Static Friction between Elastic Solids due to Random Asperities

J. B. Sokoloff
Physics Department and Center for Interdisciplinary Research on Complex Systems, Northeastern University, Boston, Massachusetts 02115
(Received 14 June 2000)

Several workers have established that the Larkin domains for two three-dimensional nonmetallic elastic solids in contact with each other at a disordered but atomically flat interface are enormously large, implying that there should be negligible static friction per unit area in the macroscopic solid limit. In contrast, the present Letter argues that when the Larkin domains are calculated for disorder on the multiasperity scale, they are much smaller than the interface size. This can account for the virtual universal occurrence of static friction.

DOI: 10.1103/PhysRevLett.86.3312 PACS numbers: 46.55.+d, 62.20.Qp, 81.40.Pq

It is well known that one must apply a minimum force (i.e., static friction) in order to get two solids, which are in contact, to slide relative to each other. It has been argued, however, that there might be no static friction for nonmetallic crystalline surfaces, which are incommensurate with each other. Aubry showed this for the weak potential limit of the one-dimensional Frenkel-Kontorova model [1], and recently He *et al.* [2] and Muser and Robbins [3] have shown for clean weakly interacting two-dimensional incommensurate interfaces that the force of static friction per unit area falls to zero as $A^{-1/2}$ in the thermodynamic limit, where A is the interface area. Even identical solids are incommensurate if their crystalline axes are rotated with respect to each other. Disorder, however, can pin contacting solids, just as it pins sliding charge density waves [4,5] and vortices in a superconductor [6]. Recently, it has been shown that Larkin domains (i.e., domains over which the solids are able to distort to accommodate the disorder at the interface) for contacting three-dimensional elastic solids are enormously large compared to typical solid sizes [7–9], implying that the force of static friction per unit area due to interface disorder should also fall off as $A^{-1/2}$ in the thermodynamic limit. In contrast to Refs. [2] and [3], where it was proposed that the presence of a submonolayer film of mobile molecules at the interface is a requirement for the occurrence of static friction between incommensurate surfaces, it is argued that disorder that occurs on the multimicron scale, due to disordered asperities, results in Larkin domains that are much smaller than the interface size, even for clean interfaces, implying that there is static friction for a macroscopic surface.

Here scaling methods, like those used by Fisher for charge density waves (CDW) [10], are used to study static friction for disordered interfaces. This can be accomplished by minimizing the potential energy of the solid in contact with a rigid disordered substrate at $z = 0$ with respect to the size of a Larkin domain [4], which is expected to give qualitatively correct results for the problem of two elastic disordered solids in contact. Given that the energy density of the elastic solid is given approximately by

$$(1/2)E'|\nabla \mathbf{u}|^2 + V(\mathbf{r})\delta(z), \tag{1}$$

where E' is an effective Young's modulus and $V(\mathbf{r})$ is the potential per unit area of the disordered substrate and $\mathbf{u}(\mathbf{r})$ is the local displacement of the solid, the energy of a single Larkin domain is given by

$$E = (1/2)L'L^2E' \times [|\nabla_t' \mathbf{u}'|^2/L^2 + |\partial \mathbf{u}'/\partial z'|^2/L'^2] - V_0 bL, \tag{2}$$

where b is an atomic length scale (e.g., a lattice constant), L is the width, and L' is the height of the domain, $|\nabla_t' \mathbf{u}'|^2 = |\partial \mathbf{u}'/\partial x'|^2 + |\partial \mathbf{u}'/\partial y'|^2$, where we assume that the local displacement $\mathbf{u}$ varies on length scales L and L' in the x and y and the z directions, respectively. That is, we assume that $\mathbf{u}(x, y, z)$ has the form $\mathbf{u}'(x', y', z')$, where the function u' varies by an amount of the order of atomic length scales when x', y', and z', defined by $(x', y', z') = (x/L, y/L, z/L')$, each vary by an amount of order unity. Here V_0 is a typical value of the potential per unit area. When Eq. (2) is minimized with respect to L' one finds that $L' \approx L$ and the energy per unit area at the interface is given by

$$E/L^2 \approx [(1/2)E'|\nabla' \mathbf{u}'|^2 - V_0 b]/L \tag{3}$$

(where we use the average value of $|\nabla' \mathbf{u}'|^2$ here), whose absolute minimum occurs for infinite L (more correctly L comparable to the interface length) for $E'|\nabla' \mathbf{u}'|^2 > V_0 b$ (i.e., when elastic energy dominates), implying that the static friction per atom decreases as the reciprocal of the square root of the surface area.

Fisher [10] has shown that above the critical dimension of 4, charge density waves are not pinned for typical impurity strengths, but fluctuations in the impurity concentration and strength lead to pinning. (The critical dimension for two solids in contact at a disordered interface is 3, as seen above.) Consider the effect of fluctuations in the defect concentration for thick solids for atomic level disorder, by dividing the solid into boxes of length L and examining the percentage of blocks at the interface of sufficiently

 0031-9007/01/86(15)/3312(4)$15.00

large defect concentration to be in the "strong pinning" regime, where the substrate force on each block dominates over the interblock elastic forces. Such blocks will be referred to as "strong blocks." Consider the parameter $\lambda \approx V_1/E'b^3$, where b is of the order of a lattice spacing and V_1 is the strength of the potential due to a defect acting on an atom in the second surface. Let $n_c = c'L^2$ be the number of defects within a particular block and c' (where $c' > c$, where c is the average defect concentration for the interface), the defect concentration strong enough for the block to be considered a strong block. Then the ratio of the interaction of a typical strong block with the substrate to typical elastic energy $E'b^3$ is $\lambda(c'L^2)^{1/2}$. The interface area surrounding each strong block is the total interface area A divided by the number of strong blocks at the interface, PA/L^2, where P is the probability of a particular block being a strong one. Then L^2/P is the interface area surrounding each block and the typical length L' over which the elastic interaction between two strong blocks acts is its square root, $L' = L/P^{1/2}$. Then the total elastic energy associated with each strong block is the product of the volume per strong block $= (L')^3$ and the elastic energy density, which is proportional to $|\nabla u|^2$ [which scales as $(L')^{-2}$] or L'. The criterion for a block to be a strong one is $\lambda(c'L^2)^{1/2} \gg L'$, or $\lambda \gg (c'P)^{-1/2}$. Since $c'P < 1$, this violates our previous assumption that $\lambda \ll 1$, implying that such fluctuations cannot result in strong pinning. There are also fluctuations in the locations of the points of contact within the defect potential wells within each Larkin domain; it too does not lead to static friction [11].

The above arguments seem to imply that weakly interacting disordered surfaces cannot exhibit static friction. We shall see, however, that unlike weak atomic scale defects, for which the elastic interaction between them can dominate over their interaction with the second surface, for contacting asperities that occur when the problem is studied on the multimicron scale, the interaction of two contacting asperities from the two surfaces dominates over the elastic interaction between asperities in one solid. It is suggested here that this could be responsible for the virtual universal occurrence of static friction. Roughness due to asperities is well described by the Greenwood-Williamson (GW) model [11–14], in which there are assumed to be elastic spherical asperities on a surface with an exponential or Gaussian height distribution in contact with a rigid substrate, especially for relatively light loads. Volmer and Nattermann's discussion of static friction [14] is not qualitatively different from that of Ref. [12]. In the GW model, the total contact area is of the order of

$$A_c = 2\pi\sigma R_c N \int_h^\infty ds\, \phi(s)(s - h), \tag{4}$$

where $\phi(s)$ is the distribution of asperity heights z, where $s = z/\sigma$, where σ is a length scale associated with the height distribution, and h is the ratio of the distance of the lower part of the bulk part of the sliding solid, from the surface in which it is in contact to σ, R_c is the radius of curvature of an asperity, and N is the number of asperities above a certain size, independent of whether they are in contact [11–14]. Taking the interaction of a single asperity with the substrate equal to the product of the contact area and a shear strength for the interface, the GW model gives friction approximately proportional to the load [12,13].

The energy of the interface consists of two parts. One part is the single asperity energy, which consists of the interaction energy of an asperity with the substrate plus the elastic energy cost necessary for each asperity to seek its minimum energy, neglecting its elastic interaction with other asperities, which is independent of the asperity density. The second part includes the elastic interaction between asperities within the same solid, which depends on the asperity density. In order to determine these energies, let us model the interaction of the ℓth asperity with the substrate by a spherically symmetric harmonic potentials of force constant α_ℓ. Assume that in the absence of distortion of the solid, the ℓth asperity lies a distance $\mathbf{\Delta}_\ell$ from the center of its potential well. Let $\mathbf{u}_\ell$ be the displacement of the ℓth asperity from its initial position. We use the usual elastic Green's function tensor of the medium at a distance r from the point at which a force is applied at the interface, but for simplicity, we approximate it by the simplified form $G(r) = (E'r)^{-1}$, where E' is Young's modulus [15]. Then the equilibrium conditions on the u's are

$$\mathbf{u}_\ell = (E'a)^{-1}\alpha_\ell(\mathbf{\Delta}_\ell - \mathbf{u}_\ell) + \sum_j (E'R_{\ell,j})^{-1}\alpha_j(\mathbf{\Delta}_j - \mathbf{u}_j), \tag{5}$$

where a is a parameter of the order of the size of the asperity, and $R_{\ell,j}$ is the distance between the ℓth and jth asperities. To lowest order in the interasperity interaction, the approximate solution for $\mathbf{u}_\ell$ is

$$\mathbf{u}_\ell = \mathbf{u}_\ell^0 + [1 + (E'a)^{-1}\alpha_\ell]^{-1} \times \sum_j (E'R_{\ell,j})^{-1}\alpha_j(\mathbf{\Delta}_j - \mathbf{u}_j^0), \tag{6}$$

where $\mathbf{u}_\ell^0 = \frac{\alpha_\ell}{\alpha_\ell + E'a}\mathbf{\Delta}_\ell$ is the zeroth order approximation [i.e., the solution to Eq. (5) neglecting the second term on the right hand side of the equation]. Since the contacting asperities are randomly distributed over the interface, we can estimate the second term (i.e., the summation over j) on the right hand side of Eq. (6) by its root mean square (rms) average which is estimated by integrating the square of the summand over the position of the jth asperity, which is in contact with the substrate, over its position and multiplying by the density of asperities in contact with the substrate ρ. Since the angular integrals give only a factor of order unity, we need evaluate only the integral over the magnitude of $R_{\ell,j}$, giving an rms value of the sum over R^{-1} of order $[\rho \ln(W/a)]^{1/2}$ where here W is the width of the interface and a is the asperity size. For $W \approx 1$ cm

and $a \approx 10^{-6}$ cm, $[\ln(W/a)]^{1/2}$ is of order unity. Since the shearing of the junction at the area of contact of two asperities involves the motion of two atomic planes relative to each other, the distance over which the contact potential varies must be of the order of atomic distances. Then, if we denote the width of the asperity contact potential well by b, of the order of an atomic spacing, we must choose a typical value for α such that αb is of the order of the shear rupture strength of the asperity contact junction. Thus, $\alpha_\ell \gg E'a$. Hence, $\mathbf{u}_\ell^0 \approx \mathbf{\Delta}_\ell$, implying that in the low asperity density limit, all asperities lie in their potential minima. The asperity-asperity interaction can be estimated from the energy of the system, which can be written as

$$(1/2)\sum_j \alpha_j |\mathbf{\Delta}_j - u_j|^2 + (1/2)E' \sum_j \int d^3r\, |\nabla \mathbf{G}(\mathbf{r}) \cdot [\alpha_j(\mathbf{\Delta}_j - \mathbf{u}_j)]|^2. \tag{7}$$

It follows from Eqs. (5)–(7) that the two lowest order nonvanishing terms in an expansion of the energy of the system in powers of $\rho^{1/2}$ are the zeroth and first order ones. Zeroth order in the asperity density in Eq. (7) is of the order of $\alpha\Delta^2$, where α is a typical value of α_j, and Δ is a typical value of Δ_j. The term linear in $\rho^{1/2}$ is easily shown to be of the order of $E'a^2\Delta^2\rho^{1/2}$ to zeroth order in $E'a/\alpha$. Since it depends on ρ it represents an interaction energy between the asperities. Then, the mean interasperity interaction is proportional to the square root of the number of contacting asperities per unit surface area, given by

$$\rho(h) = (N/A)\int_h^\infty ds\, \phi(s), \tag{8}$$

where A is the total surface area and N is the total number of asperities whether in contact with the substrate or not. The integral in Eq. (4) divided by the integral in Eq. (8), which is proportional to the contact area per asperity and the square root of the integral in Eq. (8) are plotted as a function of the load, which is given in the GW model [12] by

$$F_L = (4/3)E'N(R_c/2)^{1/2}\sigma^{3/2}\int_h^\infty ds\, \phi(s)(s-h)^{3/2}, \tag{9}$$

in Fig. 1. [Since A_c and $\rho(h)$ are functions of h, and h can be determined from the functional relationship between h and F_L, they can be plotted as a function of F_L.] A Gaussian distribution is assumed here for $\phi(s)$ [i.e., $\phi(s) = (2\pi)^{-1/2}e^{-s^2/2}$]. Since the square root of Eq. (8) drops to zero in the limit of vanishing load, whereas Eq. (4) divided by Eq. (8) approaches a nonzero value, this implies that the interface will approach the regime in which the asperity-substrate interaction dominates over the interasperity interaction in the limit of vanishing load.

Let us now support our conclusions with some sample numerical calculations using typical values for the quantities which occur in the application of the GW model to this problem. Following Ref. [13], we choose $\sigma = 2.4 \times 10^{-4}$ mm and $R_c = 6.6 \times 10^{-2}$ mm, and assume that there is a density of 4.0×10^3 asperities/mm^2. Then by performing the integrals in Eqs. (4), (8), and (9), we find that for $F_L/A = 3.98 \times 10^{-4}$ N/mm^2, where A is the apparent area of the interface, the total contact area divided by A is 3.03×10^{-5}, and the contact area per asperity from the ratio of Eqs. (4) and (8) is 2.44×10^{-5} mm^2. Also, $\rho(h)^{1/2}$, which is equal to the square root of Eq. (8), is 1.11 mm^{-1}. The mean interasperity interaction force is approximately equal to the derivative of the first order term in $\rho^{1/2}$ in the energy given above Eq. (8) with respect to Δ or $E'a^2\rho(h)^{1/2}\Delta$, where a is taken as the square root of the mean contact area per asperity divided by π. The mean strength of the force acting on an asperity, due to the solid with which it is in contact, will be estimated by the product of its contact area and the shear rupture strength E_r. Then, the condition for the latter quantity to dominate over the asperity-asperity interaction, $E_r\pi a^2 > E'4\pi a^2\rho^{1/2}\Delta$ or $E_r/E' > 4\rho(h)^{1/2}\Delta$, is easily satisfied by the above calculated quantities since the right hand side is 4×10^{-7} and the left hand side cannot be too small since E_r, for typical asperities which are too small to have dislocations, is of the order of the shear modulus of the interface, which is the same order of magnitude as E'.

Although for higher loads the system appears to move towards the "weak pinning" limit, the latter conclusion is most likely incorrect because it does not take into account the fact that the distribution of asperity heights contains asperities which are much higher than average. These asperities will be compressed much more than a typical asperity, making the friction force on them considerably

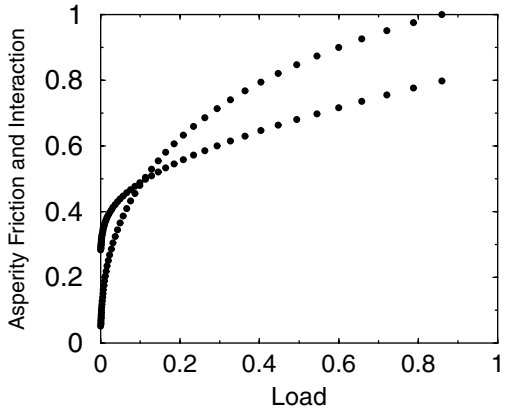

FIG. 1. The curve which is lower at the right is a plot of the integral in Eq. (4) divided by the integral in Eq. (8), which is proportional to the area per asperity. The curve which is higher on the right is a plot of the square root of the integral in Eq. (8), proportional to the square root of the asperity density, versus the integral in Eq. (9), proportional to the load. All quantities are dimensionless.

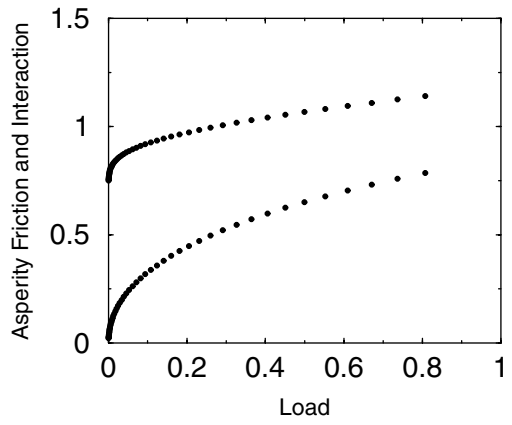

FIG. 2. Equation (11) (the higher curve), proportional to the contact area per asperity of height $h > h_L$, and the square root of Eq. (10) (the lower curve), proportional to the square root of the asperity density, are plotted versus the load, given by Eq. (9), divided by $(4/3)E(b/2)^{1/2}\sigma^{3/2}$. All quantities are dimensionless.

larger than average. Since the probability of such asperities occurring is relatively small, however, they will be typically far apart, putting them in the strong pinning limit. For example, the probability of the ratio of an asperity height to σ being greater than h by an amount h_L is

$$P(h_L) = \int_{h+h_L}^{\infty} ds\, \phi(s), \tag{10}$$

whose mean height and hence contact area is proportional to

$$P(h_L)^{-1} \int_{h+h_L}^{\infty} ds\, \phi(s)(s - h). \tag{11}$$

These two quantities are plotted in Fig. 2. It is seen that even for h_L only equal to $1/2$, Eq. (11) remains larger than the square root of Eq. (10).

We have argued that the asperities are essentially uncorrelated. They will still not produce static friction, unless they exhibit multistability [9,16]. The condition for multistability to occur at an interface [9], namely that the force constant due to the asperity contact potential be larger than that due to the elasticity of the asperity ($\approx E'a$), however, is satisfied, as noted earlier.

In conclusion, when one considers atomically smooth surfaces, arguments based on Larkin domains indicate that the disorder at an interface between two nonmetallic elastic solids in contact will not result in static friction. When one applies such arguments to the distribution of asperities that occur on multimicron length scales, however, one finds that the asperities are virtually always in the "strong pinning regime," in which the Larkin domains are comparable in size to a single asperity. This accounts for the fact that there is almost always static friction. Muser and Robbins' idea [2,3], however, is not invalidated by this argument. Their result will still apply for a smooth crystalline interface. It will also apply in the present context to the contact region between two asperities, implying that for a clean interface the shear force between contacting asperities is proportional to the square root of the contact area. The GW model predicts for this case that the average force of friction is proportional to the 0.8 power of the load [11,12] and is considerably smaller for clean than for dirty interfaces.

I thank the Department of Energy (Grant No. DE-FG02-96ER45585).

[1] S. Aubry, in *Solitons and Condensed Matter,* edited by A. R. Bishop and T. Schneider (Springer, New York, 1978), p. 264.
[2] G. He, M. H. Muser, and M. O. Robbins, Science **284**, 1650 (1999).
[3] M. H. Muser and M. O. Robbins, Phys. Rev. B **61**, 2335 (2000); M. H. Muser, L. Wenning, and M. O. Robbins, Phys. Rev. Lett. **86**, 1295 (2001).
[4] H. Fukuyama and P. A. Lee, Phys. Rev. B **17**, 535 (1978).
[5] P. A. Lee and T. M. Rice, Phys. Rev. B **19**, 3970 (1979).
[6] A. I. Larkin and Yu. N. Ovchinikov, J. Low Temp. Phys. **34**, 409 (1979).
[7] B. N. J. Persson and E. Tosatti, in *Physics of Sliding Friction,* edited by B. N. J. Persson and E. Tosatti (Kluwer Academic Publishers, Boston, 1995), p. 179.
[8] V. L. Popov, Phys. Rev. Lett. **83**, 1632 (1999).
[9] C. Caroli and Ph. Noziere, Eur. Phys. J. B **4**, 233 (1998).
[10] D. S. Fisher, Phys. Rev. B **31**, 1396 (1985); Phys. Rev. Lett. **50**, 1486 (1983).
[11] J. B. Sokoloff (unpublished).
[12] J. A. Greenwood and J. B. P. Williamson, Proc. R. Soc. London A **295**, 3000 (1966).
[13] J. I. McCool, Wear **107**, 37 (1986).
[14] A. Volmer and T. Nattermann, Z. Phys. B **104**, 363 (1997).
[15] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity* (Pergamon Press, New York, 1970), p. 30.
[16] J. B. Sokoloff, Phys. Rev. B **23**, 1992 (1981). One way of understanding this is that if there is static friction, the sliding velocity of the solid will be nonzero only if a force above the force of static friction is applied to the body. Alternatively, we can view this as implying that the force of friction approaches a nonzero value as the center of mass velocity approaches zero. The arguments in this reference and Ref. [9] tell us that there must be multistability for this to occur.