Deformation and adhesion of elastic bodies in contact

Phil Attard and John L. Parker

Department of Applied Mathematics, Research School of Physical Sciences and Engineering, Australian National University, Canberra, Australian Capital Territory, 2601, Australia (Received 6 May 1992)

The elastic deformation and adhesion of two convex bodies that interact via surface forces of finite range are calculated self-consistently. Hertz theory is compared to the result of an exponential repulsion, and it is found to be valid in the limit of short-ranged forces and high loads. A Lennard-Jones law is used to examine the classical theories of adhesion, which relate the surface energy to the pull-off force, and their regime of validity is explored. Explicit expressions are given for the displacement prior to contact, and for the jump instabilities due to elastic deformation, which occur for compliant bodies with rapidly changing surface forces. The loading-unloading cycle is shown to be hysteretic for large adhesions, and this is correlated with the onset of jumps. In these cases the pull-off force is demonstrated to depend upon the history of the sample, and it increases with increasing maximum applied loads.

PACS number(s): 03.40.Dz, 46.30.Pa, 62.20.Fe, 68.35.Md

I. INTRODUCTION

The phenomenon of deformation, whereby a solid body changes shape in response to an applied load, may be described by the classical continuum theory of elasticity. On the other hand, the application of the load is necessarily mediated by surface forces, which arise much more directly from molecular interactions, as is particularly evident when the surfaces adhere. Traditionally, surface interactions have been included by simply invoking infinitely short-ranged contact forces. This paper is concerned with the effects of realistic surface forces of finite range on the elastic deformation of solids, and also on the relationship between their adhesion and surface energy.

Historically, the theory of elastic deformation began with Hertz [1], who analyzed the shape of spherical glass lenses under gravitational load. Under the assumption of a flat, nonadhesive contact region, he derived the relationship between the contact radius, the applied load, and the central displacement. Johnson, Kendall, and Roberts (JKR) [2] modified the Hertz theory to account for adhesion between the surfaces in the flattened contact region. In particular, they obtained an expression for the force F_a required to separate adhesive spheres of radius 2R.

$$F_a = -\frac{3}{2}\pi\gamma R \quad , \tag{1}$$

where, in vacuo, $\gamma > 0$ is twice the surface free energy per unit area of the solid. Derjaguin, Muller, and Toporov (DMT) [3] assumed that the Hertz profile was unmodified by adhesion, and found for the pull-off force

$$F_a = -2\pi\gamma R \quad . \tag{2}$$

By utilizing the Hertz profile, they avoided the singularities at the boundary of the contact zone that are present in JKR theory, and which are believed to make that theory more applicable to soft solids with high surface energies [4,5].

It is difficult to resolve experimentally the contradic-

tion between the JKR and DMT expressions for the pull-off force because accurate measurements are problematic, as are independent determinations of γ . More exact theoretical approaches are required, and so Muller, Yushchenko, and Derjaguin [4] solved the elasticity equations numerically for the case of a simplified model of real, adhesive surfaces, namely a Lennard-Jones interaction. Earlier, Hughes, and White [6,7] had pointed out that the classical assumption of a flat contact region was incommensurate with the variation of stress across that region, since for surface forces of finite range the latter implies a variation in surface separation. These they termed soft-contact problems, and they obtained results for an exponentially repulsive force law [6,7]. A different theoretical approach calculates the mutual deformation of crystal lattices due to their interatomic potentials [8,9]. One prediction is that for small separations and attractive forces the deformation can be unstable, and the surfaces can jump into contact [8,9].

In this paper the consequences of finite-ranged surface forces are considered, and relatively extensive results are presented for two basic forces, an exponential repulsion and a Lennard-Jones adhesion. Exact results are obtained by the self-consistent numerical solution of the elasticity equations. Support is provided for the conclusion of Hughes and White [6,7] that Hertz theory is valid in the limit of short-ranged repulsive forces and large applied loads. An explicit estimate of the relative accuracy of the Hertz theory for a range of characteristic parameter values is given. The findings of Muller, Yushchenko, and Derjaguin [4,5] are confirmed, namely that the DMT expression is valid for small adhesions, and for small, hard particles. In the intermediate regime, the pull-off force approaches the JKR prediction. It is also found that for rapidly varying attractions and compliant bodies, an instability in the force-deformation relation means that the surfaces jump into contact from a finite separation. Similar behavior has previously been predicted theoretically [8,9] and measured experimentally [10]. Approximate analytic expressions are given for the deformation prior to contact, and for position of the jump into contact. For soft bodies with large adhesions, hysteresis is found in the load-displacement cycle, similar to that observed experimentally [11–16], and neither of the classical theories correctly relates the pull-off force to the surface energy in this regime.

II. ANALYSIS

After a brief review of the Hertz and JKR theories, this section outlines the application of the elastic deformation equations to the interaction of convex bodies. A description of the numerical algorithm used to obtain the results presented in this paper is included. An approximate analytic expression valid for slowly varying deformations is then given, followed by an analysis of the stability of adhesive bodies. The section concludes with a description of the particular finite-ranged interaction forces used here.

A. Hertz theory

Hertz [1] found the pressure distribution p(r) that leads to flattened contact of two elastic bodies,

$$p(r) = \begin{cases} \frac{E}{1 - v^2} \frac{a}{\pi R} (1 - r^2 / a^2)^{1/2}, & r \le a \\ 0, & r \ge a \end{cases}$$
 (3)

where a is the so-called contact radius, and r is the distance from the axis of the cylindrically symmetric system. Here E is Young's modulus, and v is Poisson's ratio. R is a geometrical factor related to the curvature of the bodies; for two identical crossed cylinders, or a sphere against a plane, it is the radius, and for two identical spheres it is half the radius. The central displacement, which measures the deformation on the axis, and which for two spheres is the sum of their radii less the distance between their centers, is, in Hertz theory,

$$\delta = \frac{a^2}{R} \ . \tag{4}$$

The applied load, which is the integral of the pressure profile, is

$$F = \frac{2}{3} \frac{E}{1 - v^2} \frac{a^3}{R} \ . \tag{5}$$

Hence the Hertz load-displacement relation is

$$\delta = \left[\frac{3}{2} \frac{1 - v^2}{E} \frac{F}{\sqrt{R}} \right]^{2/3} . \tag{6}$$

B. JKR theory

Johnson, Kendell, and Roberts (JKR) [2] added to the Hertz pressure, Eq. (3), a pressure distribution of the form

$$p(r) = \begin{cases} p_1/(1-r^2/a^2)^{1/2}, & r \le a \\ 0, & r \ge a \end{cases}$$
 (7)

Such a profile causes a uniform deformation within the contact zone, and a singularity in the curvature of the surface at the edge of the zone. The value of $p_1 < 0$ is determined by minimizing the sum of the elastic and the surface energy. In JKR theory the latter is taken to be the contact energy,

As stated above, γ is twice the solid surface free energy per unit area, for contact in vacuum or gas. For contact in a liquid, it is the negative of the interaction free energy per unit area between planar solids at the equilibrium separation under zero load, which approximately equals the solid-solid free energy less twice the solid-liquid free energy, per unit area.

The energy minimization gives [17]

$$p_1 = -\left[\frac{E}{1 - v^2} \frac{\gamma}{\pi a}\right]^{1/2},\tag{9}$$

from which the contact radius is

$$a^{3} = \frac{3(1-v^{2})R}{2E} \{ F + 3\pi\gamma R + [6\pi\gamma RF + (3\pi\gamma R)^{2}]^{1/2} \} ,$$
(10)

and the central displacement is

$$\delta = \frac{a^2}{R} - \left[4\pi\gamma a (1 - v^2)/E\right]^{1/2} . \tag{11}$$

Finally, the pull-off force in JKR theory, Eq. (1), is the maximum tension that the bodies can sustain. If instead of the load the displacement is controlled, then the pull-off force predicted by JKR is

$$F_{\delta} = \frac{5}{9} F_{\sigma} = -\frac{5}{6} \pi \gamma R \quad . \tag{12}$$

C. Elastic deformation of convex bodies

In general, the variation in stress and strain within an elastic body decays rapidly away from the point of application of a load, and one may therefore use the elastic equations derived for a semi-infinite half-space. In the linear elastic regime, a pressure distribution $p(\mathbf{r})$ applied inward to the surface causes a deformation of the surface in the direction of the outward normal given by [18]

$$u(\mathbf{r}) = -\frac{1 - v^2}{\pi E} \int \frac{p(\mathbf{t})}{|\mathbf{r} - \mathbf{t}|} d\mathbf{t} . \tag{13}$$

The problems considered here have cylindrical symmetry, and one has

$$u(\mathbf{r}) = -\frac{1 - v^2}{E} \int_0^\infty p(t)k(r, t)tdt , \qquad (14)$$

where the kernel is

$$k(r,t) = \begin{cases} \frac{4\pi}{r} K(t^2/r^2), & t < r \\ \frac{4\pi}{t} K(r^2/t^2), & t > r \end{cases}$$
 (15)

K(m) being the complete elliptic integral of the first kind of modulus m.

The deformation in one body is due to the load applied by a second body. As described above, Hertz and JKR theories solve the deformation equation assuming a particular pressure profile. In fact, the stress distribution in the region of contact, p(r), arises from the surface forces between the bodies, which in turn implies that

$$p(r) = p_s(h(r)) , \qquad (16)$$

where h(r) is the separation between the surfaces at r, measured parallel to the axis. The right-hand side is the inward force per unit area between two semi-infinite half-spaces separated by h(r). This is valid if the radii of the convex bodies is much larger than their surface separation; it is invoked in the derivation of the well-known Derjaguin approximation [14,19–21].

The separation between the surfaces of the bodies at r is

$$h(r) = h_0(r) - 2u(r)$$
, (17)

where $h_0(r)$ is the separation which would occur in the absence of any interaction between the bodies. $[h_0(r)]$ is negative if the bodies would have interpenetrated; for two spheres, $h_0 \equiv h_0(0)$ is the actual distance between their centers less the sum of their undeformed radii.] The unstressed separation depends on the geometry. For two convex bodies it has the general form

$$h_0(r) = h_0 + \frac{r^2}{2R}, \quad r \ll R \quad .$$
 (18)

R is related to the principal radii of curvature of the bodies [20]; for two identical spheres it equals half their radius, and for two identical crossed cylinders, or a sphere against a flat, it equals the radius. Because surface forces decay with separation, $p_s(h) \rightarrow 0$, $h \rightarrow \infty$, the dominant region is always $r \ll R$.

In this work results will be explicitly stated as applying to identical bodies. The more general asymmetric case follows immediately by interpreting $(1-v^2)/E$ as the arithmetic mean of this quantity for the two bodies, and by using the appropriate geometrical quantity R.

Two quantities of interest in contact mechanics are the central displacement

$$\delta \equiv h(0) - h_0 = -2u(0) \tag{19}$$

and the total load

$$F = 2\pi \int_0^\infty p(r)r \, dr \ . \tag{20}$$

Equations (14)-(18) determine the deformation, the separation, and the pressure profiles. They may be solved self-consistently for a specified surface force at a given unstressed separation h_0 . The numerical algorithm used to solve the equations mimics a typical loading-unloading experiment, except that the displacement is controlled rather than the load. The three functions u(r), h(r), and $p(r) = p_s(h(r))$ are defined on a grid, $0 \le r \le R^{\max} \ll R$, which in the results reported here was a uniform spacing of $\approx 10-100$ nm, with 200-1500 nodes. The results were not very sensitive to the choice of grid or spacing, except for compliant bodies with large adhesions.

The algorithm begins with the construction of the kernel, Eq. (15), (cf. Eqs. (3.11) and (A4) of Ref. [22]). One begins at a large separation, h_0 , such that $p_s(h_0) \approx 0$, and initially the surfaces are undeformed [u(r)=0, $h(r) = h_0(r)$]. A cycle begins with a change in h_0 to $h_0 + \Delta$ at constant surface separation. This means that the deformation is changed to $u(r) + \Delta/2$, and h(r) and p(r) are unchanged. The iteration steps in the cycle consist of calculating in turn new deformation, separation, and pressure profiles, via Eqs. (14), (17), and (16), respectively. The new deformation is mixed with the old deformation to ensure convergence. The mixing ratio is automatically decreased if oscillations or growth in the change in the central displacement was detected, and otherwise allowed to increase up to a value of 0.2. The cycle is terminated when the central displacement has converged to within about 10^{-4} nm, typically requiring 50-1000 iterations. At this time the total load is calculated, Eq. (20), and, if required, the surface profiles are stored. For loading, $\Delta < 0$. The unloading $(\Delta > 0)$ commences with the deformation from the end of the loading path. This point, which corresponds to the maximum applied load, is here called the penetration.

For repulsive surface forces, the procedure was found to be very stable, but care was required for adhesive, soft bodies. This is due in part to the very sharp change in the pressure, which is of several orders of magnitude over fractions of a nanometer, in the vicinity of the equilibrium surface separation. This is compounded when the surfaces jump into or out of contact, since in this case a discontinuous change in the deformation occurs. Any hysteresis also increases the possibility of error. When the latter occurs, the elastic equations do not possess unique solutions, and it is quite important to solve them by incrementing h_0 as described above, since this corresponds to the actual experimental procedure.

Since the numerical errors in the solution of the equations can be minimized, the theory may be described as exact. This is strictly true only in the linear elastic regime for surfaces with small curvatures. This restriction appears to be of little practical import. Both Hertz and JKR theories are similarly restricted, and so the present calculations serve as benchmarks to test the application of those classical approaches to systems with realistic surface forces.

D. Slowly varying deformation approximation

One can obtain an approximate solution to the elastic equations when the deformation varies slowly compared to the curvature. In this case one may take

$$u(r) \approx u(0) . \tag{21}$$

Under this assumption, Eq. (13) yields for the central de-

formation

$$u(0) = -\frac{1 - v^2}{\pi E} 2\pi \int_0^\infty p_s(h_0 + t^2/2R - 2u(0))dt$$

$$= -\sqrt{2R} \frac{1 - v^2}{E} \int_{h(0)}^\infty \frac{p_s(h')}{[h' - h(0)]^{1/2}} dh', \qquad (22)$$

where $h(0) = h_0 - 2u(0)$. This gives the central displacement as a function of the actual surface separation on the axis. Under the same approximation, Eq. (21), the total load is

$$F = 2\pi \int_{0}^{\infty} p_{s}(h_{0} + t^{2}/2R - 2u(0))tdt$$

$$= 2\pi R \int_{h(0)}^{\infty} p_{s}(h')dh'$$

$$= 2\pi R E_{s}(h(0)), \qquad (23)$$

where $E_s(h)$ is the interaction free energy per unit area between planar walls separated by h. This result is just the Derjaguin approximation [14,19,20,21].

These two equations relate the central deformation and the applied load to the actual surface separation. They are straightforward to evaluate numerically, since the right-hand sides are known and no iteration is involved. For particular surface forces, analytic results may be obtained (see below). The regime of validity of the approximation, Eq. (21), is for deformations that vary slowly compared to the undeformed curvature of the bodies. Hence it and the consequent results, Eqs. (22) and (23), are applicable to the situation *before* contact. After contact, the surface flattening implies that the variation in the deformation is of the same magnitude as the curvature, and the approximations are not applicable.

E. Stability analysis

Here a stability analysis of the elastic equations is performed in order to find the condition for adhesive surfaces to jump into or out of contact. If the u(r) that satisfies Eq. (13) is perturbed by $\epsilon(r)$ in the right-hand side of that equation, then the left-hand side shifts by $\mu(r)$. Using Eqs. (16) and (17) one has

$$\mu(r) = 2 \frac{1 - v^2}{\pi E} \int \frac{p_s'(h(t))}{|\mathbf{r} - \mathbf{t}|} \epsilon(t) d\mathbf{t} , \qquad (24)$$

where the prime denotes the derivative of the pressure. For the deformation to be stable, the response to the perturbation must be smaller than the perturbation itself. Thus one may take as a criterion for stability the local condition

$$|\mu(r)/\epsilon(r)| < 1 \tag{25}$$

for arbitrary perturbations. Any particular from for $\epsilon(r)$ will yield a sufficient, but not necessary, condition for stability. A convenient choice is

$$\epsilon(r) = \alpha p_s(h(r)) / p_s'(h(r)) , \qquad (26)$$

where α is small. This gives

$$\mu(r) = -2\alpha u(r) , \qquad (27)$$

and the stability condition is

$$|2u(r)p_s'(h(r))/p_s(h(r))| < 1$$
 (28)

A jump occurs when this inequality is violated. This analysis, which applies to both loading and unloading jumps, is relevant only for adhesive bodies. In this case the deformation and the derivative of the pressure are

positive, and the pressure itself is negative.

One is most interested in the stability at r=0, since this is when the bodies jump into or out of contact, and one has

$$2p_s'(h(0))u(0)/p_s(h(0)) > -1. (29)$$

The condition depends on the derivative of the pressure, and on the ratio of the deformation to the pressure, which will be large for compliant bodies. Note that this condition, and the stability criterion (25), differ from the surface rigidity condition given in Ref. [8]. This stability criterion is obviously most useful if the central displacement is known as a function of the surface separation. The approximation given above is suited for this purpose.

F. Interaction forces

This work examines both repulsive and attractive surface forces of finite range. Relatively simple models for both are chosen so that the effects of the finite range can be distinguished from other complications.

The repulsive pressure is chosen to be of the form

$$p_s(h) = P_0 \exp(-\kappa h) . \tag{30}$$

This exponential law may be considered as a generic repulsion between surfaces at small separations; in fact, it is of the same form as the force per unit area between charged surfaces in electrolyte in the linearized Poisson-Boltzmann approximation. A choice of $\kappa^{-1}=1$ nm and $P_0=10$ MN/m² is applicable to mica surfaces, which have a surface potential of less than about 85 mV, in 0.1M monovalent aqueous electrolyte.

The approximate expressions for the central displacement and for the total load, Eqs. (22) and (23), may be evaluated analytically for this exponential force law. One obtains

$$\delta = 2 \frac{1 - v^2}{E} (2\pi R \kappa^{-1})^{1/2} P_0 e^{-\kappa h(0)}$$
(31)

and

$$F = 2\pi R \kappa^{-1} P_0 e^{-\kappa h(0)} , \qquad (32)$$

where again $\delta = h(0) - h_0 = -2u(0)$.

The surface force for the adhesive case is taken from a Lennard-Jones continuum model of the solids,

$$p_s(h) = \frac{A}{6\pi h^3} \left[\frac{z_0^6}{h^6} - 1 \right] . \tag{33}$$

The Hamaker constant A characterizes the van der Waals attraction of the solids across a liquid or gas. Values in the range $10^{-21}-10^{-19}$ J represent the physically realistic regime. The equilibrium separation is z_0 ; throughout this work this was taken to be 0.5 nm. The value of the surface energy in this model is

$$\gamma = \frac{A}{16\pi z_0^2} \ . \tag{34}$$

One may again analytically evaluate the displacement

and force approximations, Eqs. (22) and (23). One obtains

$$\delta = \sqrt{2R} \frac{2A}{3\pi} \frac{1 - v^2}{E} \left[\frac{z_0^6}{h^6} I_8 - I_2 \right] h^{-5/2}$$
 (35a)

$$\approx -\sqrt{2R} \frac{A}{8} \frac{1-v^2}{E} h^{-5/2}$$
, (35b)

where $I_m = \pi(2m)!/2(2^m m!)^2$. The second line follows by neglecting the repulsive part of the interaction, valid for $h \gg z_0$. Similarly,

$$F = 2\pi R \frac{A}{12\pi h^2} \left[\frac{z_0^6}{4h^6} - 1 \right] \tag{36a}$$

$$\approx -2\pi R \frac{A}{12\pi h^2} \ . \tag{36b}$$

Although it is straightforward to obtain the position of the jump from the full expression (35a) and the jump criterion (29), if only the attractive part of the potential is important, then one can obtain an explicit analytic expression for the separation at which the surfaces jump into contact, namely

$$h = \left[\sqrt{2R} \, \frac{3A}{8} \, \frac{1 - v^2}{E} \, \right]^{2/7} \,. \tag{37}$$

This is similar to a result obtained by Pethica and Sutton [8]; the different numerical coefficients appear to arise from the different stability criteria used.

In general, the form of the elastic equations for the deformation and the total load, Eqs. (13) and (20), may be cast in nondimensional form

$$\bar{u}(\bar{r}) = \Theta U[\bar{u}], \quad \bar{F} = \Theta W[\bar{u}], \quad (38)$$

where $\overline{r}=r^2/\lambda R$, $\overline{u}=u/\lambda$, and $\overline{F}=F(1-v^2)/E(R\lambda^3)^{1/2}$. Here λ is some length scale, and U and W are functionals of the dimensionless displacement that depend upon the particular form of the interaction pressure between the bodies. Θ is a dimensionless parameter which characterizes the system (elasticity, interaction, separation, etc.). For the case of the exponential repulsion, Eq. (30), $\lambda=\kappa^{-1}$, and

$$\Lambda \equiv \Theta = \sqrt{\kappa R} \frac{1 - v^2}{E} P_0 e^{-\kappa h_0} , \qquad (39)$$

which is similar to a parameter defined by Hughes and White [6,7]. For the case of the Lennard-Jones pressure, Eq. (33), $\lambda = z_0$, and

$$\sigma \equiv \Theta = \gamma \frac{1 - \nu^2}{E} \left[\frac{R}{z_0^3} \right]^{1/2} . \tag{40}$$

The parameter μ of Muller, Yushenko, and Derjaguin [4,5] is essentially $\sigma^{2/3}$. In this case U and W are also functions of h_0/z_0 .

An ambiguity arises in comparing the results of calculations using these relatively realistic interactions with the classical Hertz and JKR theories. The latter assume a precisely flat contact region, of radius a, with zero separation between the surfaces. In practice, the contact re-

gion is not exactly flat, and the surfaces nowhere contact at h=0, because of the large repulsion present at small separations. However, both the total load and the central displacement are well defined in all approaches, and this will be the primary comparison between them.

III. RESULTS

Figure 1 compares Hertz and JKR theories and the exact results for finite-ranged attractive and repulsive surface forces. As noted above, the total load and the central displacement are appropriate quantities that are free of ambiguities in the definition of the zero of surface separation or of the contact radius. In general for a given load, the displacement for adhesive surfaces is greater than that for repulsive or hard contact. This is because the pressure profile in the adhesive case contains both repulsive (compressive) and attractive (tensile) regions, and hence for the same load the repulsion and the displacement on the central axis must be greater for adhesive surfaces. One can see that both Hertz and JKR theories are relatively accurate in these cases. The finite range of the Lennard-Jones potential can be seen from the negative displacement of the surfaces before they jump into contact. Neither of these two effects is contained in JKR theory. The greatest negative load in the adhesive case is $F_a = -59 \,\mu\text{N m}^{-1}$. It follows that in this case JKR is much more accurate $(-56 \,\mu\,\text{N m}^{-1})$ than is DMT $(-75 \mu \text{Nm m}^{-1})$. Note that the numerical solu-

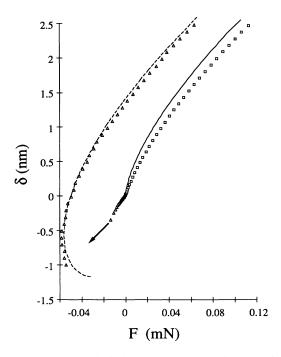


FIG. 1. The central displacement vs the applied load for two identical crossed cylinders of radius R=1.5 cm and elastic constant $E/(1-v^2)=10^{10}$ J m $^{-3}$. The solid curve is Hertz theory and the dashed curve is JKR theory, $\gamma=0.80$ mJ m $^{-2}$. The squares represent the exponential repulsion ($P_0=250$ M N m $^{-2}$, $\kappa^{-1}=0.2$ nm), and the triangles represent the Lennard-Jones force law ($A=10^{-20}$ J, $z_0=0.5$ nm). Note the jump in the latter data (arrow).

tion of the elastic equations correspond to controlled displacement, and hence the surfaces jump apart under some tension less than the maximum. On this scale the outward run was indistinguishable from the inward run.

A. Repulsive contact

Figure 2 tests Hertz theory, Eq. (6), and the slowly varying deformation approximation, Eqs. (31) and (32), for two decay lengths of the repulsive force law, and for the repulsive part of the Lennard-Jones potential [cf. Eq. (33)]. It may be seen that Hertz theory becomes more accurate for more rapidly decaying forces, and is quite good for the short-ranged Lennard-Jones repulsion. In general, Hertz theory overestimates the deformation caused by a given load. The inset to the figure shows that the new approximation given here is rather more accurate than Hertz theory in the small load regime, particularly for the more slowly decaying surface force law.

Hertz theory becomes relatively more accurate as the load is increased. This is explicitly illustrated in Fig. 3, where the relative error in the Hertz prediction for the normalized total load at a given (normalized) central displacement is shown as a function of the parameter Λ , Eq. (39). The error decreases as Λ increases (large loads, short-ranged forces, small curvatures, and soft bodies), in agreement with earlier results [6,7]. Also shown in the figure is the error in the slowly varying deformation approximations, Eqs. (31) and (32). It may be seen that at small values of Λ it is more accurate than Hertz theory.

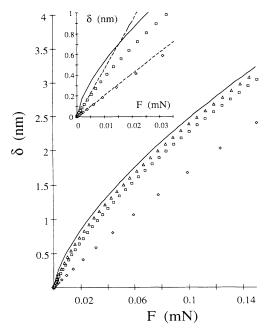


FIG. 2. The central displacement vs the applied load for R=1.5 cm and $E/(1-v^2)=10^{10}$ J m⁻³. The solid line is Hertz theory. The diamonds and squares are the exponential repulsion ($\kappa^{-1}=0.2$ and 1.0 nm, respectively, $P_0=250$ M N m⁻²). The triangles are the repulsive part of the Lennard-Jones pressure, $[p_s(h)=Az_0^6/6\pi h^9, A=10^{-20}$ J, $z_0=0.5$ nm). Inset: comparison with the slowly varying approximation, Eqs. (31) and (32), (dashed lines), for small loads.

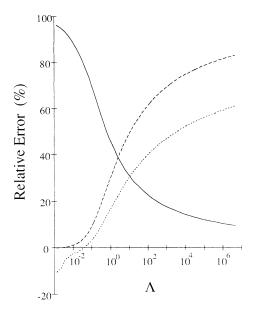


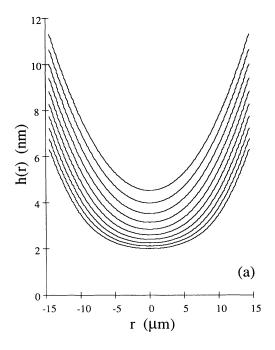
FIG. 3. The relative error in various approximations for the exponential repulsion as a function of the parameter Λ , Eq. (39). The solid line is Hertz theory for the applied load at a given central displacement, the dotted line is the displacement approximation, Eq. (31), and the dashed line is the Derjaguin approximation, Eq. (32).

The Derjaguin approximation, Eq. (32), always underestimates the load because of surface flattening. This is because flattened surfaces are almost everywhere closer together than undeformed surfaces at the same central separation h(0), and hence the deformed bodies experience a larger repulsion than bodies with the smaller, unstressed curvature R. The approximation (31) underestimates the displacement to a lesser extent than the Derjaguin approximation underestimates the load, which is why the approximate curves lie above the exact results in the inset to Fig. 2. In contrast to the Hertz approach, the approximations become worse for increasing loads (or for shorter-ranged forces), mainly because of the significant surface flattening that occurs and which they do not account for.

Figure 3 should be useful as a quantitative guide to the validity of the application of Hertz theory in any particular measurement. For example, Λ was of the order of 10^4 in the experiments of Horn, Israelachvili, and Pribac [12]. Hence their estimate of Young's modulus for their system using Hertz theory is probably accurate to about 10-20%.

Figure 4(a) shows the deformation of the surfaces for increasing loads. The crossed-cylinder geometry is equivalent to a sphere against a plane; the reason that the undeformed profiles appear parabolic is that two different scales are used on the two axes. The scale, with the separation in nm, and the radial coordinate in μ m, corresponds to that typically observed in the interferometric spectrometer used in the surface force apparatus [12]. Each curve is a cross section of the cylindrically symmetric system. The ten curves are in equal increments of

 h_0 , and represent increasing loads from top to bottom. It may be seen that the surfaces become less curved by comparison with their undeformed shape at large separations. In the cases shown, there is no region where the surfaces could be described as strictly flat, or as in contact at h=0. This is inconsistent with the Hertz model of con-



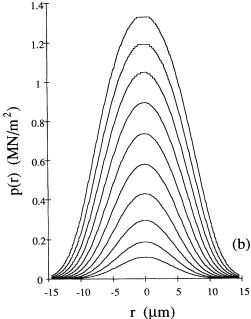


FIG. 4. (a) The surface deformation for the exponentially decaying repulsion ($P_0 = 10 \text{ M N m}^{-2}$, $\kappa^{-1} = 1 \text{ nm}$). h(r) is the local surface separation at a distance r from the central axis. The undeformed geometry (evident at large separations) is for crossed cylinders $[R = 1.5 \text{ cm}, E/(1-v^2) = 10^{10} \text{ J m}^{-3}]$. The profiles are given at equal 0.7 nm increments in h_0 , from $h_0 = 4.3$ to -2.0 nm. Note the different scales on the two axes. (b) The pressure profiles corresponding to (a).

tact deformation, and means that one cannot satisfactorily compare its predictions for the contact radius as a function of load or central displacement. For higher applied loads, or shorter decay lengths, the surfaces do become flatter, although there is still no sharp microscopic delineation of a contact zone. This distinction may be somewhat academic, since one might discern a flat contact region in practice, depending on the characteristics and resolution of the optical system employed. Even though there is this qualitative contradiction in the Hertz model of the contact region, the prediction for the displacement as a function of load can be quite good, as seen above.

Figure 4(b) shows the pressure profiles corresponding to Fig. 4(a). There is qualitative agreement with the Hertz assumption, namely that the pressure is highly compressive in the center, and that it decays monotonically to zero away from the axis.

B. Low and moderate adhesion

Turning now to the adhesive surface, Fig. 5 compares JKR theory with the numerical calculations for the Lennard-Jones surface pressure, Eq. (33), for several values of the parameter σ , Eq. (40). It can be seen that in these cases (σ < 1), JKR theory is very good, predicting the force-deformation relation quite accurately. Hertz theory would lie just below the results for the lowest adhesion. Only the exact result with the largest value of

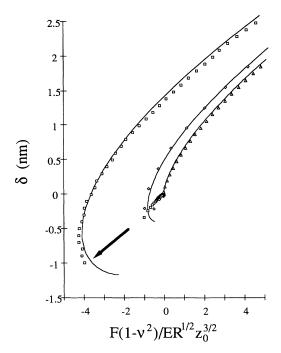


FIG. 5. The central displacement as a function of applied load for adhesive surfaces. The squares, diamonds, and triangles are for σ =0.87, 0.17, and 0.0087, respectively, which correspond to Hamaker constants of 10^{-19} , 2×10^{-20} , and 10^{-21} J, for an elasticity of $E/(1-v^2)=10^{11}$ J m⁻³, a radius of R=1.5 cm, and an equilibrium separation of z_0 =0.5 nm [cf. Eq. (40)]. The solid curves are the corresponding JKR predictions.

 σ exhibits a jump. In this case there is significant deformation before the jump (see below). On the scale of the figure, the inward and outward cycles gave identical results, including for the jump, and no hysteresis is visible.

Surface profiles are shown in Fig. 6(a). Prior to the jump into contact, one can see that the surface becomes elongated, and the curvature increases. After the jump there is a neck, due to that part of the surface under ten-

sion. In this figure there seems to be a well-defined, flattened contact region after the jump, apparently consistent with the JKR assumption. The inset to the figure compares the JKR prediction for the contact radius to that derived from the profiles. Note that the unloading radii are slightly larger than those that occur for loading at the same total force. To obtain the data for the finite-ranged Lennard-Jones interaction, the surfaces were

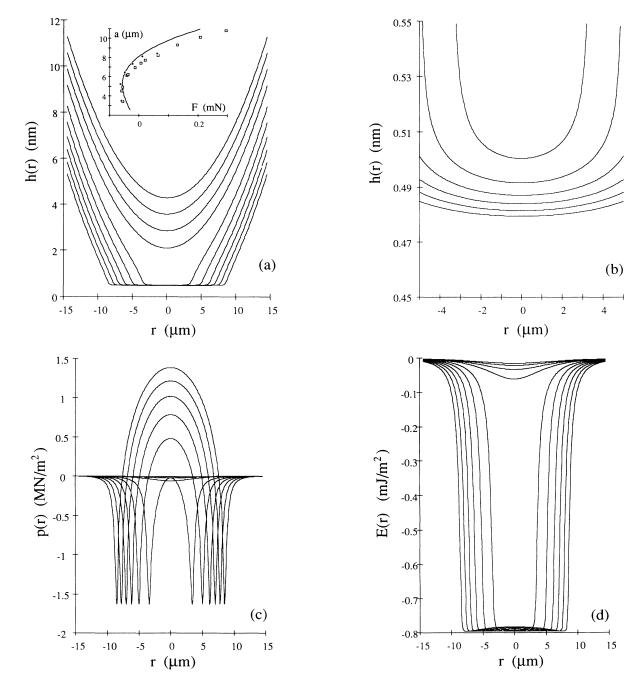


FIG. 6. (a) Surface separation profiles for equal 0.7 nm increments in h_0 (from 4.3 to -2.0 nm) for the Lennard-Jones surfaces $[A=10^{-19} \text{ J}, z_0=0.5 \text{ nm}, E/(1-v^2)=10^{11} \text{ J m}^{-3}, R=1.5 \text{ cm}]$. Inset: contact radius vs the applied load. The solid line is the JKR approximation, and the symbols represent exact results for loading (open symbols) and unloading (closed symbols). For the finite-ranged potential, contact was defined to occur when $h(r) < 1.1z_0$. (b) Expanded view of the deformed surfaces of (a). (c) The pressure distributions that correspond to (a). (d) The interaction free energy per unit area across the contact region corresponding to (a).

defined to be in contact whenever they were within 10% of the equilibrium spacing. On an expanded scale [Fig. 6(b)] it is clear that the surfaces remained curved, and that this or any definition of the contact radius is somewhat arbitrary.

The corresponding pressure profiles are shown in Fig. 6(c). Note the difference before and after contact. In the latter case the pressure is increasingly repulsive toward the center. Near the edge of the contact region the pressure becomes attractive. The regime over which this change occurs is quite small, leading to a very sharp change in the pressure. There is qualitative agreement between this pressure distribution and that assumed by JKR theory, Eqs. (3) and (7), although the pressure never becomes infinite as it does at the edge of the contact region in that theory. The pressure profiles for the Lennard-Jones surfaces contrast with the much smoother distribution for the exponential repulsion [Fig. 4(b)].

The interaction free energy per unit area is shown in Fig. 6(d). Again one can imagine a contact region, defined by some convention, in which the energy is approximately constant. Here the energy is close to its minimum possible value, which is the negative of the surface energy $\gamma = 0.80 \text{ mJ/m}^2$. The figure suggests that the JKR assumption of an energy equal to $-\gamma$ in the contact region, and zero outside of it, is not too unrealistic.

Figure 7 shows the pull-off force for controlled load (maximum tension) as a function of the parameter σ , Eq. (40). The force has been normalized so that JKR theory corresponds to unity, and DMT to $\frac{4}{3}$ (arrow). Small values of σ correspond to stiff bodies, of small radius and low-surface energies. One can see that in this regime DMT theory is more accurate than is JKR, as has been

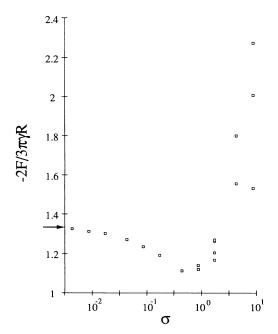


FIG. 7. The pull-off force (maximum tension for unloading) divided by the JKR prediction. The arrow indicates the DMT result. The several results for each of the larger values of the parameter correspond to different penetration depths.

found earlier [4,5]. For values of $\sigma \lesssim 1$, the accuracy of JKR theory for the pull-off force improves. Neither theory is particularly appropriate for larger values of the parameter. In this regime the pull-off force in the exact calculations has been taken to be the maximum force of each of the unloading cycles (see below).

Both JKR and DMT theories predict that the pull-off force scales linearly with the surface energy, and that it is independent of the elasticity. The first statement seems intuitively obvious, but the second is perhaps surprising. However, if the pull-off force is indeed linear in the surface energy (as Fig. 7 confirms with relative accuracy over several orders of magnitude), then the minimum of the dimensionless force, Eq. (38), must be a linear function of the parameter σ . Since the dimensionless force and σ both contain Young's modulus in the denominator, the pull-off force itself has to be independent of the elasticity whenever it is linear in γ . The same argument also shows that here it scales linearly with R. The assumption of linearity clearly breaks down at larger values of σ , Fig. 7, and here the pull-off force can no longer be independent of the elasticity.

C. Deformation and jumps prior to contact

One reason that JKR and DMT theories fail in the large-adhesion, compliant-body limit is that significant surface deformation occurs before the surfaces contact. This is due to the finite range of the attractive force. In many cases the surfaces actually jump into contact, as has been predicted previously [8,9]. The jumps represent a spontaneous, nonequilibrium transition in which energy is lost from the system in the form of heat [13]. The classical theories do not describe these phenomena, but the approximate expressions for the displacement, load, and jump, Eqs. (35)–(37), should be applicable.

Figure 8 tests the full approximate expressions for the force and displacement, Eqs. (35a) and (36a), against the exact results prior to contact, for $\sigma = 8.7$. The analytic results obtained using only the attractive part of the Lennard-Jones potential, Eqs. (35b) and (36b), are also shown, and these are in very good agreement with the results of the approximation using the full potential. It may be seen that the simple approximation is remarkably accurate, and that it is valid to neglect the repulsive part of the interaction at these separations.

The jump criterion, Eq. (29) [evaluated with the displacement and load approximations Eqs. (35) and (36)], is tested in Fig. 9 as a function of the parameter σ , Eq. (39). The analytic result that uses only the attractive part of the Lennard-Jones potential, Eq. (37), is a good representation of the full approximation. For the exact results the tension at the point prior to the jump is given. For small values of σ where jumps do not occur or are difficult to discern, the maximum tension is plotted. It can be seen that the approximations give a quite acceptable estimate of the jump.

The results confirm the validity of the stability criterion (29), which may otherwise have been criticized because its derivation was less than rigorous. The local condition (25) is not exact, since in general the change in

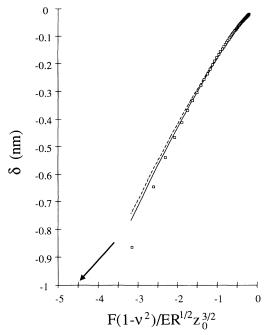


Fig. 8. A test of the deformation and load approximations, Eqs. (35) and (36), prior to the jump into contact [$A = 10^{-19}$ J, $z_0 = 0.5$ nm, $E/(1-v^2) = 10^{10}$ J m⁻³, R = 1.5 cm]. The solid curve is the complete Lennard-Jones potential, and the dashed curve is due to the attractive part only. The symbols are the exact results. The curves end at the point predicted by the stability criterion, and the final symbol represents the point in the exact calculations just prior to the jump.

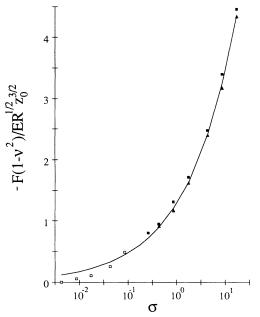


FIG. 9. The tension just prior to the jump into contact as a function of the parameter σ , Eq. (40). The squares are the exact result, and the triangles are the stability criterion, Eq. (29), using the deformation and load approximations, with the full potential, Eqs. (35a) and (36a). The solid line is the analytic expression (37), which uses only the attractive part of the Lennard-Jones pressure, and which is continued when no jump occurs. Open symbols denote the maximum tension in cases of low adhesion when no jump occurs.

deformation depends upon the perturbation of the whole surface. Equation (28) must be violated by parts of the surface just beyond the contact zone once the surfaces have jumped into contact (cf. Fig. 6). Despite this lack of rigor in the derivation of Eq. (29), Fig. 9 shows it to be a very good approximation.

D. Hysteresis

For soft bodies with high surface energies (large σ), the unloading becomes hysteretic (Fig. 10). These cases also pose greater numerical challenges, as is evident from the scatter in the data in the figures. The various runs (the loading cycles are coincident) have different penetration depths, which is the maximum displacement before unloading begins. In contrast to results given above, the unloading cycles agree neither with the inward run nor with each other. In all cases the maximum tension is much greater than is the tension immediately following the jump into contact, or anywhere on the loading cycle. For the smaller value of σ one sees that eventually the unloading paths agree, even though initially they depend upon the penetration depth. In this case the maximum tension was independent of the penetration depth, at least within the numerical noise. For larger values of the parameter, the unloading cycles do not coincide until after the point of maximum tension. This means that the pull-off force is greater for larger penetrations, and hence that there is not a unique pull-off force for a given surface

It can be seen that JKR theory is not really applicable in this hysteretic regime. That theory approximately bisects the loading and unloading cycles, lying somewhat closer to the former. JKR theory is only trivially hysteretic, since the cycles coincide everywhere except that the point of separation is beyond the point of first contact. It is somewhat surprising that JKR theory fails here, since that theory was expected to become more accurate for compliant bodies with large adhesions [4,5].

This predicted hysteretic behavior is qualitatively similar to the experimentally measured behavior of adhesive surfaces using an indenter [11], the surface-forces apparatus [12-14], and the atomic-force microscope [15,16]. A consequence of the hysteresis is that the maximum tension can depend on the past history of the sample, namely the largest load that was applied, or equivalently the penetration depth (i.e., the largest value of the central displacement prior to unloading). The measured pull-off force shown in Fig. 9 of Ref. [13] is indeed larger for greater maximum applied loads. In the hysteretic regime there is no simple direct relationship between the pull-off force and the surface energy of the bodies, such as in JKR or DMT theory; one needs to include the penetration of the samples in the equation.

The hysteresis appears to occur for values of the dimensionless variable σ greater than about unity. For the measurements of Horn, Israelachvili, and Pribac [12] of the adhesion of mica in air, $\sigma \approx 10^3$ ($A \approx 10^{-19}$ J, $z_0 \approx 10^{-1}$ nm, $\gamma \approx 10^2$ mJ/m², $R \approx 10^{-2}$ m, $E \approx 10^{10}$). This means that the application of JKR theory to their measurements is probably not valid because of hysteresis,

which in fact they observe [12].

The hysteresis is due to the jumps, as is illustrated in Fig. 11. Precritically, one can think of each element of surface as a simple spring. The relevant point is that the position the spring jumps to always lies at a smaller surface separation than the position the spring jumps from.

Mathematically, it is a consequence of the mean-value theorem. Physically, it happens because at the start and end of the jump the spring extension balances the applied force, but in the region between the jump starts the applied force changes more rapidly than the force due to the change in length of the spring. Hence for an inward

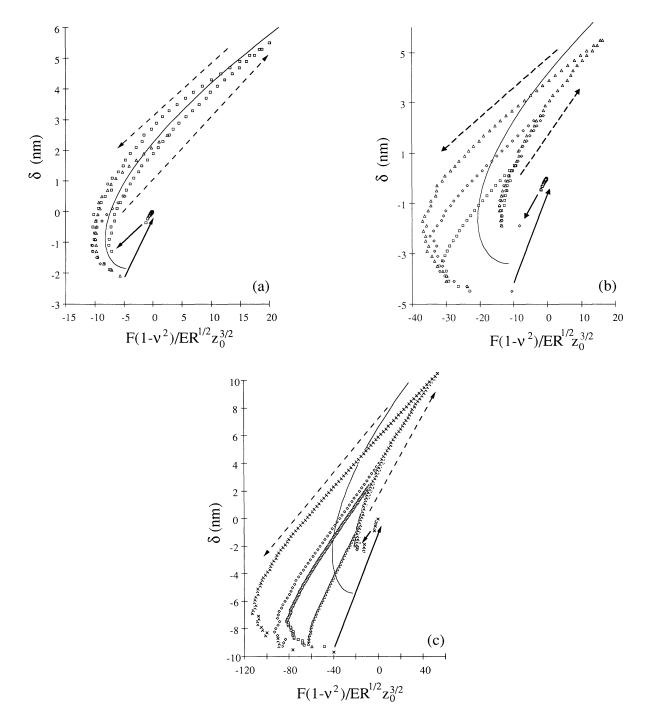


FIG. 10. (a) Displacement vs load in the case of hysteresis, $\sigma = 1.7$. The various symbols represent cycles with different penetration depths (maximum displacements), in this case $h_0 = 0$, -2, and -5 nm. The arrows indicate the direction of the loading and unloading. The solid line is the JKR prediction, which is not hysteretic. Note the inward jump, and the different outward jumps (solid arrows indicate jumps). (b) Hysteresis in the case of $\sigma = 4.4$. The penetrations are $h_0 = 0$, -2, and -5 nm. (c) Hysteresis in the case of $\sigma = 8.7$. The penetrations are $h_0 = 1$, -2, -4, and -10 nm.

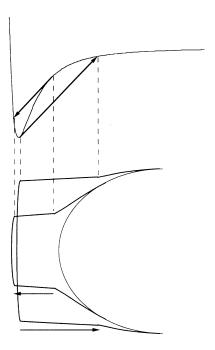


FIG. 11. Sketch of the jumps for a simple spring in a force field (upper). The arrows commence at the points where the gradient of the force equals the spring constant (the start of the jump), and point to where the applied force once more equals that due to the spring extension (the end of the jump). The region between the jump starts is inaccessible to the spring, and the jump ends always bracket the jump starts for this type of smooth attraction. The lower part of the figure is a schematic illustration of the corresponding surface deformation for loading and unloading at a given position of the undeformed body (the parabola).

jump it is not until a smaller separation than the outward jump start, where the applied force increases less rapidly than that due to the spring, that they can once more achieve a balance. The same argument applies to the outward jump.

The fact that hysteresis occurs for this simple spring indicates that it also occurs for elastic bodies which jump into adhesive contact. It is obvious that hysteresis will occur for two infinite planar surfaces, but the interaction of two convex bodies is more complicated. Because each element of surface is not independent, not only is the contact area hysteretic, but also the surface profile in the contact region. Nevertheless, the area of contact for the loading cycle is always less than the area of contact for unloading at a given central displacement (cf. the inset to Fig. 6(a); see also the measured data in Figs. 6 and 7 of Ref. [12]). As the illustration in the lower half of Fig. 11 illustrates, the separation at the edge of the contact region corresponds to the jump-in terminus, and to the jump-out origin, for loading and unloading, respectively. The surface separation just outside the contact region corresponds to the jump-in origin and to the jump-out terminus, for loading and unloading, respectively. Consequently the contact radius is smaller for loading than for unloading. The force, which is approximately the minimum of the pressure times the contact area, must be

larger for unloading than for loading at a given h_0 . For small jumps, the consequent difference in loads is not visible on the figures (cf. Fig. 5), but it was confirmed that the actual value of the unloading tension was always greater than or equal to the loading tension at a given displacement. For large adhesions and soft bodies, the hysteresis is quite marked (Fig. 10).

The fact that the hysteretic surface shape and deformation represents a jump instability means that experimental measurements can be time dependent. This is due to the finite probability that a mechanical vibration will occur that changes the contact profile and area because surface elements at the edges of the contact region jump in or out. Thus, compared to a fast experiment, where vibrations are less likely to occur during the measurement, the contact area is larger during slow loading (because here it is the separated surface elements, which are just beyond the contact region, and which are perched on an instability that will jump into contact) and it is smaller during slow unloading (because here it is the surface elements in contact at the edges of the contact region that are unstable and jump out). This time-dependent behavior has been observed experimentally (Fig. 8(a), Ref. [13]).

IV. CONCLUSION

This paper has been concerned with the deformation and adhesion of elastic bodies in contact. The emphasis has been on the finite range that all real surface forces possess. To this end a comparison has been made with the classical theories—Hertz theory for repulsive forces, and JKR and DMT theories for the case of adhesion—that assume infinitely short-ranged contact forces. Hertz theory was found to be accurate in the case of short-range surface forces and high loads [6,7], and JKR theory predicted the deformation-load relationship accurately up to moderate adhesions. For low adhesions, DMT theory was more accurate than JKR theory in relating the pull-off force to the surface energy [4,5].

As the adhesion and the compliance of the bodies increase, significant deformation can occur away from contact, and they can actually jump into contact due to the finite range of the surface forces [8,9]. This precontact deformation is not predicted by the classical theories, and here a simple but accurate approximation was given. A stability criterion for the jump was derived, and it shows that these occur for large values of the product of the derivative of the pressure (rapidly increasing attractions) and the ratio of the central displacement to the pressure (soft bodies).

For high adhesions and/or soft bodies, it was found that the displacement-load cycles were hysteretic, and that the actual value of the pull-off force could depend upon the amount of the preceding deformation. Accordingly, in this regime there is not a unique relationship between the pull-off force and the surface energy, and neither JKR nor DMT theory is applicable.

The observation of hysteresis in this simple continuum model with Lennard-Jones adhesion is quite interesting. The origin ascribed to it here—surface jump instabilities of compliant bodies due to rapidly varying attractions—

is a generally occurring mechanism. In the past, hysteresis has been attributed to molecular rearrangements, contamination and natural variability, surface heterogeneity and roughness, or viscoelasticity and plasticity. While undoubtedly these will have an effect if they occur, the present results show that it is not necessary to invoke these complicated phenomena in order to explain the experimental observation of hysteresis [11–16]. The simplest explanation is that hysteresis occurs whenever there are elastic instabilities associated with rapidly varying but finite-ranged attractions between the surfaces.

Hysteresis is also a likely explanation of the variability in measurements of the surface energy from the pull-off force [11–13], since it was shown here that the maximum tension increased with the sample penetration (the maximum load previously applied). It is worth noting that hysteresis also occurs in other well-known methods for determining surface energies, such as contact angle measurements, where the advancing angle (loading) is found to exceed the retreating angle (unloading). The present results are qualitatively relevant to this case, since the drop is elastic (in the sense that energy is required to deform its shape) and it is in contact with an elastic substrate. A theoretical analysis of the effect on the contact angle of elastic deformation of the substrate has been given [23], and possible mechanisms for contact angle hysteresis have been reviewed [24]. The jump-instability mechanism for loading and/or unloading that is discussed here represents another contribution to contact angle hysteresis. A direct quantitative application of the results obtained here to the drop problem is inappropriate because the elastic equations solved here incorporate shear stress, which a liquid cannot support. However, in the present study and in the case of drops on a substrate, the surface profile arises from the balance of the surface forces against the change in energy due to deformation, and the present conclusions about the origin of hysteresis are at least qualitatively applicable.

Tribology and friction represent other possible fields where extensions of the present analysis of adhesion and deformation due to finite-ranged surface forces may be applied. The physical origin of friction can be puzzling, particularly if one thinks in terms of equilibrium mechanics. However, the mutual sliding and rolling of bodies involves the joining and separation of surface elements before and after the area of contact. The present results indicate that even for perfectly smooth surfaces, for compliant adhesive bodies the process is not symmetrical, and it requires energy, which is manifest as a lateral force. Heat is created in traversing the hysteresis loop, due to the spontaneous, irreversible jumps into and out of contact of the surface elements.

Beyond these extensions to contact angles and to friction, the analysis of the loading and unloading of elastic bodies is relevant to any situation which involves the movement of a line of three-phase contact. In a forth-coming paper the authors will discuss in detail the effect of elastic deformations on the direct measurement of surface forces, and quantify the errors in the measurements when these effects are ignored. Recently techniques have been developed for measurement of load-displacement curves using force feedback [25], and it is intended to carry out high-precision tests of the present theories of elastic deformation and adhesion.

^[1] H. Hertz, J. Reine Angew. Math. 92, 156 (1881); *Miscellaneous Papers* (Macmillan, London, 1896), p. 146.

^[2] K. L. Johnson, K. Kendall, and A. D. Roberts, Proc. R. Soc. London Ser. A 324, 301 (1971).

^[3] B. V. Derjaguin, V. M. Muller, and Yu. Toporov, J. Colloid Interface Sci. 53, 314 (1975).

^[4] V. M. Muller, V. S. Yushchenko, and B. V. Derjaguin, J. Colloid Interface Sci. 77, 91 (1980).

^[5] V. M. Muller, V. S. Yushchenko, and B. V. Derjaguin, J. Colloid Interface Sci. 92, 92 (1983).

^[6] B. D. Hughes and L. R. White, Quart. J. Mech. Appl. Math. 32, 445 (1979).

^[7] B. D. Hughes and L. R. White, J. C. S. Faraday I 76, 963 (1980).

^[8] J. B. Pethica and A. P. Sutton, J. Vac. Sci. Technol. A 6, 2490 (1988).

^[9] J. R. Smith, G. Bozzolo, A. Banerjea, and J. Ferrante, Phys. Rev. Lett. A 63, 1269 (1989).

^[10] J. L. Parker and A. M. Stewart, Prog. Colloid Polym. Sci. (to be published).

^[11] M. D. Pashley and J. B. Pethica, J. Vac. Sci. Technol. A 3, 757 (1985).

^[12] R. G. Horn, J. N. Israelachvili, and F. Pribac, J. Colloid Interface Sci. 115, 480 (1987).

^[13] Y. L. Chen, C. A. Helm, and J. N. Israelachvili, J. Phys. Chem. 95, 10736 (1991).

^[14] P. Attard and J. L. Parker, J. Phys. Chem. 96, 5087 (1992).

^[15] N. A. Burnham and R. J. Colton, J. Vac. Sci. Technol. 7, 2906 (1989).

^[16] N. A. Burnham, D. D. Dominguez, R. L. Mowery, and R. J. Colton, J. Phys. Lett. 64, 1931 (1990).

^[17] K. L. Johnson, Contact Mechanics (Cambridge University, Cambridge, England, 1985).

^[18] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 2nd ed. (Pergamon, London, 1970).

^[19] B. V. Derjaguin, Kolloid Z. 69, 155 (1934).

^[20] L. R. White, J. Colloid Interface Sci. 95, 286 (1983).

^[21] P. Attard, D. R. Bérard, C. P. Ursenbach, and G. N. Patey, Phys. Rev. A 44, 8224 (1991).

^[22] P. Attard, D. J. Mitchell, and B. W. Ninham, J. Phys. Chem. 89, 4358 (1988).

^[23] G. R. Lester, J. Colloid Interface Sci. 16, 315 (1961).

^[24] P. G. de Gennes, Rev. Mod. Phys. 57, 827 (1985).

^[25] J. L. Parker, Langmuir 8, 551 (1992).