

## Elastoplastic Contact between Randomly Rough Surfaces

B. N. J. Persson

*IFF, FZ-Jülich, 52425 Jülich, Germany*

(Received 29 January 2001; published 27 August 2001)

I have developed a theory of contact mechanics between randomly rough surfaces. The solids are assumed to deform elastically when the stress  $\sigma$  is below the yield stress  $\sigma_Y$ , and plastically when  $\sigma$  reaches  $\sigma_Y$ . I study the dependence of the (apparent) area of contact on the magnification. I show that in most cases the *area of real contact*  $A$  is *proportional* to the load. If the rough surface is self-affine fractal (Hurst exponent  $H$ ) the whole way up to the lateral size  $L$  of the nominal contact area, then (assuming no plastic deformation)  $A \sim L^H$ .

DOI: 10.1103/PhysRevLett.87.116101

PACS numbers: 81.40.Pq, 62.20.-x

Even a highly polished surface has surface roughness on many different length scales. When two bodies with nominally flat surfaces are brought into contact, the area of real contact will usually only be a small fraction of the nominal contact area. We can visualize the contact regions as small areas where asperities from one solid are squeezed against asperities of the other solid; depending on the conditions the asperities may deform elastically or plastically.

How large is the area of *real* contact between a solid block and the substrate? This fundamental question has extremely important practical implications. For example, it determines the contact resistivity and the heat transfer between the solids. It is also of direct importance for sliding friction [1], e.g., the rubber friction between a tire and a road surface, and it has a major influence on the adhesive force between two solid blocks in direct contact.

In this Letter I develop a theory of contact mechanics, valid for randomly rough (e.g., self-affine fractal) surfaces [2]. In the context of rubber friction, which motivated this theory, mainly elastic deformation occurs. However, the theory can also be applied when both elastic and plastic deformations occur in the contact areas. This case is, of course, relevant to almost all materials other than rubber.

The contact theory can be applied to surfaces with roughness on many different length scales. The classical contact theory of Greenwood [3,4] (see also [2,5,6]) was developed for surfaces with roughness on a single length scale. Thus, in this theory the surface asperities are “approximated” by spherical caps of *identical* radius of curvature (but with a Gaussian height distribution). The Greenwood theory has been applied to real surfaces with roughness on many different length scales, by defining an average radius of curvature  $R$ . However, it turns out that  $R$  depends strongly on the resolution of the roughness-measuring instrument, or any other form of filtering, and hence is not unique [7,8]. The contact theory I have developed is based on a completely different physical approach, and gives well-defined results for surfaces with arbitrary surface roughness.

The basic idea behind the new contact theory is that it is very important not to *a priori* exclude any roughness length scale from the analysis. Consider a surface with surface roughness on two different length scales as indicated in Fig. 1. Assume that a rubber block is squeezed against the substrate and that the applied pressure is large enough to squeeze the rubber into the large “cavities” as indicated in the figure. It is clear that even if the rubber is able to make direct contact with the substrate in the large cavities, the pressure acting on the rubber at the bottom of a large cavity will be much smaller than the pressure at the top of a large asperity. Thus while, because of the high local pressure, the rubber may be squeezed into the “small” cavities at the top of a large asperity, the pressure at the bottom of a large cavity may be too small to squeeze the rubber into the small-sized cavities at the bottom of a large cavity. If  $A(\lambda)$  is the (apparent) area of contact on the length scale  $\lambda$  [more accurately, I *define*  $A(\lambda)$  to be the area of real contact if the surface would be smooth on all length scales shorter than  $\lambda$ , see Fig. 2], then I have studied the function  $P(\zeta) = A(\lambda)/A(L)$  which is the relative fraction of the rubber surface area where contact occurs on the length scale  $\lambda = L/\zeta$  (where  $\zeta \geq 1$ ), with  $P(1) = 1$ . Here  $A(L) = A_0$  denotes the macroscopic

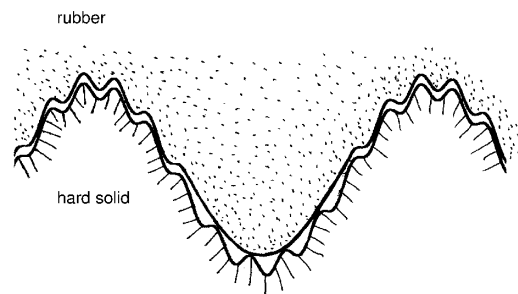


FIG. 1. A rubber block squeezed against a substrate with roughness on two different length scales. The rubber is able to fill out the long-wavelength roughness profile, but it is not able to get squeezed into the small-sized “cavities” at the bottom of a big cavity (schematic).

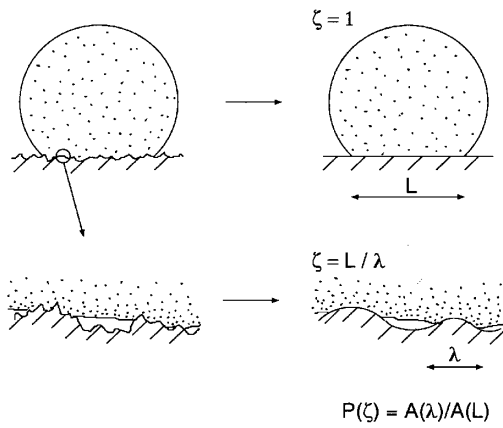


FIG. 2. A rubber ball squeezed against a hard, rough, substrate. Left: the system at two different magnifications. Right: the area of real contact  $A(\lambda)$  on the length scale  $\lambda$  is defined as the area of real contact when the surface roughness on shorter length scales than  $\lambda$  has been removed (i.e., the surface has been “smoothed” on length scales shorter than  $\lambda$ ).

contact area ( $L$  is the diameter of the macroscopic contact area so that  $A_0 \approx L^2$ ).

From contact mechanics (see, e.g., Ref. [6]) it is known that in the frictionless contact of elastic solids with rough surfaces, the contact stresses depend only upon the shape of the gap between them before loading. Thus, without loss of generality, the actual system may then be replaced by a flat elastic surface [elastic modulus  $E$  and Poisson ratio  $\nu$ , related to the original quantities via  $(1 - \nu^2)/E = (1 - \nu_1^2)/E_1 + (1 - \nu_2^2)/E_2$ ] in contact with a rigid body having a surface roughness profile which result in the same undeformed gap between the surfaces.

The detailed derivation of  $P(\zeta)$  will be presented elsewhere, and here I give only the major equations. If  $A_0$  denotes the nominal contact area, the load  $F_N = \sigma_0 A_0$ . This load must remain unchanged as we study the contact at shorter length scales. Consider the system at the length scale  $\lambda = L/\zeta$ , where  $L$  is of the order of the diameter of the nominal contact area. We define  $q_L = 2\pi/L$  and write  $q = q_L \zeta$ . Let  $P(\sigma, \zeta)$  denote the stress distribution in the contact areas under the magnification  $\zeta$ . Let us first assume complete contact between the rubber and the substrate on all length scales. We have

$$P(\sigma, \zeta) = \langle \delta[\sigma - \sigma_1(\mathbf{x})] \rangle,$$

where  $\sigma_1(\mathbf{x})$  is the stress which occurs in the contact area when the surface roughness with a wavelength shorter than  $L/\zeta$  has been smoothed out. Here  $\langle \dots \rangle$  stands for ensemble averaging, i.e., averaging over different realization of the surface roughness profile. By expanding the equation  $P(\zeta + \Delta\zeta, \sigma) = \langle \delta(\sigma - \sigma_1 - \Delta\sigma_1) \rangle$  to linear order in  $\Delta\zeta$ , and assuming that averaging over different regions in  $\zeta$  are independent processes, one obtains

$$\frac{\partial P}{\partial \zeta} = G'(\zeta) \sigma_0^2 \frac{\partial^2 P}{\partial \sigma^2}, \quad (1)$$

where  $G'(\zeta)$  denotes the  $\zeta$  derivative of the function

$$G(\zeta) = \frac{\pi}{4} \left[ \frac{E}{(1 - \nu^2)\sigma_0} \right]^2 \int_{q_L}^{\zeta q_L} dq q^3 C(q). \quad (2)$$

The surface roughness power spectra

$$C(q) = \frac{1}{(2\pi)^2} \int d^2x \langle h(\mathbf{x})h(\mathbf{0}) \rangle e^{-i\mathbf{q}\cdot\mathbf{x}},$$

where  $z = h(\mathbf{x})$  is the height of the surface above a flat reference plane (chosen so that  $\langle h \rangle = 0$ ).

Note that

$$P(\sigma, 1) = P_0(\sigma),$$

where we assume that  $P_0(\sigma) = \delta(\sigma - \sigma_0)$ , corresponding to a constant pressure in the nominal contact area.

Equation (1) is a diffusion type of equation, where time is replaced by the magnification  $\zeta$ , and the spatial coordinate with the stress  $\sigma$  (and where the “diffusion constant” depends on  $\zeta$ ). Hence, when we study  $P(\sigma, \zeta)$  on shorter and shorter length scales (corresponding to increasing  $\zeta$ ), the  $P(\sigma, \zeta)$  function will become broader and broader in  $\sigma$  space. We can take into account the fact that detachment actually will occur when the local stress reaches  $\sigma = 0$  (we assume no adhesion) via the boundary condition

$$P(0, \zeta) = 0.$$

We assume first that only elastic deformation occurs (i.e.,  $\sigma_Y \rightarrow \infty$ ). In this case

$$P(\zeta) = \int_0^\infty d\sigma P(\sigma, \zeta).$$

It is straightforward to solve (1) with the boundary conditions  $P(0, \zeta) = 0$  and  $P(\infty, \zeta) = 0$  to get

$$P(\zeta) = \frac{2}{\pi} \int_0^\infty dx \frac{\sin x}{x} \exp[-x^2 G(\zeta)]. \quad (3)$$

We consider now the limit  $\sigma_0 \ll E$ , which is satisfied in most applications. In this case, for most  $\zeta$  values of interest,  $G(\zeta) \gg 1$ , so that only  $x \ll 1$  will contribute to the integral in (3), and we can approximate  $\sin x \approx x$  and

$$P(\zeta) \approx \frac{2}{\pi} \int_0^\infty dx \exp[-x^2 G(\zeta)] = [\pi G(\zeta)]^{-1/2}. \quad (4)$$

Thus, within this approximation, using (2) and (4) we get  $P(\zeta) \propto \sigma_0$  so that the area of real contact is proportional to the load. This is the reason why the friction coefficient in most cases is independent of the load.

The theory above is valid for surfaces with arbitrary random roughness, but I now apply it to self-affine fractal surfaces. It has been found that many “natural” surfaces, e.g., surfaces of many materials generated by fracture, can be approximately described as self-affine surfaces over a rather wide roughness size region. A self-affine fractal surface has the property that if we make a scale change that is different for each direction, then the surface does not change its morphology [9]. Recent studies have shown that even asphalt road tracks (of interest for rubber friction)

are (approximately) self-affine fractal, with an upper cutoff length  $\lambda_0 = 2\pi/q_0$  of the order of a few mm [10]. For a self-affine fractal surface  $C(q) = C_0$  for  $q < q_0$ , while for  $q > q_0$

$$C(q) = C_0 \left( \frac{q}{q_0} \right)^{-2(H+1)}, \quad (5)$$

where  $H = 3 - D_f$  (where the fractal dimension  $2 < D_f < 3$ ), and where  $q_0$  is the lower cutoff wave vector, and  $C_0$  is determined by the rms roughness amplitude,  $\langle h^2 \rangle \equiv h_0^2/2$  via  $C_0 = \alpha(h_0/q_0)^2 H/2\pi$  where  $\alpha = 1/[1 + H - (q_L/q_0)^2 H]$ . Note that for  $q_L/q_0 \ll 1$ ,  $\alpha \approx 1/(1 + H)$  is independent of  $L$ .

Substituting (5) in (2), defining  $q = q_L \zeta$  and assuming  $q \gg q_0$  gives

$$G(\zeta) \approx \left( \frac{q_0 h_0}{4(1 - \nu^2)} \right)^2 \frac{\alpha H}{(1 - H)} \left( \frac{E}{\sigma_0} \right)^2 \left( \frac{q}{q_0} \right)^{2(1-H)},$$

so that

$$P(\zeta) \approx \frac{4(1 - \nu^2)}{q_0 h_0} \left( \frac{1 - H}{\pi \alpha H} \right)^{1/2} \frac{\sigma_0}{E} \left( \frac{q}{q_0} \right)^{H-1}.$$

But  $\sigma_0 = F_N/A_0$  so that the (apparent) area of contact on the length scale  $\lambda = 2\pi/q$  becomes

$$\begin{aligned} A(\lambda) &= A_0 P(\zeta) \\ &= \frac{4(1 - \nu^2)}{q_0 h_0} \left( \frac{1 - H}{\pi \alpha H} \right)^{1/2} \frac{F_N}{E} \left( \frac{\lambda}{\lambda_0} \right)^{1-H}. \end{aligned} \quad (6a)$$

If  $\lambda_0 = L$  we get  $q_0 = 2\pi/L$  and

$$A(\lambda) = \frac{2L(1 - \nu^2)}{\pi h_0} \left( \frac{1 - H}{\pi \alpha H} \right)^{1/2} \frac{F_N}{E} \left( \frac{\lambda}{L} \right)^{1-H}. \quad (6b)$$

If  $\lambda_1$  denotes the low-distance cutoff in the self-affine fractal distribution (which cannot be smaller than an atomic dimension), then (6) shows that the *area of real contact*  $A(\lambda_1)$  is *proportional* to the load. If the upper cutoff length  $\lambda_0$  is independent of the size  $L$  of the system, then  $A(\lambda_1)$  is also independent of  $L$ . However, if  $\lambda_0 = L$  then the area of real contact  $A(\lambda_1)$  depends on the size  $L$  of the system, increasing as  $\sim L^H$  with increasing  $L$ .

Let us note that in some problems, e.g., in the contact resistivity between two metallic bodies, the only thing which matters is the area of real contact  $A_0 P(\zeta_1)$  for  $\zeta_1 = L/\lambda_1$ . However, in other problems the whole function  $P(\zeta)$  matters. This is the case in rubber friction, where viscoelastic deformation of the rubber block on all length scales contributes in an equally important manner, and where the kinetic friction coefficient can be written as an integral over all  $\zeta$ , with the factor  $P(\zeta)$  occurring in the integrand. In other cases, what matters is not  $P(\zeta)$  on the shortest (cutoff) length  $\lambda_1 = L/\zeta_1$ , but  $P(\zeta)$  for some  $\zeta < \zeta_1$ . This is the case, for example, in sealing processes where a significant proportion of the surfaces needs to be in contact [11]. In the absence of adhesion (and plastic yielding), perfect

contact requires infinitely high pressure and can therefore not be obtained. However, for gas leakage, the gas mean free path length  $l$  might be an appropriate order of magnitude for the shortest length scale of any relevance, in which case  $P(\zeta)$  for  $\zeta = L/l$  would be the important quantity to focus on. This agrees with the general trend of experimental results that leakage rates diminish with increasing nominal contact pressure  $\sigma_0$ , but that leak tightness is a matter of degree or sensitivity of measurement, rather than an absolute state [11].

In the study above we assumed that only elastic deformation occurs. However, the theory can be generalized to the case where also plastic deformation occurs simply by replacing the boundary condition  $P(\infty, \zeta) = 0$  with  $P(\sigma_Y, \zeta) = 0$ , which describes that plastic deformation occurs in the contact area when the local stress has reached  $\sigma_Y$ . Let us introduce the functions  $P_{\text{non}}(\zeta)$  and  $P_{\text{pl}}(\zeta)$  which describe the fraction of the original (for  $\zeta = 1$ ) macrocontact area where, under the magnification  $\zeta$ , non-contact, and contact with plastic yield has occurred, respectively. Thus we have

$$P_{\text{el}}(\zeta) + P_{\text{non}}(\zeta) + P_{\text{pl}}(\zeta) = 1, \quad (7)$$

where  $P_{\text{el}}(\zeta) = P(\zeta)$  describes the fraction of the macrocontact area where elastic contact occurs on the length scale  $L/\zeta$ . At this point we note that the present theory is strictly valid only as long as  $|\nabla h(\mathbf{x})| \ll 1$ , which is satisfied in most engineering applications. If this condition is not satisfied, the tangent area can be larger than the projected [on the  $(x, y)$  plane] area, and the ‘‘conservation law’’ (7) is broken.

It is straightforward to solve (1) with the boundary conditions  $P(0, \zeta) = 0$  and  $P(\sigma_Y, \zeta) = 0$  to get

$$\begin{aligned} P_{\text{non}} &= \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{\sin \alpha_n}{n} \{1 - \exp[-\alpha_n^2 G(\zeta)]\}, \\ P_{\text{pl}} &= -\frac{2}{\pi} \sum_{n=1}^{\infty} (-1)^n \frac{\sin \alpha_n}{n} \{1 - \exp[-\alpha_n^2 G(\zeta)]\}, \end{aligned}$$

where  $\alpha_n = n\pi\sigma_0/\sigma_Y$ , and where  $G(\zeta)$  is given by (2). In the elastic limit,  $\sigma_Y \rightarrow \infty$ ,  $P_{\text{pl}} = 0$ , and  $P_{\text{el}} = 1 - P_{\text{non}}$  reduces to Eq. (3).

When  $C(q)$  is given by (5), the functions  $P_{\text{el}}$  and  $P_{\text{pl}}$  depend only on  $H$  (or, equivalently, on the fractal dimension  $D_f = 3 - H$ ), on  $\sigma_0/\sigma_Y$ , and on the *plasticity index*  $\psi = (E/\sigma_Y)q_0 h_0$ . In Fig. 3, I show the dependence of  $P_{\text{el}}$  and  $P_{\text{pl}}$  on the magnification  $\zeta$ . I have used parameters which correspond (roughly) to a cubic steel block ( $L = 10$  cm), on a steel substrate. I assume  $\sigma_0 = 10^4$  Pa,  $\sigma_Y = 10^9$  Pa, and  $E = 10^{11}$  Pa. The surface roughness of the substrate is assumed to be self-affine fractal with  $q_0 h_0 = 0.001$  (solid lines) and 0.01 (dashed lines). The theory does not depend on  $q_0$  directly (but only on the product  $q_0 h_0$ ), but if we choose the cutoff wave vector  $q_0 = 10^4 \text{ m}^{-1}$  (corresponding to the typical cutoff length  $\lambda_0 = 2\pi/q_0$  of order  $\approx 1$  mm), then  $q_0 h_0 = 0.001$  and 0.01 correspond to the

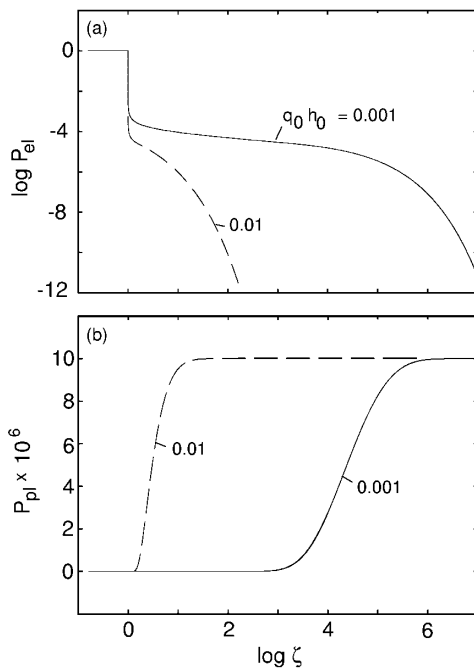


FIG. 3. The functions (a)  $P_{el}$  and (b)  $P_{pl}$  describe the fraction of the macroscopic contact area where elastic and plastic contacts occur, when the system is studied at different magnifications  $\zeta$ . For  $H = 0.8$ ,  $q_0 = 10^4 \text{ m}^{-1}$  and  $q_0 h_0 = 0.001$  (solid lines), and  $q_0 h_0 = 0.01$  (dashed lines). Results are shown for  $E = 10^{11} \text{ Pa}$ ,  $\sigma_Y = 10^9 \text{ Pa}$ , and  $\sigma_0 = 10^4 \text{ Pa}$ . Note that in the present case  $\zeta = 1$  corresponds to the length scale  $\lambda_0 = 2\pi/q_0 \approx 1 \text{ mm}$  so that the  $\log \zeta < 0$  corresponds to length scales  $\lambda > \lambda_0$ , and on these length scales the solid block makes (apparent) contact with the substrate over the whole block-substrate interface.

rms roughness  $h_0 = 0.1$  and  $1 \mu\text{m}$ , respectively. In the calculations we used the fractal exponent  $H = 0.8$ . Note that for the case  $q_0 h_0 = 0.01$  plastic deformation starts already at the cutoff length  $\lambda_0 \approx 1 \text{ mm}$ , and on the length scale  $\lambda_0/10 \approx 0.1 \text{ mm}$  all junctions have yielded plastically. However, when  $q_0 h_0 = 0.001$  plastic yield starts when  $\zeta$  is of the order of a few 1000, corresponding to distances of order  $\lambda_0/\zeta \approx 0.1 \mu\text{m}$ . On the length  $\lambda \approx 20 \text{ \AA}$  (corresponding to  $\zeta \approx 3 \times 10^5$ ) all asperities have yielded plastically. However, on this short length scale steel may

be much harder than the macroscopically observed yield stress; thus, for “real” steel mainly elastic deformation is likely to prevail when  $q_0 h_0 = 0.001$ .

We are at present generalizing the theory presented above to take into account adhesion. It has already been shown experimentally by Fuller and Tabor [12] that surface roughness can completely remove the effect of the attractive block-substrate interaction. As described in Ref. [13], this result can be understood if one compares the (roughness induced) elastic energy  $U_{el}$  stored in the elastic deformation field at the interface, with the adhesion energy  $U_{ad}$  due to the attractive block-substrate interaction: when  $U_{el} + U_{ad} = 0$  no “external” energy is needed to break the block-substrate bond.

This work has been supported by a research and development grant from Pirelli Pneumatici. I also thank BMBF for a grant related to the German-Israeli Project Cooperation “Novel Tribological Strategies from the Nano-to Meso-Scales.”

- [1] B. N. J. Persson, *Sliding Friction: Physical Principles and Applications* (Springer, Heidelberg, 2000), 2nd ed.
- [2] S. Roux, J. Schmittbuhl, J.P. Vilotte, and A. Hansen, *Europhys. Lett.* **23**, 277 (1993); A. Majumdar and B. Bhushan, *J. Tribol.* **113**, 1 (1991).
- [3] J. A. Greenwood, in *Fundamentals of Friction, Macroscopic and Microscopic Processes*, edited by I.L. Singer and H. M. Pollack (Kluwer, Dordrecht, 1992).
- [4] J. A. Greenwood and J.B.P. Williamson, *Proc. R. Soc. London A* **295**, 300 (1966).
- [5] J. F. Archard, *Proc. R. Soc. London A* **243**, 190 (1957).
- [6] K. L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985).
- [7] A. W. Bush *et al.*, *Mech. Res. Commun.* **3**, 169 (1976).
- [8] See page 68 in Ref. [3] and page 591 in Ref. [10].
- [9] J. Feder, *Fractals* (Plenum Press, New York, 1988).
- [10] M. Klüppel and G. Heinrich, *Rubber Chem. Technol.* **73**, 578 (2000).
- [11] W. Manners, *Int. J. Mech. Sci.* **42**, 2321 (2000).
- [12] K.N.G. Fuller and D. Tabor, *Proc. R. Soc. London A* **345**, 327 (1975).
- [13] B. N. J. Persson and E. Tosatti, *J. Chem. Phys.* (to be published).