Surface roughness and dry friction

J. B. Sokoloff
Department of Physics and Center for Interdisciplinary Research in Complex Systems, Northeastern University, Boston, Massachusetts 02115, USA.

(Received 26 August 2011; revised manuscript received 6 February 2012; published 24 February 2012)

Persson's multiscale contact mechanics theory combined with a multiscale Brillouin-Prandtl-Tomlinson model is used to show that on the basis of these models "dry friction" [i.e., kinetic friction that remains at exceedingly small velocities (but still above the creep range) close to its value at higher velocities] should almost always occur for self-affine surfaces when the dominant interaction between two surfaces in contact is due to interatomic hard core repulsion, except for extremely smooth surfaces (i.e., surfaces with a Hurst index very close to 1).

DOI: 10.1103/PhysRevE.85.027102 PACS number(s): 81.40.Pq, 68.35.Af, 46.55.+d

I. INTRODUCTION

It is difficult to conceive of a model for "dry friction," i.e., sliding friction that is nonzero in the limit of zero sliding velocity, without postulating that even in that limit there is a good deal of rapid motion taking place at the interface. The possibility of explaining the occurrence of dry friction due to Prandtl-Brillouin-Tomlinson model [1–3] instabilities was discussed in some detail by Caroli and Noziere [4] and Tanguy and Noziere [5]. It was pointed out in their work that since asperities at an interface are generally short and fat, rather than tall and thin [6], they are unlikely to possess the multistable equilibrium that is required for the existence of Brillouin-Prandtl-Tomlinson instabilities [1–3]. They are also unlikely to interlock. As the load increases, the contribution to the friction coefficient from hard-core interactions will dominate over adhesive interactions because the former increase with load but the latter do not grow with increasing load.

It is quite likely that almost all surfaces possess roughness on multiple length scales, and there is experimental evidence that many surfaces are self-affine [7,8]. In Ref. [9], an explanation for the occurrence of dry friction was proposed based on a scaling approach and a multiscale contact mechanics theory put forward by Persson [10]. Here it will be pointed out that the argument given in Ref. [9] is stronger than what was presented there, allowing us to conclude that the approaches of Refs. [9,10] predict that there should almost always be dry friction for self-affine surfaces, as is commonly observed, unless the surfaces are extremely smooth.

Models for dry friction based on earthquake models [11,12] have been proposed as microscopic models for the occurrence of dry friction [13,14], which postulate the existence of local bonds that break, resulting in instabilities like those in the Prandtl-Brillouin-Tomlinson model [1–3], but since these bonds are introduced without providing a microscopic mechanism for them, they do not provide a true microscopic explanation for the occurrence of dry friction. The explanation that will be proposed here deals with a situation (likely to occur in many applications) in which the interactions at the interface are dominated by hard core interactions between atoms. The proposed model uses Persson's multiple-length-scale contact mechanics theory [10] to show that rough surfaces posses asperities on all length scales down to atomic scales. As a result, all of the load must be supported by the smallest length-scale asperities. Consequently, there is a very large contribution to the normal force provided by each of these smallest length-scale asperities. This contribution is provided by the hard core repulsion of the atoms that are in contact. Since these hard core repulsions must also have components along the interface of comparable magnitude, the smallest length-scale asperities are highly likely to exhibit the multistability required for the occurrence of dry friction resulting from the Brillouin-Prandtl-Tomlinson model [1–3]. Here the focus is on the friction that occurs for very slow sliding velocity (but still above the thermally activated creep regime), in contrast to the small but finite velocity regime considered in Ref. [15].

II. A BRIEF SUMMARY OF HOW MULTISTABILITY CAN RESULT FROM MULTISCALE ROUGHNESS

In order to present an explanation for why dry friction almost always occurs, in the interest of clarity, let us first summarize the treatment discussed in Ref. [9] briefly. The improvements to Ref. [9] of this Brief Report are presented in the next section. Let L_{n_m} represent the length (and width) of the interface as viewed at the largest length scale, meaning that the length and width of the interface are of the order of L_{n_m}. We use the usual approach of replacing the problem of two rough surfaces in contact by a rough surface in contact with a smooth surface [6]. Then, the interface possesses a collection of hills or asperities of length scale L_{n_m-1}, less than L_{n_m}, a fraction c_{n_m-1} of which are in contact. Each of these asperities is covered with asperities of length scale L_{n_m-2} ($L_{n_m-2} < L_{n_m-1}$), a fraction c_{n_m-2} of which are in contact. This is illustrated schematically if Fig. 1. We continue going down to smaller length scales until we reach L_0 and then a, the atomic length scale. In contrast to Ref. [10] in which L_n for smaller values of n represents larger length scales, here we take L_n for smaller values of n to represent smaller length scales. It was shown in Ref. [10] that the contact area at the nth length scale is given by

$$\frac{A(\zeta_n)}{A} = P(\zeta_n), \tag{1}$$

where $\zeta_n = L_{n_m}/L_n$, $A = L_{n_m}^2$, and

$$P(\zeta_n) = \int_0^{\sigma_y} d\sigma\, P(\zeta_n,\sigma), \tag{2}$$

where $P(\zeta_n,\sigma)$ is the distribution of stresses σ over the interface at the length scale denoted by ζ_n and σ_y is the yield

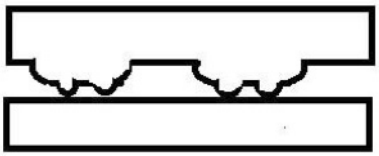

FIG. 1. This is a schematic illustration of the asperity hierarchy on the top surface sliding on a flat substrate (i.e., the bottom block). (Real asperities have arbitrary shapes, as opposed to the rectangular shapes shown in this schematic representation.) Each asperity of a given order has a number of (smaller) asperities of one order lower on its surface. In turn, each of these lower order asperities has a number of (smaller) asperities of one order lower. This continues until we reach the zeroth order asperity, whose surface consists of atoms. Only three orders of asperities are illustrated here.

strength of the material. In the above treatment of multiscale roughness, there are a total of

$$c_{n_m-1}\left(L_{n_m}/L_{n_m-1}\right)^2 c_{n_m-2}\left(L_{n_m-1}/L_{n_m-2}\right)^2 \times c_{n_m-3}\left(L_{n_m-2}/L_{n_m-3}\right)^2 \cdots c_n (L_{n+1}/L_n)^2 \quad (3)$$

nth length-scale asperities in contact with the second surface. (What we mean by being in contact is that all L_{n-1} length-scale asperities have atoms on them that are in contact with atoms on the second surface.) Then, $A(\zeta_n)/A$ is the product of the number of nth length-scale asperities given in Eq. (3) and L_n^2/A, and since $A = L_{n_m}^2$, we obtain

$$P(\zeta_n) = A(\zeta_n)/A = c_{n_m-1}c_{n_m-2}c_{n_m-3}\cdots c_n. \quad (4)$$

In order to write down a condition for the occurrence of multistability, it is argued in Ref. [9] that the elastic energy of a typical nth length-scale asperity is of the order of $L_n' L_n^2 K(\Delta x_n/L_n')^2$, where L_n' and L_n are the mean height and length of a typical nth length-scale asperity, respectively, Δx_n is the displacement at the top of the asperity due to shearing, and K is the shear modulus of the material. In order to treat the interaction of the asperities with the second surface, let us first discuss the interaction of the $n = 0$ length-scale asperities with the second surface due to hard core interactions. The product of the component normal to the surface of the hard core repulsive force between the atoms belonging to the two surfaces that are in contact F times the number of atoms in contact N_a must be equal to the total load supported by the surface, which is equal to σA, where σ is the total load divided by the surface area (at the n_m length scale), and hence $FN_a \sim \sigma A$. Then, clearly, $N_a \sim c(A/a^2)$, and hence $F \sim \sigma a^2/c$, where c, the fraction of surface atoms in contact with the second surface, is equal to $c_{n_m-1}c_{n_m-2}\cdots c_0 c_a$, where c_a is the fraction of atoms on the zeroth order length-scale asperities that are in contact with second surface. Then the mean potential energy of interaction between an atom on an $n = 0$ asperity and the second surface must have the approximate form $V_0 f_0(\Delta x_0/a)$, where f_0 is a function whose value is of order unity that varies over a range of its argument of order unity, and $V_0 \sim \sigma a^3/c$ since the work done by the component of the hard core interaction along the interface is of the order of Fa. To determine the net force due to the second surface on an $n = 0$ asperity, we have two possibilities. Either the $n = 0$ order asperities are in the strong pinning regime, in which each of the atoms can displace enough to lie close to the potential minima of the second surface, in which case, the potential acting on an $n = 0$ asperity due to the second surface is $c_a(L_0/a)^2 V_0 f_0(\Delta x_0/a)$, or the asperity is in the weak pinning regime [16], in which the atoms cannot sink to their potential minima and hence the hard core forces acting on them cannot act together. (In the strong pinning regime, the atoms can displace under the action of the substrate potential so that each atom lies in a substrate potential minimum. In contrast, for weak pinning the interaction of the atoms with each other is strong enough to prevent the atoms from sinking into substrate potential minima. As a result, the interaction of an $n = 0$ asperity with the second surface will be proportional to the number of asperity atoms in contact with the second surface in the strong pinning regime and proportional to the square root of the number of asperity atoms in contact with the second surface for weak pinning.) It will be shown later that the small length-scale asperities are almost always in the strong pinning regime, and hence, weak pinning will not be considered. Applying the same arguments to the $n = 1$ asperities, the elastic energy is given by $L_1' L_1^2 K(\Delta x_1/L_1')^2$, and the interaction of the asperities on the second surface is $c_a(L_0/a)^2 c_0(L_1/L_0)^2 V_0 f_1(\Delta x_1/a)$. Continuing to the $n = 2$ length-scale asperities and so on, we have, for the elastic energy and interaction with the second surface for the nth order asperities,

$$L_n' L_n^2 K(\Delta x_n/L_n')^2 \quad (5)$$

and

$$[c_a(L_0/a)^2 c_0(L_1/L_0)^2 \cdots c_{n-1}(L_n/L_{n-1})^2] V_0 f_n(\Delta x_n/a) \quad (6)$$

for the interaction of the asperities with the second surface for strong pinning.

In order to determine whether there is multistability at the nth length scale, we minimize the sum of the asperity elastic potential energy and its interaction with the second surface, which gives

$$(Ka^3/V_0)(c_a c_0 \cdots c_{n-1})^{-1}(a/L_n')\Delta x_n/a \sim f_n'(\Delta x_n/a) \quad (7)$$

if the nth order asperities are in the strong pinning regime. The solution of Eq. (7) gives the equilibrium values of Δx_n, as is illustrated in Fig. 2.

For strong pinning, the straight line representing the left hand side of Eq. (7) intersects the curve represented by the right hand side more than once. This condition becomes $(V_0/Ka^3)(c_a c_0 \cdots c_{n-1})(L_n'/a) > 1$, which gives

$$\sigma > (a/L_n')\left(c_n c_{n+1} \cdots c_{n_m-1}\right)K = (a/L_n')\left(\frac{A(\zeta_n)}{A}\right)K. \quad (8)$$

Since as n decreases (i.e., we consider smaller and smaller length scales) $A(\zeta_n)$ decreases, it is clear that as we move to smaller and smaller length scales, it becomes easier to satisfy the above conditions for the occurrence of the multistability of the asperities on that length scale.

It has been argued in the context of friction force microscope tips sliding over surfaces that thermal activation out of potential wells in which the tip resides can sometimes

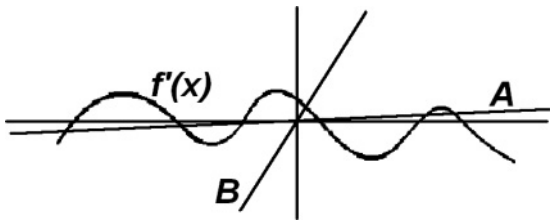

FIG. 2. Illustration of the solution of Eq. (7) for Δx_n. The function $f'(x)$ is a schematic illustration of the functions f'_0, f'_1, and f'_n, and the x axis denotes Δx_n. Lines A and B represent the left hand side of Eq. (7) for $(Ka^3/V_0)(c_a c_0 \cdots c_{n-1})^{-1}(L_n/L'_n)$ less than and greater than 1, respectively. For the situation illustrated by line A, there are multiple solutions, while for the situation illustrated by line B, there is only one.

allow the tip to slide without rapid stick-slip motion [17]. From the discussion following Eq. (4), we know that in the strong pinning regime a typical L_0 length-scale asperity lies in a potential due to the second surface with a corrugation strength of the order of $V_0 c_a (L_0/a)^2 \sim (\sigma/c)a^3 c_a (L_0/a)^2$, where $c = c_{n_m-1}c_{n_m-2}\cdots c_0 c_a = c_a A(\zeta_0)/A$, from Eqs. (1) and (4). Then, the corrugation felt by an L_0 length-scale asperity can be written as $[F_n/A(\zeta_0)]L_0^2 a$ since the total normal force $F_n = \sigma A$. Equation (C14) in the first paper cited in Ref. [10] gives $A(\zeta_0)/A = [2\sigma L/(\pi E^* h_0)][(1-H)/\pi H]^{1/2}(\zeta_0)^{-(1-H)}$. Substituting this in the above expression, we obtain, for the corrugation, $(1/2)(\pi h_0/L)E^* a^3[\pi H/(1-H)]^{1/2}(L/a)^{1-H}(L_0/a)^2$ since $\zeta_0 = L/a$ (where $L = L_{n_m}$), $E^* = E/(1-\nu^2)$, where E is Young's modulus and ν is Poisson's ratio. For $E^* \sim 10^{11}$ dyn/cm^2, $a = 3\times 10^{-8}$ cm, $\zeta_0 \sim 10^8$, $h_0/L \sim 10^{-2}$, $L_0/a \sim 10$, and $H = 0.8$, the latter expression gives, for the corrugation of an L_0 length-scale asperity, about 10 eV. For such a large well depth, the Boltzmann factor is much too small at room temperature for thermal activation of the well to play a significant role in the sliding motion.

III. AN EXPLANATION FOR WHY DRY FRICTION ALMOST ALWAYS OCCURS

Using Eq. (C14) of Persson's contact mechanics theory [10] to determine $A(\zeta_n)/A$, the condition for multistability given in Eq. (8) becomes

$$\left(\frac{L}{L_n}\right)^{1-H} > \frac{2K}{\pi E^*}\left(\frac{1-H}{\pi H}\right)^{1/2}\frac{L}{h_0}\frac{a}{L'_n}, \tag{9}$$

where $L = L_{n_m}$. Since L/L_n can be as large as 10^8, the inequality in Eq. (9) is not difficult to satisfy if H is not too close to 1 because, although $L/h_0 \approx 100$ by Persson's estimates, a/L'_n can be quite small, except for very small n asperities. (The ratios of the other quantities on the right hand side of the inequality are of order 1.) The physical reason for why the inequality in Eq. (8) is satisfied under most conditions is that $A(\zeta_n)/A$ decreases as L_n/L decreases but is proportional to σ, implying that as σ increases, small-scale asperities do not flatten out by a larger percentage than larger length-scale asperities. The reason that surfaces with an index H close to 0 are rough and those H close to 1 are smooth is that Persson [10] treats the surfaces as self-affine [8], which means that if we look at the surfaces on a length scale L_{n-1} instead of L_n, the coordinates x and y (along the surface) get reduced by L_{n-1}/L_n, whereas distances normal to the surface, such as the asperity heights, only get reduced by a factor $(L'_{n-1}/L'_n)^H$, with $H < 1$. Consequently, when viewed on smaller length scales, the asperities appear taller compared to their spacing. In contrast, the inequality in Eq. (9) can be difficult to satisfy, except for values of H not too close to 1 (which are very rough surfaces), because $L/h_0 \approx 100$.

From the results of Ref. [9] we conclude that the $n = 0$ asperities will be in the weak pinning regime if $\sigma_0 < c_a^{1/2}K$, where σ_0 is the mean load carried by an $n = 0$ asperity, which is given by $\sigma_0 = \sigma/(c_0 c_1 \cdots c_{n_m-1})$, or $\sigma < c_a^{1/2}(c_0 c_1 \cdots c_{n_m-1})K$. By similar arguments, it follows that the condition for the n^{th} asperity to be in the weak pinning regime is $\sigma_n < c_{n-1}^{1/2}K$, where σ_n is the mean load carried by an n^{th} length scale asperity, or $\sigma \ll c_{n-1}^{1/2}(c_n c_{n+1}\cdots c_{n_m-1})K$. Since the condition for weak pinning for $n = 0$ is obviously easier to satisfy than that for higher values of n, in order to determine whether any asperities are likely to be weak pinning, let us examine the condition for weak pinning of the $n = 0$ asperities. The condition for weak pinning of the $n = 0$ asperity, using the fact that from the relationship given above

$$P(\zeta_n) = A(\zeta_n)/A = c_{n_m-1}c_{n_m-2}c_{n_m-3}\cdots c_n, \tag{10}$$

and hence the condition for the $n = 0$ asperity being in the weak pinning regime can be written as

$$\sigma < c_a^{1/2}\frac{A(\zeta_0)}{A}. \tag{11}$$

Using Persson's expression [10] for $A(\zeta_0)/A$, Eq. (11) can be written as

$$\left(\frac{L}{a}\right)^{1-H} < c_a^{1/2}\frac{2K}{\pi E^*}\left(\frac{1-H}{\pi H}\right)^{1/2}\frac{L}{h_0}, \tag{12}$$

which, except for H very close to 1 and c_0 extremely small, is difficult to satisfy since L/a can be as large as 10^8. Thus, we conclude that, in most cases, the asperities will all be in the strong pinning regime, and for that regime, Eq. (9) tells us that at least the smallest length-scale asperities will be multistable, except for the case of extremely smooth surfaces.

Because the small length-scale asperities often carry very high loads, we expect that most of them will deform plastically. If an nth length-scale asperity deforms plastically under load, shearing of the asperity should also occur plastically [18]. As a consequence, the restoring force when the asperity is sheared by an amount Δx_n will no longer be linear in Δx_n as it was when the asperity was assumed to deform elastically. Instead, the left hand side of Eq. (7) will now be a sublinear function of Δx_n, and the solution of that equation illustrated in Fig. 2 will be more likely to have multiple solutions, implying that the asperity is more likely to exhibit multistability. Thus, when plasticity is taken into account, the condition for the occurrence of dry friction is even more likely to be satisfied. In Appendix C of Ref. [10], it is shown that about 50% of the

asperities at a length scale L_n will have yielded plastically if the condition

$$\frac{H}{1-H}\left(\frac{\pi^2 E h_0}{2(1-\nu^2)L\sigma_Y}\right)^2 [(L/L_n)^{2(1-H)} - 1] \approx 1 \quad (13)$$

is satisfied. It was argued in Ref. [10] that this condition will certainly be satisfied for L_n at atomic length scales (i.e., L_n close to a).

IV. CONCLUSIONS

It is demonstrated on the basis of the Brillouin-Prandtl-Tomlinson model [1–3] combined with a multiscale contact mechanics theory [10] that when the interaction between two surfaces is due to interatomic hard core interaction, self-affine surfaces will exhibit friction that remains close to its value at higher velocity for exceedingly low velocities (but still above the thermally activated creep regime), except possibly those with a Hurst index close to 1.

[1] M. Brillouin, *Notice sur les Travaux Scientifiques* (Gautier-Villars, Paris, 1909).
[2] L. Prandtl, Z. Angew. Math. Mech. **8**, 85 (1928).
[3] G. A. Tomlinson, Philos. Mag. **7**, 905 (1929).
[4] C. Caroli and Ph. Nozieres, Eur. Phys. J. B **4**, 233 (1998).
[5] A. Tanguy and Ph. Noziere, J. Phys. I France **6**, 1251 (1996).
[6] K. L. Johnson, in *Contact Mechanics* (Pergamon, New York, 1985), pp. 90–95.
[7] J. Krim and G. Palasantzas, Int. J. Mod. Phys. B **9**, 599 (1995).
[8] A.-L. Barabasi and H. E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995).
[9] J. B. Sokoloff, Phys. Rev. E **78**, 036111 (2008).
[10] B. N. J. Persson, J. Chem. Phys. **115**, 3840 (2001); Phys. Rev. Lett. **87**, 116101 (2001).
[11] R. Burridge and L. Knopoff, Bull. Seismol. Soc. Am. **57**, 341 (1967).
[12] Z. Olami, H. J. S. Feder, and K. Christensen, Phys. Rev. Lett. **68**, 1244 (1992).
[13] O. M. Braun and M. Peyrard, Phys. Rev. Lett. **100**, 125501 (2008); Phys. Rev. E **82**, 036117 (2010).
[14] B. N. J. Persson, Phys. Rev. B **51**, 13568 (1995).
[15] J. R. Rice and A. L. Ruina, J. Appl. Mech. **50**, 343 (1983).
[16] J. B. Sokoloff, Phys. Rev. E **73**, 016104 (2006).
[17] K. B. Jinesh, S. Yu. Krylov, H. Valk, M. Dienwiebel, and J. W. M. Frenken, Phys. Rev. B **78**, 155440 (2008); D. G. Abel, S. Yu. Krylov, and J. W. M. Frenken, Phys. Rev. Lett. **99**, 166102 (2007); M. Muser, M. Urbakh, and M. O. Robbins, Adv. Chem. Phys. **126**, 187 (2003).
[18] A. Anandarajah, *Computational Methods in Elasticity and Plasticity: Solids and Porus Media* (Springer, New York, 2010).