Journal of Colloid and Interface Science **211,** 395–400 (1999)
Article ID jcis.1998.6027, available online at http://www.idealibrary.com on IDEAL®

A General Equation for Fitting Contact Area and Friction vs Load Measurements

Robert W. Carpick,[1] D. Frank Ogletree, and Miquel Salmeron

Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Received August 31, 1998, accepted November 30, 1998

The variation of contact area with load between adhesive elastic spheres depends upon the effective range of attractive surface forces. Relatively simple forms to describe limiting cases exist, but the general intermediate case requires a more complex analysis. Maugis, using a Dugdale model [D. Maugis, *J. Colloid. Interf. Sci.* 150, 243 (1992)], provides an analytic solution, but the resulting equations are cumbersome if one wishes to compare with experimental data such as atomic force microscope measurements. In this paper we present a simpler general equation that approximates Maugis' solution extremely closely. 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.

Key Words: **contact mechanics; contact area; adhesion; Dugdale model; adhesive spheres; atomic force microscope; friction force microscope; surface forces apparatus.**

1. INTRODUCTION

An understanding of contact, adhesion, and friction between surfaces requires knowledge of the area of contact between them. Continuum models which predict the contact area for various geometries have been worked out starting with the pioneering work of Hertz (1). Experimental techniques that measure contact area for elastic single asperities, namely the surface forces apparatus (SFA) (2, 3) and the atomic force microscope (AFM) (4), can be described by the geometry of contacting spheres (approximated as paraboloids). In the case of the AFM, one sphere represents the sample, with an infinite radius of curvature, while the other sphere represents the tip, which in many cases conforms to a paraboloidal shape (5). In the absence of adhesion, the Hertz model has been shown to accurately describe the contact area between elastic spheres (6). However, at small scales the surface-to-bulk ratio becomes significant. Therefore, adhesion arising from attractive surface forces is generally not negligible and must be included in any description of contact area.

The spatial range over which surface forces act depends upon the chemistry of the materials in contact, and may or may not be long range compared to the scale of elastic deformations due to these forces (7) (Fig. 1). Two limiting cases are apparent. When the surface forces are short range in comparison to the elastic deformations they cause (i.e. compliant materials, strong adhesion forces, large tip radii), the contact area is described by the Johnson–Kendall–Roberts (JKR) model (8). The opposite limit (i.e., stiff materials, weak adhesion forces, small tip radii) is referred to as the Derjaguin–Müller–Toporov (DMT) regime (9) and the form of the contact area is presented in the work of Maugis[2] (10). It is convenient to utilize a nondimensional physical parameter to quantify these limits and cases in between. Often referred to as Tabor's parameter μ, this transition parameter is defined as (11)

$$\mu = \left(\frac{16R\gamma^2}{9K^2z_0^3}\right)^{1/3}, \qquad [1]$$

where z_0 is the equilibrium separation of the surfaces, R is the tip curvature radius, γ is the interfacial energy per unit area (work of adhesion), and K is the combined elastic modulus of tip and sample, given by $K = 4/3 \cdot ((1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2)^{-1}$ where E_1 and E_2 are the tip and sample Young's moduli, respectively, and ν_1 and ν_2 are the tip and sample Poisson ratios, respectively. The quantity μ is in fact equal to the ratio of the elastic deformation just before the surfaces separate to the equilibrium separation z_0.

To approximate an actual interaction potential like that depicted in Fig. 1, Maugis (10) considers a "Dugdale" (square well) potential to describe attractive forces between contacting spheres (Fig. 2), where a constant adhesive 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, given by

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

By choosing σ_0 to match the minimum adhesive stress of a Lennard–Jones potential (with equilibrium separation z_0), it follows that $\delta_t = 0.97z_0$, and so $\lambda = 1.1570\mu$. Thus, λ and μ are roughly equivalent. For convenience we shall refer to λ in this paper as the "transition parameter." If $\lambda > 5$, the JKR

[1] Present address: Mailstop 1413, Sandia National Laboratories, Albuquerque, NM 87185-1413.

[2] As discussed by Greenwood (10), it is more appropriate to refer to the DMT regime as the "Bradley limit," which is a model for completely rigid spheres. Here we refer to "DMT" in a conventional sense, as with Maugis (9).

0021-9797/99 $30.00

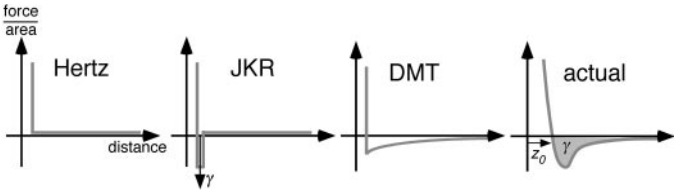

FIG. 1. Interaction forces (per unit area) for the Hertz, JKR, and DMT models, compared to a realistic interaction. There is no attractive force in the Hertz model, only hard wall repulsion at contact. The JKR model includes short range adhesion which is essentially a delta function with strength γ and thus acts only within the contact zone. The DMT curve shown represents a long-range surface force. A volume integrated force, like the van der Waal's 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 force–distance attractive well corresponds to the work of adhesion, γ.

model applies, and if $\lambda < 0.1$, the DMT model applies. Values between 0.1 and 5 correspond to the "transition regime" between JKR and DMT. A summary of different conventions used for defining this "transition parameter" is provided by Greenwood (11). The Hertz model applies when there are no attractive surface forces ($\gamma = 0$).

2. DEPENDENCE OF CONTACT AREA UPON LOAD FOR THE JKR AND DMT REGIMES

The variation of contact area with load for various values of λ, and for the Hertz model, is plotted in Fig. 3. In the non-zero adhesion cases, there is a well-defined "pull-off force" or negative "critical load" at which the surfaces separate when being pulled apart. This applies for the case of "fixed loads" where one of the spheres is attached to a compliant spring (as in the case of an AFM cantilever). We shall refer to this negative critical load as L_c, and for the two limiting cases it is given by

$$L_{c(JKR)} = -\frac{3}{2}\pi\gamma R \qquad [3a]$$

$$L_{c(DMT)} = -2\pi\gamma R. \qquad [3b]$$

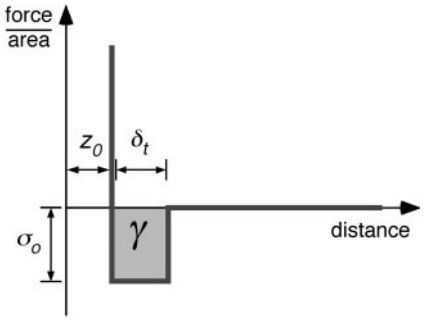

FIG. 2. The force–distance relation for the Dugdale model used by Maugis. A constant adhesive stress (force per unit area) σ_0 acts between the surfaces over a range δ_t. At greater separations, the attractive force is zero. The work of adhesion is thus $\gamma = \sigma_0\delta_t$.

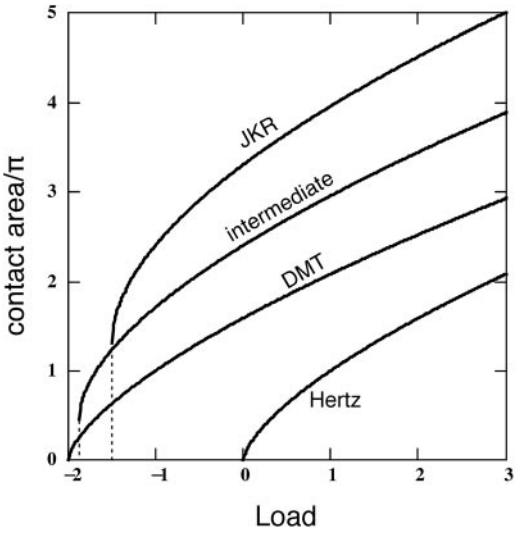

FIG. 3. The Hertz area–load curve, and the JKR–DMT transition, plotted in Maugis' units (Eq. [9]). Area–load curves for the JKR limit, the DMT limit, and an intermediate case are shown. These approach the Hertz curve in the limit $\gamma \to 0$ (no adhesion). Adhesion increases the contact area from the Hertz case for a given load by an amount dependent upon the range of attractive forces.

The models also predict particular values for the contact radius at zero load a_0, given by

$$a_{0(JKR)} = \left(\frac{6\pi\gamma R^2}{K}\right)^{1/3} \qquad [4a]$$

$$a_{0(DMT)} = \left(\frac{2\pi\gamma R^2}{K}\right)^{1/3}. \qquad [4b]$$

The variation of contact radius a with load L for the JKR and DMT cases are each described by relatively simple equations:

$$\frac{a}{a_{0(JKR)}} = \left(\frac{1 + \sqrt{1 - L/L_{c(JKR)}}}{2}\right)^{2/3} \qquad [5a]$$

$$\frac{a}{a_{0(DMT)}} = (1 - L/L_{c(DMT)})^{1/3}. \qquad [5b]$$

Since the tip is axially symmetric, the contact area A is simply given by

$$A = \pi a^2. \qquad [6]$$

It is apparent that Eqs. [5a] and [5b] can be generalized to form an equation which describes the contact radius for both cases,

$$\frac{a}{a_{0(\alpha)}} = \left(\frac{\alpha + \sqrt{1 - L/L_{c(\alpha)}}}{1 + \alpha}\right)^{2/3}, \qquad [7]$$

where $\alpha = 1$ corresponds exactly to the JKR case, and $\alpha = 0$ corresponds exactly to the DMT case. Included in Eq. [7] is the fact that L_c and a_0 depend on α as well. We shall refer to Eq. [7] as the generalized transition equation. We now show that for intermediate cases ($0 < \alpha < 1$), the generalized transition equation corresponds very closely to solutions for the transition regime ($0.1 < \lambda < 5$) elegantly worked out by Maugis using the Dugdale model (10).

3. DEPENDENCE OF CONTACT AREA UPON LOAD FOR THE MAUGIS–DUGDALE MODEL

Utilizing the Dugdale model is a reasonable method for estimating the value of the contact radius (and other quantities) as a function of load. The solution is referred to as the "MD" solution (Maugis–Dugdale). Two equations are needed to relate a and L,

$$\frac{\lambda\hat{a}^2}{2}\left[\sqrt{m^2-1}+(m^2-2)\cos^{-1}\left(\frac{1}{m}\right)\right] + \frac{4\lambda^2\hat{a}}{3}\left[\sqrt{m^2-1}\cos^{-1}\left(\frac{1}{m}\right)-m+1\right]=1 \quad [8a]$$

$$\hat{L}=\hat{a}^3-\lambda\hat{a}^2\left[\sqrt{m^2-1}+m^2\cos^{-1}\left(\frac{1}{m}\right)\right], \quad [8b]$$

where $\hat{L}$ and $\hat{a}$ are simple parameterizations of L and a,

$$\hat{L}=\frac{L}{\pi\gamma R} \quad [9a]$$

$$\hat{a}=a\cdot\left(\frac{K}{\pi\gamma R^2}\right)^{1/3}, \quad [9b]$$

and the parameter m represents the ratio between the contact radius a and an outer radius c at which the gap between the surfaces reaches δ_t (i.e., where the adhesive stress no longer acts). Maugis' equations properly predict the JKR and DMT limits.

The difficulty in utilizing the MD equations lies in the lack of a single expression relating only a and L. To plot the MD solution or fit it to data, one needs to simultaneously solve Eqs. [8a] and [8b] by letting m vary appropriately between limits which depend upon λ. Furthermore, the relation for the pull-off force must be determined through iteration (12) if the value of λ is not known a priori (the usual case with experimental measurements). In practice, this is rather cumbersome if not impossible to carry out with common software programs that utilize automated statistical fitting procedures.

TABLE 1
Conversion Table between α and λ and the Associated Values of $\hat{L}_c(\lambda)$ and $\hat{a}_0(\lambda)$

α	λ	$\hat{L}_c(\lambda)$	$\hat{a}_0(\lambda)$
DMT: 0	0	−2	$\sqrt[3]{2} = 1.260\cdots$
0.074	0.1	−1.951	1.336
0.158	0.2	−1.881	1.394
0.256	0.4	−1.816	1.442
0.433	0.5	−1.718	1.517
0.609	0.8	−1.634	1.598
0.692	1.0	−1.601	1.636
0.817	1.5	−1.556	1.700
0.886	2.0	−1.535	1.738
0.922	2.5	−1.523	1.760
0.944	3.0	−1.517	1.775
0.958	3.5	−1.513	1.785
0.967	4.0	−1.510	1.792
0.979	5.0	−1.506	1.800
JKR: 1	∞	−3/2	$\sqrt[3]{6} = 1.817\cdots$

4. COMPARISON OF THE MD SOLUTION AND THE GENERALIZED TRANSITION EQUATION

In Fig. 4, we plot the MD solution for four different values of λ. The generalized transition equation is overlaid, where the value of α in each case was optimized to provide the best fit. Each pair of plots share identical values of a_0 and L_c, respectively. For all values of λ, the difference between the generalized transition equation and the MD solution is less than 1% (and less than 0.1% in most cases) for the relevant load regime. Furthermore, the generalized transition equation is exact for the JKR and DMT limits. Therefore, the generalized transition equation is an excellent approximation to the MD solution. The optimization of α was performed by minimizing the integrated square deviation between the MD solution and the generalized transition equation for the load range $-|L_{c(\alpha)}|$ to $+2|L_{c(\alpha)}|$. This load range was chosen because it is typical of the range employed in AFM measurements (4), where loads are kept low to avoid wear. The low load regime also exhibits the most substantial curvature of the a *vs* L relationship in general, thus making fitting more reliable.

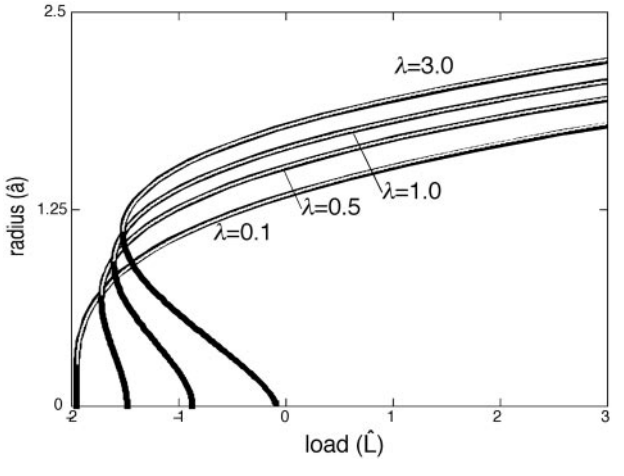

FIG. 4. The MD solution (heavy black lines) and the generalized transition equation, Eq. [7] (thin white lines), compared for four different values of λ. The generalized transition equation is plotted for values of $\hat{a}$ above the critical load, while the MD solution includes the unstable lower branch of solutions.

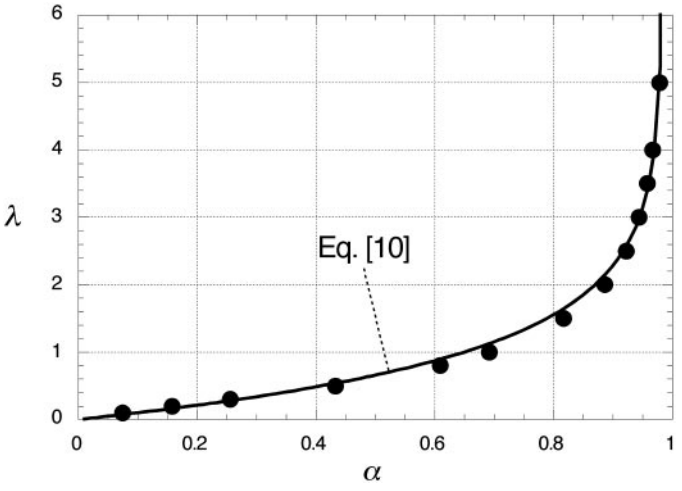

FIG. 5. Plot of λ vs α from Table I (filled circles) and the empirical conversion equation, Eq. [10] (solid line), fit to the data.

Table 1 displays the optimized values of α for various values of λ. An empirical formula was fit to this table of values to provide a conversion equation from α to λ (Fig. 5), which is given by

$$\lambda = -0.924 \cdot \ln(1 - 1.02\alpha). \qquad [10]$$

For $\alpha > 1.02^{-1} = 0.980$, one can simply assume the JKR limit ($\alpha = 1$). The steep slope as α approaches the JKR limit means a larger uncertainty in determining λ from a value of α in this range. Essentially this applies for values of α greater than ~0.95 or equivalently for values of λ greater than ~3. This is not problematic since the change in the contact radius as λ ranges from 3 and up is reasonably small, apparent from Fig. 5a of Maugis (10). In other words, a substantial uncertainty in λ in this range does *not* lead to a correspondingly substantial uncertainty in the contact radius. The generalized transition equation is therefore reliable for all values of λ, and Eq. [10] provides a reasonable and simple conversion from α to λ. Importantly, the generalized transition equation does not introduce any additional free parameters compared to the MD solution.

L_c and a_0 can be represented in Maugis' nondimensional units given in Eqs. [9a] and [9b] as follows:

$$\hat{L}_c(\lambda) = \frac{L_c}{\pi\gamma R} \qquad [11a]$$

$$\hat{a}_0(\lambda) = a_0 \cdot \left(\frac{K}{\pi\gamma R^2}\right)^{1/3}. \qquad [11b]$$

These nondimensional values are uniquely determined by the value of λ and are listed in Table 1 and plotted vs λ in Figs. 6 and 7, respectively. We have determined empirical equations

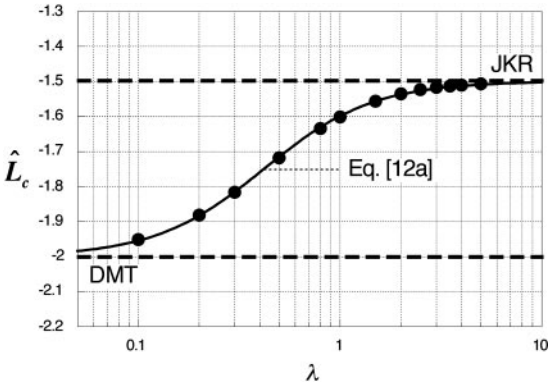

FIG. 6. Plot of $\hat{L}_c(\lambda)$ vs λ determined from the MD solution (filled circles), and an empirical fit, Eq. [12a] (solid line). The values of $\hat{L}_c(\lambda)$ for the JKR and DMT limits are indicated by the dotted lines.

that fit these values reasonably well (within ~1% accuracy or better), which are also plotted in Figs. 6 and 7, respectively:

$$\hat{L}_c(\lambda) = -\frac{7}{4} + \frac{1}{4} \cdot \left(\frac{4.04 \cdot \lambda^{1.4} - 1}{4.04 \cdot \lambda^{1.4} + 1}\right) \qquad [12a]$$

$$\hat{a}_0(\lambda) = 1.54 + 0.279 \cdot \left(\frac{2.28 \cdot \lambda^{1.3} - 1}{2.28 \cdot \lambda^{1.3} + 1}\right). \qquad [12b]$$

Eqs. [12a] and [12b] provide fast and simple ways of estimating $\hat{L}_c(\lambda)$ and $\hat{a}_0(\lambda)$. These quantities can instead be determined exactly from the MD equations using numerical methods if preferred.

5. FITTING PROCEDURE

The above discussion leads to the following procedure for curve fitting:

(1) Obtain a measurement of contact radius a or friction F_f as a function of load. For friction measurements, one must

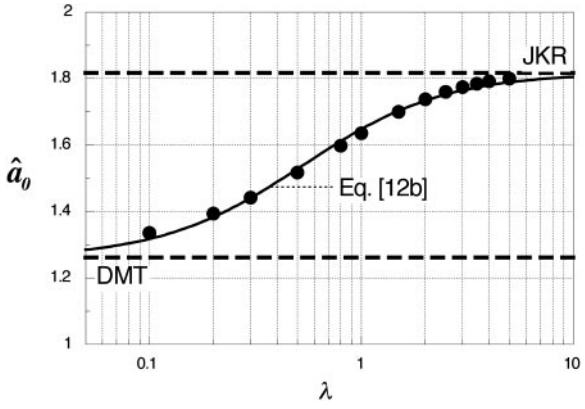

FIG. 7. Plot of $\hat{a}_0(\lambda)$ vs λ determined from the MD solution (solid circles) and an empirical fit, Eq. [12b] (solid line). The values of $\hat{a}_0(\lambda)$ for the JKR and DMT limits are indicated by the dotted lines.

assume that friction is directly proportional to the contact area ($F_f = \tau \cdot \pi a^2$, where τ is the constant interfacial shear strength) (4, 5). This is not expected to be true for multiple asperity contacts, or in the presence of wear (see Section 6). In the case of friction measurements, replace all occurrences of a with $\sqrt{F_f}$ in the relevant equations and discussion below.

(2) Use a curve fit optimization routine such as those available in commercial software programs to fit the generalized transition equation, Eq. [7], to the data, letting a_0, L_c, and α be free parameters whose values will then be extracted from the fit. Alternately, measure L_c independently from force–distance curves (4) and constrain L_c to this value for the fit. Be aware that occasionally the last data point before separation in a sliding friction measurement where the load is being decreased may not correspond to L_c if the tip and sample have prematurely separated.

(3) Use the above conversion equation, Eq. [10], or the values in Table 1, to determine λ from the value of α extracted from fitting the generalized transition equation.

(4) The interfacial energy, γ, can now be determined. Use Eq. [12a] to solve for $\hat{L}_c(\lambda)$ now that λ is known. Plug this value into Eq. [11a] along with the value of L_c extracted from the generalized transition equation fit to solve for γ (this can be done only if the tip radius R is known). This can also be done numerically using the MD equations.

(5) If measuring contact radius, the elastic modulus of the contact K can be determined. Use Eq. [12b] to determine the value of $\hat{a}_0(\lambda)$. Plug this value and the value of a_0 extracted from the generalized transition equation fit into Eq. [11b] to determine K. This requires knowing R and using the value of γ determined above. This can also be done numerically using the MD equations.

(6) If measuring friction instead of contact radius, determine τ by using $\tau = F_0/\pi a_0^2$, where F_0 is extracted from the generalized transition equation fit, and a_0 is determined by combining Eqs. [12b] and [11b]. This requires knowing values for R and K.

6. DISCUSSION

One must be careful in utilizing the generalized transition equation for the case of friction measurements since it assumes direct proportionality between friction and contact area. In general, SFA, AFM or other measurements where this model can be applied require elastic, single-asperity contacts. This is readily determined in the case of SFA, where the smooth mica sheets ensure a single contact, and optical measurements can reveal if wear is occurring (13). More care is required with AFM measurements. Tip shapes must be characterized to verify that they consist of a single, well-defined paraboloidal protrusion (4, 5). Loads must be kept low to avoid wear, especially since atomic-scale wear may not be readily visible with standard AFM imaging capabilities (4).

The study of nanometer-scale single asperity contacts is still relatively new, and it is not certain that friction and contact area will always be directly proportional even in the elastic single asperity regime. For example, there could be a significant pressure dependence of the shear strength τ, or a significant change in the contact area due to the presence of lateral forces, as opposed to purely normal forces. The latter issue has been discussed in further detail by Johnson (14). Facile comparison of friction data with the generalized transition equation could help to elucidate deviations due to these effects. In general, experiments where both contact area and friction are measured are ideal for such purposes (12, 15, 16).

Although the generalized transition equation is an excellent approximation for the MD solution, it is important to consider that the MD solution is exact for an *idealized model potential*. A more physically realistic potential is the Lennard–Jones potential, but this can be solved only numerically. Greenwood has performed this calculation for contacting spheres (11). At this point the limits of continuum solutions become apparent. It is not clear exactly where to define the edge of the contact zone, i.e., the contact radius. For real materials, the contact zone edge corresponds to the point where interfacial bonds are no longer formed. In a continuum model, one can define it to be where the gap between the continuum surfaces first displays an infinitesimal increase, or possibly somewhere further out from the center of the contact zone. For parameters in a typical AFM contact, one may need to span radially outward a few nanometers to find a change in the gap of only 1 Å (i.e., less than an atomic bond length). Continuum mechanics cannot address such a question in an exact fashion when the atomistic nature of the materials must dominate. Therefore, it is possible that the MD solution for the contact radius is not an accurate prediction for real materials at nanometer length scales. Johnson, for example, has proposed an alternative based on the MD model (12, 14) which closely resembles Greenwood's result. The only way to test all of this is to compare experimental measurements with the MD model, Johnson's, and others. The generalized transition equation facilitates this. A simple analytic equation for Johnson's or Greenwood's models analogous to the generalized transition equation would be useful.

In summary, we have provided a simple analytic equation which can be used to fit the contact radius (or friction) as a function of load to scanning probe measurements and others. The difference between this generalized transition equation and the MD solution is insignificant. A table of values and an empirical equation to determine the transition parameter λ from the generalized transition equation parameter α is provided. This approach requires a parabolic tip, a flat sample, and wearless elastic interactions. Application of the generalized transition equation to friction or contact radius data can help identify the accuracy of particular continuum models to describe the contact radius for nanometer scale contacts and others.

ACKNOWLEDGMENTS

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Materials Sciences Division, of the U.S. Department of Energy under Contract DE-AC03-76SF00098. R.W.C. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Hertz, H., *J. Reine Angew. Math.* **92,** 156 (1881).
2. Israelachvili, J. N., and Tabor, D., *Proc. R. Soc. London A* **331,** 19 (1972).
3. Tabor, D., and Winterton, R. H. S., *Proc. R. Soc. London A* **312,** 435 (1969).
4. Carpick, R. W., and Salmeron, M., *Chem. Rev.* **97,** 1163 (1997).
5. Carpick, R. W., Agraït, N., Ogletree, D. F., and Salmeron, M., *J. Vac. Sci. Technol. B* **14,** 1289 (1996).
6. Johnson, K. L., "Contact Mechanics" Univ. Press, Cambridge, 1987.
7. Johnson, K. L., and Greenwood, J. A., *J. Colloid Interface Sci.* **192,** 326 (1997).
8. Johnson, K. L., Kendall, K., and Roberts, A. D., *Proc. R. Soc. London A* **324,** 301 (1971).
9. Derjaguin, B. V., Muller, V. M., and Toporov, Y. P., *J. Colloid Interface Sci.* **53,** 314 (1975).
10. Maugis, D., *J. Colloid Interface Sci.* **150,** 243 (1992).
11. Greenwood, J. A., *Proc. R. Soc. London A* **453,** 1277 (1997).
12. Lantz, M. A., O'Shea, S. J., Welland, M. E., and Johnson, K. L., *Phys. Rev. B* **55,** 10776 (1997).
13. Israelachvili, J. N., *in* "Fundamentals of Friction" (I. L. Singer and H. M. Pollock, Eds.). Kluwer, Dordrecht, 1992.
14. Johnson, K. L., *Proc. R. Soc. London A* **453,** 163 (1997).
15. Enachescu, M., van den Oetelaar, R. J. A., Carpick, R. W., Ogletree, D. F., Flipse, C. F. J., and Salmeron, M., *Phys. Rev. Lett.* **81,** 1877 (1998).
16. Carpick, R. W., Ogletree, D. F., and Salmeron, M., *Appl. Phys. Lett.* **70,** 1548 (1997).