Contact Mechanics Analysis of Elastic-Plastic Rough Surfaces Characterized by Fractal Geometry

W. Yan, Graduate Student

K. Komvopoulos, Associate Professor (Mem. ASME)

Department of Mechanical Engineering
University of California
Berkeley, CA 94720

Abstract

Random rough surfaces are characterized by fractal geometry using a modified two-variable Weierstrass-Mandelbrot function. This surface model is incorporated into an elastic-plastic contact mechanics analysis of two approaching rough surfaces, and closed form solutions for the elastic and plastic components of the normal contact force and real contact area are derived in terms of fractal parameters, material properties, and mean surface separation distance. The effects of surface topography and material properties on the total deformation force are investigated by comparing results from two-dimensional and three-dimensional contact mechanics analyses and elastic and elastic-plastic material behaviors. The developed algorithm yields three-dimensional fractal surfaces that are representative of engineering rough surfaces. For elastic-plastic silica surfaces in normal contact and the range of surface interference examined, the interfacial force is predominantly elastic and the real area of contact is approximately one percent of the apparent area of contact or less, depending on the mean interfacial distance.

Submitted to the ASME *Journal of Applied Mechanics* on May 1, 1997.

1 Introduction

The topography of engineering surfaces has been traditionally considered to be a stationary random process (Nayak, 1971). Consequently, statistical parameters of the surface height function, such as the root-mean-square (rms) of the surface height (σ), slope (σ'), and curvature (σ''), have been used for topography characterization (Greenwood and Williamson, 1966; Nayak, 1973). Greenwood and Williamson (1966) analyzed elastic contact of two rough surfaces by considering an equivalent rough surface comprising spherical asperities of constant radius (equal to the average radius of curvature of the original asperities) in normal contact with a flat surface. Assuming a Gaussian distribution for the asperity heights and uniform distribution of asperities over the projected in-plane area according to a known areal density, the number of asperity contacts, mean microcontact area, real contact area, and interfacial distance were obtained in terms of the normal load. Nayak (1973) considered rigid-plastic contact of two Gaussian rough surfaces and noticed that the Greenwood-Williamson model was not applicable to all microcontact sizes; hence, the rough surfaces were represented by a narrow-band rough surface and a wide-band smooth surface. A review of classical elastic-plastic contact models for isotropic and anisotropic surfaces presented by McCool (1986) has indicated that surface characterization has been primarily based on the surface height distribution function and values of σ, σ', and σ''.

Experiments by Sayles and Thomas (1978) have shown that the surface topography is a nonstationary process, and that the corrugation of the surface height may contain a broad bandwidth. In addition, the determination of σ, σ', and σ'' depends on the sample size and instrument resolution or any other experimental filter used in topography data acquisition. For example, in the case of fractal surfaces, σ' and σ'' depend on the instrument resolution but not on the sample size, whereas σ depends on the sample size but is insensitive to the instrument resolution (Majumdar and Bhushan, 1995). In view of the dependence of previous classical models on scale-dependent parameters, to perform objective contact mechanics analyses of rough surfaces, it is essential to develop surface models that are invariant at all length scales.

The power spectra of engineering surfaces produced by random processes, such as cleavage, solidification, vapor deposition, and directionally unbiased machining, have been observed to follow inverse power laws over several decades of length scales (Majumdar and Tien, 1990). This is an inherent behavior of fractal geometry revealing its potential to represent surface fea-

tures from the microscale to the nanoscale. Fractal geometry has been recently introduced in the field of contact mechanics to perform scale-independent analyses. Majumdar and Bhushan (1995) reviewed the traditional and fractal contact mechanics models and attempted to describe both fractal and nonfractal surfaces by using the structure function and magnitude of σ'' determined in two orthogonal in-plane directions. However, this method is relatively complex and is based on difficult-to-obtain experimental data.

An especially difficult problem in contact mechanics analyses of rough surfaces is the change of the surface shape during deformation. Among the few models dealing with this problem is the rigid-perfectly plastic contact model of Warren and Krajcinovic (1995) and Warren et al. (1996) that uses a deterministic Cantor set to represent the effective fractal surface and accounts for volume conservation. Assuming that the material truncated by the opposing smooth surface flows into the valleys of the noncontacting regions of the rough surface, analytical results for the contact load that were supported by experimental measurements were obtained for relatively small surface interference; however, discrepancies between analytical and experimental results were found at relatively large surface interference.

The main objectives of this publication are to introduce a comprehensive contact mechanics analysis of elastic-plastic rough surfaces characterized by fractal geometry and to present numerical results for the variation of the interfacial contact force and real contact area during quasi-static approach of the surfaces that elucidate the significance of fractal parameters and material properties in the deformation at the contact interface

2 Fractal Surface Characterization

Fractal geometry, pioneered by Mandelbrot (1982), is observed in various natural phenomena, such as precipitation, turbulence, and surface topography, and is characterized by continuity, nondifferentiability, and self-affinity. These mathematical properties are satisfied by the Weierstrass-Mandelbrot (W-M) function (Berry and Lewis, 1980) given by

$$w(x) = \sum_{n=-\infty}^{\infty} \gamma^{(D-2)n} (1 - e^{i\gamma^n x}) e^{i\phi_n} \tag{1}$$

where w is a complex function of the real variable x. A fractal profile $z(x)$ can be obtained as the real part of $w(x)$,

$$z(x) = \mathrm{Re}\ \left[w(x)\right] = \sum_{n=-\infty}^{\infty} \gamma^{(D-2)n}\left[\cos\phi_n - \cos(\gamma^n x + \phi_n)\right] \tag{2}$$

where D ($1 < D < 2$) is the fractal dimension of the profile, ϕ_n is a random phase, and γ ($\gamma > 1$) is a parameter that determines the density of frequencies in the profile. The right hand side of Eq. (2) is a superposition of a series of cosine functions with geometrically increasing frequencies. The random phase ϕ_n is introduced to prevent the coincidence of different frequencies at any point of the surface profile. Based on surface flatness and frequency distribution density considerations (Komvopoulos and Yan, 1997), γ is set equal to 1.5 in the present study. The approximate continuous power spectrum, $P(\omega)$, of $z(x)$,

$$P(\omega) = \frac{1}{\omega^{(5-2D)} \ln\gamma}, \tag{3}$$

is an inverse power function of the spatial frequency, ω, which has been found to hold for many engineering surfaces.

Ausloos and Berman (1985) generalized the W-M function (Eq. (1)) by introducing multiple variables to account for higher-dimension stochastic processes. In this generalization, the homogeneity and scaling properties of the single-variable function are preserved, i.e., the mean square increment of function $w(x)$, $V(\tau) = \left\langle \left| w(x+\tau) - w(x) \right|^2 \right\rangle$, remains independent of x and $V(\gamma\tau) = \gamma^{2(2-D)}V(\tau)$ still holds. Two-variable functions can be used to model fractal surfaces exhibiting corrugations in all directions. Such a function satisfying the requirements of homogeneity and scaling has been constructed by Ausloos and Berman (1985) by using a weighted random phase superposition of ridge-like surfaces with cross sections similar to that given by Eq. (1). The height function of a fractal surface exhibiting randomness in all planar directions, hereafter referred to as a three-dimensional (3-D) surface, can be obtained as the real part of the Ausloos-Berman function,

$$z(\rho,\theta) = \left(\frac{\ln\gamma}{M}\right)^{1/2} \sum_{m=1}^{M} A_m \sum_{n=-\infty}^{\infty} \left(k\gamma^n\right)^{(D-3)} \left\{\cos\phi_{m,n} - \cos\left[k\gamma^n \rho\cos\left(\theta - \alpha_m\right) + \phi_{m,n}\right]\right\} \tag{4}$$

where ρ and θ are the planar polar coordinates of a point on the surface with height z, and are related to the planar Cartesian coordinates, x and y, by

$$\rho = \left(x^2 + y^2\right)^{1/2} \qquad \text{and} \qquad \theta = \tan^{-1}\left(y/x\right),$$

γ has the same physical meaning and magnitude as in Eq. (2), and D ($2 < D < 3$) is the fractal dimension of the surface. The physical significance of D is the extent of space filling by the rough surface, with larger values corresponding to denser fillings. For isotropic surfaces, the value of D can be determined from the slope of the log-log plot of the power spectrum $P(\omega)$ given by Eq. (3) (Wang and Komvopoulos, 1994a), and for anisotropic (and isotropic) surfaces by using a method proposed by Gagnepain and Roques-Carmes (1986). In both methods, D is invariant of spatial frequencies. The parameter M represents the number of superposed ridges used to construct the surface. When $M = 1$, surfaces with cylindrical corrugations (hereafter termed two-dimensional (2-D) surfaces) are produced. The anisotropy of the surface geometry is controlled by the magnitude of A_m. For isotropic surfaces, $A_m = A$ for all m values; for anisotropic surfaces, A_m assumes different values for different m's. The arbitrary angle α_m is used to offset the ridges in the azimuthal direction. Since in the present study the ridges are equally offset, $\alpha_m = \pi m/M$. The values of the random phase $\phi_{m,n}$ are uniformly distributed in the interval $[0, 2\pi]$ based on a random number generator (Press et al., 1992). The parameter k is a wavenumber which is related to the sample size by $k = 2\pi/L$. In practice, the frequency index n assumes finite values. Since the lowest frequency of the sample is $1/L$, the lower bound of n is limited by the sample length, L, and can therefore be set equal to zero. In addition, because fractal geometry cannot be extended below the atomic scale, the highest frequency is set equal to $1/L_s$, where L_s is of the order of the atomic distance and is thus assumed equal to 1 Å. Hence, the maximum value of n is

$$n_{\max} = \text{int}\left[\frac{\log\left(L/L_s\right)}{\log\gamma}\right]$$

where int [·] denotes the integer part of the number in the brackets. The surface height function of 3-D isotropic surfaces can be obtained by substituting the above relations of A_m, α_m, ρ, θ, k, and n into Eq. (4),

$$z(x,y) = A\left(\frac{L}{2\pi}\right)^{(3-D)}\left(\frac{\ln\gamma}{M}\right)^{1/2}\sum_{m=1}^{M}\sum_{n=0}^{n_{\max}}\gamma^{(D-3)n}\left\{\cos\phi_{m,n} - \cos\left[\frac{2\pi\gamma^n\left(x^2+y^2\right)^{1/2}}{L}\cos\left(\tan^{-1}\left(\frac{y}{x}\right) - \frac{\pi m}{M}\right) + \phi_{m,n}\right]\right\} \tag{5}$$

Since A, L, and 2π all have length dimensions, a length parameter G can be defined such that $A = 2\pi\left(2\pi/G\right)^{(2-D)}$. Introducing this relationship into Eq. (5) gives

$$z(x,y) = L\left(\frac{G}{L}\right)^{(D-2)}\left(\frac{\ln\gamma}{M}\right)^{1/2}\sum_{m=1}^{M}\sum_{n=0}^{n_{\max}}\gamma^{(D-3)n}\left\{\cos\phi_{m,n} - \cos\left[\frac{2\pi\gamma^{n}\left(x^{2}+y^{2}\right)^{1/2}}{L}\cos\left(\tan^{-1}\left(\frac{y}{x}\right) - \frac{\pi m}{M}\right) + \phi_{m,n}\right]\right\} \tag{6}$$

As can be seen from Eq. (6), G is a height scaling parameter that is independent of the frequency; thus, it is termed the fractal roughness.

Equation (6) will be used to represent a 3-D isotropic fractal surface in the elastic-plastic contact mechanics analysis presented below. This function of the surface height provides a deterministic means for generating stochastic rough surfaces. The only unknown variables in Eq. (6) are the scale-independent fractal parameters G and D, which can be determined experimentally. Therefore, this fractal approach has the inherent capability of representing surfaces at various length scales, different from those at which the measurements were made.

Figure 1 shows a $1\,\mu\text{m} \times 1\,\mu\text{m}$ fractal surface generated from Eq. (6) for D = 2.4, σ = 7 nm ($G = 1.36 \times 10^{-11}$ m), and M = 10. Although a visually random surface is produced when $M > 3$, a circular heterogeneity can be detected in the power spectrum when a relatively small ridge number is used. The homogeneity of the simulated surface shown in Fig. 1 can be evaluated by examining the contour plot of the corresponding power spectrum, shown in Fig. 2, obtained by fast Fourier transform. It is known that isotropic surfaces possess axially symmetric power spectra (Nayak, 1973). With the exception of two orthogonal high intensity bands along the two frequency axes, the power spectrum shown in Fig. 2 exhibits axial symmetry. These artificial bands are due to the unmatchness of the opposing boundaries of the simulated surface, and can be eliminated by adopting a simple technique proposed by Anguiano et al. (1994) that uses the mirror images of a fractal surface with respect to its boundaries to obtain an axially symmetric power spectrum.

3 Elastic-Plastic Contact Model

In previous elastic-plastic contact mechanics models dealing with fractal surfaces possessing single-slope and double-slope power spectra (Majumdar and Bhushan, 1991; Bhushan and Majumdar, 1992), classical Hertz theory was used to determine the elastic contact load at asperity microcontacts, whereas the load at fully plastic microcontacts was assumed to be proportional to the real area of contact. However, a clear distinction between truncated and real microcontact ar-

eas was not provided in these studies. In recent contact mechanics analyses of fractal surfaces, Wang and Komvopoulos (1994a, 1994b, 1995) distinguished and related the truncated and real microcontact areas and considered elastic, elastoplastic, and fully plastic deformation of the asperities in order to analyze frictional heating at sliding surfaces. In all the above models, rough surfaces have been represented by 2-D profiles, which is a reasonable approximation for isotropic surfaces. For anisotropic surfaces, however, a 3-D surface characterization must be used. Thus, the present study is conducted for random 3-D rough surfaces. To simplify the problem, the system of two opposing rough surfaces is replaced by a deformable rough surface with roughness equivalent to the effective roughness of the two original surfaces and effective Young's modulus, $E^* = \left[\left(1-\nu_1^{\,2}\right)/E_1 + \left(1-\nu_2^{\,2}\right)/E_2\right]^{-1}$, where ν_1, ν_2, and E_1, E_2 are the Poisson's ratios and Young's moduli of the original surfaces, respectively, and a flat (smooth) rigid countersurface. For isotropic surfaces, the height function of the equivalent rough surface is given by Eq. (6) and for anisotropic surfaces by Eq. (4). The basic approach for determining the deformation force at the contact interface comprises the calculation of the forces at asperity microcontacts and the integration of microcontact forces using the size distribution function developed by Mandelbrot (1982).

To obtain the deformation force at a microcontact, its interference with the opposing rigid plane must be determined first. However, in view of Eqs. (4) and (6), it is impossible to obtain a simple solution for the asperity interference in terms of the microcontact size due to the summation over *m* and the presence of random phases. It is therefore necessary to assume spherical asperity contacts and derive single-variable equivalent relationships for the two-variable surfaces represented by Eqs. (4) and (6). This can be accomplished by introducing a factor $C(M)$ to account for the superposition of *M* ridges and enforcing the rms roughness of the original 3-D surfaces to be equal to that of the equivalent 2-D surfaces. Although the following discussion is for isotropic surfaces, a similar approach can be adopted to analyze anisotropic surfaces. This will be discussed briefly at the end of this section.

After introducing $C(M)$ and assuming only one ridge, Eq. (6) becomes

$$z(x) = C(M)L\left(\frac{G}{L}\right)^{(D-2)}(\ln\gamma)^{1/2}\sum_{n=0}^{n_{\max}}\gamma^{(D-3)n}\left[\cos\phi_{1,n} - \cos\left(\frac{2\pi\gamma^n x}{L} - \phi_{1,n}\right)\right] \tag{7}$$

Computations showed that, for a typical value of D = 2.5, $C(M)$ assumes values between 0.85 and 1.15 when M varies in the range of 1 to 100. Thus, in view of the small variation of $C(M)$, it may be considered that $C(M) \approx 1$ and Eq. (7) reduces to

$$z(x) = L\left(\frac{G}{L}\right)^{(D-2)} \left(\ln \gamma\right)^{1/2} \sum_{n=0}^{n_{\max}} \gamma^{(D-3)n} \left[\cos \phi_{1,n} - \cos\left(\frac{2\pi \gamma^n x}{L} - \phi_{1,n}\right)\right]$$

Since the right hand side of the above equation is a series of cosine functions, a profile with smaller asperities residing on larger asperities is produced. For an asperity with truncated micro-contact radius r' (see Fig. 3), the longest wavelength in the asperity waveform is $2r'$. Thus, the corresponding frequency index is

$$n_0 = \frac{\ln\left(L/2r'\right)}{\ln \gamma}$$

and the cosine term is

$$z_0(x) = G^{(D-2)} \left(\ln \gamma\right)^{1/2} (2r')^{(3-D)} \left[\cos \phi_{1,n_0} - \cos\left(\frac{\pi x}{r'} - \phi_{1,n_0}\right)\right]$$

The asperity interference, δ, is determined by $z_0(x)$ and is equal to the peak-to-valley amplitude of the cosine function,

$$\delta = 2G^{(D-2)} \left(\ln \gamma\right)^{1/2} (2r')^{(3-D)} \tag{8}$$

The interference of a spherical asperity with the flat surface is related to its radius of curvature, R, by $(R-\delta)^2 + r'^2 = R^2$. Since the asperity radius of curvature is typically orders of magnitude greater than the asperity height, the latter relationship can be reduced to $r'^2 = 2R\delta$. In order to calculate an equivalent radius of curvature of the cosine-shaped asperity, its shape can be approximated by a circular profile. Then substituting Eq. (8) into the above relationship gives

$$R = \frac{a'^{(D-1)/2}}{2^{(5-D)} \pi^{(D-1)/2} G^{(D-2)} \left(\ln \gamma\right)^{1/2}} \tag{9}$$

where a' is the truncated area of the microcontact ($a' = \pi r'^2$). From Hertz contact theory, the elastic force at a microcontact is given by (Johnson, 1987)

$$\Delta F_e = \frac{4E^* r^3}{3R} \tag{10}$$

where r is the radius of the real area of the microcontact, $a = \pi r^2$. For a circular elastic microcontact, $a' = 2a$ (Johnson, 1987), thus $a' = 2\pi r^2$. Then the elastic force at a microcontact can be obtained as a function of a' by substituting this relationship and Eq. (9) into Eq. (10),

$$\Delta F_e = \frac{2^{(11-2D)/2}}{3\pi^{(4-D)/2}} \left(\ln \gamma\right)^{1/2} G^{(D-2)} E^* a'^{(4-D)/2} \tag{11}$$

Since the critical interference at yielding, δ_c, is proportional to $R\left(H/E^*\right)^2$ (Majumdar and Bhushan, 1995), the yielding criterion for plastic microcontact can be written as

$$\delta_c = \lambda R \left(\frac{H}{E^*} \right)^2 \tag{12}$$

where λ is a coefficient. Substituting Eqs. (8) and (9) into Eq. (12), the critical microcontact truncated area, a_c', corresponding to δ_c is obtained as

$$a_c' = \left[2^{(9-2D)} \pi^{(D-2)} \lambda^{-1} \ln \gamma G^{(2D-4)} \left(\frac{E^*}{H} \right)^2 \right]^{1/(D-2)} \tag{13}$$

Since elasto-plastic deformation is not considered in this study, asperities with truncated areas larger than a_c' are elastically deformed, whereas those with truncated areas smaller than a_c' are fully plastically deformed. For elastic microcontacts, the contact force is given by Eq. (11), whereas for fully plastic microcontacts, the contact force is

$$\Delta F_p = Ha$$

Experiments by Pullen and Williamson (1972) involving a rough metal surface pressed by a rigid plane demonstrated that plastic flow at microcontacts yielded a uniform increase in surface height at noncontacting regions. Since for light and intermediate loads the total real contact area is much smaller than the apparent contact area, a should be slightly greater than a'. Thus, it may be assumed that

$$\Delta F_p \approx Ha' \tag{14}$$

The magnitude of a_c' is determined based on the condition that the microcontact force is continuous during the approaching process of the surfaces. Hence, at $a' = a_c'$, it is set $\Delta F_e(a) = \Delta F_p(a)$. Thus, from Eqs. (11) and (14), it is found that

$$2^{(11-2D)} \ln \gamma G^{(2D-4)} E^{*2} = 9\pi^{(4-D)} a_c'^{(D-2)} H^2$$

Combining Eq. (13) with the above equation yields

$$\lambda = \frac{9\pi^2}{4}$$

Substituting this value of λ into Eq. (13) gives

$$a_c' = \left[\frac{2^{(11-2D)}}{9\pi^{(4-D)}} \ln \gamma G^{(2D-4)} \left(\frac{E^*}{H} \right)^2 \right]^{1/(D-2)} \tag{15}$$

Mandelbrot (1982) has found that the number of islands, N, with areas greater than a particular area, s', follows the power-law relationship,

$$N(s') = \left(\frac{s_L'}{s'} \right)^{D_s/2}$$

where D_s is the fractal dimension of coastlines and s_L' is the largest island area. This relationship has been used in previous contact mechanics analyses of fractal surfaces with $D_s = D - 2$, where D is the fractal dimension of the rough surface (Majumdar and Bhushan, 1991; Wang and Komvopoulos, 1994a, 1994b, 1995) due to the close fractal resemblance between islands over the sea level and truncated asperities of rough surfaces. Hence, the truncated asperity size distribution function can be written as

$$n(a') = -\frac{dN(a')}{da'} = \frac{(D-1)}{2a_L'} \left(\frac{a_L'}{a'} \right)^{(D+1)/2} \tag{16}$$

where a_L' is the size of the largest truncated microcontact area at a given mean surface separation distance. The number of microcontacts with truncated areas between a' and a'+da' is given by $n(a')da'$.

The total contact force can be determined by integrating the forces $\Delta F_e(a')$ and $\Delta F_p(a')$ using the size distribution function, $n(a')$. If at a given interfacial distance $a_L' \le a_c'$, all the asperities are fully plastically deformed. Then the total deformation force, F_t, is equal to the total plastic deformation force, F_p,

$$F_t = F_p = HS'$$

where S' is the total truncated area at a given mean surface separation distance, and the total elastic deformation force, F_e, is zero. If $a_L' > a_c'$, both elastic and fully plastic microcontacts exist; thus,

$$F_e = \int_{a_c'}^{a_L'} \Delta F_e(a') n(a') da' \tag{17}$$

$$F_p = \int_0^{a_c'} \Delta F_p(a') n(a') da' \tag{18}$$

and

$$F_t = F_e + F_p$$

Substituting Eqs. (11) and (16) into Eq. (17) yields

$$F_e = \frac{2^{(9-2D)/2}}{3\pi^{(4-D)/2}} (D-1)(\ln\gamma)^{1/2} G^{(D-2)} E^* a_L'^{(D-1)/2} \int_{a_c'}^{a_L'} a'^{(3-2D)/2} da' \tag{19}$$

Integration of Eq. (19) leads to

$$F_e = \frac{2^{(11-2D)/2}}{3\pi^{(4-D)/2}} \left(\frac{D-1}{5-2D}\right) (\ln\gamma)^{1/2} G^{(D-2)} E^* a_L'^{(4-D)/2} \left[1 - \left(\frac{a_c'}{a_L'}\right)^{(5-2D)/2}\right] \quad \text{for} \quad D \neq 2.5 \tag{20}$$

and

$$F_e = 2\pi^{-3/4} (\ln\gamma)^{1/2} G^{1/2} E^* a_L'^{3/4} \ln\left(\frac{a_L'}{a_c'}\right) \qquad \text{for} \quad D = 2.5 \tag{21}$$

Introducing Eqs. (14) and (16) into Eq. (18) gives

$$F_p = \left(\frac{D-1}{3-D}\right) H a_L' \left(\frac{a_c'}{a_L'}\right)^{(3-D)/2} \tag{22}$$

As can be seen from Eqs. (20), (21), and (22), the contact forces are functions of a_L' and a_c'. At a given mean surface separation distance, the truncated area of the largest microcontact, a_L', can be determined from the total truncated area of the equivalent rough surface, S', by

$$S' = \int_0^{a_L'} a' n(a') da'$$

Substituting Eq. (16) into the above relationship yields

$$S' = \left(\frac{D-1}{3-D}\right) a_L' \tag{23}$$

The total truncated area, S', can be determined from numerical integration of the truncated areas of the rough surface, and a_L' can be obtained from Eq. (23).

In contact analyses of rough surfaces, the total real area of contact, S, and its elastic and plastic components, S_e and S_p, are of significant importance. When $a_L' \le a_c'$, all microcontacts are in the fully plastic contact state; thus,

$$S = S_p = \int_0^{a_L'} an(a')da' = \int_0^{a_L'} a'n(a')da' = \left(\frac{D-1}{3-D}\right)a_L' = S' \qquad \text{and} \qquad S_e = 0$$

When $a_L' > a_c'$, both elastic and fully plastic microcontacts exist at the interface, and the total elastic and plastic contact areas are given by

$$S_e = \int_{a_c'/2}^{a_L'/2} an(a')da' = \int_{a_c'}^{a_L'} \frac{a'}{2} n(a')da' = \left(\frac{D-1}{6-2D}\right)\left[1-\left(\frac{a_c'}{a_L'}\right)^{(3-D)/2}\right]a_L'$$

and

$$S_p = \int_0^{a_c'} an(a')da' = \int_0^{a_c'} a'n(a')da' = \left(\frac{D-1}{3-D}\right)\left(\frac{a_c'}{a_L'}\right)^{(3-D)/2} a_L'$$

Hence,

$$S = S_e + S_p = \left(\frac{D-1}{6-2D}\right)\left[1+\left(\frac{a_c'}{a_L'}\right)^{(3-D)/2}\right]a_L'$$

From the previous analysis, it can be seen that both the total contact force and the total real area of contact are functions of a_c', a_L', and S'. According to Eq. (15), a_c' depends on the surface topography and material properties and is invariant of the interfacial distance, whereas Eq. (23) indicates that a_L' is proportional to S', which depends on the surface topography and interfacial distance.

Although the previous analysis for the contact force and contact area is for isotropic surfaces, a similar derivation can be followed for anisotropic surfaces. To elaborate on this, the expression for the anisotropic surface (Eq. (4)) is first written as

$$z(\rho,\theta) = \left(\frac{\ln\gamma}{M}\right)^{1/2} A_1 \sum_{m=1}^{M} \frac{A_m}{A_1} \sum_{n=-\infty}^{\infty} \left(k\gamma^n\right)^{(D-3)} \left\{\cos\phi_{m,n} - \cos\left[k\gamma^n \rho\cos\left(\theta-\alpha_m\right)+\phi_{m,n}\right]\right\}, \quad (24)$$

and, subsequently, Eq. (24) is modified by introducing a factor $C(M)$ and maintaining only one ridge function such that the rms roughness of the surface represented by Eq. (24) is identical to that of the modified relationship, i.e.,

$$z(x) = C(M)A_1(\ln\gamma)^{1/2}\sum_{n=-\infty}^{\infty}\left(k\gamma^n\right)^{(D-3)}\left[\cos\phi_{1,n} - \cos\left(k\gamma^n x - \phi_{1,n}\right)\right] \tag{25}$$

Computations similar to those performed for isotropic surfaces can then be performed for the anisotropic surface assuming $2 < D < 3$ and $1 < M < 100$ to determine the range of $C(M)$. After the median value of $C(M)$ is substituted into Eq. (25), the asperity interference, contact force, and real contact area can be obtained following a procedure similar to that presented above for isotropic surfaces.

4 Numerical Results

Results for the elastic and plastic components of the deformation force and real contact area based on the above analysis are presented in this section for rough silica surfaces, the basic structural material of microelectromechanical systems where contact phenomena involving high adhesion (stiction) forces are of great concern. Unless otherwise stated, simulation results are given for 1 μm × 1 μm surfaces with $M = 10$, equivalent rms roughness $\sigma_e = 7$ nm ($G = 1.36 \times 10^{-11}$ m), and fractal dimension $D = 2.4$. The material properties of silica are $E = 72$ GPa, $\nu = 0.17$, and $H = 5.5$ GPa (Lin and Pugacz-Muraszkiewicz, 1972; McLellan and Shand, 1984). The forces are converted to average contact pressures by dividing them by the apparent contact area, and the real contact areas are normalized by the apparent contact area. In a previous publication by Komvopoulos and Yan (1997), a 2-D contact mechanics analysis of elastic rough surfaces was presented, and important insight into the evolution of the repulsive force due to deformation of the asperities and the van der Waals, capillary, and electrostatic attractive forces at the contact interface was obtained. In the present study, however, elastic-plastic material behavior and 3-D surface topography are introduced to develop a more rigorous contact mechanics analysis of rough surfaces. Results from both analyses are presented below for comparison.

In order to examine the effect of the material behavior on the resulting contact force, numerical results for the average contact pressure versus mean surface separation distance are presented in Fig. 4 for elastic and elastic-plastic 2-D surfaces possessing identical geometry ($M = 1$, $G = 1.16 \times 10^{-11}$ m). A lower contact pressure is produced in the case of elastic-plastic response. For both material behaviors, the incipient mean surface separation distance is approximately equal to 2.6 σ_e and the contact pressure increases rapidly by about two orders of magnitude in

the early stage of contact and levels off at distances less than about $2\sigma_e$. This behavior is attributed to the rapid increase of the real contact area at the incipience of surface contact and its relatively gradual variation at smaller surface separation distances.

Figures 5 and 6 show results for the average contact pressure and real contact area versus mean surface separation distance for 3-D fractal surfaces consisting of silica. The corresponding elastic and plastic components are also plotted to reveal the dominant deformation mode. Figure 5 shows that, for the simulated range of surface distance, elastic deformation dominates at the interface of contacting rough silica surfaces, except at the initial contact distances. The plastic deformation force component comprises only about 10 percent of the total contact force. Although the pressure distributions shown in Fig. 5 exhibit an overall trend similar to that shown in Fig. 4, there are two profound differences worth mentioning. The incipient mean surface separation distance is equal to about $3.6\sigma_e$ for the 3-D surfaces and the contact pressure is markedly lower than those shown in Fig. 4. Figure 6 shows that the real contact area is less than 1.1 percent of the apparent contact area when the interfacial pressure is less than 40 MPa, as evidenced from Fig. 5. These results indicate that the real contact area is a very small fraction of the apparent contact area. This is generally true for relatively low and moderate contact pressures that are typically encountered at interfaces of microsystems, such as magnetic recording devices and microelectromechanical systems.

5 Conclusions

A contact mechanics analysis of elastic-plastic rough surfaces characterized by fractal geometry was presented. Based on the power spectrum of the surface height function, it was demonstrated that random rough surfaces can be generated by using the proposed fractal model. The theoretical analysis yields relationships of the total contact force and real contact area in terms of surface separation distance, fractal parameters, and material properties, and provides new insight into the effects of surface topography and material behavior on the evolution of elastic and plastic deformation at the contact interface. Elastic-plastic material behavior yields lower contact pressures compared to purely elastic behavior. The contact pressure and incipient mean surface separation distance corresponding to 3-D elastic-plastic rough surfaces are smaller and greater, respectively, than those of 2-D surfaces possessing the same material properties and rms rough-

ness. In view of 3-D elastic-plastic contact simulation results for rough silica surfaces, the real contact area is a remarkably small fraction of the apparent contact area and, for the range of surface separation distance considered, elastic deformation is predominant at the contact interface.

Acknowledgment

This research was partially supported by the Surface Engineering and Tribology Program of the National Science Foundation under Grant No. CMS-9504403 and the Computer Mechanics Laboratory at the University of California at Berkeley.

References

Anguiano, E., Pancorbo, M., and Aguilar, M., 1994, "Pitfalls in the Fractal Characterization of Real Microscopic Surfaces by Frequency Analysis and Proposal of a New Method," *Fractals in the Natural and Applied Sciences*, Novak, M. M., ed., Elsevier Science, North-Holland, pp. 37-46.

Ausloos, M., and Berman, D. H., 1985, "A Multivariate Weierstrass-Mandelbrot Function," *Proceedings of the Royal Society of London, Series A*, Vol. 400, pp. 331-350.

Berry, M. V., and Lewis, Z. V., 1980, "On the Weierstrass-Mandelbrot Fractal Function," *Proceedings of the Royal Society of London, Series A*, Vol. 370, pp. 459-484.

Bhushan, B., and Majumdar, A., 1992, "Elastic-Plastic Contact Model for Bifractal Surfaces," *Wear*, Vol. 153, pp. 53-64.

Gagnepain, J. J., and Roques-Carmes, C., 1986, "Fractal Approach to Two-Dimensional and Three-Dimensional Surface Roughness," *Wear*, Vol. 109, pp. 119-126.

Greenwood, J. A., and Williamson, J. B. P., 1966, "Contact of Nominally Flat Surfaces," *Proceedings of the Royal Society of London, Series A*, Vol. 295, pp. 300-319.

Johnson, K. L., 1987, *Contact Mechanics*, Cambridge University Press, Cambridge, U.K., pp. 90-104.

Komvopoulos, K., and Yan, W., 1997, "A Fractal Analysis of Stiction in Microelectromechanical Systems," *ASME Journal of Tribology*, Vol. 119, in press.

Lin, S. C. H., and Pugacz-Muraszkiewicz, I., 1972, "Local Stress Measurement in Thin Thermal SiO_2 Films on Si Substrates," *Journal of Applied Physics*, Vol. 43, pp. 119-125.

Majumdar, A., and Bhushan, B., 1991, "Fractal Model of Elastic-Plastic Contact Between Rough Surfaces," *ASME Journal of Tribology*, Vol. 113, pp. 1-11.

Majumdar, A., and Bhushan, B., 1995, "Characterization and Modeling of Surface Roughness and Contact Mechanics," *Handbook of Micro/Nanotribology*, Bhushan, B., ed., CRC Press, New York, NY, pp. 109-165.

Majumdar, A., and Tien, C. L., 1990, "Fractal Characterization and Simulation of Rough Surfaces," *Wear*, Vol. 136, pp. 313-327.

Mandelbrot, B. B., 1982, *The Fractal Geometry of Nature*, Freeman, San Francisco, CA.

McCool, J. I., 1986, "Comparison of Models for the Contact of Rough Surfaces," *Wear*, Vol. 107, pp. 37-60.

McLellan, G. W., and Shand, E. B., 1984, *Glass Engineering Handbook*, McGraw-Hill, New York, NY, pp. 2-23.

Nayak, P. R., 1971, "Random Process Model of Rough Surfaces," *ASME Journal of Lubrication Technology*, Vol. 93, pp. 398-407.

Nayak, P. R., 1973, "Random Process Model of Rough Surfaces in Plastic Contact," *Wear*, Vol. 26, pp. 305-333.

Press, W. H., Teukolsky, S. A., Vettering, W. T., and Flannery, B. P., 1992, *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, Cambridge University Press, Cambridge, U.K., pp. 123-158 and 266-319.

Pullen, J., and Williamson, J. B. P., 1972, "On the Plastic Contact of Rough Surfaces," *Proceedings of the Royal Society of London, Series A*, Vol. 327, pp. 159-173.

Sayles, R. S., and Thomas, T. R., 1978, "Surface Topography as a Nonstationary Random Process," *Nature*, Vol. 271, pp. 431-434.

Wang, S., and Komvopoulos, K., 1994a, "A Fractal Theory of the Interfacial Temperature Distribution in the Slow Sliding Regime: Part I — Elastic Contact and Heat Transfer Analysis," *ASME Journal of Tribology*, Vol. 116, pp. 812-823.

Wang, S., and Komvopoulos, K., 1994b, "A Fractal Theory of the Interfacial Temperature Distribution in the Slow Sliding Regime: Part II — Multiple Domains, Elastoplastic Contacts and Applications," *ASME Journal of Tribology*, Vol. 116, pp. 824-832.

Wang, S., and Komvopoulos, K., 1995, "A Fractal Theory of the Temperature Distribution at Elastic Contacts of Fast Sliding Surfaces," *ASME Journal of Tribology*, Vol. 117, pp. 203-215.

Warren, T. L., and Krajcinovic, D., 1995, "Fractal Models of Elastic-Perfectly Plastic Contact of Rough Surfaces Based on the Cantor Set," *International Journal of Solids and Structures*, Vol. 32, pp. 2907-2922.

Warren, T. L., Majumdar, A., and Krajcinovic, D., 1996, "A Fractal Model for the Rigid-Perfectly Plastic Contact of Rough Surfaces," *ASME Journal of Applied Mechanics*, Vol. 63, pp. 47-54.

List of Figures

Fig. 1 Simulated three-dimensional fractal surface with $M = 10$, $D = 2.4$, and $G = 1.36 \times 10^{-11}$ m ($\sigma = 7$ nm)

Fig. 2 Contour plot of the power spectrum of the surface shown in Fig. 1

Fig. 3 Schematic of a microcontact established between an asperity on the composite rough surface and the opposing rigid plane

Fig. 4 Average contact pressure versus mean surface separation distance for two-dimensional elastic and elastic-plastic silica surfaces with $M = 1$, $D = 2.4$, and $G = 1.16 \times 10^{-11}$ m ($\sigma_e = 7$ nm)

Fig. 5 Elastic, plastic, and total average contact pressures versus mean surface separation distance for three-dimensional elastic-plastic silica surfaces with $M = 10$, $D = 2.4$, and $G = 1.36 \times 10^{-11}$ m ($\sigma_e = 7$ nm)

Fig. 6 Elastic, plastic, and total real contact areas versus mean surface separation distance for three-dimensional elastic-plastic silica surfaces with $M = 10$, $D = 2.4$, and $G = 1.36 \times 10^{-11}$ m ($\sigma_e = 7$ nm)

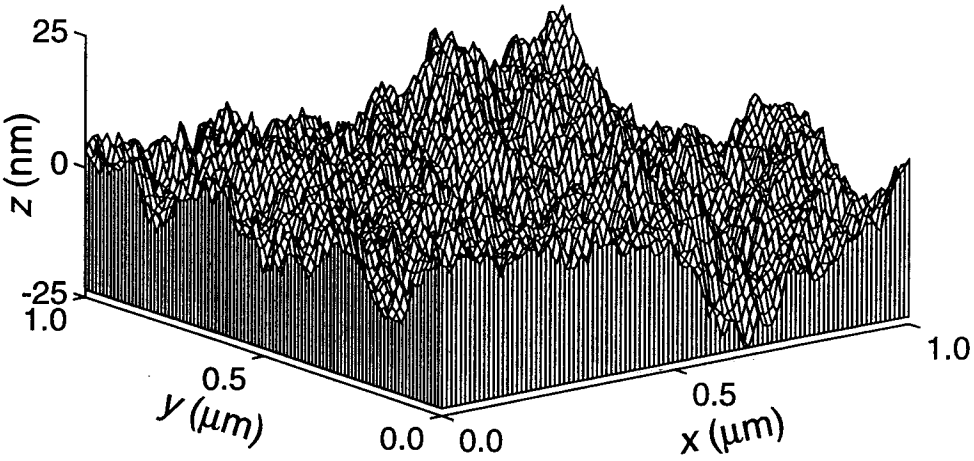

Fig. 1

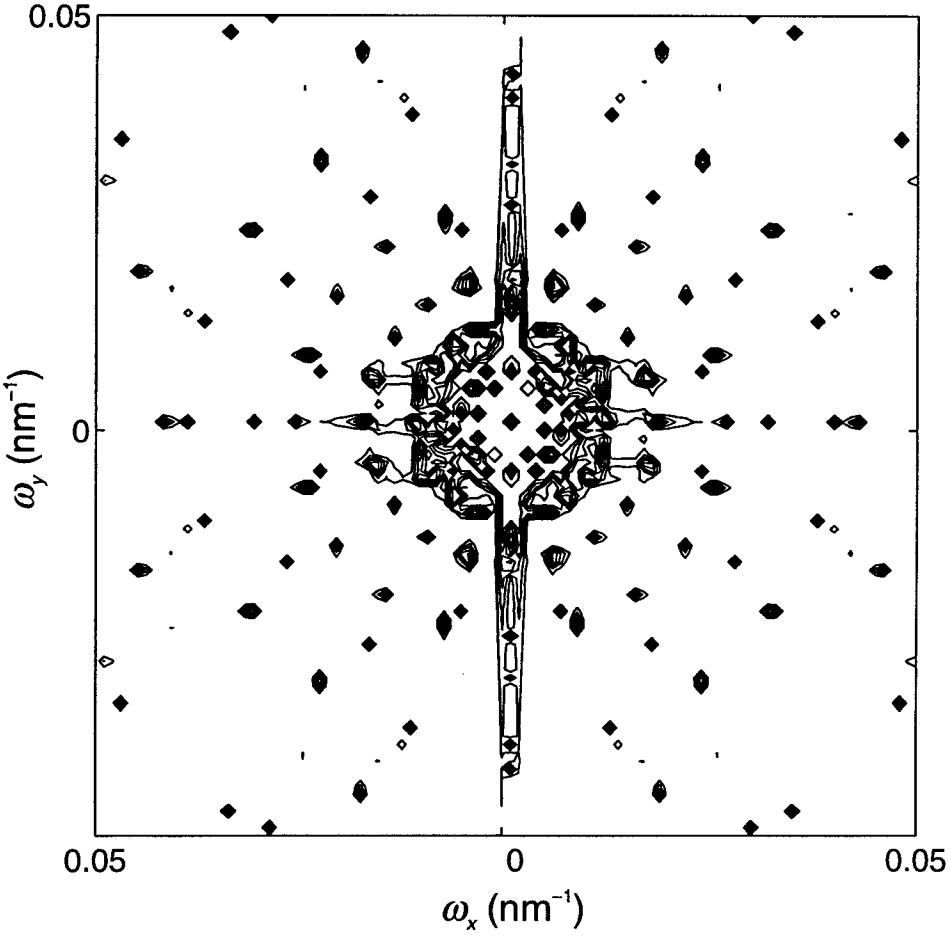

Fig. 2

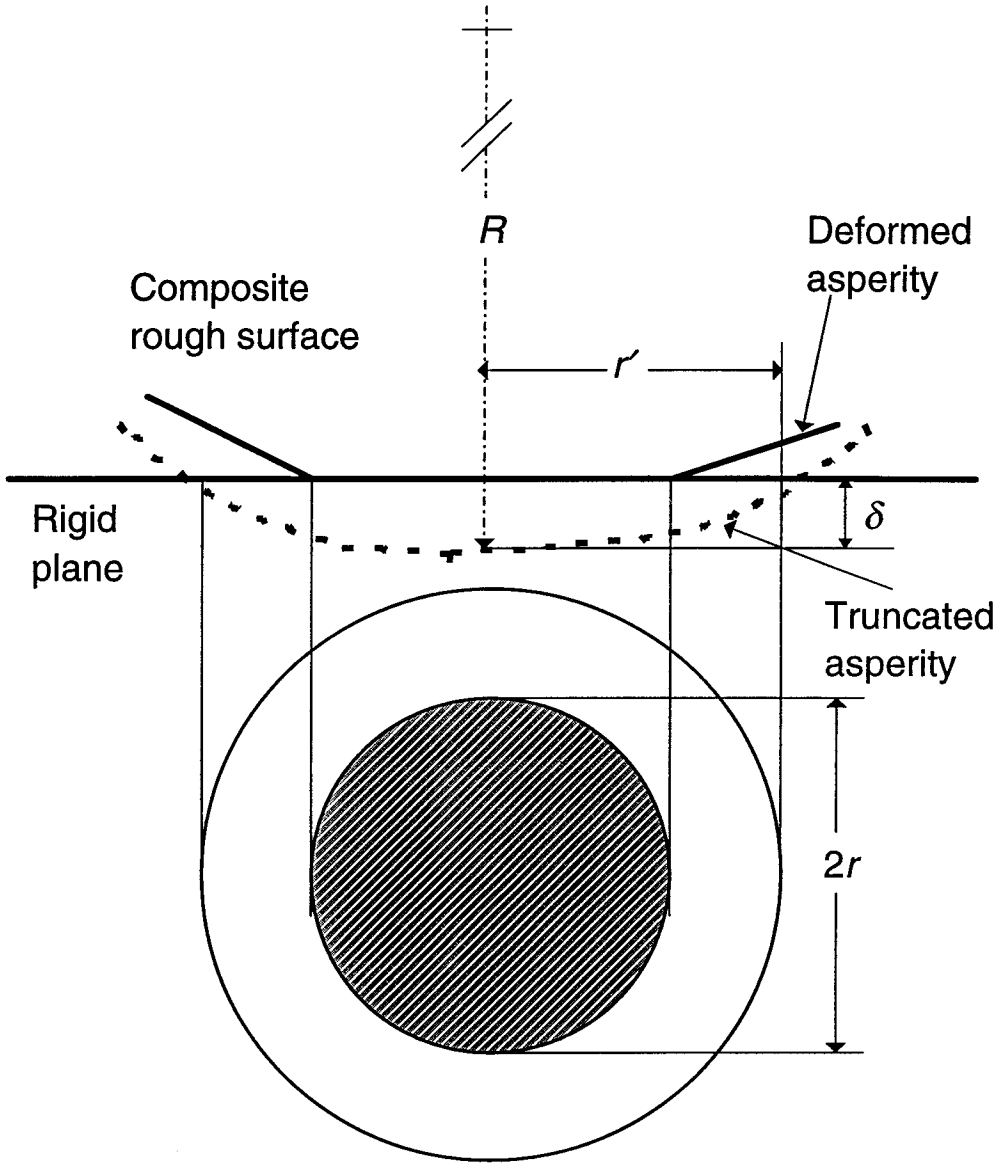
R
Deformed
asperity
Composite
rough surface
r'
Rigid
plane
δ
Truncated
asperity
2r

Fig. 3

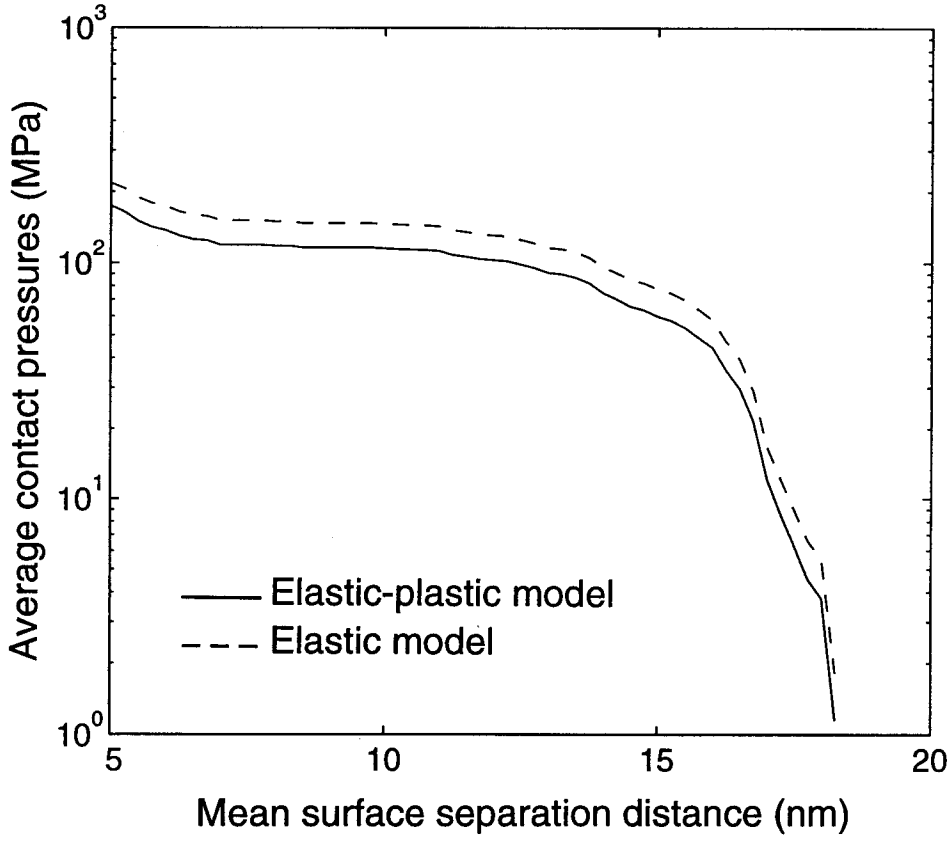
10^3
10^2
10^1
10^0
Average contact pressures (MPa)
5
10
15
20
Mean surface separation distance (nm)
Elastic-plastic model
Elastic model

Fig. 4

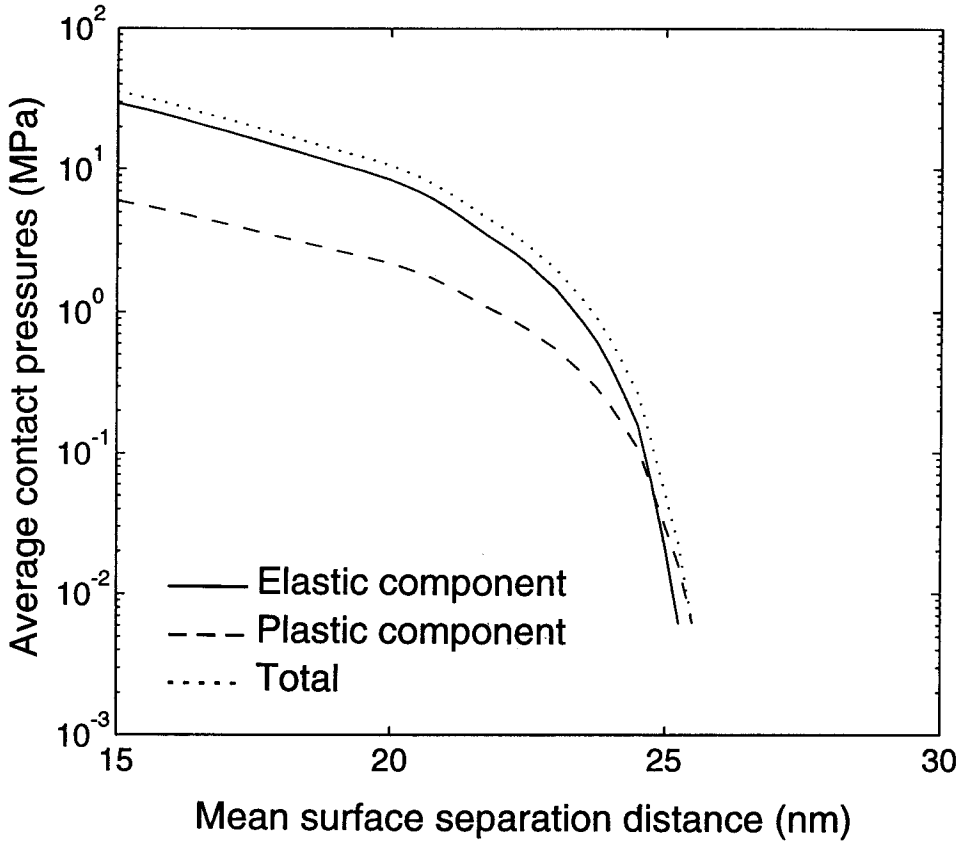
Average contact pressures (MPa)
10^2
10^1
10^0
10^{-1}
10^{-2}
10^{-3}
15
20
25
30
Mean surface separation distance (nm)
Elastic component
Plastic component
Total

Fig. 5

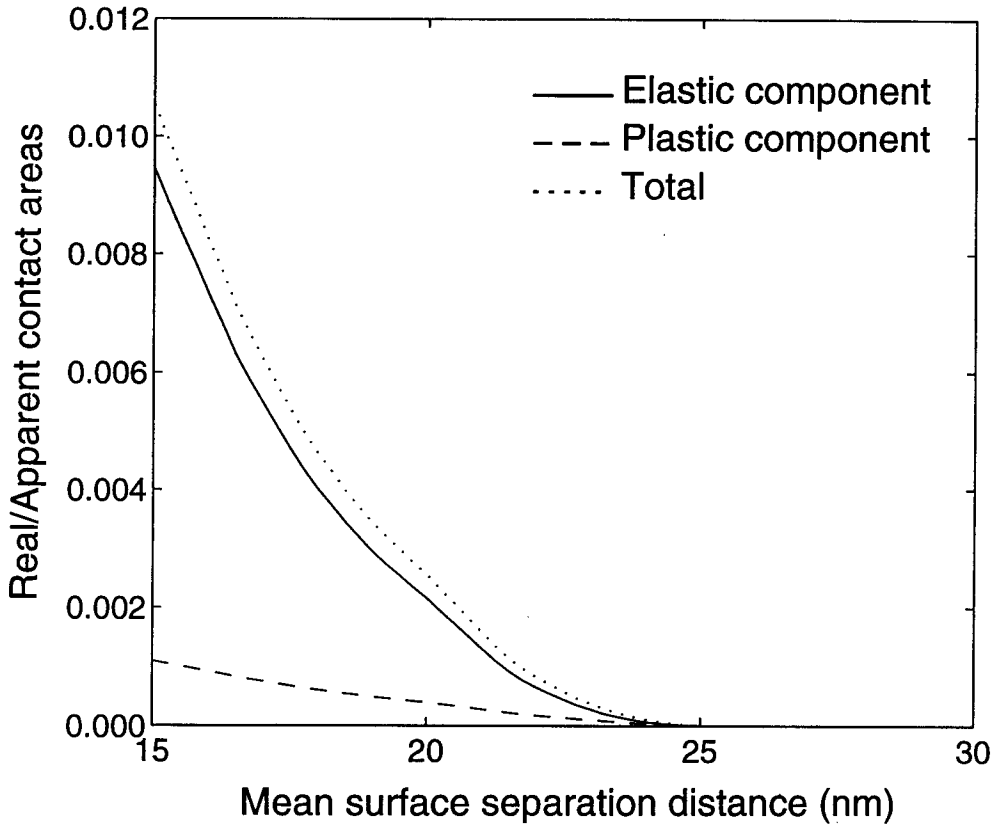
0.012
0.010
0.008
0.006
0.004
0.002
0.000
Real/Apparent contact areas
15
20
25
30
Mean surface separation distance (nm)
Elastic component
Plastic component
Total

Fig. 6