

Accounting for the JKR–DMT transition in adhesion and friction measurements with atomic force microscopy

D. S. GRIERSON, E. E. FLATER and R. W. CARPICK *

Department of Engineering Physics, University of Wisconsin-Madison, 1500 Engineering Drive, Madison, WI 53706, USA

Received in final form 5 February 2005

Abstract—Over the last 15 years, researchers have applied theories of continuum contact mechanics to nanotribology measurements to determine fundamental parameters and processes at play in nanometer-scale contacts. In this paper we discuss work using the atomic force microscope to determine nanoscale adhesion and friction properties between solids. Our focus is on the role that continuum contact mechanics plays in analyzing these measurements. In particular, we show how the JKR-to-DMT transition is taken into account, as well as limitations involved in using these models of contact in the presence of adhesion.

Keywords: Atomic force microscopy (AFM); adhesion; friction; nanotribology; contact mechanics.

1. INTRODUCTION

With the rising interest in nanotechnology, new fields of research have emerged in the attempt to understand nanoscale science and engineering. One of these fields is nanotribology, the nanoscale analogue to the macroscale study of friction and wear [1, 2]. At the nanoscale, surface forces become dominant, and surface properties related to adhesion, friction and wear become critical design considerations. Of particular note are micro/nano-electromechanical systems (MEMS/NEMS). Because of the problematic effects of friction and wear [3, 4], there are currently no MEMS devices on the market that involve sliding interfaces in contact. Studies in this area are important not just for nanotechnology applications, but for fundamental science as well. Although many practical, macroscopic engineering applications involving tribology have been successfully addressed, relatively little is understood

*To whom correspondence should be addressed. Tel.: (1-608) 263-4891. Fax: (1-608) 263-7451.
E-mail: carpick@engr.wisc.edu

thus far about the underlying fundamental mechanisms of friction and energy dissipation between sliding surfaces.

There are several experimental approaches to nanotribology, including the use of the atomic force microscope (AFM) [5] and the surface forces apparatus (SFA) [6]. In this paper we focus on work using the AFM to determine nanoscale adhesion and friction between solid surfaces. Our particular focus is on the role that contact mechanics plays in analyzing these measurements.

Over the last 15 years or so, researchers have applied theories of continuum contact mechanics to nanotribology measurements to determine fundamental parameters and processes at play in nanometer-scale contacts. In 1999, Unertl concisely discussed several of these issues [7]. An increasing amount of evidence supports the seemingly surprising conclusion that a continuum description of contact is accurate down to contacts as small as a few nanometers in size. Our primary focus will be on how to properly apply continuum mechanics theories to nanotribology measurements. The important question of where exactly continuum representations break down has yet to be addressed comprehensively. We will discuss this briefly at the conclusion of this paper. We will first briefly review contact mechanics, and then discuss its application to adhesion and friction measurements.

2. REVIEW OF CONTACT MECHANICS

Heinrich Hertz is said to be the pioneer of the field of contact mechanics as we know it today [8]. In 1882, part of his graduate work involved studying interference patterns between glass lenses when pressed together. One relation he determined was that contact between a flat plane and a sphere pressed together with a normal load P occurred within a circular contact area with contact radius a according to the following equation:

$$a = \left(\frac{PR}{K} \right)^{1/3}, \quad (1)$$

where R is the radius of the sphere,

$$K = \frac{4}{3} \left(\frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2} \right)^{-1},$$

E_1 , E_2 are the sphere and flat plane Young's moduli and ν_1 , ν_2 are the sphere and flat-plane Poisson ratios, respectively. This theory describes the contact area for smooth macroscopic contacts, and is used to this day in many applications. It is applicable only to homogeneous, isotropic, linear elastic materials that exhibit no attractive surface forces (adhesion) and assumes that the contact radius is much smaller than the sphere radius, so that the sphere can be approximated as a paraboloid.

In 1971, Johnson, Kendall and Roberts (JKR) [9] proposed a new theory that accounted for adhesion between two elastic bodies. They were motivated by

experimentally measured contact areas that were larger than predicted by the Hertz theory at low loads, and by the observation of finite contact area at zero applied load. They derived the following equation to describe the contact radius between a sphere and a plane with adhesion acting:

$$a = \left(\frac{R}{K} (P + 3\gamma\pi R + \sqrt{6\gamma\pi RP + (3\gamma\pi R)^2}) \right)^{1/3}, \quad (2)$$

where γ is the Dupré energy of adhesion, or work of adhesion. It is, in fact, an energy per unit area and it represents the work done in completely separating a unit area of the interface. Characteristics of this theory include infinite tensile stress at the contact area perimeter, a non-zero contact area at zero load and a minimum stable normal load which still exhibits a non-zero contact area. This minimum load can be called the pull-off force or critical load P_c , and is given by:

$$P_{c(\text{JKR})} = -\frac{3}{2}\pi\gamma R. \quad (3)$$

Shortly thereafter, Derjaguin, Muller and Toporov (DMT) [10] derived a separate expression to include adhesion in the contact of elastic bodies. They assumed that the deformed contact profile remained the same as in the Hertz theory, but with an overall higher load due to adhesion. This is equivalent to attractive interactions acting at all separations between the sphere and the plane, like a dead weight, thus:

$$a = \left(\frac{R}{K} (P + 2\pi\gamma R) \right)^{1/3}, \quad (4)$$

and now the pull-off force is:

$$P_{c(\text{DMT})} = -2\pi\gamma R. \quad (5)$$

In the DMT model, the contact area goes to zero at the pull-off force, and there is no singularity in the contact stresses.

Although these two seemingly competing theories were the cause of much heated debate, the controversy was settled when it was determined that they were valid at opposite ends of the same spectrum of contact behavior. When surface forces are short range compared to the resulting elastic deformations (i.e. compliant materials, large sphere radii, and strong, short-range adhesion forces), the JKR model describes the contact area accurately. The opposite limit (i.e. stiff materials, small sphere radii and weak, long-range adhesion forces) corresponds to the DMT regime. The form of the contact area in the DMT model was, in fact, first presented in the work of Maugis [11] and is often referred to as the ‘Hertz-plus-offset’ model. The development of the theories is provided in a more detailed manner by Maugis in that paper.

It is convenient to utilize a non-dimensional physical parameter to quantify these limits and the cases in between. Often referred to as Tabor’s parameter μ_T , this

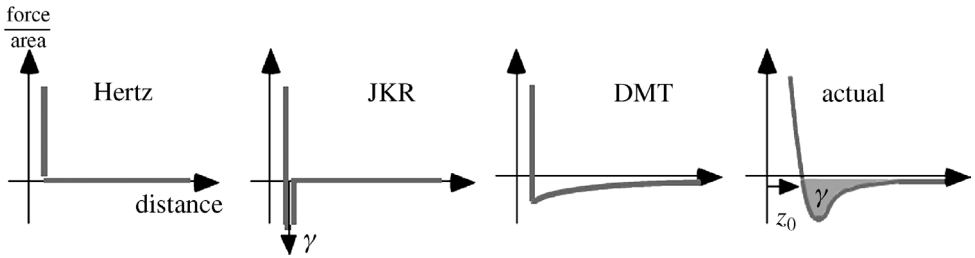


Figure 1. Interaction forces (per unit area) for the Hertz, JKR and DMT models, compared to an actual interaction. There is no attractive force in the Hertz model, only hard wall repulsion at contact. The JKR model includes short-range adhesion that is essentially a delta function with strength γ , and thus only acts within the contact zone. The DMT curve shown represents a long-range surface force. A volume integrated force, like the van der Waals force, can also lead to a DMT dependence, where the contact profile remains Hertzian and the attractive forces act like an additional external load. For an actual interaction force, the integral of the attractive well corresponds to the work of adhesion, γ .

transition parameter is defined as [12]:

$$\mu_T = \left(\frac{16R\gamma^2}{9K^2z_0^3} \right)^{1/3}, \quad (6)$$

where z_0 is the equilibrium separation of the surfaces. The spatial range of the adhesion forces is assumed to scale with z_0 , as in the case of the Lennard–Jones potential, for example, where z_0 is the only length scale in the definition of the potential. **Tabor’s parameter is physically equivalent to the ratio between the normal elastic deformation caused by adhesion (i.e. in the absence of applied load) and the spatial range of the adhesion forces themselves.**

Figure 1 illustrates the different interaction forces (per unit area) as a function of separation for the Hertz, JKR and DMT models, and a more ‘realistic’ interaction like a Lennard–Jones potential.

To approximate such an actual interaction, Maugis [11] elegantly showed that it was possible to connect the JKR and DMT limits and determine contact parameters for the entire range of materials parameters. He considered a ‘Dugdale’ (square-well) potential to describe attractive forces between contacting spheres, as is illustrated in Fig. 2. In this model, a constant adhesional stress σ_0 acts over a range δ_t . Thus, the work of adhesion is $\gamma = \sigma_0 \cdot \delta_t$. Maugis defines a parameter, λ , which is similar to μ_T , given by:

$$\lambda = 2\sigma_0 \left(\frac{R}{\pi\gamma K^2} \right)^{1/3}. \quad (7)$$

By choosing σ_0 to match the minimum adhesional stress for a Lennard–Jones potential (with equilibrium separation z_0), it follows that $\delta_t = 0.97z_0$, and so $\lambda = 1.1570\mu_T$. Thus, λ and μ_T are nearly equivalent. For convenience we shall refer to λ as the ‘transition parameter’. If $\lambda > 5$, the JKR model applies and if $\lambda < 0.1$ the DMT model applies. Values between 0.1 and 5 correspond to the ‘transition

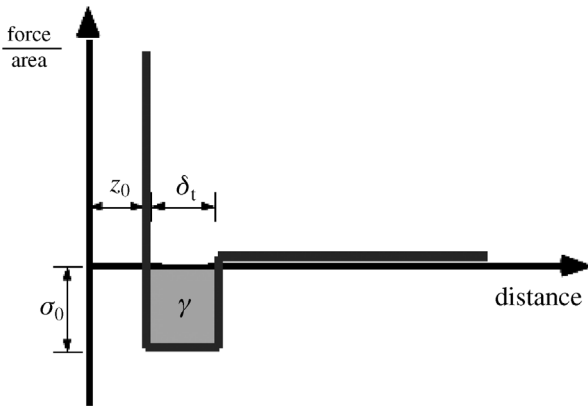


Figure 2. The force (per unit area)–distance relation for the Dugdale model used by Maugis. A constant adhesional stress (force per unit area) σ_0 acts between the surfaces over a range δ_t . At greater separations, the attractive force is zero. The work of adhesion is thus $\gamma = \sigma_0\delta_t$.

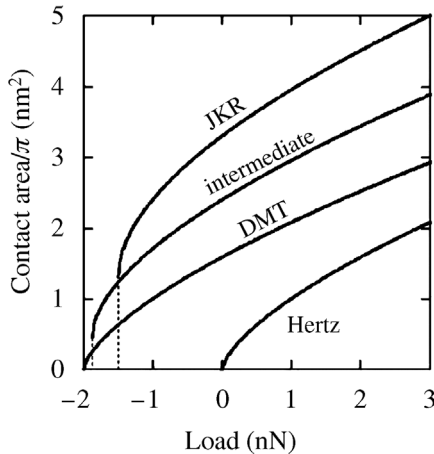


Figure 3. The Hertz area–load curve and the JKR–DMT transition, plotted by setting $K = 1$ GPa, $R = 1$ nm and $\pi\gamma = 1$ J/m². Area–load curves for the JKR limit, the DMT limit and an intermediate case are shown. These approach the Hertz curve in the limit $\gamma \rightarrow 0$ (no adhesion). Adhesion increases the contact area from the Hertz case for a given load by an amount dependent upon the range of attractive forces.

regime’ between JKR and DMT models. A summary of different conventions used for defining this ‘transition parameter’ is provided by Greenwood [12]. When there are no attractive surface forces ($\gamma = 0$), the Hertz theory applies. The variation of contact area with load in these cases is shown in Fig. 3.

The assumption of a square-well potential may seem somewhat limiting. However, Barthel has shown that the contact behavior is generally insensitive to the actual shape of the interaction potential, provided there is only one length scale involved [13]. However, the behavior can vary significantly if an additional length scale to the interaction is introduced [14]. It may be appropriate to describe, for ex-

ample, the simultaneous effect of short-range solid–solid adhesion and longer range electrostatic, van der Waals', or solvation forces.

The Maugis–Dugdale equations are somewhat difficult to utilize, since Maugis' formulation lacks a single expression relating only a and P . To plot the Maugis–Dugdale solution or fit it to data, one needs to simultaneously solve two equations by letting certain parameters vary between appropriate limits that depend upon λ . Furthermore, the relation for the pull-off force must be determined through iteration [15] if the value of λ is not known *a priori* (the usual case with experimental measurements). In practice, this is rather cumbersome to carry out with common software programs that utilize automated statistical fitting procedures.

In 1999, Carpick, Ogletree and Salmeron (COS) [16] provided an approximate general equation for easily describing the contact area. The general equation is amenable to conventional curve fitting software routines and provides a rapid method of determining the value of the transition parameter which describes the range of surface forces. They showed that the Maugis formulation could be approximated using the following formula to determine the contact radius:

$$a = a_{0(\alpha)} \left(\frac{\alpha + \sqrt{1 - P/P_c(\alpha)}}{1 + \alpha} \right)^{2/3}, \quad (8)$$

where α is another transition parameter and a_0 is the contact area at zero load. Note that $\alpha = 1$ corresponds exactly to the JKR case, and $\alpha = 0$ corresponds exactly to the DMT case. Included in equation (8) is the fact that P_c and a_0 depend on α as well. Equation (8) is referred to as the generalized transition equation or COS equation. For intermediate cases ($0 < \alpha < 1$), the generalized transition equation corresponds very closely to Maugis' solution for the transition regime ($0.1 < \lambda < 5$). An approximate relation is used to relate α and λ :

$$\lambda = -0.924 \cdot \ln(1 - 1.02\alpha). \quad (9)$$

They also outlined a step-by-step procedure to fit this equation to determine the regime of contact for AFM or SFA experiments.

In an analogous fashion, Piétrement and Troyon subsequently presented approximate general equations for the elastic indentation depth and the normal contact stiffness [17] that are also easier to handle than the corresponding Maugis–Dugdale equations. Schwarz then showed that equation (8) could be derived from physically-based arguments by combining an infinitely short-range attractive interaction (essentially a JKR-type interaction) with a long-range force [18] of adjustable relative weight. This further supports the validity of the simple, approximate equations presented in Refs [16, 17].

All models described above assume homogeneous, isotropic, linear elastic contact between a sphere and a plane (or, equivalently, two spheres) where the contact radius is much smaller than the sphere radius (allowing it to be described mathematically as a paraboloid), and the loading is purely in the normal direction. In practice, there will be deviations from these assumptions that may be significant. Fortunately,

solutions for many of these cases have been worked out. Maugis worked out the adhesional contact problem for large contact radii (i.e. using spheres, not paraboloids) [19]. The solution for adhesional contact for an axisymmetric power-law indenter shape was presented by Carpick *et al.* [20, 21]. Johnson has discussed the effect that lateral forces may have on the contact [22]. Viscoelastic effects have been discussed in a number of cases [23–28] and were reviewed recently by Shull [29]. Plastic effects, while often dealt with through finite element modeling, have also been modeled analytically, including the effects of phase transitions [30] and adhesion [31]. However, the effects of anisotropy have been discussed in considerably less detail, see for example Refs [32, 33].

Some particularly interesting work has been performed in applying contact mechanics to the surface forces apparatus (SFA), which can be viewed as a single asperity contact of larger scale than the AFM, and with important capabilities including the direct measurement of the contact area [34, 35]. The SFA typically consists of thin mica layers on a compliant (glue) backing. The consequences of this geometry have been studied by Johnson and Sridhar [36]. They find that the JKR model works well for a sufficiently thick layer, as well as for a sufficiently thin layer, but layers of an intermediate thickness violate the conditions of the JKR theory. In between these limits the contact behavior is modeled by the numerical results. Violations of the JKR model were found to be due to the small size of the contact radius compared to the layer thickness.

Barthel and Roux have studied the velocity dependence of the work of adhesion when studying organic monolayers with the SFA [37]. They find that lower velocities produce consistent pull-off force values and, at higher velocities, viscoelastic effects give rise to higher pull-off force values. They derive analytical results by considering viscous dissipation, and they compare the results with numerical treatments. Stability of the system is significantly dependent on the chemical composition of the organic layer, and they highlight the importance of formulating generalized contact mechanics to intermediate ranges of interaction (i.e. between DMT and JKR regimes).

In summary, contact mechanics models can be powerful tools in extracting fundamental parameters from studies of surfaces in contact. However, they must be used cautiously, with attention paid to the appropriate limits and assumptions they entail. In what follows, we discuss how measurements of adhesion and friction with the AFM can be treated using contact mechanics approaches under conditions (as specified above) appropriate for behavior in the JKR-to-DMT range.

3. ADHESION MEASUREMENTS

Fundamental mechanisms of adhesion at the micro/nanoscale are a topic of substantial current interest. Much of this work is motivated by the ‘stiction’ problems seen in MEMS/NEMS devices [3, 38] and it is postulated that a fundamental understanding of adhesion at the nanoscale may give rise to guidelines for designing

and fabricating reliable NEMS/MEMS devices. In this regard, one can consider single-asperity adhesion to be at the smallest possible scale and, therefore, understanding single-asperity adhesion may be the key to understanding the true work of adhesion between two complex, multi-asperity surfaces at larger scales. The AFM is an ideal tool to quantitatively characterize the true work of adhesion at this length scale. Contact mechanics models are often used to analyze data, and considerable efforts have been made to enhance these models based on understanding how surface interactions can deviate from the ideal cases treated by DMT or JKR theories.

3.1. AFM methodology

To use the AFM as a quantitative tool to investigate interfacial adhesion, cantilever calibration procedures and contact mechanics modeling need to be applied consistently. Adhesion data taken with the AFM can vary significantly due to seemingly subtle changes in the environmental conditions, AFM tip material and/or geometry, and surface contamination and/or modification. To ensure that calculated values for work of adhesion are intrinsic to the interface tested, rigorous characterization techniques of the tip and surface materials should be performed.

Typically, the pull-off forces are measured by AFM in the following way. A laser beam is focused onto the backside of the cantilever and reflected onto a position-sensitive photodetector. The cantilever is deflected by interaction with the surface material, which moves the reflected laser spot vertically on the photodetector. This motion is measured by the photodetector and is recorded in a so-called ‘force–distance’ plot. This plot is then analyzed to determine the magnitude of the pull-off force, which is subsequently converted into a value for the work of adhesion. The process of obtaining the data needed to calculate the work of adhesion generally involves three stages: acquiring reliable force–distance curves, calibrating force constants to obtain pull-off force values and compiling the necessary contact mechanics parameters.

3.2. Acquiring reliable force–distance curves

The operation of an AFM is fairly straightforward, which is one reason why adhesion data are often reported without appreciation for the afore-mentioned subtleties. Assuming that the environmental conditions (temperature, relative humidity, atmospheric composition) are kept constant, several issues that often arise can prevent reliable, reproducible data from being collected. First, if the tip geometry is altered by contacting the surface (e.g. due to wear of the tip), pull-off force values can vary significantly. There are two primary techniques for measuring the shape of an AFM tip. One technique involves a tip reconstruction by analyzing AFM images with computer software [39, 40]. The other involves imaging the tip in a high-resolution scanning electron microscope (SEM) or transmission electron microscope (TEM). It is recommended that the AFM tips be analyzed both before and after scanning to check for tip modification.

Second, pull-off forces taken on a homogenous surface can vary if the surface chemical composition of the tip or sample is modified during the experiment. Even if geometrical changes of the tip do not occur, it is still possible that chemical changes may, due to surface contamination, material transfer, or modification involving surface treatments or coatings that are present. Contamination can be addressed by improving sample preparation procedures or conducting experiments in a clean, dry environment including ultrahigh vacuum. If the chemistry taking place is due to the presence of a surface coating, a more in-depth analysis of the surface interactions and chemical exchanges needs to be applied. Unfortunately, there is no technique for perfectly characterizing the chemical state of the tip, although measurements of contact conductance can at least be used to reveal changes that have occurred [41].

Third, when multiple pull-off measurements are taken at different locations on a heterogeneous surface, variations in both the surface roughness and surface chemistry can affect the interaction forces [42, 43]. If the surface topography and composition vary, the contact between the tip and sample changes based on local slope, curvature, elasticity, and adhesion. ‘Force-volume’ imaging is often performed, which involves acquiring a topography image and a pull-off force map simultaneously, to correlate pull-off force values with specific regions of a sample. In this way, high surface energy features can be observed (e.g. step edges, defects, other inhomogeneities). As well, regions of differing chemistry can be identified.

3.3. *Extracting the pull-off force values from force–distance curves*

When it has been determined that the tip and sample are not being altered, the force–distance plots can then be analyzed to obtain quantitative pull-off force values using a cantilever calibration procedure. There are several methods in the literature, but two procedures have emerged as the most reliable. One method is Cleveland’s ‘added mass method’ [44] and the other is Sader’s ‘unloaded resonance technique’ [45]. For the added mass method, a known mass is attached to a cantilever, the resulting change in resonant frequency is measured, and this frequency shift is correlated with the normal stiffness of the cantilever. For the Sader method, the length, width, resonant frequency and quality factor of a cantilever are measured in ambient conditions. These measured properties are related to the cantilever’s normal stiffness through the hydrodynamic function due to the air damping of the cantilever. The Sader method is straightforward to apply *in situ* and is non-destructive. Therefore, this method is preferable for extracting calibrated pull-off force values from uncalibrated force–distance plots.

3.4. *Compiling the necessary parameters for contact mechanics modeling*

Before work of adhesion values can be calculated from pull-off force data, several parameters must be either measured or determined from literature values. As discussed above, the contact between an AFM tip and the sample can range from

DMT-like to JKR-like, depending on the elastic properties (Young's modulus and Poisson's ratio) of the tip and sample, the tip radius, the work of adhesion and the equilibrium spacing (or range of attractive forces) between the tip and sample when in contact. The elastic properties can be determined from nanoindentation experiments or from literature values. However, the equilibrium spacing is not directly known. To handle this, upper and lower bounds can be used initially as described below.

3.5. Fitting adhesion data based on contact mechanics modeling

As mentioned above, the work of adhesion calculation depends upon the appropriate contact regime, which ranges from the JKR to the DMT regime. Equations (3) and (5) for the pull-off forces can be rearranged and then generalized as:

$$\gamma = \frac{-P_c}{\chi \pi R}, \quad (10)$$

where χ ranges monotonically from 1.5 (for the JKR limit) to 2 (for the DMT limit), depending on Tabor's parameter, μ_T . This full range of uncertainty in χ corresponds to a possible error in γ of 33%. A plot, table and approximate formulae for χ as a function of μ_T are given in Ref. [16]. Still, μ_T is not known without knowledge of z_0 and γ . However, γ is not known without knowledge of μ_T . This is the adhesion 'Catch-22'.

3.6. Experimental example

One way to get around this 'Catch-22' is to first check if upper (lower) bounds can be placed on μ_T by assuming the smallest (largest) reasonable values of z_0 and any other uncertain parameters in the experiment. This is best illustrated with an experimental example. An experiment was conducted using commercial AFM tips fabricated from silicon and coated with a partially oxidized tungsten carbide film. This presents an additional problem, since the elastic properties of the tungsten carbide are uncertain. Note that while the elastic constants do not enter into the calculation of the adhesion energy using equations (3) and (5), they do affect the value of χ . We measured the pull-off force between these tips and a single crystal diamond sample (Diamond Innovations) with a (111) orientation, terminated with hydrogen by an RF microwave plasma treatment. X-ray photoelectron spectroscopy and near-edge X-ray absorption fine structure spectroscopy were used to verify the expected chemical composition and structure of the diamond surface. AFM measurements were carried out with two different cantilevers to check for reproducibility. We used the Sader method to calibrate each cantilever. The tip shape and radius were measured *in situ* by the inverse imaging method using a TipCheck calibration standard (Aurora Nanodevices) and deconvolution software. We reduced the possibility of modification of the AFM tip by taking pull-off force measurements directly without scanning the sample, and

Table 1.

Tip radii, normal force constants and average pull-off forces measured on different samples. Measurements taken using two separate AFM cantilevers are reported

Cantilever	R (nm)	k (N/m)	$P_c \pm \sigma$ (nN)
#1	25 ± 5	0.108 ± 0.022	6.2 ± 1.1
#2	30 ± 6	0.146 ± 0.029	9.2 ± 5.0

Errors are the standard deviations from each set of measurements.

took at least 35 measurements per tip. In another work (data not shown), we have successfully used a TEM to characterize AFM tips from the same wafer under the same experimental conditions to verify that the tips did remain unchanged while taking such measurements. Cantilever parameters and calculated values of the pull-off force for the two levers are presented in Table 1.

To determine the appropriate contact regime for these experiments, we used extreme (i.e. low) values for the elastic constants and equilibrium spacing (z_0) to set an upper bound on Tabor's parameter. The test for an upper bound was chosen because it was likely that Tabor's parameter would be low given the stiffness and expected low interfacial energy of diamond and tungsten carbide. Our lower-bound 'test' value of z_0 was chosen to be 0.154 nm, which is the C–C bond distance in diamond. We used a conservatively low modulus for the tip of 357 GPa (50% of the value of pure single-crystal tungsten carbide), and used the smallest possible value of $\chi = 1.5$. We used literature values for the modulus and Poisson's ratio of diamond. Poisson's ratio for tungsten carbide was also taken from the literature, which is a safe choice, since the dependence of Tabor's parameter on Poisson's ratio is rather weak. As stated above, the only other parameters in the problem, the pull-off force and the tip radius, are experimentally determined. Choosing the cases that exhibited the largest values of P_c , we determined 'test' adhesion energies from equation (10). These are then used in equation (6), where we find that even using the most extreme assumptions, Tabor's parameter does not exceed approx. 0.09. Thus, we are firmly in the DMT regime, i.e. $\chi = 2$, and the correct work of adhesion can now be directly determined from the pull-off force using equation (10). All values used in calculating μ_T and the resulting work of adhesion values are given in Table 2. A significant error is associated with the measurement from cantilever #2. Despite the care taken with the measurement, there still exists some heterogeneity on the sample, and the possibility of tip changes may have also contributed. Nevertheless, the two values for the work of adhesion are comparable.

In the case of an interface expected to be in the JKR regime, an analogous procedure is used: assume $\chi = 2$, use the largest reasonable values for any uncertain elastic moduli and choose an upper bound for z_0 . This upper bound for z_0 could be extracted, for example, by examining the force-distance curve. z_0 can be assumed to be no greater than the snap-in distance of the tip, which would be calculated by

Table 2.

Values used for determining the upper bound to Tabor's parameter and the subsequent work of adhesion values assuming the DMT model holds

Cantilever	$P_c \pm \sigma$ (nN)	χ	γ (test) (J/m ²)	E_{tip} (GPa)	ν_{tip}	E_{sample} (GPa)	ν_{sample}	K (GPa)	z_0 (nm)	μ_T	γ (actual, DMT) (J/m ²)
#1	6.2 ± 1.1		0.053 ± 0.020							0.079	0.0397 ± 0.0072
#2	9.2 ± 5.0	1.5	0.065 ± 0.048	357	0.24	967	0.08	364	0.154	0.089	0.049 ± 0.027

converting the change in the normal load signal at snap-in to the distance the tip has moved. If one finds that μ_T , even under these extreme assumptions, exceeds approx. 5, then one can safely assume that JKR limit applies.

Intermediate cases remain problematic without further information. At the very least, the methodology outlined above allows one to set upper and lower bounds for possible values of the work of adhesion. Also, as we show in the next section, if additional information about the contact area can be determined, such as from friction measurements, then Tabor's parameter can be determined.

The general methodology adopted here is well-suited for many tribological interfaces studied. This approach to characterization provides insight into why results can vary significantly between the DMT and JKR regimes, and offers guidelines on how to embody this understanding when analyzing AFM pull-off force data. If true work of adhesion values are to be trusted and are comparable between systems, experiments, and laboratories, this type of methodology needs to be applied consistently to AFM data and reported accordingly in publications and presentations.

4. FRICTION MEASUREMENTS

As with adhesion measurements, friction measurements need to be carefully calibrated and analyzed, taking into account the necessary contact mechanics and the assumptions that go with them. The tip shape and the cantilever's normal force constant must be calibrated properly, as described above. In addition, the lever must be calibrated to quantitatively describe lateral forces. Frictional properties are also quite sensitive to tip and environmental conditions. The process for friction data takes a similar form as adhesion measurements: acquire reliable and reproducible friction data, calibrate the AFM lever to obtain quantitative values for normal load and friction during an experiment, and use known characteristics of the lever and sample along with contact mechanics fits, taking into account the JKR–DMT transition, to determine quantitatively the interfacial shear strength and adhesion energy.

An AFM cantilever can be scanned across a surface so that the direction of motion is perpendicular to the long axis of the cantilever. In this way the lever twists due to the frictional resistance of the tip and the laser spot reflected from the back of the cantilever moves laterally on the photodetector. Friction is determined by taking half the difference of the forward and backward scans along a given line on the surface. In topographic feedback control, the cantilever is scanned at a constant applied load. Alternately, the load can be varied in a specific way while scanning the tip across the sample as described below. We refer to this type of experiment as determining friction as a function of load, or more simply ‘friction *versus* load’.

Friction *versus* load curves can be obtained by the following procedure. First, the AFM control electronics must be able to vary the applied load on the cantilever while simultaneously monitoring and recording the normal and lateral signals while scanning across the surface. This can be accomplished in a number of different ways. Some software and controller packages, such as RHK’s SPM 100 controller and SPM 32 software (RHK Technology), allow the applied load to be automatically ramped in a stepwise fashion so that an image can be acquired with every scan line obtained at a different load. Some other AFM controller/software packages do not allow the user sufficient flexibility to complete such an experiment. In this case, the user may apply an external signal to the AFM itself. One way is to supply a voltage signal directly to the z -piezo of the AFM, thus altering the load on the cantilever. This may be accomplished by either modifying the electronics of the AFM controller, or by using an external connection, such as a breakout box. The downside to applying a signal directly to the z -piezo is that the feedback system needs to be turned off, so no line-by-line feedback can be maintained. A better method is to send a signal to the AFM that changes the feedback setpoint value, i.e. the applied normal load in feedback control. An external signal is sent directly into the AFM electronics to vary the setpoint, again either by modification of the AFM electronics or using a breakout box. This method allows for feedback control while the load is varied during scanning, which is beneficial for measurements on rough or sloped surfaces. Load and friction are continuously measured by monitoring the position of the laser spot on the photodetector in the normal and lateral directions, respectively. Usually it is easier to apply a continuously varying voltage, as opposed to the stepwise voltage ramp as in the RHK system for example. However, instead of each line being a separate load value, the load is continuously varied even during a single scan line. To account for this, the load and friction variations on a single line are averaged (excluding the edges of the image where the tip is sticking, and not sliding, with respect to the surface) to obtain a single data point of average load and friction for every scan line. This is reasonable to do if the variation in voltage (and, thus, force) during a single scan line is small compared to the range of voltage (force) applied during the experiment.

4.1. Acquiring reliable friction versus load curves

To monitor the reproducibility of the data, the applied load may be started at a high value, ramped down until pull-off occurs and then increased back to the starting load. This way, the same normal applied load is measured twice within a short span of time, for both increasing and decreasing load (except for most of the negative load regime, which is only accessed during the decreasing load ramp). If the data overlap in these two cases, the data have been successfully reproduced. This can be checked by repeating the friction measurements at different locations on the sample surface to check for any variation. Note that for viscoelastic materials, hysteresis between loading and unloading portions may occur and is, in fact, a useful probe for the viscoelastic properties of the sample [23].

For consistent friction results one should obtain friction versus load data on a relatively small area of the sample to avoid topographic and chemical inhomogeneities, steps, grain boundaries, etc. However, to ensure that the AFM tip overcomes the static friction force and actually slides across the surface, the width of the line being scanned should be sufficiently large, which is easily verified by confirming that the friction loops contain a reasonable sliding portion. When analyzing the friction data, only the portion of the friction traces that corresponds to this steady sliding situation should be considered, thereby excluding the static friction regimes. In addition, it may be beneficial to scan the same line for the entire duration of the experiment, instead of rastering a square region on the surface. This allows for the sampling of the same region of the surface for all of the scan lines, except for a slight deviation from the region due to thermal drift and any in-plane translation of the tip due to the bending of the tilted lever. These effects can be compensated if desired [46].

4.2. Relation between contact area and friction

Several studies have shown that the friction force in solid-solid nanocontacts below the wear threshold is proportional to the true contact area, i.e. number of interfacial atoms [1, 20, 47–50]. This was demonstrated by measuring both friction and the nano-scale contact area experimentally. In other words, friction force F_f for a single-asperity contact is given by:

$$F_f = \tau \cdot A, \quad (11)$$

where A is the interfacial contact area and τ is the interfacial shear strength. As it represents the friction force per unit area of a pair of materials, one can also consider τ as corresponding to the intrinsic frictional dissipation per interfacial atom.

The interfacial shear strength is not necessarily constant, and may, as a first attempt, be more generally described as a constant plus a pressure-dependent term:

$$\tau = \tau_0 + \zeta \cdot p, \quad (12)$$

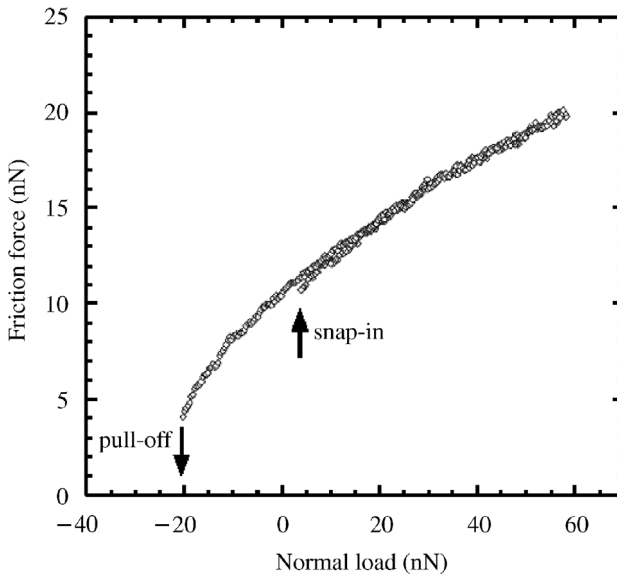
where τ_0 is the constant interfacial shear strength term, p is the nominal contact pressure and ζ is a dimensionless coefficient. The magnitude and pressure

dependence of the interfacial shear strength will depend on the materials and the sliding conditions (environment and temperature), and determining this behavior is a key goal in nanotribology, as is the determination of the general validity of equation (11), which may break down for sufficiently small contacts or under different sliding conditions (e.g. when wear takes place). Equation (12) has been shown to apply very well to systems involving molecular films in contact in the SFA [51], for some soft solids in macroscopic contact [52] and for some systems measured with the AFM [53]. For bare solid surfaces, the majority of studies report that ζ is negligible [1, 15, 20, 22, 47–50, 53–60], but the range of materials for which this has been tested is limited. In any event, the determination of the contact area and how it relates to friction is a key challenge, and once again the JKR-to-DMT transition must be taken into account to properly describe the contact mechanics.

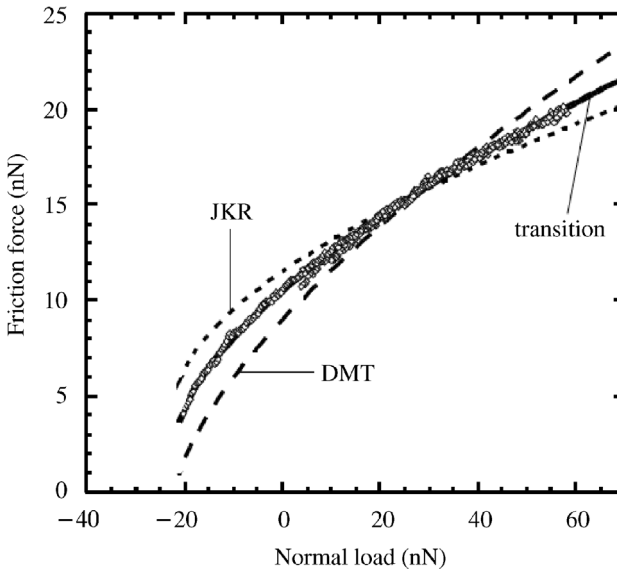
4.3. Fitting friction versus load data to determine adhesion energy and interfacial shear strength

Although Carpick *et al.* outlined a procedure for applying equation (11) to a given set of contact area or friction data [16], the technique of determining a correct fit to data requires experience and knowledge of the characteristics of the different contact regimes. We have already discussed that the JKR and DMT limits represent two ends of the spectrum, but how does one determine from a given set of data whether it would obey the JKR model, the DMT model, or something in between, or neither? Figure 3 shows the contact area *versus* load relations for the various models, which again, we emphasize, are only truly valid under several assumptions including homogeneous, isotropic, linear elastic materials. One distinction of the JKR model is that the Hertz and DMT models predict zero contact area at pull-off, while the JKR, as well as the intermediate models, have a finite contact area at pull-off. This gives the experimentalist an easy way to make a qualitative estimate regarding the contact regime of the interface, and highlights the importance of acquiring a dense set of data in the adhesive regime (where the loads are negative).

As an example, a set of friction versus load data from an AFM experiment is shown in Fig. 4a. This set of data was taken using a Digital Instruments MultiMode AFM (Nanoscope IV controller) on a silicon substrate coated with a thin organic hydrocarbon monolayer film, using a silicon tip with native oxide termination. To be clear, the organic film will alter the contact mechanics substantially since the system is no longer homogeneous axially, it may be anisotropic, and the organic film may be somewhat non-linear and anelastic. This renders our example a case where no model in the JKR–DMT spectrum ought to apply quantitatively. However, for illustration purposes only, we carry out JKR–DMT analysis, assuming that the measured data are obtained for a homogeneous, isotropic, linear elastic system that obeys all the assumptions of the JKR–DMT models. A more complete and detailed analysis that goes beyond these assumptions will be presented elsewhere (data not shown).



(a)



(b)

Figure 4. (a) An example set of AFM friction force *versus* load data, taken on a silicon substrate coated with a thin organic film, using a silicon tip with native oxide termination. This set of data is comprised of two overlapping traces, ramping from high load to low load, and then low load to high load. The instability points of pull-off and snap-in are shown on the graph. (b) Fit of data from (a) to the DMT, JKR and generalized transition equations, setting the same pull-off force for all fits. The two free parameters are the interfacial shear strength τ and the transition parameter α . (c) Fit of data from (a) to the DMT, JKR and generalized transition equations, where the pull-off force is a free parameter in addition to the interfacial shear strength τ and the transition parameter α .

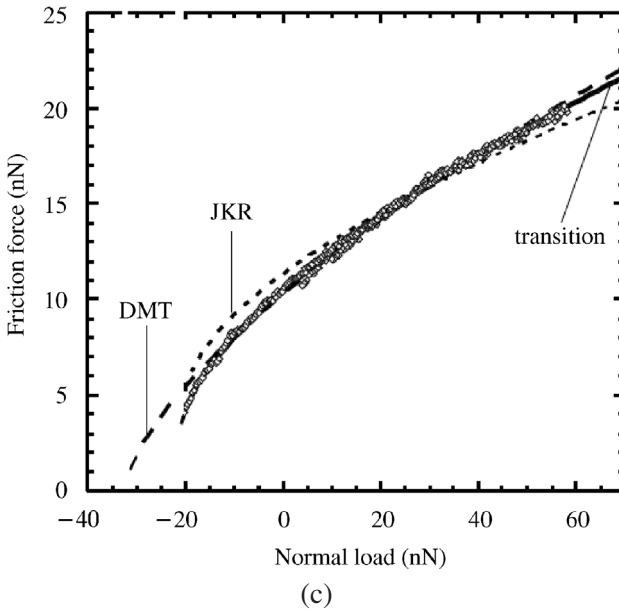


Figure 4. (Continued).

The data were acquired according to the procedure outlined above, where a ramped external voltage was applied to AFM electronics to continuously vary the normal force setpoint. There is minimal scatter in the data and no hysteresis other than a slight difference at the snap-in point.

Note that the normal load axis is defined such that zero corresponds to zero externally applied load. That is, a normal load of zero corresponds to the deflection of the lever when well out of contact with the sample, and no load is acting. This is the correct way to plot normal loads. The load axis then simply represents whether the lever is applying a tensile or compressive load to the sample. In some cases in the literature, researchers shift the normal load axis so that the pull-off force is at zero. This unfortunate practice is inappropriate and misleading, because it shifts the load axis by different amounts for different interfaces, and the unique amount of adhesion for each measurement cannot be determined from the graph as plotted.

The example data set has been fitted to the DMT, JKR, and generalized transition equation, as shown in Fig. 4b. The fit for the generalized transition equation is determined by letting the pull-off force P_c , friction at zero load F_{f0} , and transition parameter, α , all be free parameters in a least squares minimization. The optimization was performed using KaleidaGraph (Synergy Software). The DMT and JKR fits were obtained by setting the pull-off force to be equal to that found from the transition fit. Equivalently, one could set this force equal to the pull-off force obtained from force–distance measurements, if one is confident that the adhesion measurements are consistent and uniform. Thus, there is only one free parameter, the friction force at zero load, for each of these two fits. Clearly, the JKR

and DMT fits are not appropriate. The transition fit is the most convincing, and the pull-off force from this fit is very close to that obtained directly from force-distance measurements.

The JKR and DMT fits can be improved somewhat by allowing the pull-off force to be a second free parameter. These fits are shown in Fig. 4c. The JKR fit is not substantially changed, but the DMT fit now matches the data better. However, this is at the expense of requiring a much larger adhesion force, whereby a portion of data at low loads appears to be ‘missing’. One could attribute the apparent ‘missing’ data to premature pull-off during friction scanning. However, as stated above, the pull-off force from the DMT fit does not match those from force-distance measurements. The transition fit, therefore, remains the most convincing.

The values from the transition fit are $P_c = 22.3$ nN, $F_{T0} = 10.4$ nN and $\alpha = 0.42$. From equation (6) and the relation $\lambda = 1.1570\mu_T$, this corresponds to $\mu_T = 0.45$, which indeed is in between the JKR and DMT limits.

Now that we know the value for μ_T , the pull-off force and friction force at zero load can be converted into values for the work of adhesion and the interfacial shear strength by using the formulae in Ref. [16]. In this example, $\gamma = 119$ mJ/m² and $\tau = 550$ MPa. A pure two-parameter JKR fit would give $\gamma = 123$ mJ/m² and $\tau = 436$ MPa, and a pure two-parameter DMT fit would give $\gamma = 148$ mJ/m² and $\tau = 706$ MPa. Clearly, using the appropriate fit within the JKR–DMT spectrum makes a substantial difference, particularly for the interfacial shear strength.

The data can also be fitted by allowing ζ to be non-zero. The fits are less accurate in this case, suggesting that ζ is indeed zero for this system. However, the number of free parameters, four, is substantial. It, therefore, becomes desirable to measure the contact area independently of the friction force. This is done by measuring either the lateral contact stiffness [15, 48, 53, 57, 61] or the contact conductance [55, 62] as a function of load. Both of these quantities are related to the contact area, and analogous fitting procedures can be carried out to determine the transition parameter. These are important approaches that we strongly endorse, but do not discuss in detail in this publication.

5. FUTURE DIRECTIONS

It is remarkable that continuum mechanics is able to provide such a reasonable description of the contact properties at the nanometer scale. That is not to say that continuum mechanics tells the whole story; the system is, of course, made of atoms, and atomic-scale behavior like stick-slip friction is readily observed in many systems. Furthermore, recent work has indicated significant deviations from continuum behavior. Socoliuc *et al.*, for example, have found that the lateral contact stiffness at the lowest accessible loads for AFM tips with radii nominally below 15 nm cannot be accounted for by a continuum approach [63]. Specifically, the contact stiffness is load-independent, suggesting that a fixed number of atoms are interacting across the interface, and that this structure is not changing with the load.

Cha *et al.* have simulated single-asperity contact during loading and unloading using molecular dynamics [64]. The simulations and experiments both show that elastic contact mechanics should not be applied to describe asperity loading because the loading process involves generation and motion of dislocations. However, they find that the JKR model is applicable for unloading, where the deformation is almost completely elastic [64]. Luan and Robbins [65] have recently shown in simulations that the JKR model breaks down for sufficiently small contacts. The stress distribution no longer smoothly varies with location, and the contact area becomes difficult to determine. Wenning and Müser [66] have shown in simulations that for small contacts, atomic-scale contamination and lattice commensurability are determining factors, and the dependence of friction upon load is not necessarily monotonic nor strictly determined by the contact area.

These examples suggest that the description of the contact area, if it can even be defined in the atomistic limit, now poses the next set of challenges for researchers in nanotribology.

6. CONCLUSIONS

We have discussed the development of contact mechanics, from Hertz to the present day. In the case of homogeneous, isotropic, linear elastic normal adhesional contact between a sphere and a plane, the spectrum of interactions between the JKR and DMT limits can be applied to describe the contact area. We have shown how adhesion and friction measurements by AFM can be understood in light of the transition. We have also given examples of how the fits are performed, and how the fits significantly affect the calculations of fundamental interfacial parameters, namely the interfacial shear strength and the work of adhesion. Future challenges could encompass more complex contact mechanics situations, such as anisotropy, plasticity, viscoelasticity, inhomogeneity, the presence of thin films, and the breakdown of the continuum description of contact at the atomic scale.

Acknowledgements

We thank Rachel J. Cannara and Dr. Anirudha V. Sumant for many fruitful interactions related to this work. We acknowledge Dr. William Robert Ashurst, Dr. Maarten P. de Boer and Dr. Alex D. Corwin for the organic-film-coated silicon sample. We gratefully acknowledge Prof. R. J. Hamers and Wensha Yang for use and assistance with the hydrogen plasma system. EEF acknowledges support from the National Science Foundation for a Graduate Research Fellowship. RWC acknowledges support from the National Science Foundation CAREER Program, grant #CMS-0134571, and from the Department of Energy, grant #DE-FG02-02ER46016.

REFERENCES

1. R. W. Carpick and M. Salmeron, *Chem. Rev.* **97**, 1163–1194 (1997).
2. I. L. Singer and H. M. Pollock (Eds), *Fundamentals of Friction: Macroscopic and Microscopic Processes*. Kluwer, Dordrecht (1992).
3. R. Maboudian and R. T. Howe, *J. Vac. Sci. Technol. B* **15**, 1–20 (1997).
4. R. Maboudian, W. R. Ashurst and C. Carraro, *Tribol. Lett.* **12**, 95–100 (2002).
5. R. W. Carpick, PhD Thesis, Department of Physics, University of California at Berkeley, Berkeley, CA (1997).
6. A. M. Homola, J. N. Israelachvili, P. M. McGuiggan and M. L. Gee, *Wear* **136**, 65–83 (1990).
7. W. N. Unertl, *J. Vac. Sci. Technol. A* **17**, 1779–1786 (1999).
8. K. L. Johnson, *Contact Mechanics*. Cambridge University Press, Cambridge (1987).
9. K. L. Johnson, K. Kendall and A. D. Roberts, *Proc. R. Soc. Lond. A* **324**, 301–313 (1971).
10. B. V. Derjaguin, V. M. Muller and Yu. P. Toporov, *J. Colloid Interf. Sci.* **53**, 314–326 (1975).
11. D. Maugis, *J. Colloid Interf. Sci.* **150**, 243–269 (1992).
12. J. A. Greenwood, *Proc. R. Soc. Lond. A* **453**, 1277–1297 (1997).
13. E. Barthel, *J. Colloid Interf. Sci.* **200**, 7–18 (1998).
14. E. Barthel, *Colloid. Surf. A* **149**, 99–105 (1999).
15. M. A. Lantz, S. J. O’Shea, M. E. Welland and K. L. Johnson, *Phys. Rev. B* **55**, 10776–10785 (1997).
16. R. W. Carpick, D. F. Ogletree and M. Salmeron, *J. Colloid Interf. Sci.* **211**, 395–400 (1999).
17. O. Piétrement and M. Troyon, *J. Colloid Interf. Sci.* **226**, 166–171 (2000).
18. U. D. Schwarz, *J. Colloid Interf. Sci.* **261**, 99–106 (2003).
19. D. Maugis, *Langmuir* **11**, 679–682 (1995).
20. R. W. Carpick, N. Agraït, D. F. Ogletree and M. Salmeron, *J. Vac. Sci. Technol. B* **14**, 1289–1295 (1996).
21. R. W. Carpick, N. Agraït, D. F. Ogletree and M. Salmeron, *J. Vac. Sci. Technol. B* **14**, 2772 (1996).
22. K. L. Johnson, *Proc. R. Soc. Lond. A* **453**, 163–179 (1997).
23. K. J. Wahl, S. V. Stepanowski and W. N. Unertl, *Tribol. Lett.* **5**, 103–107 (1998).
24. D. Maugis and M. Barquins, *J. Phys. D* **11**, 1989–2023 (1978).
25. V. S. Mangipudi and M. Tirrell, *Rubber Chem. Technol.* **71**, 407–448 (1998).
26. W. N. Unertl, *J. Adhesion* **74**, 195–226 (2000).
27. M. Giri, D. B. Bousfield and W. N. Unertl, *Langmuir* **17**, 2973–2981 (2001).
28. E. Barthel, G. Haiat and M. C. Phan Huy, *J. Mech. Phys. Solids* **51**, 69–99 (2003).
29. K. R. Shull, *Mater. Sci. Eng. R* **36**, 1–45 (2002).
30. B. A. Galanov, V. Domnich and Y. Gogotsi, *Exp. Mech.* **43**, 303–308 (2003).
31. S. D. Mesarovic and K. L. Johnson, *J. Mech. Phys. Solids* **48**, 2009–2033 (2000).
32. D. Shi, Y. Lin and T. C. Ovaert, *J. Tribol.* **125**, 223–231 (2003).
33. V. I. Fabrikant, *J. Strain Anal. Eng. Design* **39**, 55–70 (2004).
34. J. N. Israelachvili and D. Tabor, *Proc. R. Soc. Lond. A* **331**, 19–38 (1972).
35. J. N. Israelachvili, in: *Fundamentals of Friction*, I. L. Singer and H. M. Pollock (Eds), p. 351. Kluwer, Dordrecht (1992).
36. K. L. Johnson and I. Sridhar, *J. Phys. D* **34**, 683–689 (2001).
37. E. Barthel and S. Roux, *Langmuir* **16**, 8134–8138 (2000).
38. M. P. de Boer and T. M. Mayer, *Mater. Res. Bull.* **26**, 302–304 (2001).
39. J. S. Villarrubia, *Surf. Sci.* **321**, 287–300 (1994).
40. J. S. Villarrubia, *J. Res. (Natl. Inst. Stand. Technol.)* **102**, 425–454 (1997).
41. M. Enachescu, R. W. Carpick, D. F. Ogletree and M. Salmeron, *J. Appl. Phys.* **95**, 7694–7700 (2004).
42. G. W. Tormoen, J. Drelich and E. R. Beach, III, *J. Adhesion Sci. Technol.* **18**, 1–17 (2004).

43. E. R. Beach, G. W. Tormoen, J. Drelich and R. Han, *J. Colloid Interf. Sci.* **247**, 84–99 (2002).
44. J. P. Cleveland, S. Manne, D. Bocek and P. K. Hansma, *Rev. Sci. Instrum.* **64**, 403–405 (1993).
45. J. E. Sader, J. W. M. Chon and P. Mulvaney, *Rev. Sci. Instrum.* **70**, 3967–3969 (1999).
46. R. J. Cannara and M. J. Brukman, *Rev. Sci. Instrum.* (submitted).
47. R. W. Carpick, N. Agraït, D. F. Ogletree and M. Salmeron, *Langmuir* **12**, 3334–3340 (1996).
48. R. W. Carpick, D. F. Ogletree and M. Salmeron, *Appl. Phys. Lett.* **70**, 1548–1550 (1997).
49. R. W. Carpick, M. Enachescu, D. F. Ogletree and M. Salmeron, in: *Fracture and Ductile vs. Brittle Behavior – Theory, Modeling and Experiment*, G. Beltz, K.-S. Kim and R. L. Selinger (Eds), pp. 93–103. Mater. Res. Soc., Warrendale, PA (1999).
50. M. Enachescu, R. J. A. van den Oetelaar, R. W. Carpick, D. F. Ogletree, C. F. J. Flipse and M. Salmeron, *Tribol. Lett.* **7**, 73–78 (1999).
51. B. J. Briscoe and D. C. B. Evans, *Proc. R. Soc. Lond. A* **380**, 389–407 (1982).
52. I. L. Singer, R. N. Bolster, J. Wegand, S. Fayeulle and B. C. Stupp, *Appl. Phys. Lett.* **57**, 995–997 (1990).
53. O. Piétrement and M. Troyon, *Langmuir* **17**, 6540–6546 (2001).
54. A. R. Burns, J. E. Houston, R. W. Carpick and T. A. Michalske, *Phys. Rev. Lett.* **82**, 1181–1184 (1999).
55. M. Enachescu, R. J. A. van den Oetelaar, R. W. Carpick, D. F. Ogletree, C. F. J. Flipse and M. Salmeron, *Phys. Rev. Lett.* **81**, 1877–1880 (1998).
56. R. Lüthi, E. Meyer, M. Bammerlin, L. Howald, H. Haefke, T. Lehmann, C. Loppacher, H.-J. Güntherodt, T. Gyalog and H. Thomas, *J. Vac. Sci. Technol. B* **14**, 1280–1284 (1996).
57. M. A. Lantz, S. J. O’Shea, A. C. F. Hoole and M. E. Welland, *Appl. Phys. Lett.* **70**, 970–972 (1997).
58. O. Piétrement and M. Troyon, *Surf. Sci.* **490**, L592–596 (2001).
59. Z. Wei, C. Wang and C. Bai, *Langmuir* **17**, 3945–3951 (2001).
60. I. L. Singer, in: *Fundamentals of Friction*, I. L. Singer and H. M. Pollock (Eds), p. 237. Kluwer, Dordrecht (1992).
61. O. Piétrement, J. L. Beaudoin and M. Troyon, *Tribology Lett.* **7**, 213–220 (1999).
62. M. A. Lantz, S. J. O’Shea and M. E. Welland, *Phys. Rev. B* **56**, 15345–15352 (1997).
63. A. Socoliuc, R. Bennewitz, E. Gnecco and E. Meyer, *Phys. Rev. Lett.* **92**, 134301–134304 (2004).
64. P.-R. Cha, D. J. Srolovitz and T. K. Vanderlick, *Acta Mater.* **52**, 3983–3996 (2004).
65. B. Luan and M. O. Robbins, Personal Communication.
66. L. Wenning and M. H. Müser, *Europhys. Lett.* **54**, 693–699 (2001).