than the error in the amplitude.

Stiffness values that were calculated from the thermal noise of the sensor are summarized in Table 3.

Comparison of the methods

We found good agreement between calculated stiffness values using the three methods, as can be seen from a direct comparison in Table 4. The variation in the calculated stiffness values stays within the 10% range for both the continuum theory of elasticity and Cleveland's method; and within 20% range for the thermal-excitation method (see Table 4). Surprisingly, we did not observe a systematic error between results provided by the added-mass method and calculated stiffness values from the continuum theory of elasticity. This finding suggests that the thin gold wire attached to the end of the prong has only a minor effect on the overall stiffness. In the case of a straight Au wire, the stiffness can be obtained from

$$k = \frac{E\pi d^4}{4l^3}, \tag{7}$$

where d is the diameter, l is the length of the wire and E is the Young's modulus (79 GPa for Au). The negligible role of the wire is mainly due to the relatively small diameter (25 µm) of the Au wire. In our case, the typical length of the Au wire is ≈500–600 µm. Using Equation 7, we obtain the stiffness of the gold wire ≈200–160 N/m. This value is much lower than the stiffness of the tuning fork itself (especially of shortened tuning forks). To test the negligible effect of the gold wire on the sensor stiffness in more detail, we repeated the added-mass measurement after the wire had been removed (see Table 4,

Table 3: Stiffness values for home-built sensors measured by thermal-excitation method. Commercial sensors were not tested, because our qPlus system is modified for our home-made sensors only.

Sensor	hb_1	hb_2	hb_3	hb_4
k [N/m]	3650 ± 369	3702 ± 367	3872 ± 382	1779 ± 179

Table 4: Summary of stiffness values obtained by the introduced methods.

Sensor	*l* [mm]	*k* [N/m]		
		Continuum theory of elasticity	Cleveland's method	Thermal excitation
hb_1	1.802	3818 ± 83	3835 ± 211	3650 ± 369
hb_2	1.769	3887 ± 78	3846 ± 218	3702 ± 367
hb_3	1.816	4056 ± 101	3945 ± 217	3872 ± 382
hb_3–w/o wire	–	–	3991 ± 220	–
hb_4	2.314	1943 ± 56	1857 ± 105	1779 ± 179
com	2.400 (2.385)	1757 ± 43 (1872 ± 43)	1782 ± 101 (1829 ± 104)	–

sensor hb_3 and hb_3–w/o wire). From the data we obtained it is evident that the calculated stiffness does not show any significant changes.

From Table 4 we see that the thermal-peak method systematically underestimates the stiffness compared to two other methods. We attribute this error to additional (e.g., mechanical) excitation presented during the measurements. It was shown by Welker et al. [34] that the mechanical noise becomes dominant at low temperature, which would imply a very large error in stiffness calculations based on thermal-noise analysis. However, we performed our calibration procedure at RT, where the thermal excitation is much larger and the quality factor of the tuning fork is also significantly lower (in our case $Q \approx$ 1500–4500). Thus the tuning fork is less responsive to mechanical vibration, although it still cannot be omitted completely. The piezo scanner was connected to the ground during the thermal-noise measurement, therefore we can rule out the possibility of additional excitation by electrical noise.

Finally, we would like to address the question of how the precision of the stiffness would affect experimental force measurements. The force can be expressed by using the Sader formula [35] as follows

$$F(z)=\frac{2k}{\nu_r}\int_z^\infty\left(\left(1+\frac{A^{1/2}}{8\sqrt{\pi(x-z)}}\right)\Delta\nu(x)-\frac{A^{3/2}}{\sqrt{2(x-z)}}\frac{\mathrm{d}\Delta\nu(x)}{\mathrm{d}x}\right)\mathrm{d}x, \quad (8)$$

where $\Delta\nu$ is the frequency shift, ν_r is the resonant frequency, k is the stiffness of the sensor, A is the amplitude of oscillation, F is the tip–surface interaction force, x is the tip–surface distance, and z being the closest tip approach towards the sample.

From Equation 8 it is clear, that an error in the stiffness leads to a proportional but systematic error in the force measurement. In the large amplitude limit, there is a second dominant source of errors in quantitative force analyses, namely the amplitude calibration (see Equation 8).

The other two methods do not depend on the amplitude calibration. They depend only on size measurements, which can be calibrated with very high accuracy by using optical or electron microscopes. Therefore, we suppose that the overall error in the force measurement is smaller with these methods. Especially with the added-mass method because this can take into account other influences on the stiffness such as improper gluing during sensor construction [22]. We have to note here that in the case of our home-built sensors neither the gluing nor the length of the tip seem to have a strong effect on the stiffness. However, due to the lack of usability of the original, commercial sensors, we do not have enough samples to be able to contribute to the previously reported statistics.

Finally the main benefit of the thermal-noise measurement is that it can be performed any time during the actual AFM measurement, and at least at RT we have seen quite a good agreement with the other two methods.

Conclusion

In this paper we applied and compared three different methods for the estimation of the stiffness of qPlus sensors. Our analysis showed that all three methods give very similar results varying within a range of 10% or 14% for the thermal-excitation method. Surprisingly, the added-mass (Cleveland's) method gave very similar results to calculated values from the continuum theory of elasticity. This finding suggests that a gold wire does not strongly alter the stiffness. Furthermore, we have proved that the stiffness can be obtained with reasonably small error from the thermal noise measurement even at RT.

In addition we designed and built a device for testing qPlus sensors. It can be used for testing sensors before inserting into the chamber and to obtain information about mechanical properties of sensors such as the resonant frequency, quality factor and stiffness. We also discussed a fast and cost-effective way to perform the added-mass method under ambient conditions. This method is based on adding small pieces of tungsten wire whose mass was determined from the volume of the wire. The wire can

be easily removed after the measurement without destroying the sensor.

Acknowledgements

The authors thank F. J. Giessibl and A. Bettac for fruitful discussions. This work was financially supported by GAAV under project no. IAA100100905 and GAČR under project no. 204/10/0952. We also acknowledge the financial support of the Moremit project FR-TI2/736 financed by the Czech Ministry of Industry and Trade and LNSM infrastructure financed by the Ministry of Education.

References

1. Binnig, G.; Rohrer, H.; Gerber, C.; Weibel, E. *Phys. Rev. Lett.* **1982,** *49,* 57–61. doi:10.1103/PhysRevLett.49.57
2. Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986,** *56,* 930–933. doi:10.1103/PhysRevLett.56.930
3. Albrecht, T. R.; Grütter, P.; Horne, D.; Rugar, D. *J. Appl. Phys.* **1991,** *69,* 668–673. doi:10.1063/1.347347
4. Giessibl, F. J. *Science* **1995,** *267,* 68–71. doi:10.1126/science.267.5194.68
5. Barth, C.; Reichling, M. *Nature* **2001,** *414,* 54–57. doi:10.1038/35102031
6. Sugimoto, Y.; Pou, P.; Abe, M.; Jelinek, P.; Pérez, R.; Morita, S.; Custance, Ó. *Nature* **2007,** *446,* 64–67. doi:10.1038/nature05530
7. Setvín, M.; Mutombo, P.; Ondráček, M.; Majzik, Z.; Švec, M.; Cháb, V.; Ošťádal, I.; Sobotik, P.; Jélinek, P. *ACS Nano* **2012,** *6,* 6969. doi:10.1021/nn301996k
8. Teobaldi, G.; Lämmle, K.; Trevethan, T.; Watkins, M.; Schwarz, A.; Wiesendanger, R.; Shluger, A. L. *Phys. Rev. Lett.* **2011,** *106,* 216102. doi:10.1103/PhysRevLett.106.216102
9. Ternes, M.; Lutz, C. P.; Hirjibehedin, C. F.; Giessibl, F. J.; Heinrich, A. J. *Science* **2008,** *319,* 1066–1069. doi:10.1126/science.1150288
10. Sugimoto, Y.; Pou, P.; Custance, O.; Jelinek, P.; Abe, M.; Perez, R.; Morita, S. *Science* **2008,** *322,* 413–417. doi:10.1126/science.1160601
11. Hirth, S.; Ostendorf, F.; Reichling, M. *Nanotechnology* **2006,** *17,* 148–154. doi:10.1088/0957-4484/17/7/S08
12. Dürig, U.; Gimzewski, J. K.; Pohl, D. W. *Phys. Rev. Lett.* **1986,** *57,* 2403–2406. doi:10.1103/PhysRevLett.57.2403
13. Howald, L.; Meyer, E.; Lüthi, R.; Haefke, H.; Overney, R.; Rudin, H.; Güntherodt, H.-J. *Appl. Phys. Lett.* **1993,** *63,* 117–120. doi:10.1063/1.109732
14. Loppacher, C.; Bammerlin, M.; Guggisberg, M.; Schär, S.; Bennewitz, A.; Baratoff, A.; Meyer, E.; Güntherodt, H.-J. *Phys. Rev. B* **2000,** *62,* 16944. doi:10.1103/PhysRevB.62.16944
15. Giessibl, F. J. *Appl. Phys. Lett.* **2000,** *76,* 1470–1472. doi:10.1063/1.126067
16. Ternes, M.; González, C.; Lutz, C. P.; Hapala, P.; Giessibl, F. J.; Jelínek, P.; Heinrich, A. J. *Phys. Rev. Lett.* **2011,** *106,* 016802. doi:10.1103/PhysRevLett.106.016802
17. König, T.; Simon, G. H.; Rust, H.-P.; Heyde, M. *Appl. Phys. Lett.* **2009,** *95,* 083116. doi:10.1063/1.3189282
18. Fournier, N.; Wagner, C.; Weiss, C.; Temirov, C.; Tautz, F. S. *Phys. Rev. B* **2011,** *84,* 035435. doi:10.1103/PhysRevB.84.035435
19. Sun, Y.; Mortensen, H.; Schär, S.; Lucier, A.-S.; Miyahara, Y.; Grütter, P.; Hofer, W. *Phys. Rev. B* **2005,** *71,* 193407. doi:10.1103/PhysRevB.71.193407
20. Majzik, Z.; Setvín, M.; Bettac, A.; Feltz, A.; Cháb, V.; Jelínek, P. *Beilstein J. Nanotechnol.* **2012,** *3,* 249–259. doi:10.3762/bjnano.3.28
21. Hembacher, S.; Giessibl, F. J.; Mannhart, J.; Quate, C. F. *Phys. Rev. Lett.* **2005,** *94,* 056101. doi:10.1103/PhysRevLett.94.056101
22. Sweetman, A.; Jarvis, S.; Danza, R.; Bamidele, J.; Gangopadhyay, S.; Shaw, G. A.; Kantorovich, L.; Moriarty, P. *Phys. Rev. Lett.* **2011,** *106,* 136101. doi:10.1103/PhysRevLett.106.136101
23. Shaw, G. A.; Pratt, J. R.; Kubarych, Z. J. Small mass measurements for tuning fork-based force microscope cantilever spring constant calibration. In *Proceedings of the Society for Experimental Mechanics,* Indianapolis, IN, June 7–10, 2010; 2011. NIST.
24. Kennedy, S. J.; Cole, D. G.; Clark, R. L. *Rev. Sci. Instrum.* **2009,** *80,* 125103. doi:10.1063/1.3263907
25. Cleveland, J. P.; Manne, S.; Bocek, D.; Hansma, P. K. *Rev. Sci. Instrum.* **1993,** *64,* 403. doi:10.1063/1.1144209
26. Gao, S.; Zhang, Z.; Wu, Y.; Herrmann, K. *Meas. Sci. Technol.* **2009,** *21,* 015103. doi:10.1088/0957-0233/21/1/015103
27. Cole, D. G. *Meas. Sci. Technol.* **2008,** *19,* 125101. doi:10.1088/0957-0233/19/12/125101
28. Lübbe, J.; Doering, L.; Reichling, M. *Meas. Sci. Technol.* **2012,** *23,* 045401. doi:10.1088/0957-0233/23/4/045401
29. Wutscher, E.; Giessibl, F. J. *Rev. Sci. Instrum.* **2011,** *82,* 093703. doi:10.1063/1.3633950
30. Morita, S.; Giessibl, F. J.; Wiesendanger, R. *Noncontact Atomic Force Microscopy;* Springer, 2009.
31. Chen, C. J. *Introduction to Scanning Tunneling Microscopy;* Oxford University Press, 1993.
32. Golovko, D. S.; Haschke, T.; Wiechert, W.; Bonaccurso, E. *Rev. Sci. Instrum.* **2007,** *78,* 043705. doi:10.1063/1.2720727
33. Gibson, C. T.; Weeks, B. L.; Abell, C.; Rayment, T.; Myhra, S. *Ultramicroscopy* **2003,** *97,* 113–118. doi:10.1016/S0304-3991(03)00035-4
34. Welker, J.; de Faria Elsner, F.; Giessibl, F. J. *Appl. Phys. Lett.* **2011,** *99,* 084102. doi:10.1063/1.3627184
35. Sader, J. E.; Jarvis, S. P. *Appl. Phys. Lett.* **2004,** *84,* 1801. doi:10.1063/1.1667267

Beilstein Journal of Nanotechnology

Calculation of the effect of tip geometry on noncontact atomic force microscopy using a qPlus sensor

Julian Stirling*[1] and Gordon A. Shaw[2]

Full Research Paper

Open Access

Address:
[1]School of Physics and Astronomy, The University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom and [2]Physical Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA

Email:
Julian Stirling* - ppxjs1@nottingham.ac.uk

* Corresponding author

Keywords:
atomic force microscopy; force spectroscopy; lateral forces; mechanical vibrations; qPlus

Beilstein J. Nanotechnol. **2013,** *4,* 10–19.
doi:10.3762/bjnano.4.2

Received: 10 October 2012
Accepted: 02 December 2012
Published: 08 January 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

In qPlus atomic force microscopy the tip length can in principle approach the length of the cantilever. We present a detailed mathematical model of the effects this has on the dynamic properties of the qPlus sensor. The resulting, experimentally confirmed motion of the tip apex is shown to have a large lateral component, raising interesting questions for both calibration and force-spectroscopy measurements.

Introduction

From imaging of individual chemical bonds [1] to subatomic resolution of the structure of the tip apex [2], many experiments have demonstrated the ability of qPlus atomic force microscopy (AFM) to produce unprecedented imaging resolution. Other qPlus studies have measured both the forces necessary to perform atomically precise manipulation [3-5], and the strength of both atomic and molecular interactions [6,7]. As with all forms of AFM, image resolution and force measurements ultimately depend on the structure of the last few angstroms of the tip apex [1,2,4,6-9]. The qPlus sensor is unusual for an AFM sensor in that it is constructed from a quartz tuning fork and has a tip comparable in length to the tine [10]. Hence, the macroscopic geometry of the tip also cannot be ignored [11].

The spring constant of the sensor must be known for any conversion from raw data to meaningful force measurements [12]. However, despite the quoted piconewton precision of qPlus measurements [3], the spring constant is often left unmeasured and is assumed to be $k \approx 1800\ \mathrm{N{\cdot}m^{-1}}$ from the geometry of the bare tine [10]. Measurements of the spring constants of qPlus sensors have produced conflicting results [4,13], which

highlights the need for more detailed analysis. Tung et al. [11] have shown that the dimensions of conical tips also have large effects on the dynamics of higher eigenmodes, suggesting that careful consideration of tip geometry is necessary for sensors operated in the second eigenmode or above.

In this paper we go further, providing a detailed analytical solution for the deflection, elastic potential energy, and spring constant of the tine of a qPlus sensor, for an arbitrary tip geometry. Furthermore, we derive the resulting lateral component of the motion of the tip apex, showing it, both theoretically and experimentally, to be comparable to the amplitude of motion normal to the surface even in the first eigenmode. This differs from previous work considering the lateral motion of qPlus sensors, such as Heyde et al. [14], as the lateral motion arises directly from the eigenmode used for imaging, rather than a parallel or torsional eigenmode. As such the lateral motion will not reveal itself by producing double images. The lateral motion has major implications for force measurements: whereas standard analysis assumes sensitivity to normal forces only [15], our analysis shows that both normal and lateral tip–sample interactions are sampled during qPlus imaging in the first eigenmode. We calculate the effect the lateral motion has on the measured frequency shift, and hence, how this affects calibration, imaging, and spectroscopy.

Results and Discussion

Modelling the tine of the qPlus sensor as an Euler–Bernoulli beam [16] of length L, we can write

$$EI\frac{\partial^4 Z(x,t)}{\partial x^4} = f(x,t) - \rho A\frac{\partial^2 Z(x,t)}{\partial t^2}, \tag{1}$$

where E, I, ρ, and A are the Young's modulus, second moment of area, density, and cross-sectional area of the tine, respectively. $f(x,t)$ is the external force per unit length acting on the tine, and $Z(x,t)$ is the deflection along the length of the tine. Separating the spatial ($\Phi_i(x)$) and temporal ($\mathcal{T}_i(t)$) components of the deflection for all eigenmodes, i, we can write

$$Z(x,t) = \sum_{i=1}^{\infty}\Phi_i(x)\mathcal{T}_i(t), \tag{2}$$

where

$$\mathcal{T}_i = C_i\sin(\omega_i t + \delta_i). \tag{3}$$

As Equation 1 is fourth-order spatially, we get the following general spatial solution:

$$\Phi_i(x) = b_1\cos(\beta_i x) + b_2\sin(\beta_i x) + b_3\cosh(\beta_i x) + b_4\sinh(\beta_i x) \tag{4}$$

where β_i is the spatial frequency of the ith mode governed by

$$\beta_i^4 = \frac{\rho A\omega_i^2}{EI}. \tag{5}$$

The tip connected to the tine, with mass m_{tip} and moment of inertia about the point of rotation of the tip $\mathcal{I}_{\text{tip}}$ (A stylised $\mathcal{I}$ is used to differentiate between moments of inertia and moments of area I), will produce a resulting force of

$$m_{\text{tip}}\frac{\partial^2 Z(L,t)}{\partial t^2}$$

and torque of

$$-\mathcal{I}_{\text{tip}}\frac{\partial^3 Z(L,t)}{\partial x\partial^2 t}.$$

Thus, we can write the spatial boundary conditions as:

$$\Phi_i(0) = 0 \tag{6}$$

$$\frac{\partial\Phi_i(0)x}{\partial x} = 0 \tag{7}$$

$$EI\frac{\partial^2\Phi_i(L)}{\partial x^2} = \mathcal{I}_{\text{tip}}\omega_i^2\frac{\partial\Phi_i(L)}{\partial x} \tag{8}$$

$$EI\frac{\partial^3\Phi_i(L)}{\partial x^3} = -m_{\text{tip}}\omega_i^2\Phi_i(L) \tag{9}$$

as the tine is fixed at $x = 0$. Under the assumption of uniform tip density we can write $\mathcal{I}_{\text{tip}} = Jm_{\text{tip}}$, where J is a constant with units of length squared. The very small effect of the motion, normal to the tine, of the center of mass of the tip due to the tip rotation has been ignored, as this significantly increases the complexity of the problem for a negligible component of the motion unless the width of the tip becomes significant compared to its entire length, including the portion attached to the end of the tine.

In Appendix A (see Supporting Information File 1), we apply these conditions to solve the spatial equation. From Equation 36 and Equation 37, the conditions $b_1 = -b_3$ and $b_2 = -b_4$ are apparent, and the ratio of b_1 to b_2 can be found from Equation 38. From this ratio we obtain the analytical form of the spatial solution as

$$\begin{aligned}\Phi_i(x) = &\left(\sin(\beta_i L)+\sinh(\beta_i L)-\frac{J\beta_i^2}{\gamma_i}\left(-\cos(\beta_i L)+\cosh(\beta_i L)\right)\right)\\ &\times\left(\cos(\beta_i x)-\cosh(\beta_i x)\right)\\ &-\left(\cos(\beta_i L)+\cosh(\beta_i L)-\frac{J\beta_i^2}{\gamma_i}\left(\sin(\beta_i L)+\sinh(\beta_i L)\right)\right)\\ &\times\left(\sin(\beta_i x)-\sinh(\beta_i x)\right)\end{aligned} \tag{10}$$

where γ_i is defined in Equation 40.

Effects on dynamic properties

As shown by Butt and Jaschke [17], properties of a cantilever (or tine) such as the dynamic spring constants and the proportion of energy in each eigenmode at thermal equilibrium can be found by considering the elastic potential energy of the tine. This becomes particularly important if calibration is done by thermal tuning, or if force measurements are performed by using higher modes [11]. The elastic potential energy of the tine is given by

$$W = \frac{EI}{2}\int_0^L\left(\frac{\partial^2 Z(x,t)}{\partial x^2}\right)^2 \mathrm{d}x\,. \tag{11}$$

Again, separating into spatial and temporal components gives

$$W = \frac{EI}{2}\int_0^L\sum_{i=1}^{\infty}\left(\mathcal{I}_i{}^2(t)\frac{\partial^2\Phi_i(x)}{\partial x^2}\right)^2 \mathrm{d}x\,. \tag{12}$$

In Appendix B (see Supporting Information File 1), we show that, just as for bare cantilevers [17], this reduces to

$$W = \frac{EI}{2}\int_0^L\sum_{i=1}^{\infty}\left(\mathcal{I}_i{}^2(t)\frac{\partial^2\Phi_i(x)}{\partial x^2}\right)^2 \mathrm{d}x\,. \tag{13}$$

For brevity we define $\Lambda_i = \int_0^L\left(\frac{\partial^2\Phi_i(x)}{\partial x^2}\right)^2 \mathrm{d}x$.

This allows us to write the average elastic potential energy for each eigenmode as

$$\langle W_i\rangle = \frac{EI}{2}\left\langle\mathcal{I}_i{}^2(t)\right\rangle\Lambda_i\,. \tag{14}$$

In thermal equilibrium, the equipartition theorem requires that $\langle W_i\rangle = \frac{1}{2}k_\mathrm{B}T$. But, as $\left\langle Z_i^2(L,t)\right\rangle = \left\langle\mathcal{I}_i{}^2(t)\right\rangle\Phi_i^2(L)$, and as we include the static spring constant from Euler–Bernoulli beam theory, $k_\mathrm{stat} = 3EI/L^3$, we obtain the mean square deflection of each mode in thermal equilibrium as

$$\left\langle Z_i^2(L,t)\right\rangle = \frac{3k_\mathrm{B}T\Phi_i^2(L)}{L^3k_\mathrm{stat}\Lambda_i}\,. \tag{15}$$

The full analytical form of $L^3\Lambda_i/\Phi_i^2(L)$ is derived in Appendix C (see Supporting Information File 1). Combined with values for β_i (solutions of Equation 42 in Appendix A, see Supporting Information File 1), this could be used to measure the static spring constant of a qPlus sensor with a well-defined tip geometry by thermal tuning. However, thermal tuning of qPlus sensors with spring constants on the order of 2 kN·m^{-1} remains a challenging experimental task as the rms amplitude of thermal excitation at 300 K is ≈1.4 pm.

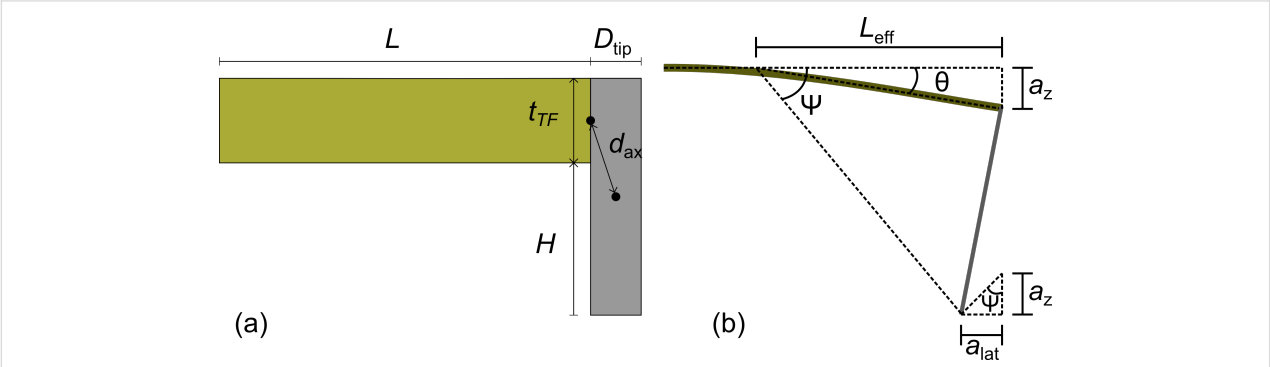

Figure 1: (a) Diagram of theoretical model. d_{ax} is the distance from the centre of mass of the tip to the axis of tip rotation. The tine of the tuning fork is assumed to be clamped at the left-hand side. (b) Geometrical diagram of the tip motion as the cantilever deforms. Note the two ψ angles are not identical in this diagram as θ is too large due to the exaggerated amplitude of deflection, a_z.

For well-calibrated force measurements, the dynamic spring constant for the excited eigenmode k_i must be calculated. By considering the equipartition theorem again, but with the dynamic spring constant and Hooke's law, $\frac{1}{2}k_{\mathrm{B}}T=\frac{1}{2}\left\langle Z_i^2(L,t)\right\rangle k_i$. Combining with Equation 15 this gives

$$\frac{k_i}{k_{\mathrm{stat}}}=\frac{L^3\Lambda_i}{3\Phi_i^2(L)}. \tag{16}$$

Thus, from the results in Appendix C (see Supporting Information File 1), as with the thermal tuning, the dynamic spring constant for any eigenmode can be calculated, provided the tip geometry is well defined, and the static spring constant is known.

The tine of a qPlus sensor was modeled as a quartz beam of length L = 2.4 mm, width w = 130 µm, and thickness t_{TF} = 214 µm (Figure 1a). The tip was modeled as a tungsten cylinder of diameter D_{tip} attached to the end of the tine of the tuning fork. The dynamics of the tip bending have not been included as the model is only of interest at the eigenfrequencies of the loaded tine. The tip protrudes H from the tine giving the tip a total length of $H + t_{\mathrm{TF}}$. The axis of rotation is located at the center of the join between the tip and cantilever. By the parallel-axis theorem, the moment of inertia, $\mathcal{I}_{\mathrm{tip}}$, can be calculated as $\mathcal{I}_{\mathrm{tip}}=\mathcal{I}_{\mathrm{tip}}^{\mathrm{COM}}+d_{\mathrm{ax}}^2m_{\mathrm{tip}}$, where $\mathcal{I}_{\mathrm{tip}}^{\mathrm{COM}}$ is the moment of inertia through the center of mass of the tip and d_{ax} is the distance from the axis of rotation to the center of mass. The distance to the axis is $d_{\mathrm{ax}}=\left[\left(H/2\right)^2+\left(D_{\mathrm{tip}}/2\right)^2\right]^{1/2}$. Thus, the moment of inertia about the axis of rotation is

$$\mathcal{I}_{\mathrm{tip}}=m_{\mathrm{tip}}J=\frac{m_{\mathrm{tip}}}{48}\left(15D_{\mathrm{tip}}^2+4\left(4H^2+2Ht_{\mathrm{TF}}+t_{\mathrm{TF}}^2\right)\right). \tag{17}$$

Using this model, dynamic spring constants were calculated for the first four eigenmodes, relative to k_{stat} (Figure 2). The first two eigenmodes agree qualitatively with experimentally verified Hamiltonian calculations by Tung et al. for the first two eigenmodes of conical tips [11]. The sudden rises to infinite spring constant correspond to when a node of the vibrational mode is located at the end of the cantilever.

Resulting lateral motion

The model presented above allows us to calculate the angle the tip rotates through during the oscillation by using

$$\theta=\arctan\left(\frac{\partial^2\Phi(L)}{\partial x^2}\frac{a_z}{\Phi(L)}\right) \tag{18}$$

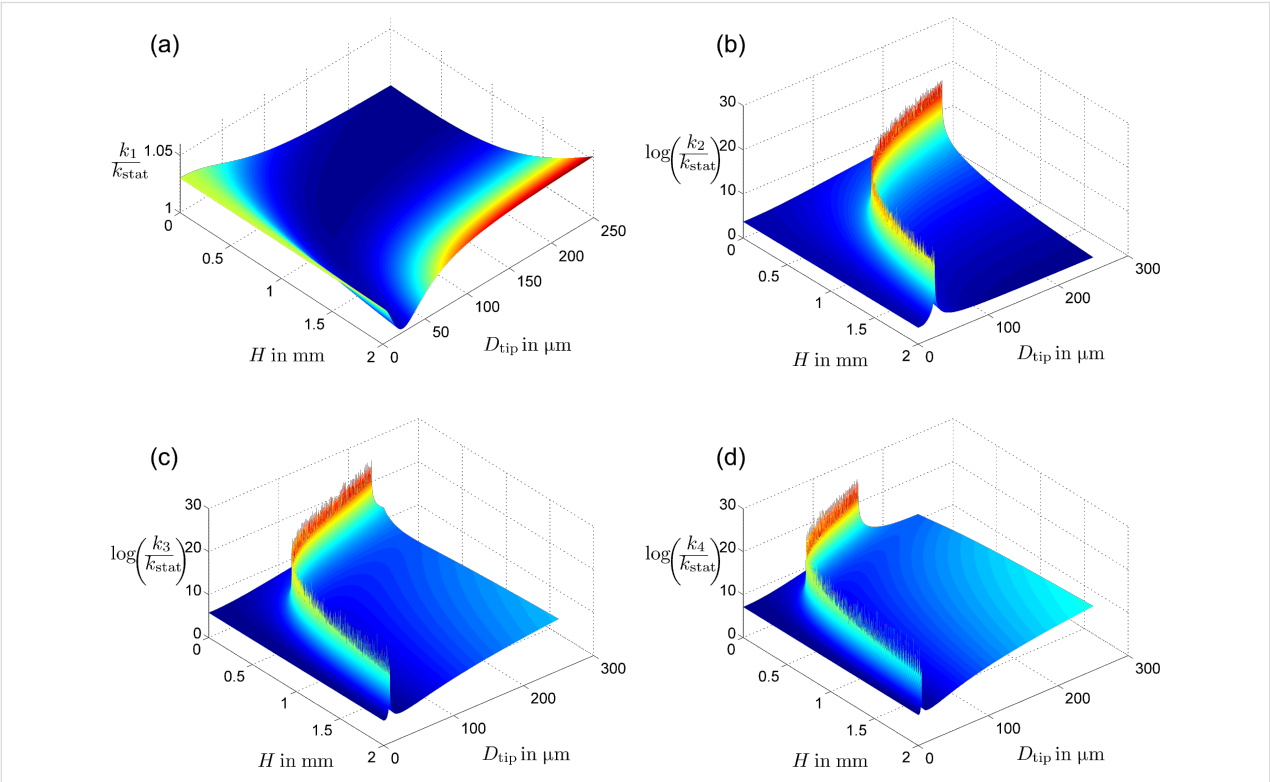

Figure 2: The ratio of dynamic spring constants k_n to the cantilever static spring constant k_{stat} for n = 1,2,3,4, plotted for a range of tip lengths and diameters. The sudden increases in the higher eigenmodes result from nodes positioned at the end of the tip resulting in infinite spring constants.

$$= \arctan\left(a_z \frac{\beta_i \sin(\beta_i L)\sinh(\beta_i L)}{\Xi}\right) \quad (19)$$

where a_z is the amplitude of oscillation, and Ξ is a dimensionless parameter defined as

$$\Xi = -\beta_i^3 JLm^* - \cos(\beta_i L)\sinh(\beta_i L) + \cosh(\beta_i L)\left(\beta_i^3 JLm^* \cos(\beta_i L) - \sin(\beta_i L)\right). \quad (20)$$

m^* is the ratio of tip mass to the mass of the tine.

This has been plotted in Figure 3a for the first eigenmode with an amplitude of 0.5 nm. The angle is extremely small, about $(2 \times 10^{-5})°$, and nearly constant for different tip geometries. However, the lateral motion at the end of the tip will have an amplitude of

$$a_{\text{lat}} = \left(H + \frac{t_{\text{TF}}}{2}\right)\sin\theta = \left(H + \frac{t_{\text{TF}}}{2}\right)\sin\left(\arctan\left(\frac{\partial\Phi(L)}{\partial x}\frac{a_z}{\Phi(L)}\right)\right) \quad (21)$$

$$\approx \frac{\partial\Phi(L)}{\partial x}\frac{a_z\left(H + t_{\text{TF}}/2\right)}{\Phi(L)} \quad (22)$$

$$= a_z\left(H + \frac{t_{\text{TF}}}{2}\right)\frac{\beta_i \sin(\beta_i L)\sinh(\beta_i L)}{\Xi} \quad (23)$$

The calculated amplitude of lateral vibrations has been plotted as a function of H in Figure 3b, for D_{tip} = 50 µm, as the angular dependence on the tip diameter is relatively small. It is apparent that this lateral motion can be significant compared with the normal motion, even for relatively short tips, reaching an equal amplitude at H = 1.389 mm. We can define a, the total ampli-

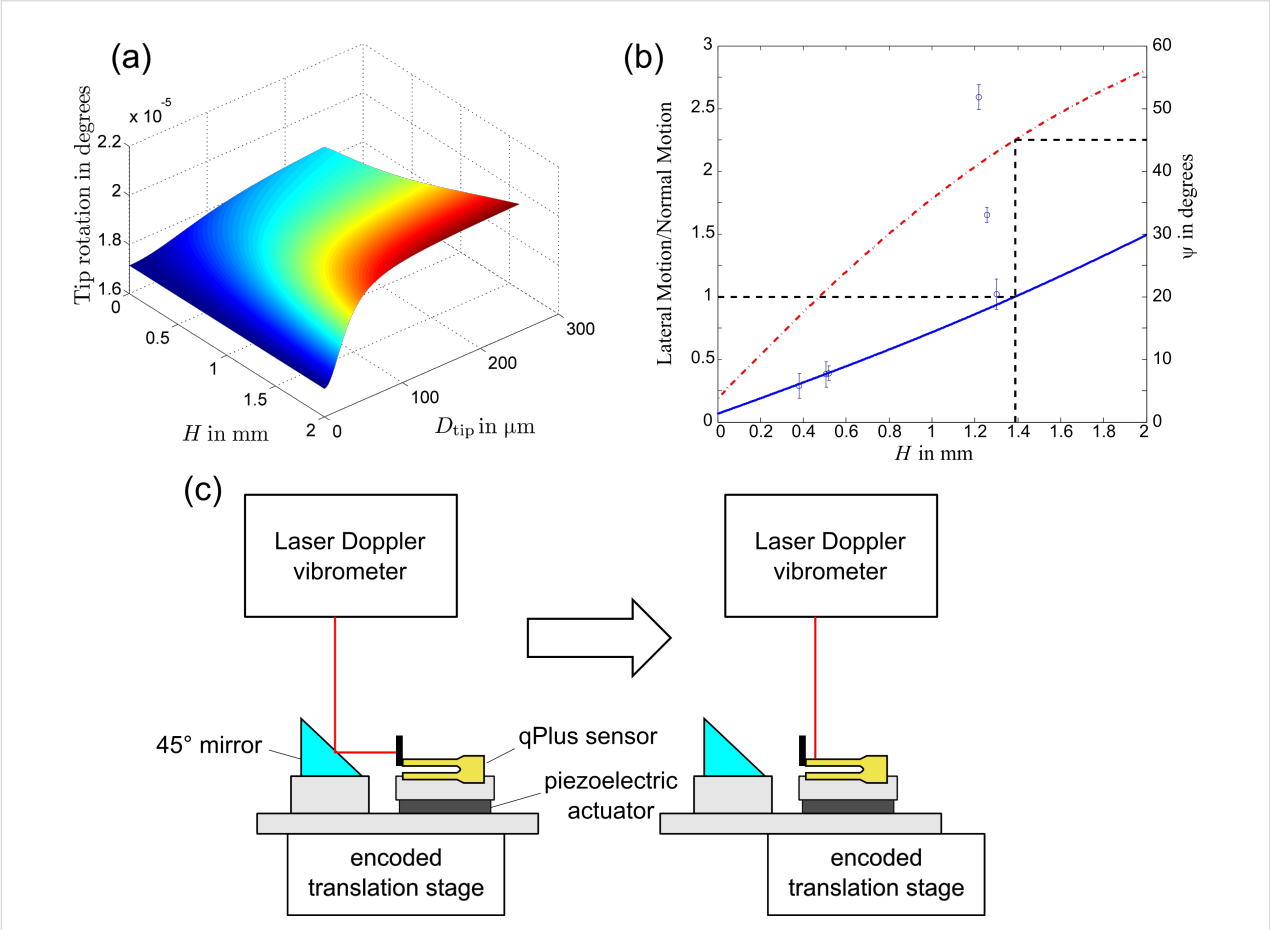

Figure 3: Effect of tip geometry on tip movement for an arbitrary oscillation amplitude. (a) Angle of tip rotation plotted for varying tip sizes. (b) Amplitude of lateral motion of the tip apex plotted (solid line) for varying lengths with a tip diameter of 50 µm. Circular data points represent experimentally measured values. Angle of resulting motion, ψ of the tip apex is also plotted (dot-dashed line). The dashed lines represent where the lateral amplitude equals the normal amplitude. Error bars represent 1 standard error. (c) Schematic of the experimental setup for measuring both normal and lateral motion of the qPlus sensor.

tude of oscillation at the apex, as

$$a = \sqrt{a_z^2 + a_{\text{lat}}^2}\ . \tag{24}$$

Defining the angle of tip motion as ψ, where ψ = 0 corresponds to oscillation normal to the surface and ψ = 90° to oscillation parallel to the surface, we obtain

$$\psi = \arctan\left(\frac{a_{\text{lat}}}{a_z}\right). \tag{25}$$

It is important to note that as the angle of tip rotation is so small, the motion of the tip apex should not be thought of as circular motion, but instead the tip apex is moving over a linear trajectory at an angle ψ to the surface, with negligible rotation.

Experimental validation

The resulting lateral motion was measured experimentally for Omicron Nanotechnology qPlus sensors excited mechanically from below by a piezoelectric actuator (Figure 3c). The actuator was driven by a digital lock-in amplifier (Perkin Elmer 7280 DSP). The deflection of the qPlus sensor was measured by using a laser Doppler vibrometer (Polytec OFV-522) connected to a lock-in amplifier. The qPlus sensor and piezoelectric actuator was mounted on an encoded translation stage allowing deflection measurements to be recorded in multiple positions on the sensor. A 45° mirror positioned near the qPlus sensor allowed lateral deflection measurements of the tip to be made without remounting the sensor or interrupting the excitation, thus limiting changes to the transfer function.

Tungsten wire of 50 μm in diameter was attached to the end of the bare qPlus sensor with EPO-TEK H21D electrically conductive silver epoxy. The tip wire was then etched to different lengths by using a Gamry potentiostat operating as a chronocoulometer with a platinum mesh counter electrode and a saturated calomel reference electrode (SCE). The integrated etch current allowed some degree of control over the quantity of material removed from the tip [18]. The sensor was positioned with a micromanipulator such that the 1 M KOH electrolyte contacted only the tungsten probe tip, then a potential of 0.3 V versus SCE was applied until the desired charge was accumulated from the Faradaic etch current, and the sensor was then removed from the solution. A subsequent five-minute rinse in deionized water was performed in a similar fashion, without application of an external potential.

For measurements of both normal and lateral motion the translation stage was used to collect multiple deflection readings at different positions along the sensor and tip. These were then extrapolated to get the deflection at the end of the sensor and tip despite poor reflectivity at both regions of the sensor not covered by gold electrodes and at the tip apex where it was etched.

The ratio of lateral to normal motion was measured for six sensors. Tip lengths were measured with an optical microscope and an encoded translation stage. These results are presented in Figure 3b. Four of the six sensors show good agreement with the theoretical curve. The two remaining sensors show significantly higher lateral motion. This is most likely due to compliance in the epoxy used to connect the tip, allowing some rotation of the tip relative to the end of the sensor. These results confirm the prediction of large lateral motions on qPlus sensors.

Lateral amplitudes of this magnitude, at first glance, could be thought to limit the resolution of the AFM. However, it is important to consider that the lateral motion is perfectly correlated to the normal motion, unlike lateral motion resulting from torsional modes of the cantilever, which will oscillate with much higher frequency causing blurring of the image. Simulated images and spectra can be generated to theoretically calculate the effect; however, one must carefully consider both the amplitude calibration and the methods for calculating frequency shifts from a potential before continuing.

Effect on frequency shift

Under the assumption that the direction of motion of the tip apex is parallel to the motion of the end of the cantilever (or tine) it can be shown [9] that

$$\Delta f = -\frac{f_0}{k a_z^2}\left\langle F_{\text{ts}} Z\left(L\right)\right\rangle, \tag{26}$$

where $Z(L) = a_z\cos(\omega_0 t)$, and F_{ts} is the force due to the tip–sample interaction.

In the case that the motion of the tip is not parallel to the oscillation of the cantilever, more care must be taken. Equation 26 can be derived from Newton's second law in the reference frame of the end of the cantilever

$$m_{\text{eff}}\ \frac{\mathrm{d}^2 Z\left(L\right)}{\mathrm{d}t^2} = -k_1 Z\left(L\right) + F_{\text{ts}}^{\text{C}}\left(Z\left(L\right)\right), \tag{27}$$

where F_{ts}^{C} is the force due to the tip–sample interaction as experienced at the end of the cantilever, and m_{eff} is the effective mass of the tip and cantilever. Thus, if Equation 26 is modified

to include lateral oscillations, the amplitude terms will remain as a_z, as this is the oscillation amplitude of the beam. However, the tip–sample force must be modified from the interaction at the tip apex to the resulting force at the end of the cantilever.

For an amplitude of a_z, the end of the cantilever is angled by θ at its maximum deflection, as shown in Figure 1b, which can be treated as circular motion about an effective pivot at a distance of $L_{\text{eff}} = a_z/\tan\theta$ from the end of the cantilever. Combining with Equation 21 and Equation 25, and equating sin θ to tan θ due to the very small angle we can show that

$$L_{\text{eff}} = \frac{H}{\tan\psi}. \tag{28}$$

Thus, the bending can be described by a torque of

$$\tau_B = \frac{H}{\tan\psi} F_{\text{ts}}^{\text{C}}. \tag{29}$$

In the reference frame of the tip apex the distance to the effective pivot is $H/\sin\psi$ by simple geometry, and the measured tip–sample force F_{ts} is perpendicular to the vector from the tip apex to the effective pivot, such that

$$\tau_B = \frac{H}{\sin\psi} F_{\text{ts}}. \tag{30}$$

Hence,

$$F_{\text{ts}}^{\text{C}} = \frac{F_{\text{ts}}}{\cos\psi}, \tag{31}$$

and so the frequency shift can be calculated as

$$\Delta f = -\frac{f_0}{k a_z^2}\left\langle \frac{F_{\text{ts}}}{\cos\psi} Z(L) \right\rangle. \tag{32}$$

In the case that the lateral force is zero,

$$F_{\text{ts}} = F_{\text{ts}}^{z} \cos\psi \tag{33}$$

where F_{ts}^{z} is the z component of the tip–sample force. Thus, if the calibrated amplitude of the oscillation is a_z, rather than the total amplitude of oscillation at the tip apex a, then Δf is equal to the expected result for tip motion parallel to the cantilever oscillation. However, if lateral forces are present, then these will also affect the frequency shift.

Effect on calibration

Amplitude calibration in qPlus AFM is usually performed by measuring the z extension needed to maintain a constant value for $\Delta f a_z^{3/2}$ [10]. As it can be shown for large amplitudes, by inserting

$$F_{\text{ts}}(q) = -Cq^{-n} \tag{34}$$

into Equation 32, where q is the position of the tip apex relative to the surface, that

$$\Delta f \approx \frac{f_0 C}{\sqrt{2}\pi\cos\psi k a_z^{3/2} d^{n-1/2}} I_1(n). \tag{35}$$

This follows from using the force conversion in Equation 31, and applying the method from [15]. $I_1(n)$ is an integral, dependent on n, but constant for the experiment, and d is the point of closest approach.

To maintain a constant $\Delta f a_z^{3/2}$, the point of closest approach must be kept constant. The recorded z extension to meet this condition will be equal to the change in a_z. Thus, the calibrated amplitude is not the amplitude of the complete motion of the tip apex, but rather a_z, the z component of this amplitude.

Effect on imaging and spectroscopy

Simulated AFM data were produced by creating a Lennard-Jones potential for a simple 2-D square lattice, with a lattice constant of 3 Å, and a minimum potential of −3 eV at a distance of 0.5 Å (Figure 4a). For simplicity the simulated AFM was run in constant height mode to collect Δf images. These were calculated by using Equation 32. Images were collected for both tips oscillating normal and at 45° to the sample (i.e., a qPlus sensor with a tip length of 1.389 mm). An oscillation amplitude of a_z = 0.5 Å was used for both motions (Figure 4b and Figure 4c), thus mimicking a calibration performed using the method described in the previous section. This leads to a total amplitude of a = 0.707 Å for the angled motion of the finite tip. Theses scans were aligned such that the position of closest approach of the tip apex is the same for both scans.

Qualitatively the images for both angled and normal tip motion look almost identical, and there is no noticeable reduction in resolution. Quantitatively the difference between the scans (Figure 4d) is ±2.5 Hz for images with a Δf range of approximately 34 Hz, giving rise to a relative error of less than 8%.

As the model potential has only one decay power rather than a combination of long-range and short-range forces, the relative changes between angled and normal oscillations are largely

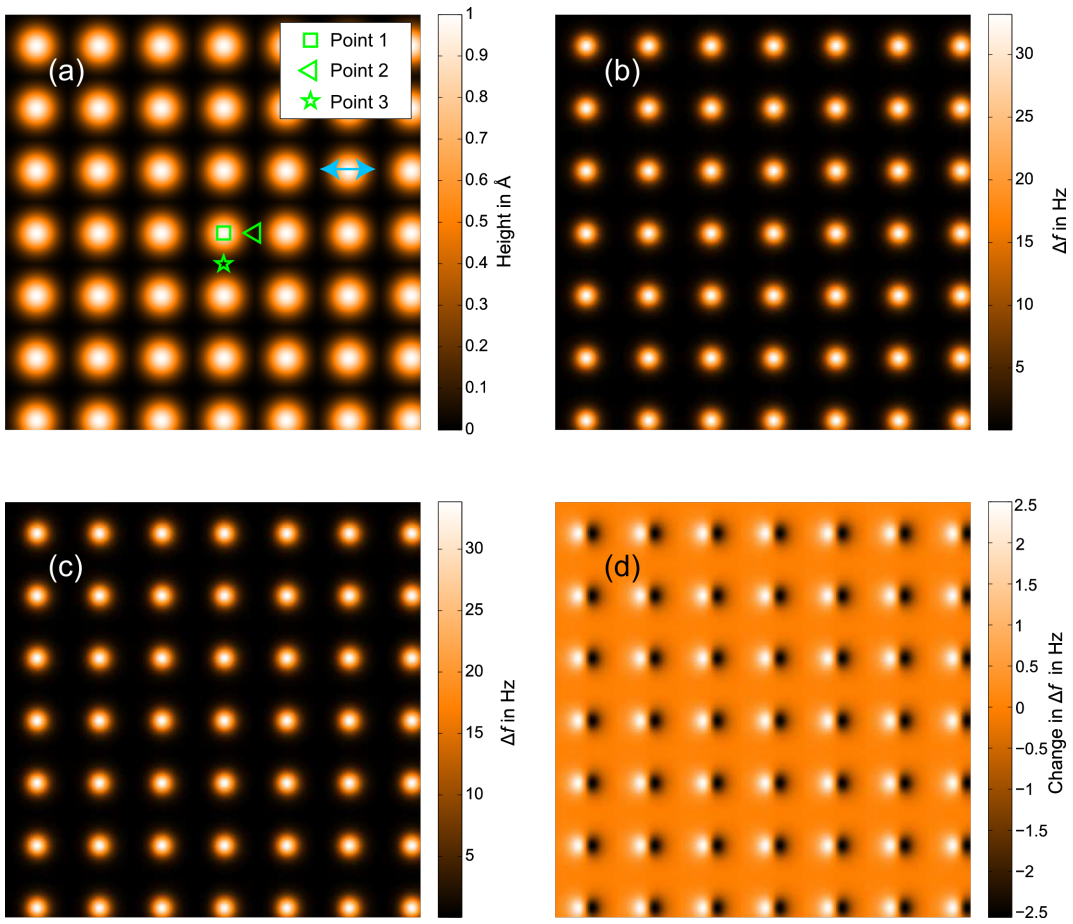

Figure 4: Effect of tip motion on imaging for an oscillation amplitude of a_z = 5 Å. All image widths are 2 nm × 2 nm. (a) shows the model surface. (b) and (c) show theoretical Δ*f* images for constant height scans with tip motion normal and at 45° to the sample surface, respectively. Lateral motion of the tip in (c) is aligned with the *x* axis of the scan, as indicated by the light blue arrow in (a). (d) is the difference between scans (b) and (c).

independent of the amplitude of oscillation. The change in sensitivity that arises from the different amplitudes will affect the absolute values of the frequency shift; however, as the simulation is not subject to limits in frequency resolution from experimental noise, these absolute values are of little interest.

Following this, simulated *z*-spectroscopy measurements were taken over three points (marked on Figure 4a) for both images. The results presented in Figure 5 show that for points 1 and 3, where the lateral force is near zero throughout the tip oscillation, the spectra align with a relative error of less than 3% at the point of highest interaction. However, for point 2 where the lateral force is significant as the tip moves diagonally over the adjacent atom, there is a difference of 23% between the curves at the point of highest interaction.

Conclusion

We have shown that using a simple Euler–Bernoulli model for the tine of a qPlus sensor, and inserting boundary conditions that account for both the moment of inertia and mass of the tip, we were able derive analytical results for a range of dynamic sensor properties. When the moment of inertia of the tip is zero our results agree with previous results found in the literature for point-mass weighted cantilevers [19]. After including the moment of inertia our results are in agreement with Tung et al. [11].

Further analysis of our model revealed large lateral motion at the tip apex when the tip length is comparable to the cantilever length. Our model was validated by direct experimental measurement of the lateral tip motion of a resonant qPlus sensor. To maintain a ratio of lateral to normal motion of under 10% the tip must protrude less that 57 µm from the tine for qPlus sensors of the modelled geometry. The methods currently used to calibrate amplitude in these systems accurately describe the *z*-component of amplitude; however, the overall amplitude of motion at the tip apex is larger. If the lateral force is near zero throughout the tip oscillation then this has minimal effect

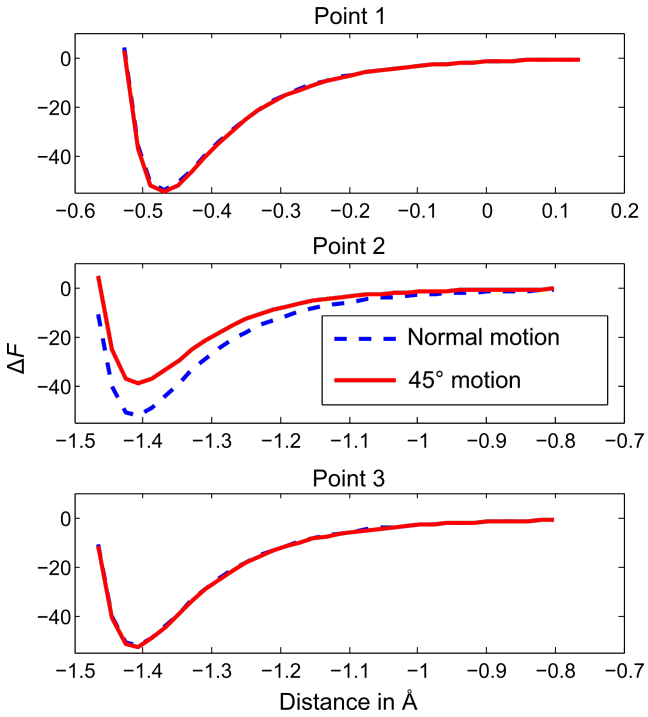

Figure 5: Effect tip motion on spectroscopy for an oscillation amplitude of a_z = 5 Å. Simulated spectroscopy measurements taken at three points as marked on Figure 4a. These show that only when the tip apex moves through significant lateral forces does the angle of scan have a noticeable effect on the *z*-spectroscopy.

on either imaging or spectroscopy measurements, thus explaining the ability of qPlus AFM to gain both subangstrom spatial resolution, and agreement with theoretical force measurements [7]. However, the lateral motion has a large effect on any data when lateral forces are present, requiring both careful analysis of experimental results and knowledge of the sensor and tip geometry.

Supporting Information

Supporting Information File 1

Further details of the presented theoretical model
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-4-2-S1.pdf]

Acknowledgements

The authors would like to thank P. Moriarty, R. A. J Woolley, and S. Jarvis for many fruitful discussions. This work was financially supported by a doctoral training grant from the EPSRC.

NIST Disclaimer: This article is authored in part by an employee of the U.S. federal government. Commercial equipment and materials are identified in order to adequately specify certain procedures. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

References

1. Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. *Science* **2009,** *325,* 1110–1114. doi:10.1126/science.1176210
2. Giessibl, F. J.; Hembacher, S.; Bielefeldt, H.; Mannhart, J. *Science* **2000,** *289,* 422–425. doi:10.1126/science.289.5478.422
3. Ternes, M.; Lutz, C. P.; Hirjibehedin, C. F.; Giessibl, F. J.; Heinrich, A. J. *Science* **2008,** *319,* 1066–1069. doi:10.1126/science.1150288
4. Sweetman, A.; Jarvis, S.; Danza, R.; Bamidele, J.; Gangopadhyay, S.; Shaw, G. A.; Kantorovich, L.; Moriarty, P. *Phys. Rev. Lett.* **2011,** *106,* 136101. doi:10.1103/PhysRevLett.106.136101
5. Sweetman, A.; Jarvis, S.; Danza, R.; Bamidele, J.; Kantorovich, L.; Moriarty, P. *Phys. Rev. B* **2011,** *84,* 085426. doi:10.1103/PhysRevB.84.085426
6. Chiutu, C.; Stannard, A.; Sweetman, A. M.; Moriarty, P. *Chem. Commun.* **2011,** *47,* 10575–10577. doi:10.1039/C1CC14147C
7. Chiutu, C.; Sweetman, A. M.; Lakin, A. J.; Stannard, A.; Jarvis, S.; Kantorovich, L.; Dunn, J. L.; Moriarty, P. *Phys. Rev. Lett.* **2012,** *108,* 268302. doi:10.1103/PhysRevLett.108.268302
8. Sweetman, A.; Jarvis, S.; Danza, R.; Moriarty, P. *Beilstein J. Nanotechnol.* **2012,** *3,* 25–32. doi:10.3762/bjnano.3.3

9. Giessibl, F. J. *Rev. Mod. Phys.* **2003,** *75,* 949–983. doi:10.1103/RevModPhys.75.949
10. Giessibl, F. J. *Appl. Phys. Lett.* **2000,** *76,* 1470–1472. doi:10.1063/1.126067
11. Tung, R. C.; Wutscher, T.; Martinez-Martin, D.; Reifenberger, R. G.; Giessibl, F.; Raman, A. *J. Appl. Phys.* **2010,** *107,* 104508. doi:10.1063/1.3407511
12. Sader, J. E.; Jarvis, S. P. *Appl. Phys. Lett.* **2004,** *84,* 1801. doi:10.1063/1.1667267
13. Shaw, G. A.; Pratt, J. R.; Jabbour, Z. J. Small mass measurements for tuning fork-based atomic force microscope cantilever spring constant calibration. In *MEMS and Nanotechnology, Volume 2;* Proulx, T., Ed.; Springer: New York, 2011; Vol. 2, pp 49–56.
14. Heyde, M.; Kulawik, M.; Rust, H.-P.; Freund, H.-J. *Rev. Sci. Instrum.* **2004,** *75,* 2446. doi:10.1063/1.1765753
15. Giessibl, F. J. *Phys. Rev. B* **1997,** *56,* 16010–16015. doi:10.1103/PhysRevB.56.16010
16. Rao, S. *Mechanical vibrations;* Addison-Wesley: Reading, MA, USA, 1995.
17. Butt, H.-J.; Jaschke, M. *Nanotechnology* **1995,** *6,* 1. doi:10.1088/0957-4484/6/1/001
18. Nowakowski, B. K.; Smith, S. T.; Pratt, J. R.; Shaw, G. A. *Rev. Sci. Instrum.* **2012,** *83,* 105115. doi:10.1063/1.4758125
19. Lozano, J. R.; Kiracofe, D.; Melcher, J.; Garcia, R.; Raman, A. *Nanotechnology* **2010,** *21,* 465502. doi:10.1088/0957-4484/21/46/465502

License and Terms

Beilstein Journal of Nanotechnology

Thermal noise limit for ultra-high vacuum noncontact atomic force microscopy

Jannis Lübbe[1], Matthias Temmen[1], Sebastian Rode[2,3], Philipp Rahe[2,4], Angelika Kühnle[2] and Michael Reichling[*1]

Full Research Paper

Open Access

Address:
[1]Fachbereich Physik, Universität Osnabrück, Barbarastraße 7, 49076 Osnabrück, Germany, [2]Institut für Physikalische Chemie, Johannes Gutenberg-Universität Mainz, Duesbergweg 10-14, 55099 Mainz, Germany, [3]now at: SmarAct GmbH, Schütte-Lanz-Strasse 9, 26135 Oldenburg, Germany and [4]now at: Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA

Email:
Michael Reichling[*] - reichling@uos.de

* Corresponding author

Keywords:
Cantilever; feedback loop; filter; noncontact atomic force microscopy (NC-AFM); noise

Beilstein J. Nanotechnol. **2013,** *4,* 32–44.
doi:10.3762/bjnano.4.4

Received: 04 October 2012
Accepted: 14 December 2012
Published: 17 January 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

The noise of the frequency-shift signal Δf in noncontact atomic force microscopy (NC-AFM) consists of cantilever thermal noise, tip–surface-interaction noise and instrumental noise from the detection and signal processing systems. We investigate how the displacement-noise spectral density d^{z} at the input of the frequency demodulator propagates to the frequency-shift-noise spectral density $d^{\Delta f}$ at the demodulator output in dependence of cantilever properties and settings of the signal processing electronics in the limit of a negligible tip–surface interaction and a measurement under ultrahigh-vacuum conditions. For a quantification of the noise figures, we calibrate the cantilever displacement signal and determine the transfer function of the signal-processing electronics. From the transfer function and the measured d^{z}, we predict $d^{\Delta f}$ for specific filter settings, a given level of detection-system noise spectral density d^{z}_{ds} and the cantilever-thermal-noise spectral density d^{z}_{th}. We find an excellent agreement between the calculated and measured values for $d^{\Delta f}$. Furthermore, we demonstrate that thermal noise in $d^{\Delta f}$, defining the ultimate limit in NC-AFM signal detection, can be kept low by a proper choice of the cantilever whereby its Q-factor should be given most attention. A system with a low-noise signal detection and a suitable cantilever, operated with appropriate filter and feedback-loop settings allows room temperature NC-AFM measurements at a low thermal-noise limit with a significant bandwidth.

Introduction

In this contribution, we discuss noise in frequency-modulation noncontact atomic force microscopy (NC-AFM) using cantilevers as force sensors and optical beam deflection (OBD) for signal detection. Figure 1 shows a schematic diagram of an NC-AFM setup based on OBD to illustrate the signal path and the quantities describing noise. Measured quantities discussed

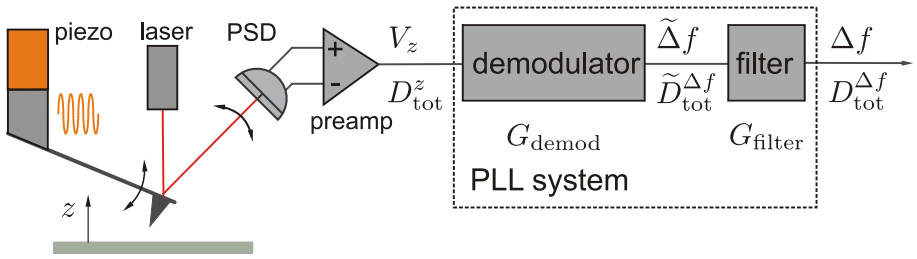

Figure 1: Schematic representation of the signal path in an NC-AFM system based on optical beam deflection with frequency demodulation using a PLL. The amplitude response of the PLL $G_{\mathrm{PLL}} = G_{\mathrm{filter}} \times G_{\mathrm{demod}}$ is split into demodulation and filtering parts, which are described by G_{demod} and G_{filter}. The quantities V_z and Δf describe the input and output signals of the PLL in NC-AFM operation while D_{tot}^{z} and $D_{\mathrm{tot}}^{\Delta f}$ are the corresponding noise power spectral densities superimposed to the signals.

here are often electrical signals that are equivalent to quantities describing the mechanical oscillation of the cantilever. The calibration procedure described in Section 1 of Supporting Information File 1 establishes a relation between the representation in mechanical and electrical units. During NC-AFM operation, the cantilever with eigenfrequency f_0 is excited to oscillation at the resonance frequency f_r, which differs from its eigenfrequency by the frequency shift $\Delta f = f_r - f_0$ when there is a tip–surface interaction. The mechanical oscillation, i.e., a periodic displacement $z(t)$ of the cantilever with amplitude A, is converted into the oscillation signal $V_z(t)$ by the position-sensitive detector (PSD) connected to the preamplifier. The amplitude A of this signal is determined and stabilised to a preset value by the amplitude feedback loop. Signal processing in NC-AFM involves the demodulation of the periodic cantilever-displacement signal $V_z(t)$ as well as filtering in the frequency domain to yield the frequency shift $\Delta f(t)$ carrying the information on the tip–surface interaction [1]. Demodulation is commonly performed by a phase-locked loop (PLL) circuit [2]. As schematically depicted in Figure 1, the amplitude response of the PLL unit can formally be decomposed into the amplitude response G_{demod} of the demodulator and the amplitude response G_{filter} of an in-loop or output filter. The characteristics of G_{filter} can be set by the user according to the needs of the experiment.

Noise in NC-AFM consists mainly of three contributions: noise arising from the thermal excitation of a cantilever or another force sensor, noise caused by the detection system and signal processing electronics [3,4], and instabilities arising from the interaction of the force microscopy tip with the surface as well as arising from the feedback loops stabilising the cantilever oscillation amplitude and the tip–surface distance [5]. Here, we investigate noise for the case of negligible tip–surface interaction and discuss the cantilever-displacement thermal-noise spectral density $d_{\mathrm{th}}^{z}(f)$ as well as the displacement-equivalent noise spectral density $d_{ds}^{z}(f)$ introduced by the detection system. This is carried out here in search of the ultimate limits of detection defined by thermal noise, while a systematic study of the tip–sample interaction noise that is present in any NC-AFM imaging or spectroscopy experiment will be the subject of forthcoming work. Here, we entirely focus the discussion on cantilever-based NC-AFM; however, the concepts, theoretical framework, and experimental strategies for the noise analysis can easily be transferred to systems based on other force sensors and detection schemes.

Under ultrahigh-vacuum (UHV) conditions, the thermal noise of the cantilever is usually small compared to the noise of the detection system due to the high Q-factor of the cantilever in vacuum [6]. The instrumental noise sources in an optical beam deflection (OBD) setup were recently discussed in detail [3] and it was found that the major noise sources are shot noise arising from the photodetector as well as Johnson noise originating from the resistors in the preamplifier. Further noise is generated in the laser diode that is mainly quantum noise for small output power and mode-hopping noise for large output power [3]. Back reflections of the laser beam into the laser optical resonator may increase mode hopping. The laser spot on the photodiode may further be disturbed by optical interference, creating time-varying speckle patterns due to temperature fluctuations and mechanical instability. It has been shown, however, that by operating the laser diode with radio-frequency modulation, the contribution of the light source to the total noise can be reduced to a negligible minimum [3].

The issue of noise is intimately related to the requirements of the NC-AFM system to process signals varying in time. The detection bandwidth B needed to retrieve the full information present in the $\Delta f(t)$ signal at the output of the PLL system depends on the spectral components produced during a scanning or spectroscopy experiment. Practically, the maximum usable bandwidth B_{max} is defined by the total displacement-noise spectral density $d_{\mathrm{tot}}^{z} = \sqrt{(d_{\mathrm{th}}^{z})^2 + (d_{\mathrm{ds}}^{z})^2}$ as schematically illustrated in Figure 2. In this figure, we show the displacement spectral density $d^z(f)$ present at the input of the frequency demodulator with contributions of the measurement signal and

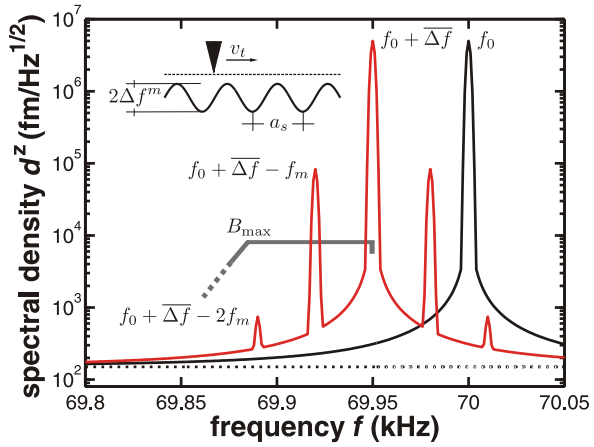

Figure 2: Illustrative representation for the spectral density of the displacement of a cantilever excited to oscillation with 10 nm amplitude at its eigenfrequency f_0 = 70 kHz without tip–sample interaction (black curve) and with tip–sample interaction resulting in a frequency modulation (red curve). Data is drawn for a modulation frequency of f_m = 30 Hz, a modulation amplitude of Δf^m = 1 Hz and a mean frequency shift of $\overline{\Delta f}$ = −50 Hz. A typical detection-system noise floor of d^z_{ds} = 150 fm/$\sqrt{\text{Hz}}$ (dotted line) as well as thermal noise based on the cantilever properties (k = 2.5 N/m, Q_0 = 100000) are added to the signal. The inset schematically illustrates how scanning the tip over the sample having a spatial periodicity a_s with a scan speed of v_t yields a modulation at frequency $f_m = v_t/a_s$. The surface corrugation Δz yields a modulation amplitude Δf^m where the modulation index is $\Delta f^m/f_m$ = 1/30 for this example.

noise (see Figure 1) as a function of the frequency f. This quantity is the root of the one-sided power spectral density $D^z(f)$, which is derived from the displacement signal $V_z(t)$ via a Fourier transform as

$$D^z(f) = \left(d^z(f)\right)^2 = 2\left|S\mathcal{F}\{V_z(t)\}(f)\right|^2$$

where S is the calibration factor converting voltage into displacement as defined in Section 1 of Supporting Information File 1 and $\mathcal{F}$ the Fourier transform of the displacement signal V_z with:

$$\mathcal{F}\{V_z(t)\}(f) = \int_{-\infty}^{\infty} V_z(t)\exp(-2\pi i f t)dt\,.$$

For the case of absent tip–surface interaction, d^z is a sharp peak centred at the cantilever eigenfrequency f_0 (f_0 = 70 kHz in Figure 2) including noise contributions from d^z_{th} and d^z_{ds}, which will be described in detail below. In the presence of a tip–surface interaction, the resonance peak is shifted by the amount $\overline{\Delta f}$ ($\overline{\Delta f}$ = −50 Hz in Figure 2) caused by the time-invariant part of the interaction. Additionally, sidebands appear that represent spectral components in $V_z(t)$ created during scanning or spectroscopy. For simplicity, we assume here a scanning of the tip over the surface with a speed v_t where a periodic corrugation (period a_s) of the surface Δf^m creates a sinusoidal modulation at the frequency $f_m = v_t/a_s$ (f_m = 30 Hz in Figure 2), i.e., $\Delta f(t) = \overline{\Delta f} + \Delta f^m \sin(2\pi f_m + \varphi)$. Effectively, this is a frequency modulation of $V_z(t)$ with a modulation index $\Delta f^m/f_m$ producing an infinite number of higher harmonics with rapidly decreasing power [4]. How many of these side peaks can be detected depends on the modulation index of the signal and the noise characteristics of the measurement system. For the hypothetical measurement illustrated in Figure 2, only two sideband peaks are well above the noise floor. Here, the suitable bandwidth B_{max} is defined by the frequency of the second sideband peak.

The frequency demodulator extracts the frequency shift $\Delta f(t)$ from the periodic displacement signal $V_z(t)$ and, for an arbitrary signal, projects the power in the sidebands of $D^z(f)$ into the frequency-shift power spectral density $D^{\Delta f}(f_m)$, which can be represented as:

$$D^{\Delta f}(f_m) = \left(d^{\Delta f}(f_m)\right)^2 = 2\left|\mathcal{F}\{\Delta f(t)\}(f_m)\right|^2$$

$$\mathcal{F}\{\Delta f(t)\}(f_m) = \int_{-\infty}^{\infty} \Delta f(t)\exp(-2\pi i f_m t)dt\,.$$

The frequency shift $\Delta f(t)$ varies on a time scale that in an imaging experiment is determined by the spatial periodicity of the scanned structure and the scanning speed, rather than by the period of the cantilever oscillation. Therefore, the spectrum of the frequency shift signal present at the output of the demodulator has significant power only in a limited spectral range of f_m. The detection bandwidth B of the demodulator is, therefore, usually restricted to a value of the order of 100 Hz to 1 kHz. As the noise is transformed by the demodulator in a similar way, we define $d^{\Delta f}_{tot}(f_m)$ and $D^{\Delta f}_{tot}(f_m)$ as the frequency-shift-noise spectral density and the frequency-shift-noise power spectral density, respectively, and discuss separate noise contributions $d^{\Delta f}_{th}$ and $d^{\Delta f}_{ds}$ to the frequency-shift signal Δf, as the noise contributions of the thermal cantilever excitation and the detection system yield different spectral characteristics. The detection bandwidth B and, consequently, the noise propagation characteristics depend on the PLL amplitude response $G_{PLL} = G_{filter} \times G_{demod}$, which can usually be influenced by the operator through the filter settings (see Figure 1).

To understand the influence of various experimental parameters and the settings of the PLL filter on $D^{\Delta f}_{tot}(f_m)$, which is the most relevant noise figure in the NC-AFM experiment, we derive noise models based on system parameters. Hypotheses and conclusions are tested against the reality of NC-AFM

experiments, by comparing the noise figures and filter settings for three NC-AFM systems based on the OBD scheme and comparing experimental results to the predicted settings for noise-optimised operation. We find that by the correct choice of the cantilever, by using optimised detection electronics and by appropriate PLL filter settings, the frequency-shift signal Δf can be detected at a low thermal-noise limit over a bandwidth B that is more than 100 Hz for room temperature operation under UHV conditions. The dependence of the thermal limit and other noise figures on relevant experimental parameters is discussed in detail.

Displacement noise

Here, we discuss the displacement noise superimposed on the displacement signal $V_z(t)$ in the case of negligible tip–surface interaction. Usually, the signal $V_z(t)$ is a noisy sinusoidally oscillating voltage and the noise can be described in the frequency domain by the displacement-noise spectral density $d^z_{\rm tot}(f)$. This is the square root of the displacement-noise power spectral density $D^z_{\rm tot}(f)$, which is proportional to the unwanted energy per frequency interval stored in the oscillating system.

A cantilever that is not deliberately excited but in equilibrium with a thermal bath at temperature T exhibits random fluctuations resulting in measurable noise in the cantilever displacement signal. This noise can be predicted by a model outlined in Section 2 of Supporting Information File 1. Furthermore, all electrical and optical components that are part of the detection system produce noise, superimposed on the displacement signal. Therefore, the power spectral density of the total displacement signal noise $D^z_{\rm tot}$ can be described as

$$D^z_{\rm tot} = D^z_{\rm th} + D^z_{\rm ds} \tag{1}$$

where $D^z_{\rm th}$ and $D^z_{\rm ds}$ represent the thermal and the detection-system contributions. The quantity $D^z_{\rm th}(f)$ as derived in Section 2 of Supporting Information File 1 can be represented as:

$$D^z_{\rm th} = \frac{2k_{\rm B}T/(\pi k_0 f_0 Q_0)}{\left(1-(f/f_0)^2\right)^2 + \left(f/(f_0Q_0)\right)^2}. \tag{2}$$

Here, $D^z_{\rm th}$ is calculated only for the fundamental cantilever oscillation mode with eigenfrequency f_0, stiffness k_0 and Q-factor Q_0 as the contribution of higher harmonics to the total noise power spectral density is small; the fundamental mode contains 97% or more of the total power extracted by the cantilever from the thermal bath. For the investigation of noise at higher harmonics, f_0, k_0 and Q_0 would have to be replaced by the respective modal values f_n, k_n and Q_n (see Section 2 of Supporting Information File 1). The noise spectral density $d^z_{\rm tot}(f)$ as defined in Equation 1 is displayed in Figure 3a for typical experimental parameters exhibiting a sharp peak at f_0 = 70 kHz (note the logarithmic scale of the ordinate) sticking far out of the white-noise floor $d^z_{\rm ds}(f)$ when using low-noise detection electronics. The sharp peak in $d^z_{\rm th}(f)$ results from the cantilever resonance. Especially high Q-factor cantilevers strongly amplify the white spectral power of thermal excitation only in a narrow range of frequencies around f_0 according to Equation 2. The detection-system noise represented by $d^z_{\rm ds}(f)$ is governed by the quality of the optical and electronic compo-

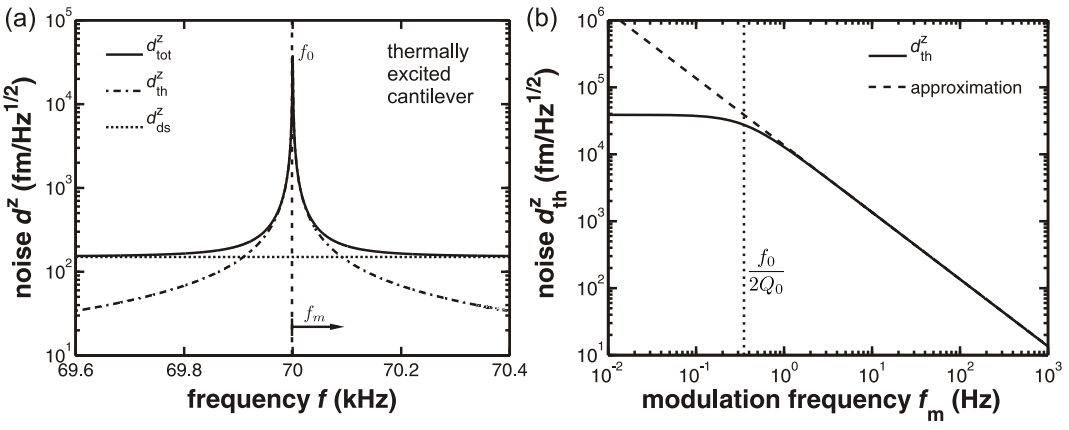

Figure 3: Illustrative representation of noise properties for a cantilever with f_0 = 70 kHz, k = 2.5 N/m and Q_0 = 100000 that is solely excited by its contact to a thermal bath at room temperature. (a) Calculated total-displacement noise spectral density $d^z_{\rm tot}$ (solid line) compared to the thermal-noise contribution $d^z_{\rm th}$ (dash-dotted line) and the detection-system noise $d^z_{\rm ds}$ = 150 fm/$\sqrt{\rm Hz}$ (dotted line). (b) Comparison between the thermal-displacement noise spectral density $d^z_{\rm th}(f_0 \pm f_m)$ as given in Equation 2 (solid line) and the approximation of Equation 3 (dashed line) for the cantilever with a corner frequency of $f_0/(2Q_0)$ = 0.35 Hz. Considering the oscillating cantilever as a mechanical low-pass filter for the displacement noise close to f_0, the corner frequency defines the point at which the noise is attenuated by 3 dB. At modulation frequencies larger than the corner frequency, $d^z_{\rm th}$ decreases essentially as $1/f_m$.

nents used in the detection system. In contrast to thermal noise, which is a fixed quantity for a given cantilever and temperature, the detection-system noise floor can be reduced by technical improvements of the detection system [3,7,8].

Frequency-shift noise

The frequency demodulator of the NC-AFM system extracts the cantilever response to the tip–surface interaction from the sidebands of the cantilever-oscillation frequency spectrum (see Figure 2) and yields the signal power spectral density present in the sidebands, i.e., the displacement power spectral density $D^z(f)$ is transformed to the frequency-shift power spectral density $D^{\Delta f}(f_m)$ by the demodulation process. The noise contribution $D_{\text{tot}}^{\Delta f}\left(f_m\right)$ in this spectrum is the most relevant noise figure in NC-AFM measurements and can be calculated from the demodulator input noise by applying the appropriate demodulator transfer function and an approximation to obtain a simple yet accurate expression for the thermal-displacement power spectral density D_{th}^z. As the frequency noise is represented as a function of the modulation frequency f_m, it is desirable to represent the displacement noise as a function of $f_0 \pm f_m$. For D_{th}^z, we use the following approximation [1] instead of the precise result from Equation 2:

$$D_{\text{th}}^z\left(f_0 \pm f_m\right) \simeq \frac{k_{\text{B}}Tf_0}{2\pi k_0 Q_0}\frac{1}{f_m^2}. \tag{3}$$

This expression is a very good approximation for modulation frequencies f_m exceeding the cantilever corner frequency $f_0/(2Q_0)$ as seen in Figure 3b. This approximation covers most of the practically relevant spectral range as the corner frequency is smaller than 1 Hz for high-Q cantilevers. Combining Equation 1 and Equation 3 yields a simple yet accurate expression for the power spectral density of the total displacement noise in an FM-AFM system operated under high-Q conditions [4]:

$$D_{\text{tot}}^z\left(f_0 \pm f_m\right) \simeq \frac{k_{\text{B}}Tf_0}{2\pi k_0 Q_0}\frac{1}{f_m^2} + D_{\text{ds}}^z. \tag{4}$$

To obtain the noise power spectral density of the frequency-shift signal present at the demodulator output, the demodulator amplitude response for noise $G_{\text{demod}} = \sqrt{2}f_m/A$ is applied, and we find [4]

$$\widetilde{D}_{\text{tot}}^{\Delta f} = G_{\text{demod}}^2 D_{\text{tot}}^z = \frac{f_0 k_{\text{B}}T}{\pi k_0 Q_0 A^2} + \frac{2f_m^2}{A^2}D_{\text{ds}}^z. \tag{5}$$

As apparent from Equation 5, the contribution of the thermal noise to the total noise is independent of the modulation frequency f_m, whereas the detection-system-noise power contribution is amplified by the square of the modulation frequency. We further note that the total noise power in Δf depends on the reciprocal of the squared cantilever oscillation amplitude.

The frequency-shift noise spectral density $\widetilde{d}_{\text{tot}}^{\Delta f}$ and its components as described in Equation 5 are shown as a function of the modulation frequency f_m in Figure 4 for typical experimental conditions neglecting bandwidth limitations. This result clearly points to the experimental parameters determining the frequency-shift noise: the thermal limit is defined by the temperature T and cantilever properties, namely the ratio $f_0/(k_0Q_0)$. For a cantilever with given f_0 and k_0, it is most important to yield a high effective Q-factor that may considerably differ from the intrinsic Q-factor [9] if one is interested in reducing the thermal-noise limit to the lowest possible value. The noise contribution from the detection system depends on the required bandwidth B (range of f_m) and the quality of the detection system represented by d_{ds}^z. Overall, $\widetilde{d}_{\text{tot}}^{\Delta f}$ scales with the inverse of the cantilever oscillation amplitude A. In Figure 4, the thermal noise limit is shown for typical cantilever properties and for T = 300 K as dash-dotted lines representing different Q-factors. From Figure 4 we can deduce the displacement noise floor of the detection system d_{ds}^z that must not be exceeded for a thermal-noise-limited measurement. We define the bandwidth $B_{\text{th}}^{\text{lim}}$ for a thermal-noise-limited measurement by the frequency where the contributions of d_{th}^z and d_{ds}^z to the total frequency-shift noise spectral density $\widetilde{d}_{\text{tot}}^{\Delta f}$ are equal. This frequency $f_{\text{m}}^{\text{lim}}$ corresponds to the crossing point between the dashed and dash-dotted lines in Figure 4.

It follows that operation at the thermal noise limit can only be obtained if the bandwidth B of the demodulator is set close to

$$B_{\text{th}}^{\text{lim}} = \sqrt{\frac{f_0 k_{\text{B}}T}{2\pi k_0 Q_0}} \times \frac{1}{d_{\text{ds}}^z} \tag{6}$$

where the noise spectral density d_{ds}^z is treated here as a constant. This is fully justified by its white-noise character around the cantilever resonance. Note that this bandwidth limitation is solely based on noise considerations and does not reflect other bandwidth requirements, such as the stable operation of the PLL. However, there is a bandwidth limitation in any real system and $\widetilde{D}^{\Delta f}$ has to be considered as a hypothetical quantity that is rarely accessible. In any PLL system of practical use, the detection bandwidth is defined by internal filters, loop-gain settings and time constants that are normally accessible to the user for an optimisation of the signal processing.

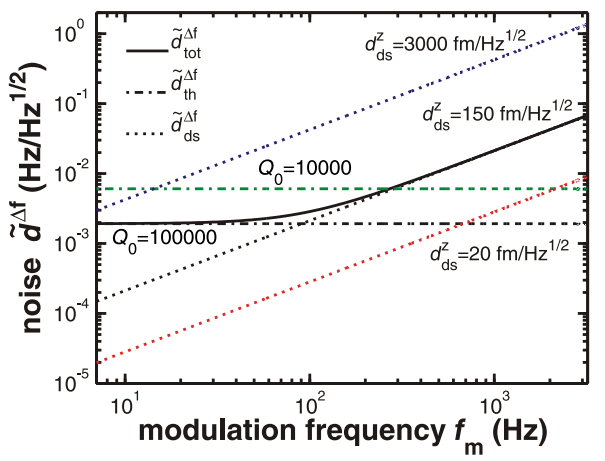

Figure 4: Illustrative representation of the noise spectral density for the total frequency-shift noise $\tilde{a}^{\Delta f}_{tot} = \sqrt{\tilde{D}^{\Delta f}_{tot}}$ for a system without bandwidth limitations. The total noise is composed of contributions from the thermal noise $\tilde{a}^{\Delta f}_{th} = \sqrt{\tilde{D}^{\Delta f}_{th}}$ plotted for different *Q*-factors and the noise of the frequency-shift detection system $\tilde{a}^{\Delta f}_{ds}$ plotted for different values of the noise floor $d^z_{ds} = \sqrt{D^z_{ds}}$. Cantilever and oscillation parameters are f_0 = 70 kHz, k_0 = 2.5 N/m and A = 10 nm.

Thus, a complete PLL is modelled by using $G_{PLL} = G_{filter} \times G_{demod}$, with G_{filter} being the amplitude response for the aforementioned filters (see Figure 1). Taking the amplitude response of the full PLL system into account, we obtain for the accessible noise power spectral density

$$\begin{aligned} D^{\Delta f}_{tot} &= D^{\Delta f}_{th} + D^{\Delta f}_{ds} \\ &= G^2_{filter}\tilde{D}^{\Delta f}_{tot} \\ &= G^2_{filter}\left(\frac{f_0 k_B T}{\pi k_0 Q_0 A^2} + \frac{2 f_m^2}{A^2} D^z_{ds}\right) \end{aligned} \tag{7}$$

at the output of the bandwidth-limited PLL system. The experimental determination of an unknown amplitude response G_{filter} is described in Section 3 of Supporting Information File 1. To characterise the demodulator output noise with a single number, we define δf_{tot} as the root mean square (RMS) of the overall frequency-shift noise:

$$\delta f_{tot} = \left(\int_0^\infty D^{\Delta f}_{tot}(f_m)\, df_m\right)^{1/2} \tag{8}$$

where the integration can practically be limited to an upper frequency limit related to the detection bandwidth B. This is fully justified as filtering in the demodulator always yields a low-pass characteristic. A discussion of the RMS noise figure and its calculation by using approximations for the demodulator bandwidth is presented in Section 4 of Supporting Information File 1.

Experimental

Noise measurements are performed with three NC-AFM systems, named systems A, B and C in the following. All systems are well decoupled from mechanical vibrations by spring suspension and eddy-current damping systems. As an additional precaution, connections between the electronics and piezos are removed during noise measurements to ensure that measurements are not affected by any spurious electrical signals exciting the cantilever. All systems investigated here are based on the optical beam-deflection scheme for measuring the cantilever displacement. Therefore, the laser-light power P_{pd} reaching the photodetector is a parameter characterising the system. P_{pd} is calculated from the sum signal of the PSD, which in turn depends on the spectral sensitivity of the photodiode, the used laser light wavelength and the DC transimpedance of the preamplifier. The calibration of the detection system is described in Section 1 of Supporting Information File 1.

System A is a room-temperature UHV AFM/STM (Omicron NanoTechnology GmbH, Taunusstein, Germany) equipped with an easyPLL (Nanosurf AG, Liestal, Switzerland) for frequency demodulation. The AFM/STM setup has been modified by replacing the light source (light-emitting diode exchanged with a laser diode) and using optimised preamplifiers. Preamplifiers have been optimised for low-noise operation at frequencies around 100 kHz and 300 kHz, respectively, and are exchanged depending on the eigenfrequency of the cantilever. Details on this modification and the frequency response of the preamplifiers can be found in [7]. The light source is a 48TE-SOT (Schäfter+Kirchhoff GmbH, Hamburg, Germany) and emits light at a wavelength of 685 nm, while the PSD has a spectral sensitivity of 0.45 A/W at this wavelength. Noise spectra are recorded with an SR770 spectrum analyser (Stanford Research Systems, Inc., Sunnyvale, CA, USA).

System B is a UHV VT AFM/STM (Omicron NanoTechnology GmbH, Taunusstein, Germany) equipped with an easyPLL plus (Nanosurf AG, Liestal, Switzerland) as the demodulator. This system uses a light source having a wavelength of 830 nm, while the spectral sensitivity of the PSD is 0.57 A/W at this wavelength. Noise spectra are measured using the zoom FFT module of a HF2LI lock-in detector (Zurich Instruments AG, Zurich, Switzerland) for spectral analysis.

System C is a UHV 750 variable temperature STM/AFM with a PLLPro2 (software version 0.20.0) as the demodulator (RHK Technology, Inc., Troy, MI, USA). The light source is a laser source type 51nanoFCM (Schäfter+Kirchhoff GmbH,

Hamburg, Germany) operated in the constant-power mode with radio-frequency modulation to reduce the coherence length to about 300 µm. The laser-light wavelength is 639 nm, and a maximum output power of 5 mW is available at the fibre end while the PSD has a spectral sensitivity of 0.4 A/W at this wavelength. A home-built preamplifier (low-bandwidth preamplifier) or the preamplifier supplied by the manufacturer (high-bandwidth preamplifier) is used depending on the bandwidth requirements. The frequency response of both preamplifiers is shown in Figure 5. To measure noise spectra, the SR770 spectrum analyzer (Stanford Research Systems, Inc., Sunnyvale, CA, USA) is used.

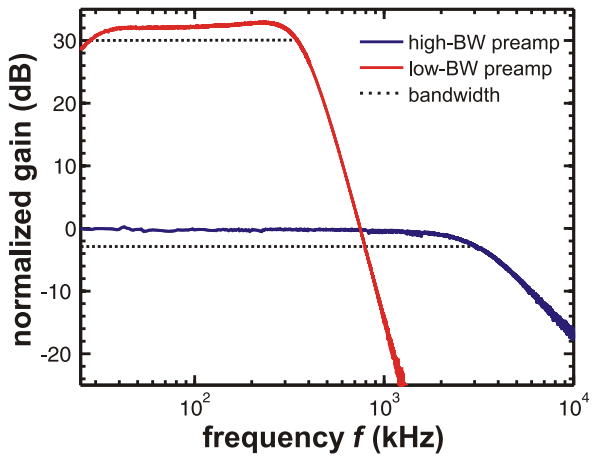

Figure 5: Frequency response of the high-bandwidth preamplifier (bandwidth 3.1 MHz) and the low-bandwidth preamplifier (bandwidth 320 kHz) for System C. The gain is normalised to the DC gain of the high-bandwidth preamplifier. Instead of connecting to the PSD, a sine wave of 0.5 V RMS amplitude was fed into a single quadrant input with a 100 kΩ resistor resulting in 5 µA RMS current.

Force sensors are commercial silicon cantilevers (Nanoworld AG, Neuchâtel, Switzerland). For our investigations, we use a set of cantilevers covering a large range of resonance frequencies, Q-factors and stiffness, to explore the impact of these parameters on the noise figures. Cantilever specifications are compiled in Table 1 and Table 2 (cantilevers D and AO are type FM, cantilevers AQ, AR and V are type NCH, cantilever AP is type NCVH and cantilever AL is type Arrow™ according to the commercial classification scheme).

The cantilever eigenfrequencies of the fundamental and the first and second harmonic mode are determined by measuring resonance curves and fitting the amplitude response function to the data as described in [9]. This procedure also yields quality factors Q_n, while the stiffness k_{dim} is calculated from cantilever dimensions and material properties [10] and used as a good approximation to the modal stiffness k_0 [12].

Table 1: Fundamental properties of the cantilevers used for noise analysis. Length l, mean width $\bar{w}$ and thickness t are provided by the manufacturer. The stiffness k_{dim} is calculated from the cantilever dimensions [10]. Typical properties of a qPlus sensor are taken from [11] for comparison.

cantilever	l (µm)	$\bar{w}$ (µm)	t (µm)	k_{dim} (N/m)
AO 3	224	30	3.0	3.0 ± 0.9
D 5	229	30	2.9	2.5 ± 0.8
AR 17	127	27	3.6	26 ± 5
V 15	125	26	3.7	29 ± 6
AQ 10	123	29	4.5	60 ± 10
AP 5	40	24	2.0	130 ± 50
AL 3	35	42	0.7	9 ± 3[a]
qPlus	2400	126	214	1800

[a]Value provided by the manufacturer.

Results and Discussion

The noise analysis is performed in two steps. First, we measure the displacement noise spectral density d^z_{tot} and, second, we investigate how it is propagated to the frequency-shift noise spectral density $d^{\Delta f}_{\mathrm{tot}}$. The displacement noise is measured by a spectrum analyser connected directly to the output of the preamplifier (see Figure 1). The measurement range of the spectrum analyser is set to a few kilohertz around the cantilever resonance frequency to obtain high spectral resolution. The spectral density of the noise in the signal V_z is measured and converted to the displacement-noise spectral density d^z_{tot} in units of fm/$\sqrt{\mathrm{Hz}}$ by the calibration procedure outlined in Section 1 of Supporting Information File 1. Figure 6 shows a representative result obtained with system C. The measured displacement noise spectral density is shown in Figure 6a (solid lines), together with the thermal noise contribution d^z_{th} (dash-dotted line) calculated from the given cantilever properties by using Equation 2. The noise floor of the detection system d^z_{ds} (dotted lines) is measured beside the resonance peak where thermal noise becomes negligible (solid and dashed lines are identical). For a study on how the noise of the detection system d^z_{ds} propagates through the demodulation system, different noise levels are artificially created by using white noise from a waveform generator DS345 (Stanford Research Systems, Inc., Sunnyvale, CA, USA) added to the displacement signal V_z. The curve with d^z_{ds} = 108 fm/$\sqrt{\mathrm{Hz}}$ represents the noise floor of the setup while the other curves show artificially increased noise levels.

To measure the frequency-shift noise $d^{\Delta f}_{\mathrm{tot}}$, the cantilever is excited to an oscillation with typically 10 nm amplitude and the spectrum analyser is connected to the output of the demodulator (see Figure 1) to measure the voltage noise in the Δf signal. The demodulator is adjusted to zero mean frequency

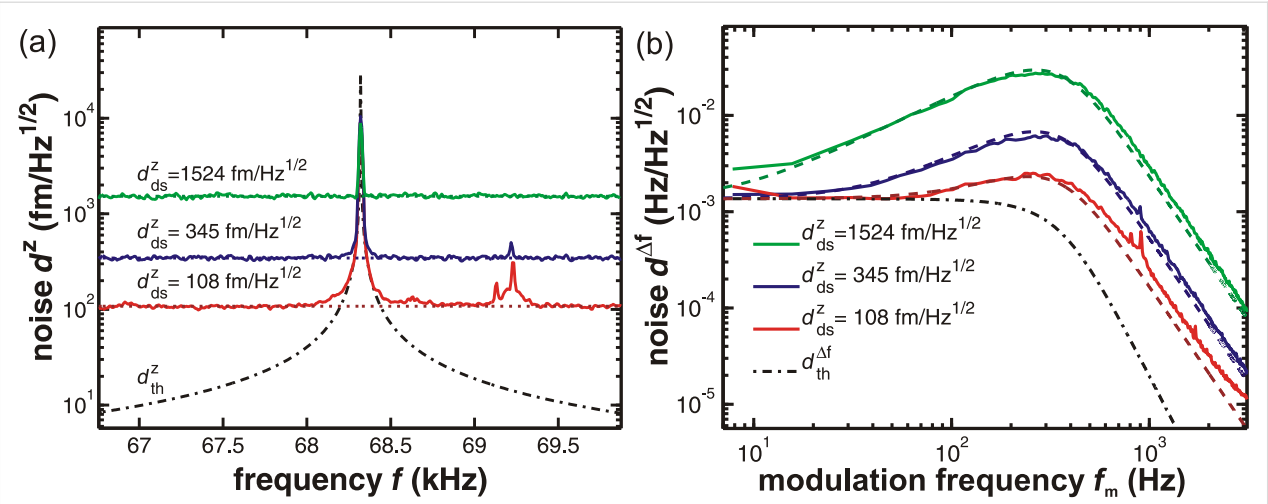

Figure 6: Measured and modelled noise figures for system C. (a) Different levels of displacement noise spectral density d^z_{tot} at the output of the low-bandwidth preamplifier for a thermally excited cantilever. Solid lines represent measured data while dotted lines indicate the corresponding detection-system noise floor d^z_{ds}. The dash-dotted line is a calculation of the displacement-thermal-noise spectral density d^z_{th} of the cantilever. (b) Calculated noise spectral density $d^{\Delta f}_{tot}$ at the PLL output for a cantilever oscillation amplitude of 5 nm and different noise floor levels (dashed lines) compared to measured data (solid lines). The dash-dotted line represents the modelled thermal noise contribution $d^{\Delta f}_{th}$ to the noise in the Δf signal. Measurements are performed with cantilever D 5 (see Table 1 and Table 2 for cantilever properties). Filter settings are f_c = 500 Hz, o = 3, P = −2.0 Hz/deg and l = 1 Hz (see Section 3 of Supporting Information File 1 for a detailed explanation).

shift and the measurement range of the spectrum analyser is set to the frequency region between 0 and 3 kHz. The measured voltage noise is multiplied by the known conversion factor of the demodulator (e.g., 30 Hz/V) to obtain the frequency-shift-noise spectral density $d^{\Delta f}_{tot}$ in units Hz/$\sqrt{\text{Hz}}$. In Figure 6b, this quantity is shown for the same three levels of artificial detection-system displacement noise d^z_{ds} at the input of the demodulator as supplied for the measurement in Figure 6a. Measurements (solid lines) are compared to calculated curves (dashed lines) based on the d^z_{tot} values obtained from the measurements shown in Figure 6a. The curve $d^{\Delta f}_{th}$ (f_m) is determined from the cantilever properties and the filter settings of the PLL demodulator (thermal contribution in Equation 7) and represents the ideal case of the thermal noise of the cantilever without any detection-system noise. The trailing edge on the right side is caused by the attenuation through the low-pass filter with amplitude response G_{filter} (see Section 3 of Supporting Information File 1 for details).

The dashed lines are model curves calculated using Equation 7 with the measured noise contribution d^z_{ds} and the calculated thermal noise contribution d^z_{th}. The measured noise curves (solid lines) are in good agreement with the model (dashed lines). A disturbing side peak, which can be observed on top of d^z_{tot} in Figure 6a for a low detection-system noise floor, similarly appears in the corresponding curve $d^{\Delta f}_{tot}$ in Figure 6b. Such peaks are due to electromagnetic emission from switching power supplies and other devices present in the laboratory environment. As the propagation of displacement noise d^z_{tot} to frequency-shift noise $d^{\Delta f}_{tot}$ is well reproduced by experimental data, the latter can be used to obtain the level of the noise floor d^z_{ds} from a noise measurement in the low-frequency region of the Δf signal. By inversion of Equation 7, the displacement-noise spectral density $d^z_{ds} = \sqrt{D^z_{ds}}$ can be obtained from $D^{\Delta f}_{tot}$ if the system frequency response and $\tilde{D}^{\Delta f}_{th}$ are known:

$$D^z_{ds} = \frac{1}{G^2_{demod}}\left(\frac{D^{\Delta f}_{tot}}{G^2_{filter}} - \tilde{D}^{\Delta f}_{th}\right). \tag{9}$$

At the frequency $f_m = f_m^{max}$, where $f_m \times G_{filter}(f_m)$ has its maximum, the total noise is dominated by the noise from the detection system. At this point, Equation 9 can be simplified when assuming $D^{\Delta f}_{th} << D^{\Delta f}_{tot}$ yielding

$$\begin{aligned} D^z_{ds} &\approx \left.\frac{D^{\Delta f}_{tot}}{G^2_{demod}G^2_{filter}}\right|_{f_m=f_m^{max}} \\ &\approx \left.\frac{A^2 D^{\Delta f}_{tot}}{2 f_m^2 G^2_{filter}}\right|_{f_m=f_m^{max}}. \end{aligned} \tag{10}$$

This approximation defines an upper limit $d^z_{ds,max} = \sqrt{D^z_{ds,max}}$ for the detection-system noise spectral density.

In this manner, we investigate the noise characteristics of the three NC-AFM systems using different preamplifiers and

various cantilevers at different eigenmodes; the corresponding results are listed in Table 2. The detection noise measured directly in the displacement signal V_z as shown in Figure 6a is denoted as d^z_{ds} while the same quantity obtained from the frequency-shift noise $d^{\Delta f}_{\mathrm{tot}}$ by using Equation 9 is denoted as $d^z_{\mathrm{ds},\Delta f}$. The upper limit derived from Equation 10 is denoted as $d^z_{\mathrm{ds,max}}$. The latter is a useful approximation that can easily be calculated without knowledge of the cantilever properties.

Table 2 allows a comparison of the noise floor for different NC-AFM systems and demonstrates the influence of cantilever properties on the noise figures. The best values for the noise floor achieved here are around 100 fm/$\sqrt{\mathrm{Hz}}$ as measured for cantilever V 15 in system B and cantilevers D 5 and AR 17 in system C. These cantilevers have a length in the range of 100 µm to 250 µm. Exchanging the preamplifier may cause a large difference in the noise floor. This can be observed for cantilevers AO 3 and AR 17 in system C where the noise floor is doubled by changing from the low-bandwidth to the high-bandwidth preamplifier. The benefit of the high-bandwidth amplifier is the possibility to operate cantilevers at their higher resonance frequencies, where the displacement-noise floor significantly decreases even for a similar voltage noise caused by the laser-power-dependent photodiode shot noise due to different amplitude calibration factors for the corresponding modes. However, due to the length of only 40 µm and 35 µm, for the high-frequency cantilevers AP 5 and AL 3, the laser adjustment becomes difficult, yielding only 77 µW and 18 µW laser power on the PSD compared to about 100 µW for other cantilevers. Therefore, the detection-noise floor for these canti-

Table 2: Cantilever properties and noise figures for systems A, B and C. f_n and Q_n are the eigenfrequencies and Q-factors for the nth eigenmode of the cantilever. Noise-floor values d^z_{ds} are directly determined from the displacement signal V_z, while $d^z_{\mathrm{ds},\Delta f}$ and $d^z_{\mathrm{ds,max}}$ are extracted from the Δf noise at the demodulator output as described in the main text. P_{pd} is the total light power on the PSD. For system C, measurements are performed with two different preamplifiers. Missing d^z_{ds} values are due to frequency-range limitations of the spectrum analyser. In the case of higher harmonics, we cannot easily calculate the modal cantilever stiffness, as it strongly depends on the tip mass, which is generally not known [13]. Therefore, determining $d^z_{\mathrm{ds},\Delta f}$ requires the knowledge of the stiffness and is, thus, only calculated for the measurements at the fundamental resonance frequency. Typical properties of a system operated with a qPlus sensor are taken from [11] for comparison.

cantilever	f_n	Q_n	d^z_{ds} (fm/$\sqrt{\mathrm{Hz}}$)	$d^z_{\mathrm{ds},\Delta f}$ (fm/$\sqrt{\mathrm{Hz}}$)	$d^z_{\mathrm{ds,max}}$ (fm/$\sqrt{\mathrm{Hz}}$)	P_{pd} (µW)
System A						
AQ 10	f_0 = 361,599 Hz	Q_0 = 21,200		275	278	97
System B						
V 15	f_0 = 279,451 Hz	Q_0 = 47,200	125	119	124	105
System C, low-bandwidth preamplifier						
D 5	f_0 = 68,353 Hz	Q_0 = 118,000	115	122	130	120
AO 3	f_0 = 68,183 Hz	Q_0 = 173,700	237	223	226	106
AR 17	f_0 = 276,360 Hz	Q_0 = 39,200		97	98	120
System C, high-bandwidth preamplifier						
AO 3	f_0 = 68,183 Hz	Q_0 = 173,700		416	417	105
AO 3	f_1 = 437,086 Hz	Q_1 = 48,500			93	105
AO 3	f_2 = 1,235,138 Hz	Q_2 = 15,200			51	105
AR 17	f_0 = 276,360 Hz	Q_0 = 39,200		258	259	120
AR 17	f_1 = 1,730,811 Hz	Q_1 = 6,300			99	120
AP 5	f_0 = 1,996,199 Hz	Q_0 = 32,400		302	309	77
AL 3	f_0 = 1,316,757 Hz	Q_0 = 16,600		845	892	18
qPlus system						
qPlus	f_0 = 32,768 Hz	Q_0 = 5,000	62			

levers is much higher than for larger cantilevers. With an improved laser-spot adjustment, however, a noise floor close to 100 fm/$\sqrt{\text{Hz}}$ should be possible.

In Figure 7, we illustrate the choice of optimum filter settings for a thermal-noise-limited detection. For that purpose, the frequency noise originating from the same detection system but passed through different PLL filters is shown. As these measurements are performed in system C, the filters are modelled as a closed loop where the settings of the PI controller have a significant effect on the frequency response and need to be individually adjusted for each setting of the loop filter order o and cutoff frequency f_c (see Section 3 of Supporting Information File 1 for details). The optimum settings for each loop filter used in the following are listed in Table S4 in Section 3 of Supporting Information File 1. In Figure 7a, we display the noise spectral characteristics of the Δf signal, while Figure 7b shows a plot of the total noise represented by the RMS value of the Δf signal as a function of the cantilever oscillation amplitude. Using a bandwidth of $B_{-3\text{dB}}$ = 385 Hz (f_c = 1 kHz, o = 5), the total noise exceeds the thermal noise level by half an order of magnitude. Choosing a much lower bandwidth of $B_{-3\text{dB}}$ = 48 Hz (f_c = 125 Hz, o = 5) decreases the frequency range where the signal is not attenuated below the one defined by the thermal noise limit. The optimum filter setting for the d_{ds}^{z} noise floor present in this measurement is a filter setting with a bandwidth of $B_{-3\text{dB}}$ = 103 Hz (f_c = 125 Hz, o = 1), where the total noise does not significantly exceed the thermal noise and the signal is not unnecessarily attenuated. For all filter settings investigated here, experiment (solid lines) and model (dashed lines) agree well with each other.

In Figure 7b, measured values δf_{tot} (circles, squares, triangles) are compared to calculated values δf_{tot} derived from Equation 8 (solid lines) and δf_{th} defining the thermal limit of the RMS frequency-shift noise (dash-dotted lines). While there is a large difference between thermal noise and total noise for the large-bandwidth filter setting (f_c = 1 kHz, o = 5), this discrepancy becomes smaller and finally negligible on further reduction of the bandwidth. Note, however, that the settings yielding the smallest RMS noise are not the optimum as the corresponding filter does not only reduce the noise but attenuates the NC-AFM signal more than necessary for thermal-noise-limited operation.

The RMS value of the total noise is an important figure of merit of the NC-AFM detection system, as it defines the minimum detectable frequency shift. Figure 7b is an excellent demonstration of the potential of small amplitudes for atomic resolution measurements as it is known that the atomic contrast increases with reduced cantilever oscillation amplitude [11,14]. For a measurement with $B_{-3\text{dB}}$ = 385 Hz (red line, f_c = 1 kHz, o = 5), one would choose an amplitude of 5 nm or above to reduce the noise; however, this would also reduce the atomic contrast compared to a lower amplitude measurement. For a measurement with $B_{-3\text{dB}}$ = 103 Hz (blue line, f_c = 125 Hz, o = 1), one can take full advantage of the increased atomic corrugation for the smaller amplitude, as the total noise is even below the thermal noise for the larger amplitude.

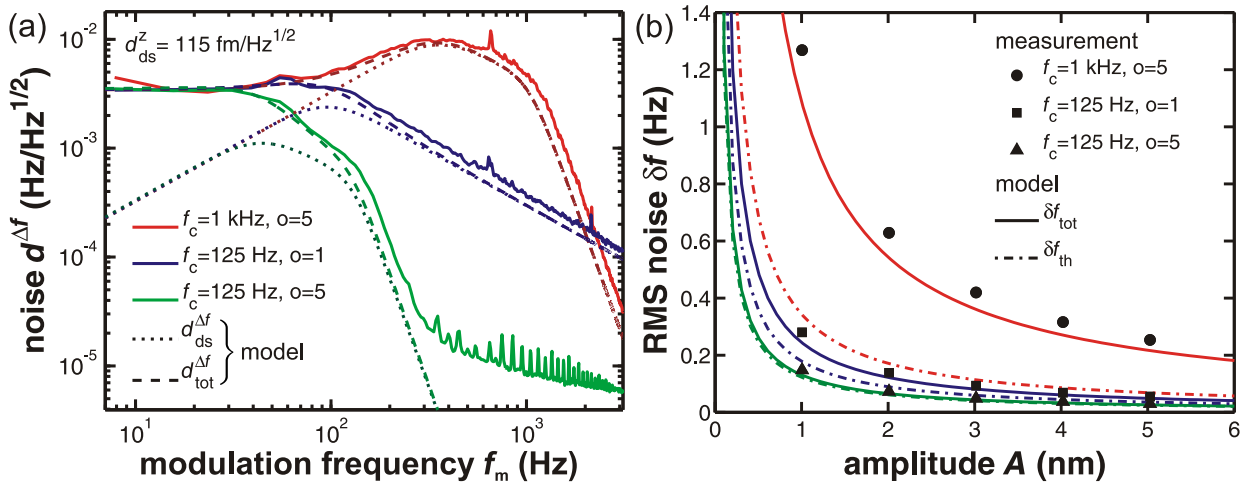

Figure 7: (a) Measured (solid) and modelled (dashed) frequency-shift-noise spectral density $d_{\text{tot}}^{\Delta f}$ using three different filter settings. Dotted lines show the contribution $d_{\text{ds}}^{\Delta f}$ of the detection-system noise to the total frequency-shift-noise spectral density for a noise floor of d_{ds}^{z} = 115 fm/$\sqrt{\text{Hz}}$ at the input of the demodulator. The oscillation amplitude is 5 nm. (b) Noise figures for different cantilever oscillation amplitudes. Measured RMS frequency-shift noise δf (circles, squares, triangles) for different PLL filter settings compared to predictions from Equation 8 (solid lines) using the measured detection-system noise d_{ds}^{z} in the cantilever displacement signal V_z. Dash-dotted lines represent calculations of the thermal-noise contribution δf_{th}. Measurements are performed with cantilever D 5 in system C (see Table 1 and Table 2 for cantilever properties).

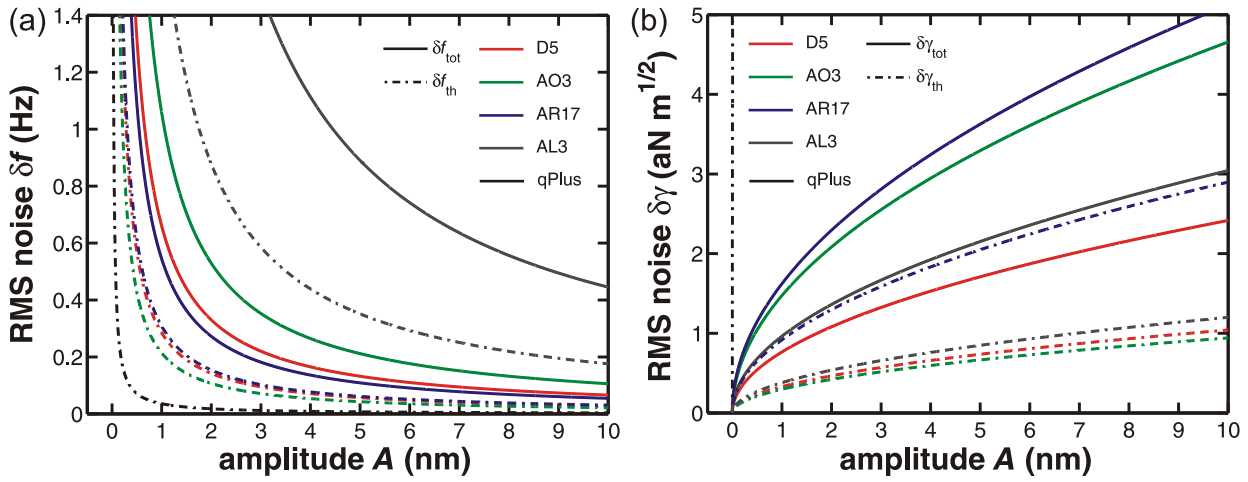

Figure 8: (a) RMS frequency-shift noise δ*f* and (b) normalised RMS frequency-shift noise δ*γ* in the limit of purely thermal noise (dash-dotted lines) as well as in combination with the corresponding detection noise (solid lines) for different cantilevers (T = 300 K, filter settings f_c = 500 Hz, o = 3, P = −2.0 Hz/deg and I = 1 Hz yielding B_{-3dB} = 258 Hz). Calculations are performed for the fundamental eigenfrequency f_0 of the cantilever. Cantilever properties and the corresponding values of the displacement-noise floor of the detection system d_{ds}^{z} are given in Table 1 and Table 2.

In Figure 8a, different cantilevers are compared regarding their total RMS frequency shift noise δf_{tot} (solid lines) as well as the thermal frequency noise δf_{th} (dash-dotted lines). Here, the same bandwidth of B_{-3dB} = 258 Hz is chosen for all simulations to facilitate the comparison of the cantilevers. Regarding thermal noise, all cantilevers except AL 3 exhibit an RMS noise below 0.5 Hz for amplitudes larger than 1 nm. The total noise values are ordered by the level of the corresponding noise floor d_{ds}^{z}, dominating the total noise for a bandwidth larger than the thermal-limit bandwidth. Note that the thermal-noise contribution of AL 3 is even larger than the total noise of the other cantilevers. These results are compared to typical values for a qPlus sensor with parameters taken from [11]. The thermal noise δf_{th} of the qPlus sensor is an order of magnitude below the values for the cantilevers. Including the noise of the detection system, δf_{tot} of the qPlus sensor is nearly identical to the thermal noise δf_{th} obtained for cantilever D 5 (curve not shown) and, therefore, only half of the noise level of the best cantilevers.

For a valid comparison of measurements obtained under different experimental conditions, however, it is important to compare limits in the normalised frequency shift γ rather than the plain frequency shift Δf. Based on the concept of the normalised frequency shift [15], we define a normalised-frequency-shift RMS noise as

$$\delta\gamma = kA^{3/2}\delta f / f_0 \qquad (11)$$

to compare the noise characteristics of cantilevers independently of their stiffness and resonance frequency and display the corresponding data as a function of the cantilever oscillation amplitude in Figure 8b. Regarding the thermal contribution $\delta\gamma_{th}$ to the normalised frequency-shift noise, cantilevers D 5 and AO 3 exhibit the best performance but are closely followed by cantilever AL 3. The $\delta\gamma_{th}$ value of AR 17 is even larger than the total noise $\delta\gamma_{tot}$ of cantilever D 5. This is presumably due to the large ratio k/f_0. Although cantilever AL 3 has the largest detection-system noise floor, its $\delta\gamma_{tot}$ is quite close to that of cantilever D 5. On the other hand, the qPlus sensor has a noise level $\delta\gamma$ more than two orders above the results for the cantilevers due to its exceptional k/f_0 ratio. Therefore, the advantageous noise figures of the qPlus sensor documented in Figure 8a can only be exploited if the sensor is operated at very low amplitudes.

As cantilevers D 5 and AL 3 have thermal bandwidth limits of B_{th}^{lim} = 95 Hz and B_{th}^{lim} = 85 Hz according to Equation 6, they are best suited for thermal-noise-limited operation. Operating them with a filter B_{-3dB} = 103 Hz (f_c = 125 Hz, o = 1) yields noise limits of $\delta\gamma_{tot}$ = 0.69 aN$\sqrt{m}$ and $\delta\gamma_{tot}$ = 0.84 aN$\sqrt{m}$, respectively, for an oscillation amplitude of A = 5 nm. Assuming, the detection noise floor of AL 3 could be decreased to 130 fm/$\sqrt{\text{Hz}}$ as for cantilever D 5, thermal-noise-limited operation with a bandwidth of B_{th}^{lim} = 586 Hz and $\delta\gamma_{tot}$ = 1.62 aN$\sqrt{m}$ would be possible for an oscillation amplitude of 5 nm and a filter setting of B_{-3dB} = 646 Hz (f_c = 1000 Hz, o = 3). This means that by switching from cantilever D 5 to AL 3, the usable bandwidth could be increased by a factor of six at the cost of

increasing $\delta\gamma_{tot}$ by a factor of two. In comparing such numbers, one should, however, consider that the assumed oscillation amplitude of 5 nm may be at the limit of stable operation [15], specifically for the soft cantilever D 5. In conclusion, the high-frequency and relatively stiff cantilever AL 3 represents an excellent choice for high-speed measurements with small amplitudes and good noise performance, while the larger and softer cantilever D 5 is the better choice for slower measurements with best possible noise performance.

Conclusion

We investigated the relation between the displacement noise in NC-AFM measurements and the corresponding frequency-shift noise at the output of the demodulator and demonstrated that predictions based on the demodulator transfer function and filtering are well reproduced by experiments. For a quantitative analysis of the noise, a precise amplitude calibration of the detection system relating electrical signals to the mechanical oscillation of the cantilever is inevitable. The displacement noise of an NC-AFM system can be measured directly with a spectrum analyser at the output of the detection system, and the thermal component of the displacement noise extracted from such spectra agrees well with spectra derived from a model of thermal cantilever excitation. The noise contribution of the detection system can be obtained from the white-noise floor of the measured spectra. The knowledge of the detection-system transfer functions allows one to predict the frequency-shift noise from the measured displacement noise, and by inversion, a measurement of the detection-system noise from the frequency-shift noise is possible. While the former analysis requires a spectrum analyser with very high resolution and an operating range that includes the eigenfrequency of the cantilever, the latter procedure requires only a measurement of the frequency-shift noise with a device covering the frequency range between a few hertz and about 10 kHz at moderate frequency resolution. Therefore, a rather complete noise characterisation with a simple spectrum analyser as integrated in many NC-AFM systems is possible for a calibrated system.

The framework of modelling noise in the NC-AFM system in combination with the experimental practice described here provides a clear guideline for system design and the choice of experimental parameters for thermal-noise-limited operation. The analysis shows that for a noise-optimised NC-AFM measurement, the right choice of the cantilever is most important, and obtaining a high effective Q-factor should be given great attention to keep the level of thermal noise at a minimum. The bandwidth of thermal-noise-limited operation is determined by the noise generated in the detection system. By an appropriate choice of PLL filter settings, one can make full use of this bandwidth without attenuating the NC-AFM signal while very efficiently eliminating most of the detection system noise. We find that with a technically optimised system and an appropriate choice of experimental parameters, room-temperature thermal-noise-limited NC-AFM measurements are possible over a bandwidth of 100 Hz and a detection limit smaller than 0.7 aN$\sqrt{m}$ for the normalised frequency shift operating at an amplitude of 5 nm.

Supporting Information

Supporting Information File 1

Experimental details and theory.
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-4-4-S1.pdf]

Acknowledgements

The authors gratefully acknowledge experimental support from Stefan Kuhn and Hans-Hermann Pieper and most helpful discussions with Steffen Porthun, Sadik Hafizovic, Kei Kobayashi and Hirofumi Yamada. Substantial help from Alfred Ziegler with the mathematics of Section 2 in Supporting Information File 1 is gratefully acknowledged. The project has been generously supported by Nanoworld Services GmbH.

References

1. Albrecht, T. R.; Grütter, P.; Horne, D.; Rugar, D. *J. Appl. Phys.* **1991,** *69,* 668–673. doi:10.1063/1.347347
2. Kobayashi, K.; Yamada, H.; Itoh, H.; Horiuchi, T.; Matsushige, K. *Rev. Sci. Instrum.* **2001,** *72,* 4383–4387. doi:10.1063/1.1416104
3. Fukuma, T.; Kimura, M.; Kobayashi, K.; Matsushige, K.; Yamada, H. *Rev. Sci. Instrum.* **2005,** *76,* 053704. doi:10.1063/1.1896938
4. Kobayashi, K.; Yamada, H.; Matsushige, K. *Rev. Sci. Instrum.* **2009,** *80,* 043708. doi:10.1063/1.3120913
5. Polesel-Maris, J.; Venegas de la Cerda, M. A.; Martrou, D.; Gauthier, S. *Phys. Rev. B* **2009,** *79,* 235401. doi:10.1103/PhysRevB.79.235401
6. Lübbe, J.; Temmen, M.; Schnieder, H.; Reichling, M. *Meas. Sci. Technol.* **2011,** *22,* 055501. doi:10.1088/0957-0233/22/5/055501
7. Torbrügge, S.; Lübbe, J.; Tröger, L.; Cranney, M.; Eguchi, T.; Hasegawa, Y.; Reichling, M. *Rev. Sci. Instrum.* **2008,** *79,* 083701. doi:10.1063/1.2964119
8. Rode, S.; Stark, R.; Lübbe, J.; Tröger, L.; Schütte, J.; Umeda, K.; Kobayashi, K.; Yamada, H.; Kühnle, A. *Rev. Sci. Instrum.* **2011,** *82,* 073703. doi:10.1063/1.3606399
9. Lübbe, J.; Tröger, L.; Torbrügge, S.; Bechstein, R.; Richter, C.; Kühnle, A.; Reichling, M. *Meas. Sci. Technol.* **2010,** *21,* 125501. doi:10.1088/0957-0233/21/12/125501
10. Lübbe, J.; Doering, L.; Reichling, M. *Meas. Sci. Technol.* **2012,** *23,* 045401. doi:10.1088/0957-0233/23/4/045401
11. Giessibl, F. J.; Pielmeier, F.; Eguchi, T.; An, T.; Hasegawa, Y. *Phys. Rev. B* **2011,** *84,* 125409. doi:10.1103/PhysRevB.84.125409
12. Lübbe, J.; Temmen, M.; Rahe, P.; Kühnle, A.; Reichling, M. *Beilstein J. Nanotechnol.* **2013,** submitted.

13. Melcher, J.; Hu, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 053101. doi:10.1063/1.2767173
14. Giessibl, F. J.; Bielefeldt, H.; Hembacher, S.; Mannhart, J. *Appl. Surf. Sci.* **1999,** *140,* 352–357. doi:10.1016/S0169-4332(98)00553-4
15. Giessibl, F. J. *Phys. Rev. B* **1997,** *56,* 16010–16015. doi:10.1103/PhysRevB.56.16010

License and Terms

Beilstein Journal of Nanotechnology

Interpreting motion and force for narrow-band intermodulation atomic force microscopy

Daniel Platz*[1], Daniel Forchheimer[1], Erik A. Tholén[2] and David B. Haviland[1]

Full Research Paper

Open Access

Address:
[1]Royal Institute of Technology (KTH), Section for Nanostructure Physics, Albanova University Center, SE-106 91 Stockholm, Sweden and [2]Intermodulation Products AB, Vasavägen 29, SE-169 58 Solna, Sweden

Email:
Daniel Platz* - platz@kth.se

* Corresponding author

Keywords:
atomic force microscopy; AFM; frequency combs; force spectroscopy; high-quality-factor resonators; intermodulation; multifrequency

Beilstein J. Nanotechnol. **2013,** *4,* 45–56.
doi:10.3762/bjnano.4.5

Received: 30 July 2012
Accepted: 26 November 2012
Published: 21 January 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

Intermodulation atomic force microscopy (ImAFM) is a mode of dynamic atomic force microscopy that probes the nonlinear tip–surface force by measurement of the mixing of multiple modes in a frequency comb. A high-quality factor cantilever resonance and a suitable drive comb will result in tip motion described by a narrow-band frequency comb. We show, by a separation of time scales, that such motion is equivalent to rapid oscillations at the cantilever resonance with a slow amplitude and phase or frequency modulation. With this time-domain perspective, we analyze single oscillation cycles in ImAFM to extract the Fourier components of the tip–surface force that are in-phase with the tip motion (F_I) and quadrature to the motion (F_Q). Traditionally, these force components have been considered as a function of the static-probe height only. Here we show that F_I and F_Q actually depend on both static-probe height and oscillation amplitude. We demonstrate on simulated data how to reconstruct the amplitude dependence of F_I and F_Q from a single ImAFM measurement. Furthermore, we introduce ImAFM approach measurements with which we reconstruct the full amplitude and probe-height dependence of the force components F_I and F_Q, providing deeper insight into the tip–surface interaction. We demonstrate the capabilities of ImAFM approach measurements on a polystyrene polymer surface.

Introduction

Since its invention [1] atomic force microscopy (AFM) has developed into one of the most versatile techniques in surface science. At length scales ranging from micrometers down to the level of single atoms, AFM-based techniques are used to image [2-4], measure [5,6] and manipulate matter [7-9] at an interface. As an imaging tool, the goal of AFM development has been to increase spatial resolution and minimize the back-action force from the probe on the sample surface. A major advancement in this regard was the development of dynamic AFM [10] in which a sharp tip at the free end of the AFM cantilever oscillates close to the sample surface, as depicted in Figure 1. In order to achieve stable oscillatory motion, an external drive force is

applied to the cantilever, which is usually purely sinusoidal in time with a frequency that is close to the resonance frequency of the first flexural eigenmode of the cantilever. The high-quality factor of the resonance ensures that the responding motion of the tip is approximately sinusoidal in time, with the same frequency as the drive signal [11,12]. Such periodic motion is best analyzed in the frequency or Fourier domain, where the motion is well described by one complex-valued Fourier coefficient at the drive frequency. This motion has a corresponding Fourier coefficient of the tip–surface force, which can be expressed in terms of two real-valued components, F_I, which is in-phase with the motion, and F_Q, which is quadrature to the motion. At a fixed probe height h above the surface, the two force quadratures F_I and F_Q give only qualitative insight into the interaction between the tip and the surface [13] and most quantitative force reconstruction methods are based on a measurement of F_I and F_Q at different h [14-19].

In order to increase the accessible information while imaging with AFM, a variety methods have been put forward in which amplitude and phase at more than one frequency are analyzed. These multifrequency methods can be divided in to two general groups: those using only Fourier components with frequencies close to a cantilever resonance, and those that use off-resonance components. Off-resonance techniques typically measure higher harmonics of the tip motion, which allows for a reconstruction of time-dependent surface forces acting on the tip. Due to the lack of transfer gain off resonance, these off-resonance components have small signal-to-noise ratio and their measurement requires special cantilevers [20], high interaction forces [21] or highly damped environments [22]. To increase the number of Fourier components with good signal-to-noise ratio, on-resonance techniques utilize multiple eigenmodes of the cantilever [23-26]. However, accurate calibration of higher cantilever modes remains complicated since additional knowledge about the cantilever is required. Both on- and off-resonance techniques require broad-band detection of the cantilever motion, which implies a sacrifice in the sensitivity and gain of the motion-detection system.

To mitigate these problems, we have developed narrow-band intermodulation AFM (ImAFM), which analyzes the response only near the first flexural eigenmode. In general ImAFM utilizes frequency mixing due to the nonlinear tip–surface inter-

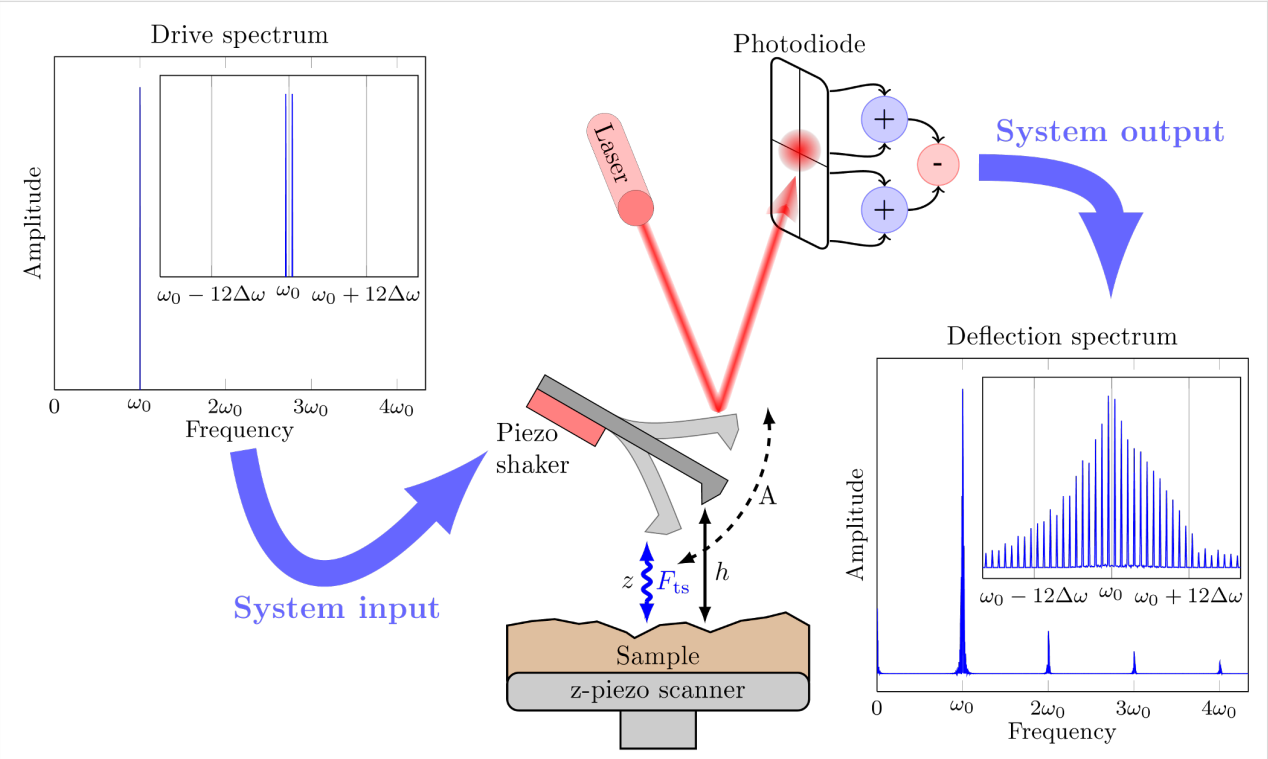

Figure 1: Sketch of the basic experimental setup in narrow-band ImAFM. In the absence of a drive signal the tip is at rest at the static probe height h. The spectrum of the drive comb consists of two frequency components spaced by $\Delta\omega$ and centered at the first flexural resonance frequency ω_0 of the cantilever, which is much higher than the comb base frequency $\Delta\omega$ (here $\omega_0 = 600 \times \Delta\omega$). The driven tip oscillates with amplitude A and interacts with the sample surface. The instantaneous tip position z is measured in the rest frame of the sample surface. The corresponding deflection signal is detected by an optical lever system and is concentrated to a narrow band around ω_0, as the drive signal. In this band, new frequency components spaced by $\Delta\omega$ are present, which are generated by the nonlinear tip–sample interaction F_{ts}. Outside the narrow band at ω_0 there is only a small response in bands at integer multiples of ω_0.

action. A drive signal that comprises multiple frequency components is used for exciting the cantilever, which will exhibit response not only at the drive frequencies, but also at frequencies that are linear, integer combinations of the drive frequencies:

$$\omega_{\text{imp}} = m_1\omega_1 + m_2\omega_2 + \ldots + m_M\omega_M \qquad (1)$$

where $m_1, m_2, \ldots, m_M \in \mathbb{Z}$ and $\omega_1, \omega_2, \ldots, \omega_M$ are the drive frequencies. These new frequency components are called intermodulation products (IMPs) and one usually defines an order for each IMP that is given by $|m_1| + |m_2| + \ldots + |m_M|$. If all frequencies in a signal are integer multiples of a base frequency, $\Delta\omega$, the signal is called a frequency comb. The nonlinear tip–surface interaction maps a drive-frequency comb to a response-frequency comb, both having the same base frequency $\Delta\omega$. Different drive-frequency combs can be used to place many response frequency components close to a resonance of the cantilever where they can be detected with good signal-to-noise ratio. In general the drive and response frequency combs could encompass more than one eigenmode of the cantilever. For a drive signal consisting of only two frequencies symmetrically placed around the first flexural resonance frequency, as illustrated in Figure 1, the response is concentrated in the narrow band around the first resonance, for which very accurate calibration methods exist [27-29].

In what follows, we will focus on this particular case, which we call narrow-band ImAFM. However, we want to emphasize that drive schemes that generate response in more than one frequency band are also possible. We have previously shown how the individual amplitudes [26] and phases [30] of the IMPs in the narrow band around the first flexural resonance can be used for imaging. Furthermore, a polynomial reconstruction of the tip–surface force [31,32] and a numerical fit of the parameters of a force model [33] are possible by analysis of the data in the frequency domain. Here, we consider the meaning of the narrow-band intermodulation response comb in the time domain, which leads to a physical interpretation of the intermodulation spectrum in terms of the in-phase force component F_I and the quadrature force component F_Q.

Results and Discussion

Time-domain interpretation of narrow-band frequency comb

Figure 2a portrays the amplitudes of the components of a narrow-band frequency comb. Whereas we only plot the amplitude of each component, it is understood that each component also has a phase. The frequency comb is characterized by the center frequency $\bar{\omega}$ of the comb, a base frequency $\Delta\omega$ and a finite number N of Fourier components at discrete frequencies. Without loss of generality, we assume N to be even. The center frequency $\bar{\omega}$ can be described in terms of the ratio $\bar{n} = \bar{\omega}/\Delta\omega$ and the bandwidth of the comb is given by $\Delta N = N - 1$. The N discrete frequencies ω_n in the band are represented by an integer frequency index n such that

$$\omega_n = n\Delta\omega \qquad (2)$$

where n takes consecutive integer values between $\bar{n} - \Delta N/2$ and $\bar{n} + \Delta N/2$. In the time domain, the corresponding real-valued signal $x(t)$ is then given by the Fourier series

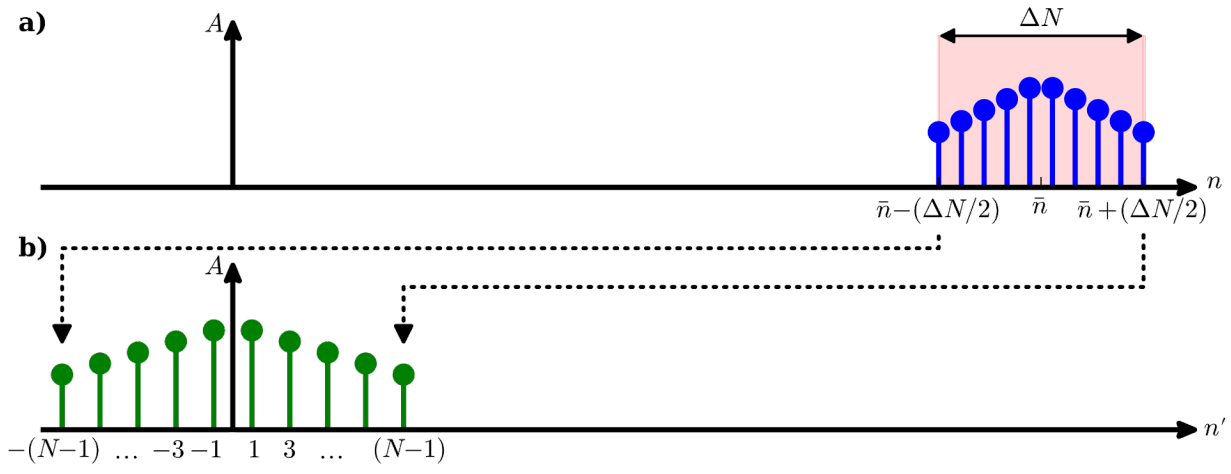

Figure 2: The amplitude spectrum of a narrow band signal as a function of the Fourier index (a) is characterized by a finite number of Fourier components in a frequency band around the center frequency $\bar{\omega} = \bar{n}\,\Delta\omega$, in a bandwidth that is given by the integer number ΔN. The spectrum of the corresponding time-dependent envelope function (b) is obtained by down-shifting the original spectrum in frequency space such that the shifted center frequency is zero.

$$x(t)=\sum_{n=\bar{n}-\Delta N/2}^{\bar{n}+\Delta N/2}\hat{x}_n e^{in\Delta\omega t}+\sum_{n=\bar{n}-\Delta N/2}^{\bar{n}+\Delta N/2}\hat{x}_n^* e^{-in\Delta\omega t} \tag{3}$$

where $\hat{x}_n$ are the complex Fourier components in the narrow frequency band and the star denotes complex conjugation. The center frequency $\bar{\omega}$ is usually much bigger than the base frequency $\Delta\omega$,

$$\bar{\omega} >> \Delta\omega. \tag{4}$$

Therefore, the time-domain signal $x(t)$ exhibits two different time scales: a fast time scale $T_{\text{fast}} = 2\pi/\bar{\omega}$ and a slow time scale $T_{\text{slow}} = 2\pi/\Delta\omega$. To separate these two time scales, we factor out a rapidly oscillating term at the frequency $\bar{\omega}$ from the Fourier series in Equation 3,

$$\begin{aligned}x(t)=&\sum_{n=\bar{n}-\Delta N/2}^{\bar{n}+\Delta N/2}\hat{x}_n e^{i(n\Delta\omega-\bar{\omega})t}e^{i\bar{\omega}t}\\&+\sum_{n=\bar{n}-\Delta N/2}^{\bar{n}+\Delta N/2}\hat{x}_n^* e^{-i(n\Delta\omega-\bar{\omega})t}e^{-i\bar{\omega}t},\end{aligned} \tag{5}$$

$$\begin{aligned}=&\sum_{n=\bar{n}-\Delta N/2}^{\bar{n}+\Delta N/2}\hat{x}_n e^{i(2n-2\bar{n})\frac{\Delta\omega}{2}t}e^{i\bar{\omega}t}\\&+\sum_{n=\bar{n}-\Delta N/2}^{\bar{n}+\Delta N/2}\hat{x}_n^* e^{-i(2n-2\bar{n})\frac{\Delta\omega}{2}t}e^{-i\bar{\omega}t}.\end{aligned} \tag{6}$$

Since N is even, $2\bar{n}$ is an odd integer number and we can define a new sum index

$$n' \equiv 2n-2\bar{n} \tag{7}$$

which increases in steps of $\Delta n' = 2$ and the summation limits become

$$n_1' = 2\left(\bar{n}-\frac{\Delta N}{2}\right)-2\bar{n} = -(N-1) \tag{8}$$

$$n_N' = 2\left(\bar{n}+\frac{\Delta N}{2}\right)-2\bar{n} = N-1. \tag{9}$$

Since N is even and n' increases only in steps of 2, n' can only take odd values. Additionally, we define new Fourier coefficients

$$\hat{x}_{n'}' \equiv \hat{x}_{\frac{1}{2}(n'+2\bar{n})} \tag{10}$$

such that the signal Fourier series becomes

$$\begin{aligned}x(t)=&\left(\sum_{\substack{n'=-(N-1)\\ n'\text{ odd}}}^{N-1}\hat{x}_{n'}' e^{in'\frac{\Delta\omega}{2}t}\right)e^{i\bar{\omega}t}\\&+\left(\sum_{\substack{n'=-(N-1)\\ n'\text{ odd}}}^{N-1}\hat{x}_{n'}'^* e^{-in'\frac{\Delta\omega}{2}t}\right)e^{-i\bar{\omega}t}.\end{aligned} \tag{11}$$

We identify the terms in parentheses as the Fourier series of a complex-valued time-dependent envelope function $\hat{E}(t)$ expanded in the base frequency $\Delta\omega/2$,

$$\hat{E}(t)=\sum_{\substack{n'=-(N-1)\\ n'\text{ odd}}}^{N-1}\hat{x}_{n'}' e^{in'\frac{\Delta\omega}{2}t} \tag{12}$$

and write the original signal $x(t)$ as

$$x(t)=\hat{E}(t)e^{i\bar{\omega}t}+\hat{E}^*(t)e^{-i\bar{\omega}t}. \tag{13}$$

The envelope function $\hat{E}(t)$ was obtained by down-shifting the narrow intermodulation frequency band to a center frequency of zero (see Figure 2). If the maximum frequency in the Fourier series of $\hat{E}(t)$ is much smaller than $\bar{\omega}$, the envelope function $\hat{E}(t)$ varies slowly compared to the term $e^{i\bar{\omega}t}$ in Equation 13. When we represent $\hat{E}(t)$ by a time-dependent amplitude $A(t)$ and a time-dependent phase $\phi(t)$, such that

$$\hat{E}(t)=A(t)e^{i\phi(t)}, \tag{14}$$

the signal $x(t)$ is completely described by a modulated oscillation amplitude and a modulated oscillation phase:

$$\begin{aligned}x(t)&=A(t)e^{i\phi(t)}e^{i\bar{\omega}t}+A(t)e^{-i\phi(t)}e^{-i\bar{\omega}t}\\&=2A(t)\cos\left(\bar{\omega}t+\phi(t)\right).\end{aligned} \tag{15}$$

We would like to emphasize that the narrow-band frequency comb can also describe amplitude- and frequency-modulated

signals. For frequency modulation we define an instantaneous oscillation phase

$$\theta(t) = \bar{\omega}t + \phi(t) \tag{16}$$

and an instantaneous oscillation frequency

$$\omega(t) = \frac{d\theta}{dt} = \bar{\omega} + \frac{d\phi}{dt}. \tag{17}$$

The instantaneous frequency shift δω compared to $\bar{\omega}$ is then simply

$$\delta\omega(t) = \frac{d\phi}{dt}. \tag{18}$$

In a small region around the time t' the signal $x(t)$ can be obtained by a Taylor expansion to first order of the instantaneous phase θ:

$$\begin{aligned} x(t) &= A(t)\cos\big(\theta(t)\big) \\ &\approx A(t)\cos\Big(\big(\bar{\omega} + \delta\omega(t')\big)t + \big(\phi(t') - \delta\omega(t')t'\big)\Big). \end{aligned} \tag{19}$$

Thus, a narrow-band frequency comb can describe signals with frequency shifts that are periodic in time.

To illustrate the complete description of a narrow-band signal by its envelope function, Figure 3 shows the spectrum of an artificially constructed signal with sinusoidally modulated amplitude and frequency. Typical parameters from AFM experiments have been chosen for the amplitude and frequency modulation. The spectrum of the signal shown in Figure 3 shows significant amplitudes at discrete frequencies in only a narrow band around 300 kHz. We down-shift the spectrum to determine the slowly varying envelope function from which we compute the time-dependent oscillation amplitude and frequency, both of which are in excellent agreement with the actual amplitudes and frequencies used for the signal generation.

To summarize, we have introduced a time-domain interpretation of narrow-band frequency combs. If the center frequency of the band is much higher than the base frequency of the comb, we can separate a fast and a slow time scale in the time domain. On the fast time scale the signal rapidly oscillates at the center frequency. The slow-time-scale dynamics is given by the down-shifted intermodulation spectrum, which describes a slow amplitude modulation and a slow phase or frequency modulation of the signal in the time domain.

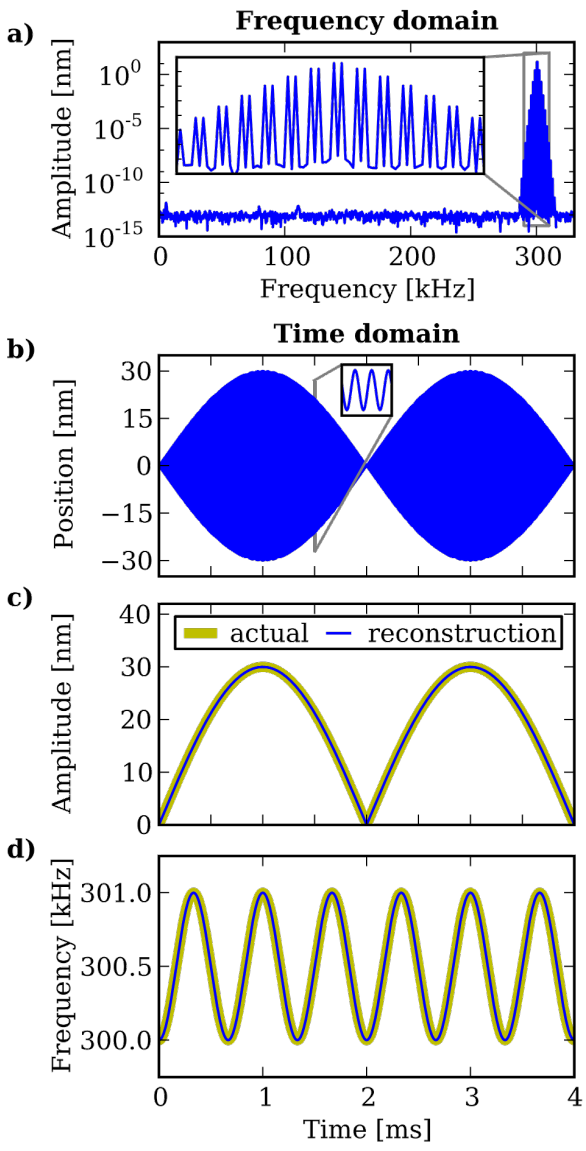

Figure 3: (a) The amplitude spectrum of an amplitude and frequency modulated signal. The response is concentrated in a narrow-band frequency comb with a center frequency much higher than the comb base frequency. In the time domain (b) the signal rapidly oscillates on a short time scale and the slow amplitude modulation is clearly visible. The time-dependent amplitude (c) and frequency (d) reconstructed from the envelope function are in excellent agreement with the actual modulation used for the signal generation in the time domain.

Physical interpretation of the tip motion and force envelope functions in ImAFM

In ImAFM the measured frequency comb corresponds to a vertical motion $z(t)$ of the tip, which undergoes rapid oscillations at frequency $\bar{\omega}$ with slowly varying amplitude and phase,

$$z(t) = A(t)\cos\big(\bar{\omega}t + \phi(t)\big) + h \tag{20}$$

where h is the static probe height above the surface, and the amplitude $A(t)$ and the phase $\phi(t)$ are determined from the complex-valued motion envelope function $\hat{E}_z(t)$ as

$$A(t) = \left|\hat{E}_z(t)\right|, \tag{21}$$

$$\phi(t) = \arg\left(\hat{E}_z(t)\right). \tag{22}$$

The envelope function $\hat{E}_z(t)$ was obtained directly from the measured motion spectrum by using Equation 12.

Knowledge of the cantilever transfer function $\hat{G}$ and the applied drive force allows for converting the measured motion spectrum into the spectrum of the time-dependent tip–surface force acting on the tip. However, the force spectrum is incomplete since higher frequency components of the force are filtered out from the motion spectrum by the sharply peaked cantilever transfer function. The time-dependence of the corresponding partial force signal is described by the force envelope function $\hat{E}_F(t)$:

$$F_{\text{partial}}(t) = \hat{E}_F(t)\mathrm{e}^{i\bar{\omega}t} + \hat{E}_F^*(t)\mathrm{e}^{-i\bar{\omega}t} \tag{23}$$

where $\hat{E}_F(t)$ is determined by applying Equation 12 to the partial force spectrum. To understand the physical meaning of the partial force, we analyze the signals at the level of single rapid oscillation cycles in the time domain. During each oscillation cycle the tip interacts with the sample surface. This interaction is very localized (a few nanometers above the surface) compared to the oscillation amplitude (tens of nanometers), and thus the interaction time T_{inter} is short compared to the fast oscillation period T_{fast}, which itself is much shorter than the period of the beat T_{slow} (see Figure 4),

$$T_{\text{inter}} < T_{\text{fast}} << T_{\text{slow}} \tag{24}$$

Therefore, amplitude $A(t)$, phase $\phi(t)$ and force envelope function $\hat{E}_F(t)$ can be considered to be constant during each interaction cycle, and the motion and the partial force during the i-th tip oscillation cycle are given by

$$z^{(i)}(t) = A^{(i)}\cos\left(\bar{\omega}t + \phi^{(i)}\right) + h \tag{25}$$

$$F_{\text{partial}}^{(i)}(t) = \hat{E}_F^{(i)}\mathrm{e}^{i\bar{\omega}t} + \hat{E}_F^{(i)*}\mathrm{e}^{-i\bar{\omega}t} \tag{26}$$

where $A^{(i)}$, $\Phi^{(i)}$ and $\hat{E}_F^{(i)}$ are constant and are determined at the time $t^{(i)}$ of the ith lower turning point of the tip motion

$$A^{(i)} = A\left(t^{(i)}\right) = \left|\hat{E}_z\left(t^{(i)}\right)\right| \tag{27}$$

$$\phi^{(i)} = \phi\left(t^{(i)}\right) = \arg\left(\hat{E}_z\left(t^{(i)}\right)\right) \tag{28}$$

$$\hat{E}_F^{(i)} = \hat{E}_F\left(t^{(i)}\right). \tag{29}$$

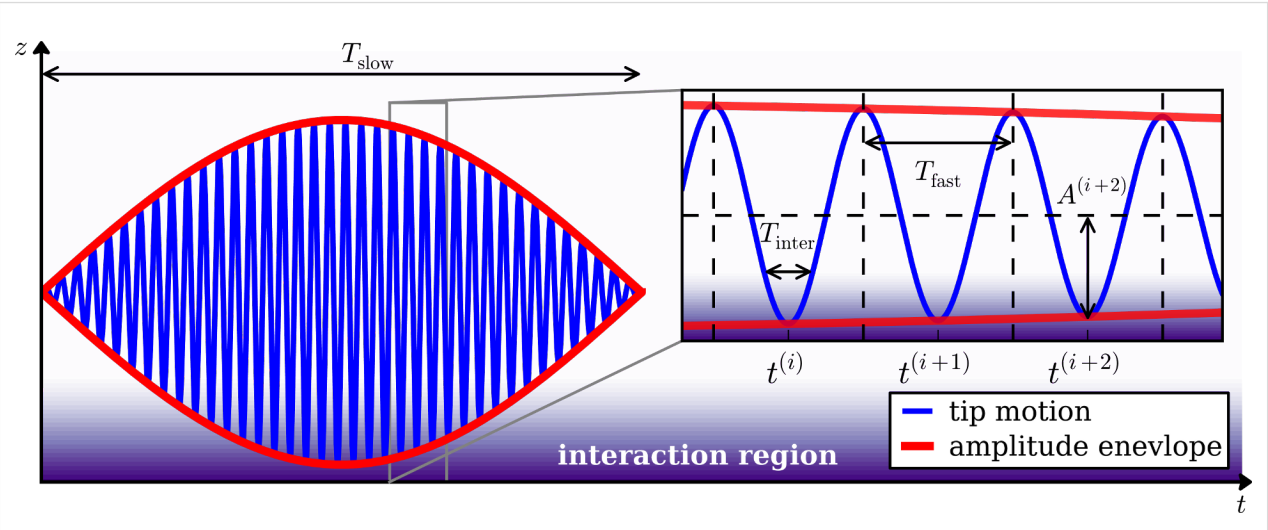

Figure 4: Sketch of a narrow-band signal. On the slow time scale T_{slow} the tip motion shows an amplitude modulation. On the fast time scale T_{fast} the signal rapidly oscillates. During each oscillation cycle the tip interacts with the surface for the time T_{inter} during which the oscillation amplitude and phase are approximately constant.

The complete time-dependent tip–surface force during an interaction cycle is a force pulse that can be written as a Fourier series in the oscillation frequency $\bar{\omega}$ as

$$F^{(i)}_{\text{complete}}(t) = \sum_{n=-\infty}^{\infty} \hat{F}^{(i)}_n \mathrm{e}^{in\bar{\omega}t} \tag{30}$$

where the complex Fourier components $\hat{F}^{(i)}_n$ fulfill the relation $\hat{F}^{(i)}_{-n} = \hat{F}^{(i)*}_n$. Comparison of Equation 26 with Equation 30 yields

$$\hat{F}^{(i)}_1 = \hat{E}^{(i)}_F = \hat{E}_F\left(t^{(i)}\right) \tag{31}$$

which reveals that the first Fourier component $\hat{F}^{(i)}_1$ of the force pulse during the ith oscillation cycle is given by the force envelope function $\hat{E}_F$ determined from the partial force spectrum. Since each lower motion turning point is associated with a unique amplitude $A^{(i)}$ we can consider $\hat{F}_1$ as a function of the continuous variable A,

$$\hat{F}^{(i)}_1 = \hat{F}_1(A^{(i)}) \Rightarrow \hat{F}_1 = \hat{F}_1(A). \tag{32}$$

The amplitude dependence of $\hat{F}_1$ can then be uncovered by the analysis of all oscillation cycles during the time T_{slow}.

To better relate motion and the force we compute the components of $\hat{F}_1(A)$ that are in phase with the motion (F_I) and quadrature to the motion (F_Q). For a tip–surface force that only depends on the instantaneous tip position and velocity,

$$F_{\text{ts}} = F_{\text{ts}}(z, \dot{z}), \tag{33}$$

we approximate the tip motion to be purely sinusoidal at frequency $\bar{\omega}$ with amplitude A and without an additional phase. At fixed probe height h, the components F_I and F_Q are given by two integral equations

$$F_I(A,h) = \frac{1}{T_{\text{fast}}}\int_0^{T_{\text{fast}}} \left\{ \begin{array}{r} F_{\text{ts}}\left(A\cos(\bar{\omega}t)+h, -\bar{\omega}A\sin(\bar{\omega}t)\right) \\ \times\cos(\bar{\omega}t) \end{array} \right\} \mathrm{d}t \tag{34}$$

$$F_Q(A,h) = \frac{1}{T_{\text{fast}}}\int_0^{T_{\text{fast}}} \left\{ \begin{array}{r} F_{\text{ts}}\left(A\cos(\bar{\omega}t)+h, -\bar{\omega}A\sin(\bar{\omega}t)\right) \\ \times\sin(\bar{\omega}t) \end{array} \right\} \mathrm{d}t. \tag{35}$$

With these assumptions F_I becomes the so-called virial of the tip–surface force, which is only affected by the conservative part of the tip–surface interaction [34], whereas F_Q is related to the energy dissipated by the tip–surface interaction [11]. We note that, through their dependence on tip position z and velocity $\dot{z}$, the force components F_I and F_Q are functions of both probe height h and oscillation amplitude A. However, they are usually considered as being functions of the probe height h only.

For an ImAFM measurement at fixed probe height, the amplitude dependence of F_I and F_Q can readily be obtained by defining a new force envelope function that is phase-shifted with respect to the motion by the angle ϕ:

$$\hat{E}'_F(t) = \hat{E}_F(t)\mathrm{e}^{-i\phi(t)} \tag{36}$$

which we evaluate as real and imaginary parts at the times of the lower turning points of the motion

$$F_I(A^{(i)}) = \mathrm{Re}\left(\hat{E}'_F(t^{(i)})\right), \tag{37}$$

$$F_Q(A^{(i)}) = \mathrm{Im}\left(\hat{E}'_F(t^{(i)})\right). \tag{38}$$

With this interpretation of an intermodulation spectrum we are able to reconstruct the amplitude dependence of the force quadratures F_I and F_Q, which are independent of details of the tip motion on the slow time scale. Due to this independence, F_I and F_Q are the input quantities for nearly all force spectroscopy methods in dynamic AFM and thereby they form the basis of quantitative dynamic AFM.

Force quadrature reconstruction from simulated data

To demonstrate the accuracy of the F_I(A) and $F_Q(A)$ reconstruction from ImAFM data we simulate the tip motion in a model force field. We excite the tip with two frequencies close to the first flexural resonance frequency of $\omega_0 = 2\pi \cdot 300$ kHz, which allows us to model the cantilever as a single eigenmode system for which the tip dynamics are described by an effective harmonic oscillator equation [35,36]

$$\ddot{z} + \frac{\omega_0}{Q}\dot{z} + \omega_0^2(z-h) = \frac{\omega_0^2}{k_c}\left(\begin{array}{r} F_1\cos(\omega_1 t) + F_2\cos(\omega_2 t) \\ +F_{\text{ts}}(z,\dot{z}) \end{array} \right) \tag{39}$$

where Q = 400 is the quality factor of the resonance, k_c = 40 N/m is the mode stiffness h = 20 nm is the static probe height above the surface. The drive strengths F_1 and F_2 at the frequencies $\omega_1 = 2\pi \cdot 299.75$ kHz and $\omega_2 = 2\pi \cdot 300.25$ kHz are chosen such that in the absence of a tip–surface force,

the tip oscillation amplitude is sinusoidally modulated between 0 and 30 nm. For the tip–surface force F_{ts} we assume a van-der-Waals–Derjaguin–Muller–Toporov (vdW-DMT) force with additional exponential damping, which is defined as

$$F_{ts}(z,\dot{z}) = \begin{cases} \left[-\dfrac{HR}{6(a_0+z)^2} - \gamma \exp\left(-z/z_\gamma\right)\dot{z} \right] & z \geq 0 \\ \left[-\dfrac{HR}{6a_0^2} + \dfrac{4}{3}\sqrt{R}E^*(-z)^{3/2} \right] & \\ \left[-\gamma \exp\left(-z/z_\gamma\right)\dot{z} \right] & z < 0 \end{cases} \quad (40)$$

where $H = 2.96 \cdot 10^{-7}$ J is the Hamaker constant, $R = 10$ nm is the tip radius, $\gamma = 2.2 \cdot 10^{-7}$ Ns/m is the damping constant, $z_\gamma = 1.5$ nm is the damping decay length and $E^* = 2.0$ GPa is the effective stiffness. For the numerical integration of Equation 39 we use the adaptive step-size integrator CVODE [37] with root detection to properly treat the piecewise definition of the tip–surface force in Equation 40. From the simulated tip motion we determine the motion and the force envelope functions $\hat{E}_z$ and $\hat{E}_F$ and reconstruct F_I(A) and $F_Q(A)$ according to Equation 37 and Equation 38. As shown in Figure 5 the reconstructed curves are in excellent agreement with the curves directly computed with Equation 34 and Equation 35 from the model force used in the simulations.

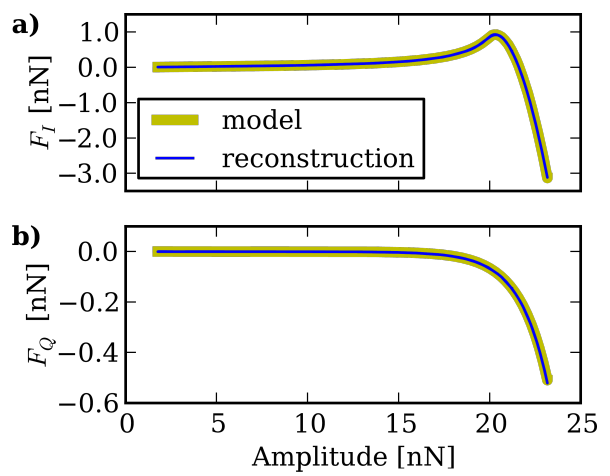

Figure 5: The $F_I(A)$ and $F_Q(A)$ curves reconstructed from simulated tip motion in ImAFM. The reconstructed curves are in good agreement with the actual curves directly determined from the model force used in the simulations.

Probing the force quadratures

The force quadratures F_I and F_Q are the basic input quantities for a variety of force-reconstruction techniques [14-19]. Over the past decade, the dominate paradigm was to consider F_I and F_Q as functions of the static probe height h only, and only one oscillation amplitude A was considered at each probe height. F_I and F_Q are, however, functions of both h and A as seen in the two-dimensional color maps shown in Figure 6 for the vdW-DMT force with exponential damping used in the previous section. In order to emphasize the interaction region near the point of contact, data in the h–A plane with $F_I < -8$ nN are masked with white.

In both frequency-modulation AFM (FM-AFM) and amplitude-modulation AFM (AM-AFM) F_I and F_Q are usually probed by a slow variation of the probe height h with fixed oscillation

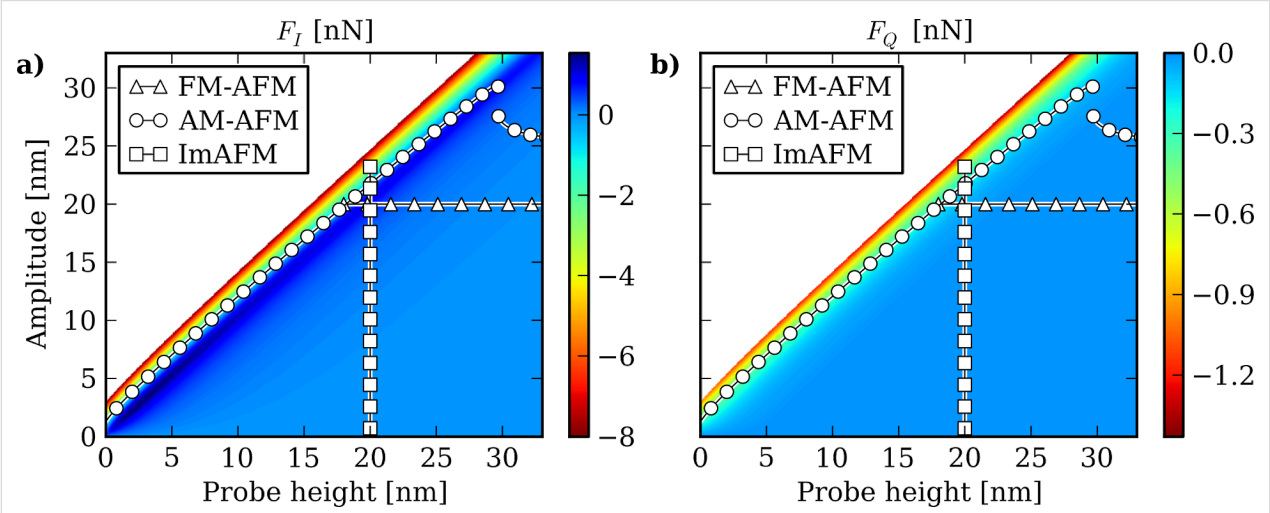

Figure 6: Model $F_I(h,A)$ and $F_Q(h,A)$ maps for the vdW-DMT force with exponential damping introduced in Equation 40. The displayed measurement paths correspond to a frequency-shift–distance curve in FM-AFM, an amplitude–phase–distance curve in AM-AFM and an ImAFM measurement. In contrast to FM-AFM and AM-AFM, the static probe height is constant during an ImAFM measurement and the h–A plane is explored along a path parallel to the A axis. One should also note the amplitude jump along the AM-AFM path at a probe height of h = 30 nm.

amplitude at each height. To measure F_I and F_Q in FM-AFM the oscillation frequency shift and the drive force are recorded as the static probe height is slowly varied (frequency-shift–distance curves). Active feedback is used to adjust both the drive power and drive frequency, to keep the response amplitude and phase constant. The obtained frequency shifts and drive forces can then be converted into the force quadratures [38,39] so that the measurement corresponds to a measurement of F_I and F_Q along a path parallel to the h-axis in the h–A plane (see Figure 6).

In AM-AFM the oscillation amplitude and phase with respect to the drive force are measured as a function of the static probe height h (amplitude–phase–distance curves) and are then converted into values of the force quadratures [13]. In contrast to FM-AFM, the oscillation amplitude is free to change during the measurement and thus the AM-AFM measurement path in the h–A plane is more complicated. The path shown in Figure 6 was obtained by simulating the AM-AFM tip dynamics with cvode. In the simulations we used the same cantilever and force parameters as in the previous section, with a drive signal at only one frequency of ω_d = 300 kHz. As is often the case with AM-AFM, the amplitude–phase–distance curve exhibits an abrupt amplitude jump due to the existence of multiple oscillation states [40]. This instability is frequently observed in experiments, and it makes the reconstruction of tip–surface forces rather difficult.

In contrast to FM-AFM and AM-AFM, ImAFM allows for a measurement of F_I and F_Q at fixed static probe height, along a straight path parallel to the A axis in the h–A plane, as shown in Figure 6 for the simulation of the previous section. Each of these three measurement techniques probes the tip–surface interaction along a different path in the h–A plane. With ImAFM, however, the measurement can be rapidly performed at each point of an image, while scanning with normal speed [26], allowing for unprecedented ability to analyze the tip–surface force while imaging. The ImAFM spectral data, which is concentrated to a narrow band near resonance, is a complete representation of the measurable tip motion, because there is only noise outside this narrow frequency band. Thus the method optimally extracts the signal for compact storage and further analysis.

We note that the ImAFM path provides an equivalent amount of information to that in frequency-shift–distance or amplitude–phase–distance curves. This implies that, for a single scan, ImAFM image information equivalent to a frequency-shift–distance curve or an amplitude–phase–distance curve is available in every image point. Moreover, the ImAFM measurement does not suffer from amplitude jumps since the stiffness of the cantilever resonance prevents big amplitude changes from one single oscillation cycle to the next single oscillation in the beat tip motion.

ImAFM approach measurements

It is possible to acquire maps of F_I and F_Q in the full h–A plane with a protocol we call ImAFM approach measurements. Similar to the measurement of frequency-shift or amplitude–phase curves, the static probe height above the surface is varied by slowly extending the z-piezo toward the surface. However, in contrast to FM-AFM and AM-AFM measurements, the oscillation amplitude is rapidly modulated as the probe slowly approaches the surface. Because the height variation is much slower (order of seconds) than the amplitude modulation (order of milliseconds), the probe height can be considered to be constant during each amplitude modulation. In this case each amplitude modulation reveals the amplitude dependence of F_I and F_Q at a constant probe height. From the different probe heights F_I and F_Q can be reconstructed in the full h–A plane. With FM-AFM or AM-AFM such a measurement would require much longer measurement time since multiple surface approaches with different amplitudes would be required. With ImAFM all the data is acquired during a single surface approach.

We use ImAFM approach curves to reconstruct F_I and F_Q maps on a polystyrene (PS) polymer surface. We perform a slow surface approach, and from the acquired data we reconstruct the F_I and F_Q maps shown in Figure 7. On the h axis we show the piezo extension since the absolute probe height cannot be defined unambiguously in an experiment. The areas in the h–A plane that were not explored are displayed as white areas.

The boundary to the white area in the upper part of the plots represents the maximum oscillation achievable for the fixed drive power. The boundary to the white area in the lower part of the plot corresponds to the minimum oscillation amplitude during one modulation period (one beat). This lower boundary shows interesting variations with the piezo extension. Further away from the surface at larger piezo extension, only positive values of F_I are achieved, which corresponds to the tip oscillating in a region where the net conservative force is purely attractive. As the surface is further approached, F_I also takes on negative values, when the net force becomes repulsive. At probe heights between 7 and 13 nm, the strongest repulsive force is experienced. Around h = 13 nm and below, the attractive region vanishes, which may be the result of a change in the cantilever dynamics due to a relatively long interaction time, or a change in the hydrodynamic damping forces due to the surrounding air close to the sample surface. One should also note that at this piezo extension the minimum oscillation ampli-

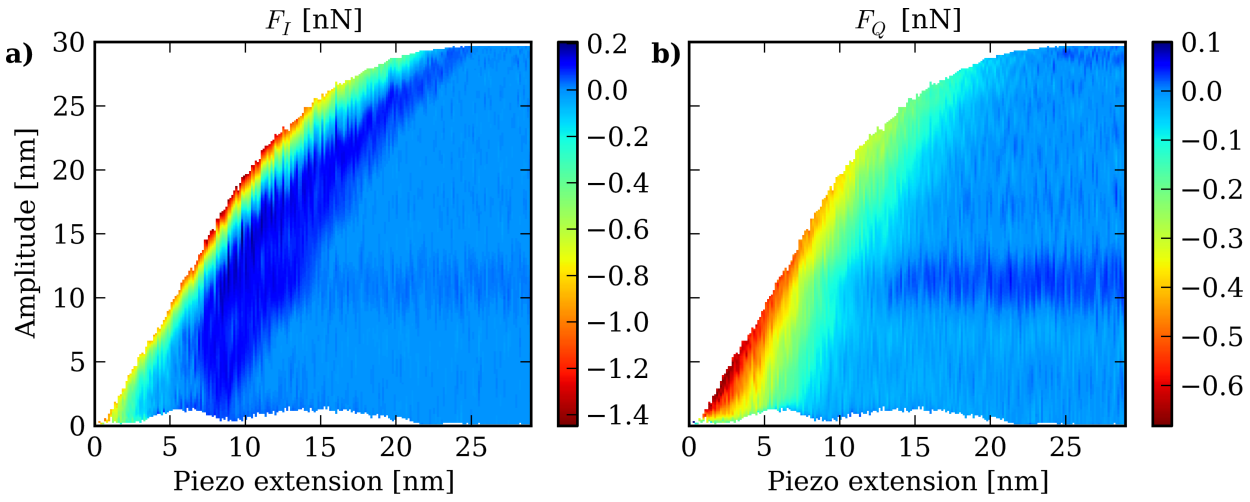

Figure 7: $F_I(h,A)$ and $F_Q(h,A)$ maps reconstructed from an ImAFM approach measurements on a PS surface. The *z*-piezo extension corresponds to a relative change of the probe height *h* above the sample surface.

tude begins to increase again. A possible artifact of the measurement method may result in this low amplitude region, if the motion spectrum is no longer confined to a narrow frequency band, as assumed in the analysis.

The map of F_Q characterizes the dissipative interaction between tip and sample. The dissipative tip–surface force can be much more complex than the conservative part of the interaction since dissipative forces do not only depend on the instantaneous tip position, as with the conservative force. The F_Q map can provide detailed insight into the nature of the dissipative interaction since the full dependence of F_Q on probe height h and oscillation amplitude A is measured. The F_Q map shows a lower level of force than the F_I map and it therefore appears more noisy, because the tip–surface interaction is predominately conservative. The small positive values of F_Q that occur far from the surface would imply that the tip gained energy from the interaction with the surface. This may be an artifact, but another possible explanation is some sort of hydrodynamic mode above the surface [41]. In both the attractive and repulsive region of the in-phase force F_I, the quadrature force F_Q is predominantly negative and it decreases as the surface is indented, corresponding to a increasingly dissipative tip–surface interaction. However, the maximum dissipation does not coincide with the maximum repulsive conservative force, and the energy dissipation is largest at peak amplitude for piezo extensions between 2 and 6 nm. Another interesting feature of the F_I map is the fine structure in the contact region, between h = 10 and 20 nm piezo extension. These small step-like changes of conservative force are not present in the smooth force model function, and could be an indication that the dissipative forces result in small, irreversible modifications of the sample surface.

Conclusion

We presented a physical interpretation of tip motion when described by a narrow-band frequency comb in ImAFM. We showed by separation of time scales that the time domain signal of a narrow-band frequency comb is completely characterized by a complex-valued envelope function and a rapidly oscillating term. The application of this time-domain picture to ImAFM allows for the reconstruction of two force quadratures F_I and F_Q as functions of the oscillation amplitude A. The quantities F_I and F_Q can be considered as two-dimensional functions, depending on both the probe height and the oscillation amplitude. Within this framework we find a connection between frequency-shift–distance curves in FM-AFM, amplitude–phase–distance curves in AM-AFM, and ImAFM measurements. Moreover, we introduced ImAFM approach measurements, which allow for a rapid and complete reconstruction of F_I and F_Q in the full h–A plane, providing detailed insight into the interaction between tip and surface. We demonstrated the reconstruction of F_I and F_Q maps experimentally on a PS polymer surface. We hope that the physical interpretation of narrow-band dynamic AFM presented here, will inspire new force-spectroscopy methods in the future that will take advantage of the high signal-to-noise ratio and the high acquisition speed of ImAFM.

Experimental

The PS sample was spin cast from toluene solution on a silicon oxide substrate. Both PS (M_w = 280 kDa) and toluene were obtained from Sigma-Aldrich and used as purchased. The measurements where performed with a Veeco Multimode II and a Budget Sensor BS300Al-G cantilever with a resonance frequency of f_0 = 311.838 kHz, a quality factor of Q = 539.9,

and a stiffness of k_c = 29.5 N/m, which was determined by thermal calibration [29]. We chose the two drive frequencies f_1 = 311.585 kHz and f_2 = 312.085 kHz symmetrically around the resonance frequency and the drive strengths such that the free oscillation amplitude was modulated between 0.0 and 29.7 nm. The probe height was changed with a speed of 5.0 nm/s.

References

1. Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986,** *56,* 930–933. doi:10.1103/PhysRevLett.56.930
2. Ohnesorge, F.; Binnig, G. *Science* **1993,** *260,* 1451–1456. doi:10.1126/science.260.5113.1451
3. Giessibl, F. J. *Science* **1995,** *267,* 68–71. doi:10.1126/science.267.5194.68
4. Hansma, P. K.; Cleveland, J. P.; Radmacher, M.; Walters, D. A.; Hillner, P. E.; Bezanilla, M.; Fritz, M.; Vie, D.; Hansma, H. G.; Prater, C. B.; Massie, J.; Fukunaga, L.; Gurley, J.; Elings, V. B. *Appl. Phys. Lett.* **1994,** *64,* 1738–1741. doi:10.1063/1.111795
5. Rugar, D.; Yannoni, C. S.; Sidles, J. A. *Nature* **1992,** *360,* 563–566. doi:10.1038/360563a0
6. Butt, H.-J.; Cappella, B.; Kappl, M. *Surf. Sci. Rep.* **2005,** *59,* 1–152. doi:10.1016/j.surfrep.2005.08.003
7. Pérez-Murano, F.; Abadal, G.; Barniol, N.; Aymerich, X.; Servat, J.; Gorostiza, P.; Sanz, F. *J. Appl. Phys.* **1995,** *78,* 6797–6801. doi:10.1063/1.360505
8. García, R.; Calleja, M.; Rohrer, H. *J. Appl. Phys.* **1999,** *86,* 1898–1903. doi:10.1063/1.370985
9. Sugimoto, Y.; Abe, M.; Hirayama, S.; Oyabu, N.; Custance, O.; Morita, S. *Nat. Mater.* **2005,** *4,* 156–159. doi:10.1038/nmat1297
10. Martin, Y.; Williams, C. C.; Wickramasinghe, H. K. *J. Appl. Phys.* **1987,** *61,* 4723–4729. doi:10.1063/1.338807
11. Cleveland, J. P.; Anczykowski, B.; Schmid, A. E.; Elings, V. B. *Appl. Phys. Lett.* **1998,** *72,* 2613–2615. doi:10.1063/1.121434
12. Salapaka, M. V.; Chen, D. J.; Cleveland, J. P. *Phys. Rev. B* **2000,** *61,* 1106–1115. doi:10.1103/PhysRevB.61.1106
13. San Paulo, A.; García, R. *Phys. Rev. B* **2002,** *66,* 041406(R). doi:10.1103/PhysRevB.66.041406
14. Dürig, U. *Appl. Phys. Lett.* **2000,** *76,* 1203–1205. doi:10.1063/1.125983
15. Sader, J. E.; Uchihashi, T.; Higgins, M. J.; Farrell, A.; Nakayama, Y.; Jarvis, S. P. *Nanotechnology* **2005,** *16,* S94–S101. doi:10.1088/0957-4484/16/3/018
16. Hölscher, H. *Appl. Phys. Lett.* **2006,** *89,* 123109. doi:10.1063/1.2355437
17. Lee, M.; Jhe, W. *Phys. Rev. Lett.* **2006,** *97,* 036104. doi:10.1103/PhysRevLett.97.036104
18. Hu, S.; Raman, A. *Nanotechnology* **2008,** *19,* 375704. doi:10.1088/0957-4484/19/37/375704
19. Katan, A. J.; van Es, M. H.; Oosterkamp, T. H. *Nanotechnology* **2009,** *20,* 165703. doi:10.1088/0957-4484/20/16/165703
20. Sahin, O.; Magonov, S.; Su, C.; Quate, C. F.; Solgaard, O. *Nat. Nanotechnol.* **2007,** *2,* 507–514. doi:10.1038/nnano.2007.226
21. Stark, M.; Stark, R. W.; Heckl, W. M.; Guckenberger, R. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 8473–8478. doi:10.1073/pnas.122040599
22. Legleiter, J.; Park, M.; Cusick, B.; Kowalewski, T. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103,* 4813–4818. doi:10.1073/pnas.0505628103
23. Rodríguez, T. R.; García, R. *Appl. Phys. Lett.* **2004,** *84,* 449–451. doi:10.1063/1.1642273
24. Solares, S. D.; Chawla, G. *J. Appl. Phys.* **2010,** *108,* 054901. doi:10.1063/1.3475644
25. Kawai, S.; Glatzel, T.; Koch, S.; Such, B.; Baratoff, A.; Meyer, E. *Phys. Rev. Lett.* **2009,** *103,* 220801. doi:10.1103/PhysRevLett.103.220801
26. Platz, D.; Tholén, E. A.; Pesen, D.; Haviland, D. B. *Appl. Phys. Lett.* **2008,** *92,* 153106. doi:10.1063/1.2909569
27. Hutter, J. L.; Bechhoefer, J. *Rev. Sci. Instrum.* **1993,** *64,* 1868–1873. doi:10.1063/1.1143970
28. Sader, J. E.; Chon, J. W. M.; Mulvaney, P. *Rev. Sci. Instrum.* **1999,** *70,* 3967–3969. doi:10.1063/1.1150021
29. Higgins, M. J.; Proksch, R.; Sader, J. E.; Polcik, M.; Mc Endoo, S.; Cleveland, J. P.; Jarvis, S. P. *Rev. Sci. Instrum.* **2006,** *77,* 013701. doi:10.1063/1.2162455
30. Platz, D.; Tholén, E. A.; Hutter, C.; von Bieren, A. C.; Haviland, D. B. *Ultramicroscopy* **2010,** *110,* 573–577. doi:10.1016/j.ultramic.2010.02.012
31. Hutter, C.; Platz, D.; Tholén, E. A.; Hansson, T. H.; Haviland, D. B. *Phys. Rev. Lett.* **2010,** *104,* 050801. doi:10.1103/PhysRevLett.104.050801
32. Platz, D.; Forchheimer, D.; Tholén, E. A.; Haviland, D. B. *Nanotechnology* **2012,** *23,* 265705. doi:10.1088/0957-4484/23/26/265705
33. Forchheimer, D.; Platz, D.; Tholén, E. A.; Haviland, D. B. *Phys. Rev. B* **2012,** *85,* 195449. doi:10.1103/PhysRevB.85.195449
34. San Paulo, Á.; García, R. *Phys. Rev. B* **2001,** *64,* 193411. doi:10.1103/PhysRevB.64.193411
35. Rodríguez, T. R.; García, R. *Appl. Phys. Lett.* **2002,** *80,* 1646–1648. doi:10.1063/1.1456543
36. Melcher, J.; Hu, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 053101. doi:10.1063/1.2767173
37. Hindmarsh, A. C.; Brown, P. N.; Grant, K. E.; Lee, S. L.; Serban, R.; Shumaker, D. E.; Woodward, C. S. *ACM Trans. Math. Software* **2005,** *31,* 363–396. doi:10.1145/1089014.1089020
38. Giessibl, F. J. *Phys. Rev. B* **1997,** *56,* 16010–16015. doi:10.1103/PhysRevB.56.16010
39. Giessibl, F. J.; Herz, M.; Mannhart, J. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 12006–12010. doi:10.1073/pnas.182160599
40. García, R.; San Paulo, A. *Phys. Rev. B* **1999,** *60,* 4961–4967. doi:10.1103/PhysRevB.60.4961
41. Fontaine, P.; Guenoun, P.; Daillant, J. *Rev. Sci. Instrum.* **1997,** *68,* 4145–4151. doi:10.1063/1.1148359

Beilstein Journal of Nanotechnology

Towards 4-dimensional atomic force spectroscopy using the spectral inversion method

Jeffrey C. Williams and Santiago D. Solares*§

Full Research Paper

Open Access

Address:
Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, USA

Email:
Santiago D. Solares* - ssolares@umd.edu

* Corresponding author
§ Phone: (301) 405-5035, Fax: (301) 314-9477

Keywords:
atomic force microscopy; spectral inversion; spectroscopy; torsional harmonic cantilever; viscoelasticity

Beilstein J. Nanotechnol. **2013,** *4,* 87–93.
doi:10.3762/bjnano.4.10

Received: 07 November 2012
Accepted: 24 January 2013
Published: 07 February 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We introduce a novel and potentially powerful, yet relatively simple extension of the spectral inversion method, which offers the possibility of carrying out 4-dimensional (4D) atomic force spectroscopy. With the extended spectral inversion method it is theoretically possible to measure the tip–sample forces as a function of the three Cartesian coordinates in the scanning volume (x, y and z) and the vertical velocity of the tip, through a single 2-dimensional (2D) surface scan. Although signal-to-noise ratio limitations can currently prevent the accurate experimental implementation of the 4D method, and the extraction of rate-dependent material properties from the force maps is a formidable challenge, the spectral inversion method is a promising approach due to its dynamic nature, robustness, relative simplicity and previous successes.

Introduction

Besides topographical imaging, a popular application of atomic force microscopy (AFM) is the measurement of probe–sample interaction force curves (force spectroscopy), generally based on contact and frequency-modulation methods [1-6]. The procedure is generally time-consuming because the acquisition of the force curve for each (x,y) location on the surface requires that the cantilever approaches and retracts from the sample at relatively low vertical speed, without traveling horizontally. If one wishes to fully characterize a 2D sample, one must slowly acquire force curves throughout the surface, one location at a time. In 2002, Stark et al. [7] introduced an AFM method for performing real-time spectroscopy simultaneously with topographic imaging (that is, while the cantilever images the surface at typical scan speeds), through acquisition and inversion of the spectral response of the tip motion. This approach permits extraction of the tip–sample interaction force as a component of the driving force acting on the cantilever and was demonstrated with standard cantilevers, although the low signal-to-noise ratio

of certain regions of the spectrum limited its accuracy. In 2007, Sahin and co-workers [8] introduced a T-shaped cantilever with an off-centered tip located on one of the arms of the "T", the so-called *torsional harmonic cantilever* (THC), on which the tip–sample interaction generates a torsional oscillation whose amplitude is enhanced by the soft and highly detectable fundamental torsional frequency. Such enhancement provided a more accurate means to implement the method of Stark et al. [7] and the improved technique has been validated experimentally [8-10], studied computationally [11,12] and also commercialized [Harmonix™, Bruker Corporation (formerly Veeco Instruments)].

In the current implementations of the spectral inversion method, the user performs a 2-dimensional (2D) scan of the surface to acquire the topography plus a tip–sample force curve, $f(z)$, at every (x,y) pixel, which effectively results in a 3D description of the tip–sample forces, $f(x,y,z)$. The force curves for each sample location are plotted as 2D graphs depicting the force as a function of the tip–surface separation, as is customary in atomic force spectroscopy (see Figure 1, traditional representation). The purpose of this paper is to introduce an extension of the method, such that the forces can be acquired in 4 dimensions (4D), as a function of the three Cartesian coordinates of the scanning volume, plus the vertical velocity of the tip (see Figure 1, expanded representation). This capability could be useful in the study of samples whose response depends on the rate of application of strains or stresses, such as polymers, composites, soft metals and biological materials, as well as in the study of other rate-dependent phenomena, such as binding or folding/unfolding events in complex biomolecules.

Methods

The details of the spectral inversion method using the torsional harmonic cantilever have been described in detail elsewhere [8,11], so they are discussed only briefly here. In order to implement the mathematics of the force inversion procedure, it is assumed that the fundamental torsional eigenmode follows the dynamics of a damped harmonic oscillator, whose transfer function is

$$T(\omega) = \frac{\omega_T^2 / k_T}{\omega_T^2 - \omega^2 + \mathrm{i}\omega\omega_T / Q_T}, \qquad (1)$$

where ω is the angular frequency, ω_T the torsional resonance angular frequency, k_T the torsional force constant, which has been linearized in the vertical direction, and Q_T is the torsional quality factor. As the cantilever base is oscillated in the vertical direction by the piezo shaker, the fundamental flexural eigenmode is excited such that the tip undergoes intermittent contact with the surface, which in turn excites the torsional eigenmode

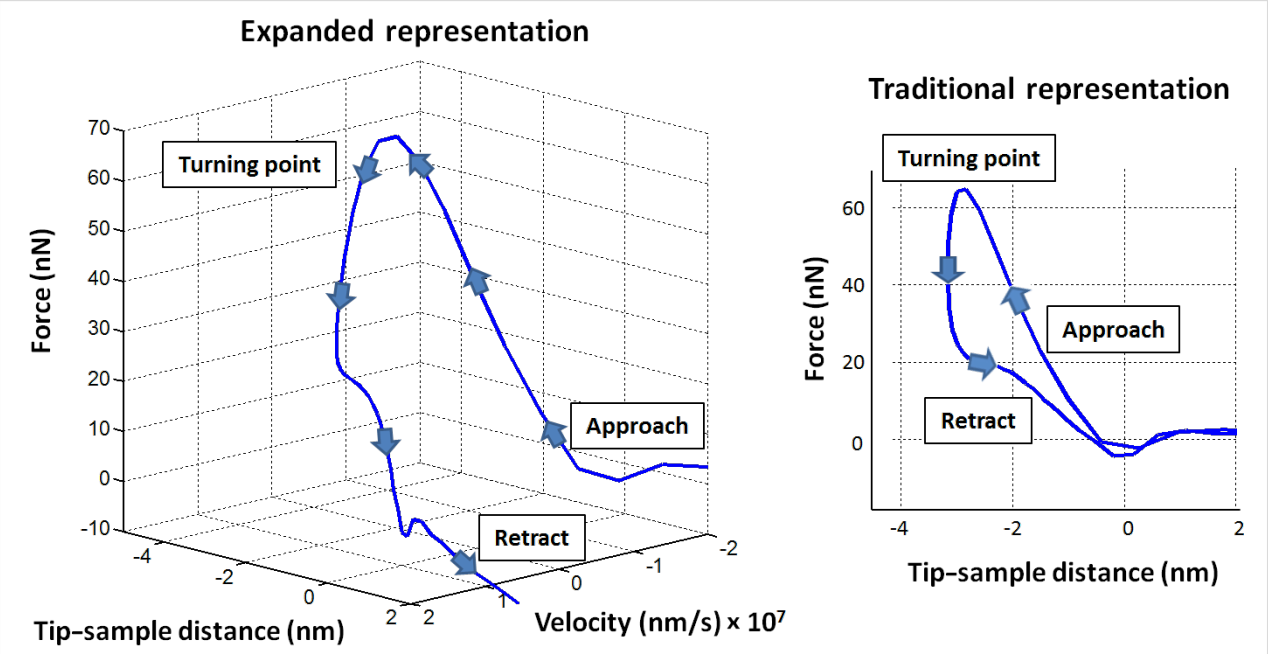

Figure 1: Simulated reconstructed tip–sample interaction force curve for a typical polymer, for a fixed (x,y) point on the surface for which the force is plotted as a function of the vertical tip–sample distance and the velocity of the cantilever tip (expanded representation, left), and only as a function of the vertical tip–sample distance (traditional representation, right). All force curves shown in this paper were constructed through simulation of the spectral inversion method [11]. In all cases the origin of the vertical position (tip–sample distance) axis is located at the relaxed position of the surface. The "turning point" is the lowest vertical position reached by the tip during tip–sample impact.

(recall that the tip is located on one of the arms of the "T" and not on the center of the cantilever end). Since the torsional eigenmode is not *directly* driven by the piezo shaker, the only driving force acting on it is the time-dependent tip–sample interaction, $f_d(t) = F_{ts}[z_c(t) + z_p(t)]$, which generates a torsional response that can be linearized in the z-direction, $z_p(t)$. Here F_{ts} is the tip–sample interaction force, which is a function of the distance between the tip and the sample (tip position). The tip–sample distance, in turn, is determined by the flexural cantilever position, $z_c(t)$, plus the deflection of the torsional "paddle" (arms of the "T") with respect to the flexural position, $z_p(t)$. Invoking the definition of the transfer function, one can write the response of the torsional oscillator in Fourier space as

$$Z_p(\omega) = T_p(\omega)F_d(\omega)\,, \tag{2}$$

where $Z_p(\omega)$ is the spectrum of $z_p(t)$, $F_d(\omega)$ is the spectrum of $f_d(t)$, and $T_p(\omega)$ is the transfer function, defined in Equation 1. Since it is possible to measure the deflection of the torsional oscillator, $z_p(t)$, in real time as the cantilever taps on the sample, one can easily calculate its spectrum, $Z_p(\omega)$, through application of the fast Fourier transform to a sequence of values of $z_p(t)$ recorded at regular intervals. Additionally, one can also obtain the spectrum of the driving force by rewriting Equation 2 as

$$F_d(\omega) = Z_p(\omega) / T_p(\omega)\,. \tag{3}$$

Next, one can apply the inverse fast Fourier transform to $F_d(\omega)$ in order to obtain $f_d(t)$, that is, the time-dependent tip–sample interaction force acting on the cantilever tip. Finally the force curves are obtained by plotting $f_d(t)$ as a function of the vertical tip position, which is generally approximated by the position of the cantilever flexural oscillation, $z_c(t)$ (that is, neglecting $z_p(t)$), under the assumption of negligible torsional oscillation compared to the length scale over which the tip–sample forces vary.

In the ideal application of the spectral inversion method, it would be desirable to acquire $Z_p(\omega)$ for the widest possible frequency range, such that all features of the force curve (e.g., sharp turns at the location of maximum attractive force) can be accurately recovered. However, this is difficult because the spectrum of the torsional eigenmode, $Z_p(\omega)$, follows the general shape of the transfer function, $T_p(\omega)$, which rapidly decays as the frequency deviates from the resonance frequency (recall that harmonic oscillators exhibit a Lorentzian frequency response). As a result, the high-frequency peaks in $Z_p(\omega)$ become smaller and smaller as ω increases, such that their signal-to-noise ratio decreases rapidly as ω increases. Thus, in order to prevent high-frequency noise in the data from being magnified into the recovered force curves through division by very small values of $T_p(\omega)$ (see Equation 3), one applies a cutoff to $Z_p(\omega)$ [7,8]. This cutoff is generally set to include only a few harmonics (often only one) above the torsional resonance frequency.

In order to obtain a 4D representation of the tip–sample forces, it is necessary to also measure the tip velocity in real time, which can be easily recovered from the flexural position, $z_c(t)$ (again, neglecting $z_p(t)$), by using the well-known property of Fourier analysis that states that

$$\dot{z}_c(t) = F^{-1}\left\{i\omega * F\left\{z_c\left(t\right)\right\}\right\}, \tag{4}$$

where the operators $F\{\}$ and $F^{-1}\{\}$ are the Fourier transform and the inverse Fourier transform, respectively. Recovery of the velocity in the Fourier domain allows averaging of multiple oscillation cycles as well as filtering of the data in order to reduce the noise typically seen in the photodetector signal and in order to filter out the effect of flexural–torsional cross-talk [8,11,12]. Upon completion of the scan and post-processing of the data [11], the user will have acquired $f_d(z,\dot{z})$ for every (x,y) pixel on the surface (see Figure 1, expanded representation), which is equivalent to the 4D representation $f_d(x,y,z,\dot{z})$ for the full scan. It is also possible to perform spectroscopy at a fixed (x,y) position on the surface, acquiring successive force curves while varying a particular imaging parameter, such that multiple force curves following different $(z,\dot{z})$ trajectories can be combined in representing the probe–sample interaction as a corrugated plane instead of a single force curve. An example is provided in Figure 2. We point out that although the proposed expanded method provides data in four dimensions, it is not an *unrestricted* 4D force-measurement method. This is because the velocity cannot be varied arbitrarily, but is instead directly related to the tip position and both are governed by the dynamics and properties of the cantilever–sample system. Thus, depending on the system, there will be regions of the $(x,y,z,\dot{z})$ phase space for which the force cannot be measured because not all four coordinates can be varied independently of one another.

The numerical integration methods and equations of motion used to simulate the spectral inversion reconstruction of the force curves presented in this paper are described in detail in [11]. Briefly, the procedure for simulating the acquisition of each individual force curve consists of (i) defining the system parameters (cantilever eigenfrequencies, force constants and quality factors, free flexural amplitude and amplitude setpoint, tip–sample force model, etc.); (ii) numerical implementation of the amplitude-modulation imaging scheme; (iii) recording of the flexural and torsional eigenmode positions as a function of time at regularly spaced intervals (digitally) for several flexural periods; and (iv) application of the inversion procedure

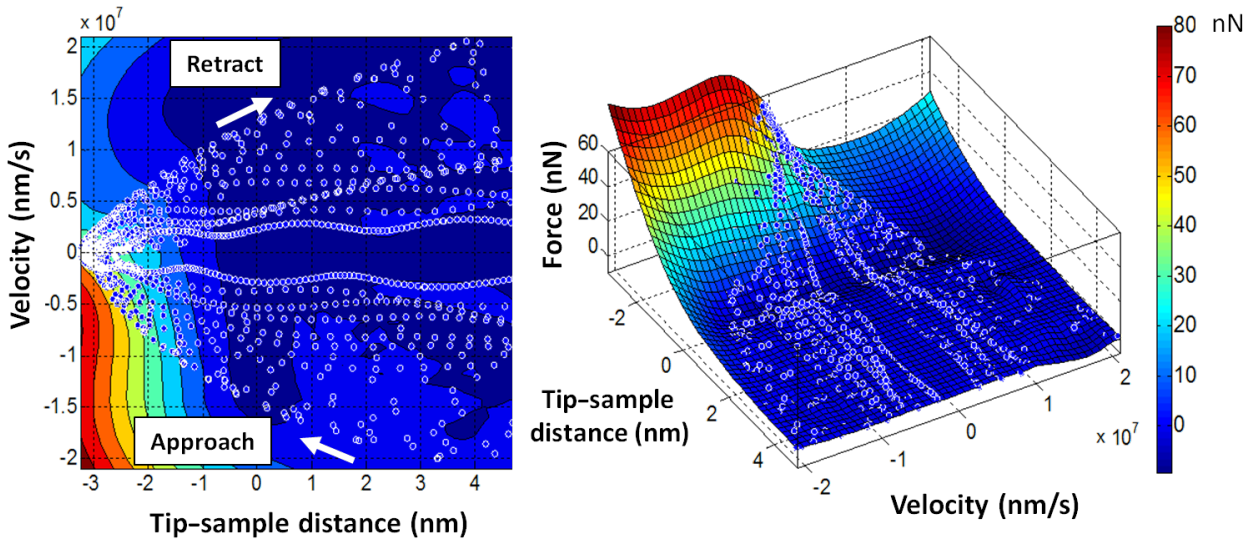

Figure 2: Collection of reconstructed tip–sample interaction force curves for a fixed (x,y) point on the surface, each similar to that shown in Figure 1 (expanded representation), plotted together as a plane. In this simulation, different force curves were acquired for different rest positions of the cantilever above the surface and were all plotted together. The plots show two different views of the same data. The white open circles represent $(z,\dot{z},f_d)$ points and each force curve is a succession of such points that follows the direction indicated by the white arrows. That is, during the approach, the tip travels towards and into the surface (with negative velocity) down to the turning point, and then travels upward (with positive velocity) away from the sample. Different force curves have different incoming and outgoing velocity.

described above to recover the tip–sample forces. The tip–sample interaction was simulated as the combination of attractive van der Waals forces (modeled through the Hamaker equation [13]) plus repulsive and dissipative interactions. In most cases, the repulsive forces were modeled by using a Hertzian contact [13], while the dissipative interactions were modeled by using a viscous force term proportional to the tip speed with a coefficient that decayed exponentially with the tip position [14] (see equations and further details in [11]). We also conducted simulations in which the conservative and dissipative interactions were accounted for through the standard linear solid model (see Figure 3 and the next section), in combination with attractive van der Waals interactions.

Results and Discussion

Characterization of rate-dependent phenomena

The ability to recover rate-dependent signatures of the tip–sample forces presents a unique opportunity to study phenomena such as plasticity, viscoelasticity and biomolecular binding and folding/unfolding. For example, it should be possible to develop methods for fitting experimental data to increasingly elaborate viscoelastic models that go beyond the Kelvin–Voigt model used in the current state of the art in contact-resonance AFM [15,16]. In particular, the Kelvin–Voigt model is not well suited to study stress relaxation. (While this paper is not intended to be a study of surface viscoelasticity, we briefly illustrate the use of slightly more elaborate surface models.) Instead, one could, for example, use the standard linear solid (SLS) model, which is a combination of the Maxwell and Kelvin–Voigt models. In the SLS configuration a Maxwell element is connected in parallel with a second spring (this setup is also known as the Zener model). The SLS approximation provides the simplest form of a linear viscoelastic approximation that can reproduce both stress relaxation and creep compliance, which are observed in the response of real viscoelastic surfaces. Figure 3a provides a schematic of our application of the 4D force mapping concept, modeling the surface as a simple SLS that also experiences van der Waals interactions with the tip. Figure 3b illustrates a "virtual experiment" in which, using the same cantilever, the user acquires multiple force curves at a single pixel by varying the amplitude setpoint. One can also conceive other types of studies in which one varies other parameters, such as the cantilever stiffness, mass, eigenfrequency ratios, etc., keeping all other parameters constant. While not all these studies are experimentally feasible, they can provide significant insight into the evolution of the 4D force representation, individual $f_d(z,\dot{z})$ trajectories, and optimum imaging parameters for different types of systems. Due to the nature of the SLS, one standard feature of the corresponding $f_d(z,\dot{z})$ trajectories, which is evident in Figure 3b, is that the vertical tip position at which the tip impacts the surface is not the same location at which the tip leaves the surface (that is, the surface can remain temporarily indented after being impacted by the tip). In fact, depending on the horizontal scan speed of the cantilever and the number of taps that take place at

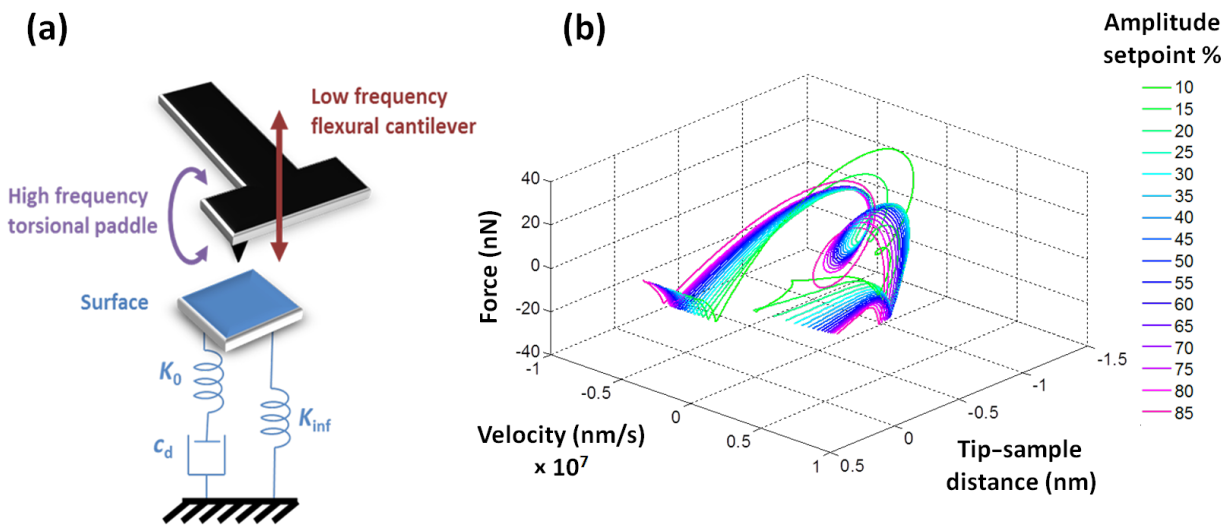

Figure 3: (a) Schematic of a torsional harmonic cantilever interacting with a surface modeled as a standard linear solid (K_o and K_{inf} represent linear springs and c_d represents a linear damper); (b) collection of force curves acquired for different values of the amplitude setpoint by using the same cantilever. Note that the force curves in (b) exhibit a loop, indicating that the tip travels into and towards the surface twice during each flexural cycle resulting in a "double tap". This can occur whenever two eigenmodes with different natural frequencies are active. Similar phenomena take place when imaging samples in high-damping (liquid) environments [18] or in multifrequency AFM characterization [19].

every pixel, the tip may contact the surface at different heights during successive impacts, as illustrated in Figure 4. Due to the simplicity of the SLS model, analytical expressions exist for various aspects of its behavior, such as the rate of recovery of the surface (for example, the surface height recovery from position Z'_1 to position Z_2 between times t_1 and t_2 in Figure 4 can be expressed analytically). As a result, it is possible to fit the three parameters of the model from reconstructed force curves, even without explicitly acquiring velocity information, provided that the force curve inversion is accurate (a detailed analysis of parameter recovery from force curves for the SLS model can be found in [17]).

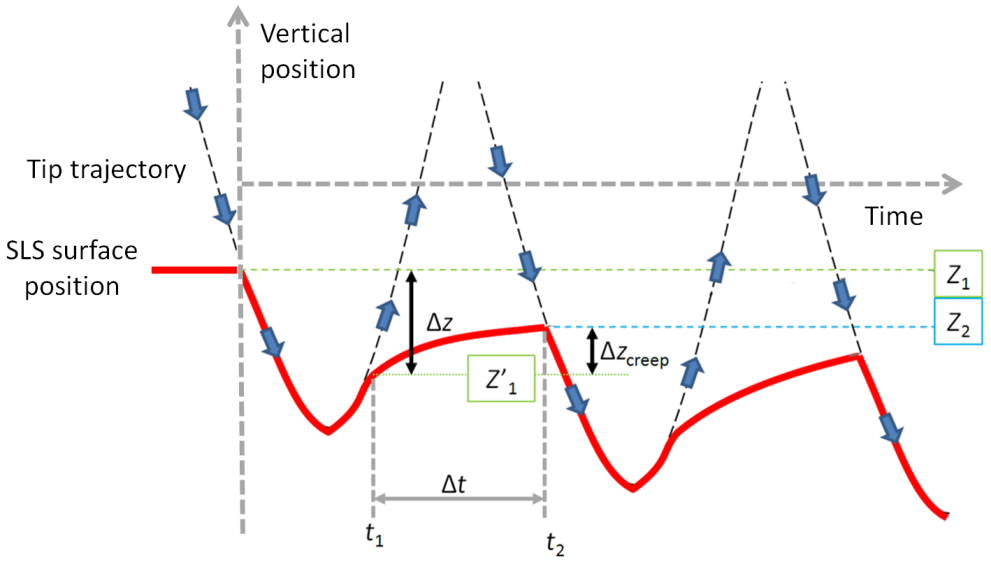

Figure 4: Illustration of the surface depression by the tip–sample impact, and successive recovery within the standard linear solid model. Z_1 is the undisturbed surface position, before any impacts have taken place. Z'_1 is the vertical tip position where tip–sample contact is lost after the first impact (loss of contact occurs because the surface recovery speed is lower than the tip speed). Between the first and second impacts the tip recovers to position Z_2. If the tip continues tapping at the same surface pixel, the impact and loss-of-contact locations asymptotically approach constant values.

Experimental feasibility

The additional post-processing demands required to extend the spectral inversion method from three to four dimensions are relatively minor, since the tip position data is already recorded. Furthermore, all Fourier analysis is carried out during a post-processing step and the calculation of the velocity does not represent an excessive computational burden. Thus, the 4D implementation is relatively straightforward, requiring no new technology. However, important limitations still exist in performing the method upgrade accurately and using it meaningfully, depending on the application. In particular, the study of viscoelastic models in intermittent contact AFM demands very high accuracy in order to reproduce the sharp curvatures and intricate features of the force curves [17]. While the torsional harmonic cantilever leads to an enhanced implementation of the original spectral inversion procedure, which is sufficiently accurate to estimate the effective Young's modulus of soft samples, it does not solve the issues of signal-to-noise ratio for the *higher* frequencies in the spectrum (that is, for frequencies that are appreciably higher than the torsional eigenfrequency). This challenge becomes more significant as the sample stiffness increases (see for example, Figure 7 in [11]) and can compound itself with distortions in the force curve that may emerge in the presence of dissipation, whereby the hysteresis loops in the tip–sample force curve can change shape or shift along the tip-position axis as harmonics are removed from the spectrum (see Figure 5). Even for a simple SLS surface, it is not possible to recover the model parameters unless the impact and loss-of-contact tip positions, both of which are located at sharp minima on the force curve, can be determined accurately (see Figure 3, Figure 4 and [17]). A second important consideration concerning accuracy is that when the torsional oscillation amplitude becomes significant with respect to the length scale over which the tip–sample forces vary, accuracy is lost when the z-position in the force curves is approximated by $z_c(t)$ in place of $z_c(t) + z_p(t)$ (this is also discussed in detail in [11]). The use of the correct expression describing the time-dependent tip position is mathematically straightforward, but is not trivial in an experiment, where there can be signal cross-talk [8,11] and where calibration and noise limitations are present. For all these reasons, it is recommended to apply the method in combination with simulations [11,12], such that the user has a theoretical estimate of the errors incurred in the characterization. Finally, as already implied above, there remain unsolved signal-to-noise ratio challenges even when using the torsional harmonic cantilever, so hardware and cantilever development opportunities still exist, especially in terms of enhancing and expanding the region of the cantilever spectrum that can be recovered with a high signal-to-noise ratio.

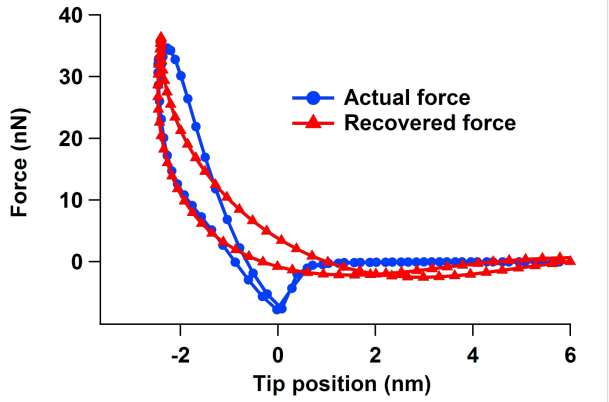

Figure 5: Example of force curve distortion for a case in which both conservative and dissipative interactions are present. The recovered curve uses a frequency cutoff of one harmonic above the torsional resonance frequency, similar to what is customary in experiments, and contains a very shallow and shifted minimum which differs for the approach and retract (compare to Figure 6b in [11], which shows the high-quality force curve recovered by using all available harmonics). The sample modulus of elasticity in this simulation is 2 GPa. The force curve was simulated as the sum of a conservative Hamaker–Hertzian contact with a height-dependent viscous term [14] (this combination of models does not consider surface depression or recovery, so the location of the force minima is the same for the tip approach and retract).

Conclusion

We have presented a simple, yet potentially powerful upgrade to the spectral inversion method of atomic force microscopy, which makes possible the mapping of the tip–sample interaction forces in four dimensions (the three Cartesian coordinates in the scanning volume plus the vertical velocity of the probe). We have also highlighted important limitations that still exist in the accurate experimental implementation of this procedure. Despite the unsolved challenges, the proposed approach could, in combination with future instrumentation and cantilever upgrades, enable studies in which rate-dependent phenomena, such as viscoelasticity and plasticity, are characterized in real time by using tapping-mode atomic force microscopy.

Acknowledgements

J.C.W. gratefully acknowledges support from the U.S. National Science Foundation through the Louis Stokes Alliances for Minority Participation, Bridge to the Doctorate Program.

References

1. Albrecht, T. R.; Grütter, P.; Horne, D.; Rugar, D. *J. Appl. Phys.* **1991,** *69,* 668–673. doi:10.1063/1.347347
2. Burnham, N. A.; Colton, R. J. *J. Vac. Sci. Technol., A* **1989,** *7,* 2906–2913. doi:10.1116/1.576168
3. Hölscher, H.; Langkat, S. M.; Schwarz, A.; Wiesendanger, R. *Appl. Phys. Lett.* **2002,** *81,* 4428–4430. doi:10.1063/1.1525056
4. Hölscher, H.; Gotsmann, B.; Allers, W.; Schwarz, U. D.; Fuchs, H.; Wiesendanger, R. *Phys. Rev. B* **2001,** *64,* 075402. doi:10.1103/PhysRevB.64.075402

5. Albers, B. J.; Schwendemann, T. C.; Baykara, M. Z.; Pilet, N.; Liebmann, M.; Altman, E. I.; Schwarz, U. D. *Nat. Nanotechnol.* **2009,** *4,* 307–310. doi:10.1038/nnano.2009.57
6. Fukuma, T.; Ueda, Y.; Yoshioka, S.; Asakawa, H. *Phys. Rev. Lett.* **2010,** *104,* 016101. doi:10.1103/PhysRevLett.104.016101
7. Stark, M.; Stark, R. W.; Heckl, W. M.; Guckenberger, R. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 8473–8478. doi:10.1073/pnas.122040599
8. Sahin, O.; Magonov, S.; Su, C.; Quate, C. F.; Solgaard, O. *Nat. Nanotechnol.* **2007,** *2,* 507–514. doi:10.1038/nnano.2007.226
9. Dong, M.; Sahin, O. *Nat. Commun.* **2011,** *2,* 247. doi:10.1038/ncomms1246
10. Husale, S.; Persson, H. H. J.; Sahin, O. *Nature* **2009,** *462,* 1075–1078. doi:10.1038/nature08626
11. Solares, S. D.; Hölscher, H. *Nanotechnology* **2010,** *21,* 075702. doi:10.1088/0957-4484/21/7/075702
12. Solares, S. D.; Hölscher, H. *J. Vac. Sci. Technol., B* **2010,** *28,* C4E1. doi:10.1116/1.3305338
13. García, R.; Pérez, R. *Surf. Sci. Rep.* **2002,** *47,* 197–301. doi:10.1016/S0167-5729(02)00077-8
14. Gotsmann, B.; Seidel, C.; Anczykowski, B.; Fuchs, H. *Phys. Rev. B* **1999,** *60,* 11051–11061. doi:10.1103/PhysRevB.60.11051
15. Yuya, P. A.; Hurley, D. C.; Turner, J. A. *J. Appl. Phys.* **2008,** *104,* 074916. doi:10.1063/1.2996259
16. Yuya, P. A.; Hurley, D. C.; Turner, J. A. *J. Appl. Phys.* **2011,** *109,* 113528. doi:10.1063/1.3592966
17. Williams, J. C. Surface characterization of viscoelastic materials through spectral intermittent contact atomic force microscopy. M.S. Thesis, University of Maryland, College Park, USA, 2012.
18. Melcher, J.; Xu, X.; Raman, A. *Appl. Phys. Lett.* **2008,** *93,* 093111. doi:10.1063/1.2976438
19. Solares, S. D.; Chawla, G. *Meas. Sci. Technol.* **2010,** *21,* 125502. doi:10.1088/0957-0233/21/12/125502

License and Terms

Beilstein Journal of Nanotechnology

High-resolution dynamic atomic force microscopy in liquids with different feedback architectures

John Melcher*[1], David Martínez-Martín[2], Miriam Jaafar[3], Julio Gómez-Herrero[3] and Arvind Raman[4]

Full Research Paper

Open Access

Address:
[1]Department of Engineering Mathematics, University of Bristol, Bristol BS8 1TR, United Kingdom, [2]ETH Zürich, Department of Biosystems Science and Engineering, CH-4058 Basel, Switzerland, [3]Departamento de Física de la Materia Condensada C-III, Universidad Autónoma de Madrid, 28049 Madrid, Spain and [4]School of Mechanical Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907

Email:
John Melcher* - j.melcher@bris.ac.uk

* Corresponding author

Keywords:
atomic force microscopy; dAFM; high-resolution; liquids

Beilstein J. Nanotechnol. **2013,** *4,* 153–163.
doi:10.3762/bjnano.4.15

Received: 03 December 2012
Accepted: 24 January 2013
Published: 27 February 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

The recent achievement of atomic resolution with dynamic atomic force microscopy (dAFM) [Fukuma et al., *Appl. Phys. Lett.* **2005**, *87*, 034101], where quality factors of the oscillating probe are inherently low, challenges some accepted beliefs concerning sensitivity and resolution in dAFM imaging modes. Through analysis and experiment we study the performance metrics for high-resolution imaging with dAFM in liquid media with amplitude modulation (AM), frequency modulation (FM) and drive-amplitude modulation (DAM) imaging modes. We find that while the quality factors of dAFM probes may deviate by several orders of magnitude between vacuum and liquid media, their sensitivity to tip–sample forces can be remarkable similar. Furthermore, the reduction in noncontact forces and quality factors in liquids diminishes the role of feedback control in achieving high-resolution images. The theoretical findings are supported by atomic-resolution images of mica in water acquired with AM, FM and DAM under similar operating conditions.

Introduction

Since its inception [1], dynamic atomic force microscopy (dAFM) has proven to be a powerful yet versatile tool capable of operating in media ranging from vacuum to liquids and interrogating samples ranging from stiff inorganic materials to soft biological samples, with nanoscale resolution. Recently, the achievement of atomic-resolution imaging in liquids [2-6] has challenged the accepted belief that high quality factors, which are a hallmark of microcantilever probes in vacuum, are necessary for atomic-resolution imaging [7]. However, atomic-resolution images have now been obtained with several dAFM imaging modes in liquids despite the quality factors being several orders of magnitude smaller than in vacuum.

Several prior works have been dedicated to the understanding of imaging resolution and the role of feedback control in dAFM. Prior efforts to analyze imaging resolution in dAFM have typically focused on the small-amplitude limit in order to establish a relationship between various noise sources in the experimental setup and the minimum detectable gradient of the tip–sample force [1,4,8,9]. However, the optimal imaging amplitude in FM has also been considered [10]. The role of feedback control in dAFM and its stability have been studied largely by using numerical simulations to solve complex systems of nonlinear, integro-differential equations governing the deflection of the oscillating probe subject to feedback control [11-13]. Kilpatrick et al. [14] neglected tip–sample forces in order to provide an estimate for stable control parameters in FM. To improve imaging resolution in liquids, Q-controlled dAFM, which uses feedback control to manipulate the effective quality factor of the oscillating probe, has been proposed [15,16]. However, the merits of this approach for improving imaging resolution are still under question [17].

In this article we present a combined theoretical and experimental study of high-resolution imaging in liquid media with various dAFM imaging modes. The method of periodic averaging [18] is used to simplify the fast-time-scale equations governing the deflection of the oscillating probe by slow-time-scale, averaged equations that govern the amplitude and phase lag of the oscillation. The averaged equations provide a natural starting point for the analysis of closed-loop dAFM imaging modes, which are ultimately designed to regulate the amplitude and phase lag of the oscillating probe rather than its time-varying deflection. From the approximate theory, we explore performance metrics for dAFM imaging modes, such as (i) force sensitivity and resolution, (ii) detection bandwidth, (iii) disturbance mitigation and (iv) imaging stability. In support of our findings, we demonstrate atomic-resolution images of mica in water with FM, AM and DAM under similar operating conditions.

Analysis of closed-loop dAFM imaging modes

Conventional dAFM imaging modes use a microcantilever probe with a sharp tip affixed to the free-end, which is made to oscillate near its fundamental resonance in close proximity to a sample. Through the influence of tip–sample forces, the presence of the sample is detected in the oscillations of the probe. Let z denote the nominal separation between the tip and sample in the absence of tip–sample forces. The realization of a dAFM imaging mode follows from the implementation of a separation regulator that uses z as a controlled input for a feedback regulator designed to maintain the amplitude and/or phase lag of the oscillation. The actuation of z is implemented by a piezo actuator. The values of z that satisfy the regulation objective are interpreted as the topography of the sample. The simplest dAFM imaging mode is AM, where the amplitude is maintained by the separation regulator, while the phase lag of the oscillation is free to vary. Imaging modes with more complex feedback architectures, such as FM and DAM, will be described later in this section.

In order to establish the performance metrics for high-resolution imaging, we start with the equation of motion describing the time-varying deflection $x(t)$ of the probe tip in the presence of tip–sample forces given by

$$\ddot{x}+\frac{\omega_0}{Q_0}\dot{x}+\omega_0^2 x=\frac{\omega_0^2}{k}\left[F\cos\omega t+F_{\rm ts}(d,\dot{d})\right], \qquad (1)$$

where ω_0, Q_0 and k are the unperturbed natural frequency, quality factor and stiffness of the probe, respectively, and F is the excitation force [19]. $F_{\rm ts}$ is the tip–sample interaction force, which depends explicitly on the tip–sample gap $d(t) = z + x(t)$ and its rate $\dot{d}$. In the absence of tip–sample forces with $\omega = \omega_0$ and $F = F_0$, the tip oscillates with an unconstrained amplitude $a_0 = F_0Q_0/k$. $F = F_0$ and $\omega = \omega_0$ are fixed in AM, but F and ω are adjusted by feedback regulators in FM and DAM.

The solution of Equation 1 can separated into two parts, the first being the equilibrium deflection x^* in the absence of the excitation force, and, the second being oscillation about x^*. At each z, x^* is found by setting $\ddot{x} = \dot{x} = 0$ in Equation 1 and satisfies

$$kx^* = F_{\rm ts}\left(z+x^*,0\right). \qquad (2)$$

The tip–sample forces are often characterized by an attractive ($\partial F_{\rm ts}/\partial d > 0$), noncontact regime when d is sufficiently large, which gives way to a repulsive ($\partial F_{\rm ts}/\partial d < 0$), contact regime as d is reduced. If $k < \partial F_{\rm ts}/\partial d$, then for some z, x^* will be bistable. This results in one stable equilibrium for both the noncontact and contact regimes. In this case, a spontaneous transition from the noncontact equilibrium to the contact equilibrium, or *snap-in*, can occur [20]. The snap-in instability is avoided if the equilibrium deflection is monostable, which occurs when either z is sufficiently large for a given k or when k exceeds the maximum gradient of $F_{\rm ts}$ [10,21].

The model for the probe dynamics in Equation 1 can be simplified through the use of the method of first-order averaging [22,23]. To this end, consider the overall motion with excitation to be represented by $x(t) = x^* + a(t)\cos[\omega t - \phi(t)]$, where

$a(t)$ and $\phi(t)$ are the time-varying amplitude and phase lag, respectively [24]. An autonomous equation describing the dynamics of a and ϕ becomes:

$$\begin{aligned}\frac{da}{d\tau} &= \frac{F}{F_0}a_0\sin\phi - a\left[1+e_{ts}(z,a)\right],\\ \frac{d\phi}{d\tau} &= \frac{F}{F_0}\frac{a_0}{a}\cos\phi + \sigma + v_{ts}(z,a),\end{aligned} \tag{3}$$

where $\tau = \omega_0 t/2Q$ and $\sigma = Q_0[(\omega/\omega_0)^2 - 1]$ is the frequency shift scaled by the half-power bandwidth of the resonance. The nonlinear tip–sample forces are captured in Equation 3 by the functionals

$$e_{ts}\left(z,a,x^*\right) = \frac{Q_0}{\pi k a}\int_0^{2\pi} F_{ts}\left(d(\theta),d'(\theta)\right)\sin\theta\,d\theta, \tag{4}$$

$$v_{ts}\left(z,a,x^*\right) = \frac{Q_0}{\pi k a}\int_0^{2\pi} F_{ts}\left(d(\theta),d'(\theta)\right)\cos\theta\,d\theta, \tag{5}$$

where $d(\theta) = z + x^* + a\cos\theta$ and $d'(\theta) = -a\sin\theta$. e_{ts} is the energy dissipated during the tip–sample interaction and v_{ts} is the virial of the tip–sample interaction [25]. The virial is related to the kinetic energy stored in the oscillating probe through the virial theorem [26] and is a measure of the maximum potential energy stored in the tip–sample interaction during an oscillation. Moreover, by introducing a specific model for F_{ts}, a relationship between v_{ts} and the interaction potential can be established [27]. Finally, both e_{ts} and v_{ts} have been nondimensionalized by the energy dissipated by the media during an oscillation cycle $E_{med} = \pi k a^2/Q_0$. In one form or another, these parameters are ubiquitous in perturbation analysis of dAFM [22,23,26-28].

Equation 3 captures the transient response $a(\tau)$ and $\phi(\tau)$ of the oscillating probe. In addition to providing an approximate relationship between the experimental observables and the tip–sample forces, this feature accommodates the study of stability and detection bandwidth. The transient response of both the amplitude and phase lag have a nominal characteristic time scale of $2Q_0/\omega_0$ in the absence of tip–sample forces or feedback control. The equilibrium solutions a^* and ϕ^* of Equation 3 approximate steady-state, harmonic-oscillation solutions to Equation 1 with constant amplitude and phase lag, oscillating about the equilibrium deflection x^*. Note that we have included the dependence on x^* in Equation 4 and Equation 5 for completeness. However, unless the equilibrium deflection is bistable, i.e., near snap-in, x^* can be neglected in the analysis. This assumption is carried forward, and, in the subsequent analysis, we write $v_{ts}(z,a)$ and $e_{ts}(z,a)$.

Setting $da/d\tau = d\phi/d\tau = 0$ in Equation 3, we arrive at the equations governing the steady-state amplitude and phase lag

$$\begin{aligned}\frac{a^*}{a_0} &= \frac{F}{F_0}\left\{\left[\sigma + v_{ts}\left(z,a^*\right)\right]^2 + \left[1+e_{ts}\left(z,a^*\right)\right]^2\right\}^{-\frac{1}{2}},\\ \phi^* &= -\tan^{-1}\left(\frac{1+e_{ts}\left(z,a^*\right)}{\sigma + v_{ts}\left(z,a^*\right)}\right),\end{aligned} \tag{6}$$

where the asterisk denotes the steady-state. The stability of the steady-state oscillation is determined by Equation 3.

The nonlinear terms v_{ts} and e_{ts} introduce the possibility of coexisting stable solutions to Equation 6, even when the equilibrium deflection is monostable [21,29,30]. When the equilibrium is bistable, three coexisting stable oscillation states are possible [31]. Such nonlinear phenomena are of considerable practical importance in dAFM and have been studied extensively in the literature [32-34].

Next, we introduce the feedback architectures that define the AM, FM and DAM imaging modes. AM is modeled with Equation 3 by setting $F = F_0$, $\sigma = 0$, and introducing a separation regulator that manipulates z:

$$z = -K_1 a - K_2 w; \quad \frac{dw}{d\tau} = a - a_{sp}, \tag{7}$$

where K_1 and K_2 are gain parameters. The control effort in Equation 7 consists of a proportional controller $K_1 a$ and an integral regulator $K_2 w$, which ensures $a^* = a_{sp}$. By substituting $F = F_0$, $\sigma = 0$ and $a^* = a_{sp}$ into Equation 6, it can be shown that the AM topography reflects a combination of the e_{ts} and v_{ts} and the resulting phase lag reveals the relative magnitude of the two. However, it is important that the issue of co-existing oscillation states persists in AM allowing the controller to spontaneously switch between stable states [30].

The FM and DAM imaging modes have more complex feedback architectures than AM. The original implementation of FM conceived by Albrecht et al. [8] used a self-excitation scheme where the excitation signal was generated by applying a phase shift to the deflection signal and the oscillation amplitude was maintained by a regulator. Alternatively, an externally generated, excitation signal and lock-in amplifier can be implemented to maintain a resonant excitation with constant oscillation amplitude [14]. The latter will be the focus of the present analysis. Both FM and DAM incorporate these auxiliary regulators that operate independently from the separation regulator:

$$F = -K_1 a - K_2 w_1 \; ; \frac{dw_1}{d\tau} = a - a_{sp} \; ,$$
$$\sigma = -K_3 \phi - K_4 w_2 \; ; \frac{dw_2}{d\tau} = \phi - \phi_{sp} \; , \tag{8}$$

where $K_1 - K_4$ are gain parameters, $a_{sp} = a_0$ and $\phi_{sp} = \pi/2$. The auxiliary regulators ensure that at steady-state, e_{ts} is captured in F and v_{ts} is captured in σ. While both the self-excited and externally excited implementations achieve the same objective, there are some differences in the detection bandwidth and measurement noise. These issues will be discussed briefly in the following section.

The auxiliary feedback regulators essentially have complete control over the oscillations of the probe through their independent manipulation of amplitude and phase lag. When equilibrium deflection x^* is monostable, coexisting of stable oscillation states, a^* and ϕ^* are eliminated by the integral regulators, while the stability and transient settling time can be controlled completely by the proportional controllers. It is straightforward to prove these results by substituting Equation 8 into Equation 3 and solving for the equilibrium points and the eigenvalues of the Jacobian matrix or by using standard tools for control theory [18]. Limitations in the control of the amplitude and phase lag are introduced only after incorporating the finite bandwidth of the amplitude and phase lag measurements into the model [14]. However, we note that instabilities persist when the equilibrium deflection x^* is bistable since the auxiliary regulators control the amplitude and phase lag but have no control over x^*.

The separation regulator in FM actuates z in order to maintain the frequency shift σ according to

$$z = -K_5 \sigma - K_6 w_3; \quad \frac{dw_3}{d\tau} = \sigma - \sigma_{sp}, \tag{9}$$

where K_5 and K_6 are gain constants, and σ_{sp} is the set-point frequency shift. At equilibrium in FM, the topography is purely a reflection of the virial of the interaction and the dissipation is measured in the corresponding excitation force signal.

The separation regulator in DAM actuates z in order to maintain the excitation force according to

$$z = -K_5 F - K_6 w_3; \quad \frac{dw_3}{d\tau} = F - F_{sp}, \tag{10}$$

where K_5 and K_6 are gain parameters and F_{sp} is the force at the set-point. At equilibrium in DAM, the topography is purely a reflection of the dissipation, and the virial is captured by the corresponding frequency shift. In this respect, DAM can be regarded as the complementary mode to FM.

At this juncture, it is instructive to introduce some experimental data highlighting some of the key differences between dAFM operation in vacuum, air and liquid. In Figure 1, e_{ts} and v_{ts} are measured under typical operating conditions in vacuum, air and liquid with an oscillating probe that is controlled by the auxiliary feedback regulators in Equation 8 while z is displaced by a piezo actuator (see Methods for additional information). The coordinate z is shifted such that $z = 0$ is located approximately at the boundary between the contact and noncontact regimes. For high-resolution imaging, z is maintained in the neighborhood of $z = 0$. In vacuum, large long-range noncontact tip–sample forces result in $e_{ts} >> 1$ and $|v_{ts}| >> 1$ at imaging distances from the sample. Consequently, the oscillations of the probe are strongly influenced by the presence of tip–sample forces in vacuum. On the other hand, e_{ts} and v_{ts} are on the order of unity in air and small compared to unity in liquid.

Before proceeding, we will briefly address the issue of higher harmonics in liquids. Early work on dAFM in liquids showed

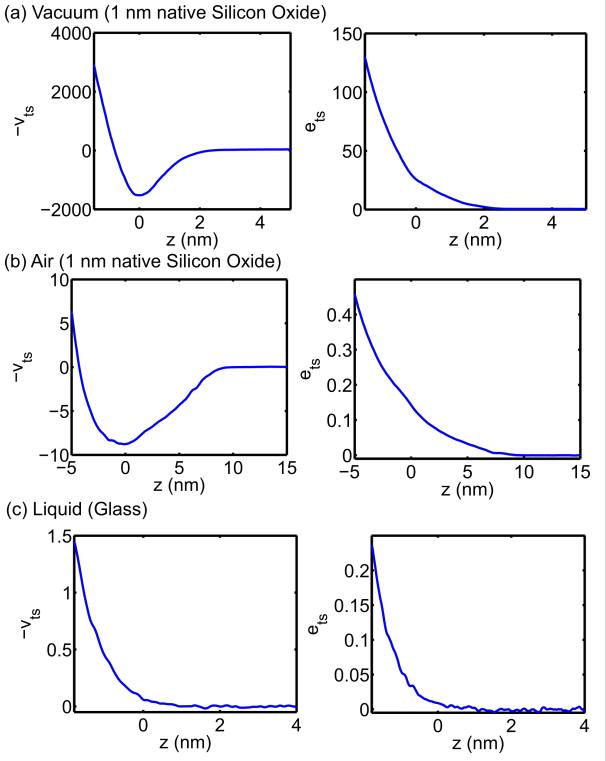

Figure 1: Experimental tip–sample virial $v_{ts}(z,a_{sp})$ and dissipation $e_{ts}(z,a_{sp})$. Data acquired in vacuum on a silicon sample with 1 nm native silicon oxide (a) in vacuum and (b) in air. (c) Data acquired in deionized water on a glass substrate. The set-point amplitudes were 8.5 nm, 8.6 nm, and 1.5 nm, respectively. See the Methods section for additional information.

that significant higher harmonic distortions in the oscillation waveform could provide additional channels for compositional mapping [35]. More recently, it was discovered that the use of soft microcantilevers (≤1 N/m) with quality factors close to unity resulted in higher harmonics from higher eigenmodes [36,37]. The present theory does not extend to soft microcantilevers in liquids. However, from prior work, we can expect that the primary difference for soft microcantilevers is that the dissipation reflects the energy lost to higher harmonics [38,39].

Performance metrics for high-resolution imaging in dAFM

Using the mathematical framework developed in the previous section based on the method of first-order averaging, we now address the question of high-resolution imaging in liquid, despite the low quality factors. We note that while the chemical makeup and atomic configuration of the tip and sample are important considerations for high-resolution imaging, the focus of this article is on the dAFM instrumentation. Specifically, we investigate the performance metrics for high-resolution imaging in dAFM, including (i) force sensitivity and resolution, (ii) detection bandwidth, (iii) disturbance mitigation and (iv) stability in dAFM modes.

Force sensitivity and resolution

To understand how atomic-resolution imaging is possible in liquids despite the low quality factors, we first examine the sensitivity of the oscillating probe to tip–sample forces. For high-resolution imaging in all dAFM imaging modes, the effect of small tip–sample forces between the foremost atom of the tip and the substrate must be detected in the steady-state amplitude and/or phase lag. Therefore, we are interested in the sensitivity of the steady-state amplitude and phase lag to a small perturbation to the total tip–sample force.

The tip–sample forces are limited in magnitude for a given length scale. These considerations are captured elegantly by an exponential function given by [10]

$$F_{\mathrm{ts}}(d) = F_{ts0}e^{-d/\lambda},\ d \geq 0\,, \tag{11}$$

where λ is the characteristic length scale and F_{ts0} is the magnitude corresponding to $d = 0$. The magnitude of F_{ts} is limited by requiring $d \geq 0$. Following [10], an approximate expression for the v_{ts}, which holds for an arbitrary amplitude, is given by

$$v_{\mathrm{ts}}(z,a) = -\frac{Q_0 F_{\mathrm{ts}0} e^{(a-z)/\lambda}}{k\lambda\left[1+\sqrt{\frac{\pi}{2}}\left(\frac{a}{\lambda}\right)^{3/2}\right]}\,. \tag{12}$$

Choosing λ appropriately allows Equation 11 and Equation 12 to approximate a variety of tip–sample forces. Forces at the atomic scale are captured by λ ≈ 1 Å.

The simple model for the tip–sample force in Equation 11 is conservative. Dissipative components of the interaction are more complex in nature and less understood. At the atomic scale, energy may be dissipated from the bulk motion of the tip due to spontaneous transitions between multistable configurations of the nearest atoms of the tip and sample [40,41]. To incorporate dissipation in a simplistic manner, we allow $e_{\mathrm{ts}} = -\mu v_{\mathrm{ts}}$, where μ is a proportionality constant. From the data in Figure 1, this appears to be a reasonable approximation when the force is unidirectional, which is the case for Equation 11.

Next, we perturb the steady-state amplitude and phase lag about their equilibrium by introducing a small variation in the tip–sample force. This is achieved by perturbing F_{ts0} by a small amount δF_{ts}. Using a first-order Taylor approximation of Equation 6 establishes a relationship between the amplitude and phase lag resolution, δa and $\delta\phi = \delta a/a$, respectively, and the force resolution, given by

$$\delta F_{\mathrm{ts}}^{(a)} = S^{-1}\left[\frac{\left(\frac{|\sigma+v_{\mathrm{ts}}|}{1+e_{\mathrm{ts}}}\right)+\mu}{\left(\frac{\sigma+v_{\mathrm{ts}}}{1+e_{\mathrm{ts}}}\right)^2+1}\right]^{-1}\delta a\,, \tag{13}$$

$$\delta F_{\mathrm{ts}}^{(\phi)} = S^{-1}\left[\frac{1+\left(\frac{|\sigma+v_{\mathrm{ts}}|}{1+e_{\mathrm{ts}}}\right)\mu}{\left(\frac{\sigma+v_{\mathrm{ts}}}{1+e_{\mathrm{ts}}}\right)^2+1}\right]^{-1}\delta a\,, \tag{14}$$

where

$$S = \frac{Q}{k}\frac{\left(\frac{a^*}{\lambda}\right)}{1+\sqrt{\frac{\pi}{2}}\left(\frac{a^*}{\lambda}\right)^{\frac{3}{2}}}\,, \tag{15}$$

and $\delta F_{\mathrm{ts}}^{(a)}$ and $\delta F_{\mathrm{ts}}^{(\phi)}$ are the force resolutions in the amplitude and phase-lag measurements, respectively, and $Q = Q_0/(1 + e_{\mathrm{ts}})$ is the effective quality factor.

Equation 13 and Equation 14 can be combined into a single approximation for force resolution in dAFM with the following approximation. Note that conventional dAFM modes are designed to excite the probe near its effective resonance frequency in the presence of the tip–sample forces. Thus, we argue that the term $(\sigma + \nu_{ts})/(1 + e_{ts})$ should be on the order of unity or smaller. It follows that the bracketed terms in Equation 13 and Equation 14 are on the order of unity allowing S to approximate the sensitivity of both the amplitude and phase-lag measurements to tip–sample forces. The force resolution can be approximated simply by $\delta F_{ts} = S^{-1}\delta a$.

Figure 2 plots the normalized force sensitivity $S \times k/Q$ versus the normalized amplitude a^*/λ. We note that much of the prior work of sensitivity and resolution in dAFM has linearized the tip–sample force to determine the minimum detectable force gradient [8]. Such analyses predict that force resolution in dAFM improves as the amplitude is increased. However, Equation 15, which holds for an arbitrary amplitude, predicts a global maximum in the force sensitivity for $a^* \approx \lambda$ (see Figure 2). This result is consistent with the analysis of FM by Giessibl et al. [10] but applies to all conventional dAFM modes.

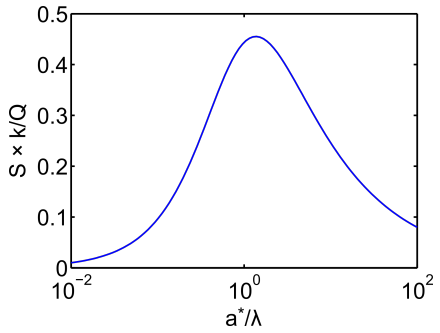

Figure 2: Plot of the normalized force sensitivity $S \times k/Q$ versus normalized amplitude a^*/λ, where λ is the characteristic length scale of the interaction. The maximum of S occurs at $a^* \approx \lambda$.

To estimate the force resolution we must also obtain some estimate of the resolution of the amplitude measurement δa. Two important noise sources that contribute to δa are the thermal noise and deflection-sensor noise. In light of recent efforts to reduce the deflection-sensor noise [42], we will focus on the thermodynamic lower limit of δa, which can be approximated by [43]

$$\delta a^2 = \frac{2k_B T Q_0 B}{\pi k \omega_0}, \quad B < \frac{\pi\omega_0}{2Q} \tag{16}$$

where B is the measurement bandwidth, k_B is the Boltzmann constant and T is the temperature. Setting $B = \pi\omega_0/2Q_0$ extends the approximation in Equation 16 to large measurement bandwidths. However, the bandwidth restriction does not apply to self-excited FM [8,44]. Allowing $B >> \omega_0/Q$ suggests that self-excited FM has the potential to be noisier than externally excited FM; however, this is a topic of ongoing debate [45,46].

Using Equation 16 to approximate δa yields to the following expression for δF_{ts}:

$$\delta F_{ts} = \frac{1+\sqrt{\frac{\pi}{2}}\left(\frac{a^*}{\lambda}\right)^{\frac{3}{2}}}{\left(\frac{a^*}{\lambda}\right)} \sqrt{\frac{2kk_B TB}{\pi\omega_0 Q_0}}, \quad B < \frac{\pi\omega_0}{2Q_0}. \tag{17}$$

In the case of small amplitudes $a^* << \lambda$, substitution of

$$S \approx \frac{Q}{k}\left(\frac{a^*}{\lambda}\right)$$

and $\delta F_{ts} \approx \lambda \delta k_{ts}$ Equation 17, where k_{ts} is the tip–sample gradient, yields

$$\delta k_{ts} = \sqrt{\frac{2kk_B TB}{\pi\omega_0 Q_0 \left(a^*\right)^2}}, \quad a^* << \lambda, \tag{18}$$

which is essentially the result obtained by Albretch et al. [8]. However, of the two expressions, only Equation 17 captures the range of amplitudes applicable to high-resolution imaging.

The present analysis of the force sensitivity in dAFM begins to shed light on how high-resolution imaging is possible with dAFM in liquids despite the low quality factors. Table 1 lists the force sensitivity S and resolution δF_{ts} for several prior works demonstrating atomic resolution in vacuum and liquid environments. While Q_0 degrades by four orders of magnitude in liquid, the force resolution, in some cases, is of the same order of magnitude. Moreover, the force resolution approximated from Giessibl et al. [47] is in accordance with all the measurements in liquids. This result is the direct consequence of the small attractive forces in liquids on clean, hard surfaces, such as mica, that allow $a_{sp} \approx \lambda$ with a probe that is much softer than the tuning fork [2]. In vacuum, strong attractive forces cause the tip to snap into the surface when oscillation amplitudes are small.

From the data in Table 1, we must also entertain the possibility that force resolution is not necessarily the limiting factor for

Table 1: Parameters in atomic-resolution imaging in vacuum and liquid medium. S (Equation 15) and $\delta F_{ts} = S^{-1}\delta a$ (Equation 17) where calculated assuming λ = 1 Å, T = 300 K and B = 1 kHz, and $Q = Q_0$ and $a^* = a_{sp}$.

Mode	Med.	k (N/m)	a^* (nm)	Q_0	B_{osc} (Hz)	$\log_{10}S$	$\log_{10}\delta F_{ts}$	Ref.
FM	Vac.	17	34	28,000	2	1.9	−11.4	[49]
FM	Vac.	41	14.8	38,000	2	1.8	−11.6	[50]
AM	Vac.	60	0.2	550	15	0.61	−11.5	[51]
AM	Vac.	1600	0.28	18,000	50	0.58	−12.2	[52]
FM	Vac.	1800	0.8	4000[a]	15	−0.22	−10.8	[47]
FM	Liq.	37	0.33	23	3000	−0.62	−10.7	[2]
PM[b]	Liq.	19	0.59	5.8	12,000	−1.0	−10.5	[3]
AM	Liq.	0.76	0.5	2[c]	1300	−0.057	−10.3	[5]
FM	Liq.	30	0.59	8	8100	−1.1	−10.4	[4]
FM	Liq.	26	0.11	8.3	8400	−0.84	−10.6	[6]

[a]taken from [48]. [b]phase modulation. [c]taken from our own data.

imaging resolution. The highest resolution images are achieved with the qPlus sensor in [47], which has the lowest force resolution amongst the references in vacuum. If the force sensitivity meets some minimal requirements, the imaging resolution may be limited by other factors, such as the imaging stability [48]. Furthermore, the minimal requirement for force resolution may be less in liquids compared to vacuum where stable images can be acquired in the contact regime.

Detection bandwidth

The overall detection bandwidth in dAFM can be limited by the bandwidth of the amplitude and phase measurements, or the transient response, or the response of the oscillating probe. Drift in dAFM, for example arising from the piezo actuators controlling the image raster, imposes a minimum scan speed and corresponding detection bandwidth requirement for high-resolution imaging. Giessibl et al. [10] approximate the required detection bandwidth as 1 kHz. From our own experiments (see Methods section), we also estimate that the required bandwidth is on the order of 1 kHz for high-resolution imaging; however, we can expect variability depending on the experimental setup.

Detection of tip–sample forces in dAFM requires that the oscillating probe reaches a steady-state and the separation regulator achieves its objective. Thus, the detection bandwidth is limited by the transient settling time of the oscillating probe. As we have discussed, the amplitude and phase lag evolve on a characteristic time scale of $2Q_0/\omega_0$, which corresponds to the ring-down time of the probe in the absence of tip–sample forces and without feedback control. The corresponding bandwidth (rad/s) is

$$B_{\text{osc}} = \frac{\omega_0}{2Q_0}. \tag{19}$$

However, it is important to note that in the presence of tip–sample forces, the settling time can potentially be much longer. Such is the case when operating close to a bifurcation point between stable and unstable amplitude branches [21,31].

The feedback control plays an important role in determining the detection bandwidth in dAFM. As we have discussed, the auxiliary regulators in FM and DAM essentially have complete control over the oscillations of the probe, including the transient settling time. In [53], it was determined experimentally that FM and DAM can achieve similar detection bandwidths in vacuum. We remark that self-excited FM has a detection bandwidth of ω_0 [8], but only when operating in the linear regime ($a^* << \lambda$), which is rarely the case in high-resolution imaging. In a nonlinear regime, the frequency shift is coupled to the amplitude response [12], and the detection bandwidth is limited accordingly. Thus, the amplitude regulator in self-excited FM determines the overall detection bandwidth for high-resolution imaging when B_{osc} is small. On the other hand, the measurement bandwidth in AM is limited roughly by the oscillator bandwidth B_{osc}.

Table 1 lists B_{osc} for high-resolution images in vacuum and liquid. The two AM references in vacuum achieve a relatively high bandwidth in vacuum by using nonstandard probes. Kawai and Kawakatsu [52] exploited a higher eigenmode of a silicon cantilever that had an unperturbed resonance frequency of 1.8 MHz. Erlandsson et al. [51] used a tungsten wire with an unperturbed quality factor of just 550 in vacuum. The references for FM in vacuum in Table 1 rely on feedback control to improve the detection bandwidth by about one order of magnitude over standard probes, yet still fall far short of our 1 kHz estimation. On the other hand, the low Q's in liquids ensure that the bandwidth requirement is met without including the auxil-

iary feedback regulators (again, see Table 1). Consequently, AM is more successful at high-resolution imaging in liquids than in vacuum.

Disturbance mitigation

A critical function of the feedback regulation in dAFM is to sustain the probe oscillations in the presence of unknown tip–sample forces, i.e., to mitigate disturbances from the tip–sample forces. In vacuum, the auxiliary regulators used in FM and DAM are essential for sustaining the oscillations in the presence of large noncontact tip–sample forces. Giessibl et al. [10] postulate that $e_{ts} < 1$ is required to maintain stable oscillations. However, the data in Figure 1 show that the auxiliary regulators are capable of maintaining stable oscillations when the magnitudes of e_{ts} and v_{ts} are much larger than unity. On the other hand, the approach taken in AM is simply to limit the magnitudes of e_{ts} and v_{ts}, in order to keep the amplitude from being attenuated. For example, choosing $a_{sp}/a_0 = 1/2$ in AM requires that the magnitudes of e_{ts} and v_{ts} do not exceed unity according to Equation 6. This limited approach to feedback regulation in AM can be problematic in vacuum and air where noncontact forces are large, but it is generally sufficient for imaging in liquids.

Stability

A final issue surrounding the high-resolution imaging in dAFM is stability. We have already discussed the importance of the auxiliary regulators in eliminating bistable oscillation states and maintaining stable oscillations in the presence of tip–sample forces. We turn now to the issue of global stability of the separation regulator. We first consider a stability issue that is inherent in FM, commonly referred to as "tip crash" [54]. To simplify matters, we require that z is manipulated slowly by the separation regulator such that the auxiliary feedback regulators maintain the probe oscillations in a quasi-steady state. Setting $K_5 = 0$ and requiring K_6 to be small in Equation 9, the dynamics of z in FM are approximated by

$$\frac{dz}{d\tau} = -K_6\left[\sigma\left(z, a_{sp}\right) - \sigma_{sp}\right] \qquad (20)$$

where $\sigma_{sp} = -v_{ts}(z,a_{sp})$.

The schematic in Figure 3 shows the typical behavior of $\sigma(z,a_{sp})$ in vacuum or air where attractive forces are significant. The arrows indicate the direction in which z is instructed to move by the feedback regulator according to Equation 20. The equilibrium points z^* are the zero crossings of $\sigma(z,a_{sp}) - \sigma_{sp}$. It is shown in Figure 3 that equilibrium in the attractive regime is locally stable but lacks global stability. A perturbation in the tip–sample forces can cause the separation regulator to approach the sample indefinitely. The closer σ_{sp} is to the onset of repulsive forces, the more likely is the onset of this instability. On the other hand, separation regulator in DAM is designed to maintain e_{ts}. When the amplitude is constant, e_{ts} typically increases monotonically with respect to z (see Figure 1); a point that was originally made in [55]. Consequently, DAM allows the oscillating probe tip to pass through the boundary between attractive and repulsive forces without necessarily resulting in a loss of stability.

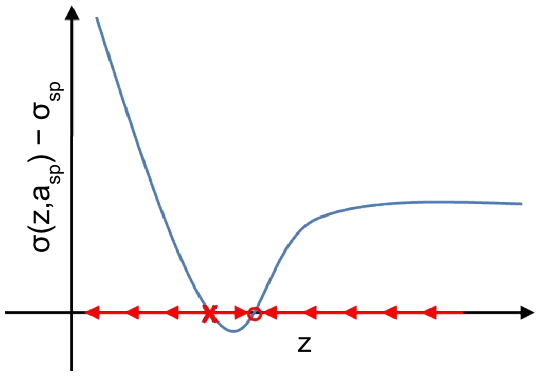

Figure 3: Stability of the z control in FM in vacuum. Arrows indicate the direction of the motion of z when placed under feedback control according to $dz/d\tau = -K_6[\sigma(z,a_{sp}) - \sigma_{sp}]$, where $K_6 > 0$.

Results and Discussion

In the previous sections, we investigated the role of the feedback control in high-resolution imaging. It is important to note that the task of regulating the oscillations of the probe under imaging conditions is greatly simplified in liquids due to the small long-range forces and low quality factors. Consequently, imaging modes with limited feedback control, such as AM, can be successful at high-resolution imaging in liquids [5]. In this section we present high-resolution images of mica in liquids with FM, DAM and AM acquired with the same probe and under similar operating conditions.

Figure 4 shows the topography and compositional images of freshly cleaved muscovite mica in water. Superimposed on the FM topography image is the theoretical structure of freshly cleaved muscovite mica, which exposes a plane of oxygen atoms (blue), which is slightly offset from a plane of silicon atoms (green) [2]. The FM dissipation image reveals a strong correlation with the FM topography image revealing dissipation on the atomic scale. The atomic-scale features in the FM dissipation image entertain the possibility of atomic-resolution imaging with DAM through regulation based on the tip–sample dissipation. This is indeed shown to be the case in Figure 4c. Similarly, atomic resolution is demonstrated in AM in Figure 4e. While topography images in AM reflect a combina-

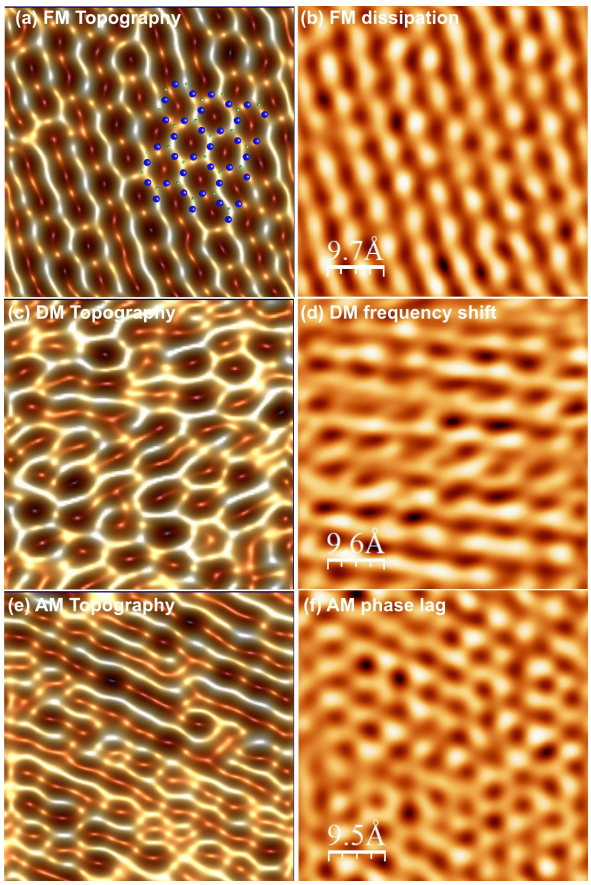

Figure 4: High-resolution images of mica in water taken with FM, DAM and AM. FM (a) topography and (b) dissipation. DAM (c) topography and (d) frequency-shift images. AM (e) topography and (f) phase-lag image. The variation in height in FM, DAM and AM topography images is 20, 80 and 40 pm, respectively, while the root-mean-squared surface roughness is 4, 13, and 6 pm, respectively. See Methods for additional specifications.

tion of the dissipation and virial, AM is more sensitive to the dissipation at high amplitude set-points, in which case AM resembles DAM.

The primary difference between the topography images in FM, DAM and AM in liquids is the treatment of tip–sample forces by the separation regulator. Of the three imaging modes, it appears that FM most faithfully reproduces the expected theoretical structure of freshly cleaved mica. From the data presented, it appears that regulation of the conservative component of the interaction captured by the virial is more favorable for atomic-resolution imaging. Furthermore, from Figure 1, we note that the magnitude of the virial is often larger than the dissipation for stiff inorganic samples such as mica. However, we stress that it was possible to obtain high-resolution images of mica in water in each of the three imaging modes under similar operating conditions.

Methods

The approach curves in Figure 1 were obtained by using the auxiliary regulators described in Equation 8, which is the typical precursory experiment to imaging with FM. The experimental data consist of F and σ versus z. Reconstruction of e_{ts} and v_{ts} versus z is achieved by substituting $a^* = a_0 = a_{sp}$ and $\phi^* = \pi/2$ into Equation 6 to yield

$$e_{ts}(z, a_{sp}) = a_{sp}\frac{F}{F_0} - 1, \tag{21}$$

$$v_{ts}(z, a_{sp}) = -\sigma. \tag{22}$$

Measurements were made on a silicon substrate with 1 nm native silicon oxide in vacuum with k = 27 N/m, Q_0 = 28,000 and a_{sp} = 8.5 nm and in ambient air with k = 36 N/m, Q_0 = 620 and a_{sp} = 8.6 nm. Data were acquired on glass in deionized water with k = 0.6 N/m, Q_0 = 1.6 and a_{sp} = 1.5 nm.

The high-resolution images of freshly cleaved mica in deionized water in Figure 4 were acquired with a Nanosenors™ PPP-NCH probe (k = 40 N/m, Q_0 = 11). Images were obtained in FM with a_{sp} = 0.7 nm, σ_{sp} = 0.01, in AM with a_0 = 0.6, $a_{sp} = 0.86a_0$, and in DAM with a_{sp} = 0.4 and $F_{sp} = 1.3F_0$. A wavelet filter with a scale of 0.13 nm was applied to each image by using the WSxM software [56]. The scan rate in the fast scan direction of the image raster is 440 nm/s, which was necessary to compensate for thermal drifts. For this scan rate, we calculate the required measurement bandwidth to be about 1 kHz for high-resolution imaging. All data were acquired with Nanotec Electrónica microscopes (Nanotec Eletronica S.L., Madrid, Spain) by using the WSxM software.

Conclusion

Through analysis and experiment, we have studied the performance metrics for high-resolution imaging in liquids with different dAFM imaging modes. In general, we find that while the quality factors of probes in liquids are typically low, the force sensitivity can be preserved by using soft probes with small amplitudes. Remarkably, it is possible for a probe in liquid to have a force sensitivity on par with the qPlus sensor in vacuum. Moreover, we find that the reduction in both attractive forces and quality factors that occurs in liquids decreases the importance of feedback control in obtaining stable, high-resolution images in liquids. Thus, the considerable advantages of FM over AM in obtaining high-resolution images in vacuum are not reproduced in liquids. These findings are supported by high-resolution images of mica obtained with FM, AM and DAM in liquid under similar operating conditions. From the data, it does

appear that FM still has some advantage over AM and DAM in atomic-resolution imaging. On the other hand, DAM offers robust stability for a range of environments and applications [53].

Acknowledgements

This research was supported by the National Science Foundation under Grant No. CMMI-0927648.

References

1. Martin, Y.; Williams, C. C.; Wickramasinghe, H. K. *J. Appl. Phys.* **1987,** *61,* 4723–4729. doi:10.1063/1.338807
2. Fukuma, T.; Kobayashi, K.; Matsushige, K.; Yamada, H. *Appl. Phys. Lett.* **2005,** *87,* 034101. doi:10.1063/1.1999856
3. Fukuma, T.; Kilpatrick, J. I.; Jarvis, S. P. *Rev. Sci. Instrum.* **2006,** *77,* 123703. doi:10.1063/1.2405361
4. Fukuma, T. *Jpn. J. Appl. Phys.* **2009,** *48,* 08JA01. doi:10.1143/JJAP.48.08JA01
5. Voïtchovsky, K.; Kuna, J. J.; Contera, S. A.; Tosatti, E.; Stellacci, F. *Nat. Nanotechnol.* **2010,** *5,* 401–405. doi:10.1038/nnano.2010.67
6. Suzuki, K.; Oyabu, N.; Kobayashi, K.; Matsushige, K.; Yamada, H. *Appl. Phys. Express* **2011,** *4,* 125102. doi:10.1143/APEX.4.125102
7. Giessibl, F. J.; Quate, C. F. *Phys. Today* **2007,** *60,* 17.
8. Albrecht, T. R.; Grütter, P.; Horne, D.; Rugar, D. *J. Appl. Phys.* **1991,** *69,* 668–673. doi:10.1063/1.347347
9. Wutscher, E.; Giessibl, F. J. *Rev. Sci. Instrum.* **2011,** *82,* 093703. doi:10.1063/1.3633950
10. Giessibl, F. J.; Bielefeldt, H.; Hembacher, S.; Mannhart, J. *Appl. Surf. Sci.* **1999,** *140,* 352–357. doi:10.1016/S0169-4332(98)00553-4
11. Nony, L.; Boisgard, R.; Aimé, J.-P. *Eur. Phys. J. B: Condens. Matter Complex Syst.* **2001,** *24,* 221–229. doi:10.1007/s100510170009
12. Gauthier, M.; Pérez, R.; Arai, T.; Tomitori, M.; Tsukada, M. *Phys. Rev. Lett.* **2002,** *89,* 146104. doi:10.1103/PhysRevLett.89.146104
13. Payton, O.; Champneys, A. R.; Homer, M. E.; Picco, L.; Miles, M. J. *Proc. R. Soc. A* **2010,** *467,* 1801–1822. doi:10.1098/rspa.2010.0451
14. Kilpatrick, J. I.; Gannepalli, A.; Cleveland, J. P.; Jarvis, S. P. *Rev. Sci. Instrum.* **2009,** *80,* 023701. doi:10.1063/1.3073964
15. Tamayo, J.; Humphris, A. D. L.; Owen, R. J.; Miles, M. J. *Biophys. J.* **2001,** *81,* 526–537. doi:10.1016/S0006-3495(01)75719-0
16. Hölscher, H.; Schwarz, U. D. *Appl. Phys. Lett.* **2006,** *89,* 073117. doi:10.1063/1.2336723
17. Ashby, P. D. *Appl. Phys. Lett.* **2007,** *91,* 254102. doi:10.1063/1.2824576
18. Khalil, H. K. *Nonlinear Systems;* Prentice Hall: Upper Saddle River, NJ, 2002.
19. Melcher, J.; Hu, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 053101. doi:10.1063/1.2767173
20. Butt, H.-J.; Cappella, B.; Kappl, M. *Surf. Sci. Rep.* **2005,** *59,* 1–152. doi:10.1016/j.surfrep.2005.08.003
21. Lee, S. I.; Howell, S. W.; Raman, A.; Reifenberger, R. *Phys. Rev. B* **2002,** *66,* 115409. doi:10.1103/PhysRevB.66.115409
22. Wang, L. *Appl. Phys. Lett.* **1998,** *73,* 3781–3783. doi:10.1063/1.122893
23. Hu, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 123106. doi:10.1063/1.2783226
24. Eq. 3 is strictly valid for $Q \gg 1$ and $\omega \approx \omega_0$.
25. The virial, as defined in Eq. 5, refers only to the oscillatory component of the force and includes a additional scaling of 2π for convenience.
26. San Paulo, Á.; García, R. *Phys. Rev. B* **2001,** *64,* 193411. doi:10.1103/PhysRevB.64.193411
27. Giessibl, F. J.; Bielefeldt, H. *Phys. Rev. B* **2000,** *61,* 9968–9971. doi:10.1103/PhysRevB.61.9968
28. Anczykowski, B.; Gotsmann, B.; Fuchs, H.; Cleveland, J. P.; Elings, V. B. *Appl. Surf. Sci.* **1999,** *140,* 376–382. doi:10.1016/S0169-4332(98)00558-3
29. Anczykowski, B.; Krüger, D.; Fuchs, H. *Phys. Rev. B* **1996,** *53,* 15485–15488. doi:10.1103/PhysRevB.53.15485
30. García, R.; San Paulo, A. *Phys. Rev. B* **1999,** *60,* 4961–4967. doi:10.1103/PhysRevB.60.4961
31. Rützel, S.; Lee, S. I.; Raman, A. *Proc. R. Soc. London, Ser. A* **2003,** *459,* 1925–1948. doi:10.1098/rspa.2002.1115
32. García, R.; Pérez, R. *Surf. Sci. Rep.* **2002,** *47,* 197–301. doi:10.1016/S0167-5729(02)00077-8
33. Raman, A.; Melcher, J.; Tung, R. *Nano Today* **2008,** *3,* 20–27. doi:10.1016/S1748-0132(08)70012-4
34. Stark, R. W. *Mater. Today* **2010,** *13,* 24–32. doi:10.1016/S1369-7021(10)70162-0
35. van Noort, S. J. T.; Willemsen, O. H.; van der Werf, K. O.; de Grooth, B. G.; Greve, J. *Langmuir* **1999,** *15,* 7101–7107. doi:10.1021/la990459a
36. Basak, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 064107. doi:10.1063/1.2760175
37. Xu, X.; Melcher, J.; Basak, S.; Reifenberger, R.; Raman, A. *Phys. Rev. Lett.* **2008,** *102,* 060801. doi:10.1103/PhysRevLett.102.060801
38. Melcher, J.; Carrasco, C.; Xu, X.; Carrascosa, J. L.; Gómez-Herrero, J.; de Pablo, P. J.; Raman, A. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106,* 13655–13660. doi:10.1073/pnas.0902240106
39. Martinez-Martin, D.; Carrasco, C.; Hernando-Perez, M.; de Pablo, P. J.; Gomez-Herrero, J.; Perez, R.; Mateu, M. G.; Carrascosa, J. L.; Kiracofe, D.; Melcher, J.; Raman, A. *PLoS One* **2012,** *7,* e30204. doi:10.1371/journal.pone.0030204
40. Martínez, N. F.; Kamiński, W.; Gómez, C. J.; Albonetti, C.; Biscarini, F.; Pérez, R.; García, R. *Nanotechnology* **2009,** *20,* 434021. doi:10.1088/0957-4484/20/43/434021
41. Kamiński, W.; Pérez, R. *Tribol. Lett.* **2010,** *39,* 295–309. doi:10.1007/s11249-010-9662-9
42. Fukuma, T.; Kimura, M.; Kobayashi, K.; Matsushige, K.; Yamada, H. *Rev. Sci. Instrum.* **2005,** *76,* 053704. doi:10.1063/1.1896938
43. Colchero, J.; Cuenca, M.; Martínez, J. F. G.; Abad, J.; García, B. P.; Palacios-Lidón, E.; Abellán, J. *J. Appl. Phys.* **2011,** *109,* 024310. doi:10.1063/1.3533769
44. Ekinci, K. L.; Yang, Y. T.; Roukes, M. L. *J. Appl. Phys.* **2004,** *95,* 2682. doi:10.1063/1.1642738
45. Gauthier, S. *J. Appl. Phys.* **2011,** *110,* 036107. doi:10.1063/1.3619776
46. Colchero, J.; Francisco González Martínez, J.; Abad, J.; Pérez García, B.; Cuenca, M.; Palacios-Lidón, E.; Abellán, J. *J. Appl. Phys.* **2011,** *110,* 036108. doi:10.1063/1.3619795
47. Giessibl, F. J.; Hembacher, S.; Bielefeldt, H.; Mannhart, J. *Science* **2000,** *289,* 422–425. doi:10.1126/science.289.5478.422
48. Giessibl, F. J.; Hembacher, S.; Herz, M.; Schiller, C.; Mannhart, J. *Nanotechnology* **2004,** *15,* S79–S86. doi:10.1088/0957-4484/15/2/017
49. Giessibl, F. J. *Science* **1995,** *267,* 68–71. doi:10.1126/science.267.5194.68

50. Uchihashi, T.; Sugawara, Y.; Tsukamoto, T.; Ohta, M.; Morita, S.; Suzuki, M. *Phys. Rev. B* **1997,** *56,* 9834–9840. doi:10.1103/PhysRevB.56.9834
51. Erlandsson, R.; Olsson, L.; Mårtensson, P. *Phys. Rev. B* **1996,** *54,* R8309–R8312. doi:10.1103/PhysRevB.54.R8309
52. Kawai, S.; Kawakatsu, H. *Appl. Phys. Lett.* **2006,** *88,* 133103. doi:10.1063/1.2189193
53. Jaafar, M.; Martínez-Martín, D.; Cuenca, M.; Melcher, J.; Raman, A.; Gómez-Herrero, J. *Beilstein J. Nanotechnol.* **2012,** *3,* 336–344. doi:10.3762/bjnano.3.38
54. Giessibl, F. J. *Rev. Mod. Phys.* **2003,** *75,* 949–983. doi:10.1103/RevModPhys.75.949
55. Jarvis, S. P.; Yamada, H.; Kobayashi, K.; Toda, A.; Tokumoto, H. *Appl. Surf. Sci.* **2000,** *157,* 314–319. doi:10.1016/S0169-4332(99)00545-0
56. Horcas, I.; Fernández, R.; Gómez-Rodríguez, J. M.; Colchero, J.; Gómez-Herrero, J.; Baro, A. M. *Rev. Sci. Instrum.* **2007,** *78,* 013705. doi:10.1063/1.2432410

License and Terms

Beilstein Journal of Nanotechnology

Bimodal atomic force microscopy driving the higher eigenmode in frequency-modulation mode: Implementation, advantages, disadvantages and comparison to the open-loop case

Daniel Ebeling and Santiago D. Solares*§

Full Research Paper

Open Access

Address:
Department of Mechanical Engineering, University of Maryland, College Park, MD 20742, USA

Email:
Daniel Ebeling - debeling@umd.edu; Santiago D. Solares* - ssolares@umd.edu

* Corresponding author
§ Phone: +1 (301) 405-5035; Fax: +1 (301) 314-9477

Keywords:
amplitude-modulation; atomic force microscopy; frequency-modulation; phase-locked loop; spectroscopy

Beilstein J. Nanotechnol. **2013,** *4,* 198–207.
doi:10.3762/bjnano.4.20

Received: 24 January 2013
Accepted: 04 March 2013
Published: 18 March 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We present an overview of the bimodal amplitude–frequency-modulation (AM-FM) imaging mode of atomic force microscopy (AFM), whereby the fundamental eigenmode is driven by using the amplitude-modulation technique (AM-AFM) while a higher eigenmode is driven by using either the constant-excitation or the constant-amplitude variant of the frequency-modulation (FM-AFM) technique. We also offer a comparison to the original bimodal AFM method, in which the higher eigenmode is driven with constant frequency and constant excitation amplitude. General as well as particular characteristics of the different driving schemes are highlighted from theoretical and experimental points of view, revealing the advantages and disadvantages of each. This study provides information and guidelines that can be useful in selecting the most appropriate operation mode to characterize different samples in the most efficient and reliable way.

Introduction

Atomic force microscopy (AFM) emerged in the mid-1980s as a powerful tool for measuring topography and forces on micro- and nanoscale surfaces [1]. Over the years, a number of new challenges have arisen in the implementation of such characterization, which have led to highly sophisticated approaches. In 2004 Garcia and co-workers [2] reported on computational simulations of a bimodal AFM technique for the simultaneous imaging of topography and mapping of compositional contrast across the sample. Within their method the fundamental cantilever eigenmode was used to acquire the sample topog-

raphy through the amplitude-modulation (AM) scheme while the second eigenmode was driven with a much smaller amplitude in open loop (OL, that is, only the first mode amplitude signal was used to control the tip–sample distance feedback loop. The second eigenmode drive signal had a constant amplitude and frequency like in standard AM-AFM, but its response was not considered in the control logic). The key advantages of this approach were (i) the ability to vary and optimize the parameters of the higher eigenmode without being restricted by the topographical acquisition control loops, and (ii) higher sensitivity of the second phase contrast to material properties in the small-amplitude regime. This method, which was later implemented experimentally [3] and studied further theoretically and computationally [4,5], gave birth to a new host of multifrequency AFM techniques, which nowadays include a wide variety of complementary methods for characterization in liquids, air and vacuum [6].

We describe here a previously introduced technique [7] similar to the original method of Garcia and co-workers (henceforth referred to as the AM-OL method), but in which the higher eigenmode is driven by using the frequency-modulation (FM, [8,9]) method, as had been previously done for vacuum operations [10,11]. Our multifrequency technique was originally introduced for ambient air operation within a trimodal scheme [12,13], in which a third active eigenmode was added to the AM-OL method. Since the dynamics of this trimodal approach are quite complex and the technique is still in the early stages of development, we focus here on a more in-depth presentation of the bimodal AM-FM method. In particular, we discuss advantages, disadvantages and differences in contrast with respect to the AM-OL approach, as well as its general applicability. The aim is to provide sufficient background to help users discern the most appropriate of these two methods for specific applications, rather than making generalizations that place one technique above the other. When appropriate we also offer brief comparisons with other imaging modes.

Control scheme of the AM-FM mode

Figure 1 shows a diagram of the experimental setup used, which consists of a commercial AFM system (MFP3D with ARC2 controller, Asylum Research Corporation, Santa Barbara, CA, USA) equipped with external phase-locked loop (PLL) electronics (PLL Pro 2, RHK Technology, Troy, MI, USA). As is customary, the oscillation of the cantilever near the sample surface was tracked with the laser beam deflection method. In the case of AM-OL operation the drive signal was generated by adding the signals of two function generators, which were set to the frequencies of the 1st and 2nd eigenmodes of the cantilever, respectively. The amplitudes and phase shifts for each eigenmode are read out by two separate lock-in amplifiers. This operation mode can be accomplished by either using the two built-in function generators and lock-in amplifiers from the MFP3D electronics or using the MFP3D electronics for the 1st eigenmode excitation only and providing the 2nd eigenmode drive signal through the RHK electronics. We decided in favor of the second option in order to be able to switch quickly between AM-OL and AM-FM operation.

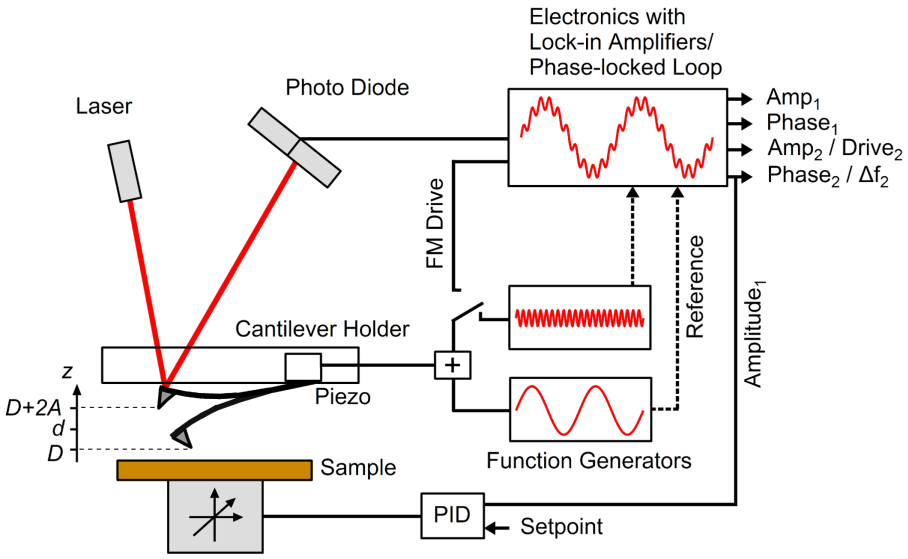

Figure 1: Schematic setup of our AFM operated in AM-OL or AM-FM mode. AM-OL mode can be accomplished by adding the drive signals of two function generators for exciting the cantilever beam and using two lock-in amplifiers to track the amplitudes and phase shifts at the corresponding eigenmodes. For AM-FM mode external PLL electronics were used, which provided a drive signal with a constant phase shift of 90° between excitation and response while tracking changes in the instantaneous resonant frequency.

In the case of AM-FM operation the 1st eigenmode of the cantilever was excited by the MFP3D controller in AM-mode, while the 2nd eigenmode was driven by the PLL in FM-mode. The PLL electronics continuously measure the instantaneous frequency of the oscillation signal and generate an excitation signal at this frequency, which is fed back to the dithering piezo (see, e.g., [14] for details) in order to keep the corresponding eigenmode always oscillating at its actual resonance frequency. The PLL can be operated in two different ways. It can either keep the drive amplitude of the excitation signal constant (this approach is often denoted as constant excitation (CE) mode [15]) or it can provide a signal with variable drive amplitude to maintain a constant oscillation amplitude of the cantilever eigenmode (accordingly denoted as constant amplitude (CA) mode [9]). The latter case is internally realized by running an additional feedback loop that controls the oscillation amplitude. Besides, since AM-OL operation can be implemented by using the MFP3D electronics alone, the setup of Figure 1 can also be employed to simultaneously excite three eigenmodes of the cantilever as described in [12,13].

Results and Discussion

Comparison of AM-FM and AM-OL methods

Comparisons of the AM-OL and AM-FM techniques are difficult because the two methods do not carry out identical tasks and also because there are a number of tradeoffs involved, which may or may not be advantageous depending on the particular application and level of skill of the user. We focus here on three angles of comparison: (i) the general appeal of frequency shift and relationship to phase contrast, (ii) the amplitude control capability and its implications (especially with regards to sensitivity), and (iii) complexity and stability. Additionally, we offer a brief discussion on selection criteria in terms of sample type, instrumentation availability and user skill level.

General appeal of frequency shift and relationship to phase contrast

The first question that emerges when discussing AM-FM concerns the reasoning behind the use of FM-AFM, which has in the past been mostly reserved for vacuum operation, with a few exceptions in liquid imaging [16,17] and spectroscopy experiments in air as well as in liquid [18-21]. Historically, FM-AFM addressed the limitation brought about by the large transient times observed in classical AM-AFM, where the oscillation amplitude is used as an input signal for the tip–sample-distance feedback loop. These transient times scale as $2Q/\omega_0$, with Q being the quality factor and ω_0 the natural frequency [22]. Clearly, imaging becomes impractical when Q increases significantly (as in vacuum operations). In FM-AFM, this drawback is overcome by using the frequency shift as a feedback input, which shows an instantaneous response to variations in the tip–sample forces (on the order of the oscillation period). However, FM-AFM can also be attractive for nonvacuum operations, primarily because of its suitability for spectroscopy experiments. For CA-mode operation it can be shown that the frequency shift signal is, at a first approximation, only affected by conservative interactions while the measured drive amplitude is mainly influenced by dissipation [23-25]. However, in AM-AFM both measured signals (amplitude and phase shift) depend on both types of interactions (conservative and dissipative) [26] (note that the measured frequency shift is also *indirectly* affected by dissipation in large-amplitude intermittent-contact experiments, in that dissipative forces can limit penetration of the probe tip into the repulsive region of the tip–sample interaction potential, thus leading to lower frequency shifts [13]). The above fact complicates the reconstruction of tip–sample interactions when performing spectroscopy in AM-AFM (see [26]). Furthermore, known bistabilities from AM operation [22,27,28] do not occur in FM-AFM, ensuring smoother characterization in some cases, as well as facilitating mathematical reconstruction of the force curves and ensuring continuous acquisition of data without any jumps in the signals or cantilever response [29,30]. This is highly relevant in multifrequency operation, where one seeks to integrate imaging and spectroscopy. Finally, FM-AFM also has the potential advantage to enable real-time 3D force spectroscopy in multifrequency operation, *in the limit of small response time*. As previously simulated [31,32], if a sufficiently high eigenmode were *self-excited* while performing intermittent contact imaging with the fundamental eigenmode, such that each higher-mode oscillation remained at the instantaneous resonance frequency, one could reconstruct the tip–sample force gradient in the volume above the sample, defined by the raster scan and the oscillation amplitude of the fundamental mode. It is straightforward to carry out integration of the force gradient to calculate the forces as a function of the *xyz*-coordinates. This results in three dimensional force fields, as they are usually obtained in time-consuming volume-scanning applications [33,34]. Although the approach described in [31,32] is not yet experimentally feasible, it represents a promising theoretical limit.

The next question concerns the relationship between the frequency shift and the phase contrast, which can be easily answered by using the damped harmonic oscillator model [22]. In the absence of tip–sample dissipation, the phase of the oscillator's response with respect to the excitation, ϕ, can be calculated through the expression,

$$\tan\phi = \frac{\omega\omega_r / Q}{\omega_r^2 - \omega^2} \quad (1)$$

where ω and ω_r are the excitation and resonance angular frequencies, respectively. The angular resonance frequency is, at a first approximation, related to the *effective* force gradient experienced by the oscillator through the equation,

$$\omega_r = \sqrt{\frac{k - \langle \nabla F_e \rangle}{m}} \quad (2)$$

where k is the oscillator force constant, F_e is the external force (the tip–sample force in this case) and m is its effective mass, equal to k / ω_0^2. Substituting Equation 2 into Equation 1 and setting the excitation frequency to be equal to the free resonance frequency, one obtains an equation that relates the response phase to the effective force gradient, which is plotted in Figure 2 together with the resonance frequency, as a function of the effective force gradient [12,13]. The graphs show that the phase and frequency shift vary in opposite (antiparallel) directions when the magnitude of the force gradient is small [12,13]. Figure 2 also highlights one additional advantage of using the frequency shift in that its behavior remains closer to a linear response for much larger magnitudes of the external force gradient, as compared to the phase response. The nearly antiparallel relationship between phase and frequency can be easily observed in a trimodal experiment in which one higher eigenmode is driven in OL and another one is driven in FM [12,13], or in separate bimodal experiments conducted in AM-OL or AM-FM [7].

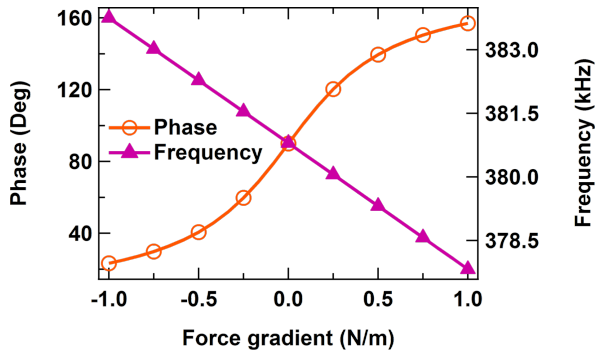

Figure 2: Phase and frequency shift calculated analytically for a higher eigenmode with f_2 = 380.8 kHz, k_2 = 64.2 N/m and Q_2 = 450 as a function of the effective tip–sample force gradient.

Figure 3 shows a comparison of bimodal AFM measurements of DNA strands adsorbed on a mica surface obtained in three different operation modes under ambient air (sample preparation details are provided below). The images were in each case captured over nearly the same sample position in AM-OL (left), AM-FM (CE) (middle), and AM-FM (CA) (right) modes, respectively. The 3D topography images are overlaid by color-scaled images of the corresponding 2nd eigenmode channels. To demonstrate the antiparallel contrasts of phase and frequency shifts we inverted the phase image (top left image) resulting in a very similar contrast as observed for the two frequency-shift channels (top middle and right). The 2nd eigenmode oscillation amplitude (AM-OL and AM-FM (CE)) and drive amplitude (AM-FM (CA)) channels are depicted in the second row of Figure 3. For these channels a clear difference between AM-OL and AM-FM operation is observed revealing the strongest contrast in the AM-mode. To be able to compare the contrast in the different modes more quantitatively, we converted the measured phase/frequency shifts and amplitudes into virial and energy dissipation, which are the time averages of the conservative and dissipative tip–sample interactions, respectively (see, e.g., equations (1a) and (1b) in [7] and references therein. Here we converted the dissipated power into energy dissipation per oscillation cycle by multiplying the power by the period length). These quantities are depicted in the two bottom rows of Figure 3 for one scan line. For the specific sample system and the actual imaging parameters used, both virial and dissipation channels reveal significant differences between AM and FM operation. While in AM mode the image contrast shows up in both channels almost equally, the FM data show highly diminished contrast in the dissipation channel but in return a slightly increased contrast for the virial.

The cause for this behavior lies in the complex inherent characteristics of the different excitation mechanisms. For guaranteeing better comparability of the data from diverse operation modes the average 2nd eigenmode oscillation amplitudes were adjusted to the same value (approx. 0.33 nm) during the experiments. To accomplish the same engaged amplitude, three different free amplitudes (that is, amplitudes without tip–sample interaction) had to be used, which were 0.75 nm (AM), 0.58 nm (CE) and 0.33 nm (CA) [35]. In light of this fact it becomes clear that a direct comparison of the various excitation schemes is difficult and has to be conducted very carefully, especially for highly dissipative sample systems. The comparison of OL, CE and CA presented in [7] seems more straightforward than for the results presented here, because in those experiments the engaged amplitude of the spectroscopic eigenmode does not differ significantly from the free oscillation amplitude for OL and CE. This is primarily a consequence of using the 3rd cantilever eigenmode in that study instead of the 2nd eigenmode used here (the 3rd eigenmode is significantly stiffer than the 2nd, and is thus influenced by the tip–sample forces to a lesser degree). However, this is not generally the case. Instead, the ratio of engaged to free amplitude can be small and differ between OL and CE throughout the sample.

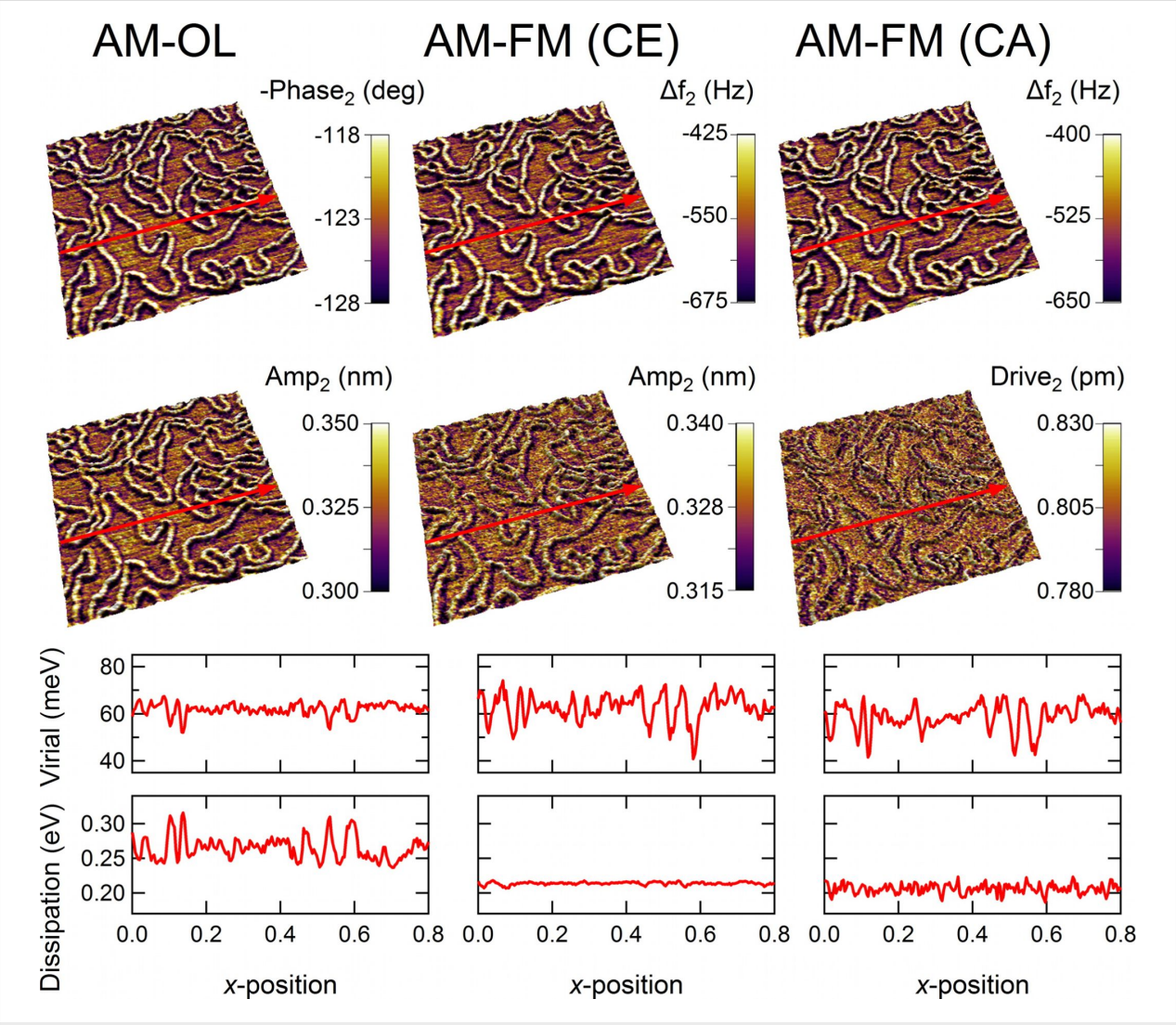

Figure 3: Comparison of 2nd eigenmode contrasts for different operation modes (left: AM-OL, middle: AM-FM (CE), right: AM-FM (CA). The two top rows show 3D topography images (800 × 800 nm^2) of lambda DNA strands adsorbed onto a mica substrate. For each operation mode the topography images are overlaid by their corresponding 2nd eigenmode signal channels, which are phase shift (inverted) and oscillation amplitude (AM-OL), frequency shift and oscillation amplitude (AM-FM (CE)), or frequency shift and drive amplitude (AM-FM (CA)). In the two bottom rows the virial and energy dissipation for the scan line indicated are depicted. Parameters: rectangular silicon cantilever (PPP-NCSTAuD, Nanosensors), f_1 = 92.4 kHz, k_1 = 2.0 N/m, Q_1 = 208, f_2 = 586.8 kHz, k_2 = 97.6 N/m, Q_2 = 678).

In Figure 4 all data channels that were captured during bimodal operation are presented for the different operation modes. Shown are 500 × 800 nm^2 2D images of the same sample location as in Figure 3. The rows from top to bottom show the height, 1st amplitude, 1st phase, 2nd (oscillation or drive) amplitude, and 2nd phase or frequency-shift channels. Whenever possible, images in the same rows were scaled to the same ranges. Here we observed an increased contrast for the 2nd eigenmode signals in comparison to the corresponding 1st mode signals in agreement with previous results reported in the literature (see, e.g., [3,10,36-38]). It is interesting to note that the 2nd mode phase and frequency-shift images reveal a significantly improved lateral resolution in comparison to the height channels. Especially at locations where DNA strands lay close to each other the height contrast becomes ambiguous, not allowing for an identification of single strands. However, in the 2nd phase and frequency-shift channels the borderlines of the single DNA strands are clearly visible. Furthermore, comparing the 1st and 2nd phase channels in AM-OL mode it becomes apparent that the observed image contrast in the 2nd eigenmode channel is about one order of magnitude higher than in the 1st mode channel for the chosen parameters. On closer examination of the 2nd eigenmode channels in Figure 4 one notices some asymmetry in the image contrast around the DNA strands in all three operation modes. We checked whether this effect is influenced by the scanning direction but this was not the case. We believe

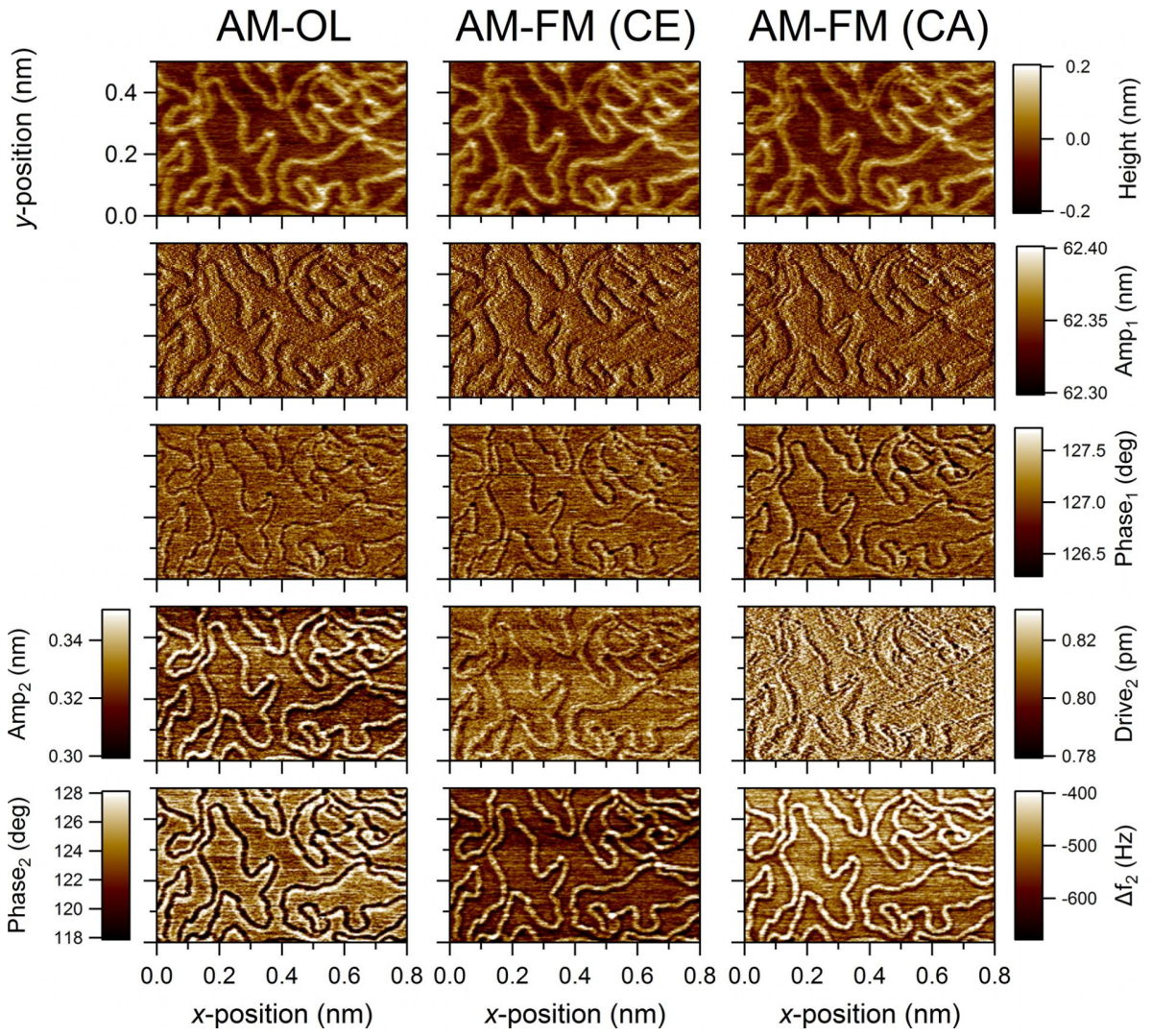

Figure 4: 500 × 800 nm^2 sized images of DNA/mica samples in three different multifrequency AFM modes (AM-OL, AM-FM (CE), AM-FM (CA)). Data channels in rows from top to bottom show: height, 1st amplitude, 1st phase shift, 2nd amplitude (oscillation or drive), 2nd phase or frequency shift (see Figure 3 for parameters).

that this asymmetry is likely to be related to the shape of the tip apex.

Amplitude control capability and its implications

An important advantage of the AM-FM scheme with respect to the AM-OL method is that the spectroscopy eigenmode operates at a fixed phase of 90 degrees, corresponding to the natural frequency, where the cantilever is generally most sensitive to external forces and thus permits characterization with gentler impacts. Although the maximum amplitude (peak in the Lorentzian response) does not occur exactly at the natural frequency due to the influence of damping, which can be more significant when characterizing highly dissipative samples, the natural frequency is a well-defined condition, which allows the relatively easy implementation of amplitude control. That is, one can control the response amplitude by adjusting the drive amplitude, using a simple control loop, such that FM-AFM can be carried out either by using the constant excitation (CE) mode [15] or the constant amplitude (CA) mode [9]. In the former case the controller simply adjusts the drive frequency proportionally, letting the response amplitude vary when dissipation is present, while in the latter case the drive is also adjusted to keep the response amplitude constant. Due to the Lorentzian behavior of each eigenmode, which leads to amplitude and phase responses that depend *nonlinearly* on the ratio of the excitation frequency to the *instantaneous* resonance frequency (that is, the resonance frequency under the influence of tip–sample forces),

implementing a constant amplitude scheme when the phase is not locked is not as simple as rescaling the drive amplitude, although it is possible and has been demonstrated experimentally [39-41]. (Although our analysis does not address the use of phase modulation in multifrequency operations, it is worth mentioning that the phase-modulation approach has been proposed to be potentially very sensitive and fast due to relying on a phase detector, which has a faster response than a PLL and is not subject to loss of oscillation as in the case when a PLL unlocks [40]). The ability to keep the amplitude constant has direct implications on whether all regions of the sample are characterized with the same probe sensitivity. This can be understood by making dimensionless the equation of motion of a damped harmonic oscillator [22,42,43]:

$$\frac{d^2\underline{z}}{d\underline{t}^2} = -\underline{z} + \frac{1}{\mathrm{Q}}\left[-\frac{d\underline{z}}{d\underline{t}} + \cos(\underline{t})\right] + \frac{F_{\mathrm{ts}}(\underline{z}_{\mathrm{ts}})}{kA_0} \qquad (3)$$

where A_0 is the free oscillation amplitude, $\underline{z} = z(t)/A_0$ is the dimensionless tip position with respect to the cantilever base position, $\underline{z}_{\mathrm{ts}} = z_{\mathrm{ts}}/A_0$ is the dimensionless tip–sample distance ($z_{\mathrm{ts}} = z + z_{\mathrm{eq}}$, where z_{eq} is the equilibrium tip position with respect to the sample surface), $\underline{t} = \omega_0 t$ is the dimensionless time, k is the cantilever force constant (stiffness) and F_{ts} is the tip–sample interaction force. We have also used the approximation $A \approx A_0 = F_0\mathrm{Q}/k$ [22], where F_0 is the amplitude of the inertial excitation force, and have grouped the damping and excitation terms together in brackets with the coefficient 1/Q. It can be inferred from the last term on the right hand side of this equation that the tip–sample forces are normalized by the product of the force constant times the free oscillation amplitude, such that the external force term becomes more or less relevant to the dynamics when the product kA_0 becomes smaller or larger, respectively. Since the cantilever becomes more sensitive to external forces when this term becomes more dominant, sensitivity increases with decreasing amplitude. Thus, unless the amplitude is kept constant across the surface, regions of the sample where dissipation is high will lead to smaller amplitudes and, thus, will be characterized with higher cantilever sensitivity, leading to unequal treatment of all regions [7]. Clearly this is not a concern for samples that exhibit low dissipation or for cases where the user is able to set the amplitude to a value that is large enough to prevent the external force term from becoming dominant (see the trace for a free second eigenmode amplitude of $A_{2\text{-}0}$ = 20 nm in Figure 5, for which the influence of the external forces is small in comparison to the trace for $A_{2\text{-}0}$ = 1 nm), although the use of larger amplitudes leads to lower sensitivity and reduced material contrast, so it is not necessarily desirable. Similarly, there may be situations where the user is interested in increasing the sensitivity of the instrument across the entire sample. For example, in the detection of atomic-scale features in liquid environments, the use of increasingly smaller oscillation amplitudes leads to gradually increased dominance of the tip–sample forces in Equation 3, allowing the user to find optimum conditions that balance cantilever sensitivity with the ability of the instrumentation to detect changes in the signals. Figure 6 shows an example of this approach for imaging the mica–water interface using AM-OL. The top row in Figure 6 depicts high-resolution topography as well as 1st and 2nd eigenmode phase-shift images of a mica surface imaged in ultrapure water. The images reveal clearly the hexagonal structure of the mica surface and, at certain positions, single atomic/molecular adsorbates/defects. In the bottom row single scan lines for three different 2nd mode amplitudes are compared with each other. The red scan lines were taken from the images in the top row (position indicated by red arrow). The 2nd mode phase-shift scan lines reveal a significant increase in contrast for decreasing 2nd mode amplitudes. Here it is important to realize that Equation 3 is valid for any harmonic oscillator, so the ability to tune its sensitivity is not limited to bimodal operations. In fact, we have also observed similar trends as in Figure 6 when performing the characterization using a single eigenmode [38].

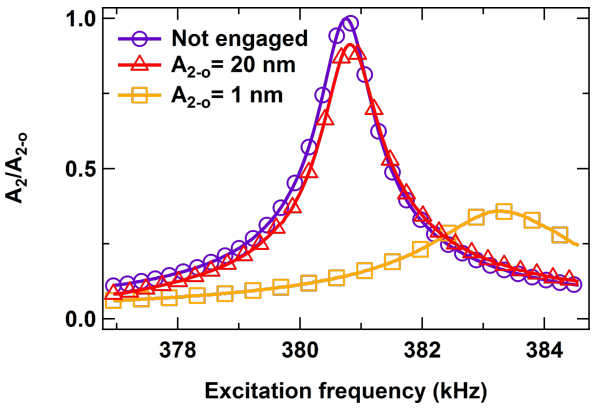

Figure 5: Simulation of the change in frequency response for the second eigenmode of a cantilever with fundamental eigenfrequency of 60.8 kHz (Q ≈ 150), second eigenfrequency of 380.8 kHz (Q ≈ 450), and fundamental force constant 1.6 N/m as a function of the second eigenmode's free oscillation amplitude. The fundamental free oscillation amplitude and amplitude setpoint are 100 nm and 70%, respectively. The sample is a soft dissipative polymer [13]. The graph illustrates the increased sensitivity of the second eigenmode as its free amplitude drops from 20 nm to 1 nm. In the former case the frequency shift and decrease in response amplitude are small compared to the free response. In the latter case, the eigenmode is much more sensitive, leading to a significantly larger frequency shift and much smaller oscillation amplitude.

Complexity and stability

As expected from a more sophisticated method having more "knobs," AM-FM offers useful advantages with respect to

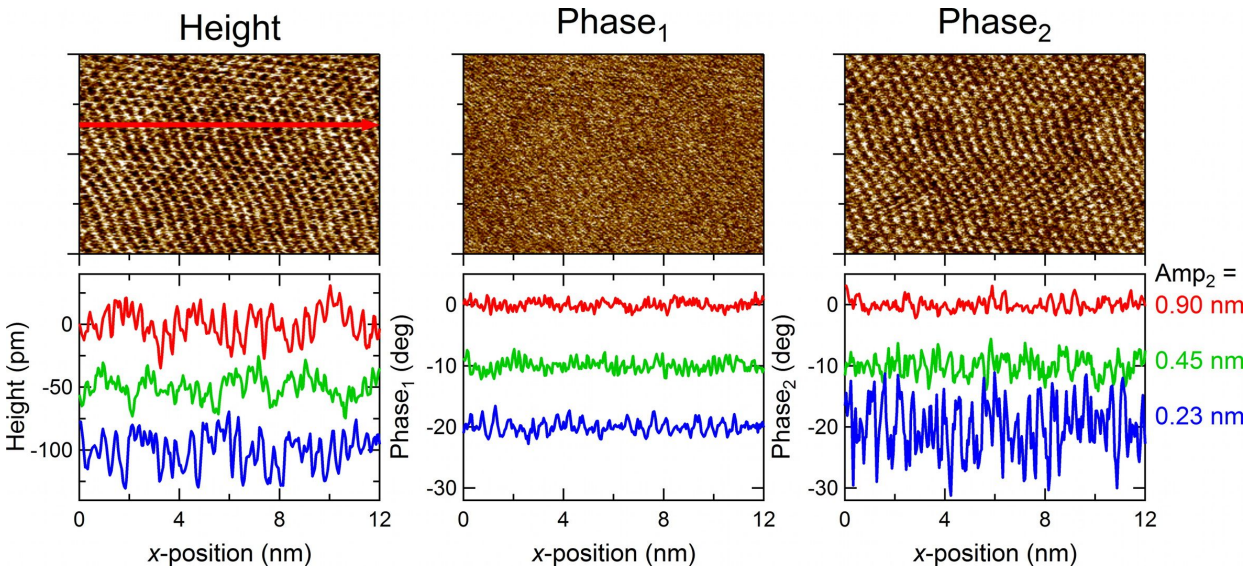

Figure 6: Top row: 8 × 12 nm^2 height, 1st phase shift, and 2nd phase shift images of a mica surface imaged in ultrapure water in AM-OL mode. Bottom row: single scan lines for three different 2nd mode amplitudes (0.90, 0.45, 0.23 nm). Parameters: rectangular silicon cantilever (PPP-NCSTAuD, Nanosensors), 1st free amplitude = 1.5 nm, amplitude setpoint = 0.75 nm.

AM-OL. However, this also comes at a price, which may or may not be worth paying. Specifically, the controls on the spectroscopy eigenmode are significantly more complex and the instrumentation more expensive. In AM-OL this eigenmode is driven at constant amplitude and frequency but not controlled. In this case "simple" lock-in amplifiers are sufficient to rapidly measure oscillation amplitudes and phase shifts at each eigenmode of the cantilever. In contrast, the CE version of AM-FM requires either a phase-locked-loop (PLL) or a self-excitation (phase-shift-based) loop to keep the phase locked at 90 degrees (so far we have observed that self-excitation loops are less stable in tapping-mode experiments using the same setup as in Figure 1, as it is very easy for the system to lose resonance during the tip–sample impact, where the instantaneous resonance frequency, and thus the length and required phase shift of each successive oscillation, changes rapidly). The CA version of AM-FM requires an additional control loop to keep the amplitude constant. This added complexity is not necessary for many samples of interest, for which AM-OL is sufficient. Furthermore, due to its robustness, AM-OL can be more advantageous when characterizing samples with properties (dissipation, stiffness or adhesive forces) that exhibit sharp variations across a wide range. Such variations pose challenges in AM-FM because the PLL control loops need to be tuned to the expected type of external forces, such that when these forces vary considerably, the operation may be detuned (and thus, less responsive or unstable at least to some degree) part of the time. If this leads to loss of lock in the PLL the user will be unable to obtain an image, in contrast to AM-OL, which will still produce an image. As an example consider the influence of dissipative forces, which can significantly lower the effective quality factor of the spectroscopy eigenmode. As discussed in [13] and illustrated by the 1 nm trace in Figure 5 (see orange trace for $A_{2\text{-}0}$ = 1 nm, which shows a drastic drop in the eigenmode's quality factor with respect to the red trace for $A_{2\text{-}0}$ = 20 nm), it is not unlikely that the effective quality factor of a higher eigenmode can become lower than the free quality factor of the fundamental eigenmode when dissipation is significant or small amplitudes are selected (as stated above, one may do this in order to measure with higher sensitivity). If the user tunes the PLL for free response, which exhibits a sharper Lorentzian curve than when tip–sample dissipative forces are present, it will necessarily be detuned upon engaging the sample with the cantilever, leading to a sluggish response.

On the other hand, the AM mode is known to be frequently accompanied by bistabilities [22,27,28] which may, depending on the sample and cantilever properties, significantly impair its imaging and spectroscopy capabilities. These bistabilities result from the existence of different oscillation states of the cantilever in AM-mode. Under certain imaging conditions it can happen that the oscillation jumps back and forth between those states, making proper tracking of the sample surface impossible. During spectroscopy measurements this phenomenon leads to discontinuities in the corresponding amplitude- and phase-versus-distance curves. Usually, these bistability issues do not occur in FM operation, since in this case, the cantilever is always driven at its actual resonance frequency [29].

Selection criteria

As can be gathered from the previous paragraphs, the different excitation mechanisms have their advantages and disadvantages both in terms of complexity and ease of data interpretation. It is well-known from single-mode operation that AM-mode is quite robust since it neither involves "locking" the phase with the PLL electronics nor setting up parameters (e.g., gains) to tune the system prior to characterization. On the other hand, it can be subject to bistabilities that may in some cases impair imaging as well as spectroscopy capabilities. For FM-operation, PLL or self-oscillation electronics are needed, which are more expensive and complex to use. Properly "locking" the phase and additional feedback loops for keeping the amplitude constant can be handicaps of this mode, especially when the user is not sufficiently experienced. However, the benefits are operation without bistabilities, straightforward way of separating and/or calculating the conservative and dissipative tip–sample interactions, and the ability to image with constant amplitude and thus uniform sensitivity. The advantages and disadvantages of each mode are similar in bimodal operation, although their relative importance can vary. For example, bistability issues are not as detrimental as in single-mode operation. Direct comparisons are more difficult in multi-frequency operation because, although some correspondence can be established between the contrast obtained from the different methods, there is not yet a *direct quantitative* relationship between them, especially when dissipation is present. The decision on which imaging mode to use should in general depend on the type of application and the purpose of the experiment. If, for example, topographical imaging of the sample surface with a general idea of the compositional contrast is the main objective (for example in the identification of two components in a sample), then it may be advantageous to profit from the robustness and simplicity of AM-OL mode. However, if inhomogeneous samples are involved, which contain wide variations in elastic and dissipative properties, and more "quantitative" data is sought, it may be advantageous to use the AM-FM scheme, which can guarantee constant sensitivity across the sample, even if the characterization is more complex and time-consuming (AM-OL can also be used with uniform sensitivity when the sample can be characterized with small variations in the engaged amplitude of the higher eigenmode with respect to the free response).

Conclusion

We have presented a theoretical and experimental comparison of three different bimodal AFM operation schemes, namely AM-OL, AM-FM (CE), and AM-FM (CA), which differ in the type of control scheme used to drive the higher eigenmode (open loop, constant-excitation frequency-modulation, and constant-amplitude frequency-modulation, respectively). The corresponding higher eigenmode channels exhibit clear differences in each case, which are closely related to dynamics and complexity of their corresponding driving mechanisms. In general, the AM-OL operation mode comes with ease of use, low requirements for special equipment, and robustness of operation. On the other hand, AM-FM is advantageous in enabling spectroscopy without "jumps", straightforward reconstruction of the tip–sample interaction force and the ability to operate with constant response amplitude, thus ensuring uniform sensitivity across the sample. Although generalized comparisons are not possible and it remains a challenge for the experimentalist to select the operation mode that best fits the actual application, we provide guidelines and physical insight that are helpful in making this selection.

Experimental

Lambda DNA (New England Biolabs) was obtained in buffer solution at a concentration of 500 μg/mL. This was diluted by adding ultrapure water (Sigma) and 1M $CaCl_2$ solution (Sigma) to a final concentration of 10 μg/mL DNA in 10 mM $CaCl_2$. Small amounts of this solution were stored at −20 °C for later use. Directly before imaging, 20 μL of the DNA in $CaCl_2$ solution was dropped onto a freshly cleaved muscovite mica (Ted Pella) surface. After waiting for approximately one minute the samples were first blown dry by air before being rinsed with ultrapure water and subsequently blown dry again. After this the samples were imaged in ambient air. For the high-resolution experiments (Figure 6) the mica samples were freshly cleaved and imaged in a droplet of ultrapure water. In this case the cantilever and its holder were rinsed by isopropanol, ethanol and water prior to imaging.

Acknowledgements

The authors gratefully acknowledge support from the U.S. National Science Foundation and the Department of Energy, through awards CMMI-0841840 and DESC0008115, respectively. We also acknowledge the support of the Maryland NanoCenter and its NispLab. The NispLab is supported in part by the NSF as a MRSEC Shared Experimental Facility. Finally, we acknowledge Dr. Roger Proksch from Asylum Research Corporation for coining the acronym "AM-FM", which we had not used in our previous presentations of this method.

References

1. Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986,** *56,* 930. doi:10.1103/PhysRevLett.56.930
2. Rodríguez, T.; Garcia, R. *Appl. Phys. Lett.* **2004,** *84,* 449. doi:10.1063/1.1642273
3. Martinez, N. F.; Patil, S.; Lozano, J. R.; Garcia, R. *Appl. Phys. Lett.* **2006,** *89,* 153115. doi:10.1063/1.2360894
4. Lozano, J. R.; Garcia, R. *Phys. Rev. Lett.* **2008,** *100,* 076102. doi:10.1103/PhysRevLett.100.076102

5. Lozano, J. R.; Garcia, R. *Phys. Rev. B* **2009,** *79,* 014110. doi:10.1103/PhysRevB.79.014110
6. Garcia, R.; Herruzo, E. T. *Nat. Nanotechnol.* **2012,** *7,* 217. doi:10.1038/nnano.2012.38
7. Chawla, G.; Solares, S. D. *Appl. Phys. Lett.* **2011,** *99,* 074103. doi:10.1063/1.3626847
8. Giessibl, F. J. *Rev. Mod. Phys.* **2003,** *75,* 949. doi:10.1103/RevModPhys.75.949
9. Albrecht, T. R.; Grütter, P.; Horne, D.; Rugar, D. *J. Appl. Phys.* **1991,** *69,* 668. doi:10.1063/1.347347
10. Kawai, S.; Glatzel, T.; Koch, S.; Such, B.; Baratoff, A.; Meyer, E. *Phys. Rev. Lett.* **2009,** *103,* 220801. doi:10.1103/PhysRevLett.103.220801
11. Naitoh, Y.; Ma, Z. M.; Li, Y. J.; Kageshima, M.; Sugawara, Y. *J. Vac. Sci. Technol., B* **2010,** *28,* 1210. doi:10.1116/1.3503611
12. Solares, S. D.; Chawla, G. *J. Appl. Phys.* **2010,** *108,* 054901. doi:10.1063/1.3475644
13. Solares, S. D.; Chawla, G. *Meas. Sci. Technol.* **2010,** *21,* 125502. doi:10.1088/0957-0233/21/12/125502
14. Nony, L.; Baratoff, A.; Schär, D.; Pfeiffer, O.; Wetzel, A.; Meyer, E. *Phys. Rev. B* **2006,** *74,* 235439. doi:10.1103/PhysRevB.74.235439
15. Ueyama, H.; Sugawara, Y.; Morita, S. *Appl. Phys. A* **1998,** *66,* S295. doi:10.1007/s003390051149
16. Fukuma, T.; Kobayashi, K.; Matsushige, K.; Yamada, H. *Appl. Phys. Lett.* **2005,** *87,* 034101. doi:10.1063/1.1999856
17. Rode, S.; Oyabu, N.; Kobayashi, K.; Yamada, H.; Kühnle, A. *Langmuir* **2009,** *25,* 2850. doi:10.1021/la803448v
18. Schmutz, J.-E.; Hölscher, H.; Ebeling, D.; Schäfer, M. M.; Anczykowski, B. *Ultramicroscopy* **2007,** *107,* 875. doi:10.1016/j.ultramic.2007.04.015
19. Ebeling, D.; Oesterhelt, F.; Hölscher, H. *Appl. Phys. Lett.* **2009,** *95,* 013701. doi:10.1063/1.3152771
20. Fukuma, T.; Ueda, Y.; Yoshioka, S.; Asakawa, H. *Phys. Rev. Lett.* **2010,** *104,* 016101. doi:10.1103/PhysRevLett.104.016101
21. Ebeling, D.; van den Ende, D.; Mugele, F. *Nanotechnology* **2011,** *22,* 305706. doi:10.1088/0957-4484/22/30/305706
22. García, R.; Pérez, R. *Surf. Sci. Rep.* **2002,** *47,* 197. doi:10.1016/S0167-5729(02)00077-8
23. Gotsmann, B.; Seidel, C.; Anczykowski, B.; Fuchs, H. *Phys. Rev. B* **1999,** *60,* 11051. doi:10.1103/PhysRevB.60.11051
24. Dürig, U. *New J. Phys.* **2000,** *2,* 5. doi:10.1088/1367-2630/2/1/005
25. Hölscher, H.; Gotsmann, B.; Allers, W.; Schwarz, U. D.; Fuchs, H.; Wiesendanger, R. *Phys. Rev. B* **2001,** *64,* 075402. doi:10.1103/PhysRevB.64.075402
26. Hölscher, H. *Appl. Phys. Lett.* **2006,** *89,* 123109. doi:10.1063/1.2355437
27. Anczykowski, B.; Krüger, D.; Fuchs, H. *Phys. Rev. B* **1996,** *53,* 15485. doi:10.1103/PhysRevB.53.15485
28. Gleyzes, P.; Kuo, P. K.; Boccara, A. C. *Appl. Phys. Lett.* **1991,** *58,* 2989. doi:10.1063/1.104690
29. Hölscher, H.; Gotsmann, B.; Allers, W.; Schwarz, U. D.; Fuchs, H.; Wiesendanger, R. *Phys. Rev. Lett.* **2002,** *88,* 019601. doi:10.1103/PhysRevLett.88.019601
30. Giessibl, F. J. *Phys. Rev. B* **1997,** *56,* 16010. doi:10.1103/PhysRevB.56.16010
31. Chawla, G.; Solares, S. D. *Meas. Sci. Technol.* **2009,** *20,* 015501. doi:10.1088/0957-0233/20/1/015501
32. Solares, S. D.; Chawla, G. *Meas. Sci. Technol.* **2008,** *19,* 055502. doi:10.1088/0957-0233/19/5/055502
33. Hölscher, H.; Langkat, S. M.; Schwarz, A.; Wiesendanger, R. *Appl. Phys. Lett.* **2002,** *81,* 4428. doi:10.1063/1.1525056
34. Albers, B. J.; Schwendemann, T. C.; Baykara, M. Z.; Pilet, N.; Liebmann, M.; Altman, E. I.; Schwarz, U. D. *Nat. Nanotechnol.* **2009,** *4,* 307. doi:10.1038/nnano.2009.57
35. The amplitude calibration for each eigenmode was accomplished by performing amplitude-versus-piezo-movement spectroscopy measurements while only exciting one single eigenmode at a time. The corresponding amplitude sensitivities for each eigenmode were determined separately by fitting the slopes of the amplitude-versus-piezo-movement curves.
36. Proksch, R. *Appl. Phys. Lett.* **2006,** *89,* 113121. doi:10.1063/1.2345593
37. Martínez, N. F.; Lozano, J.; Herruzo, E. T.; Garcia, F.; Richter, C.; Sulzbach, T.; García, R. *Nanotechnology* **2008,** *19,* 384011. doi:10.1088/0957-4484/19/38/384011
38. Ebeling, D.; Solares, S. D. *Nanotechnology* **2013,** *24,* 135702. doi:10.1088/0957-4484/24/13/135702
39. Hölscher, H. *J. Appl. Phys.* **2008,** *103,* 064317. doi:10.1063/1.2896450
40. Li, Y. J.; Takahashi, K.; Kobayashi, N.; Naitoh, Y.; Kageshima, M.; Sugawara, Y. *Ultramicroscopy* **2010,** *110,* 582. doi:10.1016/j.ultramic.2010.02.014
41. Sugawara, Y.; Kobayashi, N.; Kawakami, M.; Li, J. Y.; Naitoh, Y.; Kageshima, M. *Appl. Phys. Lett.* **2007,** *90,* 194104. doi:10.1063/1.2737907
42. Solares, S. D. *Meas. Sci. Technol.* **2008,** *19,* 015503. doi:10.1088/0957-0233/19/1/015503
43. Solares, S. D.; Chang, J.; Seog, J.; Kareem, A. U. *J. Appl. Phys.* **2011,** *110,* 094904. doi:10.1063/1.3657940

License and Terms

Beilstein Journal of Nanotechnology

Selective surface modification of lithographic silicon oxide nanostructures by organofunctional silanes

Thomas Baumgärtel*, Christian von Borczyskowski and Harald Graaf

Full Research Paper

Open Access

Address:
Center for Nanostructured Materials and Analytics, Institute of Physics, Chemnitz University of Technology Reichenhainer Str. 70, 09126 Chemnitz, Germany

Email:
Thomas Baumgärtel* - thomas.baumgaertel@physik.tu-chemnitz.de

* Corresponding author

Keywords:
AFM lithography; amino-functionalization; local anodic oxidation; octadecyl-trichlorosilane; silicon oxide nanostructures

Beilstein J. Nanotechnol. **2013,** *4,* 218–226.
doi:10.3762/bjnano.4.22

Received: 20 December 2012
Accepted: 01 March 2013
Published: 25 March 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

This study investigates the controlled chemical functionalization of silicon oxide nanostructures prepared by AFM-anodization lithography of alkyl-terminated silicon. Different conditions for the growth of covalently bound mono-, multi- or submonolayers of distinctively functional silane molecules on nanostructures have been identified by AFM-height investigations. Routes for the preparation of methyl- or amino-terminated structures or silicon surfaces are presented and discussed. The formation of silane monolayers on nanoscopic silicon oxide nanostructures was found to be much more sensitive towards ambient humidity than, e.g., the silanization of larger OH-terminated silica surfaces. Amino-functionalized nanostructures have been successfully modified by the covalent binding of functional fluorescein dye molecules. Upon excitation, the dye-functionalized structures show only weak fluorescence, which may be an indication of a relatively low surface coverage of the dye molecules on length scale that is not accessible by standard AFM measurements.

Introduction

Local anodic oxidation (LAO) nanolithography is a reliable and convenient method for the structuring of silicon on a nanometer scale [1-3]. The generated silicon oxide nanostructures are characterized by a height of only a few nanometers and a width down to some 10 nm [4]. After the local nanostructuring of the substrate, desired materials have to be selectively immobilized on the structures in order to provide them with an adjustable functionality. A combination of LAO with a treatment of the substrate surface by organic self-assembled monolayers (SAM) is a promising approach for a versatile combined top-down/bottom-up process towards the fabrication of functional nanodevices. Depending on the structure type, a multitude of routes have been proven applicable in a substantial number of publications [2,5-9].

In general, the different binding mechanisms that can be used for the selective functionalization of local anodic oxidation patterns on monolayer-terminated silicon can be divided into

three major groups: (A) intermolecular interactions [10,11], (B) electrostatic/ionic interactions [12,13], and (C) covalent bonds [14-17]. Covalent attachment of molecules is thereby a most favorable method since covalent bonds possess comparable high dissociation energy (several 100 kJ/mol), and with the knowledge of their successful implementation on nanostructures, the whole realm of modular chemistry can be applied (e.g., "Click"-chemistry [18]). Mechanisms such as electrostatic binding or binding by intermolecular forces (e.g., polar interactions or van-der-Waals forces) on the other hand are characterized by weaker binding strengths and a lower selectivity, and thus, are not as suitable for multistep surface functionalization as covalent binding.

Although the covalent functionalization of LAO patterns has been reported several times recently, most cases involve either a local mild oxidation of the monolayer only or an etching of the silicon oxide after LAO lithography [19]. Such realizations, however, require a very fine adjustment of the oxidation parameters or an additional chemical treatment. In this study, we report on a direct covalent binding to nanolithographic silicon oxide structures. A direct binding of the functional material to the oxide nanostructure is possible if the material possesses an appropriate anchoring group that binds selectively to silicon oxide surfaces. A standard substance for chemical modification of silicon oxide surfaces are silane compounds such as trichlorosilanes or ethoxysilanes [20,21]. In order to provide the possibility for further chemical modification, functional silanes are of major interest. These molecules possess a functional head-group (e.g., carboxyl, amino or thiol group) in addition to the silane tail group. This functional head group can react with other molecules resulting in an immobilization of the desired material on the structure with the silane molecule as linker. The binding of diverse silanes to silicon oxide surfaces has been the subject of several investigations [22-24]. Though a previous study demonstrated the general feasibility to bind silane molecules to LAO nanostructures, there was strong evidence for only a partial coverage of the structure and no complete monolayer formation [25].

Here we demonstrate a successful route for a homogeneous and dense binding of functional silane molecules to silicon oxide nanostructures prepared by LAO lithography of alkyl-terminated silicon [26]. In addition, the dye fluorescein-5-isothiocyanate (FITC) was subsequently bound to the amino-terminated silane layer. FITC is a fluorescein derivative with an N=C=S functional group. This group is reactive towards nucleophiles such as amine or thiol groups. A successful large-scale binding of FITC to amino-terminated silicon surfaces [27] has been demonstrated previously by fluorescence measurements [28]. The attachment of FITC on the oxide nanostructures is realized in two steps, as depicted schematically in Figure 1. First, the functional silane aminopropyltriethoxysilane (APTES) is bound to the LAO oxide, which leads to an amino-functionalization of the structure. In a second step, FITC is bound to the amino group on the structure through its N=C=S group. The successful binding is confirmed by AFM topography measurements and spectrally resolved fluorescence microscopy. Prior to the two-step functionalization with APTES and FITC, the quality of the silane layer formation is tested and proven by the binding of long-chained silanes, which are known to form densely packed and well-ordered monolayers on silicon surfaces [29,30].

Results and Discussion

Successful routes towards silane-functionalization of silicon oxide nanostructures

Prior to the multistep attachment of FITC through APTES linkers, the controlled and dense binding of silane molecules to the silicon oxide nanostructures has been investigated in detail. OTS (octadecyltrichlorosilane) has been used for these studies as its large molecular length allows for a reliable detection by

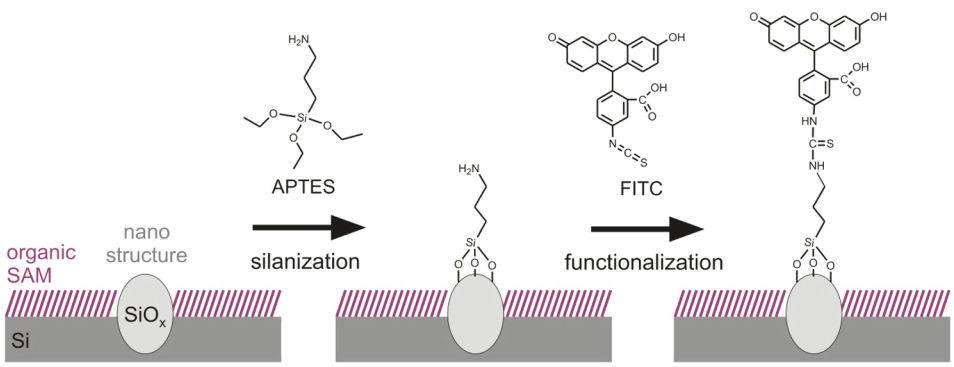

Figure 1: Schematic route for covalent binding of FITC to a silicon oxide nanostructure on alkyl-terminated silicon. After local anodic oxidation, a layer of APTES molecules is bound to the silicon oxide as linker. In a second step, FITC is bound to the amino-functionalized oxide structure.

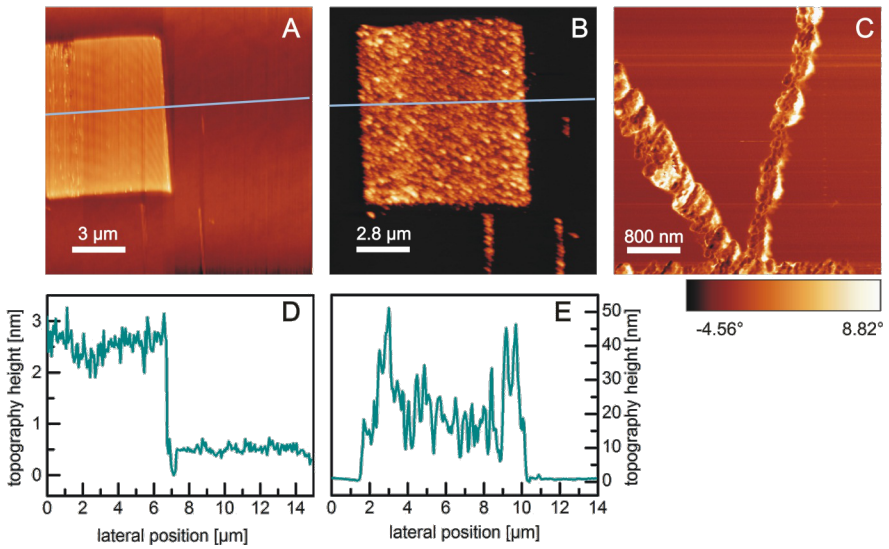

Figure 2: AFM height images of LAO oxide structures before (A) and after (B) 40 h in a 10 mM OTS solution in toluene under ambient conditions. (C) AFM phase image of line structures after 40 h in OTS. (D, E) topography profiles along the lines in A and B.

AFM measurements, and the resulting layer thickness and smoothness is an indicator of the packing density of the molecules [23,31]. Figure 2 shows the AFM images of a square oxide structure before (A) and after (B) 40 h in a 10 mM OTS solution in toluene under ambient conditions. After the OTS treatment, the measured structure height increases from about 2 nm up to 50 nm (Figure 2D and Figure 2E). Additionally this increase is very nonuniform: there is a formation of clusters of different height (Figure 2E) and the RMS roughness on the structure increases more than one order of magnitude (0.5 to 10). The lateral structure of the formed OTS layer can be observed in Figure 2C. There is a grainy substructure, which is in the range of some 10 to 100 nm, and therefore within a similar length scale as the measured topography height variations (Figure 2E). Therefore, it can be concluded that the OTS clusters possess a globular structure. As the cluster sizes are much larger than the chain length of an OTS molecule (2.6 nm), OTS evidently does not form closed monolayers on the nanostructure (Figure 3A), but rather a vertical or 3D polymerization of OTS has to be assumed (Figure 3B). For much shorter immersion times in the OTS solution (several minutes to a few hours), only a partial and also very nonuniform height increase of the LAO structures is observed. Obviously, without any special preparation conditions, the tendency towards a 3D polymerization is much stronger than that for the formation of smooth uniform layers.

A vertical polymerization is known for tri- and di-functional silanes, especially alkoxysilanes and alkylchlorosilanes [23,32,33] under certain conditions. In order to understand why 3D polymerization is preferred to monolayer formation on the LAO structure under ambient conditions, the reaction mechanism of trichlorosilanes with silica surfaces has to be considered in greater detail.

In the presence of water, the Si–Cl bonds of an OTS molecule undergo hydrolysis forming Si–OH groups. If the OTS molecule is close to the substrate, these Si–OH groups react with OH groups on the silicon oxide surface forming a Si–O–Si bond and water [20]. Such a reaction is of course also possible between two OTS molecules leading to a cross-linking of the layer molecules. Thus, the presence of a surface water layer on the silicon oxide is a necessary condition for the monolayer for-

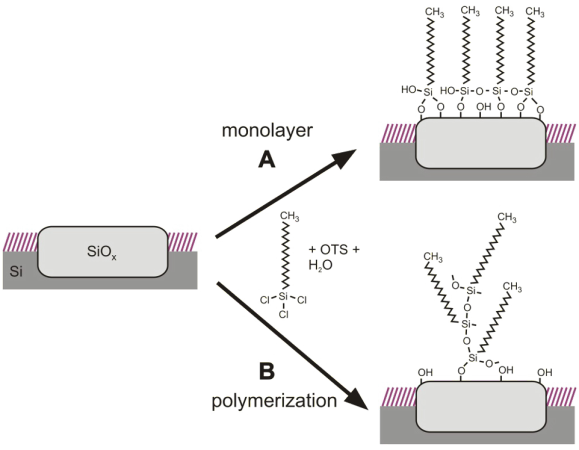

Figure 3: Schematic illustration of the two different possible covalent binding mechanisms of OTS with a silicon oxide surface. (A) formation of a smooth OTS monolayer, (B) vertical polymerization.

mation, but on the other hand too much water on the surface also promotes vertical polymerization. As the silicon oxide nanostructure is much more hydrophilic than the surrounding alkyl monolayer, the wetting of the structure is strongly favored. Recent investigations revealed that there is a pronounced water layer formation on top of the oxide structure due to ambient humidity, which takes place on the time scale of several hundred minutes [34].

Besides the availability of water, the surface density of OH groups (silanole groups) on the silica surface has a very strong impact on whether there is a monolayer formation or a vertical polymerization [23]. If the density of such groups is sufficiently large, the hydrolyzed OTS molecules should tend to react preferentially with the surface due to steric reasons. A 3D polymerization requires a certain twisting of the molecules out of their equilibrium conformation. In order to obtain high quality monolayers, silica substrates are usually cleaned in an oxidizing acid (e.g., "piranha" solution) to provide a maximum coverage of the surface with OH groups. However, such a method is not possible for LAO nanostructures on alkyl-terminated silicon as the protecting monolayer would also be oxidized. The strong tendency towards vertical polymerization in the presence of ambient water however, is indirect evidence for a rather low density of OH groups on the lithographic oxide. This is further supported by the fact that for shorter immersion times of the substrate within the OTS solution, also a very nonuniform increase in height with clear local differences was observed. The few surface hydroxy groups also lead to a much slower binding rate of OTS to the LAO oxide. Even after several hours in the OTS solution, the surface was not fully covered with OTS, whereas for "piranha"-cleaned silica a fractional surface coverage near unity was reported within a few minutes [29,35].

As the OH group density on the LAO oxide cannot be altered, the only way to prevent vertical polymerization is to completely remove the water from the silane solution in toluene. This was achieved by placing the vessel in a desiccator during reaction and further drying of the solvent (sparging of the solution with an inert gas for one hour before the reaction). Under such conditions, a much more uniform coverage of the silicon oxide nanostructures could be achieved (Figure 4). The small amount of water that is necessary for the hydrolysis of the trichlorosilane may originate from residual water in the dried solution or a very thin water film that is adsorbed on the oxide surface during transfer under ambient conditions. Since the rest of the alkyl-terminated silicon is much more hydrophobic than the local oxide, surface water will condense preferentially on the nanostructures.

Compared to the binding under ambient conditions, the height increase is much more uniform under dry conditions. In the case of a square structure, the formed silane layer (here: 11-bromoundecyltrichlorosilane) possesses a thickness of about (1.6 ± 0.1) to (2 ± 0.1) nm (Figure 4A, B and E). This is in good agreement with the thickness of 11-bromoundecyltrichlorosilane monolayers on silica substrates which was determined to a value of 1.81 nm by ellipsometry measurements [30]. There are also a few larger elevations, which may be due to polymerization already within the solution. Such impurities could be avoided by using better anhydrous conditions and controlled

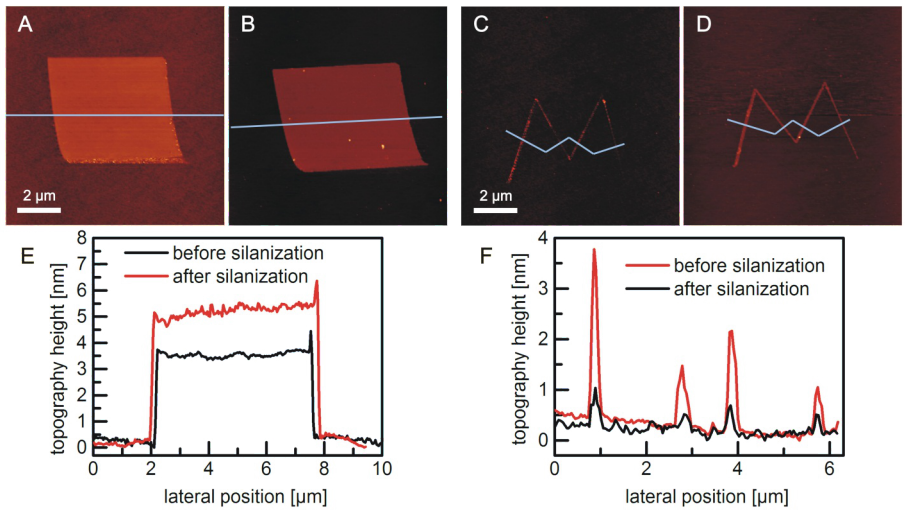

Figure 4: AFM height images of LAO oxide structures before (A, C) and after (B, D) 16 h in a 10 mM 11-bromoundecyltrichlorosilane solution in dry toluene under dry conditions. (E, F) topography profile along the lines in (A–D) averaged over a width of 10 pixels.

process conditions; however, they are not crucial for the general observations and their physicochemical interpretation. The height increase of the line structures (Figure 4C, D and F) is also within a reasonable range of (0.5 ± 0.1) to (2.5 ± 0.1) nm. Silane layer thicknesses that are smaller than the molecular length may be explained by a low surface coverage. If the monolayer is not densely packed, the molecules will not stand nearly upright (in all-trans conformation), but the alkyl chains will orient more parallel to the surface, which leads to a decrease of the measured layer thickness [29]. A height increase of more than the thickness of a densely packed monolayer, on the other hand, is a strong indication for the beginning vertical polymerization. It is striking that there seems to be a correlation between initial structure size and the thickness of the silane layer: the higher and wider the initial oxide, the larger is also the height increase by the binding of silane molecules. Whether this is caused by either an influence of the surface geometry on the reaction mechanism/the molecular order or possibly an artefact of the AFM measurement (water adsorption or the energy dissipation of the AFM tip) should be investigated in detail in further statistical studies.

Covalent binding of FITC to silicon oxide nanostructures

Once suitable conditions for the controlled silanization of LAO nanostructures had been found, functionalization with fluorescent FITC molecules was carried out according to the route displayed in Figure 1. Figure 5 shows the results of comparative AFM height measurements after different functionalization steps.

After binding of APTES, there is a height increase of approximately 0.3 ± 0.1 nm (blue dashed curve), which is in good agreement with the length of a hydrolyzed APTES molecule (0.3 nm). The subsequent binding of FITC to the amino-functionalized nanostructure leads to a further increase in height of about 0.7 ± 0.1 nm. As the length of a single FITC molecule is about 1 nm, it can be concluded that the molecules are not standing upright but are tilted from the surface normal. The tilt is most likely caused by the thiocarbonyl group (planar trigonal geometry due to binding sp^2 orbitals). A calculation of the tilt from the height measurements is not straightforward though, as too many assumptions about the 3D orientation and bending of the bound FITC molecules would have to be made. Further theoretical considerations (e.g., molecular dynamics simulations) could be employed in order to investigate the molecular orientation on the surface, but it has to be kept in mind that AFM measurements of such systems may also be influenced by many other factors, such as tip–sample interactions and the formation of water layers, that strongly depend on the chemical nature of the surface.

FITC functionalized nanostructures have been investigated using fluorescence microscopy, which further confirms the successful binding. Figure 6 shows a typical fluorescence spectrum from the FITC-terminated structure depicted in Figure 5. A confocal microscope image of the whole structure is displayed in the inset. There is a clear luminescence signal, which originates from the nanostructure only. The spectral shape of the fluorescence light resembles that of FITC in acetone solution and is shifted about 36 nm to the red due to the

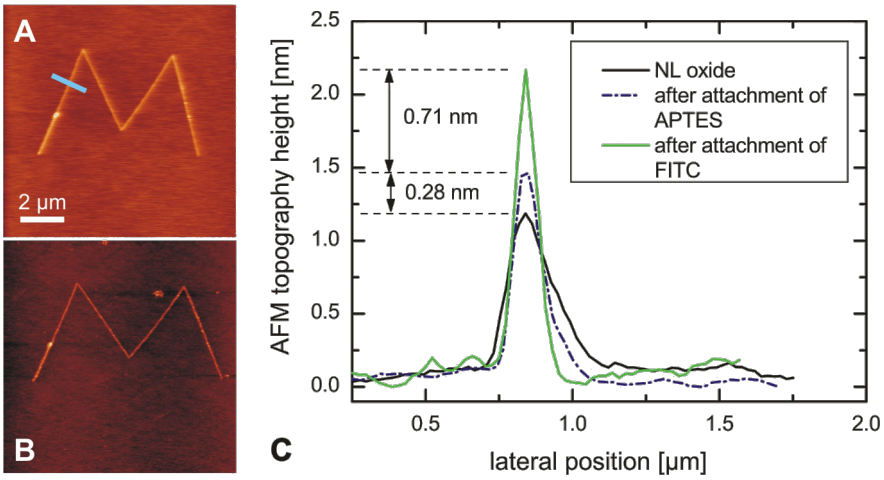

Figure 5: (A) and (B) AFM height images of LAO oxide line structures before any functionalization (A) and after binding of FITC (B). (C) Height profile of an LAO oxide structure (black curve) after 6 h in 10 mM APTES solution in dry toluene under dry conditions (blue dashed curve) and after subsequent binding of FITC in aqueous solution (green curve). The profiles were measured and averaged over a width of 20 pixels along the path indicated by the blue line in Figure (A).

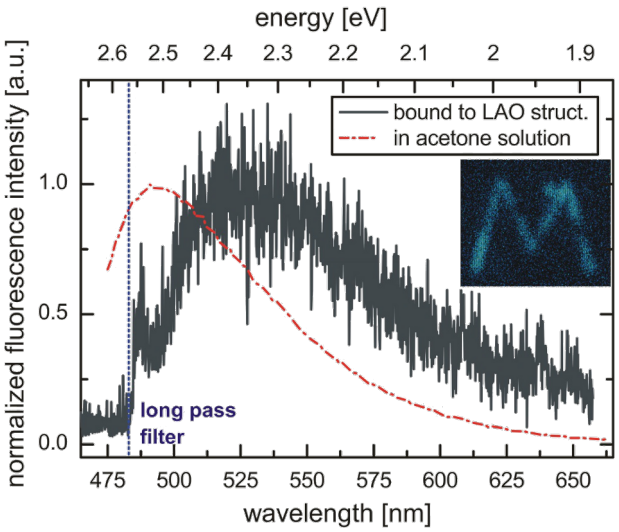

Figure 6: Fluorescence spectrum from the FITC-functionalized LAO oxide nanostructure shown in Figure 5 (dark grey curve) compared to FITC fluorescence in acetone solution (red dashed curve). The inset shows a confocal microscope image of the structure (excitation wavelength: 465 nm, excitation power: 11 μW, integration time: 1 ms/px).

different dielectric environment (the small signal close to the filter edge at 485 nm is not part of the FITC fluorescence but probably scattered light from the environment).

The fluorescence signal is comparably weak (roughly a factor of 10 above the background noise level). The main reason for this is the quenching by the underlying silicon. Nevertheless, other high-quantum-yield xanthene dyes (e.g., rhodamine 6G) that are bound to the nanostructures by electrostatic interactions, show a much higher signal-to-noise ratio for similar structure heights and excitation powers [13]. Thus, the quenching by the underlying silicon does not explain the measured low fluorescence intensities. Another important factor that influences the absorption and emission of radiation by the bound molecules is the orientation of their transition dipole moment. For xanthene derivatives, such as fluorescein or rhodamine, the transition dipole moment $\vec{\mu}$ for the $S_1 \rightarrow S_0$ transition is typically oriented along the xanthene unit [36,37]. As this unit is perpendicular to the binding axis of the FITC molecule, it should be oriented rather parallel to the surface (tilt angle $\alpha = 0°$, Figure 7). Since the directional characteristic of a dipole scales with $\cos^2\alpha$, a very large tilting (α close to 90°) would be necessary to explain the much lower fluorescence intensities. Such a large tilt is very unlikely due to steric reasons as well as the molecular structure and would also be contradictory to the AFM-investigations.

It may also be possible that the fluorescence is quenched by an efficient energy transfer between densely packed FITC molecules on the structure. However, such an effect can be excluded by comparison with the study of Imhof et al. [38]. They investigated FITC bound to silica spheres and found a quenching of the fluorescence intensity for increasing dye concentration, which was attributed to an interaction between neighboring molecules. This interaction also causes a lowering of the excited

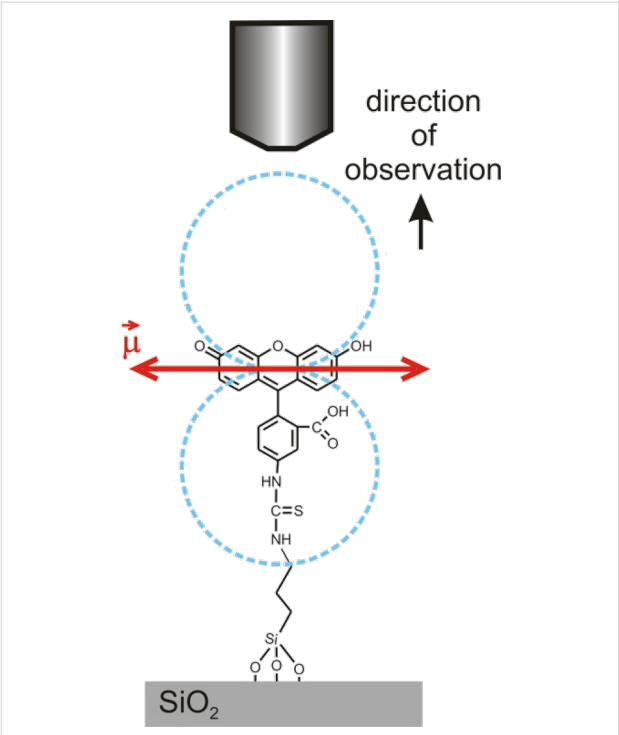

Figure 7: Schematic of the transition dipole moment orientation $\vec{\mu}$ (red arrow) of FITC bound to a SiO_2 surface.

state energy, which was observed as a 10 nm red shift of the emission spectrum from 531 nm to 541 nm for increasing dye concentration. Although it is difficult to exactly determine the peak position in Figure 6 due to the low signal-to-noise ratio of the spectrum, the maximum position clearly seems to be rather in the range of 530 nm than 540 nm. The absence of the red-shift as well as the agreement of the approximate peak position and spectral shape with the results of Imhof et al. for low dye concentrations, leads to the conclusion that the observed low fluorescence intensity is likely not caused by a dense packing of the FITC molecules on the oxide structure. Measurements of the excited state lifetime are likely not suitable for the investigation of the intermolecular interactions due to the predominant quenching by the silicon below. Consequently, the most likely reason for the strong difference in fluorescence intensity between the electrostatically bound dyes and the covalently bound FITC is a much lower surface density in the case of covalently bound FITC.

Of course this conclusion, though plausible, is based on many uncertainties and assumptions. Yet a more precise quantification is very difficult due to the unknown quenching by the underlying silicon and the unavailability of suitable measurement techniques of the surface coverage for such small geometries. If the density of FITC molecules on the structures is really much lower than that of electrostatically attracted dyes then this may be a hint that either the APTES monolayer is not closed on a nanoscopic scale or the molecules in the monolayer are not well-ordered. The latter would lead to a rather low density of surface amino groups [39] that are available for reaction with the functional group of FITC. However, the next-neighbor distance between two adjacent surface amino groups must be below the AFM tip diameter (≈30 nm) as no distinct height steps or islands could be observed in the AFM images. A resolution of the density of bound molecules may be possible by using novel techniques such as the measurement of AFM amplitude–phase–distance curves [40]. From such experiments, the dissipated energy of the AFM tip oscillation can be calculated, which depends on the local elastic and therefore structural surface properties of the substrate. The surface coverage of the relatively rigid silicon oxide with "softer" organic molecules should in principle be distinguishable by the amplitude–phase–distance curve technique.

Conclusion

In conclusion, a route for a controlled covalent functionalization of silicon oxide nanostructures with an amino-terminated silane and FITC dye molecules has been successfully realized. The formation of silane monolayers on nanoscopic silicon oxide nanostructures has proven to be much more sensitive towards ambient humidity than, e.g., silanization of larger OH-terminated silica surfaces. This is most likely due to a lower density of surface hydroxy groups, which requires a longer reaction time for the formation of a closed monolayer. Optical investigations of the bound FITC dye, on the other hand, seem to indicate that the forming APTES monolayer is not well-ordered on a scale below 20 nm. Pasternack et al. found that the ordering of APTES monolayers can be improved through in situ heating during the functionalization reaction [39]. An elevation of the reaction temperature seems to be a promising step for future experiments in order to obtain more densely packed FITC monolayers on the silicon oxide nanostructures. Also a formation of silane monolayers from the gas phase should be considered, as this technique enables a controlled formation of highly ordered films under appropriate conditions [41-43].

Experimental

The local anodic oxidation experiments were carried out on 1×1 cm^2 pieces of weakly doped n-type silicon (resistivity: >3000 Ω·cm, Fraunhofer ENAS, Chemnitz, Germany) substrates with a self-assembled monolayer (SAM) of 1-dodecene molecules as a surface coating.

The monolayer preparation was carried out in three steps: cleaning, native-oxide removal, and monolayer formation (radical chain reaction); see also [25,44]. The substrates have been cleaned by ultrasonication in acetone (C_3H_6O, "Uvasol© for spectroscopy", Merck, Germany), ethanol (C_2H_5OH, "Uvasol© for spectroscopy", Merck, Germany) and in "piranha"-solution [40% hydrogen peroxide (H_2O_2, 30% "Suprapur", Merck, Germany) and 60% sulfuric acid (H_2SO_4, 96% "Suprapur", Merck, Germany)] at 70 °C. Afterwards the samples were rinsed thoroughly with Millipore™ (ultraclean demineralized and deionized water, resistivity of > 18 MΩ·cm) and dried in a nitrogen stream. In the next step the silicon oxide samples were etched in aqueous hydrofluoric acid (HF, 40% "Suprapur", Merck, Germany) solution (3–4% in volume) for 3 min at room temperature in order to remove the native silicon oxide. Subsequently, organo-silicon monolayers were prepared according to [45] by using 1-dodecene ($C_{12}H_{25}$, for synthesis, Merck, Germany) in pure form. The alkene was deoxygenated with argon gas at least 30 min prior to the reaction and the argon flow was also maintained during the reaction. After transferring the etched substrates to the 1-dodecene solution, it was heated up to 190 °C for at least 7 h. Afterwards the samples were rinsed and sonicated for 5 min in dichloromethane (CH_2Cl_2, "Uvasol© for spectroscopy", Merck, Germany) and ethanol followed by drying in a nitrogen stream.

LAO experiments as well as topography measurements were conducted with an Anfatec Level AFM (Anfatec Instruments AG, Germany) by using platinum-coated silicon tips ("NSC18/

Pt", resonance frequency 75 Hz, spring constant 3.5 N/m, Mikromash, Estonia). As the usage of conductive ink or tape should be omitted in order to prevent contamination during the following wet-chemical steps, the substrates were fixed on a metal sample holder with a small magnet (thus providing a back-side contact to the substrates, which is sufficient for LAO). The sample holder was electrically contacted to a voltage source by conductive ink. The oxide structures were generated in contact mode operation with a voltage in the range of −8 to −11 V applied between tip and substrate at a relative humidity (RH) of around 65% and a writing velocity of about 1 μm/s. To control the humidity of the ambient atmosphere the microscope was placed under a closed PMMA dome, which can be purged with dry nitrogen or water-vapor-saturated nitrogen. The humidity was measured by using a "SHT15" digital humidity sensor (Sensirion AG, Switzerland). Lithography of diverse patterns was performed by application of different software protocols, which were written by using a homemade software user interface. For comparative height measurements, the driving frequency, driving amplitude, setpoint and the AFM tip were kept implicitly constant, as the tip–sample interaction and therefore the measured topography depends strongly on these parameters for noncontact AFM operation.

Binding of OTS and 11-bromoundecyltrichlorosilane (both ABCR, Germany) was carried out by immersion of the structures into a 10 mM solution of the silane in toluene (spectroscopic grade, Merck, Germany). After a specific amount of time, the sample was removed from the solution and rinsed thoroughly in ultrasonic baths of toluene, dichloromethane and ethanol (all were spectroscopic grade, Merck, Germany). APTES (aminopropyltriethoxysilane, Sigma Aldrich, USA) was bound by immersion under the same conditions and preparation steps as the trichlorosilanes. For sparging of the silane/toluene solutions, argon was used as the inert gas. Binding of FITC to the APTES-functionalized structures was achieved by immersion of the samples in a low-millimolar solution of FITC in acetone (spectroscopic grade, Merck, Germany) for 90 min and subsequent cleaning in ultrasonic baths of acetone, dichloromethane and ethanol. All silanization experiments were carried out at room temperature (21 °C).

Fluorescence investigations of the samples were performed with a home-built microscope setup. The 465 nm excitation light from a pulsed laser diode ("LDH-P-C-470", Picoquant GmbH) with adjustable repetition rate and a narrow pulse length of 75 ps, is focused on the sample by using an objective lens (100×, NA = 0.9, "EC Epiplan Neofluar" Carl Zeiss, Germany). The fluorescence light from the sample is collected with the same objective and separated from the reflected excitation light by a dichroic mirror ("z 470 RDC", AHF Analysetechnik GmbH, Germany) and a fluorescence filter at 480 nm (Omega Optics Inc., USA). Using a beam splitter one part of the fluorescence light is focused on an avalanche photodiode ("SPCM-AQR-14", Perkin Elmer, USA), the other part is coupled into a spectrometer ("Shamrock SR-163/SR1-GTR-600-500", Andor Technology, UK) with a thermoelectrically cooled CCD camera ("Newton DU971N-BV", Andor Technology, UK) as detector. The detection range of the spectrometer was adjusted to 480–655 nm. The triggering of the position controller of the piezo scan-stage as well as the read-out of the APD signal is realized by home-written software.

Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, GR 2695/4). The authors would like to thank Prof. Hirokazu Tada and Dr. Masato Ara (Osaka University) for stimulating ideas and fruitful discussions.

References

1. Snow, E. S.; Campbell, P. M. *Appl. Phys. Lett.* **1994,** *64,* 1932–1934. doi:10.1063/1.111746
2. Dagata, J. A. *Science* **1995,** *270,* 1625–1626. doi:10.1126/science.270.5242.1625
3. Avouris, P.; Hertel, T.; Martel, R. *Appl. Phys. Lett.* **1997,** *71,* 285–287. doi:10.1063/1.119521
4. Calleja, M.; García, R. *Appl. Phys. Lett.* **2000,** *76,* 3427–3429. doi:10.1063/1.126856
5. Wouters, D.; Hoeppener, S.; Schubert, U. S. *Angew. Chem., Int. Ed.* **2009,** *48,* 1732–1739. doi:10.1002/anie.200801013
6. Ravoo, B. J. *J. Mater. Chem.* **2009,** *19,* 8902–8906. doi:10.1039/b908564e
7. Gu, J.; Yam, C. M.; Li, S.; Cai, C. *J. Am. Chem. Soc.* **2004,** *126,* 8098–8099. doi:10.1021/ja048405x
8. Wouters, D.; Schubert, U. S. *Langmuir* **2003,** *19,* 9033–9038. doi:10.1021/la034711x
9. Sugimura, H.; Nanjo, S.; Sano, H.; Murase, K. *J. Phys. Chem. C* **2009,** *113,* 11643–11646. doi:10.1021/jp901470q
10. Mo, Y.; Wang, Y.; Pu, J.; Bai, M. *Langmuir* **2009,** *25,* 40–42. doi:10.1021/la803379e
11. Graaf, H.; Vieluf, M.; von Borczyskowski, C. *Nanotechnology* **2007,** *18,* 265306. doi:10.1088/0957-4484/18/26/265306
12. Hoeppener, S.; Schubert, U. S. *Small* **2005,** *1,* 628–632. doi:10.1002/smll.200500017
13. Baumgärtel, T.; von Borczyskowski, C.; Graaf, H. *Nanotechnology* **2010,** *21,* 475205. doi:10.1088/0957-4484/21/47/475205
14. Maoz, R.; Cohen, S. R.; Sagiv, J. *Adv. Mater.* **1999,** *11,* 55–61. doi:10.1002/(SICI)1521-4095(199901)11:1<55::AID-ADMA55>3.0.CO;2-8
15. Haensch, C.; Hoeppener, S.; Schubert, U. S. *Nanotechnology* **2009,** *20,* 135302. doi:10.1088/0957-4484/20/13/135302
16. Fresco, Z. M.; Fréchet, J. M. J. *J. Am. Chem. Soc.* **2005,** *127,* 8302–8303. doi:10.1021/ja052738s
17. Yang, M.; Wouters, D.; Giesbers, M.; Schubert, U. S.; Zuilhof, H. *ACS Nano* **2009,** *3,* 2887. doi:10.1021/nn9007059

18. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001,** *40,* 2004–2021. doi:10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.3.CO;2-X
19. Fabre, B.; Herrier, C. *RSC Adv.* **2012,** *2,* 168–175. doi:10.1039/c1ra00450f
20. Sagiv, J. *J. Am. Chem. Soc.* **1980,** *102,* 92–98. doi:10.1021/ja00521a016
21. Howarter, J. A.; Youngblood, J. P. *Langmuir* **2006,** *22,* 11142–11147. doi:10.1021/la061240g
22. Ulman, A. *Chem. Rev.* **1996,** *96,* 1533–1554. doi:10.1021/cr9502357
23. Fadeev, A. Y.; McCarthy, T. J. *Langmuir* **2000,** *16,* 7268–7272. doi:10.1021/la000471z
24. Desbief, S.; Patrone, L.; Goguenheim, D.; Guérin, D.; Vuillaume, D. *Phys. Chem. Chem. Phys.* **2011,** *13,* 2870–2879. doi:10.1039/c0cp01382j
25. Ara, M.; Graaf, H.; Tada, H. *Appl. Phys. Lett.* **2002,** *80,* 2565–2567. doi:10.1063/1.1467973
26. Baumgärtel, T.; Graaf, H.; von Borczyskowski, C.; Ara, M.; Tada, H. In *Proceedings of International conference Nanomeeting 2011, Reviews and short notes,* Physics, Chemistry and applications of nanostructures, Minsk, Belarus, May 24–27, 2011; Borisenko, V. E.; Gaponenko, S. V.; Gurin, V. S.; Kam, C. H., Eds.; World Scientific Publishing: New Jersey, London, Singapore, Bejing, Shanghai, Hong-Kong, Taipei, Chennai; pp 349–352.
27. Ara, M.; Tsuji, M.; Tada, H. *Surf. Sci.* **2007,** *601,* 5098–5102. doi:10.1016/j.susc.2007.04.213
28. Tada, H.; Ara, M.; Tanaka, S. *Mater. Res. Soc. Symp. Proc.* **2002,** *739,* H7.37. doi:10.1557/PROC-739-H7.37
29. Mirji, S. A. *Surf. Interface Anal.* **2006,** *38,* 158–165. doi:10.1002/sia.2309
30. Sawoo, S.; Dutta, P.; Chakraborty, A.; Mukhopadhyay, R.; Bouloussa, O.; Sarkar, A. *Chem. Commun.* **2008,** 5957–5959. doi:10.1039/b813296h
31. Wang, M.; Liechti, K. M.; Wang, Q.; White, J. M. *Langmuir* **2005,** *21,* 1848–1857. doi:10.1021/la048483y
32. Brandriss, S.; Margel, S. *Langmuir* **1993,** *9,* 1232–1240. doi:10.1021/la00029a014
33. Plueddemann, E. *Silane Coupling Agents,* 2nd ed.; Springer: Berlin, 1991.
34. Baumgärtel, T.; von Borczyskowski, C.; Graaf, H. *Nanotechnology* **2012,** *23,* 095707. doi:10.1088/0957-4484/23/9/095707
35. Belgardt, C.; Graaf, H.; Baumgärtel, T.; von Borczyskowski, C. *Phys. Status Solidi C* **2010,** *7,* 227–230. doi:10.1002/pssc.200982471
36. Penzkofer, A.; Wiedmann, J. *Opt. Commun.* **1980,** *35,* 81–86. doi:10.1016/0030-4018(80)90364-8
37. Arbeloa, F. L.; Martínez, V. M. *Chem. Mater.* **2006,** *18,* 1407–1416. doi:10.1021/cm051518a
38. Imhof, A.; Megens, M.; Engelberts, J. J.; de Lang, D. T. N.; Sprik, R.; Vos, W. L. *J. Phys. Chem. B* **1999,** *103,* 1408–1415. doi:10.1021/jp983241q
39. Pasternack, R. M.; Amy, S. R.; Chabal, Y. J. *Langmuir* **2008,** *24,* 12963–12971. doi:10.1021/la8024827
40. Dietz, C.; Zerson, M.; Riesch, C.; Franke, M.; Magerle, R. *Macromolecules* **2008,** *41,* 9259–9266. doi:10.1021/ma801236p
41. Dong, J.; Wang, A.; Ng, K. Y. S.; Mao, G. *Thin Solid Films* **2006,** *515,* 2116–2122. doi:10.1016/j.tsf.2006.07.041
42. Sugimura, H.; Hozumi, A.; Kameyama, T.; Takai, O. *Surf. Interface Anal.* **2002,** *34,* 550–554. doi:10.1002/sia.1358
43. Hozumi, A.; Yokogawa, Y.; Kameyama, T.; Sugimura, H.; Hayashi, K.; Shirayama, H.; Takai, O. *J. Vac. Sci. Technol., A* **2001,** *19,* 1812–1816. doi:10.1116/1.1336833
44. Graaf, H.; Baumgärtel, T.; Vieluf, M.; von Borczyskowski, C. *Superlattices Microstruct.* **2008,** *44,* 402–410. doi:10.1016/j.spmi.2008.01.015
45. Linford, M. R.; Fenter, P.; Eisenberger, P. M.; Chidsey, C. E. D. *J. Am. Chem. Soc.* **1995,** *117,* 3145–3155. doi:10.1021/ja00116a019

License and Terms

Beilstein Journal of Nanotechnology

Determining cantilever stiffness from thermal noise

Jannis Lübbe[1], Matthias Temmen[1], Philipp Rahe[2,3], Angelika Kühnle[2] and Michael Reichling[*1]

Full Research Paper

Open Access

Address:
[1]Fachbereich Physik, Universität Osnabrück, Barbarastraße 7, 49076 Osnabrück, Germany, [2]Institut für Physikalische Chemie, Johannes Gutenberg-Universität Mainz, Duesbergweg 10-14, 55099 Mainz, Germany and [3]now at: Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA

Email:
Michael Reichling[*] - reichling@uos.de

* Corresponding author

Keywords:
AFM; cantilever; noncontact atomic force microscopy (NC-AFM); Q-factor; thermal excitation; resonance; spectral analysis; stiffness

Beilstein J. Nanotechnol. **2013,** *4,* 227–233.
doi:10.3762/bjnano.4.23

Received: 13 January 2013
Accepted: 28 February 2013
Published: 28 March 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We critically discuss the extraction of intrinsic cantilever properties, namely eigenfrequency f_n, quality factor Q_n and specifically the stiffness k_n of the nth cantilever oscillation mode from thermal noise by an analysis of the power spectral density of displacement fluctuations of the cantilever in contact with a thermal bath. The practical applicability of this approach is demonstrated for several cantilevers with eigenfrequencies ranging from 50 kHz to 2 MHz. As such an analysis requires a sophisticated spectral analysis, we introduce a new method to determine k_n from a spectral analysis of the demodulated oscillation signal of the excited cantilever that can be performed in the frequency range of 10 Hz to 1 kHz regardless of the eigenfrequency of the cantilever. We demonstrate that the latter method is in particular useful for noncontact atomic force microscopy (NC-AFM) where the required simple instrumentation for spectral analysis is available in most experimental systems.

Introduction

Noise as a result of thermal fluctuations is a ubiquitous phenomenon present in any physical system kept at a finite temperature. The seminal work of Nyquist established the simple framework of thermodynamic considerations for a quantitative description of such noise for a resistor kept at a temperature T and connected to an electrical network, as an example of a dynamic system in equilibrium with a thermal bath [1]. At the same time, it was pointed out by Johnson that such understanding is of great practical relevance as it allows for an optimisation of critical electronic devices with respect to their noise figures [2]. The main conclusion from this work is that a thermal bath provides a source of excitation with a strength that is constant over the entire frequency range, while the strength and spectral characteristics of the system response depends

solely on the system transfer function. According to the equipartition theorem, the energy transferred from the thermal bath to a dynamic system equals $(1/2)k_BT$ for each degree of freedom, where k_B is the Boltzmann constant.

A cantilever is a mechanical dynamic system that is often described as a simple harmonic oscillator with a response function dominated by resonances at the eigenfrequencies f_n of the flexural cantilever oscillation modes. Each of these modes represents a mechanical degree of freedom extracting $(1/2)k_BT$ of energy if connected to a thermal bath. The corresponding response to thermal excitation, namely the resulting noise power spectral density of the cantilever displacement $(d_{\mathrm{th}}^z)^2$, is the superposition of contributions from all modes and can be derived within the framework of the Nyquist theory [3]. Provided the simple harmonic oscillator model is valid, i.e., the internal damping of the cantilever is small, $(d_{\mathrm{th}}^z)^2$ is given by

$$\left(d_{\mathrm{th}}^z(f)\right)^2 = \sum_{n=0}^{\infty} \frac{2k_BT/\left(\pi k_n f_n Q_n\right)}{\left(f^2/f_n^2 - 1\right)^2 + \left(f/\left(f_n Q_n\right)\right)^2} \qquad (1)$$

where k_n and Q_n are the modal stiffness [4] and Q-factor of the nth cantilever eigenmode [5], respectively. The relation is of relevance for any practical application involving microcantilevers and specifically important for high-resolution noncontact atomic force microscopy (NC-AFM) based on cantilever or tuning fork force sensors. We recently demonstrated how Equation 1 defines the ultimate limit of signal detection for an NC-AFM measurement performed under ultrahigh vacuum (UHV) conditions [6]. Although most NC-AFM systems are operated with cantilevers excited near their fundamental eigenfrequency f_0, higher eigenmodes [7] have been investigated in the context of noise analysis [8], and it has been debated whether the thermal noise limitations in NC-AFM measurements could be reduced by operating cantilevers at higher eigenmodes [9]. It has further been realised that the cantilever properties f_n, k_n and Q_n appear as linearly independent parameters in Equation 1. This allows their independent determination from a single measurement of the displacement noise spectral density d_{th}^z over a limited spectral range around the resonance for a cantilever kept at a known temperature [10]. A practical implementation of this notion, focused on the determination of cantilever stiffness from thermal noise, demonstrated the validity of the approach by a critical comparison of the results against corresponding results from other methods [11].

While the properties f_n and Q_n can quite easily be determined with high precision by a cantilever excitation experiment [12], the thermal method discussed here is hitherto the only one to yield valid results for the modal stiffness k_n. The strength of the thermal method is that it is solely based on the equipartition theorem, establishing the simple energy balance [6]

$$\frac{k}{2}\sum_{n=0}^{\infty}\left\langle z_n^2(t)\right\rangle = k_BT\frac{k}{2}\sum_{n=0}^{\infty}\frac{1}{k_n} \qquad (2)$$

with k being the static stiffness of the cantilever.

This implies that a precise measurement of the mean square displacement $\left\langle z_n^2(t)\right\rangle$ or the corresponding power spectral density $(d_{\mathrm{th},n}^z(f))^2$ in a region around a specific resonance n allows the determination of the modal stiffness k_n without the knowledge of any other cantilever parameters such as dimensions, shape, mass or mass distribution.

Here, we critically discuss the extraction of intrinsic cantilever properties from measurements of thermal noise and focus on the precise determination of the modal stiffness k_n as this quantity is a prerequisite for the quantitative interpretation of force imaging and spectroscopy results [13-15]. Most examples are given for the fundamental mode of the cantilever oscillation, but we also demonstrate that the method is universal and can equally well be applied to higher oscillation modes. The acquisition of noise spectra is, however, not trivial in this context as intrinsic Q-factors of the fundamental mode of high-Q cantilevers may be as high as 200,000. Thus, the accurate spectral analysis of the extremely narrow resonance peak requires expensive test equipment. Therefore, we introduce an alternative method of determining the modal stiffness by using the demodulator of an NC-AFM system to project the noise power of an excited cantilever around its resonance into the frequency range of 10 Hz to 1 kHz. Processing the resulting frequency shift signal $\Delta f(t)$ to obtain the modal stiffness in this frequency range is straightforward as a spectral analysis can be performed with simple equipment available in most NC-AFM control systems.

Experimental

Measurements are performed in two UHV systems with NC-AFM instruments based on the optical beam deflection configuration. These have been described in our previous work as system B (UHV VT AFM/STM, Omicron NanoTechnology GmbH, Taunusstein, Germany) and as system C (UHV 750 variable temperature STM/AFM, RHK Technology, Inc., Troy, MI, USA) [6]. Temperatures used for data analysis are measured directly at the NC-AFM stage of a thermally equilibrated experimental system. It is, therefore, expected that the measured temperature is identical to the cantilever temperature. As test objects, we use a selection of four cantilever types with commercial names FM, NCH, Arrow™ and NCVH (Nanoworld

AG, Neuchâtel, Switzerland). These cantilevers are chosen to cover a broad range of eigenfrequencies f_0 ranging from 50 kHz to 2 MHz, static stiffness k [16] ranging from 3 to 120 N/m, and Q-factors Q_0 [12] covering the range of 20,000 to 120,000; details are provided in Table 1. Measurements of the total displacement noise spectral density $d^z_{\mathrm{tot}}(f)$ are performed by using a spectrum analyser connected to the output of the preamplifier of the position-sensitive detector of the NC-AFM instrument. The cantilever displacement is measured as a calibrated electrical signal $V_z(t)$ and processed by the spectrum analyser [6]. For measurements of the total noise spectral density $d^{\Delta f}_{\mathrm{tot}}(f_m)$ in the frequency shift signal $\Delta f(t)$, the spectrum analyser is connected to the phase-locked-loop (PLL) demodulator output of the respective NC-AFM system. In all of these experiments, utmost care has to be taken to shield the NC-AFM system from mechanical and, specifically, from electric noise in spectral regions encompassing the cantilever eigenfrequencies. Otherwise measurements may be severely false due to nonthermal noise contributions. Furthermore, valid results using this methodology can only be expected for thermal noise-limited measurements performed with a system for which the PLL transfer function is known. The former condition requires the detection system noise floor d^z_{ds} to be so low that, at least over a significant fraction of the PLL demodulator bandwidth, the frequency shift noise spectral density $d^{\Delta f}_{\mathrm{ds}}(f_m)$ of the detection system is negligible compared to the thermal frequency-shift noise spectral density $d^{\Delta f}_{\mathrm{th}}(f_m)$ [6].

Results and Discussion

Stiffness from displacement thermal noise

In a displacement noise measurement of a cantilever with a high Q-factor, the spectrum analyser measures the total displacement noise spectral density $d^z_{\mathrm{tot},n}(f)$ for the nth cantilever oscillation mode, which can be represented as [6]

$$d^z_{\mathrm{tot},n}(f) = \sqrt{\frac{2k_{\mathrm{B}}T/(\pi k_n f_n Q_n)}{\left(f^2/f_n^2 - 1\right)^2 + \left(f/(f_n Q_n)\right)^2} + \left(d^z_{\mathrm{ds},n}\right)^2}\,. \quad (3)$$

An exemplary spectrum of cantilever V 4 (see Table 1) covering the frequency region around the fundamental resonance at $f_0 \approx 284$ kHz is shown in Figure 1. The cantilever properties are extracted from the displacement noise spectrum by applying Equation 3. In the first step, the essentially white detection-system noise floor $d^z_{\mathrm{ds},n}$ of the nth mode is determined by averaging the spectral density off resonance (see Figure 1a). In the second step, Equation 3 is fitted to the data with the cantilever properties f_0, k_0^{th} and Q_0^{th} as fitting parameters and $d^z_{\mathrm{ds},0}$ as determined in the first step (see Figure 1b). Respective measurements have been performed for many cantilevers with some results compiled in Table 1, together with information on cantilever dimensions and properties measured by other techniques. Thermal noise analysis of cantilever V 4, for instance, yields $f_0^{\mathrm{th}} = 283{,}616$ Hz, $Q_0^{\mathrm{th}} = 28{,}400$ and $k_0^{\mathrm{th}} = 22$ N/m. As a consistency check, we measure the cantilever response to excitation in the vicinity of the resonance, where the corresponding results for the amplitude and phase response are shown in Figure 2. A fit of the simple harmonic oscillator model to the amplitude response (Equation 3 in [12]) yields $f_0^{\mathrm{exc}} = 283{,}620$ Hz and $Q_0^{\mathrm{exc}} = 28{,}600$ in excellent agreement with the thermal noise analysis. Generally, the fit of the thermal noise model from Equation 3 to the measured thermal excitation displacement data is excellent. In terms of experimental uncertainties, the highest precision is obtained in determining the eigenfrequency. State-of-the-art test equipment provides an accuracy of absolute frequency measurements below 1 ppm. However, practically the reproducibility is limited by thermal drift of the cantilever resonance between repeated measure-

Table 1: Synopsis of cantilever properties. Cantilever dimensions are the length l (±2.5 µm), mean width $\bar{w}$ (±1.5 µm) and thickness t (±0.2 µm) as provided by the manufacturer. The stiffness k_{dim} is calculated from the cantilever dimensions, while k_{stat} is determined by a precision measurement of the static stiffness [16]. Eigenfrequency f_0^{exc} (standard deviation below 1 ppm) and quality factor Q_0^{exc} (standard deviation below 1%) are obtained from a fit of the simple harmonic oscillator transfer function to the measured resonance curve of the excited cantilever [12]. Q_0^{th} and k_0^{th} are the properties yielded when fitting Equation 3 to the displacement noise spectral density d^z_{tot} of a thermally excited cantilever. The value $k_0^{\Delta f}$ is extracted from the frequency shift noise $d^{\Delta f}_{\mathrm{tot},0}$ from Equation 6 by using f_0^{exc} and Q_0^{exc} as known parameters.

cant.	l (µm)	$\bar{w}$ (µm)	t (µm)	f_0^{exc} (Hz)	Q_0^{exc}	Q_0^{th}	k_{dim} (N/m)	k_{stat} (N/m)	k_0^{th} (N/m)	$k_0^{\Delta f}$ (N/m)
P 5	224	30	3.0	68,319	97,500	105,300	3.0 ± 0.9	2.73 ± 0.14	2.9	3.4
D 5	229	30	2.9	68,353	118,000	123,000	2.5 ± 0.8	2.50 ± 0.13	2.7	2.9
V 4	125	26	3.8	283,620	28,600	28,400	31 ± 6	25.2 ± 1.3	22	21
V 15	125	26	3.7	279,451	47,200	46,300	29 ± 6	—	24.3	22
AF 11	125	34	4.1	311,476	37,700	—	50 ± 13	44.6 ± 2.3	—	61
AL 3	—	—	—	1,316,757	16,600	—	9 ± 3[a]	—	—	8.7
AP 5	40	24	2.0	1,996,199	32,400	—	130 ± 50	—	—	125

[a]Value provided by the manufacturer.

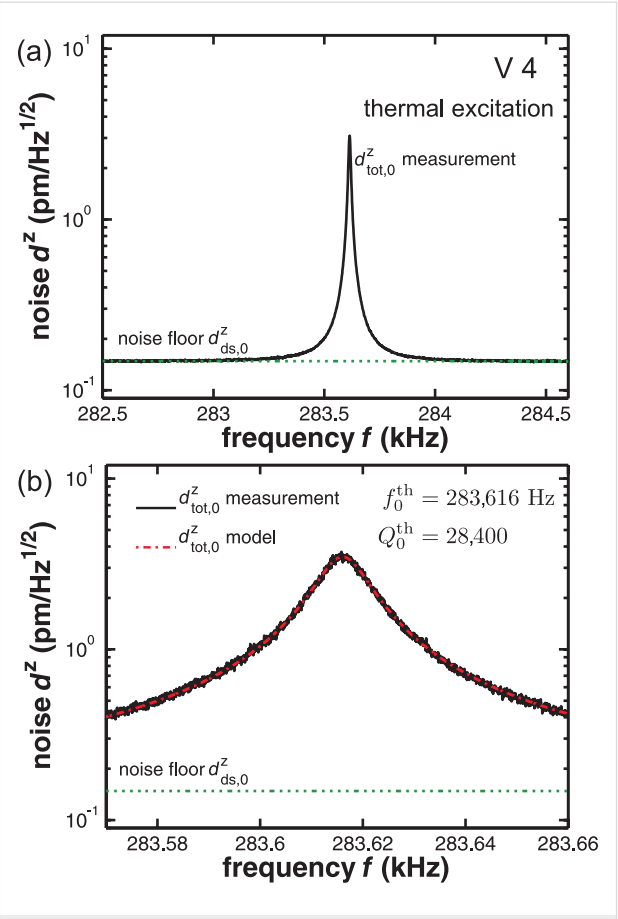

Figure 1: Displacement noise spectral density $d^z_{tot,0}$ measured for the fundamental mode of cantilever V 4. Measurements represent the average of 50,000 spectra. (a) Measurement yielding the detection-system noise floor $d^z_{ds,0}$ (dotted line). (b) Determination of f_0^{th}, Q_0^{th} and k_0^{th} by a fit of Equation 3 to the measured spectrum (dash-dotted line).

ments. This explains, for instance, the 4 Hz difference in the results for f_0^{th} and f_0^{exc} for cantilever V 4 as these measurements were performed with a time difference of several hours. The reproducibility in determining the Q-factor is determined by statistical errors and can be reduced to a standard deviation of 10% for Q_0^{th} by appropriate averaging. Great care has to be taken, however, in mounting the cantilever to ensure that the measured Q-factor is the intrinsic Q-factor of the cantilever rather than an effective Q-factor caused by improper mounting of the cantilever [12]. Determining the cantilever stiffness k_0^{th} relies on an absolute measurement of the cantilever displacement. The main limitations of precision here are the uncertainty in the calibration of the cantilever oscillation amplitude [6,17,18] and of the electronic test equipment involved. The reproducibility for the measurement of k_0^{th} is typically 5% (standard deviation), while a comparison of k_0^{th} to values k_{stat} from reference measurements [16] yields an error of about 10% for the determination of stiffness from thermal noise. Note, however, that the modal stiffness k_0 is related to the static stiffness k by $k_0 = 1.03k$ for a tipless cantilever while, for instance, a tip mass of 10% of the cantilever beam mass yields a relation of $k_0 = 1.01k$ instead [19].

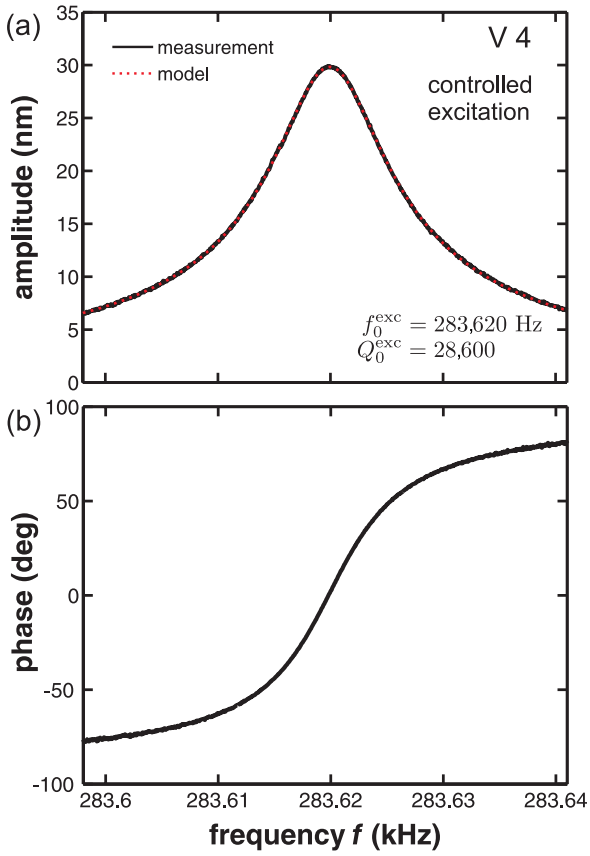

Figure 2: (a) Measured resonance curve (solid line) of the excited cantilever V 4 with a fit (dotted line) of Equation 3 from [12] to experimental data yielding f_0^{exc} and Q_0^{exc}. (b) Phase response of the excited cantilever V 4.

In summary, the analysis of the displacement-noise spectral density around resonances of a thermally excited cantilever in a UHV environment allows the extraction of intrinsic cantilever properties with high accuracy and is specifically useful for determining modal stiffness. However, such measurements require a spectral analysis with high frequency resolution.

Stiffness from frequency shift thermal noise

To circumvent the use of a high-resolution spectrum analyser and to facilitate measurements with the test equipment that is often integrated in NC-AFM control systems, we introduce an alternative method of extracting the cantilever modal stiffness from thermal noise. To apply this method, the eigenfrequency f_n and the quality factor Q_n of the nth oscillation mode have to be measured from an excited resonance curve as shown in Figure 2. To determine k_n, the cantilever is excited to oscilla-

tion in the nth mode at a well stabilised amplitude A_n. Thermal fluctuations described by the power spectral density given by Equation 1 are now superimposed to the deliberate oscillation. These fluctuations are detected by the PLL demodulator tuned to the cantilever eigenfrequency. Effectively, the PLL projects the displacement noise power spectral density in the sidebands of the mode resonance into a range of frequencies f_m starting at 0 Hz. Considering the transfer function of the demodulation and the transfer function of the PLL output or loop filter G_{filter}, the frequency shift noise spectral density at the PLL output can be represented as [6]

$$d^{\Delta f}_{\text{tot},n}\left(f_m\right) = G_{\text{filter}}\left(f_m\right)\sqrt{\frac{f_n k_{\text{B}}T}{\pi k_n Q_n A_n^2} + \frac{2f_m^2}{A_n^2}\left(d^z_{\text{ds},n}\right)^2}\ . \quad (4)$$

This allows us to obtain the modal stiffness from a measurement of $d^{\Delta f}_{\text{tot},n}(f_m)$ if all other parameters are known:

$$k_n^{\Delta f} = \frac{f_n k_{\text{B}}T}{\pi Q_n A_n^2}\left(\left(\frac{d^{\Delta f}_{\text{tot},n}}{G_{\text{filter}}(f_m)}\right)^2 - 2\left(\frac{f_m}{A_n}d^z_{\text{ds},n}\right)^2\right)^{-1}\ . \quad (5)$$

Practically, the spectral analysis can be restricted to the frequency range of 10 Hz to 1 kHz. The resulting spectra are depicted in Figure 3 for different cantilevers (namely V 4 and AL 3 as described in Table 1) excited at their fundamental resonance mode at f_0. The typical shape common to all such spectra has been explained in detail elsewhere [6]. The dotted and dash-dotted lines shown in Figure 3 represent the contributions from thermal noise and detection system frequency shift noise $d^{\Delta f}_{\text{th},n}$ and $d^{\Delta f}_{\text{ds},n}$, respectively. Here, the model curve for $d^{\Delta f}_{\text{th},n}$ is not based on an independent measurement of $d^z_{\text{th},n}$, but determined from the measured $d^{\Delta f}_{\text{tot},n}(f_m)$ curve assuming that the plateau indeed represents purely thermal noise. We focus on the plateau in $d^{\Delta f}_{\text{tot},n}(f_m)$ found in the 10 to 50 Hz region. In the plateau region labelled by a representative modulation frequency f_m^{p}, the frequency shift noise is dominated by thermal noise $d^{\Delta f}_{\text{th},n}(f_m)$ (dash-dotted line), while the noise contribution from the detection system $d^{\Delta f}_{\text{ds},n}(f_m)$ (dotted line) is negligible. Within this approximation, Equation 5 can be simplified to the following expression for the modal stiffness:

$$k_n^{\Delta f} \approx \frac{f_n k_{\text{B}}T}{\pi Q_n A_n^2}\frac{G^2_{\text{filter}}\left(f_m^{\text{p}}\right)}{\left(d^{\Delta f}_{\text{tot},n}\left(f_m^{\text{p}}\right)\right)^2}. \quad (6)$$

Prerequisite for a reliable determination of the modal stiffness is a knowledge of the cantilever properties f_n and Q_n and the PLL filter function G_{filter}. While the former parameters can precisely be determined from an in situ cantilever excitation experiment [12], the latter function can be assumed to be 1 if the filter transfer function is reasonably well-behaved as a function of frequency and a sufficiently large PLL bandwidth is chosen [6]. Results shown in Figure 3 demonstrate that the signal quality obtained under typical experimental conditions is high enough to extract a well-defined value for $d^{\Delta f}_{\text{th},n}(f_m^{\text{p}})$ from the noise data. The modal stiffness values $k_0^{\Delta f}$ obtained for seven cantilevers for fundamental mode excitation according to Equation 6 are displayed in Table 1 and compared to the stiffness results obtained by using different methods for the same cantilevers. From these and further measurements (not shown here), we find an experimental uncertainty of about 20% for the modal stiffness obtained from the frequency shift noise. We attribute the decreased accuracy to noise and uncertainty in the calibration of the additional equipment involved. Note, however, that any noise source besides the thermal excitation

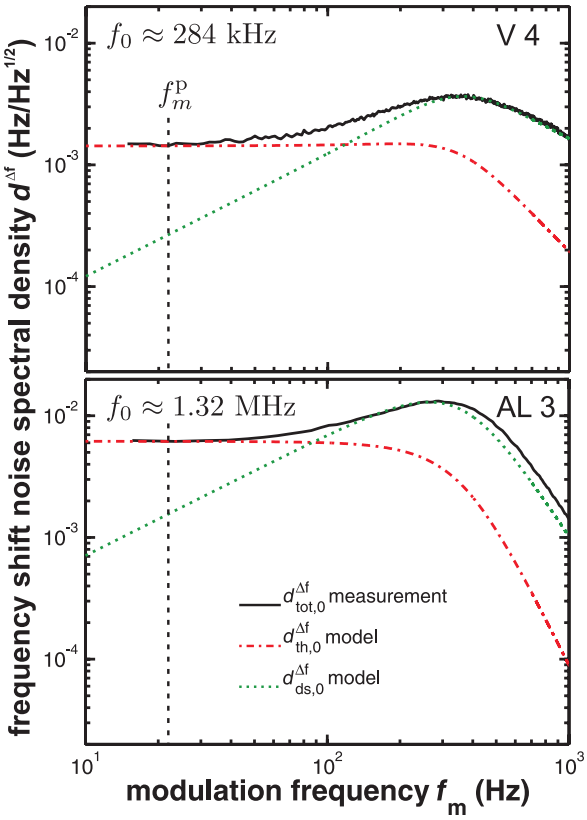

Figure 3: Frequency shift noise spectral density $d^{\Delta f}_{\text{tot},0}$ measured for cantilever V 4 (A_0 = 16.8 nm, demodulator bandwidth $B_{-3\text{dB}}$ = 415 Hz) and cantilever AL 3 (A_0 = 16.5 nm, demodulator bandwidth $B_{-3\text{dB}}$ = 258 Hz). Spectra are recorded with at least 1000 averages. Dotted lines show the contribution $d^{\Delta f}_{\text{ds},0}$ of the detection system noise floor; dash-dotted lines represent the contribution $d^{\Delta f}_{\text{th},0}$ of the thermal noise to the total frequency shift noise $d^{\Delta f}_{\text{tot},0}$. The analysis of the noise spectral density at the plateau frequency f_m^{p} yields the cantilever stiffness $k_n^{\Delta f}$ according to Equation 6.

yields a reduction in the measured modal stiffness and cannot explain values that are too high.

The determination of the effective cantilever stiffness from frequency shift noise is most interesting for cantilever excitation at higher oscillation modes where the projection of the displacement noise into the low frequency region by the PLL demodulator is especially convenient. Respective results obtained for cantilevers P 5 and AF 11 are shown in Table 2. For cantilever P 5 we obtain k_1 = 154 N/m and k_2 = 1330 N/m. As there is no reference for a cross-check of these values, we check for plausibility within the framework of the cantilever oscillation theory. Equation 5 given in [4] allows us to calculate the modal stiffness for a given ratio of tip mass to beam mass. Assuming the tip mass being 2% of the cantilever beam mass yields k_1/k_0 = 45.0 and k_2/k_0 = 397. These numbers fit well to the measured values for cantilever P 5 (see Table 2). The result for AF 11 can be explained by a tip mass of 5% of the cantilever beam mass.

Table 2: Cantilever eigenfrequencies f_n^{exc} and quality factors Q_n^{exc} of the nth oscillating mode for cantilevers P 5 and AF 11. The modal stiffness $k_n^{\Delta f}$ is obtained from the frequency shift noise spectral density $d_{\text{tot},n}^{\Delta f}$ through Equation 6. $k_n^{\Delta f}/k_0^{\Delta f}$ is the ratio of the modal stiffness of the nth oscillation mode to the stiffness of the fundamental mode.

cantilever	n	f_n^{exc} (Hz)	Q_n^{exc}	$k_n^{\Delta f}$ (N/m)	$k_n^{\Delta f}/k_0^{\Delta f}$
P 5	1	436,711	44,900	154	45.3
P 5	2	1,234,277	5841	1330	391
AF 11	1	1,934,677	9000	3420	56

Conclusion

In conclusion, we introduce a method for determining the modal stiffness k_n of a cantilever from frequency shift noise complemented by an independent measurement of the modal eigenfrequency and Q-factor. Our strategy yields valid results with an uncertainty of about 20%; however, the accuracy is expected to be increased by an improvement of the experimental setup. This method is particularly convenient as measurements can be performed with simple test equipment implemented in many NC-AFM control systems. Additionally, the involved spectral analysis is simple and can be performed over a bandwidth of only 1 kHz at maximum, irrespective of the cantilever eigenfrequency. We apply the thermal noise method to various cantilever types and find a good agreement of these cantilever parameters with those determined by using alternative methods. The strength of the methods presented here is that they directly yield the modal stiffness derived from a thermal measurement and do not require any modelling to relate geometric cantilever properties to oscillation properties.

Acknowledgements

The authors gratefully acknowledge experimental support from Stefan Kuhn (Johannes Gutenberg-Universität Mainz). This project has been generously supported by Nanoworld Services GmbH. M. T. gratefully appreciates support from the Hans-Mühlenhoff-Stiftung. P. R. gratefully acknowledges financial support from the Alexander von Humboldt-Foundation.

References

1. Nyquist, H. *Phys. Rev.* **1928,** *32,* 110–113.
2. Johnson, J. B. *Phys. Rev.* **1928,** *32,* 97–109.
3. Saulson, P. R. *Phys. Rev. D* **1990,** *42,* 2437–2445. doi:10.1103/PhysRevD.42.2437
4. Lozano, J. R.; Kiracofe, D.; Melcher, J.; Garcia, R.; Raman, A. *Nanotechnology* **2010,** *21,* 465502. doi:10.1088/0957-4484/21/46/465502
5. In literature, two different numbering schemes for the oscillation modes are commonly used starting either with n = 0,1,2,… or n = 1,2,3,…. We use n = 0 for the fundamental oscillation mode in accordance with the nomenclature found in [6],[10],[12],[13],[15-18].
6. Lübbe, J.; Temmen, M.; Rode, S.; Rahe, P.; Kühnle, A.; Reichling, M. *Beilstein J. Nanotechnol.* **2013,** *4,* 32–44. doi:10.3762/bjnano.4.4
7. Butt, H.-J.; Jaschke, M. *Nanotechnology* **1995,** *6,* 1–7. doi:10.1088/0957-4484/6/1/001
8. Paolino, P.; Tiribilli, B.; Bellon, L. *J. Appl. Phys.* **2009,** *106,* 094313. doi:10.1063/1.3245394
9. Rast, S.; Wattinger, C.; Gysin, U.; Meyer, E. *Nanotechnology* **2000,** *11,* 169. doi:10.1088/0957-4484/11/3/306
10. Hutter, J. L.; Bechhoefer, J. *Rev. Sci. Instrum.* **1993,** *64,* 1868–1873. doi:10.1063/1.1143970
11. Cook, S. M.; Schäffer, T. E.; Chynoweth, K. M.; Wigton, M.; Simmonds, R. W.; Lang, K. M. *Nanotechnology* **2006,** *17,* 2135–2145. doi:10.1088/0957-4484/17/9/010
12. Lübbe, J.; Tröger, L.; Torbrügge, S.; Bechstein, R.; Richter, C.; Kühnle, A.; Reichling, M. *Meas. Sci. Technol.* **2010,** *21,* 125501. doi:10.1088/0957-0233/21/12/125501
13. Lantz, M. A.; Hug, H. J.; Hoffmann, R.; van Schendel, P. J. A.; Kappenberger, P.; Martin, S.; Baratoff, A.; Güntherodt, H.-J. *Science* **2001,** *291,* 2580–2583. doi:10.1126/science.1057824
14. Kawai, S.; Glatzel, T.; Koch, S.; Baratoff, A.; Meyer, E. *Phys. Rev. B* **2011,** *83,* 035421. doi:10.1103/PhysRevB.83.035421
15. Welker, J.; Illek, E.; Giessibl, F. J. *Beilstein J. Nanotechnol.* **2012,** *3,* 238–248. doi:10.3762/bjnano.3.27
16. Lübbe, J.; Doering, L.; Reichling, M. *Meas. Sci. Technol.* **2012,** *23,* 045401. doi:10.1088/0957-0233/23/4/045401
17. Giessibl, F. J. *Appl. Phys. Lett.* **2000,** *76,* 1470–1472. doi:10.1063/1.126067
18. Simon, G. H.; Heyde, M.; Rust, H.-P. *Nanotechnology* **2007,** *18,* 255503. doi:10.1088/0957-4484/18/25/255503
19. Melcher, J.; Hu, S. Q.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 053101. doi:10.1063/1.2767173

Beilstein Journal of Nanotechnology

High-resolution nanomechanical analysis of suspended electrospun silk fibers with the torsional harmonic atomic force microscope

Mark Cronin-Golomb[*1] and Ozgur Sahin[*2,3]

Full Research Paper

Open Access

Address:
[1]Department of Biomedical Engineering, Tufts University, Medford, MA 02155, USA, [2]Rowland Institute at Harvard, Harvard University, Cambridge, MA 02142, USA and [3]presently with Department of Biological Sciences and Department of Physics, Columbia University, New York, NY 10027, USA

Email:
Mark Cronin-Golomb[*] - mark.cronin-golomb@tufts.edu;
Ozgur Sahin[*] - sahin@columbia.edu

* Corresponding author

Keywords:
atomic force microscopy; nanomechanical characterization; silk fibers; tissue scaffolds; torsional harmonic cantilevers

Beilstein J. Nanotechnol. **2013,** *4,* 243–248.
doi:10.3762/bjnano.4.25

Received: 29 January 2013
Accepted: 19 March 2013
Published: 05 April 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

Atomic force microscopes have become indispensable tools for mechanical characterization of nanoscale and submicron structures. However, materials with complex geometries, such as electrospun fiber networks used for tissue scaffolds, still pose challenges due to the influence of tension and bending modulus on the response of the suspended structures. Here we report mechanical measurements on electrospun silk fibers with various treatments that allow discriminating among the different mechanisms that determine the mechanical behavior of these complex structures. In particular we were able to identify the role of tension and boundary conditions (pinned versus clamped) in determining the mechanical response of electrospun silk fibers. Our findings show that high-resolution mechanical imaging with torsional harmonic atomic force microscopy provides a reliable method to investigate the mechanics of materials with complex geometries.

Introduction

Dynamic atomic force microscopy (AFM) methods provide opportunities for high-resolution compositional mapping of heterogeneous samples [1]. Recent developments in dynamic AFM methods offer the possibility of relating the measured vibration signals to the particular physical properties of the samples, such as elastic modulus, viscosity, adhesion, and chemical affinity [2-16]. These developments are accomplished by employing multiple excitation and detection frequencies during dynamic AFM imaging [17-26]. A critical element of these mechanical measurements is the physical model being used to relate the force–distance curves to parameters representing the material properties. Although contact-mechanics models can be used for a wide variety of polymer composites, block-copolymers, and biomaterials [27-34], these models are

not applicable to materials with complex geometries. For example, the use of local interaction models provides limited information in the case of suspended structures where bending modulus, geometry, and mechanical tension are the key determinants of tip–sample interactions. The interpretation of measurements on these kinds of samples requires simultaneous analysis of mechanical measurements and topography, as well as comparison of various mechanical models.

In this work we investigate the mechanical behavior of electrospun silk fibers, which are used for making scaffolds for bone-tissue engineering [35]. Mechanical characteristics of these structures are important because their primary purpose is to mimic extracellular-matrix conditions, including their rigidity. Bulk properties, while important, are not sufficient to predict the mechanical behavior of the electrospun fibers. Geometry of the network of fibers, fiber diameter, mechanical boundary conditions at the nodes of the fiber network (pinned versus clamped), and the presence of mechanical tension within the fibers can influence their mechanical behavior. We have carried out experiments to determine the relative influences of these parameters on the mechanics of electrospun silk fibers.

Results and Discussion

Electrospun silk fibers form mesh-like networks with nodes and branches. Diameters of these fibers are typically in the submicron range. Separation between the nodes, defined by intersections between two or more fibers, can be on the order of one to ten micrometers depending on the electrospinning process. This size scale is readily accessible by atomic force microscopy for topographical and mechanical characterization. When several fiber layers are deposited to form fibrous tissue scaffolds, these branches form suspended structures. We have limited our experiments to samples that are formed by two to three layers of fibers so that we can readily identify individual fibers. Although some of the branches in the first few layers appear to rest on the substrate, some branches still form suspended fibers sufficient for our study (Figure 1).

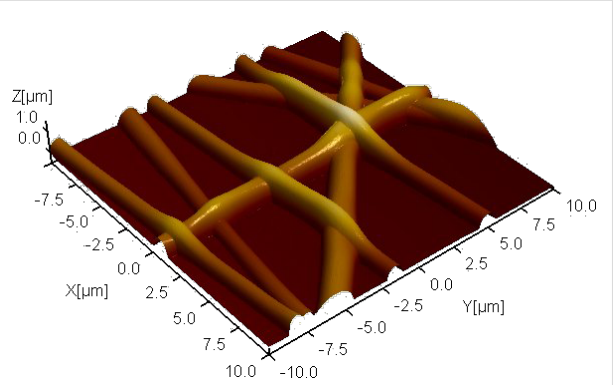

Figure 1: Topography of electrospun silk fibers on a glass substrate. The 3-D image is rendered according to the local height measured by the atomic force microscope. The scan size is 20 × 20 µm^2. The fibers form a mesh-like network. Branches between intersections occasionally form suspended fibers, allowing us to investigate their mechanical behavior.

We have used torsional harmonic AFM to determine the surface topography and local mechanical response with high spatial resolution [20,31]. This mode uses a T-shaped cantilever with an offset tip. When used in dynamic AFM, the cantilever vibrates up and down, similar to conventional cantilevers. In addition to the vertical motion, tip–sample-interaction forces twist the cantilever by a detectible amount. The high bandwidth of torsional motion allows accessing higher harmonics of the tip–sample-interaction forces to reconstruct tip–sample-force waveforms. This process involves calibration of the frequency response of the torsional mode by measuring its resonance frequency and quality factor (either by frequency sweeps or from the thermal peak in the noise spectrum). The gain of the torsional mode, defined as the photodetector signal corresponding to a unit amount of a quasi-static force acting on the tip, is determined by independently measuring the quasi-static force from vertical deflections while monitoring the torsional deflection signal. Note that the same force acts on both vertical and torsional modes. Therefore, after calibrating the vertical spring constant, the gain of the torsional mode can be determined by comparing time-average detector signals in vertical and horizontal channels during a tapping-mode AFM experiment. To minimize contributions of drift in quasi-static deflection signals, we previously developed a procedure that takes advantage of the transitions between attractive and repulsive modes [36]. The calibrated frequency response and gain of the torsional mode allows the reconstruction of the tip–sample-force waveforms. A computer program carries out these calculations in real time during the tapping-mode imaging process. The program also corrects for nonlinearities of the position-sensitive diode and for crosstalk from large vertical signals into torsional vibration signals. Once the tip–sample-force waveform is determined, the program constructs force–distance curves using the distance information in the vertical deflection signals [20]. It is possible to analyze these force–distance curves according to various physical models to obtain parameters describing the mechanical response of the sample. In the case of electrospun silk fibers, we have calculated both the local elastic modulus and the local spring constant values. The elastic modulus is calculated according to the Derjaguin–Muller–Toporov (DMT) model and the spring constant (stiffness) is calculated by fitting the unloading portion of the force–distance curve with a straight line. For the DMT model, we used a tip radius of 7 nm, which is characterized by blind reconstruction from a sample with sharp edges. Our calcu-

lations assumed that the unloading portion of the force distance-curve is the region between the peak force and the point where the force drops to 20% of the peak value.

Initially, we calculated both the elastic modulus and spring constant values regardless of their appropriateness for describing local mechanical response. We identified the appropriate parameter by analyzing the simultaneously obtained topography image. In regions where silk fibers appear to rest on the substrate, elastic modulus values are considered to be the relevant parameter. In regions where the silk fibers appear to be suspended, spring constant values are considered to be the relevant parameter. In the latter case, we further analyzed whether the measured spring constant values across the suspended fiber can be predicted by various mechanical models.

Figure 2 shows topography, elastic modulus, and stiffness images obtained from electrospun silk fibers. The sample was prepared from pure silk electrospun at 8 kV over a distance of 7 inches at pH 8. A suspended fiber branch extending from left to right is identified based on its height relative to the remaining fibers and the glass substrate. Two fibers extending from top to bottom on the left side of the image appear to be resting on the glass substrate, whereas the fibers on the right side intersect with each other and the suspended horizontal fiber at approximately the same location. The fibers that are resting on the substrate exhibit a uniform elastic-modulus profile (about 5 GPa); however the suspended horizontal fiber exhibits a gradual decrease towards its midpoint. The stiffness map and its corresponding line profile show similar trends, indicating qualitative agreement between the two parameters. Because the model used for calculating the elastic modulus does not account for the suspended geometry of the fiber, we rely on the spring constant values to understand the mechanism responsible for the observed mechanical response.

It is also important to consider the possible effect of the inertia of the fibers on the measured forces. The inertia of the fibers can be neglected if the resonance frequency of the suspended silk fibers is much higher than the frequency of cantilever vibrations analyzed in the experiment. Torsional harmonic AFM relies on force measurements at frequencies up to the torsional

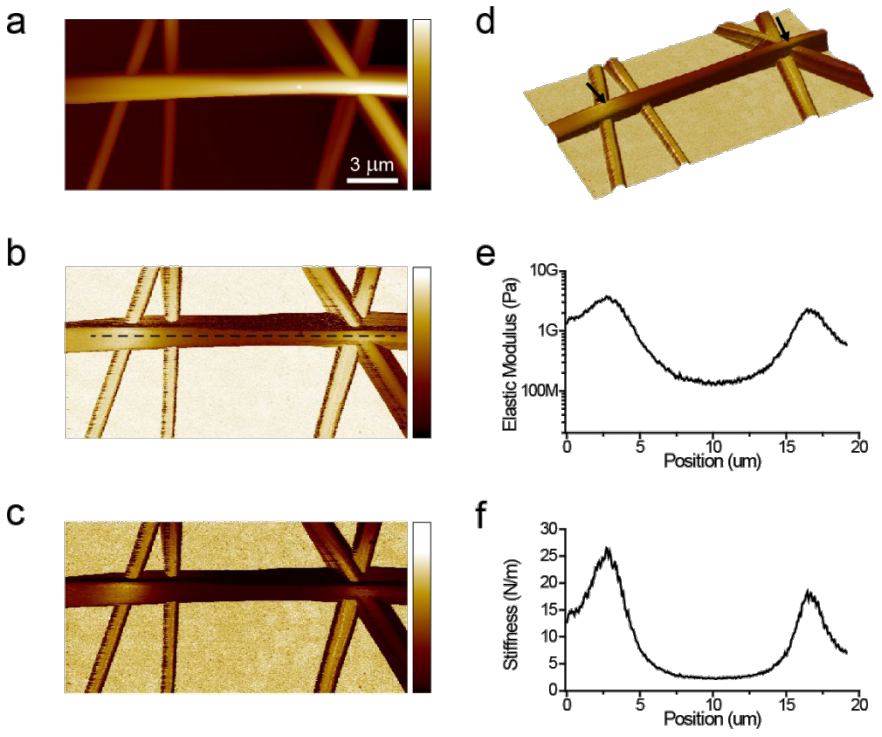

Figure 2: Simultaneously measured topography (a), elastic modulus (b), and stiffness (c) maps obtained from electrospun silk fibers. Color bars in (a–c) correspond to the ranges in height (0–1.8 µm), elastic modulus (10 MPa to 10 GPa, mapped logarithmically), and stiffness (0–5 N/m). The horizontal fiber appears to be suspended above the glass substrate. A 3-D rendering of the topography image is given in (d). The fiber is suspended between positions indicated by arrows in (d). This image is colored according to the local spring constant. Both the elastic modulus and stiffness maps show gradual variations across the suspended silk fiber. Line profiles of elastic modulus and stiffness across the dashed line in (b) are given in (e) and (f), respectively. While the local elastic modulus of the silk fiber is likely to be constant across the length of the fiber, the values in (e) show significant variation. This is because the elastic modulus values in (b,e) are calculated by the DMT contact-mechanics model, which does not take the suspended geometry of the fiber into account. Therefore, the regions of the elastic modulus image corresponding to suspended fibers are not reliable. These regions are better analyzed in the light of mechanical models describing the entire suspended structure by using the stiffness values in (c,f).

resonance frequency of the cantilever, about 1 MHz. One can estimate the resonance frequency of the suspended fiber structure using Euler–Bernoulli analysis [37]:

$$f = \frac{\beta^2}{8\pi}\sqrt{\frac{E}{\rho}}\frac{D}{L^2}. \quad (1)$$

Here the constant β^2 is equal to 22.373 for the clamped-end boundary condition. E and ρ are the elastic modulus and mass density, and D and L are the diameter and length of the silk fiber. Using E = 10 GPa, ρ = 1.3 g/cm^3, $D = 2R$ = 0.52 μm, we obtain the resonance frequency f = 129.4 MHz, which is far above the torsional resonance frequency of the AFM cantilever. Therefore, we neglect the effects of the inertia of the fibers in our experiments.

While gradual changes in stiffness of the suspended fiber are not surprising, the precise mechanism that determinines the mechanical response of the suspended fiber is not immediately clear. We identified three scenarios that can qualitatively explain the observed results. The suspended fiber can be viewed as a cantilever structure pinned at both ends, clamped at both ends, or as a string that is under tension. Graphical depictions of these three cases are given in Figure 3a–c. All three scenarios would result in variations in the local stiffness of the fiber as probed by the AFM tip. However, the stiffness values predicted by these models would have different spatial dependencies. It is worth noting that all these models assume that the displacement of the fiber at the nodes is zero, which would result in an effectively infinite spring constant at these locations. In our experiments, the spring constant at the nodes are finite and determined by both the tip–fiber contact-mechanics and the spring constant associated with fiber–fiber interactions at the nodes. To take these effects into account, we assumed a simple model depicted in Figure 3d, which we refer to as the suspended-rigid-rod model. The variables required by all four models and the equations decribing local spring constants based on these models are listed in Table 1. Note that the effective spring constant orignating from the suspended-rigid-rod model acts in series with the other three models. Additionally, a more sophisticated model could include the fiber–tip spring constant, which acts in series with the spring constants due to fiber–fiber interactions at the nodes. The two models have to give the same total spring constant at the nodes, but the model in Figure 3d results in a linear dependancy to the distance from the nodes and the model that takes the fiber–tip spring constant into account results in a nonlinear dependancy on the distance.

To determine if any of the three models in Figure 3a–c, in combination with the rigid-rod model in Figure 3d, can explain

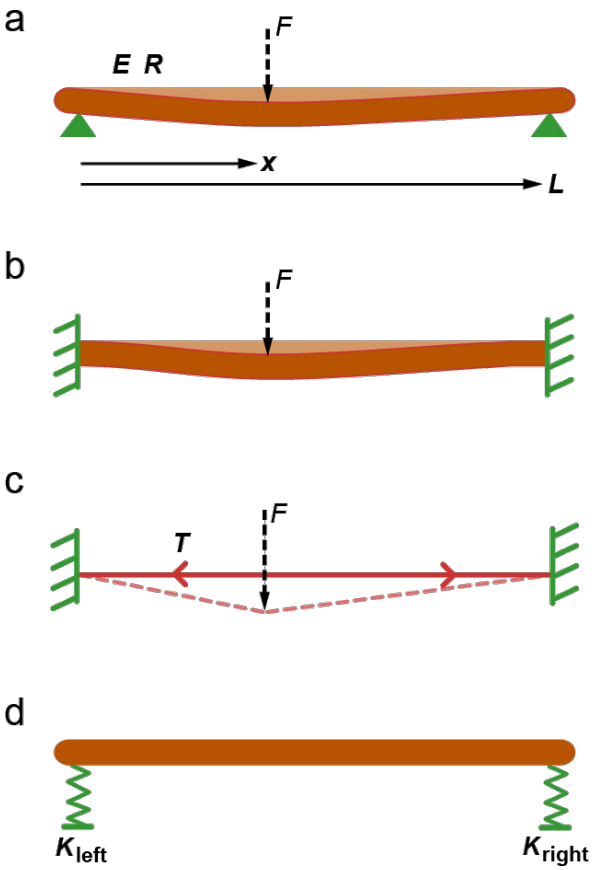

Figure 3: Illustration of possible mechanisms determining the local spring constant of suspended silk fibers. (a) "Pinned-end model" assumes that the fiber displacement is zero at the nodes; however, there is no constraint on its angle at the nodes. E elastic modulus, R is fiber radius, L is branch length. (b) "Clamped-end model" assumes that both the angle and displacement at the nodes are zero. (c) "Tension model" assumes the fiber has a built in tension T and negligible bending modulus. (d) "Suspended rigid rod model" assumes the nodes have finite spring constants K_{left} and K_{right}. Note that the spring constant according to the mechanism in (d) acts in series with the mechanisms in (a–c).

the observed variations in spring constant, we attempted to fit the data in Figure 2f with the total spring constant $K(x)$ based on equations listed in Table 1. For the pinned-end and clamped-end models, we used E and L as variables for the fitting. For the tension model, we used T and L as variables. For all three models, K_S and K_T are assumed to be constant and equal to 26.2 N/m and 18.4 N/m, respectively. In addition we used R = 0.26 μm. The values for K_S and K_T are determined from the peak spring constant values in the data plotted in Figure 2f. The value for R is determined from the topography measurements in Figure 2a. The values of the parameters used for fitting are also listed in Table 1 and the resulting curves are plotted in Figure 4.

From the results of the fitting procedures we see that all three models could reproduce qualitative trends similar to the

Table 1: Description of the variables and equations for spring constants based on the four mechanical models depicted in Figure 3. The values of variables calculated by curve fitting are given in the last column. Standard errors are given in parenthesis with the same units.

Model description	Variables	Constants	Equation	Best fit (standard error)
Pinned end	E: elastic modulus L: branch length x: position	R: fiber radius	$K_P(x) = \frac{3\pi ER^4}{4}\frac{L}{x^2(L-x)^2}$	E = 35.47 GPa (0.29) L = 13.59 µm (0.014)
Clamped end	E: elastic modulus L: branch length x: position	R: fiber radius	$K_C(x) = \frac{3\pi ER^4}{4}\frac{L^3}{x^3(L-x)^3}$	E = 10.16 GPa (0.13) L = 14.94 µm (0.021)
Tension	T: tension L: branch length x: position		$K_T(x) = \frac{TL}{x(L-x)}$	T = 16.73 µN (0.42) L = 13.47 µm (0.003)
Suspended rigid rod	L: branch length x: position	K_{left}: left spring constant K_{right}: right spring constant	$K_S(x) = K_{left}\frac{(L-x)}{L} + K_{right}\frac{x}{L}$	

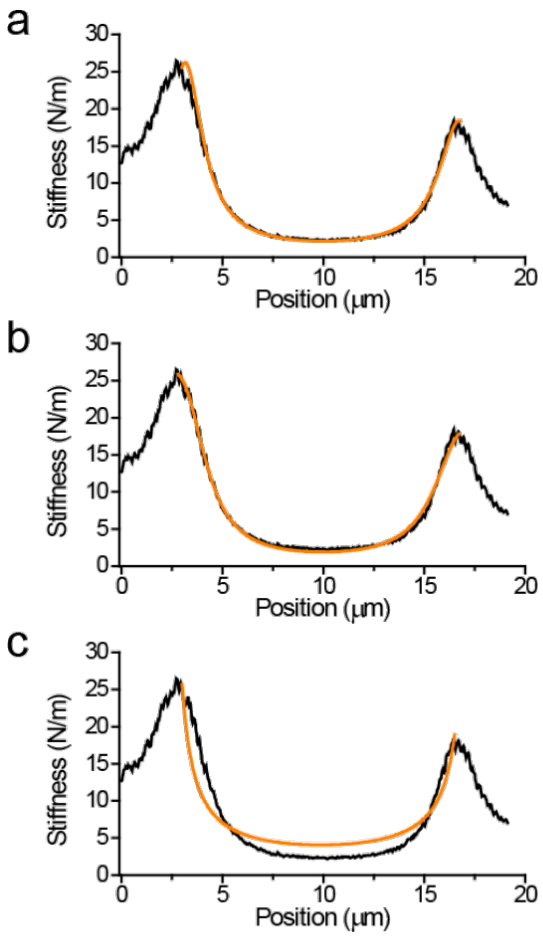

Figure 4: Curves described by equations for pinned end (a), clamped end (b), and tension (c) models fitted to the data. Values of the variables used for fitting are listed in Table 1.

measured spring-constant profile. However, the tension model did not produce a good overall fit to the data. The tension value calculated as 16.7 µN translates to a tensile stress of ≈78.6 MPa, which is a relatively high value, but silk could potentially sustain such large stresses. The other two models provided a better overall fit to the data as seen in Figure 4a and Figure 4b. However, quantitatively the pinned-end model required E to be 35.5 GPa, which is higher than even in native silk fibers ($E \approx 14$ GPa [38]). The clamped-end model predicts E to be around 10.2 GPa and therefore it is more likely that this model provides a better description of suspended silk fibers. Note that the numerical estimates depend on the fourth power of the fiber radius; however, even with 10% increase in radius, the pinned end model cannot provide an elastic modulus value that falls in a plausible range (<15 GPa).

Conclusion

In this paper we have investigated the mechanical behavior of electrospun silk fibers whose geometry does not allow the straightforward use of contact-mechanics models. We used elastic-modulus and stiffness maps determined by torsional harmonic AFM and fitted the data obtained from suspended silk fibers with models that could potentially explain the observed variations in stiffness. This analysis revealed that a clamped-end model, in which the displacements and bending of a fiber are restricted at nodes, successfully describes the observed characteristics. We expect that the applications of the general methodology used in this paper could also be extended to characterization of cytoskeletal protein networks and microelectromechanical (MEMS) devices where suspended structures are commonly encountered.

References

1. García, R.; Magerle, R.; Perez, R. *Nat. Mater.* **2007,** *6,* 405–411. doi:10.1038/nmat1925
2. Balke, N.; Jesse, S.; Morozovska, A. N.; Eliseev, E.; Chung, D. W.; Kim, Y.; Adamczyk, L.; García, R. E.; Dudney, N.; Kalinin, S. V. *Nat. Nanotechnol.* **2010,** *5,* 749–754. doi:10.1038/nnano.2010.174

3. Dokukin, M. E.; Guz, N. V.; Gaikwad, R. M.; Woodworth, C. D.; Sokolov, I. *Phys. Rev. Lett.* **2011,** *107,* 028101. doi:10.1103/PhysRevLett.107.028101
4. Dong, M.; Sahin, O. *Nat. Commun.* **2011,** *2,* 247. doi:10.1038/ncomms1246
5. Forchheimer, D.; Platz, D.; Tholén, E. A.; Haviland, D. B. *Phys. Rev. B* **2012,** *85,* 195449. doi:10.1103/PhysRevB.85.195449
6. Garcia, R.; Herruzo, E. T. *Nat. Nanotechnol.* **2012,** *7,* 217–226. doi:10.1038/nnano.2012.38
7. Jesse, S.; Kalinin, S. V.; Proksch, R.; Baddorf, A. P.; Rodriguez, B. J. *Nanotechnology* **2007,** *18,* 435503. doi:10.1088/0957-4484/18/43/435503
8. Martínez, N. F.; Lozano, J. R.; Herruzo, E. T.; Garcia, F.; Richter, C.; Sulzbach, T.; Garcia, R. *Nanotechnology* **2008,** *19,* 384011. doi:10.1088/0957-4484/19/38/384011
9. Platz, D.; Tholén, E. A.; Pesen, D.; Haviland, D. B. *Appl. Phys. Lett.* **2008,** *92,* 153106. doi:10.1063/1.2909569
10. Proksch, R. *Appl. Phys. Lett.* **2006,** *89,* 113121. doi:10.1063/1.2345593
11. Raman, A.; Trigueros, S.; Cartagena, A.; Stevenson, A. P. Z.; Susilo, M.; Nauman, E.; Contera, S. A. *Nat. Nanotechnol.* **2011,** *6,* 809–814. doi:10.1038/nnano.2011.186
12. Sikora, A.; Bednarz, L. *Cent. Eur. J. Phys.* **2011,** *9,* 372–379. doi:10.2478/s11534-010-0127-4
13. Solares, S. D.; Chawla, G. *J. Appl. Phys.* **2010,** *108,* 054901. doi:10.1063/1.3475644
14. Tetard, L.; Passian, A.; Thundat, T. *Nat. Nanotechnol.* **2010,** *5,* 105–109. doi:10.1038/nnano.2009.454
15. Tetard, L.; Passian, A.; Venmar, K. T.; Lynch, R. M.; Voy, B. H.; Shekhawat, G.; Dravid, V. P.; Thundat, T. *Nat. Nanotechnol.* **2008,** *3,* 501–505. doi:10.1038/nnano.2008.162
16. Xu, X.; Melcher, J.; Basak, S.; Reifenberger, R.; Raman, A. *Phys. Rev. Lett.* **2009,** *102,* 060801. doi:10.1103/PhysRevLett.102.060801
17. Abak, M. K.; Aktas, O.; Mammadov, R.; Gürsel, I.; Dâna, A. *Appl. Phys. Lett.* **2008,** *92,* 223113. doi:10.1063/1.2940304
18. Hutter, C.; Platz, D.; Tholén, E. A.; Hansson, T. H.; Haviland, D. B. *Phys. Rev. Lett.* **2010,** *104,* 050801. doi:10.1103/PhysRevLett.104.050801
19. Lozano, J. R.; Garcia, R. *Phys. Rev. Lett.* **2008,** *100,* 076102. doi:10.1103/PhysRevLett.100.076102
20. Sahin, O.; Magonov, S.; Su, C.; Quate, C. F.; Solgaard, O. *Nat. Nanotechnol.* **2007,** *2,* 507–514. doi:10.1038/nnano.2007.226
21. Sarioglu, A. F.; Solgaard, O. *Appl. Phys. Lett.* **2008,** *93,* 023114. doi:10.1063/1.2959828
22. Solares, S. D.; Hölscher, H. *Nanotechnology* **2010,** *21,* 075702. doi:10.1088/0957-4484/21/7/075702
23. Stark, M.; Stark, R. W.; Heckl, W. M.; Guckenberger, R. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 8473–8478. doi:10.1073/pnas.122040599
24. Stark, R. W.; Naujoks, N.; Stemmer, A. *Nanotechnology* **2007,** *18,* 065502. doi:10.1088/0957-4484/18/6/065502
25. Platz, D.; Forchheimer, D.; Tholén, E. A.; Haviland, D. B. *Nat. Commun.* **2013,** *4,* 1360. doi:10.1038/ncomms2365
26. Williams, J. C.; Solares, S. D. *Beilstein J. Nanotechnol.* **2013,** *4,* 87–93. doi:10.3762/bjnano.4.10
27. Cronin-Golomb, M.; Murphy, A. R.; Mondia, J. P.; Kaplan, D. L.; Omenetto, F. G. *J. Polym. Sci., Part B: Polym. Phys.* **2012,** *50,* 257–262. doi:10.1002/polb.23003
28. Ihalainen, P.; Järnström, J.; Määttänen, A.; Peltonen, J. *Colloids Surf., A* **2011,** *373,* 138–144. doi:10.1016/j.colsurfa.2010.10.041
29. Leung, K. M.; Wanger, G.; Guo, Q.; Gorby, Y.; Southam, G.; Lau, W. M.; Yang, J. *Soft Matter* **2011,** *7,* 6617–6621. doi:10.1039/c1sm05611e
30. Qu, M.; Deng, F.; Kalkhoran, S. M.; Gouldstone, A.; Robisson, A.; Van Vliet, K. J. *Soft Matter* **2011,** *7,* 1066–1077. doi:10.1039/c0sm00645a
31. Sahin, O.; Erina, N. *Nanotechnology* **2008,** *19,* 445717. doi:10.1088/0957-4484/19/44/445717
32. Schön, P.; Dutta, S.; Shirazi, M.; Noordermeer, J.; Vancso, G. J. *J. Mater. Sci.* **2011,** *46,* 3507–3516. doi:10.1007/s10853-011-5259-4
33. Sweers, K.; van der Werf, K.; Bennink, M.; Subramaniam, V. *Nanoscale Res. Lett.* **2011,** *6,* 270. doi:10.1186/1556-276X-6-270
34. Wang, S.; Zhao, G. *Mater. Lett.* **2012,** *79,* 14–17. doi:10.1016/j.matlet.2012.03.044
35. Li, C.; Vepari, C.; Jin, H.-J.; Kim, H. J.; Kaplan, D. L. *Biomaterials* **2006,** *27,* 3115–3124. doi:10.1016/j.biomaterials.2006.01.022
36. Sahin, O. *Rev. Sci. Instrum.* **2007,** *78,* 103707. doi:10.1063/1.2801009
37. Belov, M.; Quitoriano, N. J.; Sharma, S.; Hiebert, W. K.; Kamins, T. I.; Evoy, S. *J. Appl. Phys.* **2008,** *103,* 074304. doi:10.1063/1.2891002
38. Wang, M.; Jin, H.-J.; Kaplan, D. L.; Rutledge, G. C. *Macromolecules* **2004,** *37,* 6856–6864. doi:10.1021/ma048988v

License and Terms

Beilstein Journal of Nanotechnology

Polynomial force approximations and multifrequency atomic force microscopy

Daniel Platz*[1], Daniel Forchheimer[1], Erik A. Tholén[2] and David B. Haviland[1]

Full Research Paper

Open Access

Address:
[1]KTH Royal Institute of Technology, Section for Nanostructure Physics, Albanova University Center, SE-106 91 Stockholm, Sweden and [2]Intermodulation Products AB, Vasavägen 29, SE-169 58 Solna, Sweden

Email:
Daniel Platz* - platz@kth.se

* Corresponding author

Keywords:
AFM; atomic force microscopy; force spectroscopy; multifrequency; intermodulation; polynomial

Beilstein J. Nanotechnol. **2013,** *4,* 352–360.
doi:10.3762/bjnano.4.41

Received: 02 February 2013
Accepted: 13 May 2013
Published: 10 June 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We present polynomial force reconstruction from experimental intermodulation atomic force microscopy (ImAFM) data. We study the tip–surface force during a slow surface approach and compare the results with amplitude-dependence force spectroscopy (ADFS). Based on polynomial force reconstruction we generate high-resolution surface-property maps of polymer blend samples. The polynomial method is described as a special example of a more general approximative force reconstruction, where the aim is to determine model parameters that best approximate the measured force spectrum. This approximative approach is not limited to spectral data, and we demonstrate how it can be adapted to a force quadrature picture.

Introduction

The combination of high-resolution imaging [1-4] and high-accuracy force measurements [5-10] is a strong driving force for the development of atomic force microscopy (AFM). The advent of multifrequency AFM resulted in a variety of new measurement techniques enabling enhanced contrast and spatial mapping of surface properties on a wide range of samples [11]. However, multifrequency AFM creates more data than conventional AFM, which both complicates the interpretation of measurement results and offers the possibility of much more detailed surface analysis. One of the goals when interpreting AFM data is the reconstruction of the force between a surface and the sharp tip at the end of the oscillating cantilever, while scanning. This reconstruction is readily possible by means of the Fourier transform if the motion of the tip in response to this force and the linear response function of the cantilever are known over a wide frequency band [12-15]. However, the tip motion is often only measurable in a narrow frequency band around a cantilever resonance, since the cantilever transfer function sharply attenuates other frequency components of the tip motion, placing them below the detection noise floor. With this measured partial motion spectrum, the original force cannot be recovered with a simple Fourier transform, and additional assumptions about the functional representation of the tip–surface force are required.

These assumptions, which can be expressed with a finite set of parameters, result in a correlation of the measurable and the nonmeasurable frequency components of the motion. The parameters are chosen such that the spectrum of the reconstructed force best approximates the measured partial force spectrum. This approximative reconstruction requires the use of numerical optimization techniques if the force model is nonlinear in the parameters [16]. Analytic solutions can be obtained if the model is linear in the parameters [9,17-19]. Such a linear model of particular interest is the polynomial, as it constitutes a general expansion of the tip–surface force.

Polynomial force reconstruction methods have been proposed theoretically and tested on simulated data for intermodulation AFM (ImAFM) [17,19]. Here, we demonstrate, for the first time, polynomial force reconstruction on experimental ImAFM data and compare it with reconstruction based on amplitude-dependence force spectroscopy (ADFS) [20]. Moreover, fitting a force model to the polynomial reconstruction allows for the extraction of properties such as surface adhesion, sample stiffness or interaction geometry. We demonstrate this extraction of surface properties with high-resolution stiffness maps on a blend of polystyrene (PS) and poly(methyl methacrylate) (PMMA).

Polynomial reconstruction, and most other multifrequency methods, work directly on the measured spectral data of the tip motion. Since the tip motion can be very complicated, the interpretation of spectral data often becomes rather difficult and alternative data-representation schemes might provide a better understanding of the tip–surface force. Recently, we have shown how a narrow-band ImAFM measurement yields the oscillation-amplitude dependence of a force component F_I in-phase with the sinusoidal tip motion and a force component F_Q quadrature, or 90 degrees phase-shifted, to the tip motion [21]. Here we show how polynomial force reconstruction can be performed within the context of this picture of two force quadratures.

Results and Discussion

Polynomial force reconstruction from spectral data

In narrow-band AFM the tip dynamics as a function of time $z(t)$ is usually described by a harmonic oscillator [22,23], subject to an external drive force and a time-dependent tip–surface force

$$\ddot{z} + \frac{\omega_0}{Q}\dot{z} + k_c(z-h) = F_{\text{drive}}(t) + F_{\text{ts}}(t) \tag{1}$$

where the dot denotes differentiation with respect to time, ω_0, Q and k_c are the mode's resonance frequency, quality factor and spring constant respectively, and h is the static equilibrium position of the tip above the surface. One should note that the time dependence of the tip–surface force F_{ts} can be considered as an implicit time dependence, since it is assumed that the tip–surface interaction depends on the instantaneous tip position z and velocity $\dot{z}$, which are functions of time. In Fourier space Equation 1 becomes

$$\hat{z}(\omega) = \hat{\chi}(\omega)\left(\hat{F}_{\text{drive}}(\omega) + \hat{F}_{\text{ts}}(\omega)\right) \tag{2}$$

where the linear response function

$$\hat{\chi}(\omega) = \frac{\omega_0^2/k_c}{\omega_0^2 - \omega + i\frac{\omega_0\omega}{Q}} \tag{3}$$

with the complex unit i determines the tip response to a sinusoidal force applied at the frequency ω. The drive force can be readily determined from a measurement of the tip motion far away from the surface, $z_{\text{free}}(t)$, where the tip–surface force is zero,

$$\hat{F}_{\text{drive}}(\omega) = \hat{\chi}^{-1}(\omega)\hat{z}_{\text{free}}(\omega) \tag{4}$$

If the broad-band tip response close to the surface $\hat{z}(\omega)$ is known, one can easily solve Equation 2 for the spectrum of the tip–surface force

$$\hat{F}_{\text{ts}}(\omega) = \hat{\chi}^{-1}(\omega)\left(\hat{z}(\omega) - \hat{z}_{\text{free}}(\omega)\right) \tag{5}$$

With the inverse Fourier transform, the time-dependent force acting on the tip can be readily determined from Equation 5.

Since the result of an experiment is a vector $\underline{z}$ of time-discrete samples of the continuous signal $z(t)$ during a time window of length $T = 1/\Delta\omega$, the Fourier transform can be expressed using a unitary matrix $\underline{\underline{\mathcal{F}}}$,

$$\underline{\hat{z}} = \underline{\underline{\mathcal{F}}}\,\underline{z} \tag{6}$$

$$\underline{z} = \underline{\underline{\hat{\mathcal{F}}}}^{-1}\,\underline{\hat{z}} \tag{7}$$

where a single underline denotes a vector and a double underline a matrix. In a real experiment, only a partial motion spectrum $\underline{\hat{z}}_{\text{m}}$ can be measured since the linear transfer function of the cantilever $\hat{\chi}$ suppresses the response far away from reso-

nance. Mathematically, this can be expressed with a diagonal windowing matrix operator $\hat{\underline{\underline{W}}}$ that sets all frequency components outside the resonant detection band to zero. The measured spectrum is then given by

$$\hat{\underline{z}}_{\mathrm{m}} = \hat{\underline{\underline{W}}}\ \hat{\underline{z}} \tag{8}$$

and Equation 5 for the measured data becomes

$$\hat{\underline{\underline{\chi}}}^{-1}\left(\hat{\underline{z}}_{\mathrm{m}} - \hat{\underline{z}}_{\mathrm{free,m}}\right) = \hat{\underline{\underline{W}}}\ \hat{\underline{F}}_{\mathrm{ts}} \tag{9}$$

Since $\hat{W}$ is not invertable we cannot determine the complete force spectrum $\hat{F}_{\mathrm{ts}}$ from Equation 9, and thus the time-dependent force remains unknown. To reconstruct the complete force spectrum from the measured partial motion spectrum $\hat{z}_{\mathrm{m}}$, we expand the tip–surface force into a finite series from a set of basis functions, $f_n(z,\dot{z})$ with constant coefficients g_{n}

$$F_{\mathrm{ts}}(z,\dot{z}) = \sum_{n=0}^{N} f_n(z,\dot{z}) g_n \tag{10}$$

Our assumption that the force can be expanded in this manner results in a correlation of the unknown frequency components of the motion with the measurable components. A common choice for the functions f_n to model conservative forces are monomials [17]

$$f_n(z,\dot{z}) = (z-h)^n \tag{11}$$

but also other basis functions of the form

$$f_n(z,\dot{z}) = \begin{cases} (z-h)^{n/2} & n \text{ even} \\ \dot{z}(z-h)^{(n+1)/2} & n \text{ odd} \end{cases} \tag{12}$$

for the representation of position-dependent viscosites have been considered [19]. For a measured tip motion the force vector $\underline{F}_{\mathrm{ts}}$ can the be approximated as

$$\underline{F}_{\mathrm{ts}} \simeq \underline{\underline{H}}\ \underline{g} \tag{13}$$

where the coupling matrix H is given by

$$\underline{\underline{H}} = \underline{\underline{H}}(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) = \begin{pmatrix} \vdots & \vdots & & \vdots \\ \underline{f}_0(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) & \underline{f}_1(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) & \cdots & \underline{f}_N(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) \\ \vdots & \vdots & & \vdots \end{pmatrix} \tag{14}$$

in which the columns are formed by the vectors $\underline{f}_0, \underline{f}_1, \ldots, \underline{f}_n$ evaluated at the measured discrete tip positions and velocities. Here, we assume that the measured, or windowed tip motion $\underline{z}_{\mathrm{m}}$ is a good approximation of the complete tip motion $\underline{z}$. In Fourier space, Equation 13 becomes

$$\hat{\underline{F}}_{\mathrm{ts}} = \underline{\underline{\mathcal{F}}}\ \underline{\underline{H}}\ \underline{g} \equiv \hat{\underline{\underline{H}}}\ \underline{g} \tag{15}$$

where

$$\hat{\underline{\underline{H}}} = \hat{\underline{\underline{H}}}(z_{\mathrm{m}}, \dot{z}_{\mathrm{m}}) = \begin{pmatrix} \vdots & \vdots & & \vdots \\ \underline{\underline{\mathcal{F}}}\ \underline{f}_0(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) & \underline{\underline{\mathcal{F}}}\ \underline{f}_1(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) & \cdots & \underline{\underline{\mathcal{F}}}\ \underline{f}_N(\underline{z}_{\mathrm{m}}, \underline{\dot{z}}_{\mathrm{m}}) \\ \vdots & \vdots & & \vdots \end{pmatrix} \tag{16}$$

The force matrix Equation 9 can then be written as

$$\hat{\underline{\underline{W}}}\ \hat{\underline{\underline{H}}}\ \underline{g} = \hat{\underline{\underline{\chi}}}^{-1}\left(\hat{\underline{z}}_{\mathrm{m}} - \hat{\underline{z}}_{\mathrm{free,m}}\right). \tag{17}$$

We introduce

$$\hat{\underline{\underline{H}}}_{\mathrm{m}} = \hat{\underline{\underline{W}}}\ \hat{\underline{\underline{H}}} \tag{18}$$

and solve Equation 17 for $\underline{g}$ such that

$$\underline{g} = \hat{\underline{\underline{H}}}_{\mathrm{m}}^{+}\ \hat{\underline{\underline{\chi}}}^{-1}\left(\hat{\underline{z}}_{\mathrm{m}} - \hat{\underline{z}}_{\mathrm{free,m}}\right) \tag{19}$$

where $\hat{\underline{\underline{H}}}_{\mathrm{m}}^{+}$ denotes the pseudo-inverse of $\hat{\underline{\underline{H}}}_{\mathrm{m}}$. If a unique solution for the coefficient vector $\underline{g}$ exists, the matrix $\hat{\underline{\underline{H}}}_{\mathrm{m}}^{+}$ equals the exact inverse of $\hat{\underline{\underline{H}}}_{\mathrm{m}}$. If there is more than one solution for $\underline{g}$, Equation 19 computes the solution for which the vector $\underline{g}$ has minimum length. If no solution to Equation 17 exists the pseudo-inverse $\hat{\underline{\underline{H}}}_{\mathrm{m}}^{+}$ approximates the inverse in a least-squares sense.

The matrix $\hat{\underline{\underline{H}}}_{\mathrm{m}}$ can be rapidly computed from Equation 16 using the Fast Fourier Transform (FFT) algorithm. Therefore, Equation 19 provides an efficient way to determine the expansion coefficients $\underline{g}$ of the the tip–surface force. However, special care should be taken to avoid aliasing effects due to the finite sampling of the data. To increase the numerical stability of 19 it is advantageous to normalize $\underline{z}_{\mathrm{m}}$ and $\underline{\dot{z}}_{\mathrm{m}}$ such that the largest absolute value of any vector element is 1. This normalization can be interpreted as a preconditioning procedure for the matrix $\hat{\underline{\underline{H}}}_{\mathrm{m}}$.

To further investigate what information about the tip–surface force can be extracted, we focus on the monomial expansion basis defined in Equation 11 and the case of narrow-band ImAFM where the windowing matrix is given by

$$\hat{W}_{ij} = \begin{cases} \delta_{ij} & k_1 \le i \le k_M \\ 0 & \text{else} \end{cases} \tag{20}$$

with δ_{ij} being the Kronecker delta, $k_1\Delta\omega$ the lower frequency limit of the resonant detection band and $k_M\Delta\omega$ the upper limit. In Figure 1 we plot the absolute values of the components of the matrix $\hat{\underline{\underline{H}}}_{\mathrm{m}}$ for experimental data. One could imagine applying different windowing matrices when building $\hat{\underline{\underline{H}}}_{\mathrm{m}}$, for example one which is weighted by the signal-to-noise ratio at each frequency.

The absolute value of the element of $\hat{\underline{\underline{H}}}_{\mathrm{m}}$ with the index (i, j) is a measure of how much the j-th element g_j of the expansion coefficient vector $\underline{g}$, contributes to the force at the i-th frequency $i\Delta\omega$ in the force spectrum $\hat{F}_{\mathrm{ts}}$. It is apparent from Figure 1 that only polynomial coefficients of odd order contribute to the force measured in the resonant detection band when two drives close to resonance are used in ImAFM. Thus, Equation 19 only yields the odd coefficients in the polynomial force expansion and the resulting polynomial force is odd with respect to $z = h$. To determine the missing even coefficients we assume that the tip–surface force is zero for $z \geq z_{\text{non–interacting}}$. With this assumption we fit the even polynomial coefficients while keeping the odd coefficients constant. This reconstruction method has been extensively tested and its accuracy verified with simulated data [19]. In the following we will show results for experimental data.

Polynomial force reconstruction during slow surface approach

To demonstrate the capabilities of the polynomial force reconstruction we perform an ImAFM approach measurement on a silicon oxide surface. In this measurement two drive frequencies close to resonance result in a beat-like tip motion, with rapid sinusoidal oscillations and a slowly varying amplitude. The AFM z-piezo moves slowly towards the surface, such that during one beat period the static tip height above the surface can be considered to be constant.

Figure 2 shows one frame from a movie (Supporting Information File 1) visualizing the measurement. For four consecutive beats in the time domain the corresponding amplitude spectrum around the first resonance is displayed in Figure 2b where the components or partial spectrum used for force reconstruction are marked with red circles. The polynomial force reconstruction is plotted (yellow solid line) in Figure 2c together with an ADFS reconstruction using the same data (red circles). In Figure 2d the amplitudes of the tip motion at the lower (red) and the higher drive frequency (yellow) are shown as functions of the z-piezo extension, and the vertical blue line indicates the current z-piezo extension.

Far away from the surface the tip does not experience any surface force and the motion spectrum exhibits response only at the driven frequencies (Supporting Information File 1). Consequently, the reconstructed force is zero. As the surface is approached, the attractive force regime due to the van der Waals forces between the tip and the surface is reached. In this regime new frequency components appear in the motion spectrum, so-called intermodulation products. Note that in the time domain, the distortion of the signal is barely visible. Both poly-

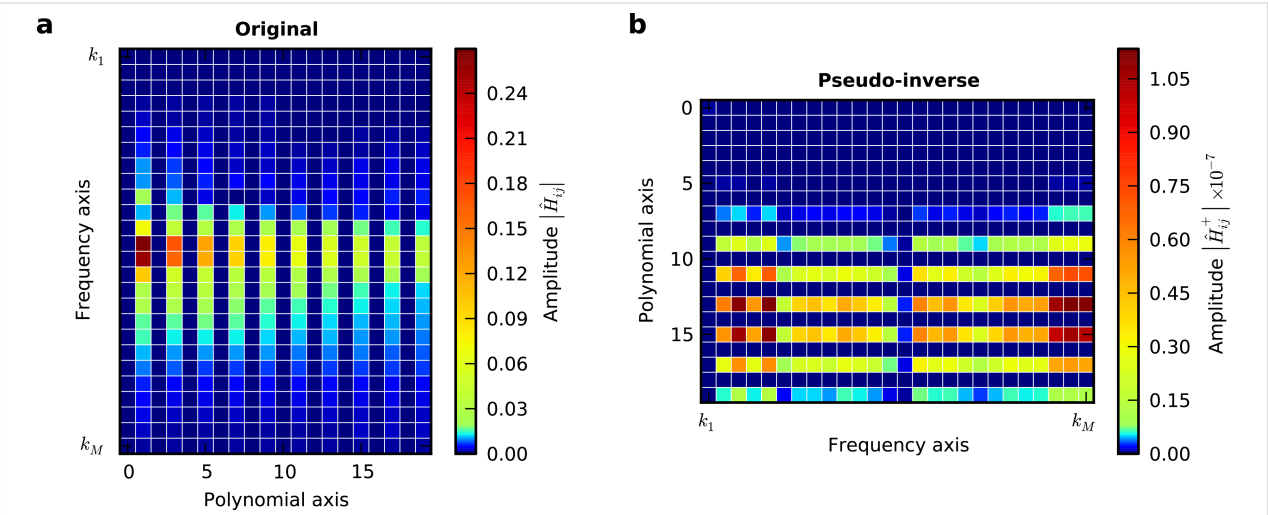

Figure 1: Visualization of the $\hat{\underline{\underline{H}}}_{\mathrm{m}}$ matrix (a) and its pseudo-inverse (b). Only rows with non-zero elements are displayed.

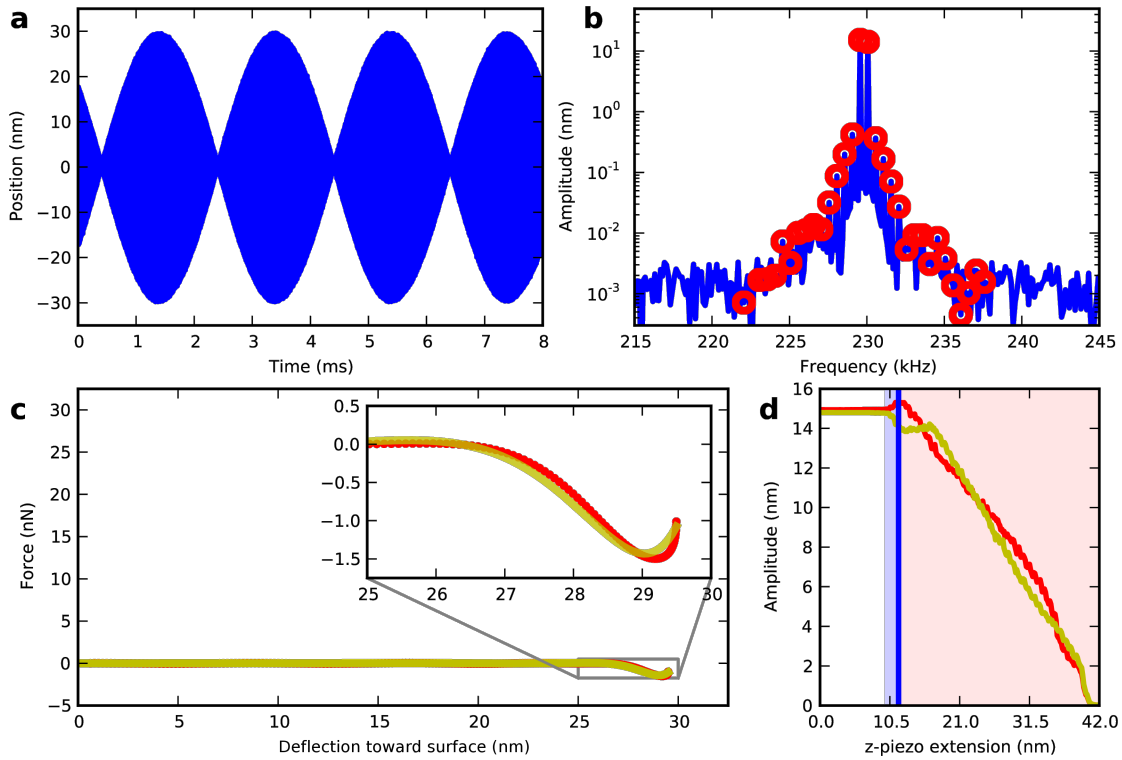

Figure 2: One frame from a surface-approach movie (see Supporting Information File 1) showing the very onset of repulsive forces. The beating waveform (a) has the intermodulation spectrum around resonance (b) where the red circles are analyzed to reconstruct the tip–surface force (c) using both polynomial (yellow) and ADFS (red) methods. The z-piezo extension z_{piezo} = 11.8 nm is indicated by the blue vertical line in (d) which displays the amplitudes at the two drive frequencies. The interaction is purely attractive in the blue shaded area, becoming repulsive in the red shaded area.

nomial and ADFS reconstruction show a growing attractive interaction until a force minimum of −1.75 nN is reached at a piezo extension of 11.8 nm. At this point the tip experiences hard mechanical impacts on the sample surface near the beat maximum, which are manifest in the sharp onset of repulsive force in the polynomial and ADFS reconstruction Figure 2c.

As the z-piezo further extends the tip indents deeper into the surface and experiences stronger repulsive forces as shown in Figure 3 where one frame of the movie at a piezo extension of 18.7 nm is shown.

In this repulsive regime the polynomial and the ADFS reconstruction agree very well. However, the force minimum has a slightly sharper shape with the ADFS reconstruction, which is not constrained to be continuous in high-order derivatives, as for a polynomial. The repulsive force reaches its maximum of 31 nN at a z-piezo extension of 23.6 nm. Moving closer to the surface the maximum force during one beat decreases until the oscillation vanishes.

During the whole surface approach the polynomial and the ADFS force reconstruction agree very well, indicating that the polynomial reconstruction accurately reproduces the force. The shape of the reconstructed force is very stable during the entire approach for both polynomial and ADFS reconstruction, which is a result of the high signal-to-noise ratio for the measured frequency components close to resonance. The stability of the reconstruction during approach gives us confidence in the method's ability to accurately reconstruct sharp features in the force curve, such as the force minimum.

Surface parameter mapping

Material scientists are often interested in determining surface properties with high spatial resolution. ImAFM measurements can be performed while scanning a sample surface [24,25], enabling polynomial force reconstruction at every point of an AFM image. A specific force model can then be fitted to the complete force curve or parts thereof, generating a map of the model parameters over the complete surface. To demonstrate this we scanned the surface of a PS/PMMA blend with ImAFM. To the repulsive part of the polynomial force reconstruction we fit a Derjaguin–Muller–Toropov (DMT) force model [26] of the form $F_{rep}(z) = F_{min} + \varepsilon(z - z_{min})^{3/2}$ where z_{min} is the position of the force minimum F_{min} and ε is the DMT stiffness factor, which depends on the tip radius and the effective stiffness of the

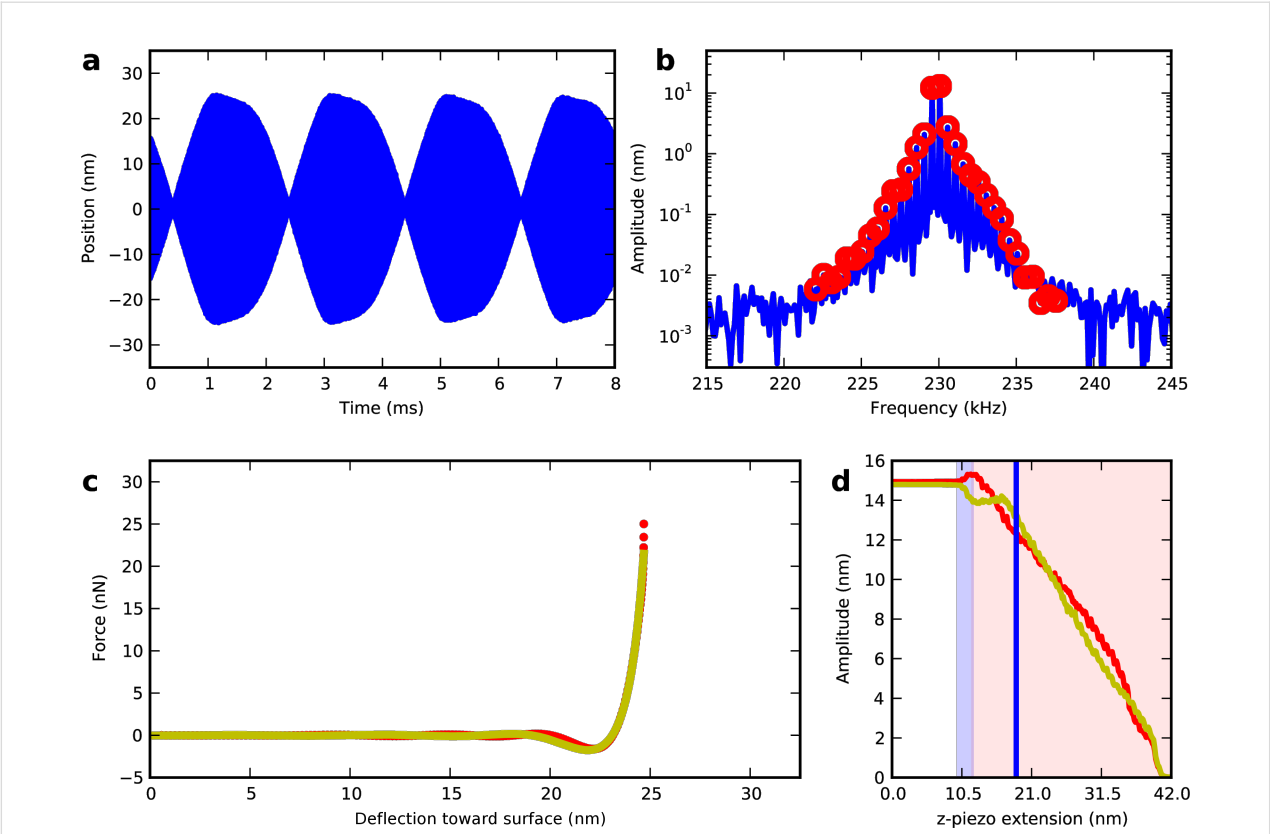

Figure 3: One frame from a surface approach movie (see Supporting Information File 1) showing strongly repulsive forces. The beating waveform (a) has the intermodulation spectrum around resonance (b) where red circles are analyzed to reconstruct the tip–surface force (c) using both polynomial (yellow) and ADFS (red) methods. The z-piezo extension z_{piezo} = 18.7 nm is indicated by the blue vertical line in (d), which displays the amplitudes at the two drive frequencies. The interaction is purely attractive in the blue shaded area, becoming repulsive in the red shaded area.

tip–surface system. One should be aware of the fact that macroscopic force models, such as the DMT model, might not be applicable on the nanoscale [27] and that tip shape and surface topography lead to an interaction geometry that is different from the model geometry of a perfect sphere and a perfectly flat plane [20]. Moreover, the DMT model does not account for adhesive forces in the contact regime, which we try to circumvent by using tips with small radii. While the DMT model provides sufficient insight into material properties, the extracted numerical values of the DMT parameters should not be expected to agree with values for the bulk material.

In Figure 4 a map of the DMT stiffness factor is shown. Even though the two polymers are very similar in stiffness at room temperature [28], two domains of different stiffness are clearly visible in the stiffness-factor map. The stiffer domains are PMMA-rich and 10 nm higher than the surrounding matrix, which is PS-rich and is a factor of two softer than the PMMA-domains. Similar results on the same model polymer system have been obtained with methods employing higher harmonics [15,29] and by ADFS [20].

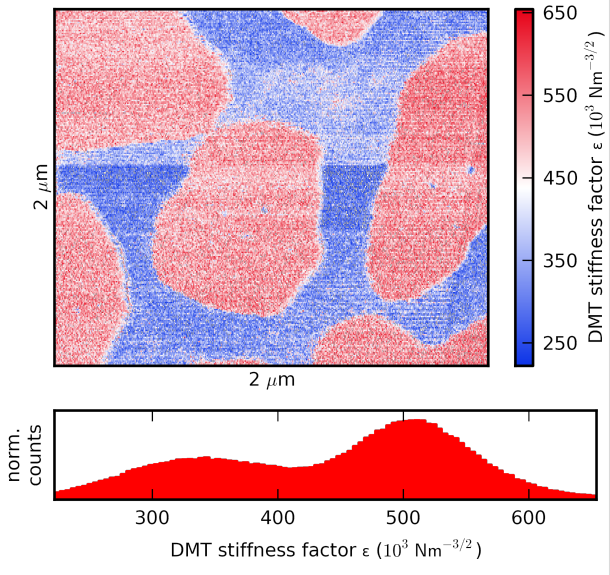

Figure 4: A surface property map showing the DMT stiffness factor ε [N m$^{-3/2}$] with a histogram of the stiffness factor.

Polynomial reconstruction from force quadrature data

Polynomial force reconstruction is an approximative reconstruction method applied to spectral data obtained from Fourier analysis of the tip motion. The general idea, to determine the parameters of a force model such that an experimental observable is best approximated, is not limited to spectral data. An alternative to the spectral representation is a picture that represents the data in terms of two force quadratures. The force quadrature F_I is the component of the (time-dependent) tip–surface force that is in phase with the sinusoidal tip motion, the quadrature F_Q is the force component that is phase-shifted 90 degrees to the tip motion [21].

$$F_I(A,\bar{\omega},h)=\frac{1}{T}\int_0^T F_{\rm ts}\left(A\cos(\bar{\omega}t)+h,-\bar{\omega}A\sin(\bar{\omega}t)\right)\cos(\bar{\omega}t)dt \quad (21)$$

$$F_Q(A,\bar{\omega},h)=\frac{1}{T}\int_0^T F_{\rm ts}\left(A\cos(\bar{\omega}t)+h,-\omega A\sin(\bar{\omega}t)\right)\sin(\bar{\omega}t)dt \quad (22)$$

where T is the measurement time. The force quadratures are functions of the oscillation amplitude A, the oscillation frequency $\bar{\omega}$ and the static tip height h, all of which are constant during each oscillation cycle. However, here we consider only the amplitude dependence of F_I and F_Q, which can be rapidly measured with ImAFM using a single-oscillation-cycle analysis that is based on a separation of time scales [21].

The representation of the measurement result in terms of the force quadratures F_I and F_Q has the advantage that they are directly connected to the tip–surface force and independent of the actual complicated multifrequency tip motion. With spectral data, certain points on the tip–surface force curve will receive greater weight if the tip spends more time at these positions. On the $F_I(A)$ and $F_Q(A)$ curves the weight at each amplitude can be controlled by design. Furthermore, distortions due to feedback artifacts can easily be removed from the $F_I(A)$ and $F_Q(A)$ curves, and both conservative and dissipative forces can be analyzed separately.

To demonstrate approximative force reconstruction on force quadrature data, we consider again a conservative polynomial force representation as in Equation 11. For such a force and $h = 0$, $F_Q(A) = 0$ and Equation 21 becomes

$$F_I(A)=\frac{1}{T}\sum_{n=0}^{N} g_n A^n\int_0^T\cos^{n+1}(\bar{\omega}t)dt \quad (23)$$

The integral is nonzero only for odd n and by using

$$\cos^m(\theta)=\frac{1}{2^m}\binom{m}{m/2}+\frac{1}{2^{m-1}}\sum_{k=0}^{\frac{m}{2}-1}\binom{m}{k}\cos\left((m-2k)\theta\right), \quad (24)$$

where $\binom{m}{k}$ is the binomial coefficent ,

Equation 23 becomes

$$F_I(A)=\sum_{\substack{n=0\\ n\ \text{odd}}}^{N}\frac{1}{2^{n+1}}\binom{n+1}{\frac{n+1}{2}}g_nA^n\equiv\sum_{\substack{n=0\\ n\ \text{odd}}}^{N}\tilde{g}_nA^n \quad (25)$$

which implies that the odd polynomial coefficients of the force expansion can be obtained by simple rescaling of the coefficients of a polynomial approximation of the $F_I(A)$ curve,

$$g_n=\left[\frac{1}{2^{n+1}}\binom{n+1}{\frac{n+1}{2}}\right]^{-1}\tilde{g}_n \quad (26)$$

To obtain the even coefficients we apply the same algorithm as for reconstruction from spectral data.

We obtain a polynomial approximation of $F_I(A)$ by a simple polynomial fit to equidistant discrete points on the $F_I(A)$ curve. Alternative methods such as expansion in orthogonal polynomials or different types of interpolation polynomials with different convergence properties can also be applied. The polynomial reconstruction based on force quadrature data can be implemented even more efficiently than the reconstruction on spectral data since multiple Fourier transforms to construct the coupling matrix $\underline{\underline{\hat{H}}}_{\rm m}$ in Equation 16 are not required.

To validate the equivalence of polynomial force reconstruction on spectral and force quadrature data, we consider the ImAFM approach measurement on silicon oxide described above. From the tip motion at a z-piezo extension of 25.6 nm we compute the $F_I(A)$ curve and remove all data points for which the oscillation amplitude was decreasing. From the polynomial approximation of the $F_I(A)$ curve we obtain the force polynomial as described above. The resulting reconstruction is shown in Figure 5 (blue line) together with the polynomial reconstruction from spectral data (yellow line) and an ADFS reconstruction (red circles). Over the full range of oscillation both curves agree well. Good agreement is also observed at all other z-piezo extensions.

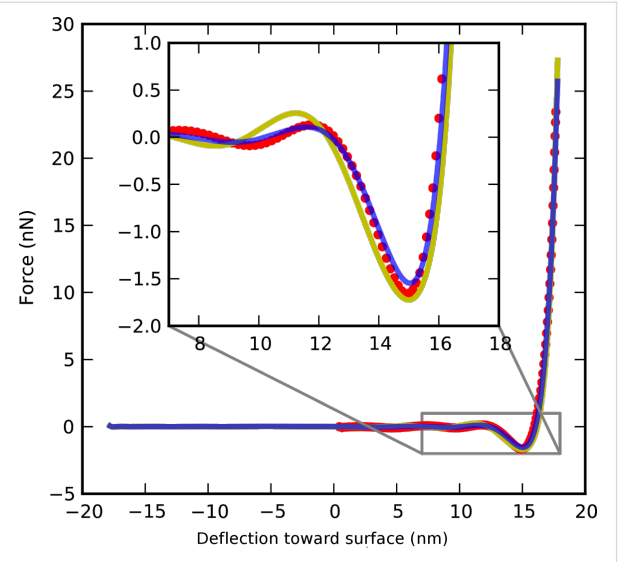

Figure 5: Comparison of force reconstructions for a tip interacting with silicon oxide. Polynomial reconstruction from spectral data (yellow), force quadrature data (blue) and ADFS reconstruction (red circles) are in good agreement.

Conclusion

Multifrequency AFM opens the window to a wide variety of novel AFM measurement techniques enabling a much improved understanding of the forces between the tip and the surface. We showed that polynomial force reconstruction is an intuitive and powerful method to approximate this interaction, and we demonstrated the method's use for accurate and detailed force measurement and for high-resolution surface-parameter mapping with experimental data.

As the field of multifrequency AFM continues to evolve, new alternative data-representation schemes can help to simplify analysis and extract more specific properties of the tip–surface interaction. The force-quadrature picture is such a scheme, which decouples information about the tip–surface interaction from the actual tip motion. We showed how the general idea behind approximative force reconstruction can be adapted to the force-quadrature picture, and we introduced an efficient way to extract the polynomial coefficients from the force quadratures.

We hope that in the future polynomial force reconstruction will be a useful method for many scientists and that new data-representation schemes will inspire innovative analysis methods.

Experimental

The silicon oxide sample was cleaned in an oxygen plasma before measurements were perfomed in a Bruker Dimension 3100 AFM system. The cantilever (Bruker MPP-11120) was calibrated by a noninvasive thermal method [30] and had a resonance frequency of f_0 = 229.802 kHz, a quality factor of Q = 396.9 and a spring constant of k_c = 16.0 N m^{-1}. The slow surface approach velocity was 2 nm s^{-1}.

PS (M_w = 280 kDa, Sigma-Aldrich) and PMMA (M_w = 120 kDa, Sigma-Aldrich) were spin-cast from toluene solution with a concentration of 0.53 %wt at a ratio of 3:1 (PMMA:PS). The sample was scanned in a Bruker Multimode 2 AFM system with a cantilever BS 300Al-G (Budget Sensors) having a resonance frequency f_0 = 343.379 kHz, quality factor Q = 556.9 and spring constant k_c = 35.1 N m^{-1}. The maximum free oscillation amplitude close to the surface was 30 nm and we scanned an image with 256 × 1024 pixels within 17 min.

For all measurements we used an intermodulation lock-in analyzer (IMP 2-32, Intermodulation Products AB) which synchronizes the signal generation and acquisition for measurement of the multifrequency response [31].

Supporting Information

Supporting Information File 1

A movie showing the tip motion and the reconstructed tip–surface during an ImAFM approach measurement.
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-4-41-S1.mp4]

Acknowledgements

The authors acknowledge financial support from the Knut and Alice Wallenberg Foundation, the Swedish Research Council (VR), the Swedish Government Agency for Innovation Systems (VINNOVA), and the Olle Engkvist Foundation.

References

1. Ohnesorge, F.; Binnig, G. *Science* **1993,** *260,* 1451–1456. doi:10.1126/science.260.5113.1451
2. Giessibl, F. J. *Science* **1995,** *267,* 68–71. doi:10.1126/science.267.5194.68
3. Fukuma, T.; Kobayashi, K.; Matsushige, K.; Yamada, H. *Appl. Phys. Lett.* **2005,** *86,* 193108. doi:10.1063/1.1925780
4. Gan, Y. *Surf. Sci. Rep.* **2009,** *64,* 99–121. doi:10.1016/j.surfrep.2008.12.001
5. Dürig, U. *Appl. Phys. Lett.* **2000,** *76,* 1203. doi:10.1063/1.125983
6. Sader, J. E.; Uchihashi, T.; Higgins, M. J.; Farrell, A.; Nakayama, Y.; Jarvis, S. P. *Nanotechnology* **2005,** *16,* S94–S101. doi:10.1088/0957-4484/16/3/018
7. Hölscher, H. *Appl. Phys. Lett.* **2006,** *89,* 123109. doi:10.1063/1.2355437
8. Lee, M.; Jhe, W. *Phys. Rev. Lett.* **2006,** *97,* 036104. doi:10.1103/PhysRevLett.97.036104
9. Hu, S.; Raman, A. *Nanotechnology* **2008,** *19,* 375704. doi:10.1088/0957-4484/19/37/375704
10. Katan, A. J.; van Es, M. H.; Oosterkamp, T. H. *Nanotechnology* **2009,** *20,* 165703. doi:10.1088/0957-4484/20/16/165703

11. García, R.; Herruzo, E. T. *Nat. Nanotechnol.* **2012,** *7,* 217–226. doi:10.1038/nnano.2012.38
12. Dürig, U. *New J. Phys.* **2000,** *2,* 5. doi:10.1088/1367-2630/2/1/005
13. Stark, M.; Stark, R. W.; Heckl, W. M.; Guckenberger, R. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 8473–8478. doi:10.1073/pnas.122040599
14. Legleiter, J.; Park, M.; Cusick, B.; Kowalewski, T. *Proc. Natl. Acad. Sci. U. S. A.* **2006,** *103,* 4813–4818. doi:10.1073/pnas.0505628103
15. Sahin, O.; Magonov, S.; Su, C.; Quate, C. F.; Solgaard, O. *Nat. Nanotechnol.* **2007,** *2,* 507–514. doi:10.1038/nnano.2007.226
16. Forchheimer, D.; Platz, D.; Tholén, E. A.; Haviland, D. B. *Phys. Rev. B* **2012,** *85,* 195449. doi:10.1103/PhysRevB.85.195449
17. Hutter, C.; Platz, D.; Tholén, E. A.; Hansson, T. H.; Haviland, D. B. *Phys. Rev. Lett.* **2010,** *104,* 050801. doi:10.1103/PhysRevLett.104.050801
18. Raman, A.; Trigueros, S.; Cartagena, A.; Stevenson, A. P. Z.; Susilo, M.; Nauman, E.; Contera, S. A. *Nat. Nanotechnol.* **2011,** *6,* 809–814. doi:10.1038/nnano.2011.186
19. Platz, D.; Forchheimer, D.; Tholén, E. A.; Haviland, D. B. *Nanotechnology* **2012,** *23,* 265705. doi:10.1088/0957-4484/23/26/265705
20. Platz, D.; Forchheimer, D.; Tholén, E. A.; Haviland, D. B. *Nat. Commun.* **2013,** *4,* 1360. doi:10.1038/ncomms2365
21. Platz, D.; Forchheimer, D.; Tholén, E. A.; Haviland, D. B. *Beilstein J. Nanotechnol.* **2013,** *4,* 45–56. doi:10.3762/bjnano.4.5
22. Rodríguez, T. R.; García, R. *Appl. Phys. Lett.* **2002,** *80,* 1646. doi:10.1063/1.1456543
23. Melcher, J.; Hu, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 053101. doi:10.1063/1.2767173
24. Platz, D.; Tholén, E. A.; Pesen, D.; Haviland, D. B. *Appl. Phys. Lett.* **2008,** *92,* 153106. doi:10.1063/1.2909569
25. Platz, D.; Tholén, E. A.; Hutter, C.; von Bieren, A. C.; Haviland, D. B. *Ultramicroscopy* **2010,** *110,* 573–577. doi:10.1016/j.ultramic.2010.02.012
26. Derjaguin, B. V.; Muller, V. M.; Toporov, Yu. P. *J. Colloid Interface Sci.* **1975,** *53,* 314–326. doi:10.1016/0021-9797(75)90018-1
27. Luan, B.; Robbins, M. O. *Nature* **2005,** *435,* 929–932. doi:10.1038/nature03700
28. Brandrup, J.; Immergut, E. H.; Grulke, E. A.; Abe, A.; Bloch, D. R. *Polymer Handbook,* 4th ed.; Wiley & Sons, 2005.
29. Sahin, O.; Erina, N. *Nanotechnology* **2008,** *19,* 445717. doi:10.1088/0957-4484/19/44/445717
30. Higgins, M. J.; Proksch, R.; Sader, J. E.; Polcik, M.; Mc Endoo, S.; Cleveland, J. P.; Jarvis, S. P. *Rev. Sci. Instrum.* **2006,** *77,* 013701. doi:10.1063/1.2162455
31. Tholén, E. A.; Platz, D.; Forchheimer, D.; Schuler, V.; Tholén, M. O.; Hutter, C.; Haviland, D. B. *Rev. Sci. Instrum.* **2011,** *82,* 026109. doi:10.1063/1.3541791

License and Terms

Beilstein Journal of Nanotechnology

Optimal geometry for a quartz multipurpose SPM sensor

Julian Stirling

Full Research Paper

Open Access

Address:
School of Physics and Astronomy, The University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom

Email:
Julian Stirling - ppxjs1@nottingham.ac.uk

Keywords:
atomic force microscopy; lateral force microscopy; lateral forces; mechanical vibrations; scanning probe microscopy; scanning tunnelling microscopy

Beilstein J. Nanotechnol. **2013,** *4,* 370–376.
doi:10.3762/bjnano.4.43

Received: 06 April 2013
Accepted: 21 May 2013
Published: 17 June 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

We propose a geometry for a piezoelectric SPM sensor that can be used for combined AFM/LFM/STM. The sensor utilises symmetry to provide a lateral mode without the need to excite torsional modes. The symmetry allows normal and lateral motion to be completely isolated, even when introducing large tips to tune the dynamic properties to optimal values.

Introduction

The heart of any scanning probe microscope (SPM) is its sensory probe. For a scanning tunnelling microscope (STM) this is simply an electrically conducting wire with an atomically sharp apex. For atomic force microscopes (AFM) and lateral force microscopes (LFM), however, the sensor is more complex. The atomically sharp probe must be combined with a force sensor, usually a cantilever, with either piezoelectric or optical deflection detection. For noncontact AFM (NC-AFM) and dynamic LFM (DLFM), where the sensor is excited at or near one of its eigenfrequencies, properties such as the Q factor, eigenfrequencies, effective spring constant [1] and other geometrical properties [2] of the eigenmodes become important.

AFM and LFM sensors have evolved from gold foil with diamond tip [3] and bent tungsten wires [4] respectively, into a wide range of specialised sensors. The most common NC-AFM sensors: silicon microcantilvers [5], and quartz sensors such as the qPlus sensor (tuning fork) [6] or KolibriSensor® [7], have all been used for combined AFM/STM [7-9]. Combined AFM/LFM sensors have been constructed from silicon cantilevers, by exciting torsional modes to generate the lateral motion needed for the LFM [10]. The qPlus sensor has been used as an LFM by rotating the tip on the end of the quartz tuning fork [11], but no combined AFM/LFM qPlus system has been developed due to the magnitude of the torsion constant for the tine of the sensor. A combined AFM/LFM sensor operated in frequency

modulation mode would enable measurements of conservative and nonconservative forces simultaneously in the normal and lateral direction. Such measurements could be used to further important investigations in single-asperity friction [12], where the relationship between normal and lateral force is of interest. In this paper, we suggest the optimum geometry of a quartz sensor to produce a combined AFM/LFM/STM from a quartz crystal resonator with many theoretical benefits over other sensors.

Combining NC-AFM and DLFM

For a sensor to image as both an NC-AFM and a DLFM the sensor must be able to oscillate both normal to and parallel to the surface it is scanning. The simplest method for achieving this is a single oscillator which will oscillate in different directions depending on the eigenmode excited. Ideally for atomic-resolution imaging the effective spring constant of the excited eigenmode should be low [13]. However, as the spring constant normal to the surface lowers, the risk of the probe snapping to contact with the surface increases. This produces a problem for combined AFM/LFM using the principal and first torsional eigenmode of a cantilever, as the torsional mode can have an effective spring constant of up to approximately two orders of magnitude higher than the principal mode [10]. This results in a difficult tradeoff. To avoid snap to contact, the following condition must be satisfied [14]:

$$A_0 k_N + F_{\mathrm{TS,N}} > 0\,, \tag{1}$$

where A_0 and k_N are the amplitude and effective spring constant of the principal eigenmode, and $F_{\mathrm{TS,N}}$ is the tip–sample force normal to the sample.

For high-resolution AFM imaging A_0 should be as low as possible [13]. However, the signal-to-noise ratio, which is a function of A_0, [15] limits the minimum amplitude. Experimentally, some groups have achieved stable imaging with amplitudes as low as 20 pm [16]. Thus, to be safe from snap to contact for atomic forces on the order of −3 nN, it is required that $k_N > 150$ N·m^{-1}. If imaging, however, is only in DLFM mode, then A_0 is ideally zero. Obviously, Equation 1 doesn't hold in this case as it would suggest that we require an infinite spring constant to stop snap to contact. In this example A_0 must be considered as the distance the tip has moved from its equilibrium position due to $F_{\mathrm{TS,N}}$. Therefore, if trying to image in DLFM mode, the error in the z-position due to normal forces is inversely proportional to k_N, requiring higher minimum normal spring constants of $k_N \geq$ 1–3 kN·m^{-1}. This would result in torsional constants on the order of hundreds of kN·m^{-1}, which is not ideal for LFM imaging.

The torque required to torsionally twist a beam of length L through an angle θ is given by

$$T = \frac{\theta JG}{L}\,, \tag{2}$$

where J and G are the torsion constant and shear modulus of the beam. In the case of a cantilever beam with a tip of length L_{tip} (measured from the central axis of the beam), the lateral displacement of the tip apex, A_{lat}, is $L_{\mathrm{tip}}\theta$. Replacing the torque with the lateral tip–sample force $F_{\mathrm{TS,L}}$ multiplied by the tip length we get

$$F_{\mathrm{TS,L}} = \frac{A_{\mathrm{lat}} JG}{LL_{\mathrm{tip}}^2}\,. \tag{3}$$

Hence, the lateral spring constant

$$k_{\mathrm{L}} = \frac{JG}{LL_{\mathrm{tip}}^2}\,, \tag{4}$$

is inversely proportional to the square of the tip length. Thus, the tip length becomes an important parameter to consider alongside the more typical geometrical constants associated with the normal spring constant.

For quartz sensors the obvious choice of cantilever is the standard qPlus sensor with a normal spring constant of approximately 1.8 kN·m^{-1}. [17,18] For commercially available silicon cantilevers the spring constants are usually less than 50 N·m^{-1}, with resonant frequencies of 200–300 kHz. The resonant frequency of the cantilever scales with L^{-4} and the spring constant with L^{-3}. Considering that the torsional eigenfrequency can be approximately two orders of magnitude larger than the normal eigenfrequency, achieving the necessary normal spring constant by length reduction could push the torsional eigenfrequency into the gigahertz range, which is impractical. We instead will consider different rectangular cross sections for a 200 μm long silicon beam.

A range of rectangular cross sections which would produce a normal spring constant of 2 kN·m^{-1} for a 200 μm long beam have been calculated, using the Euler–Bernoulli beam theory, see Table 1. This value was chosen to sit in the middle of the range suggested for the minimum normal spring constant. The frequency of the first eigenmode has also been calculated. Using Equation 4, the tip length needed for $k_{\mathrm{L}} = 2$ kN·m^{-1} was calculated, using previously tabulated values for J [19]. This tip length was also calculated for the qPlus sensor.

Table 1: Dimensions and dynamic properties of silicon microcantilevers that would provide normal spring constants of 2 kN·m^{-1}. L_{tip} is the tip length required to provide a lateral spring constant of 2 kN·m^{-1}. qPlus sensor is also included (the frequency is not 32768 Hz, as some features, such as base deformations and gold contacts, were neglected in the calculations).

Name	L (μm)	w (μm)	T (μm)	k_N (N·m^{-1})	f_N (Hz)	L_{tip} (μm)	L_{tip}/L
Si1	200	112.2	15.0	2 000	515 913	139.1	0.696
Si2	200	47.3	20.0	2 000	687 884	124.6	0.623
Si3	200	24.2	25.0	2 000	859 856	93.1	0.466
Si4	200	14.0	30.0	2 000	1 031 827	57.2	0.286
Si5	200	8.8	35.0	2 000	1 203 798	33.7	0.168
Si6	200	5.9	40.0	2 000	1 375 769	20.5	0.102
qPlus	2 400	130.0	214.0	1 763	32 246	772.9	0.322

The calculated tip lengths range from approximately 10% to 70% of the beam length. As has previously been shown, AFM sensors with tip lengths of similar scale to the length of the beam exhibit a large lateral component to the motion of the tip apex in the first eigenmode [2]. This lateral component is perpendicular to the torsional eigenmode, thus making it impossible to truly separate the normal and lateral forces. This problem is exacerbated if the tip length is further increased to increase sensitivity to lateral forces by reducing the lateral spring constant, as snap to contact is not an issue in the lateral direction. Increasing the ratio of thickness to width reduces the required tip length, but at the expense of introducing normal eigenfrequencies above 1 MHz, pushing torsional eigenfrequencies to ranges that most AFM electronics cannot handle.

Non-cantilever geometries

Due to the large difference between the operating frequencies of normal and torsional modes, and the coupling of unwanted lateral motion into the normal eigenmode for sensors with the tip lengths needed to produce low lateral spring constants, we propose a new sensor geometry. The proposed design, see Figure 1, is to attach a tungsten tip to the centre of a quartz beam. The design exploits the intrinsic symmetry of the sensor to remove any unwanted lateral motion in the principal eigenmode (Figure 1b), thus allowing for longer tips. By exciting the second eigenmode of the beam, lateral motion can be generated (Figure 1c). The effective spring constant and eigenfrequency can be calculated, and thus tuned, far more simply than for torsional modes, by solving the Euler–Bernoulli beam equation with the appropriate boundary conditions. In principle, by also exciting a torsional mode, a perpendicular lateral oscillation could be generated allowing simultaneous measurements in *all three dimensions*. This paper will, however, concentrate on just the first and second eigenmode.

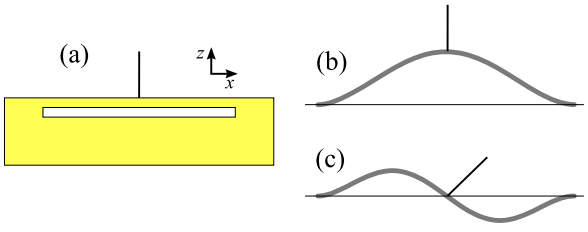

Figure 1: (a) Proposed geometry of new sensor. A tungsten tip connected to the centre of a quartz bar clamped at both ends. (b) First and (c) second eigenmode of the sensor. The symmetry provides pure normal motion in the first eigenmode and lateral motion of the tip apex in the second eigenmode due to the rotation of the tip about the antinode of the beam.

Results and Discussion

Dynamic properties of symmetric sensors

Spring constants

The two most fundamental properties to consider for dynamic force sensors are the effective spring constants and the eigenfrequencies of each imaging mode. For simplicity we will begin with effective spring constants, as the influence of the inertia of the tip has only the effect of moving the dynamic spring constant closer to the static constant [20], removing the ≈3% error. (Note that this is not true in higher eigenmodes for cantilever geometries as the inertia shifts the position of the antinodes [1]. In this system, however, the antinodes are pinned due to the symmetry of the system.)

The dynamic Euler–Bernoulli beam equation

$$EI\frac{\partial^4}{\partial x^4}\left(\sum_{i=1}^{\infty}\Phi_i(x)\mathcal{T}_i(t)\right)=f(x,t)-\rho A\,\frac{\partial^2}{\partial t^2}\left(\sum_{i=1}^{\infty}\Phi_i(x)\mathcal{T}_i(t)\right), \quad (5)$$

describes the dynamic deformations of a beam, where E and ρ are the Young's modulus and density of the material, respectively. A and I are the area and second moment of area of the cross section of the beam. $f(x,t)$ is the applied force per unit length acting on the beam, $\Phi_i(x)$ and $\mathcal{T}_i(t)$ are the spatial and temporal components of the beam's deformation for the ith eigenmode.

As any effect from the tip must be considered at the centre of the beam we will consider only one half of the beam and use symmetry (or antisymmetry in the case of even eigenmodes) to construct the full spatial solution. For both even and odd modes the boundary conditions

$$\Phi_i(0) = 0 \tag{6}$$

$$\frac{\partial \Phi_i(0)}{\partial x} = 0 \tag{7}$$

are valid. Equation 5 is spatially fourth order, therefore two further conditions are required. For odd modes:

$$\frac{\partial \Phi_{i,\mathrm{odd}}\left(L/2\right)}{\partial x} = 0 \tag{8}$$

$$EI\frac{\partial^3 \Phi_{i,\mathrm{odd}}\left(L/2\right)}{\partial x^3} = \frac{F_{\mathrm{norm}}}{2}, \tag{9}$$

and for even modes:

$$\Phi_{i,\mathrm{even}}\left(L/2\right) = 0 \tag{10}$$

$$EI\frac{\partial^2 \Phi_{i,\mathrm{even}}\left(L/2\right)}{\partial x^2} = -\frac{T}{2}. \tag{11}$$

L is the length of the entire beam, and F_{norm} and T are the force and torque applied to the centre of the beam.

Entering these boundary conditions into the general static spatial solution of Equation 5 (i.e., the final term is zero), gives the spring constant of the first eigenmode as

$$k_{\mathrm{norm}} = \frac{192EI}{L^3}. \tag{12}$$

Considering the torque on the centre of the beam $T = L_{\mathrm{tip}}F_{\mathrm{lat}}$, where F_{lat} is a lateral force applied in the x direction at the far end of the tip, the effective lateral spring constant of the second eigenmode is

$$k_{\mathrm{lat}} = \frac{16EI}{LL_{\mathrm{tip}}^2}. \tag{13}$$

Full derivations are provided in Supporting Information File 1.

Thus, just as in the case of the torsional mode, the effective spring constant in the lateral mode can be tuned by tip length. However, due to the symmetry of the sensor this will not cause unwanted lateral motion at the tip apex in the first eigenmode.

Eigenfrequencies

When considering the eigenfrequencies of the sensor, the inertia of the tip plays a very strong role, which cannot be ignored. Solving Equation 5 for the dynamic case, the same boundary conditions (Equation 6–Equation 11) hold, where

$$F_{\mathrm{norm}} = -m_{\mathrm{tip}}\omega_{i,\mathrm{odd}}^2\Phi_{i,\mathrm{odd}}\left(L/2\right) \tag{14}$$

and

$$T = -\mathcal{I}_{\mathrm{tip}}\omega_{i,\mathrm{even}}^2\frac{\partial \Phi_{i,\mathrm{even}}\left(L/2\right)}{\partial x}. \tag{15}$$

ω_i is the angular eigenfrequency of the ith eigenmode, and $\mathcal{I}_{\mathrm{tip}}$ is the moment of inertia of the tip. By combining the general spatial solution with the four boundary conditions as a matrix equation, equal to a zero vector, we see that resonance occurs when the determinant is equal to zero (See Supporting Information File 1 for full derivation). Giving the following resonance conditions:

$$\begin{aligned}&\cosh\left(\beta_{i,\mathrm{odd}}L/2\right)\sin\left(\beta_{i,\mathrm{odd}}L/2\right)+\\&\cos\left(\beta_{i,\mathrm{odd}}L/2\right)\sinh\left(\beta_{i,\mathrm{odd}}L/2\right)+\\&\frac{1}{2\gamma_{i,\mathrm{odd}}}\left(-1+\cos\left(\beta_{i,\mathrm{odd}}L/2\right)\cosh\left(\beta_{i,\mathrm{odd}}L/2\right)\right)=0,\end{aligned} \tag{16}$$

and

$$\begin{aligned}&\cosh\left(\beta_{i,\mathrm{even}}L/2\right)\sin\left(\beta_{i,\mathrm{even}}L/2\right)-\\&\cos\left(\beta_{i,\mathrm{even}}L/2\right)\sinh\left(\beta_{i,\mathrm{even}}L/2\right)+\\&\frac{\varepsilon_{i,\mathrm{even}}}{2}\left(-1+\cos\left(\beta_{i,\mathrm{even}}L/2\right)\cosh\left(\beta_{i,\mathrm{even}}L/2\right)\right)=0,\end{aligned} \tag{17}$$

where

$$\beta_i^4 = \frac{\rho A\omega_i^2}{EI},\quad \gamma_i = \frac{EI\beta_i^3}{m_{\mathrm{tip}}\omega_i^2},\ \text{and}\quad \varepsilon_i = \frac{\mathcal{I}_{\mathrm{tip}}\omega_i^2}{EI\beta_i}. \tag{18}$$

These equations can be solved numerically in terms of dimensionless quantities ($\beta_i L$, m^*, $\mathcal{I}^*$, discussed in Supporting Information File 1), and dimensions can be added later to get a value for ω_i. In the case of no tip, the ratio between the second and first eigenmode is 2.757. Whether this ratio rises or falls when a tip is added depends on the dimensions of both the tip and sensor. It is clear, however, that such a low ratio between the eigenmodes is another advantage of the symmetric sensor over torsional designs as both modes can be tuned to near the optimal frequency of the detection system.

Optimal geometry for a symmetric sensor

In order for the sensor to be used in currently available commercial UHV combined AFM/STM systems, it should be similar in size to the qPlus sensor. However, as the normal spring constant per unit beam length (with the same cross section) is 64 times higher than for a cantilever geometry, a greater length than the 2.4 mm beam of the qPlus sensor is advisable.

Choosing a 3 mm long beam and a normal spring constant of 2 kN·m^{-1}, as previously suggested, we calculate that the second moment of area of the cross section should be $I = 3.68 \times 10^{-18}$ m^4. A width (y-direction) w of 100 µm would result in a thickness (z-direction) of $t = 76.1$ µm, as $I = wt^3/12$. Such a beam would have first and second eigenfrequencies of 46.7 kHz and 128.8 kHz, respectively. These frequencies will reduce when the tip is added to the centre of the beam.

Before considering the mass or moment of inertia of the tip and its effect on the eigenfrequencies of the sensor, it is important to consider the spring constant of the tip itself. Any bending of the tip will not be detected by the piezoelectric quartz sensor. Thus, treating the tip as a cantilever, its spring constant must be much greater than the effective lateral spring constant for the sensor (k_{lat}), otherwise this will result in incorrect force measurements in the LFM mode. We consider a maximum tip length of 1.73 mm, i.e., the length that would give $k_{lat} = 500$ N·m^{-1}; thus, to keep the spring constant of the tip above 10 kN·m^{-1} the diameter of the tungsten wire must be greater than 144 µm.

We will consider a tip diameter, D_{tip}, of 150 µm, an easily available diameter of tungsten wire. The moment of inertia of the tip for the even modes should be calculated about the centre of the beam, $t/2$ from the bottom of the tip, and hence a distance of $(L_{tip} + t)/2$ from the centre of mass of the tip. Thus, the moment of inertia of the tip can simply be calculated by the parallel axis theorem as $\mathcal{I}_{tip} = m_{tip}(D_{tip}^2/16 + L_{tip}^2/3 + L_{tip}t/2 + t^2/4)$. By using Equation 13 and Equation 16–Equation 18, the spring constants and eigenfrequencies of the first two modes have been plotted in Figure 2 for a range of tip lengths. For plotted tip lengths the ratio of the spring constant of the tip to k_{lat} is at its minimum 23.5.

Examining the plot it is clear that tip lengths near 1.47 mm are unusable as the two eigenfrequencies are too close. This would make it difficult to selectively excite them, as well as require long averaging times in bimodal operation to remove any correlation between the modes. The benefit of increasing the tip length is a reduction in lateral spring constant, which comes at the price of lower eigenfrequencies. A tip length of 1mm would

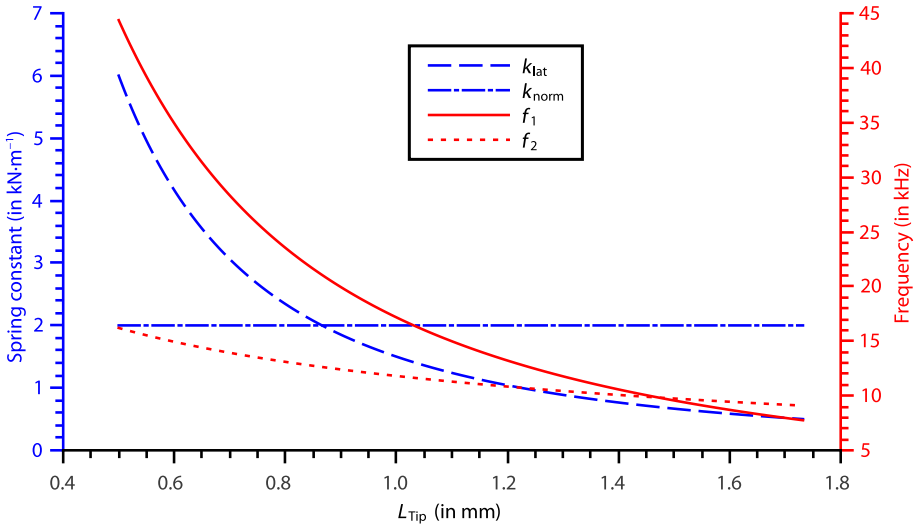

Figure 2: The effective spring constants (k_{lat} for mode 2, and k_{norm} for mode 1) and eigenfrequencies of the first two eigenmodes of a symmetric sensor. Plotted for 150 µm tungsten tips of varying lengths.

provide eigenfrequencies of f_1 = 11.8 kHz and f_2 = 17.2 kHz, with k_{lat}= 1.50 kN·m^{-1}. These frequencies are of the same order of magnitude as qPlus sensors with long tips, thus the sensor could be used in commercially available qPlus systems with no modifications to the electronics.

It is also important to consider the minimum amplitudes achievable by the sensor, particularly in the lateral mode. As little is know about the optimum amplitudes in DLFM, this issue is to be treated approximately. The lateral amplitude of the tip apex is

$$A_{\text{lat}} = L_{\text{tip}} \frac{\mathrm{d}\Phi_2\left(L/2\right)}{\mathrm{d}x}. \tag{19}$$

By considering that the tip is approximately half the length of the beam (L/2), and (dΦ_2(L/2))/dx ≈ (4A_{antinode})/L (approximating Figure 1c as three straight lines), where A_{antinode} is the amplitude of the antinode, the ratio between A_{lat} and A_{antinode} is on the order of 2. Other detection parameters are also of the same order of magnitude as for a qPlus sensor. Thus, as qPlus sensors have achieved imaging with amplitudes as low as 20 pm [16], similar amplitudes are in theory possible for the LFM mode of the symmetric sensor. Such amplitudes are an order of magnitude smaller than inter-atomic distances.

Experimental viability

The experimental viability of this method depends on the equipment available to produce the sensor. Firstly, no quartz crystal resonators of the proposed geometry are commercially available. The closest commercially available sensor is a double-ended tuning fork available from Statek (DETF Force Sensor, http://www.statek.com). By fixing the bottom tine it is possible to reproduce the required symmetry. However, these sensors are too large for most commercial qPlus systems with a total width of 15.2 mm and a beam length of 8.44 mm. Also the beams are recessed with respect to the top of the resonator by 0.86 mm, preventing tip lengths below this. A second possible option would be to attach two identical tuning forks end-to-end by using a similar method to Heyde et al. [21]; however, the glue used to attach the tines will have different mechanical properties to the quartz and also depend on the quantity, placement, and curing conditions. This will affect the repeatability as well as the shape of the eigenmodes, and hence the spring constant. Thus, ideally custom resonators would need to be made. Secondly, the correct placement of the tungsten wire is vital. The tip needs to be positioned in the centre of the 3 mm beam, which is just 100 × 76.1 µm in cross section, and needs to be mounted perfectly normal to the beam. Misplacement of the tip breaks the symmetry thus affecting operation. Reproducible tip placement requires three dimensional micromanipulators, which can be prohibitively expensive for some groups.

A final consideration should be taken regarding the connection of a separate electrode for the tunnel current. Two options are available, first a thin (about 15–50 µm) loose wire could be attached to the tungsten tip, as is often done for qPlus sensors. This is inadvisable as it also breaks the symmetry of the sensor. Another method would be to add a thin insulating layer to the top side of the resonator and on top of that a new electrode, such as the method developed by Nauga Needles [22]. This maintains the symmetry; however, great care needs to be taken to consider the possibility of capacitive cross-talk between the tunnelling and deflection channels [23].

Conclusion

We have demonstrated a new geometry of a piezoelectric sensor for use in combined AFM/LFM, which utilises symmetry to bring the eigenfrequencies and spring constants of the two modes closer together. This allows both modes be tuned to the optimal parameters for operation. The symmetry also removes issues with unwanted lateral motion in normal oscillating modes, allowing longer tips for tuning the lateral spring constant of the LFM operation. By attaching an extra electrode, the sensor can also be used for STM, providing a truly multipurpose SPM sensor.

Supporting Information

Supporting Information File 1

Full derivations of dynamic properties for a symmetric sensor
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-4-43-S1.pdf]

Acknowledgements

The author would like to thank P. Moriarty for fruitful discussions. This work was financially supported by a doctoral training grant from the EPSRC.

References

1. Tung, R. C.; Wutscher, T.; Martinez-Martin, D.; Reifenberger, R. G.; Giessibl, F.; Raman, A. *J. Appl. Phys.* **2010,** *107,* 104508. doi:10.1063/1.3407511
2. Stirling, J.; Shaw, G. A. *Beilstein J. Nanotechnol.* **2013,** *4,* 10–19. doi:10.3762/bjnano.4.2
3. Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986,** *56,* 930–933. doi:10.1103/PhysRevLett.56.930
4. Mate, C. M.; McClelland, G. M.; Erlandsson, R.; Chiang, S. *Phys. Rev. Lett.* **1987,** *59,* 1942–1945. doi:10.1103/PhysRevLett.59.1942

5. Albrecht, T. R.; Quate, C. F. *J. Vac. Sci. Technol., A* **1988,** *6,* 271. doi:10.1116/1.575441
6. Giessibl, F. J. *Appl. Phys. Lett.* **1998,** *73,* 3956. doi:10.1063/1.122948
7. Torbrügge, S.; Schaff, O.; Rychen, J. *J. Vac. Sci. Technol., B: Nanotechnol. Microelectron.: Mater., Process., Meas., Phenom.* **2010,** *28,* C4E12. doi:10.1116/1.3430544
8. Giessibl, F. J.; Trafas, B. M. *Rev. Sci. Instrum.* **1994,** *65,* 1923. doi:10.1063/1.1145232
9. Giessibl, F. J. *Appl. Phys. Lett.* **2000,** *76,* 1470–1472. doi:10.1063/1.126067
10. Kawai, S.; Kitamura, S.-i.; Kobayashi, D.; Kawakatsu, H. *Appl. Phys. Lett.* **2005,** *87,* 173105. doi:10.1063/1.2112203
11. Giessibl, F. J.; Herz, M.; Mannhart, J. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 12006–12010. doi:10.1073/pnas.182160599
12. Szlufarska, I.; Chandross, M.; Carpick, R. W. *J. Phys. D: Appl. Phys.* **2008,** *41,* 123001. doi:10.1088/0022-3727/41/12/123001
13. Giessibl, F. J.; Bielefeldt, H.; Hembacher, S.; Mannhart, J. *Appl. Surf. Sci.* **1999,** *140,* 352–357. doi:10.1016/S0169-4332(98)00553-4
14. Giessibl, F. J. *Phys. Rev. B* **1997,** *56,* 16010–16015. doi:10.1103/PhysRevB.56.16010
15. Giessibl, F. J.; Pielmeier, F.; Eguchi, T.; An, T.; Hasegawa, Y. *Phys. Rev. B* **2011,** *84,* 125409. doi:10.1103/PhysRevB.84.125409
16. Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G. *Science* **2009,** *325,* 1110–1114. doi:10.1126/science.1176210
17. van Vörden, D.; Lange, M.; Schmuck, M.; Schmidt, N.; Möller, R. *Beilstein J. Nanotechnol.* **2012,** *3,* 809–816. doi:10.3762/bjnano.3.90
18. Berger, J.; Švec, M.; Müller, M.; Ledinský, M.; Fejfar, A.; Jelínek, P.; Majzik, Z. *Beilstein J. Nanotechnol.* **2013,** *4,* 1–9. doi:10.3762/bjnano.4.1
19. Gere, J. M.; Goodno, B. J. *Mechanics of Materials;* Cengage Learning: Stamford, CT, USA, 2012.
20. Lozano, J. R.; Kiracofe, D.; Melcher, J.; Garcia, R.; Raman, A. *Nanotechnology* **2010,** *21,* 465502. doi:10.1088/0957-4484/21/46/465502
21. Heyde, M.; Kulawik, M.; Rust, H.-P.; Freund, H.-J. *Rev. Sci. Instrum.* **2004,** *75,* 2446–2450. doi:10.1063/1.1765753
22. NaugaNeedles. http://nauganeedles.com. Private Communication.
23. Majzik, Z.; Setvín, M.; Bettac, A.; Feltz, A.; Cháb, V.; Jelínek, P. *Beilstein J. Nanotechnol.* **2012,** *3,* 249–259. doi:10.3762/bjnano.3.28

License and Terms

The definitive version of this article is the electronic one which can be found at: doi:10.3762/bjnano.4.43

Beilstein Journal of Nanotechnology

Multiple regimes of operation in bimodal AFM: understanding the energy of cantilever eigenmodes

Daniel Kiracofe*[1,2], Arvind Raman[1] and Dalia Yablon[2]

Full Research Paper

Open Access

Address:
[1]Purdue University School of Mechanical Engineering, West Lafayette, IN, USA and [2]ExxonMobil Research and Engineering Co., Annandale, NJ, USA

Email:
Daniel Kiracofe* - daniel.kiracofe@gmail.com

* Corresponding author

Keywords:
atomic force microscopy; bimodal AFM; cantilever eigenmodes; polymer characterization

Beilstein J. Nanotechnol. **2013,** *4,* 385–393.
doi:10.3762/bjnano.4.45

Received: 01 March 2013
Accepted: 17 May 2013
Published: 21 June 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

One of the key goals in atomic force microscopy (AFM) imaging is to enhance material property contrast with high resolution. Bimodal AFM, where two eigenmodes are simultaneously excited, confers significant advantages over conventional single-frequency tapping mode AFM due to its ability to provide contrast between regions with different material properties under gentle imaging conditions. Bimodal AFM traditionally uses the first two eigenmodes of the AFM cantilever. In this work, the authors explore the use of higher eigenmodes in bimodal AFM (e.g., exciting the first and fourth eigenmodes). It is found that such operation leads to interesting contrast reversals compared to traditional bimodal AFM. A series of experiments and numerical simulations shows that the primary cause of the contrast reversals is not the choice of eigenmode itself (e.g., second versus fourth), but rather the relative kinetic energy between the higher eigenmode and the first eigenmode. This leads to the identification of three distinct imaging regimes in bimodal AFM. This result, which is applicable even to traditional bimodal AFM, should allow researchers to choose cantilever and operating parameters in a more rational manner in order to optimize resolution and contrast during nanoscale imaging of materials.

Introduction

Atomic force microscopy (AFM) has arisen as one of the key tools for characterization of morphology and surface properties of materials (e.g., polymer blends and composites) at the micro-/nanoscale [1]. Although there are many different operating modes in AFM, one of the most popular is amplitude modulation (AM-AFM), commonly known as tapping mode, in which the cantilever is oscillated at its first natural frequency. AM-AFM provides two basic images of the surface, a height (topography) image and the so-called "phase" image. The latter is related to material properties and is frequently used to distinguish different domains or different blend components from one another. While phase imaging often provides good contrast

between different materials, it is difficult to determine the exact mechanical property that is responsible for a particular contrast. Further, the contrast is sometimes poor between distinct domains or components within a polymer blend or composite. Finally, artifacts induced by bistable imaging in attractive and repulsive regimes often confuse the interpretation of phase images.

An extension of AM-AFM called bimodal AFM [2], a capability that has been applied to a variety of materials over the past five years, can overcome some of these limitations. Bimodal AFM oscillates the AFM cantilever at two frequencies simultaneously. This adds two additional channels of information beyond the standard AM-AFM method, namely, the amplitude and phase at the second frequency, which can be used to enhance contrast between materials. Further, this information can be assigned to specific types of interactions (i.e., conservative/elastic versus dissipative) [3]. With its recent widespread usage, especially among soft materials such as biological materials and polymers, bimodal AFM has demonstrated its capability to provide new contrast and information in the higher order mode [4-8].

The traditional choice in bimodal AFM is to oscillate the cantilever at its first two natural frequencies, namely the first and second flexural eigenmodes of the cantilever. However, modern AFMs have the frequency bandwidth to excite the third, fourth, or even fifth flexural eigenmode. This allows for the fundamental eigenmode to be paired with many different higher order eigenmodes for bimodal AFM operation (or even multiple higher eigenmodes simultaneously [9]). In this work, we wish to examine the choice of specific higher order eigenmodes for bimodal operation in order to understand if they provide any practical advantages in terms of material discrimination and identification. There are several questions of interest. First, do higher order eigenmodes probe the same types of tip–sample interactions as lower order eigenmodes? For example, polymers may be viscoelastic, so there could be a different interaction due to the frequency difference. Second, do higher order eigenmodes provide better contrast between materials and/or higher image quality? Third, can any rational guidance be provided for the selection of the higher order mode as well as operating parameters (e.g., drive amplitudes, setpoints, etc.) in order to obtain the most meaningful interpretation from the images?

In this work we show a series of bimodal experiments on a multicomponent polymer blend using different combinations of eigenmodes. Bimodal AFM shows excellent contrast between the different components. We will show that there are several interesting effects that depend on the choice of eigenmodes and operating parameters, which suggest that there are actually at least three distinct operating regimes in bimodal AFM (akin to attractive/repulsive regimes in AM-AFM). Numerical simulations are then used to provide further insight into these different regimes.

Experiment

Methods

An AFM (Asylum Research, Santa Barbara, CA) with a high frequency cantilever holder is used. Experiments were conducted using silicon Olympus cantilevers. Three different models were tried – AC240, AC200, and AC160 – which have nominal stiffness values of approximately 2, 9, and 26 N/m. Similar results were obtained for all cantilever models, and representative results for AC200 cantilevers are shown here.

The optical lever sensitivity (also known as "invOLS") of the first, second, and third eigenmodes was obtained from dynamic approach curves on a mica surface (in repulsive regime on a stiff surface, the amplitude decreases by approximately 1 nm when the *z*-piezo is displaced by 1 nm [10]). The sensitivity of the fourth eigenmode could not be obtained in this way because the modal stiffness was too high (relative to the tip–sample contact stiffness). Therefore the fourth eigenmode sensitivity was estimated based on Euler–Bernoulli beam theory and the data for the lower order modes. The stiffness of each eigenmode was calibrated by using the thermal tune method [11,12]. The thermal response of the fourth eigenmode was sometimes too small to give a meaningful calibration, and in this case the stiffness was estimated from beam theory. The natural frequency ω_4, however, can be measured precisely, and the ratio ω_4/ω_3 is within 7% of the value predicted by beam theory, suggesting that the stiffness should not be too far from beam theory predictions either.

Care was taken to tune the driving frequency exactly to the natural frequency before every experiment. The effects of squeeze film damping [13] are such that the phase can change by an appreciable amount (10 degrees) when the cantilever is moved a few micrometers away from the surface. Further, piezo resonances can distort the tuning curve. For plain AM-AFM at the first natural frequency, piezo resonances are generally only an issue in liquid [14]. However, on our instrument, piezo resonances can distort the higher eigenmode tuning curves significantly, especially for third and higher eigenmodes. Therefore, a thermally driven spectrum was obtained when the cantilever was positioned approximately 100 nm above the surface. A curve fit to the thermal spectum was used to determine the natural frequency [14]. The drive frequency was then set to this frequency, and the phase (lag) offset was set to 90 degrees.

The sample used was a ternary polymer blend consisting of isotactic polypropylene (PP, ExxonMobil Chemical Company), high density polyethylene (PE, ExxonMobil Chemical Company), and polystyrene (PS, Polysciences). A blend of 3:1:1 (by mass) of PP/PE/PS was prepared in a Brabender mixer (Brabender Instruments, South Hackensack, NJ) at 180 °C, 60 rpm, and 5 min of mixing and then compression molded into a bar. This was then cryo-cut using a microtome (Ultracut 6, Leica Mikrosysteme GmbH, Vienna, Austria) at −120 °C with a glass and a diamond knife prior to AFM imaging. This sample was chosen because the individual components are well characterized and can be easily distinguished in AFM images based on morphology, surface roughness, and height. Specifically, the matrix (dominant component) is polypropylene, with approximately circular polyethylene and polystyrene domains. The polyethylene domains appear rough because of the lamellar structure, and the polystyrene domains show fracture marks from the cryomicrotoming. Dynamic mechanical analysis using time-temperature superposition was performed on each component individually by using the method described in [15].

Results

The first experiment that was carried out was to compare two scans where all parameters were the same except for the choice of higher eigenmode (e.g., "1st + 2nd" eigenmodes versus "1st + 4th"). The objective was to determine if the two scans showed the same type of contrast between the components, and to determine if one scan showed better contrast or higher resolution. Typical results of bimodal imaging with various higher order modes on the ternary blend sample are shown in Figure 1. The cantilever parameters were k_1 = 4 N/m, Q_1 = 212 (remaining parameters given in Table 1). Two different scans are compared. The first scan is a bimodal image using the 1st and 2nd eigenmodes (left column (a,c,e)), and the second scan is a bimodal image using the 1st and 4th eigenmodes (right column (b,d,f)). In both cases, the free amplitude of the 1st eigenmode was 50 nm, the setpoint was 50%, and the free amplitude of the higher eigenmode (either 2nd or 4th) was 2.5 nm. There are several interesting differences between the "1st + 2nd" bimodal image and the "1st + 4th" bimodal image. The first observation involves relative contrast between the three different materials present in the blend. The contrast between the PP and PS domains are very similar in all the images. However, the contrast between the PE domain and the PP or PS domains changes. When considering the "1st + 2nd" scan (left column), the polyethylene has a higher (brighter) amplitude (a) and phase lag (c) than either the polystyrene or the polyproylene. However, in the "1st + 4th" scan (right column) the contrast is exactly reversed. The polyethylene has a lower (darker) amplitude (b) and phase (d) than either the polystyrene or the polypropylene.

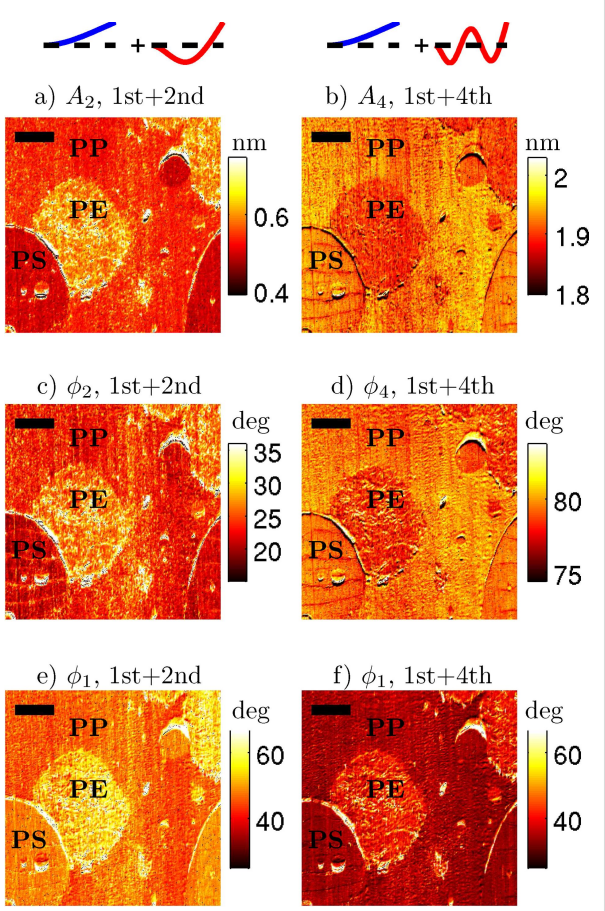

Figure 1: An 11 × 11 micrometer scan of a three component polymer blend (PS, PP, PE) imaged with an Olympus AC200 cantilever in two different modes: The left column (a,c,e) is "classical" bimodal AFM with the 1st and 2nd natural frequencies excited, and the right column (b,d,f) is bimodal AFM with the 1st and 4th natural frequencies excited. The eigenmode shapes are drawn above each column. The free amplitudes were the same in both cases (50 nm at the 1st eigenmode, and 2.5 nm at the 2nd or 4th eigenmode, respectively). Comparing the polyethylene (PE) to the other components, both the amplitude (a,b) and the phase lag (c,d) have a contrast between the modes. For the "1st + 2nd" image, PE has the highest A_2 and ϕ_2, but for the "1st + 4th" image, PE has the lowest A_4 and ϕ_4. ϕ_1 (e)(f) also shows a distinct different between the two imaging modes. The scale bar is 2 μm.

Table 1: Calibrated cantilever parameters for the experiments.

mode	1	2	3	4
stiffness (N/m)	4	78.5	366	1330[a]
quality factor	212	457	507	600[a]
natural frequency (kHz)	117	674	1758	3235

[a]Values are estimates.

A second observed difference between the bimodal images in "1st + 2nd" eigenmode versus "1st + 4th" eigenmode occurs in the first eigenmode phase (lag) channel. The overall first eigen-

mode phase (lag) has decreased considerably (darker) between the "1st + 2nd" scan (e) and the "1st + 4th" scan (f). The change in the first eigenmode phase is surprising because the bimodal AFM literature has generally treated the second frequency as an independent channel that provides additional information but does not affect the response of the first eigenmode [2,5,16,17]. The argument is that the first eigenmode is unaffected by the higher frequency oscillation because the first eigenmode amplitude is more than an order of magnitude larger than the higher eigenmode amplitude. For example, in [8] it was recently demonstrated experimentally that when $A_{1,\text{free}}/A_{2,\text{free}} = 10:1$ (or greater) there is no apparent coupling between the 1st and 2nd eigenmodes, but that when $A_{1,\text{free}}/A_{2,\text{free}} = 1:1$ there is a coupling between the eigenmodes. However, in Figure 1f, we are using a ratio of $A_{1,\text{free}}/A_{4,\text{free}} = 20:1$, but yet we see a strong change in the first eigenmode, indicating some coupling between the eigenmodes. Therefore, we conclude that a large amplitude ratio is a necessary but not sufficient condition for the two eigenmodes to be uncoupled. Later, we will attempt to determine a sufficient condition for the eigenmodes to be uncoupled.

These two features, a contrast reversal between PE and PP in the higher eigenmodes and an overall drop in the first eigenmode phase, were repeatable across multiple different cantilevers on different days, on different cantilever models with stiffness from 2 to 26 N/m, on different locations on the sample, and on different samples. Similar results to the "1st + 4th" higher order eigenmode amplitude contrast reversal and lowering of first order eigenmode phase were also observed in "1st + 3rd" bimodal imaging. We chose to focus on "1st + 4th" imaging instead of "1st + 3rd" because the 4th eigenmode was more novel experimentally.

We first discuss the PE/PP contrast reversal in the higher eigenmodes. Broadly speaking, we could imagine two possible explanations for these results. First, it could be that the tip–sample interaction probed by the fourth eigenmode is significantly different to the interaction probed by the second eigenmode (e.g., due to viscoelasticity). Alternatively, it could be that there is a difference in the cantilever dynamics at the fourth eigenmode such that it responds to the exact same tip–sample interaction in a different way. Next we show experiments designed to distinguish between these two possibilities.

Regarding the second possibility, it was pointed out in [18] that in bimodal AFM, the fundamental quantity may not be the ratio of the amplitudes of the two eigenmodes, but rather the ratio of the energy of the two eigenmodes. In our case, we compared the "1st + 2nd" scan and the "1st + 4th" scan at the exact same free amplitudes. But because $k_4 > k_2$, they were not compared at the same energy levels. Understanding the energy flow while tapping on the sample can be complicated (because of energy transfers between eigenmodes analogous to [19]). Therefore, as a first approximation, consider the kinetic energy of a freely vibrating (no tip–sample interaction) cantilever eigenmode, which is $E_i = k_i A_{i,\text{free}}^2/2$ (both kinetic and potential energy give the same result; derivation in Supporting Information File 1). For the conditions under which the data in Figure 1 were collected, $E_1/E_2 = (4(50^2))/(78.5(2.5^2)) = 20.4$. So the first eigenmode has much more energy than the second eigenmode. But when comparing the first and fourth eigenmodes we find $E_1/E_4 = (4(50^2))/(1330(2.5^2)) = 1.2$. Therefore the fourth eigenmode actually has approximately the same energy as the first eigenmode.

To evaluate which effect the eigenmode energy has on the results, we repeat the experiment except this time we consider not just the choice of eigenmode, but also adjust the drive amplitudes to control the eigenmode energy level. This leads to four bimodal scans: Figure 2a "1st + 2nd" modes with $E_2 << E_1$; Figure 2b "1st + 4th" modes with $E_4 << E_1$; Figure 2c "1st + 2nd" with $E_2 \approx E_1$; and Figure 2d "1st + 4th" modes with $E_4 \approx E_1$. The scans in (a) and (d) were taken under the same conditions as those in Figure 1. It is clear that the PE/PP contrast is similar between the "1st + 2nd" and "1st + 4th" cases as long as the energies are kept the same. That is, for

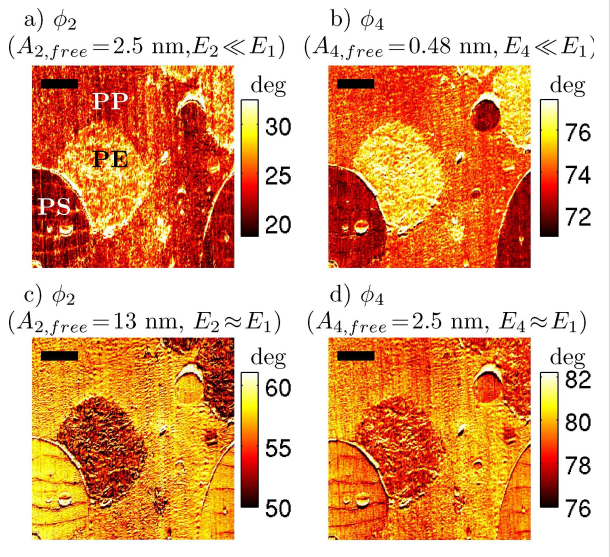

Figure 2: Using the same cantilever and sample from Figure 1, the imaging modes are compared at different energy levels. In the top row (a, b) the first eigenmode has a larger energy than the higher mode (2nd or 4th respectively). In the bottom row (c, d), the first eigenmode energy is comparable to the higher eigenmode energy. In the top row, the PE has the highest phase lag, whereas in the bottom row the PE has the lowest phase lag. This shows that the contrast reversal observed in Figure 1 is not caused by the choice of eigenmode alone, but by the energy in the eigenmode. The scale bar is 2 μm.

$E_2 << E_1$ or $E_4 << E_1$, PE is bright with respect to PP as shown in Figure 2a and Figure 2b. For $E_2 \approx E_1$ or $E_4 \approx E_1$, PE is dark with respect to PP as shown in Figure 2c and Figure 2d. This suggests that the contrast reversal is due solely to the relative eigenmode energy levels, and not to the choice of eigenmode.

Further, in Figure 3, it can be seen that the first eigenmode phase (lag) shows the same pattern: the drop in phase is due to the relative eigenmode energy levels. When the relative energy in the first eigenmode dominates (a and b), ϕ_1 is relatively high, but when the two eigenmodes have about the same energy (c and d), then ϕ_1 is relatively low. How can this change in phase be interpreted? The meaning of phase angles in AM-AFM is a difficult topic that is frequently misinterpreted. In AM-AFM imaging, it is not possible to separately identify the conservative (e.g., elastic) and dissipative (e.g., viscous) components of the tip–sample interaction by using the first eigenmode phase (in contrast to FM-AFM imaging, or higher eigenmode imaging in bimodal AFM). Instead, the first eigenmode phase gives information about the ratio of the dissipative interaction to the conservative interaction. The lower phase lag in (c) and (d) indicates a lower ratio of dissipative interaction to conservative interaction (i.e., either less dissipation, higher conservative forces, or both), as compared to (a) and (b). It appears that there are two distinct operating regimes in bimodal AFM, which have distinctly different responses to material property contrast and distinctly different energy dissipations in the first eigenmode. In the next section, numerical simulation is used to provide further insight into this second regime.

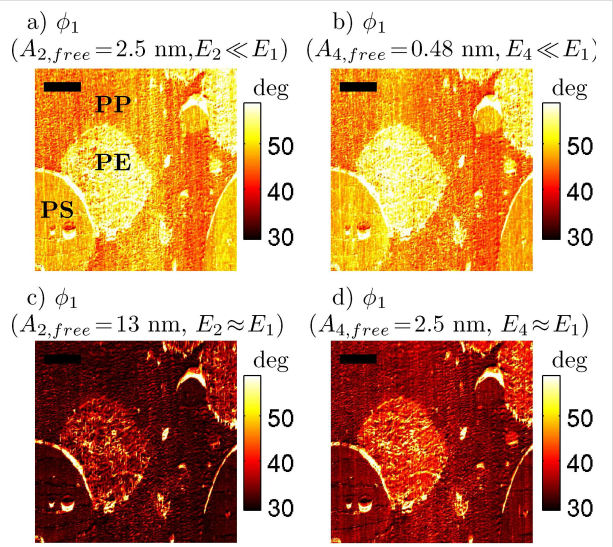

Figure 3: The same experiment as Figure 2, plotting first eigenmode phase lag ϕ_1. In the top row (a, b) the first eigenmode has a larger energy than the higher mode (2nd or 4th, respectively). In the bottom row (c, d), the first eigenmode energy is comparable to the higher eigenmode energy. The first eigenmode phase drops considerably between the top and bottom rows, indicating that in the high energy state there is a nonlinear coupling between the eigenmodes that is affecting the first eigenmode. This is clearly due to the energy ratios, and not the choice of higher eigenmode (i.e., little difference between left and right columns). The scale bar is 2 μm.

Simulation

Modeling

In order to provide insight into the physical processes at work, we use numerical simulations. The VEDA simulator (a freely available, open-source, web-based [20] AFM simulator developed by the authors) is used for numerical simulation. A full description of the simulator is given in [21,22]. Here we review the features relevant to the present work. The modeling starts with the Euler–Bernoulli partial differential equation for deflections of a slender, rectangular cantilever beam in a ground-fixed inertial frame, subject to a hydrodynamic damping force, a driving force, and a tip–sample interaction force

$$EI\frac{\partial^4 w(x,t)}{\partial x^4}+\rho_c A\frac{\partial^2 w(x,t)}{\partial t^2}=F_{\text{hydro}}(w)+F_{\text{drive}}(x,t)+F_{\text{ts}}(w(L,t)+Z)\delta(x-L) \quad (1)$$

where E, I, ρ_c, A, w, x, t, F_{hydro}, F_{ts}, F_{drive} and δ are the cantilever Young's modulus, area moment of inertia, density, cross-sectional area, deflection, axial coordinate, time, hydrodynamic force, tip–sample interaction force, driving (excitation force), and Dirac delta, respectively. The hydrodynamic forces are converted into an effective modal viscosity and added mass [23], and then the equation is discretized in the basis of cantilever eigenmodes by Galerkin's method following [24]. The method is to write w as

$$w(x,t)=\sum_{i=1}^{\infty} q_i(t)\psi_i(x)$$

where $\psi_i(x)$ is the ith eigenmode shape and $q_i(t)$ is referred to as a modal coordinate. ψ is chosen such that $\psi_i(L) = 1$ so that the modal coordinates are the deflection of the cantilever at the free end. This scaling is important because it allows the calibrated stiffnesses of the eigenmodes to be incorporated directly into the model [24]. An approximation is made by keeping only the first N eigenmodes. We take $N = 4$ in this work. This reduces the original equation to a set of four ordinary differential equations:

$$\frac{q_i}{\omega_i^2}+\frac{q_i}{\omega_i Q_i}+q_i=\frac{F_{\text{ts}}(d)}{k_i}+\frac{F_{i1}\cos\Omega_1 t}{k_i}+\frac{F_{i2}\cos\Omega_2 t}{k_i},\ i=1..4 \quad (2)$$

where $q_i(t)$, ω_i, Q_i, and k_i, are the tip deflection, natural frequency (rad/s), quality factor, and equivalent stiffness, respectively, and

$$d = Z + \sum_{i=1}^{4} q_i$$

is the tip–sample gap, where Z is the cantilever–sample separation. A bimodal excitation is used where Ω_1 is the first driving frequency, Ω_2 is the second driving frequency, F_{i1} is the force on the ith eigenmode due to the first excitation and F_{i2} is the force on the ith eigenmode due to the second excitation. In this work, we take $\Omega_1 = \omega_1$ and $\Omega_2 = \omega_2$ to simulate bimodal driving of the 1st and 2nd eigenmodes.

The tip–sample interaction force $F_{ts}(d)$ is described by a modified DMT model that includes a term for surface energy hysteresis. In other words, the force when the tip is approaching the sample is different from the force when the tip is retracting from the sample. The model is based on the one proposed by [25] and is described in detail in Supporting Information File 1.

Simulations Results

Figure 4 shows two simulations of a line scan in AM-AFM. The line scan crosses over two different materials that are located side by side. The material on the left (red lines) has a Young's modulus of 3 GPa and the one on the right (blue lines) has a modulus of 2 GPa. These values are close to the storage modulus from dynamic mechanical analysis (time–temperature superposition was used to obtain the value at 250 kHz) for polypropylene and polyethylene at 250 kHz, respectively. Both materials have a surface energy hysteresis term of 0.06 J/m^2 (chosen to approximately match the average energy dissipation in AM-AFM experiments). The same first eigenmode amplitude is used for both simulations, while two different drive amplitudes are chosen for the second eigenmode: (b) and (d) show a larger amplitude for which $E_2 \approx E_1$; (a) and (c) show a smaller amplitude for which $E_2 << E_1$. These conditions were chosen to approximately match the experiments. The full parameters for the simulation are given in Table 2.

The results in Figure 4 qualitatively match the features in the experiment. Specifically, for the smaller amplitude (1 nm, a and c), the second eigenmode phase (lag) on the softer (blue) material is higher. But for the larger amplitude (4.1 nm, b and d), the contrast between the two materials reverses and the second eigenmode phase (lag) on the softer (blue) material is lower. This contrast reversal qualitatively matches the experimentally observed contrast reversal from Figure 2. Also, the first eigenmode phase drops significantly for the larger second eigenmode amplitude, as shown in Figure 4c and Figure 4d, which is exactly the trend noted in the experiments in Figure 3.

To further explore this phenomenon, we perform a simulation in which the cantilever is tapping on a surface with the normal

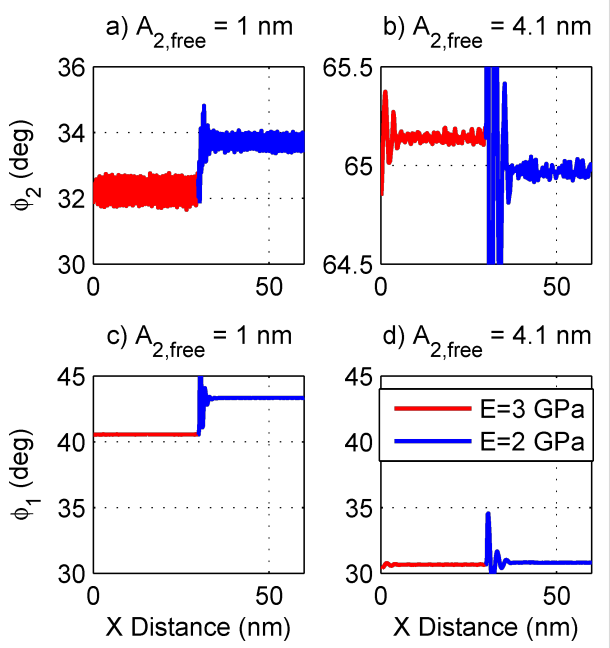

Figure 4: Simulated line scans for the parameters in Table 2 and two different second eigenmode drive amplitudes. The simulated sample is a PE domain in the center of a PP matrix. Comparing (a) versus (b), there is a clear contrast reversal of ϕ_2 between the low and high second eigenmode drive. This matches the experimental observations of Figure 2. Further, comparing (c) versus (d), ϕ_1 drops as $A_{2,init}$ is raised. This matches the experimental observation in Figure 3.

Table 2: Simulation parameters. Hamaker constant and surface energy are tuned to match the experiment. All other values are measured or nominal values.

	mode 1	mode 2
stiffness (N/m)	4	160
quality factor	200	400
natural frequency (kHz)	10	344
driving frequency (kHz)	10	344
free amplitude (nm)	40	varies
setpoint ratio	50%	
sample modulus (GPa)	2–3	
van der Waals adhesion force (nN)	1.4	
tip radius (nm)	10	
intermolecular distance (nm)	0.2	
surface energy change (J/m^2)	0.06	

feedback controller on while the second eigenmode drive amplitude (and hence second eigenmode energy) is slowly increased from zero to a maximum amplitude and then slowly decreased back to zero, as shown in Figure 5. There are two abrupt jumps in the response as the drive is changed. This plot suggests that there are not two but three distinct operating regimes in bimodal AFM, depending on the amplitudes of the eigenmodes. Further, there is a hysteresis, i.e., the jump up and jump down do not

happen at the same amplitude, indicating the presence of a bistability. Two different states are possible for the same combination of parameters (similar to the attractive/repulsive bistability in conventional AM-AFM).

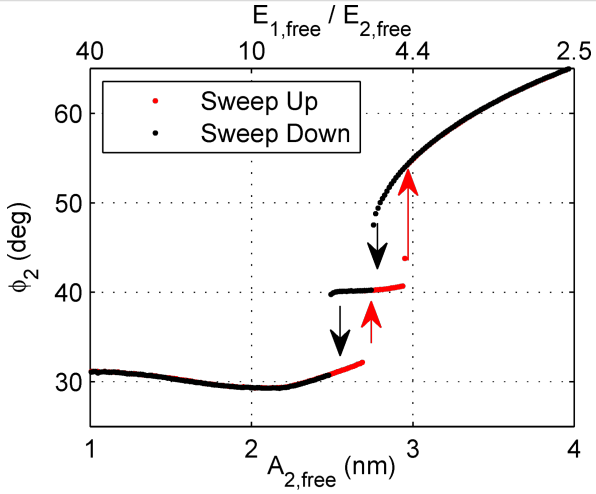

Figure 5: A simulation in which the second eigenmode drive amplitude is swept up and then down continuously. The feedback controller remains active so that a constant setpoint is maintained (i.e., A_1 is constant). The parameters are given in Table 2. The lower *x*-axis shows the second eigenmode amplitude and the upper *x*-axis the ratio of the first and second eigenmode energies (both quantities calculated based on the free case). Interestingly, there are two discrete jumps in the second eigenmode response indicating that there are three distinct dynamic states. Further, the jumps do not occur at the same amplitude on sweep up versus sweep down, indicating that the states are bistable.

A question that naturally arises is whether the amplitude and phase contrast reversals observed earlier have any correlation with the jumps between the different states. To this end, in Figure 6 amplitude sweeps are shown on two different materials (in this case, E = 2.0 GPa and E = 2.2 GPa, and the drive is swept from a high amplitude down to a low amplitude, corresponding to the range of amplitudes used in Figure 4). In fact, the contrast reversal correlates exactly to the different states. In the left state (lower $A_{2,\text{free}}$), both A_2 (Figure 6a) and ϕ_2 (Figure 6a) on the stiffer material are *lower* than on the softer material. In the right state (higher $A_{2,\text{free}}$), both A_2 and ϕ_2 on the stiffer material are *higher* than on the softer material. In the middle state, there is not much differentiation between the materials. Note that the jumps between states happen at a slightly different drive amplitude depending on the sample modulus. This is the reason that E = 2.0 GPa and E = 2.2 GPa were picked for this simulation. For E = 2.0 GPa versus E = 3.0 GPa, the jumps happen at very different amplitudes and it is more difficult to make a comparison. Because there is less contrast in modulus (as compared to the experiment), there is less contrast in phase in Figure 6 than observed in the experiment. Furthermore, Figure 6c shows a drop of ϕ_1 as the eigenmode amplitude (and hence energy) is increased as well. This is consistent with the experimental observations in Figure 3.

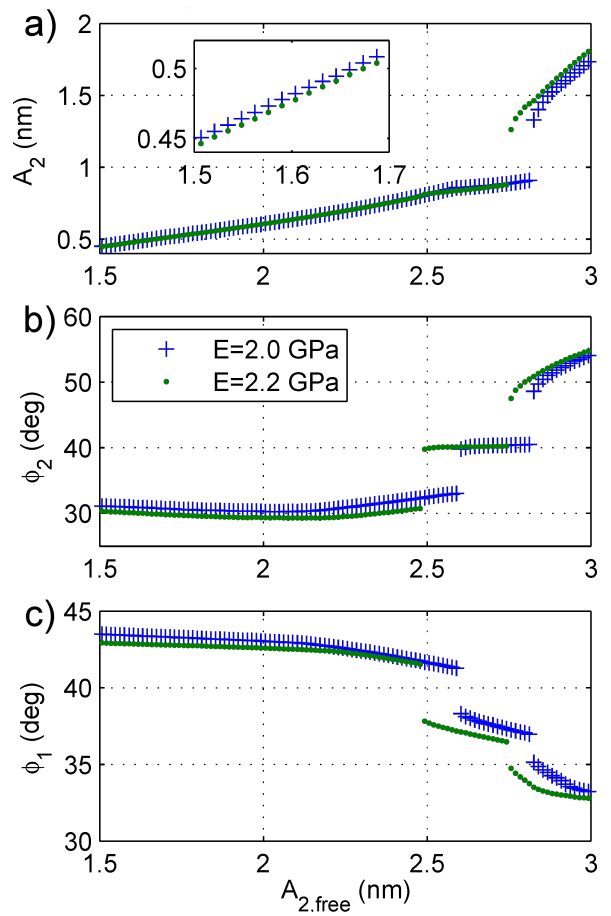

Figure 6: A drive amplitude sweep similar to Figure 5 except that two different materials are compared (both are sweeps from high amplitude down to low amplitude). The transitions between the different states happen at slightly different amplitudes for each material. It is clear from (a) and (b) that the contrast reversal observed in Figure 1, Figure 2, and Figure 4 is caused by the transition from one state to another. That is, in the state on the left (low $A_{2,\text{init}}$), A_2 and ϕ_2 on the stiffer material are higher than A_2 and ϕ_2 on the softer material. But, in the state on the right (high $A_{2,\text{init}}$), A_2 and ϕ_2 on the stiffer material are lower than A_2 and ϕ_2 on the softer material. The middle state appears to have very little contrast between the two materials. Finally, ϕ_1 drops as $A_{2,\text{init}}$ is raised, with big drops at each state transition.

The fact that the cantilever dynamics behave differently depending on the ratios of first and second eigenmode energy has been previously suggested [18]. However, that work considered only a homogenous sample, so the possibility of contrast reversal was not considered. Further, the fact that there are three states separated by a discontinuous jump was not considered. Why the contrast should reverse between the states is not immediately obvious, but it clearly happens in both experiment and simulation. To provide further insight, the simulation of Figure 6 is repeated but with a purely conservative tip–sample interaction (i.e., Hertz contact without energy dissipation). The

result is shown in Figure 7. In this case there are no discontinuous jumps. There is a point at which the slope of the amplitude and phase curves change, but there is no contrast reversal in either. Also, there is essentially no change in the first eigenmode phase (not shown).

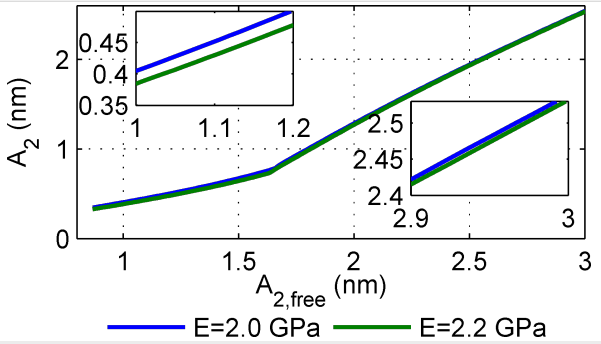

Figure 7: The simulation of Figure 6 repeated except with a pure Hertz model (no energy dissipation). In comparison to Figure 6, there are no sudden jumps and the second eigenmode contrast does not reverse.

Discussion

From a practical point of view, it is not immediately obvious if bimodal imaging using the higher states is advantageous or not. On the one hand, for certain combinations of parameters/materials, the material contrast (i.e., percent change in amplitude for a given change in Young's modulus) can be an order of magnitude higher in these states than in standard bimodal imaging. For example, in Figure 6a, the two materials are essentially indistinguishable for $A_{2,\text{free}} < 2.5$ nm, but are very clearly separated for $A_{2,\text{free}} > 2.8$ nm. This is consistent with a previous report [8] that suggested better contrast on polymers might be achieved with higher $A_{2,\text{free}}$. On the other hand, there is bistability between the different regimes. This may cause difficulties in imaging, as the attractive/repulsive regimes in conventional AM-AFM do. It may be possible to overcome the bistability by using frequency modulation, phase modulation or other newer feedback control schemes, such as drive modulation.

From a theoretical point of view, more research is needed to understand the nature of the different states and exactly why the contrast should reverse. The fact that there is no contrast reversal for the elastic case in Figure 7 suggests that the tip–sample energy dissipation plays a key role in the contrast reversal. Presumably, an energy transfer between the eigenmodes is involved.

Practically, this result reinforces the suggestion of Stark [18] that energy ratios and not amplitude ratios are the important quantity to consider in bimodal AFM. Researchers using AFM will be able to select operating conditions more intelligently if they calculate the energy ratios involved instead of amplitude ratios.

Finally, the importance of energy ratios highlights the need for better methods to calibrate stiffness and optical lever sensitivity of higher order eigenmodes. The current state of the art works well for the first few eigenmodes but becomes less reliable for third and higher modes. As multifrequency AFM evolves toward quantitative measurements using higher order eigenmodes, interferometer based AFMs, which do not suffer from these calibration problems, may become more attractive than optical lever (photodiode) based AFMs.

Conclusion

We have shown experimentally that there are multiple distinct imaging regimes in bimodal AFM. The different states were identified by contrast reversals on a multicomponent polymer blend. Higher eigenmode bimodal AFM (e.g., "1st eigenmode + 4th eigenmode") behaves essentially the same as traditional bimodal AFM ("1st + 2nd"), when operated at similar energy levels. When the energy of the higher eigenmode is much smaller than the energy of the first eigenmode, then the two eigenmodes are essentially uncoupled. This is the regime that the majority of classical bimodal AFM studies have explored. When the energy of the higher eigenmode is comparable to the energy of the first eigenmode, there are additional distinct imaging regimes involving coupling between the eigenmodes. We have shown that the experimentally observed contrast reversals can be qualitatively predicted by the use of numerical simulation. Further, the numerical simulation has shown that there are actually three distinct imaging regimes in bimodal AFM, and that the discontinuous jumps and contrast reversals are in some way caused by dissipative tip–sample interactions.

The understanding of the different imaging regimes discovered in this work will be of great help to AFM researchers by allowing them to choose their operating parameters intelligently so as to maximize material contrast.

Supporting Information

The Supporting Information contains two appendices: (1) a detailed description of the tip–sample interaction model used in the simulations and (2) a brief derivation of the kinetic/potential energy of an eigenmode.

Supporting Information File 1

Appendices
[http://www.beilstein-journals.org/bjnano/content/supplementary/2190-4286-4-45-S1.pdf]

Acknowledgements

The authors would like to thank J. Grabowski for assistance with sample preparation and R. Proksch for helpful discussions, and AR acknowledges financial support from the NSF grant CMMI 0700289 for his contribution.

References

1. Sawyer, L. C.; Grubb, D. T.; Meyers, G. F. *Polymer Microscopy,* 3rd ed.; Springer Verlag: New York, 2008.
2. Rodríguez, T. R.; García, R. *Appl. Phys. Lett.* **2004,** *84,* 449–451. doi:10.1063/1.1642273
3. Lozano, J. R.; García, R. *Phys. Rev. B* **2009,** *79,* 014110. doi:10.1103/PhysRevB.79.014110
4. Dietz, C.; Zerson, M.; Riesch, C.; Gigler, A. M.; Stark, R. W.; Rehse, N.; Magerle, R. *Appl. Phys. Lett.* **2008,** *92,* 143107. doi:10.1063/1.2907500
5. Martínez, N. F.; Lozano, J. R.; Herruzo, E. T.; García, F.; Richter, C.; Sulzbach, T.; García, R. *Nanotechnology* **2008,** *19,* 384011. doi:10.1088/0957-4484/19/38/384011
6. Li, J. W.; Cleveland, J. P.; Proksch, R. *Appl. Phys. Lett.* **2009,** *94,* 163118. doi:10.1063/1.3126521
7. Baumann, M.; Stark, R. W. *Ultramicroscopy* **2010,** *110,* 578–581. doi:10.1016/j.ultramic.2010.02.013
8. Gigler, A. M.; Dietz, C.; Baumann, M.; Martinez, N. F.; García, R.; Stark, R. W. *Beilstein J. Nanotechnol.* **2012,** *3,* 456. doi:10.3762/bjnano.3.52
9. Solares, S. D.; Chawla, G. *J. Appl. Phys.* **2010,** *108,* 054901. doi:10.1063/1.3475644
10. García, R.; Paulo, A. S. *Phys. Rev. B* **2000,** *61,* R13381–R13384. doi:10.1103/PhysRevB.61.R13381
11. Hutter, J. L.; Bechhoefer, J. *Rev. Sci. Instrum.* **1993,** *64,* 1868–1873. doi:10.1063/1.1143970
12. Butt, H.-J.; Jaschke, M. *Nanotechnology* **1995,** *6,* 1–7. doi:10.1088/0957-4484/6/1/001
13. Green, C. P.; Sader, J. E. *J. Appl. Phys.* **2005,** *98,* 114913. doi:10.1063/1.2136418
14. Kiracofe, D.; Raman, A. *Nanotechnology* **2011,** *22,* 485502. doi:10.1088/0957-4484/22/48/485502
15. Yablon, D. G.; Gannepalli, A.; Proksch, R.; Killgore, J.; Hurley, D. C.; Grabowski, J.; Tsou, A. H. *Macromolecules* **2012,** *45,* 4363–4370. doi:10.1021/ma2028038
16. Lozano, J. R.; García, R. *Phys. Rev. Lett.* **2008,** *100,* 076102. doi:10.1103/PhysRevLett.100.076102
17. Proksch, R. *Appl. Phys. Lett.* **2006,** *89,* 113121. doi:10.1063/1.2345593
18. Stark, R. W. *Appl. Phys. Lett.* **2009,** *94,* 063109. doi:10.1063/1.3080209
19. Melcher, J.; Carrasco, C.; Xu, X.; Carrascosa, J. L.; Gómez-Herrero, J.; de Pablo, P. J.; Raman, A. *Proc. Natl. Acad. Sci. U. S. A.* **2009,** *106,* 13655–13660. doi:10.1073/pnas.0902240106
20. VEDA: Virtual Environment for Dynamic AFM. http://nanohub.org/tools/veda/ (accessed Feb 28, 2013).
21. Kiracofe, D.; Melcher, J.; Raman, A. *Rev. Sci. Instrum.* **2012,** *83,* 013702. doi:10.1063/1.3669638
22. Melcher, J.; Hu, S.; Raman, A. *Rev. Sci. Instrum.* **2008,** *79,* 061301. doi:10.1063/1.2938864
23. Kiracofe, D.; Raman, A. *J. Appl. Phys.* **2010,** *107,* 033506. doi:10.1063/1.3284206
24. Melcher, J.; Hu, S.; Raman, A. *Appl. Phys. Lett.* **2007,** *91,* 053101. doi:10.1063/1.2767173
25. García, R.; Gómez, C. J.; Martinez, N. F.; Patil, S.; Dietz, C.; Magerle, R. *Phys. Rev. Lett.* **2006,** *97,* 016103. doi:10.1103/PhysRevLett.97.016103

Beilstein Journal of Nanotechnology

Ni nanocrystals on HOPG(0001): A scanning tunnelling microscope study

Michael Marz*[1,2], Keisuke Sagisaka[1] and Daisuke Fujita[1]

Full Research Paper

Open Access

Address:
[1]National Institute for Materials Science (NIMS), 1-2-1 Sengen, Tsukuba-city Ibaraki 305-0047, Japan and [2]Physikalisches Institut, Karlsruhe Institute for Technology (KIT), 76131 Karlsruhe, Germany

Email:
Michael Marz* - Michael.Marz@kit.edu

* Corresponding author

Keywords:
clusters; growth mode; Ni; nickel

Beilstein J. Nanotechnol. **2013,** *4,* 406–417.
doi:10.3762/bjnano.4.48

Received: 08 March 2013
Accepted: 02 June 2013
Published: 28 June 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Guest Editors: T. Glatzel and U. D. Schwarz

Abstract

The growth mode of small Ni clusters evaporated in UHV on HOPG has been investigated by scanning tunnelling microscopy. The size, the size distribution, and the shape of the clusters have been evaluated for different evaporation conditions and annealing temperatures. The total coverage of the surface strongly depends on the evaporation rate and time, whereas the influence of these parameters is low on the cluster size. Subsequent stepwise annealing has been performed. This results in a reduction of the total amount of the Ni clusters accompanied by a decreasing in the overall coverage of the surface. The diameter of the clusters appears to be less influenced by the annealing than is their height. Besides this, the cluster shape is strongly influenced, changing to a quasi-hexagonal geometry after the first annealing step, indicating single-crystal formation. Finally, a reproducible methodology for picking up individual clusters is reported [1].

Introduction

Metallic nanoparticles have been widely studied in the past few decades owing to their broad range of applications, such as catalysis [2-4], quantum dots [5] or chemical sensors [6]. Moreover, nano particles consisting of only some tens to a few hundred of atoms, so-called 'clusters', have unique properties since they correspond to an intermediate system between isolated atoms and bulk. This makes clusters also suitable for fundamental studies of the crossover between the characteristics of single atoms and bulk properties [7,8]. Various methods are employed for growing metallic clusters on surfaces such as ion sputtering [9], pulsed laser deposition [10], electro deposition [11-13], vapor deposition [14], aerosol deposition [15], material transfer of an STM tip [16], etc. For the formation of nanoparticles a weak interaction of the metal with the substrate

is favourable [17]. Graphite, in particular highly ordered pyrolytic graphite (HOPG), serves as an excellent substrate due to its low reactivity and large atomically flat surface. Among the transition metals the most commonly used materials for catalytic purposes are Ag, Pt, Pd, Cu, Rh and Ni [3]. Nickel clusters have recently received a lot of attention besides their catalytic properties [2-4] for the use as templates for the growth of small graphene islands [18].

The control of the size and the shape as well as the particle distribution play an important role for the aforementioned applications. Therefore, in this study we focus on the initial growing process of Ni clusters by evaporating Ni on the HOPG surface in ultrahigh vacuum (UHV). A quantitative characterization of the cluster properties, e.g., size, shape and distribution, for different experimental conditions has been performed by means of scanning tunnelling microscopy (STM) measurements. We show that Ni clusters are formed following a monomodal distribution, and that mild annealing can transform the Ni particles into single crystals. Within the annealing process two distinguishable regimes are observed. Additionally, we demonstrate the possibility to pick up single clusters from the surface in a controlled way, and we propose a model to understand the basics of the pick-up process.

Experimental

All measurements were performed in an UHV-STM system from *Unisoku (USM-1200)* at liquid nitrogen temperature. The system consists of three separated chambers, including the load lock, preparation chamber and STM (analysis) chamber. The Ni deposition and annealing of the sample was carried out in the preparation chamber. The pressure during Ni deposition was usually below $6{\cdot}10^{-8}$ Pa. The base pressure in the analysis chamber is presumed to be less than $3{\cdot}10^{-9}$ Pa at 78 K. Electrochemically etched W tips and commercially available Pt/Ir tips (Unisoku) were used. The HOPG sample was mounted on a silicon strip that serves as heater for the annealing procedure. (Different HOPG crystals with different mosaic spread were used; however, we did not observe any influence of this in our data and refer to all of them by using the term HOPG substrate.) This setup allows local heating of the sample and avoids contamination from degassing of the surrounding stainless-steel parts. In order to get a clean surface the HOPG substrate was cleaved in conventional way with adhesive tape. After loading of the substrate to the UHV chamber, it was degassed at $T \approx 500$ K for a few hours. The Ni-clusters were grown by evaporating Ni from a high purity Ni rod (5N purity) in situ by electron-beam heating. The deposition rate was controlled by monitoring the flux current. The flux ranged between 15 nA to 150 nA whereas the time was set between 30 s and 600 s. The total amount of deposited Ni is assumed to be proportional to the product of flux (f) and time (t) and the deposition coefficient ($\Pi_{f\times t}$) is defined as $\Pi_{f\times t} = f \times t$. The resulting values of $\Pi_{f\times t}$ for the aforementioned variation of flux and time range from 1.5 μAs to 36 μAs. In the following the term "experiment" refers to a single value of $\Pi_{f\times t}$ with a specific combination of flux and time.

The HOPG substrate was cooled to liquid nitrogen temperature before the Ni deposition. For the Ni deposition, the substrate had to be transferred to the preparation chamber. Inevitably, the substrate warmed up (to quantify the temperature, measurements at ambient condition were performed; assuming a slower warm-up within the UHV chamber, the maximum substrate temperature was estimated to be lower than 200 K.), but we confirm that any influence of the substrate warmup on the clusters was not detected within the deposition times used in this study, cf. section "Influence of annealing on the clusters".

A subsequent annealing process was performed in increments of approximately 100 K from 450 K up to 870 K, allowing a detailed study of the change of the clusters after each temperature step. The annealing temperature was monitored by optical pyrometers. The low annealing temperatures ($T \leq 550$ K) were measured after the whole series of measurements in a second heating cycle, in which the same heating currents were used, which should result in a similar temperatures. Corresponding to this lack of accuracy, error bars of ±50 K have been plotted in the corresponding graphs. The STM measurements had to be performed at a relatively low scanning speed to prevent uncontrolled pick-up of the clusters, resulting in roughly 30 min scanning time per image.

Analysis of the STM data was performed by using the software *Gwyddion* (http://gwyddion.net/). In order to determine the cross sections parallel to the fast raster direction, raw data were used. The cluster dimensions were determined by fitting these cross sections with a rectangular function. After fitting a large number (79–150) of clusters randomly selected from several scanning areas for each experiment, height and width histograms were plotted (not shown). The resulting histograms showed a monomodal normal distribution for the diameter and the height of the clusters. To determine the mean values for each experiment, the histograms were fitted with Gaussian curves, and details are given below. The relative coverage of the surface and the number of clusters were extracted by implementing the appropriate options in the software. After subtracting the mean plane and setting the lowest value to zero, the clusters were marked by choosing an appropriate threshold value for z above the substrate. The mean value of the height and width, and the number of clusters were calculated from these data.

Results and Discussion

Growing behaviour

Figure 1 gives an overview of the obtained Ni clusters under various evaporation conditions without any annealing; the deposition parameters are given in the figure caption. The sequence of images in Figure 1 is ordered from low to high density of Ni particles. The cluster density increases by increasing $\Pi_{f \times t}$. Figure 1a displays the freshly cleaved and degassed HOPG surface for reference. The width of the terraces was several hundreds of nanometers. The inset of Figure 1a depicts a typical image of atomically resolved HOPG. Figure 1b–Figure 1h show that the clusters form cloud-like shapes on the HOPG surface. For the nonannealed samples, the Ni-clusters appear to be randomly distributed on terraces but tend to grow preferentially along step edges (Figure 1c–Figure 1f and Figure 1h) in agreement with previous STM investigations [17,19,20]. Yang and Sacher have shown by artificially introducing defects by Ar sputtering, that defects act as nucleation centres for Ni on HOPG [20]. In our case the sticking coefficient of Ni on the basal plane for low temperatures seems to be large enough to promote the nucleation on the basal plane even without additional defects. With increasing coverage, the clusters on the terraces tend to stick together in random directions, forming conglomerates with random shapes. This clustering of the Ni clusters makes it difficult to define the number of clusters on the surface. In contrast to this, the clusters at the step edges agglomerate preferentially along the step edge, indicating a larger step–cluster interaction than cluster–cluster interaction. On the other hand, subsurface defects seem to have less influence on the clusters, as is visible in Figure 1g. This finding is consistent with the observation that the mosaic spread does not influence the cluster growth, since the mosaic spread results also from the mismatch of subsurface layers.

The overall size of the clusters is determined by the interplay of the adhesion of Ni on the clean surface and the Ni–Ni interaction, which is typical for a Vollmer–Weber growth. For a detailed statistical analysis of the results, values such as the mean cluster height and width have been calculated. Figure 2a shows the dependence of the surface coverage with Ni-clusters in percentage with respect to the total amount of evaporated Ni ($\Pi_{f \times t}$). From a linear fit, we can conclude that the HOPG surface is fully covered by clusters when the deposition of Ni reaches $\Pi_{f \times t} = 36 \pm 2$ µAs. For the calculation, the values corresponding to a very low total amount of Ni, i.e., $\Pi_{f \times t} \leq 3$ µAs (dashed blue line), have been excluded, and we will comment on this later.

Figure 2b shows that an almost constant height of the cluster, independent of any evaporation condition, is obtained. The calculated mean value of the height is $h_{\text{mean}} = 1.39 \pm 0.06$ nm (red solid line), which corresponds to at least four times the

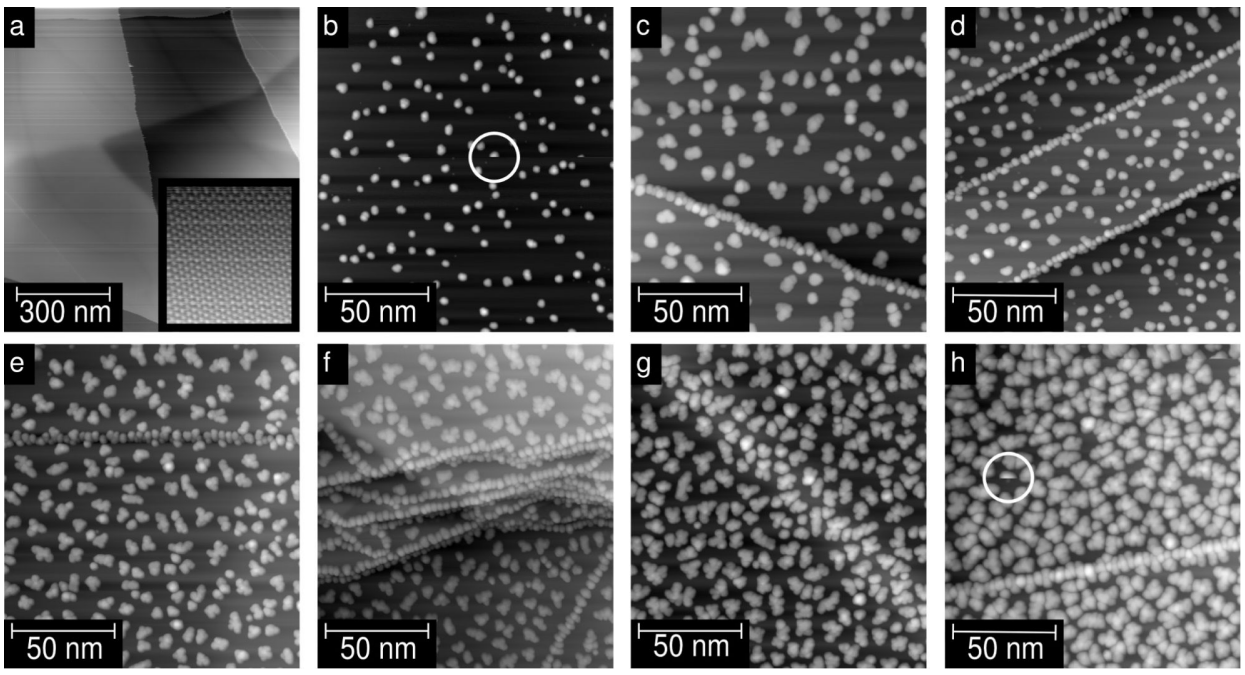

Figure 1: Evaporation time and flux dependence of the growth of the Ni-clusters without further annealing. STM topographic images of (a) freshly cleaved HOPG. The image shows a large-scale scan (V = 2 V; I = 0.2 nA), the inset shows atomic resolution (V = 0.3 V; I = 0.5 nA). (b–h) Evaporated Ni clusters on HOPG (V = 1 V; I = 0.2 nA). Evaporation parameters: (b) flux ≈50 nA, time = 30 s → $\Pi_{f \times t}$ = 1.5 µAs; (c) ≈15 nA, 300 s, 4.5 µAs; (d) ≈100 nA, 50 s, 5.0 µAs; (e) ≈50 nA, 180 s, 9.0 µAs; (f) ≈100 nA, 100 s, 10.0 µAs; (g) ≈145 nA, 100 s, 14.5 µAs; (h) ≈150 nA, 160 s, 24.0 µAs.

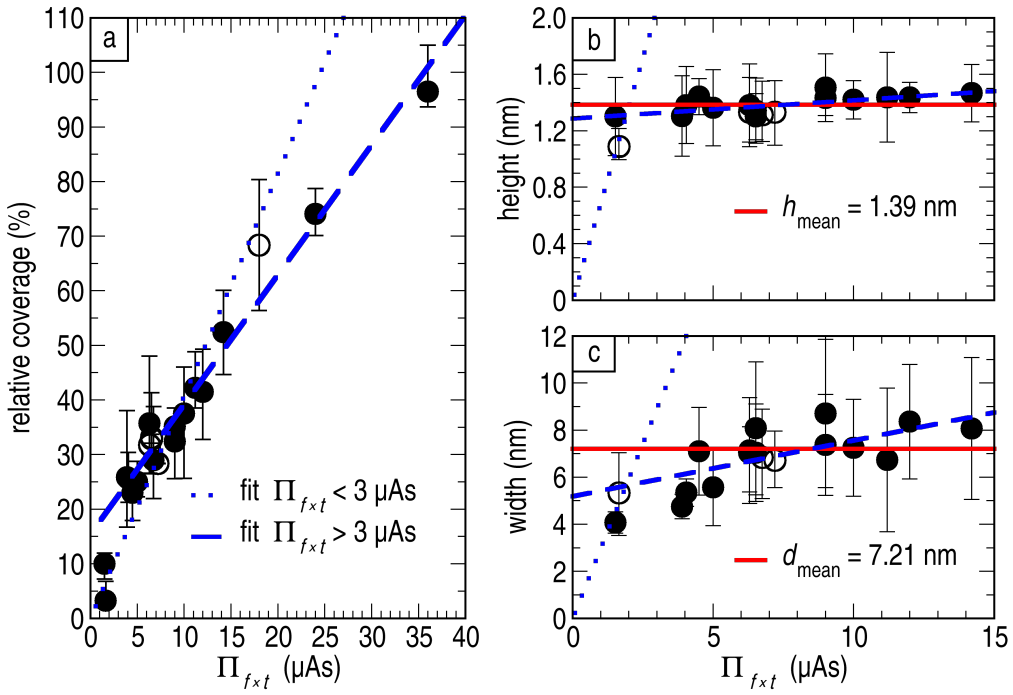

Figure 2: Analysis of the growing behaviour: (a) coverage of the surface, (b) cluster height and (c) cluster width versus $\Pi_{f\times t}$. Panel (b) and (c) have the same scale on the *x*-axis. Plotted as the filled symbols are the results obtained by fitting cross sections, as described in the text. The error bars indicate the ±2σ confidence level. Additional data points are plotted with empty symbols in the graphs to point out the lack of statistics (less than 25 cross sections), accordingly the arithmetic mean value was calculated. The error bars indicate the largest and smallest value obtained from the fit of the individual clusters. For high coverage ($\Pi_{f\times t} \geq 15$ µAs) the cluster–cluster distance was too small for getting reliable values. Thus, we do not have data for the height and the width for these experiments.

bulk lattice constant of Ni. However, a linear growth cannot be excluded from our data. Using a linear fit for the data (dashed blue line) results in a line with a very small slope. The calculated maximum height for one monolayer coverage with clusters in this case is expected to be h = 1.70 nm, i.e., five times the bulk lattice constant of Ni. Using this result, we can estimate the deposition rate as a function of the flux to be 1 µA, which corresponds to 0.038 nm/s. A similar behavior is observed in Figure 2c for the width of the clusters. The mean value of the width obtained from these data is d_{mean} = 7.21 ± 1.30 nm (red solid line). If we perform a linear fit (dashed blue line), this also results in a line with small slope. For higher coverage, the clusters tend to agglomerate making the determination of individual clusters hard, and therefore, the determination of a meaningful width for a single cluster is difficult.

We now comment on the relatively large deviation from the mean value for small $\Pi_{f\times t} \leq 3$ µAs in the coverage and the height. We presume that in the very initial state of the nanoparticle growth, the clusters cannot reach their maximum height and width due to an insufficient amount of Ni. Within this assumption, a linear function was drawn in all panels of Figure 2 (dotted blue line) to reveal a different growth behaviour with a fast increase at the very early stage. This matches well with the expected Volmer–Weber growth as initial growth state for Ni on HOPG. For the deviation of the data for low $\Pi_{f\times t}$ from the linear behaviour of the coverage the same argument is valid. The clusters grow first to their maximum height and then new clusters are formed on the substrate. Therefore, the coverage will be significantly lower for very low $\Pi_{f\times t}$ as no new clusters are formed. Another possibility for the deviations in Figure 2 could be a significantly lower sticking coefficient when only a small amount of Ni is deposited. In this case our assumption of the proportionality of $\Pi_{f\times t}$ to the total amount of deposited Ni would be no longer valid.

We have demonstrated that the size of the clusters does not depend on the evaporation condition. In order to support our claim, we plot deposition time (black circles) and rate (blue diamonds) versus height and width in Figure 3a and Figure 3b, respectively. The graphs also include the aforementioned values of the mean height and the mean width (red solid lines). Within these plots no significant dependency of the height and width on the time or the flux is found. Moreover, we do not observe any dependency in height and width by adequately varying flux and deposition time to obtain similar $\Pi_{f\times t}$ (not shown). Out of these

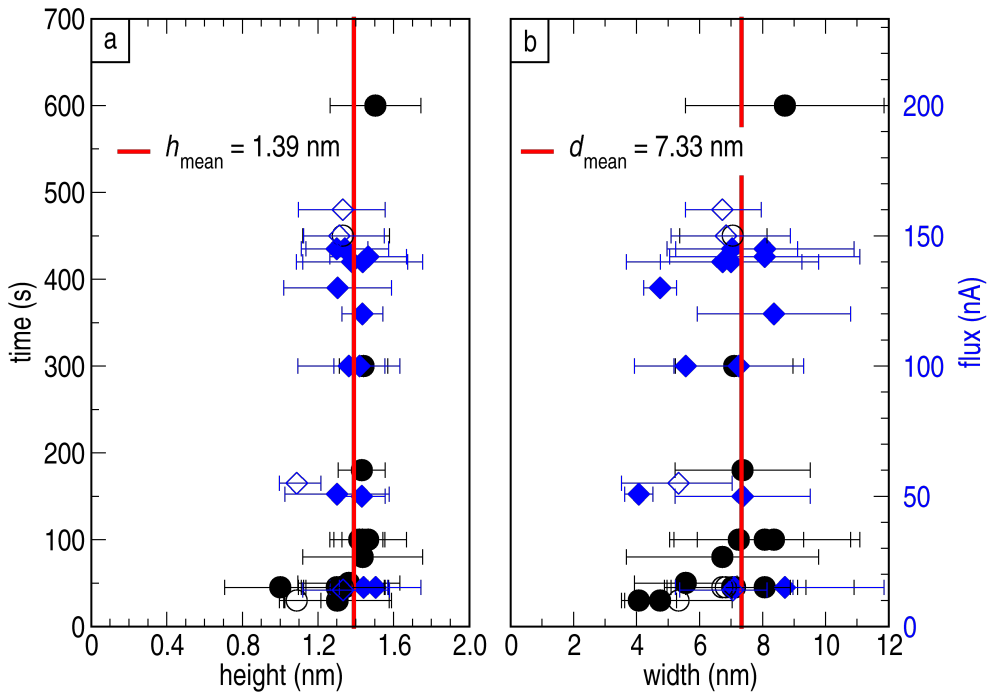

Figure 3: Analysis of the height (a) and width (b) dependence on deposition time (black circles) and flux (blue diamonds) during the evaporation process. The plots show no significant influence of the evaporation conditions on the size of the clusters.

plots, we draw two important conclusions. First, the cluster height and width can be considered independent of the evaporation conditions when the coverage of clusters is below one monolayer. Second, resulting from the first observation, the warm-up of the substrate during the deposition is negligible. The cluster diameter in the presented work is constant and roughly one order of magnitude smaller than for nickel deposition at room temperature, as done for example in the work of Bastl et al. [19]. Therefore we conclude that the temperature for the whole deposition process was significantly lower than room temperature.

Influence of annealing on the clusters

Figure 4a was recorded directly after depositing Ni on the cold surface without annealing. As previously discussed, the clusters appear randomly distributed, and bear cloud-like shapes. Thus, we assume the clusters not to exhibit single crystallinity. Figure 4b–Figure 4g show STM images of the surface after a stepwise annealing process from 450 K to 870 K. Figure 4b shows the sample after annealing for one hour with constant current and a maximum temperature of T_{max} = 450 K. It is noticeable that the cluster shape has changed drastically. Now the Ni-clusters exhibit a quasi-hexagonal structure, suggesting the formation of single crystals due to the mild annealing. For higher annealing temperatures the hexagonal structure remains visible, see Figure 4c–Figure 4f. However, for the highest used annealing temperature ($T \approx 873$ K) the hexagonal shape can no longer be resolved. In this case the height of the clusters is almost doubled compared to the one produced by the annealing at $T \approx 773$ K. Note that for such a height of the clusters the tip geometry will play an important role and is most likely the cause of the change in the observed width and shape. Also jumps of one or more clusters to the tip may have changed its resolution. Nevertheless, the change in lateral resolution has no impact on the determination of the height of the cluster and for counting the number of clusters on the surface. We notice that the height is more affected than the width of the cluster by the annealing process.

To discuss the effect of annealing on the number of cluster and their sizes, a sample with a medium coverage deposited at high flux has been used. At medium coverage, there is a large number of clusters, but it is still possible to measure their height, width and number. Figure 5a–Figure 5c shows the temperature dependence of the cluster height, the relative coverage of the surface and the number of clusters per area, respectively.

To provide a more physical insight into these results, we will make the following considerations. After annealing, the sample is transferred back to the STM chamber where it is cooled down for performing the measurements. The diffusion is allowed only

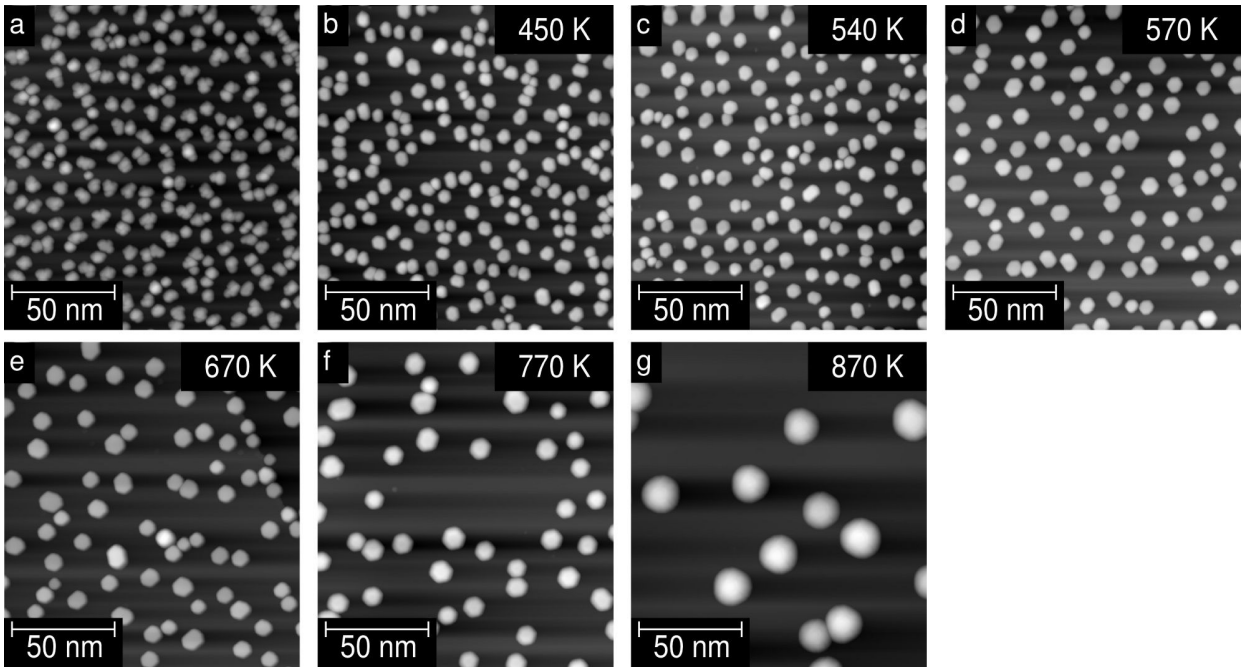

Figure 4: Temperature dependence of the annealing behaviour of the Ni clusters. STM topographic images of the HOPG sample (a) after deposition of Ni on the cold sample and (b–g) after the different annealing steps. Annealing was performed stepwise, starting at $T \approx 450$ K (b) up to $T \approx 870$ K (g). All data were obtained at $V = 1.0$ V with a tunnelling current of $I = 0.2$ nA in consistency with Figure 1.

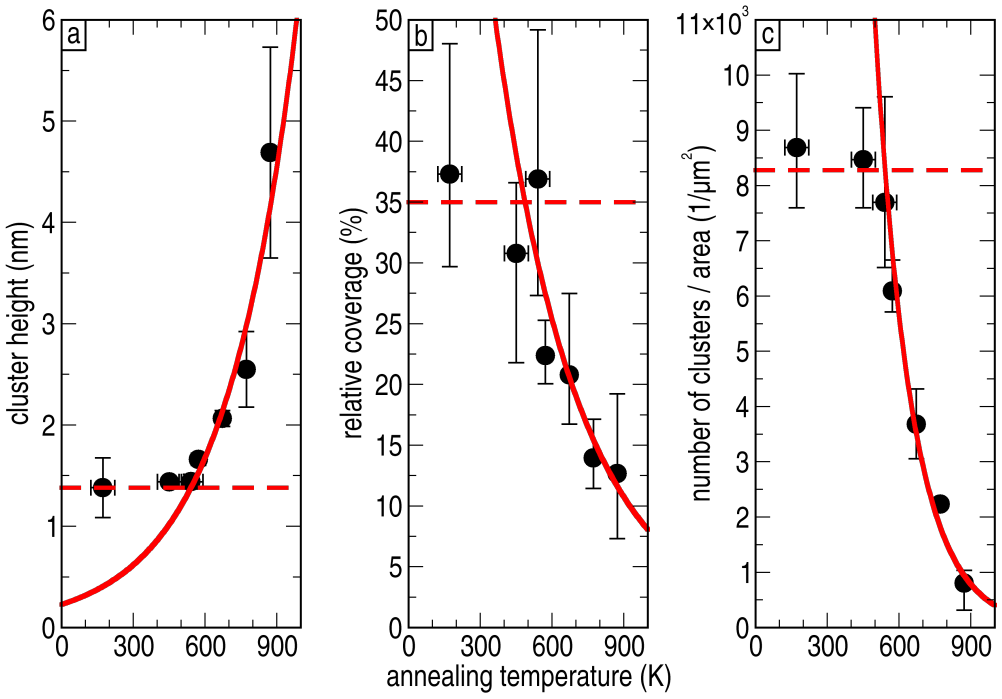

Figure 5: Temperature dependence of the annealing behaviour for (a) the height of the clusters, (b) the percentage of surface covered by the clusters, and (c) the number of clusters per area (cluster density n). The annealing process can be divided into two parts: First, for mild annealing ($T \leq 500$ K) where no diffusion takes place (indicated by the red dashed line), and second, for higher annealing temperatures with an exponential increase in height, and an exponential decay in coverage and number of clusters per area (solid red line).

at elevated temperatures during the annealing process. Therefore, the STM measurements represent a snapshot of the diffusion process that took place at the last annealing temperature. Following this assumption the diffusion energy E_d can be determined. Starting from the density of clusters n, the mean distance between two clusters L can be calculated, assuming an equidistant separation between them, according to

$$L \approx \frac{1}{\sqrt{n}}. \quad (1)$$

We consider this length L as the mean diffusion length of the clusters. The diffusion energy E_d can then be determined by an Arrhenius plot, as depicted in Figure 6, following [21], being

$$D = \alpha\left(\frac{L^2}{\tau}\right) = D_0 \cdot \exp\left[-\frac{E_d}{k_B T}\right] \quad (2)$$

$$\ln[L] = \frac{1}{2}\ln(4D_0\tau) - \frac{E_d}{2k_B}\cdot\frac{1}{T} \quad (3)$$

where τ indicates the time between two jumping events, and α is 1/4 assuming a two-dimensional diffusion. Figure 6 can be divided into two different regions that can be interpreted as two distinct diffusive regimes. First, in the low temperature regime, the data points lay on a constant curve (dashed line in Figure 6) indicating that any diffusion hardly takes place. Second, in the high annealing temperature regime, we fit a linear regression to the data (solid line in Figure 6) which results in a diffusion energy of E_d = 0.274 eV ($E_d = 4.40\cdot10^{-20}$ J). This energy is in good agreement with diffusion energies for Au on HOPG [22] ($E_d \leq 0.3$ eV) and Ni in graphite [23] (E_d = 0.807 eV), noting that the diffusion barrier at the surface is smaller than in the bulk.

These two regimes are also evident in the dependency of the number of clusters, cluster size and total coverage on the annealing temperature. The three graphs of Figure 5 show that in the low-temperature annealing range their value is almost constant, and that only the crystallinity changed due to the change in atomic arrangement within the clusters. At higher annealing temperatures, however, the number of clusters is reduced, whereas their height is increased significantly. This agrees well with the idea of Ni-atom or rather Ni-cluster diffusion. According to this interpretation, two fits are presented in the panels (a)–(c) of Figure 5. The dashed lines in all panels represent the regime where no Ni diffusion is observed. In the case of the height, panel (a), the regime of no diffusion equals the mean value obtained from the nonannealed samples. The solid lines are obtained by fitting an exponential increase for the height and an exponential decay for the coverage and number of

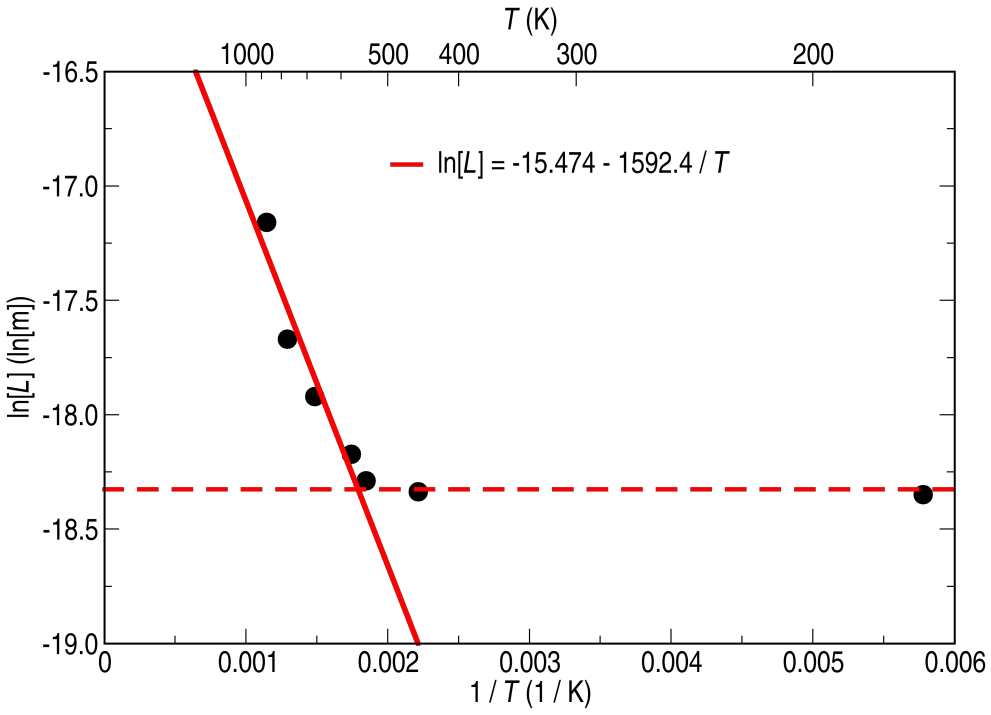

Figure 6: Arrhenius plot for analyzing the diffusion process of the Ni on HOPG. Two regimes are distinguishable in the graph and are emphasized with two linear fits to the data.

clusters to the data. The exponential decay of the coverage results directly from the exponential increase of the distance between the clusters, as shown by the fit in Figure 6, when the diameter of the clusters stays constant. The same argument is valid for interpreting the exponential increase in the cluster height, if we assume that neither evaporation from the surface nor diffusion into the substrate of the Ni clusters occurs during heating of the sample.

We have shown that annealing of the cluster can be separated into two regimes. First, single-crystalline clusters are formed. Second, with further annealing the single-crystalline clusters mainly grow in height whereas their width seems to be less affected. These two regions are also distinguishable within a simple analysis of the diffusion process.

Concluding the discussion about the growth mode and annealing behaviour of the Ni-clusters, we want to comment on the possibility of carbon interdiffusion into the clusters. Sinharoy and Levenson showed the formation and decomposition of Ni_3C in Ni-films deposited on HOPG [24]. Additionally, it is also known from Monte Carlo simulations by Diarra et al. that the solubility of carbon in Ni is significantly increased for small clusters [25]. Although we cannot exclude completely the diffusion of carbon into the clusters and the formation of nickel carbide, we have reasons to believe that it is negligible in our experiments. First, the deposition was done at low temperatures at which we expect the carbon mobility to be much lower than at room temperature. Second, the solubility of carbon in clusters of approximately 1000 atoms is already comparable to the solubility in the bulk, which is only 5–6% [25]. Finally, Ni_3C decomposes at relatively low temperatures, i.e., ≈670 K [24,26], which is below some of our annealing temperatures. However, we do not observe any significant change in the clusters, although our annealing experiments made a crossover from below to above this temperature.

Pick-up of individual clusters

The Ni-clusters are only loosely bound to the surface, and thus, it is possible to pick up individual clusters with the STM tip in a controlled way. One demonstration of the pick-up is given in Figure 7 with topographic STM images before (a) and after (b) the picking up of the cluster.

The basic idea of the process is based on gradually reducing the cluster–tip distance, until the attractive force between tip and cluster overcomes the adhesive force of the Ni nanocrystal on the HOPG surface. For this purpose, the tip was first stabilized above the center of the targeted cluster, and then the current setpoint was increased to a value up to 10 nA while the z-feedback was kept enabled. Since the tunnelling current depends, in a first-order approximation, exponentially on the distance, an increase of one order of magnitude in the current results only in about one ångström decrease in the cluster–tip distance. In order to further decrease the distance, the bias (a 1/10 divider was installed to enhance the resolution) was gradually decreased with closed *z*-feedback loop. Since we used small bias values, we can assume that the distance depends linearly on the voltage. As soon as a jump in the current or *z*-feedback signal was observed, the bias was gradually increased again. Afterwards, the usual tunnelling conditions were readjusted (V = 1.0 V, I = 0.2 nA), and the outcome of the pick-up attempt was checked with a topographic scan. In roughly 50% of the attempts the pick-up was successful. We point out that an instability in the *z*-feedback and current signal is always observed in a successful pick-up attempt. The current, bias and *z*-feedback

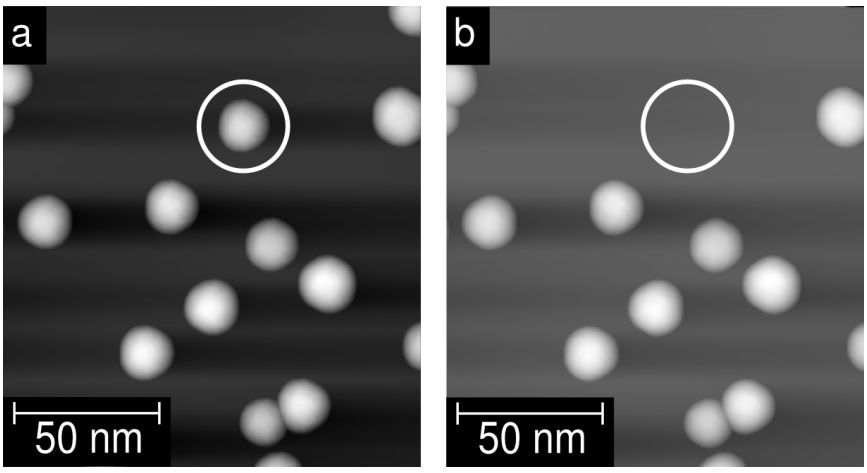

Figure 7: STM images of the HOPG surface (a) before, and (b) after the picking up of an individual cluster. The area of interest is marked with a white circle.

signals measured for a successful pick-up are given in Figure 8a. In this case, the current was set to $I = 4$ nA and the instability occurred at $V = 3.2$ mV.

Figure 8a shows that immediately after the pick-up the system becomes unstable for a relative long time. We ascribe this behaviour to instabilities at the tip, for example through the rotation or break-up of the cluster. After increasing the bias and resetting the current setpoint to the original value, the feedback-loop was usually stable again, but the resolution of the tip was often worse. Furthermore, the probability for an unintended pick-up of a second or more clusters was enhanced during topo-

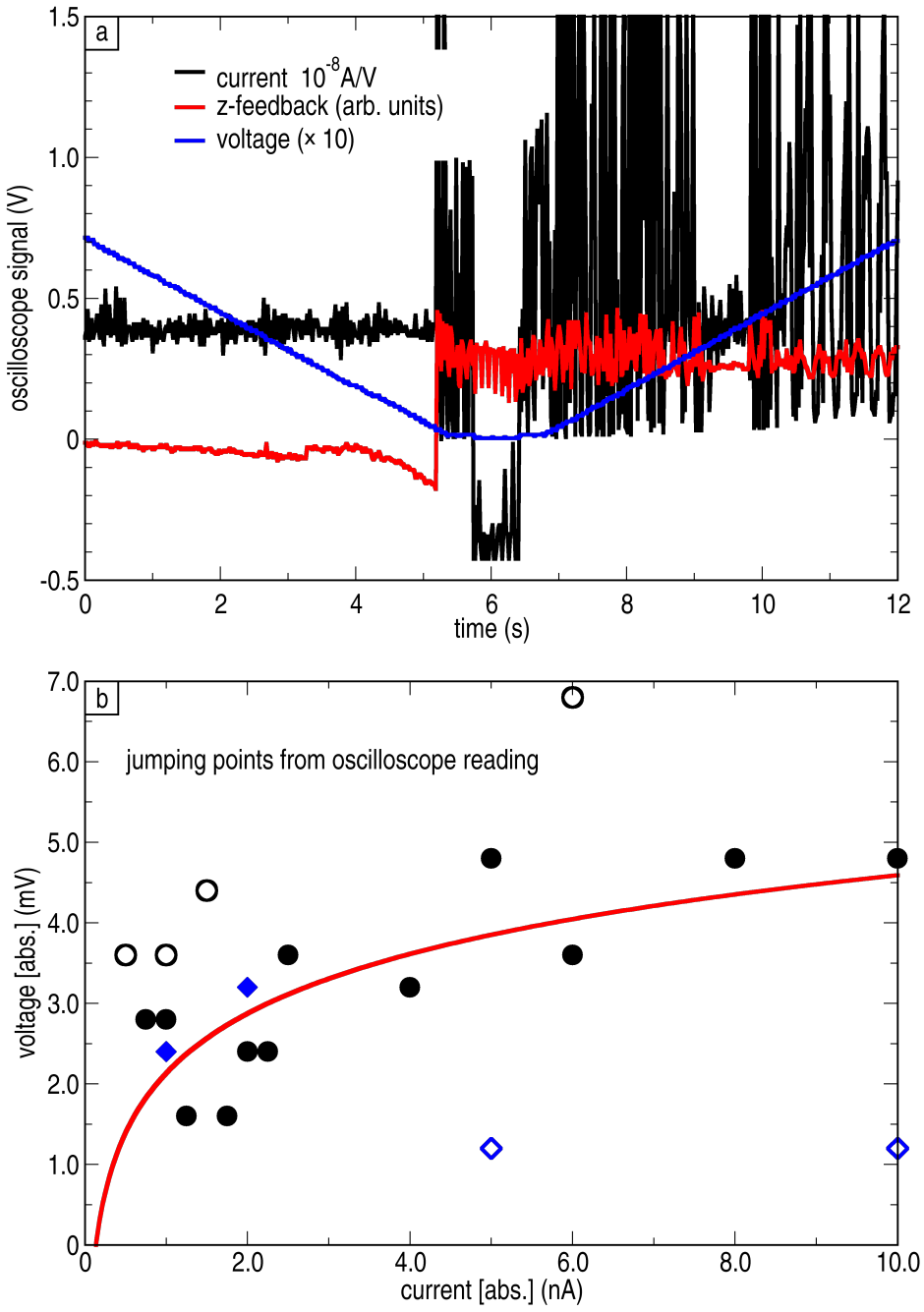

Figure 8: (a) Typical records of the current, bias and *z*-feedback signals for a successful pick-up. The jumping voltage is read out at the point where the first instability in the current or *z*-feedback signal is observed. (b) Pick-up voltage for different current setpoints. Bias ramp was started at +1 V (circles) or −1 V (diamonds). The data of the filled symbols were used for fitting the exponential curve (solid line).

graphic scans after a successful pick-up. Figure 8b reveals the correlation between the current set-point and the voltage where the pick-up happened for about twenty successful attempts.

Despite the large scattering of the data, an exponential dependence of the jumping voltage on the current set-point can be observed and the data was fitted by an exponential behaviour (solid line). For fit the data points with large deviation (open symbols) have been neglected. The result reflects the common exponential dependence of the current set-point on the tip–sample distance, when we assume that the distance decreases proportionally with the applied bias. This observation supports that mainly the tip–cluster distance is important for the pick-up process and not the applied bias, i.e., the applied electric field.

We will now take a closer look at the forces and interactions involved in the process. As mentioned before, the adhesive energy has to be overcome to remove the cluster from the surface. We can estimate the necessary force for this process in the DMT limit (Derjaguin–Muller–Toporov theory [27]), i.e., in the limit of small deformations. Thus, the adhesive force can be written as

$$F_{\mathrm{a}} = -4\pi\gamma^{*}R \tag{4}$$

$$\gamma^{*} = \gamma_{\mathrm{HOPG}} + \gamma_{\mathrm{Ni}} - 2\sqrt{\gamma_{\mathrm{HOPG}} \cdot \gamma_{\mathrm{Ni}}} \tag{5}$$

where γ_{x} is the surface tension and R the radius of the cluster, assuming a spherical particle. With γ_{HOPG} = 1.75 J/m^2, γ_{Ni} = 2.45 J/m^2, and R = 5.9 nm, a pull off force of F_{a} = −4.36 nN is expected. The attractive force between tip and cluster can be divided into three different contributions: Dielectrophoretic force, image force and van der Waals force [15,28].

$$F_{\mathrm{dep}}(z) = p\nabla E = 4\pi\varepsilon_0\varepsilon_{\mathrm{r}}R^3E\nabla E \tag{6}$$

$$F_{\mathrm{i}}(z) = -\pi\varepsilon_0\varepsilon_{\mathrm{r}}R^2E^2 \tag{7}$$

$$F_{\mathrm{vdW}}(z) = -\frac{HR}{6z^2} \tag{8}$$

where the electric field $E = E(z)$ is a function of the tip particle distance z. Since our experiments were performed in UHV, and no water interlayer is present, the Hamaker constant can be considered to be $H = 4\cdot10^{-19}$ J, and ε_{r} = 1 is valid. For simplicity, the electrical field is considered not to be influenced by the presence of the clusters and its value to be constant over the whole width of the cluster, leading to the convenient form $E(z) = V/z$. Furthermore, a good conductivity between cluster

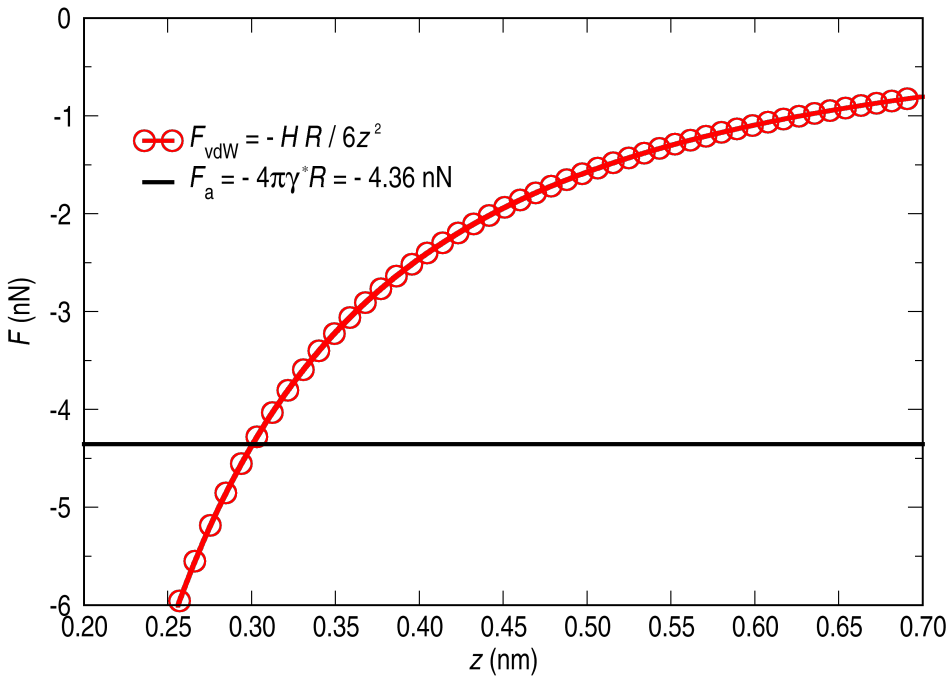

Figure 9: Forces involved in the process of picking up a Ni cluster. The tip–sample distance at which the attractive force overcomes the adhesive one is estimated to be 0.3 nm.

and substrate is assumed leading to the disappearance of the dipole moment p and consequently the dielectrophoretic force. Taking these assumptions into account, the total attractive force can be written as:

$$F_{\text{tot}}(z) = F_{\text{i}} + F_{\text{vdW}} = -\frac{6\pi\varepsilon_0 V^2 R^2 + HR}{6z^2}. \quad (9)$$

Noticing that in our experiments the bias at which the pick-up takes place is quite small, we can conclude that the whole process is dominated by the van der Waals interaction:

$$F_{\text{tot}}(z) \approx F_{\text{vdW}} = -\frac{HR}{6z^2}. \quad (10)$$

In Figure 9 the dominant forces involved in the pick-up process are plotted: the adhesive force with a constant value of $F_{\text{a}} = -4.36$ nN (solid black line) and the distance-dependent van der Waals force (open red circles). This plot provides a graphical determination of the tip–sample distance at which, following our model, the pick-up occurs. The obtained tip–cluster distance is $z = 0.30$ nm. This value is in a reasonable range taking into account the applied tunnelling conditions. Besides the controlled pick-up of the cluster, jumps of clusters to the tip occur frequently during the scan.

One example is marked with a white circle in Figure 1b and Figure 1h, respectively. The jumping rate is more enhanced when a higher tunnelling current or a faster scanning speed is used during a scan. In the case that a higher tunnelling current is used, a smaller tip–cluster distance is the reason for the rise of the jumping rate. In the case that a faster scanning speed is used, the z-feedback time constant is not fast enough to react to the steep increase of the height at the cluster edges resulting in a small tip–cluster distance. So far we have not succeeded in reversing the process, i.e., once a cluster is picked up it remains bound strongly to the tip.

Conclusion

In this study, the growth mode of small nickel metal clusters on HOPG has been investigated by scanning tunnelling microscopy. Statistical analysis of the STM images indicates that the total number of Ni clusters and the relative coverage depend linearly on the total amount of deposited Ni. The height and width of the cluster are independent of the deposition conditions, in particular of the substrate temperature (for temperatures estimated to be between 100 K and 200 K), deposition time and rate. After the annealing process, the coverage of the surface is reduced with increasing annealing temperature. Mild annealing slightly above room temperature already results in a change of the cluster shape, from a cloud-like to a quasi-hexagonal structure, indicating single-crystal formation. The lateral size of the clusters seems to be hardly affected by the annealing. In contrast to this, the cluster height continuously increases with annealing temperature. After deeply analysing the data we found that two diffusion regimes can be distinguished during the annealing process. For mild annealing temperatures ($T \leq 500$ K) only reorganization of the atoms within the clusters takes place, whereas for higher annealing temperatures ($T \geq 500$ K) Ni-atoms and even whole Ni-clusters diffuse on the HOPG surface. Finally, controlled picking-up of individual Ni clusters with the STM tip has been described.

Acknowledgements

M. Marz wishes to give thanks for support from the JSPS and AvH foundations. This work was supported by the Grant-in-Aid of JSPS(22-00740).

References

1. Parts of this work have been presented as oral or poster presentations in several national and international conferences (Meeting of the Physical Society of Japan 2011 and 2012, 6th International Symposium on Surface Science 2012, and ImagineNano 2013).
2. Aiken, J. D., III; Finke, R. G. *J. Mol. Catal. A: Chem.* **1999,** *145,* 1. doi:10.1016/S1381-1169(99)00098-9
3. Moshfegh, A. Z. *J. Phys. D: Appl. Phys.* **2009,** *42,* 233001. doi:10.1088/0022-3727/42/23/233001
4. Lebedeva, I. V.; Knizhnik, A. A.; Popov, A. M.; Potapkin, B. V. *J. Phys. Chem. C* **2012,** *116,* 6572. doi:10.1021/jp212165g
5. Schmid, G.; Bäumle, M.; Geerkens, M.; Heim, I.; Osemann, C.; Sawitowski, T. *Chem. Soc. Rev.* **1999,** *28,* 179. doi:10.1039/a801153b
6. Dupont, J.; Scholten, J. D. *Chem. Soc. Rev.* **2010,** *39,* 1780. doi:10.1039/b822551f
7. Ohgi, T.; Sheng, H.-Y.; Dong, Z.-C.; Nejoh, H.; Fujita, D. *Appl. Phys. Lett.* **2001,** *79,* 2453. doi:10.1063/1.1409585
8. Ohgi, T.; Fujita, D. *Phys. Rev. B* **2002,** *66,* 115410. doi:10.1103/PhysRevB.66.115410
9. Marcus, P.; Hinnen, C. *Surf. Sci.* **1997,** *392,* 134. doi:10.1016/S0039-6028(97)00537-2
10. Pushkin, M. A.; Lebid'ko, V. V.; Borman, V. D.; Tronin, V. N.; Troyan, V. I.; Smurov, I. *Appl. Surf. Sci.* **2006,** *252,* 4425. doi:10.1016/j.apsusc.2005.07.104
11. De Riccardis, M. F.; Carbone, D.; Martina, V.; Re, M.; Bozzini, B.; D'Urzo, L. *Appl. Surf. Sci.* **2009,** *255,* 4309. doi:10.1016/j.apsusc.2008.10.127
12. Martín, H.; Carro, P.; Hernández Creus, A.; González, S.; Salvarezza, R. C.; Arvia, A. J. *Langmuir* **1997,** *13,* 100. doi:10.1021/la960700a
13. Chen, Z.; Li, J.; Wang, E. *J. Electroanal. Chem.* **1994,** *373,* 83. doi:10.1016/0022-0728(94)03318-8
14. Anton, R.; Kreutzer, P. *Phys. Rev. B* **2000,** *61,* 16077. doi:10.1103/PhysRevB.61.16077
15. Grobelny, J.; Tsai, D.-H.; Kim, D.-I.; Pradeep, N.; Cook, R. F.; Zachariah, M. R. *Nanotechnology* **2006,** *17,* 5519. doi:10.1088/0957-4484/17/21/037

16. Fujita, D.; Kumakura, T. *Appl. Phys. Lett.* **2003,** *82,* 2329. doi:10.1063/1.1565509
17. Bäumer, M.; Libuda, J.; Freund, H.-J. *Surf. Sci.* **1995,** *327,* 321. doi:10.1016/0039-6028(94)00855-8
18. Murata, Y.; Petrova, V.; Kappes, B. B.; Ebnonnasir, A.; Petrov, I.; Xie, Y.-H.; Ciobanu, C. V.; Kodambaka, S. *ACS Nano* **2010,** *4,* 6509–6514. doi:10.1021/nn102446y
19. Bastl, Z.; Franc, J.; Janda, P.; Pelouchová, H.; Samec, Z. *Nanotechnology* **2006,** *17,* 1492. doi:10.1088/0957-4484/17/5/055
20. Yang, D.-Q.; Sacher, E. *J. Phys. Chem. B* **2005,** *109,* 19329. doi:10.1021/jp0536504
21. Henzler, M.; Göpel, W. *Oberflächenphysik des Festkörpers;* Teubner-Studienbücher: Physik: Teubner: Stuttgart, 1994. doi:10.1007/978-3-322-84875-8
22. Anton, R.; Schneidereit, I. *Phys. Rev. B* **1998,** *58,* 13874. doi:10.1103/PhysRevB.58.13874
23. Weisweiler, W.; Nageshwar, G. D. *Carbon* **1975,** *13,* 175. doi:10.1016/0008-6223(75)90228-6
24. Sinharoy, S.; Levenson, L. L. *Thin Solid Films* **1978,** *53,* 31. doi:10.1016/0040-6090(78)90367-X
25. Diarra, M.; Amara, H.; Ducastelle, F.; Bichara, C. *Phys. Status Solidi B* **2012,** *249,* 2629. doi:10.1002/pssb.201200147
26. Bahr, H. A.; Bahr, T. *Ber. Dtsch. Chem. Ges. A/B* **1930,** *63,* 99. doi:10.1002/cber.19300630114
27. Israelachvili, J. N. *Intermolecular and surface forces,* 2nd ed.; Academic Press: Amsterdam, 2006.
28. Krinke, T. J.; Deppert, K.; Magnusson, M. H.; Schmidt, F.; Fissan, H. *J. Aerosol Sci.* **2002,** *33,* 1341. doi:10.1016/S0021-8502(02)00074-5

License and Terms

Beilstein Journal of Nanotechnology

Kelvin probe force microscopy of nanocrystalline TiO_2 photoelectrodes

Alex Henning[*1,2], Gino Günzburger[1], Res Jöhr[1], Yossi Rosenwaks[2], Biljana Bozic-Weber[3], Catherine E. Housecroft[3], Edwin C. Constable[3], Ernst Meyer[1] and Thilo Glatzel[1]

Full Research Paper

Open Access

Address:
[1]Department of Physics, University of Basel, Klingelbergstrasse 82 CH4056, Switzerland, [2]School of Electrical Engineering, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv 69978, Israel and [3]Department of Chemistry, University of Basel, Spitalstrasse 51 CH4056, Switzerland

Email:
Alex Henning[*] - henning@post.tau.ac.il

* Corresponding author

Keywords:
atomic force microscopy (AFM); dye-sensitized solar cells (DSC); Kelvin probe force microscopy (KPFM); surface photovoltage (SPV); titanium dioxide (TiO_2)

Beilstein J. Nanotechnol. **2013,** *4,* 418–428.
doi:10.3762/bjnano.4.49

Received: 08 March 2013
Accepted: 06 June 2013
Published: 01 July 2013

This article is part of the Thematic Series "Advanced atomic force microscopy techniques".

Associate Editor: J. Frommer

Abstract

Dye-sensitized solar cells (DSCs) provide a promising third-generation photovoltaic concept based on the spectral sensitization of a wide-bandgap metal oxide. Although the nanocrystalline TiO_2 photoelectrode of a DSC consists of sintered nanoparticles, there are few studies on the nanoscale properties. We focus on the microscopic work function and surface photovoltage (SPV) determination of TiO_2 photoelectrodes using Kelvin probe force microscopy in combination with a tunable illumination system. A comparison of the surface potentials for TiO_2 photoelectrodes sensitized with two different dyes, i.e., the standard dye N719 and a copper(I) bis(imine) complex, reveals an inverse orientation of the surface dipole. A higher surface potential was determined for an N719 photoelectrode. The surface potential increase due to the surface dipole correlates with a higher DSC performance. Concluding from this, microscopic surface potential variations, attributed to the complex nanostructure of the photoelectrode, influence the DSC performance. For both bare and sensitized TiO_2 photoelectrodes, the measurements reveal microscopic inhomogeneities of more than 100 mV in the work function and show recombination time differences at different locations. The bandgap of 3.2 eV, determined by SPV spectroscopy, remained constant throughout the TiO_2 layer. The effect of the built-in potential on the DSC performance at the TiO_2/SnO_2:F interface, investigated on a nanometer scale by KPFM measurements under visible light illumination, has not been resolved so far.

Introduction

Dye-sensitized solar cells (DSCs) provide a promising low-cost, high-efficiency third-generation photovoltaic concept based on the spectral sensitization of a nanoporous wide bandgap semiconductor [1,2]. In the past two decades DSCs have received substantial attention from both academic and industrial communities focusing on new materials and advanced device concepts [3-8]. A typical DSC consists of a dye-coated TiO_2 photoelectrode, deposited on a fluorine-doped tin oxide (FTO) conductive-glass substrate, an I^-/I_3^- redox-couple-based electrolyte and a platinum counter electrode as depicted in Figure 1. Upon visible-light excitation, dye molecules inject electrons into the conduction band, E_{cb}, of the semiconductor; the oxidized dye is subsequently reduced by the redox couple of the surrounding electrolyte. The generated electrons diffuse toward the SnO_2:F substrate and establish the photovoltage. The most frequently used dye complexes contain less-abundant transition metal elements such as ruthenium. Complexes of earth-abundant metals such as zinc and copper are candidates to replace the more expensive ruthenium dyes [9-13]. Recently, Yella et al. reported an efficiency of over 12% with a porphyrin-sensitized DSC and a cobalt(II/III) based redox electrolyte [14]. However, many details of the hybrid organic/inorganic interface and the influence of subsequent preparation steps on the device properties, e.g., surface topography and potential, are still unclear and have the potential to increase the efficiency and long-term stability of the devices. Investigations of nanoscaled photovoltaic devices require nanometer-scale measuring methods, including time-resolved measurements of the carrier dynamics [15,16]. Although a DSC photoelectrode consists of a nanostructured TiO_2, there are few microscopic studies [17].

Surface photovoltage (SPV) spectroscopy is a non-destructive and sensitive method for determining surface potential changes upon illumination, identifying surface states, and extracting material parameters, in particular the bandgap, E_g, the minority carrier diffusion length, L_n, and the flatband potential, V_{fb} [18]. SPV spectroscopy is usually performed with a macroscopic vibrating capacitor and is hence limited by its poor lateral resolution [19,20]. Bare and dye-sensitized nanocrystalline (nc) TiO_2 have been investigated with such a macroscopic Kelvin probe (KP) revealing details about the electronic structure [21-23], trap states [24], the surface dipole [25], charge-carrier dynamics [26], and indicating changes upon chemical treatments [24,27-29]. KP studies have helped to select surface treatments that are beneficial for the DSC performance. In order to achieve a nanometer scale resolution, SPV spectroscopy can be combined with Kelvin probe force microscopy (KPFM) [30-32], an atomic force microscopy (AFM) technique that was introduced in 1991 [33]. KPFM is a surface-potential detection method that determines the contact potential difference (CPD) during scanning by compensating the electrostatic forces between a microscopic tip and the sample [34]. Figure 2a illustrates a schematic band diagram for a KPFM tip in close proximity to a semiconductor sample surface with surface states, E_{trap}. An applied dc voltage, V_{dc} = CPD, nullifies the work-function difference, $\Delta\Phi$, between both materials. The occupied surface states of the n-type semiconductor, depicted in Figure 2a, are depopulated upon illumination with an appropriate light energy. Consequently, the surface band bending of an n-type semiconductor is shifted downwards and the measured CPD decrease is equal to the SPV.

The considerably high performance in DSCs is achieved also due to the high surface-to-volume ratio of nanocrystalline TiO_2. In any case, there is a trade-off between a high surface-to-volume ratio and the carrier transport. Smaller TiO_2 particles

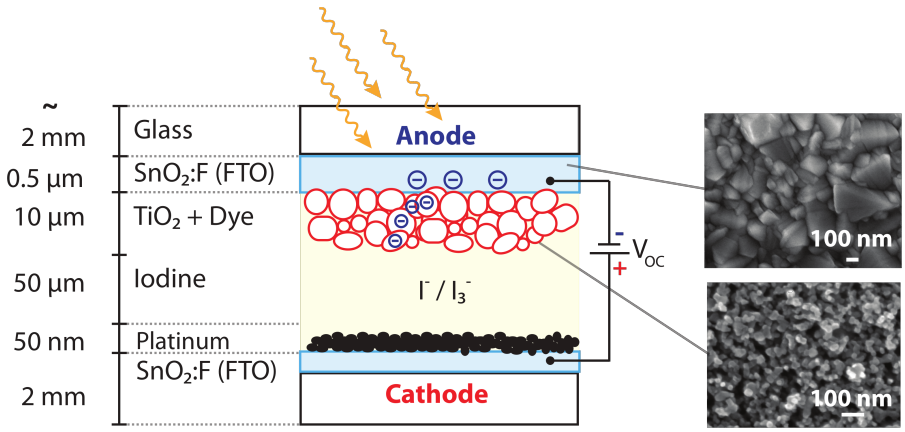

Figure 1: Different components of a DSC under illumination in an open circuit. Upon light excitation electrons are injected from the adsorbed dye molecules into the conduction band, E_{cb}, of the wide-bandgap metal oxide (nanoporous TiO_2) resulting in an open-circuit voltage, V_{oc}.

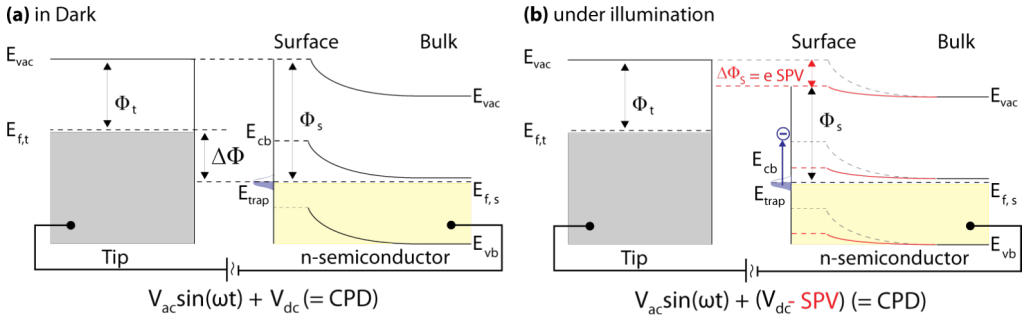

Figure 2: Schematic band diagram for a KPFM tip in close proximity to an n-type semiconductor surface (a) in the dark and (b) under illumination while the CPD is nullified by an applied dc voltage. Upon illumination the local vacuum energy level, E_{vac}, is shifted downwards and detected as a work function or CPD decrease, which results in a "negative" SPV. $E_{f,t}$ and $E_{f,s}$ are the Fermi levels of tip and sample, respectively. E_{vb} and E_{cb} are the valence and conduction band edges of the semiconductor. E_{vac} is the local vacuum and E_{trap} a surface-state energy level, respectively. The work-function shift, $\Delta\Phi_s$, upon illumination is equal to e SPV.

lead to an increase of grain boundaries and reduce the solar cell current. Hence, we have considered it as relevant to characterize the surface potential of nanostructured TiO_2 with a high-resolution method. Surface dipole changes upon dye adsorption induce a shift of the surface potential in the order of hundreds of millivolts, which is detectable by KPFM on the nanometer scale [35-37]. A direct influence of the surface dipole on the open-circuit voltage, V_{oc}, of a DSC was predicted by Angelis et al. [38] and experimentally addressed by KPFM investigations [39,40]. KPFM studies in UHV conditions of rutile TiO_2 decorated with either nanometer-sized Pt clusters [41] or single dye molecules [42] revealed a significant impact of single particles on the surface dipole. We have investigated the surface parameters of DSC photoelectrodes on the nanoscale using KPFM, which is not possible to achieve with a macroscopic KP. SPV spectra were taken on desired locations with a lateral resolution of 25 nm. Thus, the bandgap and time constants were obtained on the nanoscale. In this work, microscopic variations of the work function were observed for both sensitized and bare nc-TiO_2. To correlate the microscopic changes on a dry photoelectrode with the macroscopic DSC parameters, local surface dipole variations for a ruthenium(II)- and a copper(I)-based dye were determined. The ruthenium(II) dye chosen was the standard dye N719. The copper(I)-based dye (Figure 3) was selected from a range of complexes that we have recently prepared and screened for their potential use as sensitizers [43].

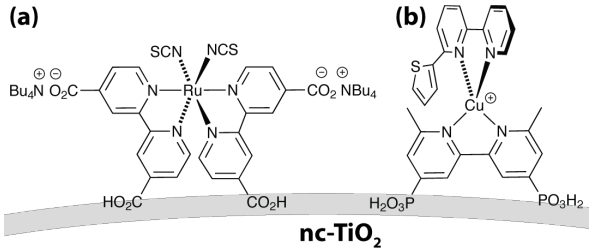

Figure 3: Schematic structures of (a) the standard dye N719 and (b) a copper(I)-based dye, assembled in situ (see text).

Results and Discussion

Work function inhomogeneities

Figure 4 shows the topography and the work function of a bare TiO_2 and an N719-sensitized TiO_2 layer measured by KPFM in a dry nitrogen glove box at room temperature. The topography images reveal, in both cases, a homogeneous surface with nanoparticles, nominal diameter of 20 nm, in the range of 20–100 nm. Work-function (Φ) variations reflecting the position of the conduction band edge E_{cb} of 80 mV on average, appear for both samples and are visible as dark regions in the measurements. They are highlighted in the cross sections in the lower part of the image. Such a local work function shift can be attributed to local variations of chemisorbed contaminants resulting in a decrease of the local vacuum energy, E_{vac}, and the electron affinity, χ. A thin water layer consisting of chemisorbed and physisorbed H_2O molecules on the nc-TiO_2 is known to be present even inside a dry nitrogen glove box [44]. Solvent residues are further possible contaminants that can be locally attached to the TiO_2 surface, or the variations may be due to varying material properties in general. In any case, such variations, which are clearly detectable by KPFM, may obstruct the optimal attachment of dye molecules and thus reduce the solar cell performance [25].

Microscopic surface photovoltage

By combining a tunable illumination system with KPFM, the surface photovoltage (SPV) can be measured on the nanometer scale and is referred to as microscopic SPV. A microscopic SPV is caused by an electron generation upon light absorption

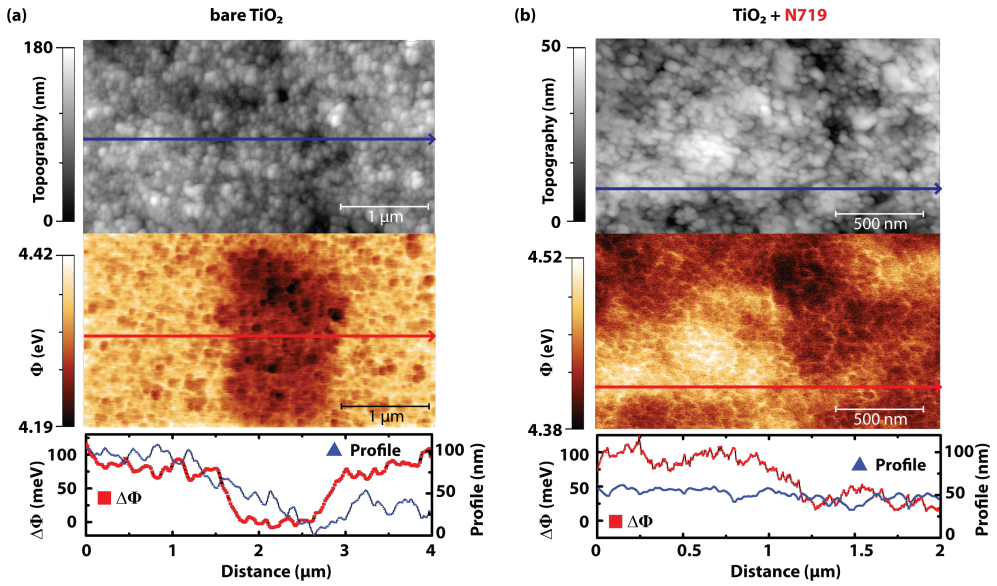

Figure 4: Topography and work function of (a) a bare TiO_2 and (b) an N719 sensitized TiO_2 layer with a thickness of ≈10 µm revealing wide-spread inhomogeneities in the work function. The measurements correspond to a scan size of (a) 2 × 4 µm and (b) 1 × 2 µm. Imaging parameters: A_{free} = 20 nm rms, A_{set} = 70%, f_{1st} = 72 kHz, f_{2nd} = 452 kHz, U_{ac} = 2 V, T = rt. The TiO_2 is a commercial product from Solaronix, Ti-Nanoxide T.

at either the surface space-charge region or at the electric field of a buried interface that is reached by the incident light [19]. In the present work, the sample was illuminated with focused light from an optical fiber or directly with a laser. The measured SPV can have two contributions, one from the TiO_2/SnO_2:F interface and the other from the TiO_2 layer depending on the energy of the incident light, i.e., super- or sub-bandgap(TiO_2) illumination. Both SPV effects are described separately in the following two sections. Time-resolved SPV measurements provide insights into charge carrier dynamics [45] and are described below.

Surface photovoltage under super-bandgap illumination

SPV spectroscopy (SPS) is a common method for measuring the bandgap, E_g, of a semiconductor by determining its dependency on the absorption coefficient, α. The obtained bandgap for nc-TiO_2 (Figure 5a) is in accordance with the literature value for bulk TiO_2, E_g = 3.2 eV [46] and validates the SPS setup. The extraction of E_g by means of SPS is superior to the usual transmission spectra since it is also applicable to thin layers, nanowires, or single nanoparticles and also for opaque samples [18]. Under illumination with a sufficiently low light intensity, the SPV can be assumed to be proportional to the absorption coefficient, implying a maximum SPV for super-bandgap illumination. Depending on the bandgap type, either direct or indirect, the SPV curve is fitted with the corresponding relation [18,47]:

$$\mathrm{SPV} \propto \alpha \propto (h\nu - E_g)^2 \text{ for indirect bandgaps,} \quad (1)$$

$$\mathrm{SPV} \propto \alpha \propto \sqrt{h\nu - E_g} \text{ for direct bandgaps,} \quad (2)$$

where h is the Planck constant and ν is the frequency of the light. For anatase TiO_2, an indirect bandgap material [48], α is therefore expected to show a quadratic dependence on the illumination wavelength for energies just above the bandgap. Figure 5a presents an SPS measurement taken on a cluster of sintered anatase particles showing a quadratic dependence on the wavelength. By linear fitting, a bandgap energy of E_g = 3.2 eV was extracted using Equation 1, assuming a phonon energy $E_p \approx 0$.

Figure 5b depicts the SPV of bare TiO_2 as a function of the light intensity for super-bandgap illumination with a wavelength of 380 nm. The negative SPV indicates an n-type behavior of the material. The SPV exhibits a linear dependency on the light intensity up to a value of −250 mV. A logarithmic dependence on the light intensity would be typical for a charge separation at a built-in potential, for instance at the surface space-charge region. However, the linear dependence indicates a charge separation, which is governed mainly by diffusion and not by drift current (electric field). Preferential trapping of electrons (holes) in defect states of the TiO_2 network leads to different diffusion coefficients for electrons and holes.

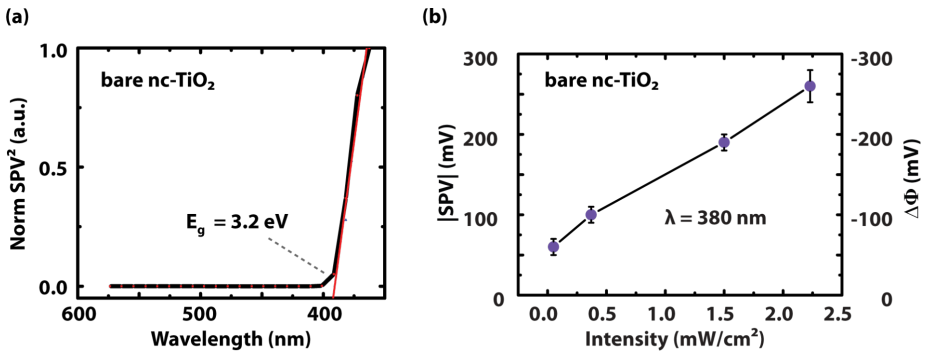

Figure 5: SPV for bare nc-TiO_2 in dependency on (a) the wavelength and (b) on the light intensity under super-bandgap illumination (380 nm). The bandgap of the material was extracted by SPV spectroscopy.

Surface photovoltage under sub-bandgap illumination

Figure 6a shows a semilogarithmic plot of the SPV versus the light intensity for three different wavelengths for the bare nc-TiO_2. An SPV of −230 mV was reached under sub-bandgap (λ = 408 nm) illumination. A negligible sub-bandgap SPV of less than 20 mV was measured for λ = 408 nm on TiO_2 layers directly deposited on glass. We assume that the SPV under sub-bandgap illumination results from a reduction of the built-in potential, V_{bi}, at the TiO_2/SnO_2-interface. This built-in electric field is screened by the photogenerated charge carriers resulting in a downward band bending of the TiO_2 conduction band edge E_{cb}.

The measured SPV of −250 mV under super-bandgap illumination (Figure 5b) provides an estimation for this downward band bending of E_{cb}. Figure 5b shows the CPD decrease with onset illumination, which decreases further with higher illumination intensities. A CPD decrease is equivalent to a work function decrease of the TiO_2. The observed logarithmic dependence demonstrates a photodiode behavior according to Equation 5 and is an indication for a built-in (Schottky barrier) potential at the interface. It should be noted that the photocurrent, J_{ph}, in Equation 5 is approximately proportional to the light intensity, I. When the incident light wavelength approaches the bandgap energy of TiO_2 higher SPVs result leading to steeper slopes of the SPV-versus-intensity curves. The SPV is proportional to the number of photogenerated charge carriers. It is evident from Figure 6a that more electrons are generated with higher illumination energies within the TiO_2 network. We assume that empty surface states just below the conduction band edge are occupied by valence band electrons. According to Howe et al. there are localized Ti^{3+}(3d) trap states just below the conduction band edge of nc-TiO_2 [49]. According to our measurements, the SPV decreases exponentially with decreasing illumination energies and we conclude, therefore, that the number of trap states also decreases exponentially with decreasing trap state energy relative to the conduction band edge of the TiO_2 nanoparticles.

The buried TiO_2/SnO_2-interface is reached by the incident light since the nanoporous TiO_2, deposited on top of the FTO-layer, is only about 10 µm thin and transparent to visible light

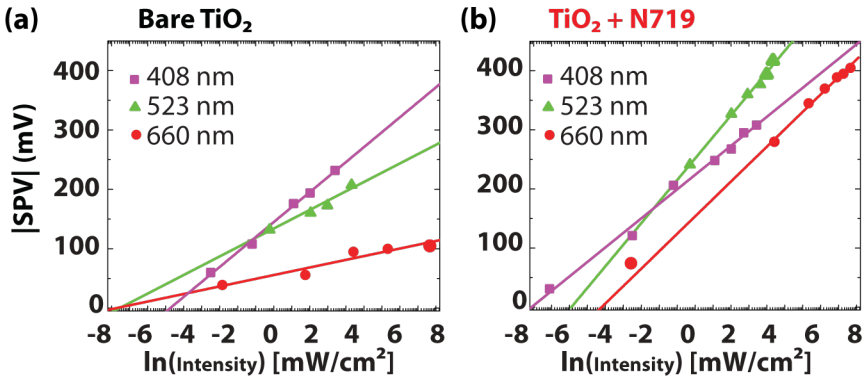

Figure 6: Semilogarithmic plot of the SPV dependence on the incident light intensity, (a) measured for three different wavelengths on bare and (b) N719-sensitized TiO_2.

(E_g = 3.2 eV). Due to the high n-dopant density of SnO_2:F ($N_D \sim 10^{20}$), it can be approximated to being nearly metallic. Generally, the SnO_2:F contact is regarded as Ohmic since electrons may tunnel through the ca. 2 nm thin barrier at the interface space charge region [50]. However, the TiO_2/SnO_2:F contact forms a heterojunction between two wide-bandgap semiconductors, degenerately doped SnO_2 and (intrinsic) nc-TiO_2. Kron et al. and Levy et al. [51,52] investigated alternative materials to SnO_2:F and found that the built-in voltage at the interface has no significant influence on V_{oc} of the DSC but does influence the fill factor, FF. With the determined work functions of 4.3 ± 0.1 eV for nc-TiO_2 and 4.7 ± 0.1 eV for SnO_2:F in our KPFM measurements, the band offset at the heterojunction allows an estimation of the energy barrier. Depending on the front electrode material, this energy barrier varies and consequently the interface contact resistance differs. As a result, the FF of the DSC can be increased with a lower interface energy barrier and a narrower space-charge region decreasing the sheet resistance at the interface.

For the sensitized TiO_2, the SPV also shows a logarithmic dependence on the light intensity (Figure 6b). We conclude that the measured SPV of the sensitized TiO_2 is created by two different effects: a change of the surface dipole (after electron donation) and a charge carrier concentration gradient between the illuminated surface and the bulk due to different diffusion coefficients for electrons and holes (photo-Dember effect). The latter effect causes a potential drop forming an electric field in the z-direction across the TiO_2 layer [29,53,54]. The Dember photovoltage is caused by a non-uniform generation or recombination of charge carriers within the sample [18]. The adsorbed dye molecules are considered as n-type "photodoping" since electrons are generated under sub-bandgap illumination. After electron injection into the conduction band of TiO_2, the dye is oxidized and charged more positive relative to its ground state. Hence, the surface dipole is reduced and detected as a change in the CPD.

Time evolution of the SPV

Time-dependent SPV measurements were performed at specific positions above single TiO_2 particles. Since KPFM is sensitive to potential drops in the entire sample, the measurements give insights into charge-carrier transport processes of the particle network. Figure 7 shows the time evolution of the measured CPD values after the turning on and off of the laser light illumination with a wavelength of 408 nm. For both sensitized and bare TiO_2 the SPV with onset illumination, t_{on}, is below the resolution limit (50 ms) of the measurement system. The photoresponse time corresponds to the required time for the charge carriers to reach a steady-state condition upon illumination. In turn the recombination time, t_{off}, is the time required to reach the initial value in the dark. Recombination times of 65 ± 6 s for bare TiO_2 and 43 ± 4 s for N719 were determined with KPFM. The recombination curve was divided into a fast and a slow component and approximated as the sum of two exponential functions. The slow component of the total recombination time is attributed to an electron diffusion process across the TiO_2 network towards the contact, whereas the fast recombination process occurs within single particles [55]. The slow electron diffusion throughout the network is due to trapping and detrapping [46] in surface and bulk defect states. TiO_2 is regarded as an insulator with a relative permittivity of $\varepsilon_r = 36$ and consequently acts as a charge storage capacitor between a metallic tip and a highly conductive SnO_2:F contact. Upon photoelectric charge injection, the redistribution of charge carriers (by diffusion through the network) is slow (seconds to minutes). A slow response time has also been reported for nanoporous TiO_2 [56,57] and for porous Si, which exhibited recombination times of up to 1 h [58].

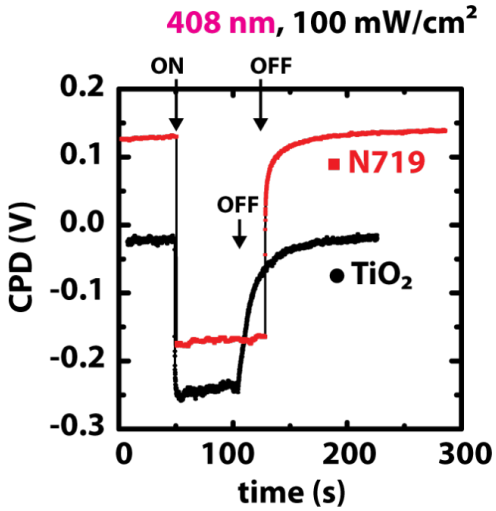

Figure 7: Time evolution of the measured CPD of TiO_2 and TiO_2 + N719 during the turning on and off of the violet (408 nm) laser light.

Microscopic surface-dipole variations

By averaging the work function values over several images on different sample spots an increase of $\Delta\Phi$ = 150 ± 40 mV for N719 and an average decrease of $\Delta\Phi$ = −180 ± 40 mV for the copper-containing dye was determined on sensitized TiO_2 films by KPFM. The values as well as a model describing the dipole moment strength and orientation are presented in Figure 8.

The surface dipole is the result of a combined effect of both anchoring domain and dye molecule. The effective electron affinity, χ^*, is affected by the surface dipole formed by adsorbed dye molecules [59]. The detailed anchoring mechanism for the N719 dye on TiO_2 is still under debate. It is widely

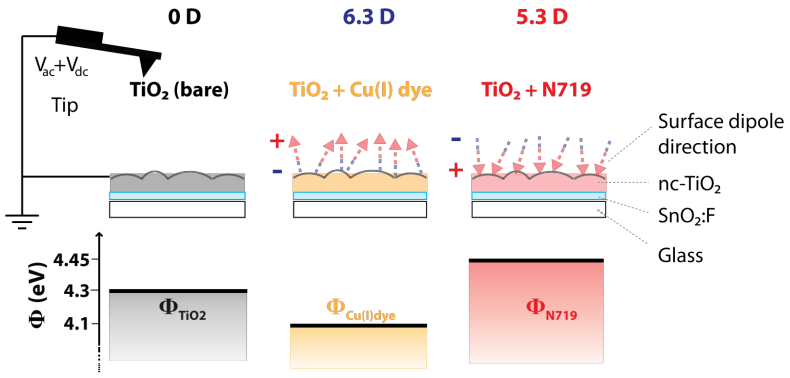

Figure 8: Schematic illustration of the KPFM measurement system and the surface dipole induced by adsorption of the ruthenium containing dye N719 (6.3 D) and the copper(I) dye (5.3 D) pointing in opposite directions. The measured work-function values are compared with the bare TiO_2 substrate without external illumination.

accepted that the N719 is chemisorbed either through two or three carboxylic acid groups. Results by Lee et al. support that an additional hydrogen bonding (physisorption) is present [60]. Molecular dynamics simulations by DeAngelis et al. show that the binding occurs through three carboxylic acid groups and that the protons initially carried by the N719 dye are transferred to the semiconductor surface [61]. After covalent attachment to the TiO_2 surface, the formed dipole might turn compared to the dipole of the free molecule in vacuum. After N719 has been chemisorbed onto TiO_2, the surface is protonated and possesses a dipole pointing from the TiO_2 surface to the negative net charge (isothiocyanato-group,-NCS) as shown in Figure 8 [38]. As a consequence the local vacuum level is bent upwards and the work function is increased compared with the bare TiO_2 surface. For the nanoporous TiO_2 surface sensitized with the Cu(I)-containing dye, a negative surface dipole pointing away from the surface leads to a decrease in work function. The copper(I) dye is a monocation in its fully protonated state, assuming that the phosphonic acid functionalities are fully protonated.

To quantify the surface dipole, Natan et al. proposed a plate capacitor model, in which the adsorbed molecules are regarded as point dipoles [62]:

$$P = \frac{P_0}{1+\frac{\alpha k}{a^3}} \equiv \frac{P_0}{\varepsilon}, \tag{3}$$

where P is the surface dipole moment, P_0 is the dipole moment of the free molecule in vacuum, k is a geometric correction factor and a is the distance between two dipoles. The change in work function, $\Delta\Phi_S$, is related to the surface dipole through the Helmholtz equation:

$$\Delta\Phi_S = 4\pi\frac{NP_0\cos\theta}{A\,\varepsilon\cdot\varepsilon_0}, \tag{4}$$

where (N/A) is the number of dipoles/molecules per surface area, $\varepsilon = (P_0/P)$ is the effective dielectric constant of a molecular monolayer and ε_0 is the permittivity in vacuum. The dipole layer is oriented at an angle, θ, relative to the surface plane normal. Due to the curved surface geometry of nc-TiO_2, the mean dipole in the z-direction is reduced (Figure 8).

A surface coverage of N/A = 1/4 molecules/nm^2 is a reasonable value found by Ikeda et al. by AFM measurements of N3 (N719 is the salt of N3) adsorbed on rutile TiO_2 in ultrahigh vacuum [42]. Using Equation 4 with $\theta = 0°$ and measured work-function shifts of $\Delta\Phi = -180 \pm 40$ mV for the Cu(I) dye and $\Delta\Phi = 150 \pm 40$ mV for N719 results in 6.3 ± 1.5 D and 5.3 ± 2 D with opposite directions, respectively. The latter value is in the same range as predicted by DFT calculations for N719 [38] and N3 [42,63] adsorbed on anatase plane-surface. However, for a complete DSC device the surface dipole may change due to screening by the surrounding electrolyte [64].

Figure 9a depicts the I–V characteristics for three different DSCs, a bare TiO_2 solar cell with electrolyte and DSCs sensitized with a Cu(I)-dye or N719, respectively. The parameters of the solar cell were extracted from the I–V data and are summarized in Table 1. A DSC sensitized with N719 is 3 times more efficient than the Cu-sensitized solar cell. We focus on the difference in the open-circuit voltages of −220 ± 20 mV between these DSCs. The deviation of the open-circuit voltages is in the range of the measured difference for the surface dipoles (350 ± 40 mV) formed by N719 and the copper(I) dye. The two sensitizers possess oppositely directed dipole moments.

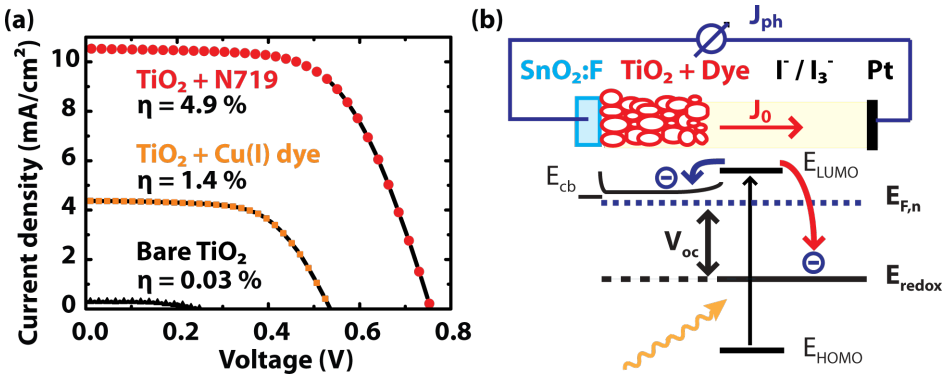

Figure 9: (a) *I–V* curves for a bare TiO_2 solar cell and DSCs sensitized with Cu(I) dye and N719. (b) A schematic band diagram for a DSC under light excitation of the dye. The desired forward reaction (blue arrow), i.e., electron transfer from E_{LUMO} into the conduction band, E_{cb} of TiO_2, is accompanied by a backward electron injection (red arrow) from E_{LUMO} into E_{redox}.

Table 1: *I–V*-characteristic values of a DSC sensitized with N719 or the copper(I)-based dye.

dye	J_{ph} [mA/cm^2]	V_{oc} [V]	FF	dipole direction	η [%]
Cu(I) dye	4	0.53	0.60	↑	1.4
N719	10	0.75	0.65	↓	4.9

Pandey et al. have investigated the surface dipoles of organic dye molecules adsorbed on TiO_2 using KPFM [40]. They observed a decrease of V_{oc} for DSCs sensitized with dye molecules, which lead to a more positive surface potential. De Angelis et al. found that the adsorption geometry of the sensitizer on the TiO_2 surface has a significant influence on V_{oc}. The desired forward reaction, i.e., electron injection from E_{LUMO} into E_{cb} of TiO_2 is accompanied by a backward electron injection from E_{LUMO} into E_{redox} (Figure 9b). This backward reaction is affected by the surface dipole of the adsorbed sensitizer [64]. Regarding the general expression for the open-circuit voltage (Equation 5), it is the reverse saturation current density, J_0, and the conduction band minimum, E_{cb}, that are influenced by the surface dipole and finally affect V_{oc}:

$$V_{oc} = A\frac{k_B T}{e}\ln\left(\frac{J_{ph}}{J_0}+1\right) \equiv E_{F,n} - E_{redox}\ , \qquad (5)$$

wherein A is the correction factor and J_{ph} the photocurrent density. A performance comparison of DSCs sensitized with N719 and the copper(I) dye is shown in Table 1.

Conclusion

Microscopic surface photovoltage and work-function measurements were performed on bare and dye-sensitized TiO_2 photoelectrodes using Kelvin probe force microscopy. Compared to a bare TiO_2 layer, the surface potential is about 150 mV higher for an N719 sensitized TiO_2 photoelectrode and about 180 mV lower for Cu-dye sensitized TiO_2 resulting in a 200 mV higher open-circuit voltage (V_{oc}) for a complete N719 DSC. We conclude that the surface dipole orientation is inverted for the two dyes and the V_{oc} of a complete DSC increases with a higher surface potential. Consequently, we assume that the detected microscopic surface potential drops/inhomogeneities on both bare and sensitized TiO_2 photoelectrodes lead to a lower V_{oc} and efficiency of the solar cell. The bandgap of 3.2 eV for anatase nanocrystalline TiO_2 particles was determined by SPV spectroscopy on different locations. It is constant throughout the TiO_2 layer and in agreement with literature values for bulk anatase. The measured SPV under sub-bandgap illumination is formed at the SnO_2:F/TiO_2 interface due to the presence of a built-in electric field. According to our results, the interface barrier is around 250 mV. Its influence on the DSC performance is not resolved. In the case of dye-sensitized photoelectrodes, three different mechanisms contribute to the measured SPV. First, there is a contribution from the SnO_2:F/TiO_2 interface, which forms a heterojunction. The band diagram at this heterojunction is still unclear; however, the influence on our measurements is expected to be negligible, and thus, we decided not to include it in this paper. Secondly, there is an SPV contribution from the nc-TiO_2 layer itself, and third, there is a reversible photochemical reaction of the dye molecule, which donates electrons under illumination and hence changes the

surface potential. To our best knowledge, the measured SPV is not understood well enough to represent it generally in a band diagram without making major assumptions. Further studies are needed to clarify this point.

Experimental

Preparation of the photoanode

A glass substrate coated with SnO_2:F (TCO 22-7, Solaronix) was cleaned in an ultrasonic bath successively with acetone, ethanol, and distilled water and subsequently treated with UV/ozone for 18 min (Model 42-220, Jelight Company). To study the SPV without a contribution from the SnO_2:F/TiO_2 interface, TiO_2 layers were also prepared on uncoated glass substrates. A colloidal TiO_2 solution (Ti-Nanoxide T, Solaronix) consisting of anatase nanoparticles (≈20 nm) was deposited by doctor blading. A mask with an area of 25 mm^2 was defined with scotch tape. The TiO_2 layer was sintered with a heating plate for 30 min at 500 °C resulting in a layer thickness of about 10 µm.

For the preparation of sensitized TiO_2, the photoanode was immersed in a sensitizer solution for 24 h after cooling it to 80 °C. The ruthenium dye solution consisted of a 1 mM ethanol solution of standard dye N719. The nc-TiO_2 was sensitized with the heteroleptic copper(I) dye shown in Figure 3 by using the procedure previously described, which involves an in situ assembly of the dye starting with the adsorbed anchoring ligand and the hexafluoridophosphate salt of the homoleptic complex $[Cu(6\text{-}(2\text{-thienyl})\text{-}2,2'\text{-bipyridine})_2]^+$ [43]. In order to remove weakly adsorbed contaminants, the sensitized TiO_2 was rinsed with ethanol and dried under nitrogen.

Kelvin probe force microscopy

AFM measurements were carried out inside a glove box (labmaster 130, mBraun) maintaining a dry nitrogen atmosphere (<1 ppm H_2O and <10 ppm O_2) on a commercial microscope (Solver PH-47, NT-MDT). Amplitude modulation (AM) KPFM was conducted with a two-scan method (lift mode) meaning that the topography and CPD were measured separately (Figure 10). During the first line scan the topography was determined in tapping mode AFM using a bias voltage of V_{dc} = 0 V applied to the tip. The cantilever oscillation amplitude was kept constant by a feedback controller at a setpoint, A_{sp}, of 20 nm rms that was pre adjusted to 75% of the free vibrational amplitude, A_0 = 27 nm rms. The second scan was performed 20 nm above the previously obtained surface profile on the same scan line. While the sample was grounded, the dc and ac voltage were applied to the cantilever, V_{ac} = 2 V at a frequency equal to its first resonance, f_{1st}. The X-component, $X = A \cdot \cos(\theta)$, of the electrically induced oscillation signal, which is proportional to the electrostatic force, F_{es}, was phase-adjusted and retrieved with a lock-in amplifier and subsequently nullified by the applied dc voltage in the Kelvin feedback. Cantilever bending and vibration were optically detected with an infrared laser (λ = 1300 nm, 1 mW, spot size 50 × 50 μm^2) and a four-quadrant photodetector. To deduce the sample work function and monitor the integrity of the tip, a highly ordered pyrolytic graphite (HOPG) reference sample was measured before and after each measured specimen. The HOPG reference with a known work function of ϕ = 4.6 eV [65] was stored inside the glove box.

Microscopic photovoltage determination

A UV-lamp (360 nm, L18W/73, Osram) was used for super-bandgap illumination. Trap and interface states were investigated under sub-bandgap illumination (λ > 385 nm) with LEDs (405 nm, 470 nm, 530 nm, 590 nm, Mightex Systems) or solid state lasers each mounted on a x-y-z positioning stage. A red laser (660 nm, 60 mW, LPM660-60C, Newport), a green laser (532 nm, 4 mW, Alpec), a violet laser (408 nm, 2.5 mW, Power Technology) and LEDs were applied for illumination from above at an angle of 10° with respect to the sample plane. The laser illumination was switched with a shutter (T132, Uniblitz). Laser light intensity was gradually reduced with different neutral density (ND) filters. Incident light intensity was

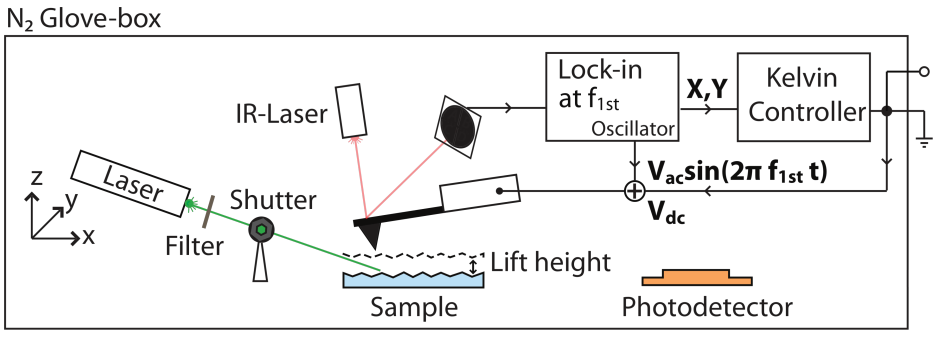

Figure 10: Lift-mode KPFM setup inside a nitrogen glove box in combination with a tunable illumination system for microscopic SPV measurements. The CPD was nullified by an applied dc voltage between tip and sample.

measured at the sample location with a calibrated silicon photodetector (PD300-UV-SH, Ophir) and a power meter head (AN/2, Ophir). SPV spectroscopy was conducted on a commercial AFM (Dimension 3100, Bruker). The cantilever vibration was optically detected with a red laser (670 nm, 1 mW, spot size ≈ 40 × 40 μm^2), which partially hit the sample surface with an intensity of ≈100 μW. In order to avoid background illumination by the red laser, the spot was positioned on the cantilever center. The light of a xenon-arc lamp (300 W, model 6258, Newport) was focused onto the entrance slit of a grating monochromator (MS257, Newport) with a resolution of 5 nm in the range between 350–700 nm. The monochromatic light was coupled into an optical light guide (3 mm, 1.83 m, N_A = 0.55, Edmund optics) which was connected to an optical microscope built-in within (inside) the AFM. Finally, the outcoming light of the optical microscope was focused onto the sample surface with a spot diameter of about 500 μm and a depth of focus of about 10 μm.

Photovoltaic characterisation of DSCs

I–*V*-curves were measured under an AM1.5G solar simulator (ABET) with an intensity of 100 mW/cm^2 and a controlled temperature of 25 °C. The voltage sweep was performed in 6 mV steps with a sampling rate of 1 Hz.

Acknowledgements

This work was supported in part by the Swiss National Science Foundation, the European Research Council (Advanced Grant 267816 LiLo) and the Nano Argovia programme of the Swiss Nanonscience Institute. We thank Peter Kopecky for the copper(I)-dye synthesis.

References

1. O'Regan, B.; Grätzel, M. *Nature* **1991,** *353,* 737–740. doi:10.1038/353737a0
2. Green, M. A.; Emery, K.; Hishikawa, Y.; Warta, W.; Dunlop, E. D. *Prog. Photovoltaics* **2012,** *20,* 606–614. doi:10.1002/pip.2267
3. Wang, H.; Liu, M.; Yan, C.; Bell, J. *Beilstein J. Nanotechnol.* **2012,** *3,* 378–387. doi:10.3762/bjnano.3.44
4. Motta, N. *Beilstein J. Nanotechnol.* **2012,** *3,* 351–352. doi:10.3762/bjnano.3.40
5. Kaneko, M.; Ueno, H.; Nemoto, J. *Beilstein J. Nanotechnol.* **2011,** *2,* 127–134. doi:10.3762/bjnano.2.15
6. Bora, T.; Kyaw, H. H.; Sarkar, S.; Pal, S. K.; Dutta, J. *Beilstein J. Nanotechnol.* **2011,** *2,* 681–690. doi:10.3762/bjnano.2.73
7. Wang, M.; Chamberland, N.; Breau, L.; Moser, J.-E.; Humphry-Baker, R.; Marsan, B.; Zakeeruddin, S. M.; Grätzel, M. *Nat. Chem.* **2010,** *2,* 385–389. doi:10.1038/nchem.610
8. Wang, M.; Xu, M.; Shi, D.; Li, R.; Gao, F.; Zhang, G.; Yi, Z.; Humphry-Baker, R.; Wang, P.; Zakeeruddin, S. M.; Grätzel, M. *Adv. Mater.* **2008,** *20,* 4460–4463. doi:10.1002/adma.200801178
9. Alonso-Vante, N.; Nierengarten, J.; Sauvage, J. *J. Chem. Soc., Dalton Trans.* **1994,** 1649–1654. doi:10.1039/DT9940001649
10. Constable, E. C.; Redondo Hernandez, A.; Housecroft, C. E.; Neuburger, M.; Schaffner, S. *Dalton Trans.* **2009,** 6634–6644. doi:10.1039/B901346F
11. Bozic-Weber, B.; Constable, E. C.; Hostettler, N.; Housecroft, C. E.; Schmitt, R.; Schönhofer, E. *Chem. Commun.* **2012,** *48,* 5727–5729. doi:10.1039/C2CC31729J
12. Bozic-Weber, B.; Brauchli, S. Y.; Constable, E. C.; Fürer, S. O.; Housecroft, C. E.; Wright, I. A. *Phys. Chem. Chem. Phys.* **2013,** *15,* 4500–4504. doi:10.1039/c3cp50562f
13. Bozic-Weber, B.; Constable, E. C.; Housecroft, C. E. *Coord. Chem. Rev.* **2013,** in press. doi:10.1016/j.ccr.2013.05.019
14. Yella, A.; Lee, H.-W.; Tsao, H. N.; Yi, C.; Chandiran, A. K.; Nazeeruddin, M. K.; Diau, E. W.-G.; Yeh, C.-Y.; Zakeeruddin, S. M.; Grätzel, M. *Science* **2011,** *334,* 629–634. doi:10.1126/science.1209688
15. Nicholson, P. G.; Castro, F. A. *Nanotechnology* **2010,** *21,* 1–26. doi:10.1088/0957-4484/21/49/492001
16. Shen, Q.; Ogomi, Y.; Park, B.-w.; Inoue, T.; Pandey, S. S.; Miyamoto, A.; Fujita, S.; Katayama, K.; Toyoda, T.; Hayase, S. *Phys. Chem. Chem. Phys.* **2012,** *14,* 4605–4613. doi:10.1039/c2cp23522f
17. Graaf, H.; Maedler, C.; Kehr, M.; Baumgaertel, T.; Oekermann, T. *Phys. Status Solidi A* **2009,** *206,* 2709–2714. doi:10.1002/pssa.200925294
18. Kronik, L.; Shapira, Y. *Surf. Sci. Rep.* **1999,** *37,* 1–206. doi:10.1016/S0167-5729(99)00002-3
19. Kronik, L.; Shapira, Y. *Surf. Interface Anal.* **2001,** *31,* 954–965. doi:10.1002/sia.1132
20. Schroder, D. K. *Meas. Sci. Technol.* **2001,** *12,* 16–31. doi:10.1088/0957-0233/12/3/202
21. Cahen, D.; Hodes, G.; Graetzel, M.; Guillemoles, J. F.; Riess, I. *J. Phys. Chem. B* **2000,** *104,* 2053–2059. doi:10.1021/jp993187t
22. Lenzmann, F.; Krueger, J.; Burnside, S.; Brooks, K.; Grätzel, M.; Gal, D.; Rühle, S.; Cahen, D. *J. Phys. Chem. B* **2001,** *105,* 6347–6352. doi:10.1021/jp010380q
23. Beranek, R.; Neumann, B.; Sakthivel, S.; Janczarek, M.; Dittrich, T.; Tributsch, H.; Kisch, H. *Chem. Phys.* **2007,** *339,* 11–19. doi:10.1016/j.chemphys.2007.05.022
24. Yang, J.; Warren, D. S.; Gordon, K. C.; McQuillan, A. J. *J. Appl. Phys.* **2007,** *101,* 023714. doi:10.1063/1.2432106
25. Rühle, S.; Greenshtein, M.; Chen, S. G.; Merson, A.; Pizem, H.; Sukenik, C. S.; Cahen, D.; Zaban, A. *J. Phys. Chem. B* **2005,** *109,* 18907–18913. doi:10.1021/jp0514123
26. Duzhko, V.; Timoshenko, V. Y.; Koch, F.; Dittrich, T. *Phys. Rev. B* **2001,** *64,* 075204. doi:10.1103/PhysRevB.64.075204
27. Rothschild, A.; Levakov, A.; Shapira, Y.; Ashkenasy, N.; Komem, Y. *Surf. Sci.* **2003,** *532–535,* 456–460. doi:10.1016/S0039-6028(03)00154-7
28. Liu, Y.; Scully, S. R.; McGehee, M. D.; Liu, J.; Luscombe, C. K.; Fréchet, J. M. J.; Shaheen, S. E.; Ginley, D. S. *J. Phys. Chem. B* **2006,** *110,* 3257–3261. doi:10.1021/jp056576y
29. Warren, D. S.; Shapira, Y.; Kisch, H.; McQuillan, A. J. *J. Phys. Chem. C* **2007,** *111,* 14286–14289. doi:10.1021/jp0753934
30. Weaver, J. M. R.; Wickramasinghe, H. K. *J. Vac. Sci. Technol., B* **1991,** *9,* 1562–1565. doi:10.1116/1.585424
31. Saraf, S.; Shikler, R.; Yang, J.; Rosenwaks, Y. *Appl. Phys. Lett.* **2002,** *80,* 2586–2588. doi:10.1063/1.1468275
32. Streicher, F.; Sadewasser, S.; Lux-Steiner, M. C. *Rev. Sci. Instrum.* **2009,** *80,* 013907. doi:10.1063/1.3072661
33. Nonnenmacher, M.; O'Boyle, M. P.; Wickramasinghe, H. K. *Appl. Phys. Lett.* **1991,** *58,* 2921–2923. doi:10.1063/1.105227

34. Glatzel, T. Measuring Atomic-Scale Variations of the Electrostatic Force. In *Kelvin Probe Force Microscopy;* Sadewasser, S.; Glatzel, T., Eds.; Springer Series in Surface Sciences, Vol. 48; Springer: Berlin Heidelberg, 2011; pp 289–327.
35. Milde, P.; Zerweck, U.; Eng, L. M.; Abel, M.; Giovanelli, L.; Nony, L.; Mossoyan, M.; Porte, L.; Loppacher, C. *Nanotechnology* **2008,** *19,* 305501. doi:10.1088/0957-4484/19/30/305501
36. Glatzel, T.; Zimmerli, L.; Koch, S.; Kawai, S.; Meyer, E. *Appl. Phys. Lett.* **2009,** *94,* 063303. doi:10.1063/1.3080614
37. Kittelmann, M.; Rahe, P.; Gourdon, A.; Kühnle, A. *ACS Nano* **2012,** *6,* 7406–7411. doi:10.1021/nn3025942
38. De Angelis, F.; Fantacci, S.; Selloni, A.; Grätzel, M.; Nazeeruddin, M. K. *Nano Lett.* **2007,** *7,* 3189–3195. doi:10.1021/nl071835b
39. Sakaguchi, S.; Pandey, S. S.; Okada, K.; Yamaguchi, Y.; Hayase, S. *Appl. Phys. Expr.* **2008,** *1,* 105001. doi:10.1143/APEX.1.105001
40. Pandey, S. S.; Sakaguchi, S.; Yamaguchi, Y.; Hayase, S. *Org. Electron.* **2010,** *11,* 419–426. doi:10.1016/j.orgel.2009.11.021
41. Hiehata, K.; Sasahara, A.; Onishi, H. *Nanotechnology* **2007,** *18,* 084007. doi:10.1088/0957-4484/18/8/084007
42. Ikeda, M.; Koide, N.; Han, L.; Sasahara, A.; Onishi, H. *J. Phys. Chem. C* **2008,** *112,* 6961–6967. doi:10.1021/jp077065+
43. Bozic-Weber, B.; Constable, E. C.; Housecroft, C. E.; Kopecky, P.; Neuburger, M.; Zampese, J. A. *Dalton Trans.* **2011,** *40,* 12584–12594. doi:10.1039/C1DT11052G
44. Tan, M.; Wang, G.; Zhang, L. *J. Appl. Phys.* **1996,** *80,* 1186–1189. doi:10.1063/1.363727
45. Łagowski, J.; Balestra, C. L.; Gatos, H. C. *Surf. Sci.* **1972,** *29,* 203–212. doi:10.1016/0039-6028(72)90080-5
46. Schwarzburg, K.; Willig, F. *Appl. Phys. Lett.* **1991,** *58,* 2520–2522. doi:10.1063/1.104839
47. Moss, T. S. *Optical properties of semi-conductors;* Butterworths scientific publications and Academic Press: London and New York, 1959.
48. Kavan, L. *Chem. Rec.* **2012,** *12,* 131–142. doi:10.1002/tcr.201100012
49. Howe, R. F.; Grätzel, M. *J. Phys. Chem.* **1985,** *89,* 4495–4499. doi:10.1021/j100267a018
50. Peter, L. M. *Phys. Chem. Chem. Phys.* **2007,** *9,* 2630–2642. doi:10.1039/B617073K
51. Levy, B.; Liu, W.; Gilbert, S. E. *J. Phys. Chem. B* **1997,** *101,* 1810–1816. doi:10.1021/jp962105n
52. Kron, G.; Rau, U.; Werner, J. H. *J. Phys. Chem. B* **2003,** *107,* 13258–13261. doi:10.1021/jp036039i
53. Kronik, L.; Ashkenasy, N.; Leibovitch, M.; Fefer, E.; Shapira, Y.; Gorer, S.; Hodes, G. *J. Electrochem. Soc.* **1998,** *145,* 1748–1754. doi:10.1149/1.1838552
54. Dittrich, T.; Duzhko, V.; Koch, F.; Kytin, V.; Rappich, J. *Phys. Rev. B* **2002,** *65,* 155319. doi:10.1103/PhysRevB.65.155319
55. Dittrich, T. *Phys. Status Solidi A* **2000,** *182,* 447–455. doi:10.1002/1521-396X(200011)182:1<447::AID-PSSA447>3.0.CO;2-G
56. Bonhôte, P.; Moser, J.-E.; Humphry-Baker, R.; Vlachopoulos, N.; Zakeeruddin, S. M.; Walder, L.; Grätzel, M. *J. Am. Chem. Soc.* **1999,** *121,* 1324–1336. doi:10.1021/ja981742j
57. Nelson, J. *Phys. Rev. B* **1999,** *59,* 15374–15380. doi:10.1103/PhysRevB.59.15374
58. Timoshenko, V. Y.; Kashkarov, P. K.; Matveeva, A. B.; Konstantinova, E. A.; Flietner, H.; Dittrich, T. *Thin Solid Films* **1996,** *276,* 216–218. doi:10.1016/0040-6090(95)08056-2
59. Cohen, R.; Bastide, S.; Cahen, D.; Libman, J.; Shanzer, A.; Rosenwaks, Y. *Opt. Mater.* **1998,** *9,* 394–400. doi:10.1016/S0925-3467(97)00065-7
60. Lee, K. E.; Gomez, M. A.; Elouatik, S.; Demopoulos, G. P. *Langmuir* **2010,** *26,* 9575–9583. doi:10.1021/la100137u
61. De Angelis, F.; Fantacci, S.; Selloni, A.; Nazeeruddin, M. K.; Grätzel, M. *J. Phys. Chem. C* **2010,** *114,* 6054–6061. doi:10.1021/jp911663k
62. Natan, A.; Zidon, Y.; Shapira, Y.; Kronik, L. *Phys. Rev. B* **2006,** *73,* 193310. doi:10.1103/PhysRevB.73.193310
63. Fantacci, S.; De Angelis, F.; Selloni, A. *J. Am. Chem. Soc.* **2003,** *125,* 4381–4387. doi:10.1021/ja0207910
64. Miyashita, M.; Sunahara, K.; Nishikawa, T.; Uemura, Y.; Koumura, N.; Hara, K.; Mori, A.; Abe, T.; Suzuki, E.; Mori, S. *J. Am. Chem. Soc.* **2008,** *130,* 17874–17881. doi:10.1021/ja803534u
65. Sommerhalter, C.; Matthes, T. W.; Glatzel, T.; Jäger-Waldau, A.; Lux-Steiner, M. C. *Appl. Phys. Lett.* **1999,** *75,* 286. doi:10.1063/1.124357

License and Terms

The definitive version of this article is the electronic one which can be found at:
doi:10.3762/bjnano.4.49