Theory of rubber friction and contact mechanics

B. N. J. Persson

Institut für Festkörperforschung, Forschungszentrum Jülich, D-52425 Jülich, Germany

When rubber slides on a hard, rough substrate, the surface asperities of the substrate exert oscillating forces on the rubber surface leading to energy "dissipation" via the internal friction of the rubber. I present a discussion of how the resulting friction force depends on the nature of the substrate surface roughness and on the sliding velocity. I consider in detail the case when the substrate surface has a self affine fractal structure. I also present a theory for the area of real contact, both for stationary and sliding bodies, with elastic or elastoplastic properties. The theoretical results are in good agreement with experimental observation. © *2001 American Institute of Physics.*
[DOI: 10.1063/1.1388626]

I. INTRODUCTION

The nature of the friction when rubber slides on a hard substrate is a topic of considerable practical importance, e.g., for the construction of tires,[1] wiper blades,[1] and in the cosmetic industry. Rubber friction differs in many ways from the frictional properties of most other solids. The reason for this is the very low elastic modulus of rubber and the high internal friction exhibited by rubber over a wide frequency region.

The pioneering studies of Grosch[2] have shown that rubber friction in many cases is directly related to the internal friction of the rubber. Thus experiments with rubber surfaces sliding on silicon carbide paper and glass surfaces give friction coefficients with the same temperature dependence as that of the complex elastic modulus $E(\omega)$ of the rubber. In particular, there is a marked change in friction at high speeds and low temperatures, where the rubber's response is driven into the so-called glassy region. In this region, the friction shows marked stick-slip and falls to a level of $\mu \approx 0.4$, which is more characteristic of plastics. This proves that the friction force under most normal circumstances is directly related to the internal friction of the rubber, i.e., it is mainly a *bulk property* of the rubber.[2]

The friction force between rubber and a rough (hard) surface has two contributions commonly described as the adhesion and hysteretic components, respectively.[1] The hysteretic component results from the internal friction of the rubber: during sliding the asperities of the rough substrate exert oscillating forces on the rubber surface, leading to cyclic deformations of the rubber, and to energy "dissipation" via the internal damping of the rubber. This contribution to the friction force will therefore have the same temperature dependence as that of the elastic modulus $E(\omega)$ (a bulk property). The adhesion component is important only for clean and relative smooth surfaces.

Because of its low elastic modulus, rubber often exhibit elastic instabilities during sliding. The most well-known involves the compressed rubber surface in front of the contact area undergoing a buckling which produces detachment waves which propagate from the front-end to the back-end of the contact area. These so called Schallamach waves[3] occur mainly at "high" sliding velocity and for very smooth surfaces, but will not be considered further in this paper.

In three earlier papers we have studied both the adhesion and hysteretic components of rubber friction.[4–6] Other studies of this topic are presented in Refs. 1, 7–9; reference 4 considered only the interaction between a flat rubber surface and a single surface asperity (or many identical asperities). In Ref. 6 we studied the hysteretic contribution to the friction for viscoelastic solids sliding on hard substrates with different types of (idealized) surface roughness.

In this paper I develop a theory of rubber friction when a rubber block is slid over a hard rough surface, with roughness on many different length scales λ. The theory is valid for arbitrary (random) surface roughness, but explicit results are presented for self affine fractal surface profiles.[10,11] Such surfaces "looks the same" when magnified by a scaling factor ζ in the xy-plane of the surface and by a factor ζ^H (where $0 < H < 1$) in the perpendicular z-direction. I note that many materials of practical importance have (approximately) self-affine fractal surfaces. Thus, for example, road surfaces and the surfaces of many cleaved brittle materials tend to be self affine fractal with the fractal dimension $D_f = 3 - H \approx 2.2 - 2.5$. In practice there is always a lower, λ_1, and upper, λ_0, cutoff length, so that the surface is self-affine fractal only when viewed in a finite length scale interval $\lambda_1 < \lambda < \lambda_0$. For surfaces produced by brittle fracture, the upper cut off length λ_0 is usually identical to the lateral size L of the fracture surface. This seems also to be the case for many surfaces of engineering importance (see, e.g., Ref. 14). However, for road surfaces the upper cutoff λ_0 is of order a few mm, which corresponds to the size of the largest sand particles in the asphalt. Less is known about the short distance cutoff λ_1, but I will argue later that in the context of rubber friction it may be taken to be of order a few μm, so that the length scale region over which the road surface may be assumed to be fractal may extend over ~ 3 orders of magnitude.

When rubber slides on a hard rough surface with roughness on the length scales λ, it will be exposed to fluctuating forces with frequencies $\omega \sim v/\lambda$. Since we have a wide distribution of length scales $\lambda_1 < \lambda < \lambda_0$, we will have a corre-

      3840

FIG. 2. Rubber sliding on 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.
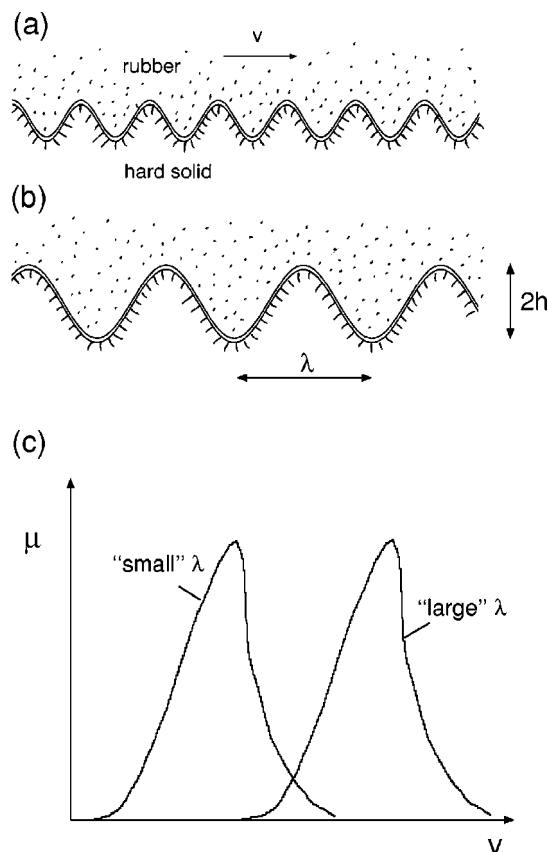

FIG. 1. Rubber (dotted area) sliding on a hard corrugated substrate. The magnitude of the contribution to the friction from the internal damping in the rubber is the same in (a) and (b) because the ratio between the amplitude and the wavelength of the corrugation is the same. (c) shows the $\mu(v)$ curves for the roughness profiles in (a) and (b) (schematic).

sponding wide distribution of frequency components in the Fourier decomposition of the surface stresses acting on the sliding rubber block. The contribution to the friction coefficient μ from surface roughness on the length scale λ, will be maximal when $v/\lambda \approx 1/\tau$, where $1/\tau$ is the frequency where $\mathrm{Im}\,E(\omega)/|E(\omega)|$ is maximal, which is located in the transition region between the rubbery region (low frequencies) and the glassy region (high frequencies). We can interpret $1/\tau$ as a characteristic rate of flips of molecular segments (configurational changes), which are responsible for the visco-elastic properties of the rubber. Since the flipping is a thermally activated process it follows that τ depends exponentially (or faster) on the temperature $\tau \sim \exp(\Delta E/k_B T)$, where ΔE is the barrier involved in the transition. In reality, there is a wide distribution of barrier heights ΔE and hence of relaxation times τ, and the transition from the rubbery region to the glassy region is very wide, typically extending over 3 orders of magnitude in frequency.

The following observation is of great importance for rubber friction. Consider the contribution to the rubber friction from surface roughness of different wavelength λ and amplitude h, see Fig. 1. If we assume that the applied pressure is so high that the rubber is squeezed into complete contact with the substrate, it follows from dimensional arguments that the magnitude of the hysteretic contribution to the friction coefficient only depends on h/λ, i.e., *surface rough-*
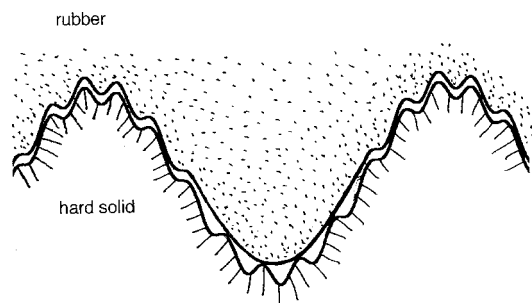
ness of different length scale contribute equally to the friction force if the ratio between the amplitude and wavelength is constant. Thus, roughly speaking, we may state that surface roughness of all length scales are equally important. Of course, the different wavelength contributions to $\mu(v)$ will peak at different sliding velocities (determined by $v/\lambda \approx 1/\tau$), i.e., the different wavelength contributions to $\mu(v)$ are shifted relative to each other along the v-axis, see Fig. 1(c). We may summarize these results by writing $\mu = f(v\tau/\lambda, h/\lambda)$.

These profound results imply that it is very important not to *a priori* exclude any roughness length scale from the analysis. The distribution of different length scales λ will broaden the $\mu(v)$ curve, and also increase the peak maximum. However, let us note the following: Consider a surface with surface roughness on two different length scales as indicated in Fig. 2. 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. Hence, during sliding the small-scale roughness may give a contribution to the pulsating deformations of the rubber (and hence to the friction force), only at the top of the big asperities. This important fact is taken into account in the analysis presented in this paper. Thus, if $A(\lambda)$ is the (apparent) area of contact on the length scale λ [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 λ, see Fig. 3], then I will study 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 contact area [L is the diameter of the macroscopic contact area so that $A(L) \approx L^2$]. I will show that for an ideal elastic body (no plasticity) squeezed against a rigid self affine fractal surface without a short-distance cut off, $P(\zeta) \rightarrow 0$ as $\zeta \rightarrow \infty$. This result is
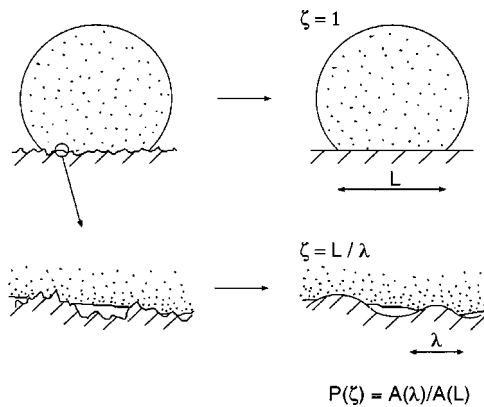
FIG. 3. A rubber ball squeezed against a hard, rough, substrate. Left: the system at two different magnifications. Right: the area of contact $A(\lambda)$ on the length scale λ is defined as the area of real contact when the surface roughness on shorter length scales than λ has been removed (i.e., the surface has been "smoothened" on length scales shorter than λ).
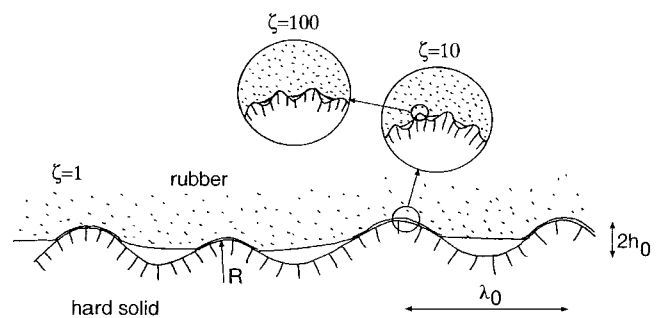

FIG. 4. Elastic contact between a flat rubber surface and a hard solid substrate. The surface is assumed to be self-affine fractal with an upper cutoff $\lambda_0 < L$. The system is shown on the length scale λ_0. Increasing the magnification shows that within an (apparent) contact area, the rubber will only make partial contact with the substrate (see text).

very important, because it shows that even without any short distance cut off it is possible for the sliding friction force to remain finite, since the rubber will not make contact with (or experience) the very short-wavelength surface roughness. Note that the longest roughness wavelength possible are of order $\sim L$. This correspond to the wave vector $q_L = 2\pi/L$. If we define $q = q_L\zeta$, we can consider $P(\zeta) = P(q/q_L)$ as a function of q; I denote this function by $P(q)$ for simplicity.

This paper focuses mainly on rubber friction, but, as indicated above, I also present a new theory of contact mechanics (see Appendices B and C), valid for randomly rough (e.g., self-affine fractal) surfaces. In the context of rubber friction, mainly elastic deformation will occur in the substrate–rubber contact areas. However, the contact theory developed in this paper can also be applied when both elastic and plastic deformation occur in the contact areas. This case is, of course, relevant to almost all materials other than rubber.

This paper is organized as follows: In Secs. II and III, I present some basic results related to self-affine fractal surfaces and contact theories, which form a necessary background for the theory developed in Secs. IV and V. In Sec. IV, I derive a general formula for the hysteretic contribution to rubber friction. This formula contains the function $P(\zeta)$ introduced above, which is derived in Sec. V and Appendix B for randomly rough (e.g., self-affine fractal) surfaces. Section VI contains numerical results for the velocity dependent friction coefficient. Section VII presents some general comments about rubber friction, and Sec. VIII is the summary and conclusion. In Appendix C, I present a new contact mechanics theory for randomly rough surfaces, when both elastic and plastic deformation occurs in the contact areas. In Appendix E, I study the contribution to rubber friction from the emission of elastic waves from the sliding interface.

II. SELF-AFFINE FRACTAL SURFACES AND CONTACT THEORIES

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,[10,11] see Fig. 4. Thus, the statistical properties of the surface are invariant under the scaling transformation,

$$x \to \zeta x, \quad y \to \zeta y, \quad z \to \zeta^H z, \tag{1}$$

where the exponent H can be related to the fractal dimension via $D_f = 3 - H$. Since we expect $2 < D_f < 3$ it follows that $0 < H < 1$. Recent studies have shown that asphalt road tracks are (approximately) self-affine in a finite surface roughness interval, with an upper cut-off of order a few mm.[12]

In order to study rubber friction on a hard self-affine fractal surface, it is first necessary to be able to describe the contact mechanics. A simple model of contact mechanics for fractal-like surfaces was studied as early as 1957 by Archard.[13,14] He showed that the area of real contact A is (nearly) proportional to the load (or normal force), $A \sim F_N$. In a recent series of papers by Roux et al.[15] and Bhushan and co-workers,[16] it is claimed that for self-affine surfaces the area of real contact depends nonlinearly on the load. Assuming only elastic deformation they found

$$A \sim F_N^{2/(1+H)}, \quad \text{or} \quad F_N \sim A^{(1+H)/2}. \tag{2}$$

Since $H < 1$ ($H = 1$ correspond to $D_f = 2$), these theories predict that the area of real contact increases faster than linear with the load. This is usually not observed experimentally. In my opinion, the theories of Roux et al. and of Bhushan and co-workers are based on questionable assumptions (see Appendix D). The contact theory developed in this paper (see Appendices B and C) predict $A \sim F_N$, unless the load F_N is so large that the contact area A is close to the nominal contact area A_0.

In the theory developed in this paper, the friction coefficient is given by a sum over different length scales. Now in most cases the upper limit in the sum is quite obvious. For example, for an asphalt road track the upper cutoff is of the order of a few mm (the typical grain sizes) as observed in surface profile measurements. In a recent measurement, an asphalt road surface was observed to be a self-affine fractal down to the shortest length-scale studied (approximately 0.03 mm).[8] The short distance cutoff in the sum over length scales may, however, not be determined by the intrinsic cut-
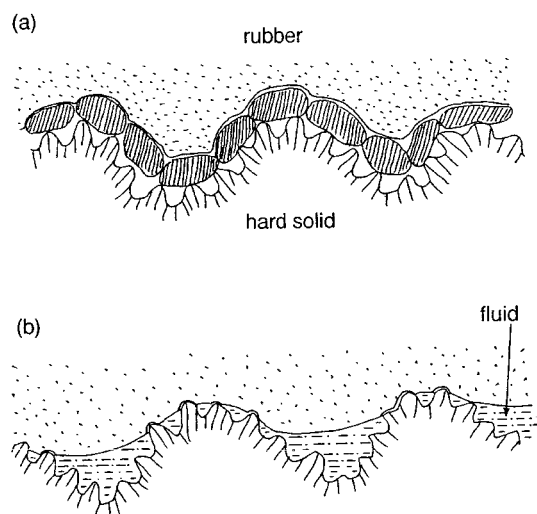
(a)

FIG. 5. Influence of contamination on the rubber–substrate interaction. Contamination particles (a), or trapped liquid (b), will inhibit the rubber to get squeezed into the small sized surface cavities.

FIG. 6. The kinetic friction coefficient for rubber sliding on a carborundum surface under different conditions (from Ref. 19).

off of the fractal nature of the surface (which could be an atomic distance), but by surface contamination,[17,18] (or compressed air pockets in small sized cavities), or by a thin "skin" on the rubber surface with strongly modified properties. For example, if the rubber surface is covered by small (uniformly sized) dust particles (e.g., talc or carbon or silica particles from the fillers, or pulverized stone from a road, or carbon particles from the automobile exhaust), then the low distance cutoff is obviously determined by the particle size [see Fig. 5(a)], since the particle covered rubber surface cannot penetrate into surface cavities smaller than the typical particle diameter. In fact, it is known that the tire-road friction increases when a road surface has dried up after a strong rain fall. Presumably, the rain washes away contamination particles from the road (and tire) surface. On the other hand, if the surface is covered by water or some other "lubrication" fluid (e.g., oil or grease), which fills out the small surface cavities, then the low distance cutoff will be determined by the smallest asperities which can penetrate above the contamination layer [Fig. 5(b)]. Thus, the contamination layer will remove the contribution to the energy dissipation from the small surface asperities and cavities, and reduce the friction force. This effect is illustrated in Fig. 6 with experimental results for a rubber block sliding on dry clean (dashed line), dusted (dashed–dotted), and wet (solid line) carborundum stone surfaces.[19] The figure also show results for wet surfaces with an added (5%) detergent. Roberts[20] has shown that polar substances like soaps prevent direct contact between track and rubber (see Sec. VII); this explain why the friction is slightly lower for the wet +5% detergent case, compared to the wet, clean carborundum surface.

III. AREA OF REAL CONTACT: QUALITATIVE DISCUSSION

I have already emphasized the importance of knowing the nature of the area of real contact when discussing rubber friction. In this section, I discuss some basic results of con-
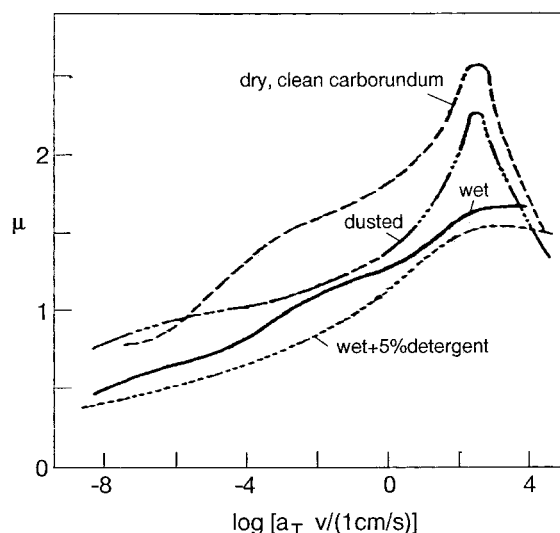
tact theory, which form a necessary background for the theory presented in Secs. IV, V, and Appendices B and C.

Consider first a flat rubber surface squeezed against a hard surface with a periodic corrugation with wavelength λ and amplitude (or height) h; see Fig. 7. If A_0 is the nominal contact area [i.e., the area of the (bottom) surface of the rubber block], and F_N the load, then we define the average perpendicular stress (or pressure) $\sigma_0 = F_N/A_0$. Let us now study under which conditions the load F_N, and the rubber–substrate adhesion forces, are able to deform the rubber so that it comes in direct contact with the substrate over the *whole* surface area A_0 [Fig. 7(b)], i.e., under which conditions the rubber is able to deform and fill out all the surface "cavities" of the substrate.

Assume first that a uniform stress σ acts within a circular area (radius R) centered at a point P on the surface of a
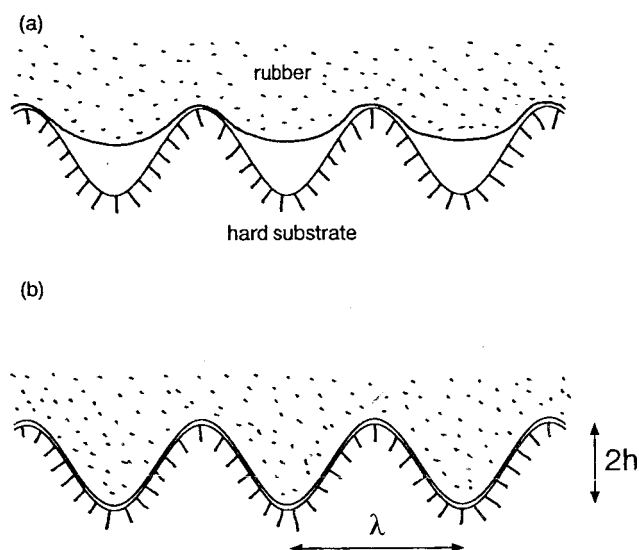
FIG. 7. A rubber block squeezed against a substrate with a cosines corrugation. In (a) the applied pressure is too small to squeeze the rubber into complete contact with the substrate, while in (b) it is high enough to do so.

semi-infinite elastic body with the elastic modulus E. This will give rise to a perpendicular displacement u of P by a distance which is easy to calculate using continuum mechanics, $u/R \approx \sigma/E$. This result can also be derived from simple dimensional arguments: First, note that u must be proportional to σ since the displacement field is linearly related to the stress field (we assume here, and in what follows, that linear elasticity theory is valid). However, the only other quantity in the problem with the same dimension as the stress σ is the elastic modulus E so u must be proportional to σ/E. Since R is the only quantity with the dimension of length we get at once $u \sim (\sigma/E)R$. Thus, with reference to Fig. 7, if $h/\lambda \approx \sigma_0/E$, the perpendicular pressure σ_0 will be just large enough to deform the rubber to make contact with the substrate everywhere.

In the case of passenger tires one typically has $\sigma_0 \approx 0.2$ MPa, and in the case of truck tires 0.8 MPa. This is at least one order of magnitude smaller than the (static or low-frequency) elastic modulus $E \approx 10$ MPa of filled rubbers (but only a little smaller than that of unfilled rubber where $E \approx 1$ MPa). We conclude that the pressure σ_0 is in general not able to deform the rubber to fill out the large surface cavities on a road, since in this case one typically has $h/\lambda \approx 1$, which according to the discussion above would require a local pressure of order $\sigma \approx E$. However, according to the contact theory of Greenwood,[14,17] the *average* pressure which acts in the rubber–substrate contact area at the *largest* asperities is of order $\approx (\Delta/R)^{1/2}E$, where Δ is the rms surface roughness amplitude and R the (average) radius of curvature of the largest surface asperities (see Fig. 4). Since for a road surface we expect $\Delta \approx R$ it is clear that the local pressure in the contact area of the large surface asperities will be of order of E, i.e., just large enough in order for the rubber to deform and *fill out at least some of the smaller sized surface cavities*. The way the (apparent) contact area varies with the observation length scale L/ζ is described by the function $P(\zeta)$.

Next, let us consider the role of the rubber–substrate adhesion interaction.[21] When the rubber deforms and fills out a surface cavity of the substrate, an elastic energy $E_{el} \approx E\lambda h^2$ will be stored in the rubber. Now, if this elastic energy is smaller than the gain in adhesion energy $E_{ad} \approx \Delta\gamma\lambda^2$ as a result of the rubber–substrate interaction (which usually is mainly of the van der Waals-type), then (even in the absence of the load F_N) the rubber will deform spontaneously to fill out the substrate cavities. The condition $E_{el} = E_{ad}$ gives[4,5] $h/\lambda \approx (\Delta\gamma/E\lambda)^{1/2}$. For the rough surfaces of interest here we typically have $h/\lambda \approx 1$, and with $E = 1$ MPa and the surface free energy change $\Delta\gamma = 3$ meV/Å^2 the adhesion interaction will be able to deform the rubber and completely fill out the cavities if $\lambda < 0.1$ μm. However, I believe that because of surface contamination there will be a low distance cut-off in the sum over length scales which is larger than 0.1 μm, and for this reason, in the context of the tire-road friction, I do not believe that the adhesion rubber–substrate interaction is important. The same conclusion has been reached by Fuller and Tabor in an experimental study of the dependence of rubber-substrate adhesion on the surface roughness.[22]

The discussion above is for stationary surfaces. During sliding we must take into account that the elastic modulus E depends on the perturbing frequency ω, and that $E(\omega)$ is a complex quantity with an imaginary part related to the internal friction of the rubber. In a first approximation we may still use the estimates presented above for the deformations induced by the largest asperities if we replace $E = E(0)$ with $|E(\omega)|$, where the frequency $\omega = v/\lambda_0$. Now, for a typical rubber at room temperature $E(\omega) \approx E(0)$ for $\omega < \omega_c = 10^5$ s^{-1}. When the frequency increases towards the glassy region $(\omega \sim 10^9$ s$^{-1})$, $|E(\omega)|$ increases by a factor of ~ 1000. In a typical case, for a tires sliding on a road with $v \approx 10$ m/s one gets $v/\omega_c \approx 0.1$ mm. Thus, the deformations induced by the largest asperities are relatively well described by using the low frequency elastic modulus $E(\omega) \approx E(0)$. However, the rubber will be much harder to deform by the small sized asperities since the effective elasticity $|E(\omega)|$ may (depending on the size of the asperities) be up to 1000 times higher than the low-frequency modulus. On the other hand, in the antilock braking system (ABS) of automobile tires on dry or wet road $v < 1$ cm/s in the incipient part of the footprint area, and in this case $v/\omega_c < 0.1$ μm so that surface cavities with linear size larger than 0.1 μm will experience relative "soft" rubber. These aspects of the (frequency-dependent) deformation of the rubber by the substrate asperities is taken fully into account in the theory developed below.

In Sec. V and Appendices B and C, I develop a new contact theory for surfaces with roughness on many different length scales. The contact theory of Greenwood was originally 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 (see, e.g., Ref. 23). 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. The contact theory developed in this paper is based on a completely different physical approach, and gives well defined results for surfaces with arbitrary surface roughness.

IV. SLIDING FRICTION

Using the theory of elasticity (assuming an isotropic elastic medium for simplicity), one can calculate the displacement field u_i on the surface $z=0$ in response to the surface stress distributions $\sigma_i = \sigma_{3i}$. Let us define the Fourier transform,

$$u_i(\mathbf{q},\omega) = \frac{1}{(2\pi)^3} \int d^2x\, dt\, u_i(\mathbf{x},t) e^{-i(\mathbf{q}\cdot\mathbf{x}-\omega t)},$$

and similar for $\sigma_i(\mathbf{q},\omega)$. Here $\mathbf{x} = (x,y)$ and $\mathbf{q} = (q_x, q_y)$ are two-dimensional vectors. In Appendix A, I have shown that

$$u_i(\mathbf{q},\omega) = M_{ij}(\mathbf{q},\omega)\sigma_j(\mathbf{q},\omega),$$

or, in matrix form,

$$\mathbf{u}(\mathbf{q},\omega) = M(\mathbf{q},\omega)\boldsymbol{\sigma}(\mathbf{q},\omega),$$

where the matrix (see Appendix A),

$$M = -\frac{i}{\rho c_T^2} \left(\frac{1}{S(q,\omega)} \left[Q(k,\omega)(\hat{z}\mathbf{q}-\mathbf{q}\hat{z}) \right. \right.$$

$$\left. \left. +\left(\frac{\omega}{c_T}\right)^2 (p_L\hat{z}\hat{z}+p_T\hat{q}\hat{q}) \right] + \frac{1}{p_T}\mathbf{ee} \right), \tag{3}$$

where $\hat{q}=\mathbf{q}/q$, $\mathbf{e}=\hat{z}\times\hat{q}$, and where

$$S = \left(\frac{\omega^2}{c_T^2}-2q^2\right)^2 + 4q^2 p_T p_L, \tag{4}$$

$$Q = 2q^2-\omega^2/c_T^2+2p_T p_L, \tag{5}$$

$$p_T = \pm\left(\frac{\omega^2}{c_T^2}\pm i\epsilon-q^2\right)^{1/2}, \quad p_L = \pm\left(\frac{\omega^2}{c_L^2}\pm i\epsilon-q^2\right)^{1/2}, \tag{6}$$

where the $+$ and $-$ sign refers to $\omega>0$ and $\omega<0$, respectively, and where ϵ is an infinitesimal positive number. In the equations above, ρ, c_T, and c_L are the mass density and the transverse and longitudinal sound velocities of the solid, respectively. Note that c_T and c_L are complex frequency dependent quantities given by

$$c_T^2 = \frac{E}{2\rho(1+\nu)}, \tag{7}$$

$$c_L^2 = \frac{E(1-\nu)}{\rho(1+\nu)(1-2\nu)}, \tag{8}$$

where $E(\omega)$ is the complex elastic modulus and $\nu(\omega)$ is the Poisson ratio.

We now assume that $|\nabla h(\mathbf{x})|<1$ [where $z=h(\mathbf{x})$ is the surface height profile] and that the surface stress $\boldsymbol{\sigma}(\mathbf{q},\omega)$ only acts in the z-direction so that

$$u_z(\mathbf{q},\omega) = M_{zz}(\mathbf{q},\omega)\sigma_z(\mathbf{q},\omega), \tag{9}$$

where

$$M_{zz} = \frac{-i}{\rho c_T^2}\frac{p_L}{S(q,\omega)}\left(\frac{\omega}{c_T}\right)^2. \tag{10}$$

Since in the present case $\omega=vq$ we get $\omega/c_T q=v/c_T\ll 1$ in most cases of practical interest. Thus, we can expand to leading order in $\omega/c_T q$. This gives

$$Q \approx \omega^2/c_L^2,$$

$$S \approx 2q^2\omega^2\left(\frac{1}{c_L^2}-\frac{1}{c_T^2}\right),$$

and

$$p_T \approx iq\left(1-\frac{\omega^2}{2c_T^2 q^2}\right),$$

and similar for p_L. Thus, we get

$$M_{zz} = \frac{-i}{\rho c_T^2}\frac{p_L}{S(q,\omega)}\left(\frac{\omega}{c_T}\right)^2 \approx -\frac{1}{2\rho c_T^2 q}\left[1-\left(\frac{c_T}{c_L}\right)^2\right]^{-1}, \tag{11}$$

so that, using Eqs. (7), (8), and (11),

$$(M_{zz})^{-1} = -\frac{Eq}{2(1-\nu^2)}. \tag{12}$$

It is interesting to note that if, instead of assuming that the surface stress act in the z-direction, we assume that the displacement $\mathbf{u}$ point along the z-direction, then

$$\sigma_z(\mathbf{q},\omega) = (M^{-1})_{zz}(\mathbf{q},\omega)u_z(\mathbf{q},\omega),$$

where in the limit $\omega/c_T q\ll 1$,

$$(M^{-1})_{zz} = -\frac{2Eq(1-\nu)}{(1+\nu)(3-4\nu)},$$

which differ from Eq. (12) only with respect to a factor $4(1-\nu)^2/(3-4\nu)$. For rubberlike materials ($\nu\approx 0.5$) this factor is of order unity. Hence, practically identical results are obtained independently of whether one assumes that the interfacial stress or displacement vector is perpendicular to the nominal contact surface. In reality, neither of these two assumptions hold strictly, but the result above indicate that the theory is not sensitive to this approximation.

Let us write

$$\mathbf{u}(\mathbf{x},t) = \int d^2q\,d\omega\,\mathbf{u}(\mathbf{q},\omega)e^{i(\mathbf{q}\cdot\mathbf{x}-\omega t)}.$$

If we assume that

$$\mathbf{u}(\mathbf{x},t) = \mathbf{u}(\mathbf{x}-\mathbf{v}t),$$

then

$$\mathbf{u}(\mathbf{q},\omega) = \frac{1}{(2\pi)^3}\int d^2x\,dt\,\mathbf{u}(\mathbf{x}-\mathbf{v}t)e^{-i(\mathbf{q}\cdot\mathbf{x}-\omega t)}$$

$$= \delta(\omega-\mathbf{q}\cdot\mathbf{v})\mathbf{u}(\mathbf{q}), \tag{13}$$

where

$$\mathbf{u}(\mathbf{q}) = \frac{1}{(2\pi)^2}\int d^2x\,\mathbf{u}(\mathbf{x})e^{-i\mathbf{q}\cdot\mathbf{x}}.$$

If σ_f denotes the frictional shear stress, then the energy dissipated during the time period t_0 equals

$$\Delta E = \sigma_f A_0 v t_0, \tag{14}$$

where A_0 is the surface area. But this energy can also be written as

$$\Delta E = \int d^2x\,dt\,\dot{\mathbf{u}}\cdot\boldsymbol{\sigma}$$

$$= (2\pi)^3\int d^2q\,d\omega\,(-i\omega)\mathbf{u}(\mathbf{q},\omega)\cdot\boldsymbol{\sigma}(-\mathbf{q},-\omega), \tag{15}$$

where $\omega=\mathbf{v}\cdot\mathbf{q}$. Substituting Eq. (9) in Eq. (15) and using Eq. (13) and that

$$[\delta(\omega-\mathbf{q}\cdot\mathbf{v})]^2 = (t_0/2\pi)\delta(\omega-\mathbf{q}\cdot\mathbf{v}),$$

gives

$$\Delta E = (2\pi)^2 t_0\int d^2q(-i\omega)[M_{zz}(-\mathbf{q},-\omega)]^{-1}$$

$$\times u_z(\mathbf{q})u_z(-\mathbf{q}).$$

Comparing this expression with Eq. (15) gives the frictional shear stress,

$$\sigma_f = \frac{(2\pi)^2}{vA_0} \int d^2q(-i\omega)[M_{zz}(-\mathbf{q}, -\omega)]^{-1}$$

$$\times \langle u_z(\mathbf{q})u_z(-\mathbf{q})\rangle, \tag{16}$$

where $\langle \cdots \rangle$ stands for ensemble averaging, i.e., averaging over different realization of the rough surface profile.

As an application, if

$$u_z(\mathbf{x}) = h_0 \cos(q_0 x) \cos(q_0 y),$$

we get

$$u_z(\mathbf{q}) = \frac{h_0}{4}[\delta(q_x - q_0) + \delta(q_x + q_0)]$$

$$\times [\delta(q_y - q_0) + \delta(q_y + q_0)].$$

Substituting this result in Eq. (16) and using that $\omega = vq_x$ (assuming sliding along the x-axis) and that

$$[\delta(\mathbf{q} - \mathbf{q}_0)]^2 = [A_0/(2\pi)^2]\delta(\mathbf{q} - \mathbf{q}_0),$$

gives

$$\sigma_f = -i\tfrac{1}{8}h_0^2 q_0([M_{zz}(q_0, q_0, -q_0 v)]^{-1}$$

$$- [M_{zz}(q_0, q_0, q_0 v)]^{-1})$$

$$= \tfrac{1}{4}h_0^2 q_0 \, \mathrm{Im}[M_{zz}(q_0, q_0, q_0 v)]^{-1}.$$

Using Eq. (12) this gives

$$\sigma_f \approx \frac{1}{8}(h_0 q_0)^2 \, \mathrm{Im}\frac{E(q_0 v)}{1 - \nu^2}. \tag{17}$$

Note that, in accordance with the discussion in Sec. I, σ_f depends only on $q_0 h_0$, so that the surface roughness profiles in Fig. 1 give equally important contributions to the sliding friction.

Let us now consider sliding on a randomly rough surface described by the function $z = h(\mathbf{x})$ [where $\mathbf{x} = (x, y)$]. Assume first that the rubber is able to deform and completely follow the substrate surface profile so that $u_z \approx h(\mathbf{x})$. Using Eq. (16) gives

$$\sigma_f = -i\frac{(2\pi)^2}{A_0} \int d^2q \, q_x \langle h(\mathbf{q})h(-\mathbf{q})\rangle$$

$$\times [M_{zz}(-\mathbf{q}, -q_x v)]^{-1}, \tag{18}$$

where we assumed that $\langle h \rangle = 0$. Now, note that

$$\langle h(\mathbf{q})h(-\mathbf{q})\rangle = \frac{A_0}{(2\pi)^4} \int d^2x \, \langle h(\mathbf{x})h(\mathbf{0})\rangle e^{-i\mathbf{q}\cdot\mathbf{x}}$$

$$\equiv \frac{A_0}{(2\pi)^2} C(q), \tag{19}$$

since $\langle h(\mathbf{x})h(\mathbf{x}')\rangle$ depends only on the difference $\mathbf{x} - \mathbf{x}'$. The *spectral density* $C(q)$ is defined by

$$C(q) = \frac{1}{(2\pi)^2} \int d^2x \, \langle h(\mathbf{x})h(\mathbf{0})\rangle e^{-i\mathbf{q}\cdot\mathbf{x}}. \tag{20}$$

We expect $C(q)$ to have the general form shown in Fig. 8.
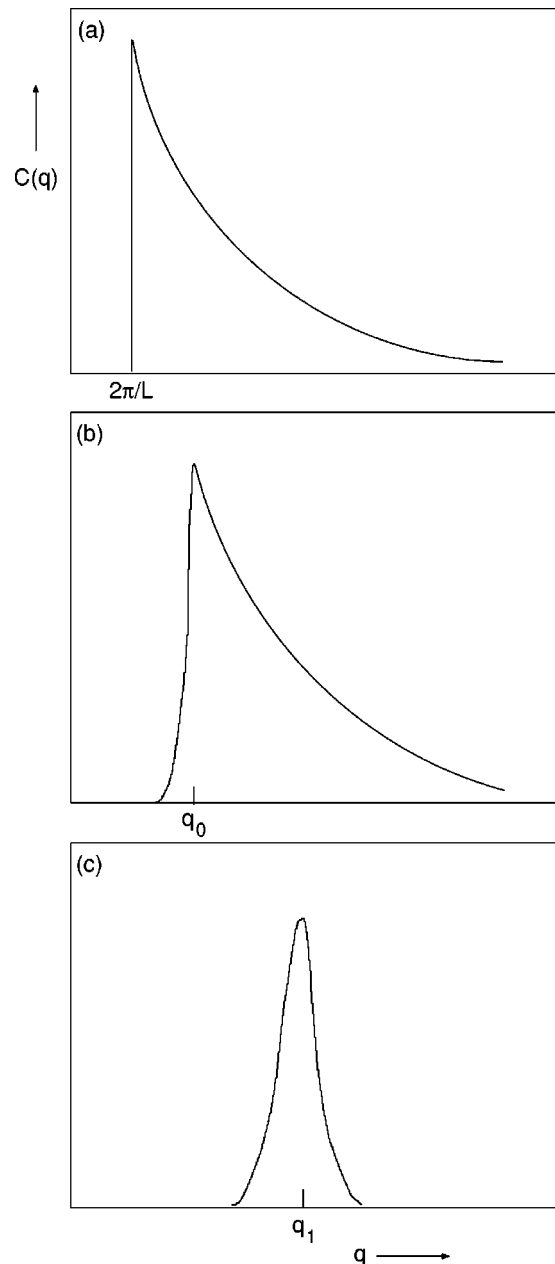

FIG. 8. The height correlation function $C(q)$ for three different (idealized) surface roughness profiles. (a) Self-affine fractal surface with the upper cut-off $\sim 2\pi/L$, determined by the lateral size L of the surface. (b) Self-affine fractal surface with the upper cutoff q_0 independent of L. (c) $C(q)$ for a rough surface, characterized by a narrow distribution of wavelength $\lambda_1 = 2\pi/q_1$ components.

The result in Fig. 8(a) corresponds to a self-affine fractal surface, where the low q-cut off (long distance) is determined by the lateral size L of the contact area, $q_0 \sim 2\pi/L$, while in Fig. 8(b), we assume that the self-affine fractal scaling only occurs for $q > q_0$, where q_0 is independent of the size of the rubber–substrate contact area (see Sec. II).

Substituting Eq. (19) in Eq. (18) and using Eq. (12) gives

$$\sigma_f = \frac{1}{2} \int d^2q \, q^2 \cos\phi \, C(q) \, \mathrm{Im}\frac{E(qv\cos\phi)}{1 - \nu^2}, \tag{21}$$

J. Chem. Phys., Vol. 115, No. 8, 22 August 2001

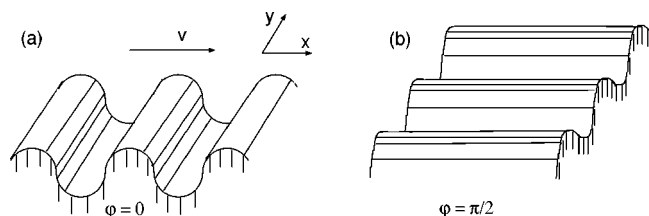Theory of rubber friction and contact mechanics   3847

FIG. 9. A cosine roughness profile with the wave vector (a) along, and (b) perpendicular to the sliding direction. Only in case (a) will the surface roughness generate time-dependent (fluctuating) deformations of the rubber block.

where we have used polar coordinates so that $q_x = q \cos \phi$, and where $E = E(\omega) = E(qv \cos \phi)$ and $\nu = \nu(\omega) = \nu(qv \cos \phi)$.

The friction coefficient μ can be obtained by dividing the frictional shear stress (21) with the pressure σ_0,

$$\mu = \frac{1}{2} \int d^2q \, q^2 \cos \phi \, C(q) P(q) \, \mathrm{Im} \frac{E(qv \cos \phi)}{(1-\nu^2)\sigma_0}. \quad (22)$$

In Eq. (22) we have introduced an additional factor $P(q)$, defined as the fraction of the original macrocontact area where contact remains when we study the contact area on the length scale $\lambda = 2\pi/q$. In principle, ν depends on frequency but the factor $1/(1-\nu^2)$ varies from $4/3 \approx 1.33$ for $\nu = 0.5$ (rubbery region) to ≈ 1.19 for $\nu = 0.4$ (glassy region) and we can neglect the weak dependence on frequency.

Since $C(q)$ and $P(q)$ only depend on the magnitude of **q**, from Eq. (22),

$$\mu = \frac{1}{2} \int dq \, q^3 \, C(q) P(q)$$
$$\times \int_0^{2\pi} d\phi \cos \phi \, \mathrm{Im} \frac{E(qv \cos \phi)}{(1-\nu^2)\sigma_0}. \quad (23)$$

Note that the factor $\cos \phi$ in the integrand vanishes when $\phi = \pi/2$, while it is maximal when $\phi = 0$. This has a simple but important physical origin: Consider two cosine-surface corrugations, where the "wave vector" points (a) along the x-axis (the sliding direction), and (b) along the y-axis, see Fig. 9. The former case corresponds to $\phi = 0$, and in this case the rubber block will experience pulsating deformations during sliding along the x-axis. The second case correspond to $\phi = \pi/2$, where the elastic deformations of the rubber *do not change* during sliding along the x-axis, and this type of surface roughness will therefore not contribute to the friction.

The present theory of rubber friction differs from the theory of Klüppel and Heinrich[9] in that it is fully 3D, and it takes into account [via the function $P(\zeta)$] how the rubber, on each length scale L/ζ, is able to follow the hard substrate profile, in contrast to Ref. 9, where this effect was only taken into account in some average way. Thus, the numerical results presented below are rather different from the prediction of the theory in Ref. 9.

V. CONTACT THEORY FOR RANDOMLY ROUGH SURFACES

We must now derive the function $P(q)$. 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 order the diameter of the nominal contact area. We define $q_L = 2\pi/L$ and write $q = q_L \zeta$. Thus, $P(q) = P(q_L \zeta)$, which we denote by $P(\zeta)$ for simplicity. If $\langle \sigma \rangle_\zeta$ denotes the *average* pressure in the (apparent) contact area on the length scale L/ζ,

$$\sigma_0 A_0 = \langle \sigma \rangle_\zeta P(\zeta) A_0, \quad (24)$$

so that

$$P(\zeta) = \sigma_0 / \langle \sigma \rangle_\zeta. \quad (25)$$

Thus, in order to determine $P(\zeta)$ we must first determine $\langle \sigma \rangle_\zeta$. If $P(\sigma, \zeta)$ denotes the stress probability distribution in the contact area on the length scale L/ζ, then

$$\langle \sigma \rangle_\zeta = \int_0^\infty d\sigma \, \sigma P(\sigma, \zeta) \Big/ \int_0^\infty d\sigma \, P(\sigma, \zeta). \quad (26)$$

Using Eqs. (25) and (26) gives

$$P(\zeta) = \sigma_0 \int_0^\infty d\sigma \, P(\sigma, \zeta) \Big/ \int_0^\infty d\sigma \, \sigma P(\sigma, \zeta). \quad (27)$$

The derivation of $P(\zeta)$ and $P(\sigma, \zeta)$ are given in Appendix B. Here I give the result for $P(\zeta)$,

$$P(\zeta) = \frac{2}{\pi} \int_0^\infty dx \, \frac{\sin x}{x} \exp\left[-x^2 \int_1^\zeta d\zeta' \, g(\zeta') \right], \quad (28)$$

where

$$g(\zeta) = \frac{1}{8} q_L q^3 C(q) \int d\phi \left| \frac{E(qv \cos \phi)}{(1-\nu^2)\sigma_0} \right|^2. \quad (29)$$

Now, assume that the macroscopic pressure σ_0 depends on the lateral position **x** in the nominal contact region, as would be the case if, e.g., a rubber ball is squeezed against a nominally flat substrate [where $\sigma_0(\mathbf{x})$ is given by the Hertz expression]. If we assume that the cut off distance λ_0 is much shorter than the diameter of the contact area [so that the variation of $\sigma_0(\mathbf{x})$ over the distance λ_0 is negligible], then, if we replace the constant σ_0 with the function $\sigma_0(\mathbf{x})$, the contact theory developed in Appendix B is still valid. We note, however, that as long as adhesion is unimportant (which is the case if the surfaces are rough enough[35]), and $\sigma_0(\mathbf{x})$ is small compared to the (low-frequency) elastic modulus E, the rubber friction coefficient is (nearly) independent of the actual pressure distribution in the nominal contact area (see below and Appendix C).

Let us reintroduce $q = q_L \zeta$, and summarize the basic results obtained above. The steady state kinetic friction coefficient for a flat rubber surface sliding on a nominally flat substrate is in the most general case is given by

$$\mu = \frac{1}{2} \int_{q_L}^{q_1} dq \, q^3 \, C(q) P(q)$$

$$\times \int_0^{2\pi} d\phi \cos \phi \, \mathrm{Im} \frac{E(qv \cos \phi)}{(1-\nu^2)\sigma_0}, \tag{30}$$

$$P(q) = \frac{2}{\pi} \int_0^\infty dx \frac{\sin x}{x} \exp[-x^2 G(q)], \tag{31}$$

where, using Eq. (29),

$$G(q) = \frac{1}{8} \int_{q_L}^q dq \, q^3 C(q) \int_0^{2\pi} d\phi \left| \frac{E(qv \cos \phi)}{(1-\nu^2)\sigma_0} \right|^2. \tag{32}$$

We consider now the limit $\sigma_0 \ll E(0)$, which is satisfied in most applications. In this case, for most q-values of interest, $G(q) \gg 1$, so that only $x \ll 1$ will contribute to the integral in Eq. (31), and we can approximate $\sin x \approx x$ and

$$P(q) \approx \frac{2}{\pi} \int_0^\infty dx \exp[-x^2 G(q)] = [\pi G(q)]^{-1/2}. \tag{33}$$

Thus, within this approximation, using Eqs. (32) and (33) we get $P(q) \propto \sigma_0$ so that μ is *independent* of the nominal stress σ_0. Similarly, note that if we scale $E(\omega) \to \alpha E(\omega)$, then from Eqs. (32) and (33), $P(q) \propto 1/\alpha$, so that μ depends only on the frequency variation of the complex elastic modulus, but not on its magnitude. We note that even if the macroscopic contact pressure $\sigma_0(\mathbf{x})$ depends on $\mathbf{x}$, and the integral (33) is still valid,

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

Thus, if $\sigma_0(\mathbf{x}) \ll E(0)$ for all $\mathbf{x}$, the friction force will be independent of $\sigma_0(\mathbf{x})$. For tires the condition $\sigma_0(\mathbf{x}) \ll E(0)$ is usually satisfied for all $\mathbf{x}$. Consequently, on a dry road track one expects the same friction for wide and narrow tires, assuming the same rubber temperature and that the rubber-road adhesional interaction is unimportant.

In order to take into account that $P(q) \to 1$ when $G(q) \to 0$, we use the interpolation formula,

$$P(q) \approx (1 + [\pi G(q)]^{3/2})^{-1/3}. \tag{34}$$

Numerical evaluation of Eq. (31) shows that Eq. (34) is an accurate representation of $P(q)$ for all q (or, equivalently, all G).

If we assume that the substrate surface is self affine fractal on all length scale between an upper and lower cutoff, $\lambda_0 \equiv 2\pi/q_0$ and $\lambda_1 \equiv 2\pi/q_1$, we have [see Fig. 8(b)] $C(q) = 0$ for $q < q_0$, while for $q > q_0$,

$$C(q) \approx k(q/q_0)^{-2(H+1)}, \tag{35}$$

where $H = 3 - D_f$ (the fractal dimension $2 < D_f < 3$). If we define $\langle h^2 \rangle = h_0^2/2$, then Eq. (20) gives $k = (h_0/q_0)^2 H/2\pi$. Using Eqs. (30) and (35) with $q = q_0 \zeta$ gives

$$\mu \approx \frac{1}{4\pi} (q_0 h_0)^2 H \int_1^{q_1/q_0} d\zeta \, \zeta^{-2H+1} P(q_0 \zeta)$$

$$\times \int d\phi \cos \phi \, \mathrm{Im} \frac{E(\zeta q_0 v \cos \phi)}{(1-\nu^2)\sigma_0} \tag{36}$$

and from Eqs. (32) and (35),

$$G(q) = \frac{1}{16\pi} (q_0 h_0)^2 H \int_1^{q/q_0} d\zeta \, \zeta^{-2H+1}$$

$$\times \int d\phi \left| \frac{E(q_0 \zeta v \cos \phi)}{(1-\nu^2)\sigma_0} \right|^2. \tag{37}$$

Note that since, to a good approximation, $P(q) \sim [G(q)]^{-1/2}$, it follows that $P \sim 1/q_0 h_0$, and thus $\mu \sim q_0 h_0$.

If we assume that $E(\omega)$ approaches a well defined limit as $\omega \to \infty$, then Eq. (37) gives for large ζ, $G(q_0 \zeta) \sim \zeta^{2-2H}$. Since $0 < H < 1$ it follows that $G(q_0 \zeta) \to \infty$ as $\zeta \to \infty$. Thus, for large ζ the integral (28) will be dominated by the small x-region and we can expand $\sin x \approx x$. Substituting this result in Eq. (46), using $G \sim \zeta^{2-2H}$ and defining a new integration variable $y = x\zeta^{1-H}$, gives $P(\zeta) \sim \zeta^{-1+H}$ as $\zeta \to \infty$. Thus for $0 < H < 1$ the contact area *goes to zero* as $\zeta \to \infty$. This will, of course, not occur in real systems, where there always exist an upper cutoff $\zeta_{max} = q_1/q_0$ in the integral over ζ. For example, the shortest possible distances are of atomic length, and this will give an upper cutoff. In practice, the cutoff is likely to occur at a much larger length scale because of contamination particles, or trapped fluid (or trapped pockets of compressed air), which will inhibit the elastic media from penetrating and fill out the small-sized roughness cavities (see Sec. III and Fig. 5). In addition, if the rubber has a thin modified surface layer (skin), this may also act as a cut off. Furthermore, when the area of real contact decreases the local pressure in the contact areas will finally reach the yield stress of the materials and beyond that point the area of real contact stays constant. However, even without an upper cutoff the friction coefficient given by Eq. (36) will (for a fixed sliding velocity v) remain finite as $\zeta_{max} = q_1/q_0 \to \infty$. This would not necessarily be the case if $P(\zeta) = 1$ for all ζ since the integrand in Eq. (36) (with $P = 1$) behaves as ζ^{-2H} for large ζ, and the integral $\sim \zeta_{max}^{1-2H}$ which diverge if $H < 0.5$. However, when the correct (asymptotic) dependence $P(\zeta) \sim \zeta^{-1+H}$ is taken into account the integral converges as $\sim \zeta_{max}^{-H}$, so that the very large ζ-contribution to the friction force will always give a small contribution. Note that $H \to 0$ corresponds to very rough surfaces (fractal dimension $D_f \approx 3$), and in this case the integral clearly converges relatively slowly.

It is possible to carry the analysis presented above further, by deriving an approximate analytical expression for $\mu(v)$. This result will be presented elsewhere.

VI. NUMERICAL RESULTS

As an example, assume that E is given by the model shown in Fig. 10. This model is, in fact, not a very good description of real rubbers, since the transition with increasing frequency from the rubbery region to the glassy region is much too abrupt, leading to a much too narrow (and too high) $\mu(v)$ peak. Nevertheless, the model gives a qualitatively correct $E(\omega)$. Later we will use experimental data for $E(\omega)$ for two different rubbers, illustrating how the results based on the present model (Fig. 10) are quantitatively modified.
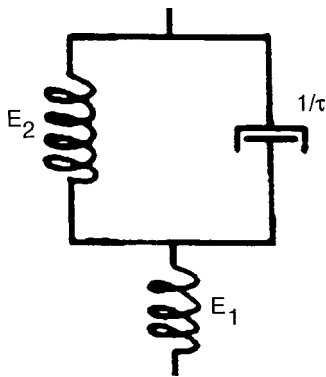
FIG. 10. Rheological model. In all calculations below $\tau = 0.001$ s, $E_1 = 10^9$ Pa, and $E_1/E_2 = a = 1000$.

FIG. 12. Variation of $P(\zeta)$ with the magnification ζ, for a few different sliding velocities for the system studied in Fig. 11.

The model in Fig. 10 corresponds to the elastic modulus,

$$E(\omega) = \frac{E_1(1 - i\omega\tau)}{1 + a - i\omega\tau}. \tag{38}$$

Note that $E(\infty) = E_1$ and $E(0) = E_1/(1+a)$ so that $E(\infty)/E(0) = 1 + a$. Since typically $E(\infty)/E(0) \approx 1000$ we take $a = 1000$ in all numerical calculations presented below. We assume $E_1 \gg \sigma_0$, in which case $\mu(v)$ is independent of E_1 and σ_0. Note that

$$E(\zeta q_0 v \cos\phi) = \frac{E_1}{1 + a} \frac{(1+a)^{-1} - i\zeta V}{1 - i\zeta V},$$

where $V = q_0 v \tau/(1+a)$. Thus, μ as a function of V depends only on H and $q_0 h_0$. However, instead of plotting μ as a function of V, we prefer to use real units corresponding to a typical case. We choose $\tau = 10^{-3}$ s, and $H = 0.85$, $q_0 = 2000$ m^{-1}, and $q_0 h_0 = 1$. Since $\mu \sim q_0 h_0$, the friction coefficient for other $q_0 h_0$ can be obtained from direct scaling.

Figure 11 shows the friction coefficient as a function of the sliding velocity, as obtained from Eqs. (30)–(32). I show results for the cutoff parameter $\zeta_{max} = 10$, 100, and 1000. We note that the inclusion of $P(\zeta)$ in Eq. (36) is very important,

since the decrease of the (apparent) contact area with increasing magnification, gives a strong reduction of the contribution to the friction force from the small-scale roughness. Figure 12 shows $P(\zeta)$ as a function of ζ, for several sliding velocities v. Figure 13 shows how $\mu(v)$ depends on the fractal dimension D_f when $\zeta_{max} = 100$, with the other parameters the same as in Fig. 11. In Fig. 14, I show $\mu_{max} = \max\{\mu(v)\}$ (from Fig. 13 and from additional calculations) as a function of H (or D_f). Note that when the fractal dimension $D_f = 3 - H$ increases towards 3, μ_{max} first increases, and then, when D_f increases beyond 2.9, μ_{max} decreases.

The rheological model used above (see Fig. 10) gives a too abrupt transition from the rubbery region to the glassy region with increasing frequency which leads to a too high μ_{max} and too narrow $\mu(v)$ peak. We therefore present some results based on experimentally measured shear modulus. Figure 15 shows the real Re $G(\omega)$ and imaginary Im $G(\omega)$ part of the shear modulus for synthetic polyisoprene (at $T = 303$ K), reticulated with dicumyl peroxide and without filler. The rubber glass transition temperature $T_g = 303$ K;
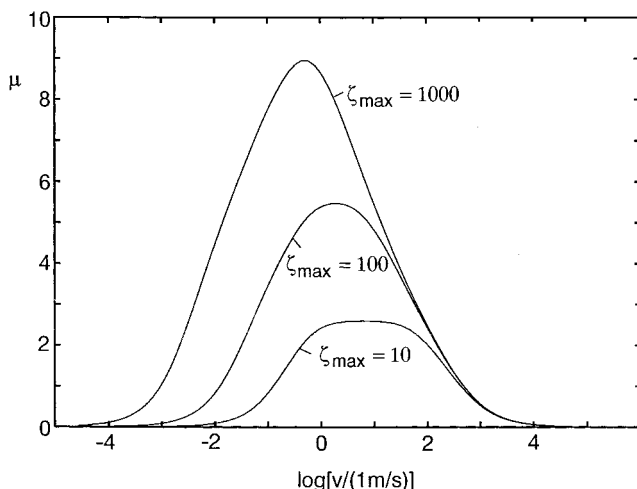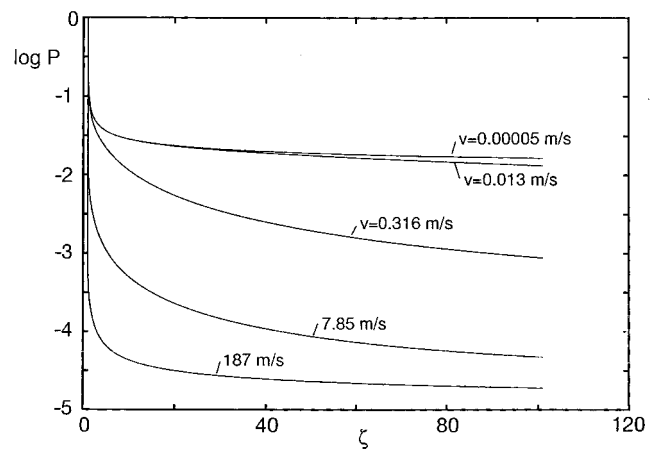

FIG. 11. The kinetic friction coefficient for rubber sliding on a substrate with a self-affine fractal surface profile characterized by the exponent $H = 0.85$. Calculations are presented for different cutoff ζ_{max} and with $q_0 h_0 = 1$ and $q_0 = 2000$ m^{-1}. Results for the rheological model shown in Fig. 10.
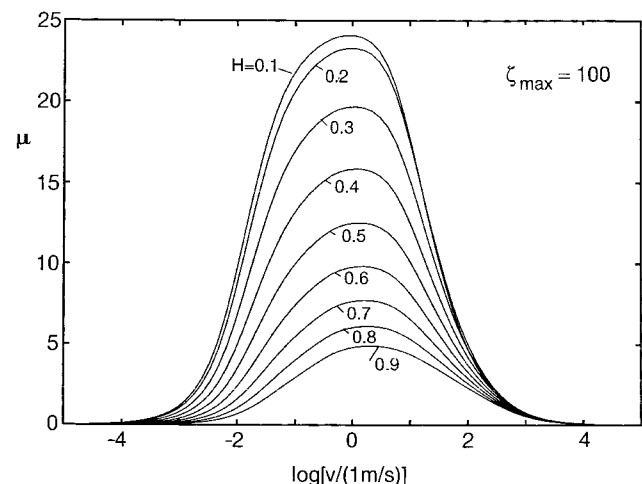

FIG. 13. The kinetic friction coefficient for rubber sliding on a substrate with a self-affine fractal surface profile with the cutoff $\zeta_{max} = 100$, and with $q_0 h_0 = 1$. Calculations are presented for different exponents H using the rheological model shown in Fig. 10 for the same parameters as in Fig. 11.
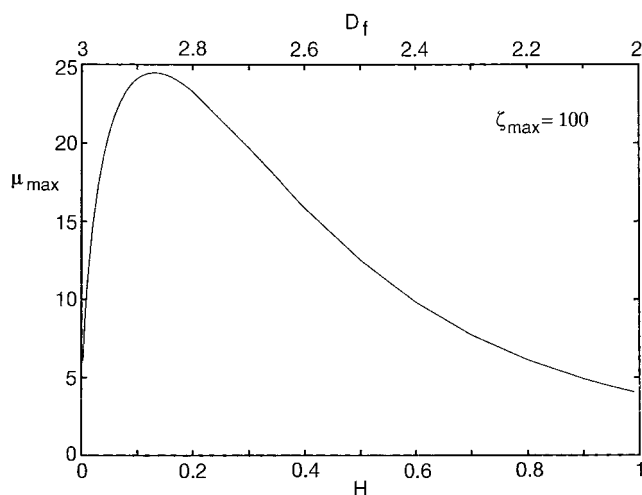
FIG. 14. The variation of the maximum friction coefficient (from Fig. 13) with the parameter H.

FIG. 16. The kinetic friction coefficient for polyisoprene rubber sliding on a substrate with a self-affine fractal surface profile characterized by the exponent $H = 0.85$. Calculations are presented for the cutoff $\zeta_{max} = 100$ and 1000, and with $q_0 h_0 = 1$ and $q_0 = 2000\ \mathrm{m}^{-1}$. For the $\zeta_{max} = 100$ case we show results for three different nominal pressures, $\sigma_0 = 0.1$, 1, and 10 MPa. For the shear modulus shown in Fig. 15.

this rubber is, of course, not used for tires, but was used in a detailed study of rubber friction.[24] In Fig. 16 we show the resulting friction coefficient $\mu(v)$ as calculated from the experimental $G(\omega)$-data given in Fig. 15, with $q_0 = 2000\ \mathrm{m}^{-1}$ and $q_0 h_0 = 1$ and with $H = 0.85$. For the $\zeta_{max} = 100$ case, I show results for three different nominal (or average) pressures: $\sigma_0 = 0.1$, 1, and 10 MPa. The nominal pressure at the tire-road interface is of order 0.3 MPa so that in that case one expect no dependence of μ on σ_0. However, when σ_0 becomes of order $E(0)$ the friction coefficient is no longer independent of σ_0, but decreases with increasing σ_0 (see 10 MPa curve in Fig. 16). Note that the effect of the applied pressure (Fig. 16) manifests itself mainly on the low-velocity side of the $\mu(v)$ peak.

Finally, let us present some results for a carbon and silica reinforced rubber compound, used by a major tire company (Pirelli) for "all-year-around" tires. One problem with applying the present theory to filled rubbers is the strongly nonlinear relation between the shear stress and the shear strain
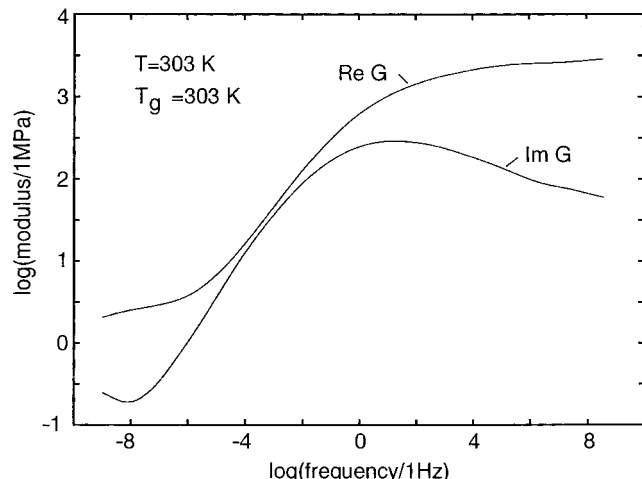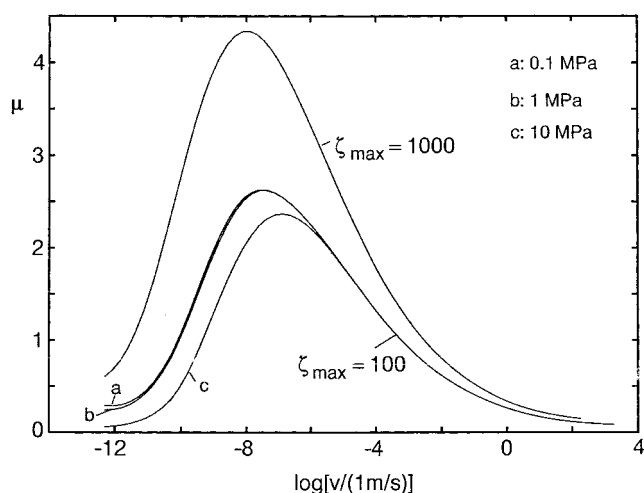
for filled rubbers. This nonlinearity is associated with the breakdown of the filler network, which occurs in the range of a few % strain amplitude. Since the stains involved in rubber friction when sliding on a road surface is of order unity (or $\sim 100\%$), when calculating the tire-road friction coefficient the effective elastic modulus $E(\omega)$ obtained from large amplitude stress-strain measurements should be used. Figure 17 shows the friction coefficient for two different temperatures, 40 °C and 70 °C, and for the cutoff $\zeta_{max} = 100$ and 1000. We have assumed a self-affine fractal substrate with $q_0 = 2000\ \mathrm{m}^{-1}$, $q_0 h_0 = 1$, and $H = 0.8$. The complex elastic modulus $E(\omega)$ used in the calculation was measured at 8% strain amplitude, which is so large that a complete break down of the filler network has occurred. Thus, further in-
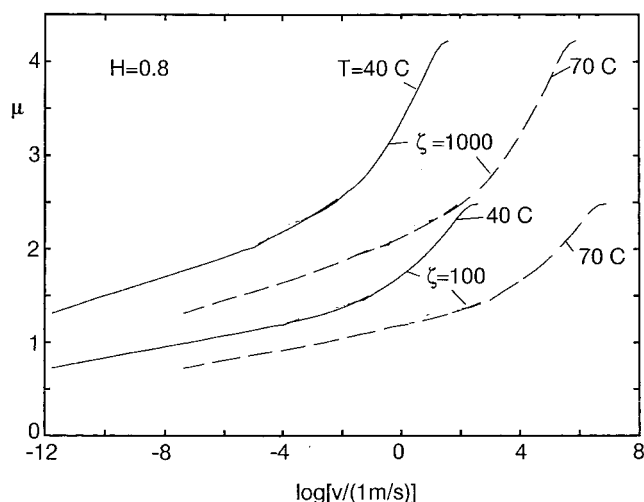

FIG. 15. The real and the imaginary part of the shear modulus of polyisoprene rubber (glass transition temperature $T_g = 303$ K) as a function of frequency, for $T = 303$ K.

FIG. 17. The kinetic friction coefficient for an "all-year" tire-rubber sliding on a substrate with a self-affine fractal surface profile characterized by the exponent $H = 0.80$. Calculations are presented for the cutoff $\zeta_{max} = 100$ and 1000, and with $q_0 h_0 = 1$ and $q_0 = 2000\ \mathrm{m}^{-1}$, and for two different temperatures, $T = 40$ °C and 70 °C.

crease in the strain amplitude gives only small modifications of the stress–strain relation, which has a negligible influence on the sliding friction. However, if $E(\omega)$ is measured at low strain amplitude, say 1%, the friction coefficient (not shown) is about half as large as when the calculation is based on the 8% stain amplitude data. In Fig. 17 we only show $\mu(v)$ up to the velocity where the friction coefficient is maximal; higher velocities are probably of no direct interest for tires since the region where $\mu(v)$ decreases with increasing v should be avoided, as interfacial stick-slip may occur when $\mu'(v) < 0$, which may result in an enhanced wear rate, and a loud noise.[25]

VII. DISCUSSION

The theory developed above can be used to estimate the kinetic friction coefficient for rubber sliding on a rough hard substrate. The input for the calculation, namely, the complex elastic modulus $E(\omega)$, and information about the substrate roughness [spectral function $C(q)$], can be obtained directly from relative simple experiments. In this section, I would like to make some additional comments related to rubber friction.

First, it would be interesting to perform experiment on systems with well defined surface roughness. Thus, it is now possible to prepare[26] surfaces covered by ordered arrays of nearly identical hemispherical ''bumps.'' Sliding of rubber on such well defined substrates would be good model systems for an accurate test of the theory developed above. It would also be interesting to perform rubber friction measurements on perfectly flat substrates to study the adhesional contribution to friction. I note that most earlier studies of the adhesional contribution have used polished glass surfaces which now are known to be very rough on the nanometer scale.[27]

By using a transparent substrate, it should be possible to study the asperity contact areas during squeezing and shearing of thin fluid films. In fact, Roberts[20] has already studied fluid films between rubber and a glass substrate. He has shown that the great flexibility of rubber surfaces leads to ready entrapment of liquid by elastic deformation. Similar effects have recently been observed for thin organic liquid films between mica surfaces,[28,29] and also observed in computer simulations.[30] Roberts also found that under certain circumstances thin (uniformly thick) fluid films remains trapped at the rubber–substrate interface. This happens when charge is introduced upon the contact surfaces leading to an electrical repulsion force between them. Such a force can support the normal load provided the contact pressure is not higher than ~ 0.1 MPa. In a typical case an equilibrium film of liquid some 200 Å thick becomes established between the surfaces. The generation of repulsive forces between rubber and glass surfaces means that the pair will make a microconforming contact with a uniform thin film of liquid between them (see Fig. 18). (Similar effects may be important in biological systems, e.g., polyelectrolyte layers are responsible for the low friction in mammalian joints.[31]) This can be used to measure the viscosity of water in thin films by squeeze action. The method has the advantage that dust or surface asperities can be tolerated without appreciably effecting the
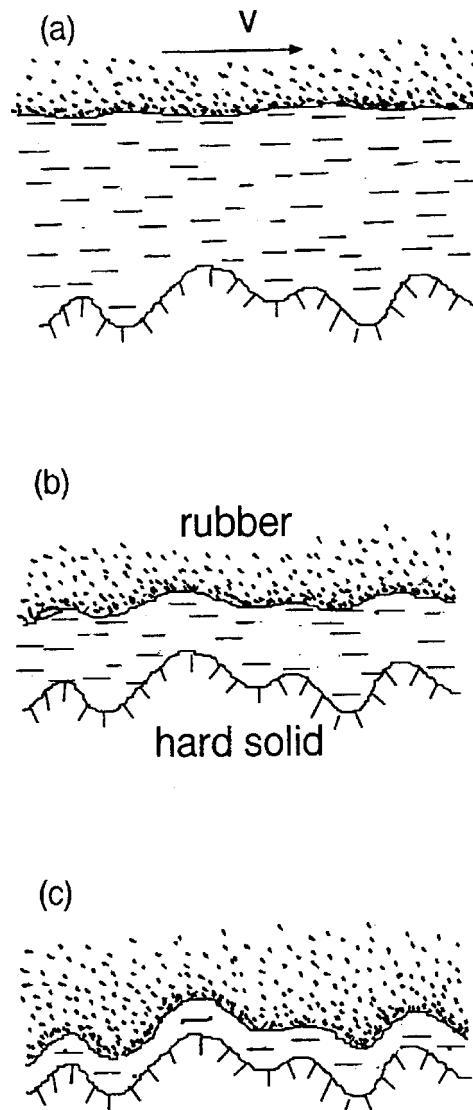

FIG. 18. A thin fluid layer between a rubber surface and a hard rough substrate. When charge is introduced upon the contact surfaces an electrical repulsion force occurs between them, which may support the normal pressure of at least ~ 0.1 MPa. In a typical case an equilibrium film of liquid some 200 Å thick becomes established between the surfaces. (a)–(c) show the system under increasing squeezing pressure. The generation of repulsive forces between rubber and glass surfaces means that the pair will make a microconforming contact with a uniform thin film of liquid between them as illustrated in (c).

result. In this way it has been shown that water (containing a trace of sodium dodecyl sulphate to generate electrical repulsive forces) in thin films 200–2000 Å thick possesses viscosity that is constant over this range of thickness and almost the same as the bulk viscosity. Roberts[32] also found that the rubber–glass contact in shear is stable (uniform film) under contact pressures of about ~ 0.1 MPa. However, friction measurements have shown that the electrolyte solution alone does not effectively lubricate the contact surfaces when films sheared are thinner than 100 Å. If, however, a surface active agent (e.g., sodium dodecyl sulphate), is included in the electrolyte solution, monolayer protection prevents surfaces from coming into intimate contact at points where the separating liquid film is locally punctured. The shear strength of the liquid film itself appears to remain constant and nearly the

same as the bulk viscosity of the liquid whether the film is 50 or 50 000 Å thick.

Another very interesting topic is rubber friction on ice. When most solids, e.g., glass, stone or metal, slide on ice the friction drop as the sliding velocity increases. As shown by Bowden,[33,5] this is caused by production of melt water as a result of frictional heating of the ice surface. For rubber on ice, however, Roberts[34] has observed the opposite effect: In the temperature range $-10\,°C<T<0\,°C$ the frictional stress *increases* with increasing sliding velocity. This can be explained by assuming that a thin liquidlike water layer occurs at the rubber–ice interface even in the absence of sliding. It is known that at the ice–vapor interface such a liquidlike layer does indeed exist as a result of *premelting*.[36] It remains to be understood why the premelted layer is absent when the ice is in contact with, e.g., glass or a metal oxide, but not when it is in contact with rubber.

I believe that the explanation of this remarkable phenomena is as follows: Water is likely to chemisorb on glass and on most metal oxide surfaces.[36] Thus, when such H_2O-monolayer (''icelike'') surfaces are brought in contact with an ice surface with a thin water layer (caused by premelting), then the situation will be similar to the case of bringing an ice block in contact with another ice block, where the water layer clearly will disappear (''refreeze'') in the contact area. On the other hand, because of the inert nature of rubber, it is unlikely that a layer of chemisorbed water molecules will occur on the rubber surface. Furthermore, the rubber surface is likely to be microscopically rough and the rubber molecules undergoes large thermal movement which may tend to break up any icelike structure at the interface. For this reason it is plausible that a liquidlike water layer may exist at the rubber–ice interface but not when ice is in contact with a hard, high energy solid surface, such as glass or metal oxides. It would be interesting to study this problem in greater detail, e.g., using molecular dynamics.

Roberts[34] also found that when a flat rubber surface slides on a flat ice surface below $-15\,°C$ at speeds less than 1 mm/s Schallamach waves are observed. At $-30\,°C$ and for speeds in excess of 100 mm/s the rubber wore rapidly and the friction fell as wear progressed. Rubber fragments were seen to form between the sliding surfaces and to become rolled together; they were then left deposited on the ice track.

Another extremely important topic is the (elastohydrodynamic) squeezing of thin liquid layers between a rubber surface and a hard rough (e.g., self-affine fractal) substrate, e.g., a water film squeezed between a tire and a road surface. This is a very complex problem related to cavity-connectivity, distribution of aperture,[37] and to the hydrodynamic pressure distribution in the liquid film at the interface, all of which depends on the local pressure and the squeeze time. This problem may be studied by computer simulations, but the very large number of length scales involved in most systems of practical interest, will make a comparison with experiment nontrivial.

The contact theory developed in this paper can be extended to the case where the rubber–substrate adhesion is important (which requires ''smooth'' surfaces). We note that

Fuller and Tabor[22] have already studied the influence of the substrate roughness on the rubber–substrate adhesion, and found (experimentally) that a relative small surface roughness (rms roughness $\sim 1\,\mu m$, or larger) is enough to remove (or kill) the effect of adhesion. However, when trying to understand this result theoretically, they employed a Greenwood-type of theory with roughness on a single length scale, while real surfaces always have roughness on many different length scales. A treatment of the adhesion contact problem within the present formalism gives a rather different picture of the role of surface roughness.[38]

Finally, for practical applications it is necessary to study the heating of the rubber during sliding. This problem is, in fact, closely related to the friction problem, since the heat source density $Q(\mathbf{x},t)$ is determined by the spatial distribution of the hysteretic energy losses in the surface region of the rubber block. The temperature field $T(\mathbf{x},t)$ must be determined by solving the heat diffusion equation with the heat source $Q(\mathbf{x},t)$, and with the appropriate boundary conditions which depend on the external conditions (e.g., road temperature).

VIII. SUMMARY AND CONCLUSION

There is at present a strong drive by tire companies to design new rubber compounds with lower rolling resistance, higher sliding friction, and reduced wear. At present these attempts are mainly based on a few empirical rules and on very costly trial-and-error procedures. I believe that a fundamental understanding of rubber friction and wear may help in the design of new rubber compounds for tires and other rubber applications, e.g., wiper blades.

In the present paper I have presented a general theory of the hysteretic contribution to rubber friction. The theory has been developed for rubber sliding on self-affine fractal surfaces, e.g., a tire on a road surface. I have shown that for stationary surfaces (or low sliding velocity), and for typical pressures in the contact area between a tire and a road, the rubber will only make (apparent) contact with about 5% of the road surface. On the longest length scale the contact involves the largest road surface asperities (which are associated with the upper cutoff length in the fractal distribution of the substrate surface roughness). However, in each such contact region the local pressure is large enough to squeeze the rubber into many of the smaller-sized ''cavities.'' I have developed a contact theory which describes how the (apparent) contact changes with the magnification.

ACKNOWLEDGMENTS

ments on the text. The author also thanks BMBF for a grant related to the German–Israeli Project Cooperation "Novel Tribological Strategies from the Nano- to Meso-Scales."

APPENDIX A: DERIVATION OF THE MATRIX M

In this Appendix, I present a short derivation of the matrix M [see Eq. (3)]. Assume that a viscoelastic solid occupate the half space $z>0$. On the surface $z=0$ of the solid acts the stress $\sigma_i=\sigma_{3i}$, where σ_{ij} is the stress tensor. We write

$$\sigma_i(\mathbf{x},t)=\int d^2q\,d\omega\,\sigma_i(\mathbf{q},\omega)e^{i(\mathbf{q}\cdot\mathbf{x}-\omega t)}$$

$$=\int d\omega\,\sigma_i(\mathbf{x},\omega)e^{-i\omega t},$$

where

$$\sigma_i(\mathbf{q},\omega)=\frac{1}{(2\pi)^3}\int d^2x\,dt\,\sigma_i(\mathbf{x},t)e^{-i(\mathbf{q}\cdot\mathbf{x}-\omega t)},$$

$$\sigma_i(\mathbf{x},\omega)=\frac{1}{2\pi}\int dt\,\sigma_i(\mathbf{x},t)e^{i\omega t}.$$

The elastic displacement field $u_i(\mathbf{x},z,t)$ satesfies the equation of motion,

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2}=\hat{\mu}\mathbf{\nabla}^2\mathbf{u}+(\hat{\mu}+\hat{\lambda})\mathbf{\nabla}\mathbf{\nabla}\cdot\mathbf{u},\tag{A1}$$

where $\hat{\mu}$ and $\hat{\lambda}$ are linear integral operators, e.g.,

$$\hat{\mu}\phi(t)=\int_{-\infty}^{\infty}dt'\,\mu(t-t')\phi(t'),$$

where $\mu(t)=0$ for $t<0$ (this is a result of causality), but this fact is not important for what follows. In what follows the time-variable will always be Fourier transformed so that Eq. (A1) takes the form,

$$-\rho\omega^2\mathbf{u}=\mu(\omega)\mathbf{\nabla}^2\mathbf{u}+[\mu(\omega)+\lambda(\omega)]\mathbf{\nabla}\mathbf{\nabla}\cdot\mathbf{u},\tag{A2}$$

where $\mathbf{u}=\mathbf{u}(\mathbf{x},z,\omega)$ and

$$\mu(\omega)=\int_{-\infty}^{\infty}dt\,\mu(t)e^{i\omega t},$$

and similarly for $\lambda(\omega)$. We *define* the complex elastic modulus $E(\omega)$ and Poisson ratio $\nu(\omega)$ via

$$\frac{\nu E}{(1+\nu)(1-2\nu)}=\lambda,\quad\frac{E}{1+\nu}=2\mu,$$

and the complex sound velocities $c_T(\omega)$ and $c_L(\omega)$ via

$$c_T^2=\frac{\mu}{\rho},\quad c_L^2=\frac{\lambda+2\mu}{\rho}.$$

In the equations above, all quantities depend on the frequency ω.

Let us now solve the boundary value problem specified above. It is convenient to introduce the vector $\mathbf{n}$ which points along the z-axis, normal to the surface of the semi-infinite solid, and write the displacement field in the solid on the form,

$$\mathbf{u}=\mathbf{p}A+\mathbf{K}B+\mathbf{p}\times\mathbf{K}C.\tag{A3}$$

Here A, B, and C are three scalar fields, and $\mathbf{p}=-i\mathbf{\nabla}$ and $\mathbf{K}=\mathbf{n}\times\mathbf{p}$ vector operators. Note that $\mathbf{K}$ is an *internal* operator on the plane $z=0$, i.e., it involves only differentation within the plane. Substituting Eq. (A3) in Eq. (A2) results in three scalar equations,

$$(\omega^2+c_L^2\mathbf{\nabla}^2)A=0,\tag{A4}$$

$$(\omega^2+c_T^2\mathbf{\nabla}^2)B=0,\tag{A5}$$

$$(\omega^2+c_T^2\mathbf{\nabla}^2)C=0.\tag{A6}$$

It is obvious from Eq. (A4) that A is associated with the longitudinal displacement field and B and C with the two transverse displacement fields. Note also that the $\mathbf{K}B$-field is parallel to the xy-plane. In the present case, the surface stress $\sigma_i(\mathbf{x},t)$ will generate viscoelastic displacement waves which propagates *into* the solid. Thus the relevant solutions to Eqs. (A4)–(A6) will have the form

$$A(\mathbf{x},z,t)=\int d^2q\,d\omega\,A(\mathbf{q},\omega)e^{i(\mathbf{q}\cdot\mathbf{x}+p_Lz-i\omega t)},\tag{A7}$$

$$B(\mathbf{x},z,t)=\int d^2q\,d\omega\,B(\mathbf{q},\omega)e^{i(\mathbf{q}\cdot\mathbf{x}+p_Tz-i\omega t)},\tag{A8}$$

$$C(\mathbf{x},z,t)=\int d^2q\,d\omega\,C(\mathbf{q},\omega)e^{i(\mathbf{q}\cdot\mathbf{x}+p_Tz-i\omega t)},\tag{A9}$$

where

$$p_T=\pm\left(\frac{\omega^2}{c_T^2}\pm i\epsilon-q^2\right)^{1/2},\quad p_L=\pm\left(\frac{\omega^2}{c_L^2}\pm i\epsilon-q^2\right)^{1/2},$$

where the $+$ and $-$ sign refers to $\omega>0$ and $\omega<0$, respectively, and where ϵ is an infinitesimal positive number, and where the square-root function has its branch cut along the negative real axis.

Now, using the equation

$$\sigma_{ij}=\hat{\mu}(u_{i,j}+u_{j,i})+\hat{\lambda}u_{k,k}\delta_{ij},$$

gives

$$\sigma_{i3}=n_j\sigma_{ij}=\hat{\mu}(\mathbf{n}\cdot\nabla u_i+\nabla_i\mathbf{n}\cdot\mathbf{u})+\hat{\lambda}n_i\nabla\cdot\mathbf{u}.$$

Thus, writing $-i\nabla=\mathbf{p}$, the boundary condition $\sigma_{3i}(\mathbf{x},0,\omega)=\sigma_i(\mathbf{x},\omega)$ takes the form

$$\mu(\omega)(\mathbf{n}\cdot\mathbf{p}\mathbf{u}+\mathbf{p}\mathbf{n}\cdot\mathbf{u})+\lambda(\omega)\mathbf{n}\mathbf{p}\cdot\mathbf{u}=-i\boldsymbol{\sigma}(\mathbf{x},\omega),\tag{A10}$$

where $\mathbf{u}=\mathbf{u}(\mathbf{x},0,\omega)$. Let us substitute Eq. (A3) in Eq. (A10). We get

$$\mathbf{p}(2p_zA+K^2C)+\mathbf{K}p_zB+\mathbf{p}\times\mathbf{K}p_zC+\mathbf{n}(\lambda/\mu)p^2A$$

$$=(-i/\mu)\boldsymbol{\sigma}.\tag{A11}$$

From Eq. (A11) we obtain three scalar equations by taking the scalar producs with the three operators $\mathbf{n}$, $\mathbf{K}$, and $\mathbf{p}_{\parallel}=(p_x,p_y,0)$. Note that all these operators are internal differential operators on the xy-plane, which is necessary since Eq. (A11) is only valid on the plane $z=0$. We get

$$(2p_z^2+(\lambda/\mu)p^2)A+2K^2p_zC=(-i/\mu)\mathbf{n}\cdot\boldsymbol{\sigma},\tag{A12}$$

$$K^2p_zB=(-i/\mu)\mathbf{K}\cdot\boldsymbol{\sigma},\tag{A13}$$

$$p_\parallel^2(2p_zA+K^2C)-p_z^2K^2C=(-i/\mu)\mathbf{p}_\parallel\cdot\boldsymbol{\sigma}. \quad\text{(A14)}$$

Using that $K^2=p_\parallel^2$ and $p^2A=(\omega^2/c_L^2)A$, and Fourier transforming the $\mathbf{x}$-dependence so that $\mathbf{p}_\parallel\rightarrow\mathbf{q}$, $p_zA=p_LA$, $p_zB=p_TB$, and $p_zC=p_TC$, Eqs. (A12)–(A14) gives

$$(2p_L^2+(\lambda/\mu)(\omega/c_L)^2)A+2q^2p_TC=(-i/\mu)\mathbf{n}\cdot\boldsymbol{\sigma}, \quad\text{(A15)}$$

$$q^2p_TB=(-i/\mu)\mathbf{K}\cdot\boldsymbol{\sigma}, \quad\text{(A16)}$$

$$q^2(2p_LA+q^2C)-p_T^2q^2C=(-i/\mu)\mathbf{q}\cdot\boldsymbol{\sigma}. \quad\text{(A17)}$$

In these equations $A=A(\mathbf{q},\omega)$ and similar for B, C, and σ, and $\mathbf{K}=\mathbf{n}\times\mathbf{q}\equiv q\mathbf{e}$. We must now solve Eqs. (A15)–(A17) for A, B, and C. From Eq. (A16) we get

$$B=-\frac{i}{\mu}\frac{1}{q^2p_T}\mathbf{K}\cdot\boldsymbol{\sigma}. \quad\text{(A18)}$$

From Eqs. (A15) and (A17) we get

$$A=-\frac{i}{\mu S}\left[2p_T\mathbf{q}+\left(\frac{\omega^2}{c_T^2}-2q^2\right)\mathbf{n}\right]\cdot\boldsymbol{\sigma}, \quad\text{(A19)}$$

$$C=-\frac{i}{\mu S}\left[2p_L\mathbf{n}-\left(\frac{\omega^2}{c_T^2}-2q^2\right)\frac{1}{q^2}\mathbf{q}\right]\cdot\boldsymbol{\sigma}, \quad\text{(A20)}$$

where

$$S=\left(\frac{\omega^2}{c_T^2}-2q^2\right)^2+4q^2p_Tp_L.$$

Now, since

$$\mathbf{p}\times\mathbf{K}=\mathbf{n}p_\parallel^2-p_z\mathbf{p}_\parallel,$$

we get from Eq. (A2),

$$\mathbf{u}(\mathbf{q},0,\omega)=\mathbf{K}B+\mathbf{q}(A-p_TC)+\mathbf{n}(p_LA+q^2C), \quad\text{(A21)}$$

where $A=A(\mathbf{q},\omega)$ and similarly for B and C. Substituting Eqs. (A18)–(A20) in this equation gives for $\mathbf{u}(\mathbf{q},0,\omega)\equiv\mathbf{u}(\mathbf{q},\omega)$,

$$\mathbf{u}(\mathbf{q},\omega)=M(\mathbf{q},\omega)\boldsymbol{\sigma}(\mathbf{q},\omega),$$

where the matrix,

$$M=-\frac{i}{\rho c_T^2}\left(\frac{1}{S(q,\omega)}\left[Q(k,\omega)(\hat{z}\mathbf{q}-\mathbf{q}\hat{z})\right.\right.$$
$$\left.\left.+\left(\frac{\omega}{c_T}\right)^2(p_L\hat{z}\hat{z}+p_T\hat{q}\hat{q})\right]+\frac{1}{p_T}\mathbf{ee}\right),$$

where $\hat{q}=\mathbf{q}/q$, $\mathbf{e}=\hat{z}\times\hat{q}$, and where

$$Q=2q^2-\omega^2/c_T^2+2p_Tp_L.$$

APPENDIX B: ELASTIC CONTACT THEORY FOR RANDOMLY ROUGH SURFACES

In this appendix we will derive the function $P(\zeta)$ introduced in Sec. V. Let us first derive an equation for the stress probability distribution in the contact area on the length scale L/ζ. We denote this function by $P(\sigma,\zeta)$. Let us first assume complete contact between the rubber and the substrate on all length scales. We have

$$P(\sigma,\zeta)=\langle\delta(\sigma-\sigma_1(\mathbf{x}))\rangle, \quad\text{(B1)}$$

where $\sigma_1(\mathbf{x})$ is the stress which occur in a contact area on the length scale L/ζ. Here $\langle\cdots\rangle$ stands for ensemble averaging, i.e., averaging over different realizations of the random process $h(\mathbf{x})$. If $\sigma_1+\Delta\sigma$ denotes the stress which occurs on the length scale $L/(\zeta+\Delta\zeta)$, then

$$P(\sigma,\zeta+\Delta\zeta)=\langle\delta(\sigma-\sigma_1-\Delta\sigma)\rangle$$
$$=\int d\sigma'\,\langle\delta(\sigma'-\Delta\sigma)\,\delta(\sigma-\sigma_1-\sigma')\rangle$$
$$=\int d\sigma'\,\langle\delta(\sigma'-\Delta\sigma)\rangle P(\sigma-\sigma',\zeta), \quad\text{(B2)}$$

where we have used that the averaging over different regions in ζ are independent processes. We can write

$$\langle\delta(\sigma'-\Delta\sigma)\rangle=\frac{1}{2\pi}\int dw\,\langle e^{iw(\sigma'-\Delta\sigma)}\rangle. \quad\text{(B3)}$$

Since $\Delta\sigma$ is small we can expand to second order in $\Delta\sigma$ to get

$$\langle\delta(\sigma'-\Delta\sigma)\rangle=\frac{1}{2\pi}\int dw\,e^{iw\sigma'}(1-w^2\langle\Delta\sigma^2\rangle/2). \quad\text{(B4)}$$

Note that $\langle\Delta\sigma^2\rangle\propto\Delta\zeta$. Substituting Eq. (B4) in Eq. (B2) and expanding the LHS to linear order in $\Delta\zeta$ gives

$$P(\sigma,\zeta)+\frac{\partial P(\sigma,\zeta)}{\partial\zeta}\Delta\zeta=\int d\sigma'\,P(\sigma-\sigma',\zeta)\left[\delta(\sigma')\right.$$
$$\left.+\frac{1}{2}\frac{\partial^2}{\partial\sigma'^2}\delta(\sigma')\langle\Delta\sigma^2\rangle\right].$$

Thus

$$\frac{\partial P}{\partial\zeta}=f(\zeta)\frac{\partial^2P}{\partial\sigma^2}, \quad\text{(B5)}$$

where

$$f(\zeta)=\frac{1}{2}\frac{\langle\Delta\sigma^2\rangle}{\Delta\zeta}.$$

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 (B5) is a diffusion type of equation, where time is replaced by the magnification ζ, and the spatial coordinate with the stress σ (and where the "diffusion constant" depends on ζ). Hence, when we study $P(\sigma,\zeta)$ on shorter and shorter length scales (corresponding to increasing ζ), the $P(\sigma,\zeta)$ function will become broader and broader in σ-space. We can take into account that detachment actually will occur when the local stress reaches $\sigma=0$ (we assume no adhesion) via the boundary condition,

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

If we multiply Eq. (B5) with σ and integrate over σ we get after some simplifications,

J. Chem. Phys., Vol. 115, No. 8, 22 August 2001

Theory of rubber friction and contact mechanics   3855

$$\frac{\partial}{\partial \zeta} \int_0^\infty d\sigma \, \sigma P(\sigma, \zeta) = 0$$

or

$$\int_0^\infty d\sigma \, \sigma P(\sigma, \zeta) = \sigma_0. \tag{B6}$$

Next, integrating Eq. (B5) over σ gives

$$\frac{\partial}{\partial \zeta} \int_0^\infty d\sigma P(\sigma, \zeta) = -f(\zeta) \frac{\partial P}{\partial \sigma}(0, \zeta)$$

or

$$\int_0^\infty d\sigma P(\sigma, \zeta) = 1 - \int_1^\zeta d\zeta' \, f(\zeta') \frac{\partial P}{\partial \sigma}(0, \zeta'). \tag{B7}$$

Using Eqs. (27), (B6), and (B7) gives

$$P(\zeta) = 1 - \int_1^\zeta d\zeta' \, f(\zeta') \frac{\partial P}{\partial \sigma}(0, \zeta'). \tag{B8}$$

Let us now calculate $\langle \Delta \sigma^2 \rangle$. Using Eqs. (12) and (19) give after some simplifications,

$$\langle \sigma_z^2 \rangle = \int d^2q \, [M_{zz}(\mathbf{q}, q_x v)]^{-1} [M_{zz}(-\mathbf{q}, -q_x v)]^{-1} C(q)$$

$$= \frac{1}{4} \int dq \, q^3 C(q) \int d\phi \left| \frac{E(qv \cos \phi)}{1 - \nu^2} \right|^2,$$

or, since $q = q_L \zeta$,

$$f(\zeta) = \frac{\langle \Delta \sigma^2 \rangle}{2 \Delta \zeta} = \frac{\langle \Delta \sigma^2 \rangle}{2 \Delta q} q_L$$

$$= \frac{1}{8} q_L q^3 C(q) \int d\phi \left| \frac{E(qv \cos \phi)}{1 - \nu^2} \right|^2.$$

Later we will need the function $g(\zeta) = f(\zeta)/\sigma_0^2$,

$$g(\zeta) = \frac{1}{8} q_L q^3 C(q) \int d\phi \left| \frac{E(qv \cos \phi)}{(1 - \nu^2)\sigma_0} \right|^2. \tag{B9}$$

Let us now solve Eq. (B5). Let us first consider a slightly more general problem, where $P(\zeta)$ again satisfies Eq. (B5),

$$\frac{\partial P}{\partial \zeta} = f(\zeta) \frac{\partial^2 P}{\partial \sigma^2},$$

but with modified boundary conditions,

$$P(0, \zeta) = P(\sigma_Y, \zeta) = 0, \tag{B10}$$

$$P(\sigma, 1) = P_0(\sigma) = \delta(\sigma - \sigma_0). \tag{B11}$$

In the equations above we consider $P(\sigma, \zeta)$ as defined only for $0 < \sigma < \sigma_Y$. Later, we will take $\sigma_Y \to \infty$. The solution to the equations above can be written as

$$P = \sum_{n=1}^\infty A_n(\zeta) \sin\left(\frac{n \pi \sigma}{\sigma_Y}\right). \tag{B12}$$

Substituting this in Eq. (B5) gives

$$\frac{dA_n}{d\zeta} = -f(\zeta) \left(\frac{n \pi}{\sigma_Y}\right)^2 A_n,$$

which is easy to integrate to get

$$A_n(\zeta) = A_n(1) \exp\left[-\left(\frac{n \pi}{\sigma_Y}\right)^2 \int_1^\zeta d\zeta' \, f(\zeta') \right].$$

Substituting this result in Eq. (B12) gives

$$P = \sum_{n=1}^\infty A_n(1) \exp\left[-\left(\frac{n \pi}{\sigma_Y}\right)^2 \int_1^\zeta d\zeta' \, f(\zeta') \right] \sin\left(\frac{n \pi \sigma}{\sigma_Y}\right). \tag{B13}$$

Using

$$\int_0^{\sigma_Y} d\sigma \, \sin\left(\frac{n \pi \sigma}{\sigma_Y}\right) \sin\left(\frac{m \pi \sigma}{\sigma_Y}\right) = \frac{\sigma_Y}{2} \delta_{nm},$$

we get

$$A_n(1) = \frac{2}{\sigma_Y} \int_0^{\sigma_Y} d\sigma \, P_0(\sigma) \sin\left(\frac{n \pi \sigma}{\sigma_Y}\right) = \frac{2}{\sigma_Y} \sin \alpha_n, \tag{B14}$$

where

$$\alpha_n = \frac{n \pi \sigma_0}{\sigma_Y} \equiv s \sigma_0, \tag{B15}$$

where we have defined $s = n \pi / \sigma_Y$. Substituting Eq. (B14) in Eq. (B13) gives

$$P = \frac{2}{\sigma_Y} \sum_{n=1}^\infty \sin \alpha_n$$

$$\times \exp\left[-\left(\frac{n \pi}{\sigma_Y}\right)^2 \int_1^\zeta d\zeta' \, f(\zeta') \right] \sin\left(\frac{n \pi \sigma}{\sigma_Y}\right). \tag{B16}$$

Let us now consider the limit $\sigma_Y \to \infty$. In this case we can replace

$$\sum_{n=1}^\infty \to \int_0^\infty dn = \frac{\sigma_Y}{\pi} \int_0^\infty ds,$$

so that Eq. (B16) reduces to

$$P = \frac{2}{\pi} \int_0^\infty ds \, \sin(s\sigma_0) \exp\left[-s^2 \int_1^\zeta d\zeta' \, f(\zeta') \right] \sin(s\sigma).$$

Now, let us consider

$$J = \int_1^\zeta d\zeta' \, f(\zeta') \frac{\partial P}{\partial \sigma}(0, \zeta')$$

$$= \frac{2}{\pi} \int_0^\infty ds \, s \, \sin(s\sigma_0) \int_1^\zeta d\zeta' \, f(\zeta')$$

$$\times \exp\left[-s^2 \int_1^{\zeta'} d\zeta'' \, f(\zeta'') \right]. \tag{B17}$$

But note that

$$\int_1^\zeta d\zeta' \, f(\zeta') \exp\left[-s^2 \int_1^{\zeta'} d\zeta'' \, f(\zeta'') \right]$$

$$= \frac{1}{s^2} \left(1 - \exp\left[-s^2 \int_1^\zeta d\zeta' \, f(\zeta') \right] \right). \tag{B18}$$

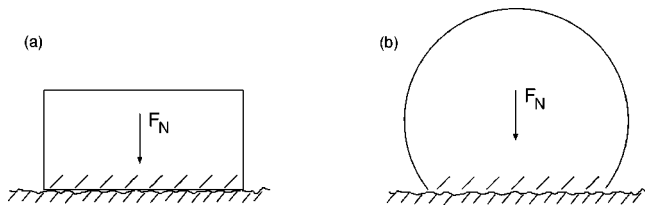Substituting Eq. (B18) in Eq. (B17) and using that

FIG. 19. (a) A rectangular block and (b) a spherical ball, squeezed against a rough substrate.

$$\frac{2}{\pi} \int_0^\infty ds \, \frac{\sin(s\sigma_0)}{s} = 1,$$

gives

$$J = 1 - \frac{2}{\pi} \int_0^\infty ds \, \frac{\sin(s\sigma_0)}{s} \exp\left[-s^2 \int_1^\zeta d\zeta' \, f(\zeta') \right].$$

Thus,

$$P(\zeta) = 1 - J = \frac{2}{\pi} \int_0^\infty ds \, \frac{\sin(s\sigma_0)}{s}$$
$$\times \exp\left[-s^2 \int_1^\zeta d\zeta' \, f(\zeta') \right].$$

Finally, let us define $g(\zeta) = f(\zeta)/\sigma_0^2$ and introduce $x = s\sigma_0$. Thus,

$$P(\zeta) = \frac{2}{\pi} \int_0^\infty dx \, \frac{\sin x}{x} \exp\left[-x^2 \int_1^\zeta d\zeta' \, g(\zeta') \right], \qquad (B19)$$

where $g(\zeta)$ is given by Eq. (B9).

APPENDIX C: ELASTOPLASTIC CONTACT THEORY FOR RANDOMLY ROUGH SURFACES

In Sec. V and Appendix B we have considered the area of real contact when an elastic body was squeezed against a hard rough surface. We assumed only elastic deformation. In this Appendix we consider the more general case when both elastic and plastic deformations occur. We consider two different cases, where the solids have (a) nominally flat surfaces as in Fig. 19(a) (e.g., a rectangular block on a flat substrate), where the macroscopic contact pressure σ_0 is constant, and (b) a curved surface as in Fig. 19(b) (e.g., an elastic spherical ball squeezed against a nominally flat substrate), where the macroscopic pressure, $\sigma_0(\mathbf{x})$, depends on $\mathbf{x}$.

1. Nominally flat surface

Consider a rectangular block on a nominally flat substrate. We define the average (or macroscopic) pressure $\sigma_0 = F_N/A_0$, where $A_0 = L^2$ is the (apparent) block-substrate contact area on the length scale L, the linear size of the block. Thus, the macroscopic pressure distribution $P_0 = \delta(\sigma - \sigma_0)$.

We assume that plastic yield occurs when the local pressure reaches $\sigma = \sigma_Y$ (the yield stress). In this case, for $0 < \sigma < \sigma_Y$, the stress probability distribution $P(\sigma, \zeta)$ satisfies Eq. (B5),

$$\frac{\partial P}{\partial \zeta} = f(\zeta) \frac{\partial^2 P}{\partial \sigma^2}, \qquad (C1)$$

with

$$P(0, \zeta) = P(\sigma_Y, \zeta) = 0,$$
$$P(\sigma, 1) = P_0(\sigma),$$

where $P_0(\sigma)$ is assumed to correspond to a constant macroscopic contact pressure $P_0(\sigma) = \delta(\sigma - \sigma_0)$. Note that $P(\sigma, \zeta)$ is defined only for $0 < \sigma < \sigma_Y$. In Appendix B we considered the limit $\sigma_Y \to \infty$, but we now keep σ_Y (the yield stress) finite.

In Appendix B we have shown that

$$P(\sigma, \zeta) = \frac{2}{\sigma_Y} \sum_{n=1}^\infty \sin \alpha_n$$
$$\times \exp\left[-\alpha_n^2 \int_1^\zeta d\zeta' \, g(\zeta') \right] \sin\left(\frac{n\pi\sigma}{\sigma_Y} \right), \qquad (C2)$$

where

$$\alpha_n = \frac{n\pi\sigma_0}{\sigma_Y}.$$

Now, let us introduce the functions $P_{\text{non}}(\zeta)$ and $P_{\text{pl}}(\zeta)$ which describe the fraction of the original (for $\zeta = 1$) macro-contact area where, under the magnification ζ, noncontact, and contact with plastic yield has occurred, respectively. Thus we have

$$P_{\text{el}}(\zeta) + P_{\text{non}}(\zeta) + P_{\text{pl}}(\zeta) = 1, \qquad (C3)$$

where $P_{\text{el}}(\zeta) = P(\zeta)$ was introduced in Appendix B and described the fraction of the macrocontact area where elastic contact occur on the length scale L/ζ. We have shown in Appendix B that

$$P_{\text{non}}(\zeta) = \int_1^\zeta d\zeta' \, f(\zeta') \frac{\partial P}{\partial \sigma}(0, \zeta'). \qquad (C4)$$

In a similar way one can show that

$$P_{\text{pl}}(\zeta) = -\int_1^\zeta d\zeta' \, f(\zeta') \frac{\partial P}{\partial \sigma}(\sigma_Y, \zeta'). \qquad (C5)$$

With these definitions it is easy to prove that the probability conservation law (C3) is satisfied. First note that the average stress $\langle \sigma \rangle_\zeta$ in the elastic contact area must be such that the total load is independent of the magnification ζ. Thus

$$\langle \sigma \rangle_\zeta A_0 P_{\text{el}}(\zeta) + \sigma_Y A_0 P_{\text{pl}}(\zeta) = \sigma_0 A_0$$

or

$$\langle \sigma \rangle_\zeta P_{\text{el}}(\zeta) = \sigma_0 - \sigma_Y P_{\text{pl}}(\zeta). \qquad (C6)$$

If we multiply Eq. (C1) with σ and integrate over σ we get after some simplifications,

$$\frac{\partial}{\partial \zeta} \int_0^{\sigma_Y} d\sigma \, \sigma P(\sigma, \zeta) = f(\zeta) \sigma_Y \frac{\partial P}{\partial \sigma}(\sigma_Y, \zeta)$$

or

$$\int_0^{\sigma_Y} d\sigma\,\sigma P(\sigma,\zeta) = \sigma_0 + \sigma_Y \int_1^{\zeta} d\zeta'\,f(\zeta')\frac{\partial P}{\partial\sigma}(\sigma_Y,\zeta')$$

$$= \sigma_0 - \sigma_Y P_{\rm pl}(\zeta).$$

Next, integrating Eq. (C1) over σ gives

$$\frac{\partial}{\partial\zeta}\int_0^{\sigma_Y} d\sigma P(\sigma,\zeta) = -f(\zeta)\left[\frac{\partial P}{\partial\sigma}(0,\zeta) - \frac{\partial P}{\partial\sigma}(\sigma_Y,\zeta)\right]$$

or

$$\int_0^{\sigma_Y} d\sigma P(\sigma,\zeta) = 1 - \int_1^{\zeta} d\zeta'\,f(\zeta')\left[\frac{\partial P}{\partial\sigma}(0,\zeta') - \frac{\partial P}{\partial\sigma}(\sigma_Y,\zeta')\right]$$

$$= 1 - P_{\rm non}(\zeta) - P_{\rm pl}(\zeta).$$

Thus

$$\langle\sigma\rangle_\zeta = \int_1^\zeta d\sigma\,\sigma P(\sigma,\zeta) \Big/ \int_1^\zeta d\sigma\,P(\sigma,\zeta)$$

$$= \frac{\sigma_0 - \sigma_Y P_{\rm pl}}{1 - P_{\rm non} - P_{\rm pl}}.$$

Substituting this in Eq. (C6) gives

$$P_{\rm el}(\zeta) = 1 - P_{\rm non}(\zeta) - P_{\rm pl}(\zeta).$$

Using Eqs. (C2), (C4), and (C5) we get

$$P_{\rm non} = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin\alpha_n}{n}\left(1 - \exp\left[-\alpha_n^2\int_1^\zeta d\zeta'\,g(\zeta')\right]\right), \tag{C7}$$

$$P_{\rm pl} = -\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n\frac{\sin\alpha_n}{n}$$

$$\times\left(1 - \exp\left[-\alpha_n^2\int_1^\zeta d\zeta'\,g(\zeta')\right]\right). \tag{C8}$$

Note that when $\zeta\to\infty$,

$$P_{\rm non} + P_{\rm pl} \to \frac{4}{\pi}\sum_{n=1,3,5,\ldots}\frac{\sin\alpha_n}{n} = 1,$$

independent of σ_0/σ_Y, while

$$P_{\rm pl} \to -\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n\frac{\sin\alpha_n}{n} = \frac{\sigma_0}{\sigma_Y}.$$

Thus, in the absence of a short-distance cutoff, at short length scale the local stress in the contact area equals σ_Y, i.e., each junction is in a state of incipient plastic flow. We note, however, that under such conditions thermally activated creep motion will be very important, and the area of real contact will increase slowly with the time of stationary contact.[39,40] This effect has a profound influence on friction dynamics (e.g., it is now believed to be the origin of earthquakes) but will not be discussed further here.

In the derivation above, $\zeta=1$ correspond to the macroscopic size L of the system. But, as mentioned earlier [see Fig. 8(b)], sometimes there is a cut off in $C(q)$ at some wave vector $q_0 > 2\pi/L$. In that case it is convenient to let $\zeta=1$ correspond to the length scale $\lambda_0 = 2\pi/q_0$. Before treating
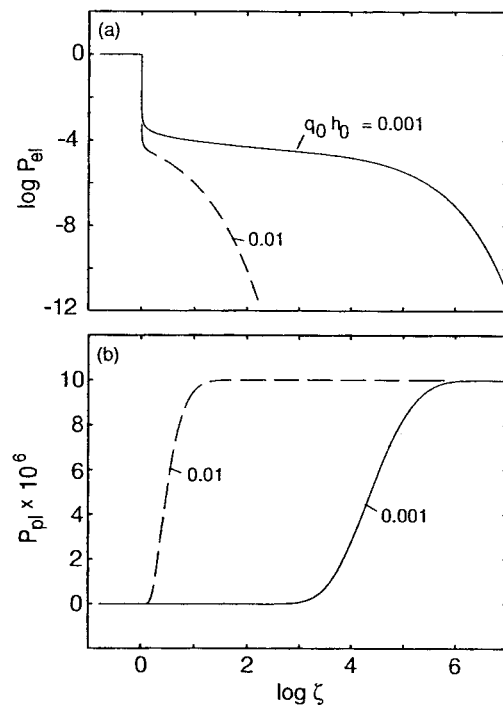

FIG. 20. The functions (a) $P_{\rm el}$ and (b) $P_{\rm pl}$ describes the fraction of the macroscopic contact area where elastic and plastic contact occur, when the system is studied at different magnifications ζ. For $H=0.8$, $q_0=10^4$ m^{-1}, and $q_0 h_0 = 0.001$ (solid lines) and $q_0 h_0 = 0.01$ (dashed lines). Results are shown for $E=10^{11}$ Pa, $\sigma_Y=10^9$ Pa, and $\sigma_0=10^4$ Pa. Note that in the present case $\zeta=1$ correspond to the length scale $\lambda_0 = 2\pi/q_0 \approx 1$ mm so that the log $\zeta<0$ correspond 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 [see Fig. 19(a)].

this case, let us summarize the results above in q-space ($q=q_L\zeta$). Equations (C7) and (C8) take the form,

$$P_{\rm non} = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{\sin\alpha_n}{n}(1 - \exp[-\alpha_n^2 G(q)]), \tag{C9}$$

$$P_{\rm pl} = -\frac{2}{\pi}\sum_{n=1}^{\infty}(-1)^n\frac{\sin\alpha_n}{n}(1 - \exp[-\alpha_n^2 G(q)]), \tag{C10}$$

where from Eq. (37),

$$G(q) = \frac{1}{8}\int_{q_0}^q dq\,q^3 C(q)\int_0^{2\pi} d\phi\left|\frac{E(qv\cos\phi)}{\sigma_0(1-\nu^2)}\right|^2. \tag{C11}$$

Now, assume zero sliding velocity and a self affine fractal surface, with $C(q)=0$ for $q<q_0$. Writing $q=q_0\zeta$ we get from Eq. (32),

$$\left(\frac{\pi\sigma_0}{\sigma_Y}\right)^2 G(q_0\zeta) = \frac{H}{1-H}\left(\frac{\pi E q_0 h_0}{4(1-\nu^2)\sigma_Y}\right)^2(\zeta^{2(1-H)} - 1).$$

For metals this formula should be (approximately) valid also for nonzero sliding velocities, since the elastic modul of metals depend much more weakly on the frequency than for rubber. Note that the functions $P_{\rm el}$, $P_{\rm pl}$, and $P_{\rm non}$ depend on H (or, equivalently, on the fractal dimension $D_f = 3-H$), on σ_0/σ_Y, and on the *plasticity index* $\psi=(E/\sigma_Y)q_0 h_0$. In Fig. 20 we show the dependence of $P_{\rm el}$ and $P_{\rm pl}$ of the magnifica-

tion ζ. We have used parameters which correspond (roughly) to a cubic steel block ($L = 10$ cm), on a steel substrate. We 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 $q_0 = 10^4$ m^{-1} (corresponding to the typical cutoff length $\lambda_0 = 2\pi/q_0$ of order ≈ 1 mm), then $q_0 h_0 = 0.001$ and 0.01 correspond to the rms roughness $h_0 = 0.1$ and 1 μm, respectively. In the calculations we have 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 $\lambda_0 \approx 1$ mm, and on the length scale $\lambda_0/10 \approx 0.1$ mm all junctions have yielded plastically. However, when $q_0 h_0 = 0.001$ plastic yield start when ζ is of order a few 1000, corresponding to distances of order $\lambda_0/\zeta \approx 0.1$ μm. On the length $\lambda \approx 20$ Å (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;[5] thus, for "real" steel mainly elastic deformation is likely to prevail when $q_0 h_0 = 0.001$.

One can easily estimate analytically the characteristic length scale λ_0/ζ, at which, say $\sim 50\%$ of the junctions have yielded plastically. According to Eq. (C8) this is the case when

$$\left(\frac{\pi\sigma_0}{\sigma_Y}\right)^2 G(q_0\zeta) = \frac{H}{1-H}\left(\frac{Eq_0h_0\pi}{4(1-\nu^2)\sigma_Y}\right)^2 (\zeta^{2(1-H)} - 1)$$

$$\approx 1$$

or

$$\zeta \approx \left[1 + \frac{1-H}{H}\left(\frac{\sigma_Y 4(1-\nu^2)}{Eq_0h_0\pi}\right)^2\right]^{1/[2(1-H)]}, \quad \text{(C12)}$$

which, in the present case, gives $\zeta \approx 1$ and 10^4 for $q_0 h_0 = 0.01$ and 0.001, respectively, in relative good agreement with Fig. 20.

Surfaces for engineering applications usually experience repeated sliding over the same area. Thus, for surfaces with large surface roughness, after repeated sliding the surface asperities will be smoothened out (by plastic deformation), and, finally, mainly elastic deformation will occur in the contact areas. The present theory can be used to estimate the length scale on which the initial plastic deformation occurs (see above).

Let us consider the "elastic limit" $\sigma_0/\sigma_Y \to 0$, where $P_{pl} = 0$. In this case we can treat $\alpha_n = x$ as a continuous variable so that

$$P_{el} = 1 - P_{non} = \frac{2}{\pi}\int_0^\infty dx \frac{\sin x}{x} \exp[-x^2 G(q)].$$

Since in most cases of interest (see Appendix B) $G(q) \gg 1$ we get

$$P_{el} \approx \frac{2}{\pi}\int_0^\infty dx \exp[-x^2 G(q)] = [\pi G(q)]^{-1/2}.$$

If we assume a frequency independent elastic modulus, then Eq. (37) gives

$$G(q_0\zeta) = \frac{(q_0 h_0)^2 E^2}{8\sigma_0^2(1-\nu^2)^2} \frac{H}{2(1-H)}(\zeta^{2(1-H)} - 1),$$

so that

$$P_{el} \approx \frac{4\sigma_0(1-\nu^2)}{q_0 h_0 E}\left(\frac{1-H}{\pi H}\right)^{1/2} \zeta^{H-1},$$

where we have assumed $\zeta \gg 1$. But $\sigma_0 = F_N/A_0$ so that the (apparent) area of contact on the length scale λ (where $\zeta = \lambda_0/\lambda$) becomes

$$A(\lambda) = A_0 P_{el}(\zeta) = \frac{4F_N(1-\nu^2)}{q_0 h_0 E}\left(\frac{1-H}{\pi H}\right)^{1/2}\left(\frac{\lambda}{\lambda_0}\right)^{1-H}. \quad \text{(C13)}$$

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

$$A(\lambda) = \frac{2F_N L(1-\nu^2)}{\pi E h_0}\left(\frac{1-H}{\pi H}\right)^{1/2}\left(\frac{\lambda}{L}\right)^{1-H}. \quad \text{(C14)}$$

Note that the contact area $A(\lambda)$ is *proportional* to the load, and that it decreases continuously towards zero as $\lambda \to 0$. If the upper cut off length λ_0 is independent of the size L of the system, then the area of contact $A(\lambda)$ is also independent of L [see Eq. (C13)]. However, if $\lambda_0 = L$ then the area of contact $A(\lambda)$ depends on the size L of the system, increasing as $\sim L^H$ with increasing L.

Finally, let us compare the prediction of the present theory with the contact theory of Greenwood. In the latter theory the surface is assumed to be covered by asperities with identical radius of curvature, R, and with a Gaussian hight distribution with the rms width h_1. We can (approximately) describe this case by assuming a surface with roughness only on a single lateral length scale $\lambda_1 = 2\pi/q_1$. We take

$$C(q) = [\langle h^2\rangle/2\pi q_1]\delta(q - q_1)$$

so that

$$\int d^2q\, C(q) = \langle h^2\rangle \equiv h_1^2.$$

Note that if $q_1 = 2\pi/\lambda_1$, then the asperity curvature $1/R \approx q_1^2 h_1$, and the asperity hight fluctuation $\Delta = h_1$ so that $(\Delta/R)^{1/2} \approx q_1 h_1$. The contact area in the Greenwood theory is of order $(\Delta/R)^{-1/2}(F_N/E) \approx (q_1 h_1)^{-1}(F_N/E)$. We will now show that essentially the same result follows from the present theory. Using Eq. (C11) we get $G(q) = 0$ if $q < q_1$ and

$$G(q) = \frac{(q_1 h_1)^2 E^2}{8\sigma_0^2}(1-\nu^2)^2 \equiv G_0$$

for $q > q_1$. The fraction of the area A_0 where contact occurs is given by

$$P = P_{el} + P_{pl} = \frac{\sigma_0}{\sigma_Y} + \frac{2}{\pi}\sum_{n=1}^\infty \frac{\sin\alpha_n}{n} e^{-\alpha_n^2 G}. \quad \text{(C15)}$$

In the elastic limit $\sigma_0/\sigma_Y \to 0$ we can treat $x = \alpha_n$ as a continuous variable, so that Eq. (C15) gives $P(q) = 1$ for $q < q_1$, while

$$P = \frac{2}{\pi} \int_0^\infty dx \, \frac{\sin x}{x} e^{-x^2 G_0}$$

$$\approx \frac{2}{\pi} \int_0^\infty dx \, e^{-x^2 G_0} = (\pi G_0)^{-1/2}$$

for $q > q_1$. Here we have assumed that $G_0 \gg 1$. Thus $A(\lambda) = A_0$ for $\lambda > \lambda_1$, while for $\lambda < \lambda_1$,

$$A(\lambda) = A_0 \frac{\sigma_0}{E} \frac{(1-\nu^2)}{q_1 h_1} \left(\frac{8}{\pi} \right)^{1/2}.$$

Since $\sigma_0 A_0 = F_N$ the normal load, we get

$$A(\lambda) = \frac{F_N (1-\nu^2)}{q_1 h_1 E} \left(\frac{8}{\pi} \right)^{1/2}$$

for $\lambda < \lambda_1$, which, except for a factor of order unity, is identical to the results of Greenwood.[17] Note, in particular, that the area of real contact is *proportional* to the load. However, in contrast to the case of self affine fractal surfaces, $A(\lambda)$ does not depend on λ (for $\lambda < \lambda_1$).

Next, let us consider the fraction of the contact area where plastic yield has occurred. As before, we expect about 50% of the contact area to have yielded plastically when

$$\left(\frac{\pi \sigma_0}{\sigma_Y} \right)^2 G_0 \approx 1$$

which gives

$$\psi = E q_1 h_1 / \sigma_Y \approx 1,$$

which, within a factor of order unity, is the same result as derived from the Greenwood theory.

2. Curved surface

Assume now that an elastic body with a nominally curved surface, e.g., a ball, is squeezed against a nominally flat substrate. In this case the macroscopic pressure in the contact region, $\sigma_0(\mathbf{x})$, will vary with the spatial location $\mathbf{x}$, from a maximum at the center to zero at the periphery. If the long-distance cut off λ_0 is much shorter than the diameter of the contact area, the theory presented above is still valid if the pressure $\sigma_0 \to \sigma_0(\mathbf{x})$. Thus, in this case the functions $P_{el}(\zeta, \mathbf{x})$ and $P_{pl}(\zeta, \mathbf{x})$ will depend on $\mathbf{x}$.

APPENDIX D: COMMENTS ON CONTACT THEORIES

Roux *et al.*[15] derived the relation (2) by scaling arguments as follows. Assume, for simplicity, that one of the surfaces is flat and elastic, while the other is rigid and has a self affine fractal surface profile. Let us rescale the spatial coordinates, $x \to \zeta x$, $y \to \zeta y$, and $z \to \zeta^H z$ and note that the surface is statistically invariant under this operation. The contact area $A \to \zeta^2 A$ as it lies in the (x, y) plane. The local perpendicular deformation u of the surface must rescale as
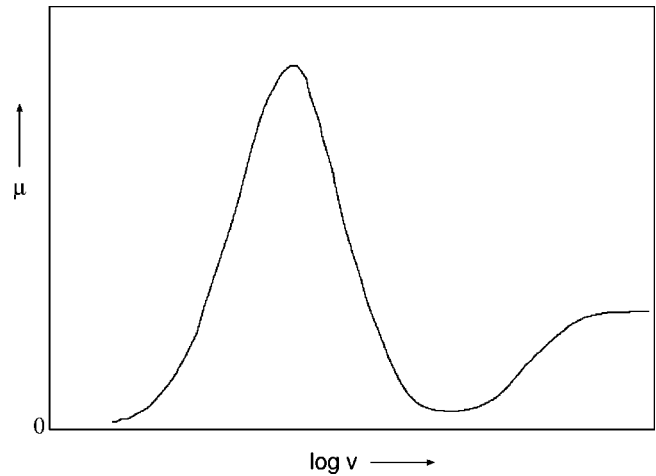

FIG. 21. The kinetic friction coefficient of rubber sliding on a rough, hard substrate (schematic).

$u \to \zeta^H u$, since u "points" in the z-direction. Since the normal surface stress σ is related to the perpendicular surface displacement u via

$$u = \int d^2 x' \, K(\mathbf{x} - \mathbf{x}') \sigma(\mathbf{x}'),$$

where, from dimensional arguments, $K \sim 1/|\mathbf{x} - \mathbf{x}'|$, it follows that $\sigma \to \zeta^{H-1} \sigma$ under rescaling. Thus, the total load $F_N = A \sigma \to \zeta^{1+H} F_N$. Assuming $F_N \sim A^n$ this gives $A^n \to \zeta^{1+H} A^n$. On the other hand we know that $A \to \zeta^2 A$ which gives $A^n \to \zeta^{2n} A^n$. Thus, $2n = 1 + H$ which gives Eq. (2). This result differs from the conclusion arrived at in this paper. However, the argument given above neglects the upper λ_0 and lower λ_1 cutoff lengths, which occur in all real systems. If there would be no low distance cutoff, the calculations presented in this paper predict that the area of real contact vanishes. This is easy to understand physically as follows: Let us first consider the system on the length scale λ_0, the upper cutoff length. On this length scale the system makes (apparent) contact with the substrate over an area $A(\lambda_0)$ (see Fig. 4). Let us now study an asperity contact area under increasing magnification. If we magnify by a factor of $\zeta = 10$ then we will observe smaller sized "asperities" and "cavities." In general, the local pressure will not be large enough to fill out all the cavities so the area of contact on the length scale $\lambda < \lambda_0$ will be smaller than the (apparent) area of contact on the length scale λ_0 (see inset in Fig. 4). This process will repeat itself as we increase the magnification further, and the area of (apparent) contact will continue to decrease with increasing magnification. If there is no short-distance cutoff the area of contact will decrease towards zero as the magnification $\zeta \to \infty$.

We note that the arguments presented by Roux do not exclude that the area of real contact may vanish, but in that case the analysis is itself irrelevant. Assume now that a short distance cutoff λ_1 exist. Let us consider the system under increasing magnification. We have to distinguish between two different cases. If the local stress which acts in the contact area reaches the yield stress σ_Y of the softer of the two solids before we have reached the cutoff length scale λ_1,

then plastic deformation will occur everywhere in the area A of *real contact*, and A will be simply determined by the load and the yield stress σ_Y via the standard relation $F_N = \sigma_Y A$. Thus, in this case the area of real contact is proportional to the load. On the other hand, if the short-distance cut-off is reached before the local stress has reached the yield stress σ_Y, then we expect mainly elastic deformation, but in this case too the area of real contact will be (nearly) proportional to the load (see Appendix C).

APPENDIX E: RUBBER FRICTION AT HIGH SLIDING VELOCITY

In this Appendix, I study the friction at high sliding velocities, where emission of sound waves from the sliding interface contributes in an important manner to the friction force. For high sliding velocities we cannot perform the expansion (11) but we have to use the full expression for M_{zz} given by Eq. (10),

$$M_{zz} = \frac{-i}{\rho c_T^2} \frac{p_L}{S(q,\omega)} \left(\frac{\omega}{c_T}\right)^2 = \frac{-i}{\rho c_T^2} \frac{\left(\frac{\omega}{c_T}\right)^2 \left[\left(\frac{\omega}{c_L}\right)^2 - q^2 + i0^+\right]^{1/2}}{\left[\left(\frac{\omega}{c_T}\right)^2 - 2q^2\right]^2 + 4q^2 \left[\left(\frac{\omega}{c_L}\right)^2 - q^2 + i0^+\right]^{1/2} \left[\left(\frac{\omega}{c_T}\right)^2 - q^2 + i0^+\right]^{1/2}}. \tag{E1}$$

Using this expression in the formulas derived in Secs. IV and V gives the friction coefficient for arbitrary sliding velocity. Here we consider only the limit when $\omega = vq_x \to \infty$, where

$$M_{zz} \sim \frac{-i}{\rho c_L \omega}. \tag{E2}$$

This equation is valid if $v \gg c_L$. For very high velocities we are the the glassy region where c_L can be treated as a constant (i.e., independent of ω). Of course, in reality we expect that high sliding velocity will result in high local temperature at the interface, and to large wear, but we neglect these effects in the present model study. Substituting Eq. (E2) in Eq. (18) gives after some simplifications,

$$\sigma_f = \pi v \rho c_L \int dq \, q^3 C(q) P(q). \tag{E3}$$

Thus, if $P(q)$ would be a constant, the friction coefficient would increase linearly with the sliding velocity v. However, using Eq. (E2) and the equations in Appendix B, it is easy to show that, if only elastic deformation occur, in the present case

$$P(q) \approx [\pi G(q)]^{-1/2} = \frac{\sigma_0}{\pi \rho c_L v} \left[\frac{1}{2} \int_{q_L}^{q} dq \, q^3 C(q)\right]^{-1/2}.$$

Substituting this in Eq. (E3), using $\mu = \sigma_f / \sigma_0$, and assuming that $C(q)$ is given by Eq. (35) gives (if $q_1 \gg q_0$)

$$\mu \approx q_0 h_0 \left(\frac{2H}{\pi(1-H)}\right)^{1/2} \left[\left(\frac{q_1}{q_0}\right)^{1-H} - 1\right],$$

so that $\mu(v)$ is velocity independent for very large v. Thus, neglecting temperature effects and wear processes, we expect $\mu(v)$ for rubberlike materials to have the qualitative form shown in Fig. 21.

[1] D. F. Moore, *The Friction and Lubrication of Elastomer* (Pergamon, Oxford, 1972); M. Barquins, Mater. Sci. Eng. **73**, 45 (1985); A. D. Roberts, Rubber Chem. Technol. **65**, 673 (1992).

[2] K. A. Grosch, Proc. R. Soc. London, Ser. A **274**, 21 (1963); see also A. D. Roberts, Rubber Chem. Technol. **65**, 3 (1992); S. P. Arnold, A. D. Roberts, and A. D. Taylor, J. Nat. Rubb. Res. **2**, 1 (1987); M. Barquins and A. D. Roberts, J. Phys. D **19**, 547 (1986).

[3] M. Barquins, Mater. Sci. Eng. **73**, 45 (1985).

[4] B. N. J. Persson, Surf. Sci. **401**, 445 (1998).

[5] B. N. J. Persson, *Sliding Friction: Physical Principles and Applications*, 2nd ed. (Springer, Heidelberg, 2000).

[6] B. N. J. Persson and E. Tosatti, J. Chem. Phys. **112**, 2021 (2000).

[7] A. Schallamach, Wear **6**, 375 (1963); Y. B. Chernyak and A. I. Leonov, *ibid.* **108**, 105 (1986).

[8] G. Heinrich, Kautsch. Gummi Kunstst. **45**, 173 (1992); Rubber Chem. Technol. **70**, 1 (1997).

[9] M. Klüppel and G. Heinrich, Rubber Chem. Technol. **73**, 578 (2000).

[10] J. Feder, *Fractals* (Plenum, New York, 1988).

[11] E. Bouchaud, J. Phys. C **9**, 4319 (1997).

[12] See, e.g., Ref. 9.

[13] J. F. Archard, Proc. R. Soc. London, Ser. A **243**, 190 (1957).

[14] J. A. Greenwood, in *Fundamentals of Friction, Macroscopic, and Microscopic Processes*, edited by I. L. Singer and H. M. Pollack (Kluwer, Dordrecht, 1992).

[15] S. Roux, J. Schmittbuhl, J. P. Vilotte, and A. Hansen, Europhys. Lett. **23**, 277 (1993).

[16] A. Majumdar and B. Bhushan, J. S. Afr. Chem. Inst. **113**, 1 (1991).

[17] J. A. Greenwood and J. B. P. Williamson, Proc. R. Soc. London, Ser. A **295**, 300 (1966).

[18] See, e.g., J. A. Schweitz and L. Ahman, in *Friction and Wear of Polymer Composites*, edited by K. Friederich (Elsevier, Amsterdam, 1986), pp. 289–327.

[19] K. A. Grosch, Rubber Chem. Technol. **69**, 495 (1997).

[20] A. D. Roberts, J. Phys. D **4**, 423 (1971).

[21] K. Mori, S. Kaneda, K. Kanae, H. Hirahara, Y. Oishi, and A. Iwabuchi, Rubber Chem. Technol. **67**, 798 (1994).

[22] K. N. G. Fuller and D. Tabor, Proc. R. Soc. London, Ser. A **345**, 327 (1975).

[23] See, e.g., Ref. 9 and A. W. Bush, R. D. Gibson, and G. P. Keogh, Mech. Res. Commun. **3**, 169 (1976).

[24] O. Ronsin and K. L. Coeyrehourcq, Proc. R. Soc. London, Ser. A **457**, 1277 (2001).

[25] B. N. J. Persson, Phys. Rev. B **63**, 104101 (2001); R. A. L. Rorrer, Rubber Chem. Technol. **73**, 486 (2000).

[26] R. Ishihara, A. Burtsev, and P. F. A. Alkemade, Jpn. J. Appl. Phys. **39**, 3872 (2000).

[27] U. D. Schwarz and R. Wiesendanger (private communication).

[28] F. Mugele and M. Salmeron, Phys. Rev. Lett. **84**, 5796 (2000).

[29] S. Zilberman, B. N. J. Persson, A. Nitzan, F. Mugele, and M. Salmeron, Phys. Rev. E **63**, 055103 (2001).

[30] B. N. J. Persson and P. Ballone, J. Chem. Phys. **112**, 9524 (2000).

[31] G. Wegner (unpublished); see also Ref. 5.

[32] A. D. Roberts, J. Phys. D **4**, 433 (1971).

[33] F. P. Bowden, Proc. R. Soc. London, Ser. A **217**, 462 (1953).

[34] A. D. Roberts and J. C. Richardson, Wear **67**, 55 (1981).

[35] D. Beaglehole and P. Wilson, J. Phys. Chem. **98**, 8096 (1994).

[36] P. A. Thiel and T. E. Madey, Surf. Sci. Rep. **7**, 211 (1987).

[37] F. Plouraboue, P. Kurowski, J.-P. Hulin, and S. Roux, Phys. Rev. E **51**, 1675 (1995).

[38] B. N. J. Persson (unpublished).

[39] B. N. J. Persson and R. Ryberg, Phys. Rev. B **32**, 3586 (1985); B. N. J. Persson and A. I. Volokitin, Surf. Sci. **457**, 345 (2000).

[40] See, e.g., Ref. 5, and B. N. J. Persson, Phys. Rev. B **61**, 5949 (2000).