Simple Microscopic Theory of Amontons's Laws for Static Friction

Martin H. Müser,[1] Ludgar Wenning,[1] and Mark O. Robbins[2]

[1]*Institut für Physik, WA 331, Johannes Gutenberg Universität, 55099 Mainz, Germany*
[2]*Department of Physics & Astronomy, Johns Hopkins University, Baltimore, Maryland 21218*

A microscopic theory for the ubiquitous phenomenon of static friction is presented. Interactions between two surfaces are modeled by an energy penalty that increases exponentially with the degree of surface overlap. The resulting static friction is proportional to load, in accordance with Amontons's laws. However, the friction coefficient between bare surfaces vanishes as the area of individual contacts grows, except in the rare case of commensurate surfaces. An area independent friction coefficient is obtained for any surface geometry when an adsorbed layer of mobile atoms is introduced between the surfaces. The predictions from our simple analytic model are confirmed by detailed molecular dynamics simulations.

PACS numbers: 81.40.Pq, 46.55.+d, 62.20.Qp

Static friction F_s is the lateral force that must be applied to initiate sliding of one object over another. Its presence implies that the objects have locked together into a local energy minimum that must be overcome by the external force. This threshold force is observed between all the objects around us, yet its molecular origins have remained baffling. How do surfaces manage to lock together, and why does the friction obey Amontons's laws that F_s increases linearly with the load L pushing the objects together, and is independent of the surface area [1]?

Early theories of friction were based on the purely geometric argument that friction is caused by interlocking of surface asperities [1,2]. The idea (Fig. 1) is that the top surface must be lifted up a typical slope $\tan\alpha$ determined by roughness on the bottom surface. If there is no microscopic friction between the surfaces, then the minimum force to initiate sliding is $F_s = L\tan\alpha$. This result satisfies Amontons's laws with a constant coefficient of friction $\mu_s \equiv F_s/L = \tan\alpha$. In 1737, Bélidor obtained a typical experimental value of $\mu_s \approx 0.35$ by modeling rough surfaces as spherical asperities arranged to form commensurate crystalline walls [2]. More recently, Israelachvili *et al.* have discussed a similar "cobblestone" model where the spherical asperities are atoms [3]. However, asperities on real surfaces do not match as well as envisioned in these models or sketched in Fig. 1. On average, for every asperity or atom going up a ramp, there is another going down. One concludes that the mean friction between rigid surfaces vanishes unless they happen to have the same periodicity and alignment. Detailed calculations show that elastic deformations are generally too small to alter this conclusion [4–7].

In this Letter, we study a simple model of the static friction in individual contacts as a function of their area A. At fixed A, the static friction between bare surfaces is proportional to load. However, as implied by previous work [4–7], the prefactor vanishes for incommensurate surfaces. We find that F_s/L also vanishes for disordered surfaces, scaling as $A^{-1/2}$ as the area increases. These results for bare surfaces are inconsistent with Amontons's

laws. Of course most surfaces around us are covered by a few angstroms of hydrocarbons, water, and other small airborne molecules [8]. We show that introducing such mobile molecules into the interface between incommensurate or disordered surfaces yields a value of $\mu_s = F_s/L$ that is independent of contact size and load. This naturally leads to Amontons's laws for any distribution of contacts. Our analytic predictions are then tested against previous [8] and new computer simulations.

In our model, two surfaces "pay" an energy penalty V_{ww} that increases exponentially as the local separation between them decreases and they begin to overlap. This form of interaction is frequently used to model the Pauli repulsion between atoms and contains the hard-sphere models of Bélidor and Israelachvili *et al.* as a limiting case. Within the contact area A the surfaces are parallel to the xy plane. The bottom wall is held fixed, has zero mean height, and local height $\delta z_b(\vec{x})$. The mean height z_t and lateral position $\vec{x}_t$ of the top wall are varied, and its local height is $z_t + \delta z_t(\vec{x} - \vec{x}_t)$. Then V_{ww} can be written as an integral over A

$$V_{ww}(\vec{x}_t, z_t) = \varepsilon \int d^2x \, e^{-[z_t + \delta z_t(\vec{x} - \vec{x}_t) - \delta z_b(\vec{x})]/\xi}, \quad (1)$$

where ξ and ε characterize the length and energy per unit area of the interaction. The dependence of V_{ww} on z_t can
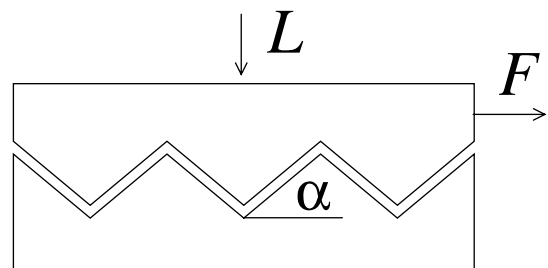

FIG. 1. Sketch of two surfaces with interlocking asperities. The top surface experiences a normal load L and a lateral force F, which attempts to pull the top surface up the slope $\tan\alpha$. The bottom wall is fixed.

be pulled out of the integral in Eq. (1). The integral then depends only on $\vec{x}_t$ and can be expressed as an effective shift in the mean wall separation by $\Delta z(\vec{x}_t)$:

$$V_{ww}(\vec{x}_t, z_t) = \varepsilon A e^{-\{z_t + \Delta z(\vec{x}_t)\}/\xi}. \qquad (2)$$

The lateral and normal forces on a static contact are equal to the derivatives of V_{ww} with respect to $\vec{x}_t$ and z_t, respectively. Because of the simple form of Eq. (2) almost all factors in the ratio of friction to normal load cancel, and the ratio is independent of load [9]. The static friction coefficient corresponds to the maximum of this ratio, taken along the direction of the applied force:

$$\mu_s = \max\left[\frac{\partial}{\partial x_t} \Delta z(\vec{x}_t)\right]. \qquad (3)$$

Physically, $\partial \Delta z(\vec{x}_t)/\partial x_t$ represents an effective ramp that the top surface must climb to escape the local potential energy minimum. It is analogous to $\tan\alpha$ in Fig. 1 but arises from many molecular-scale interactions. Note that the factorization that leads to Eq. (3) relies on the exponential form of the local energy penalty. However, our numerical tests with other interaction potentials and the simulations with Lennard-Jones potentials described below show that the result is more general.

The effect of the area of contact on the friction coefficient μ_s is most easily seen by making a cumulant expansion of Eq. (1). To leading order, Δz is proportional to the mean-squared difference between δz_t and δz_b. Only the contribution from the cross product contributes to the variation of Δz with respect to $\vec{x}_t$, yielding

$$\mu_s \approx \max_{x_t} \frac{1}{\xi} \sum_{\vec{k}} i k_x \delta\tilde{z}_b(\vec{k}) \delta\tilde{z}_t^*(\vec{k}) e^{i\vec{k}\cdot\vec{x}_t}, \qquad (4)$$

where $\delta\tilde{z}_{b,t}$ are the 2D Fourier transforms of $\delta z_{b,t}$. Explicit evaluation of Eq. (3) for various trial surfaces, and the simulations described below, show that higher order terms in δz do not alter any of the following predictions.

Equation (4) implies several relations between the geometry of the two surfaces and the friction between them: (i) If the two surfaces are crystalline and commensurate, then δz_t and δz_b share a common periodicity. The corresponding Bragg peaks in their Fourier transforms lead to an area independent μ_s. (ii) The magnitude of μ_s is largest for two identical surfaces, where all Bragg peaks contribute. For commensurate surfaces that are not identical, μ_s decreases exponentially with the length of the common period. For example, if there are p lattice constants of the bottom surface for every q lattice constants of the top surface, then the contribution to μ_s comes from the qth Bragg peak of the bottom surface and the pth Bragg peak of the top surface. The Fourier content of atomically rough surfaces [10] drops at least exponentially with $|\vec{k}|$, producing a corresponding decrease in μ_s. (iii) As a consequence of (ii), μ_s vanishes completely for infinite incommensurate contacts. Contributions from the circumference of finite contacts yield a rapidly vanishing con-

tribution to μ_s [11]. (iv) For two disordered but smooth surfaces $\delta\tilde{z}_t(\vec{k})$ and $\delta\tilde{z}_b(\vec{k})$ have rings of diffuse scattering that overlap. There will be many $\vec{k}$ that contribute to μ_s, but the phase of each contribution will be random. For interfaces with contact-area independent corrugation $\langle\delta z_{t,b}^2\rangle$, one can immediately conclude from the law of large numbers that $\mu_s \propto A^{-1/2}$. Self-similar and curved interfaces need a more careful treatment [11].

Two geometry independent predictions can also be drawn from Eqs. (3) and (4): (v) μ_s does not depend on the interaction strength ε. (vi) Allowing for elastic deformation of the surfaces typically reduces μ_s, because the roughness decreases as the surfaces become more compliant [12].

Predictions (i) to (iii) agree with previous analytic and simulation studies of the force between commensurate and incommensurate surfaces [4–6]. (v) is in agreement with recent computer simulations [8], where doubling the strength of Lennard-Jones interactions produced almost no change in μ_s. (iv) and (vi) are new results that are tested by simulations below. First we consider the implications of mobile atoms between the surfaces within our analytic model.

Figure 2 illustrates how a mobile layer of atoms between two surfaces can lead to a finite μ_s *independent* of the geometry of the two surfaces. The mobile atoms are able to move to positions where they simultaneously match the geometry of both top and bottom surfaces. This has the effect of augmenting the height of the bottom surface in a way that matches the undulations of the top surface and gives an area independent contribution to Eq. (3).

In order to incorporate mobile atoms into our theory quantitatively, we presume that wherever a mobile atom sits in the interface, the effective distance between the top and the bottom wall is reduced by the diameter d of the mobile atom. This effectively incorporates the mobile atoms into one of the walls. The position of a mobile atom in the interface is defined by a vector $\vec{x}_i$ so that the dimensionless areal density of the mobile atoms can be written as $\rho(\vec{x}) = (\pi d^2/4)\sum_i \delta(\vec{x} - \vec{x}_i)$. For the sake of simplicity, we assume that the mobile layer screens the
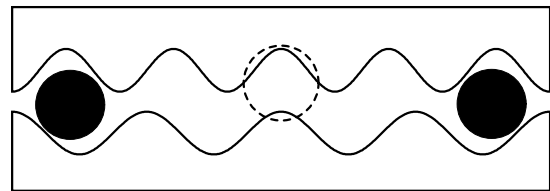

FIG. 2.   Sketch of two rigid, incommensurate surfaces that are separated by a submonolayer of mobile atoms. Circles indicate positions where the gap between the upper and the lower wall is a local maximum. The larger the gap, the smaller the energy penalty for occupation by a mobile atom. Full circles indicate $(++)$ positions where atoms fit into concave regions of both surfaces. The open circle indicates a less favorable $(+-)$ site where one wall is concave and the other is convex.

direct wall-wall interaction and that interactions between the mobile atoms can be ignored. The latter assumption is consistent with observations [8] that μ_s is insensitive to the density of mobile atoms. The validity of both assumptions was verified through the simulations described below.

With this model, the total energy of the system is given by the indirect wall-wall interaction mediated through the mobile atoms. This takes the form

$$V_{ww} = \varepsilon \int d^2 x \, \rho(\vec{x}) e^{-\{z_t + \delta z_t(\vec{x} - \vec{x}_t) - \delta z_b(\vec{x}) - d\}/\xi}$$

$$= \frac{\varepsilon \pi d^2}{4} e^{-(z_t - d)/\xi} \sum_i e^{-[\delta z_t(\vec{x}_i - \vec{x}_t) - \delta z_b(\vec{x}_i)]/\xi}. \quad (5)$$

The force on the top wall $\vec{F}_t$ can again be calculated by taking the gradient of V_{ww} with respect to the center of mass position of the top wall: $(\vec{x}_t, z_t)$. It is possible to decompose $\vec{F}_t$ into forces $\vec{F}_{t,i}$ that stem from individual atoms. For each such force, a linear relationship between the x component of the force $F_{t,ix}$ and the "local" load $F_{t,iz}$ can be established that is similar to Eq. (3):

$$F_{t,ix} = \left[\frac{\partial}{\partial x_t} [\delta z_t(\vec{x}_i - \vec{x}_t) - \delta z_b(\vec{x}_i)] \right] F_{t,iz}. \quad (6)$$

As long as the partial derivative has a well-defined average value, the total friction will rise linearly with the total load.

At this point, we turn back to a more qualitative discussion. As long as the temperature is small compared to the energy barrier for diffusion between the surfaces, atoms will sit near local energy minima. The simulations described below show that these local minima usually correspond to the $(++)$ configurations identified in Fig. 2, where the energy with respect to both surfaces is concave. Lateral displacements of the walls will ultimately limit the volume accessible to $(++)$ atoms and increase V_{ww}. The atoms have created interlocking asperities that resist sliding much like those in Fig. 1. One can show more rigorously from Eq. (5) that if all atoms sit at $(++)$ positions, the response to a lateral displacement is an opposing force that is linear in the normal load L and independent of the area of contact. Indeed, this "Amontons's law for elastic pinning" with an elastic restoring force linear in L has been observed experimentally [13]. Thus, we may conclude that mobile atoms in the interface lead to Amontons's laws with a nonvanishing friction coefficient. The pinning of two surfaces through a "between-sorbed" layer can also be interpreted in terms of a generalized Frenkel-Kontorova or Tomlinson model [7]. In these models, incommensurate layers exhibit static friction when one of the surfaces is so compliant that it can conform to the other surface. The mobile atoms act like an elastic sheet with nearly infinite compliance and thus allow locking to occur.

We have tested the above predictions for disordered walls using molecular dynamics simulations of the model described in detail in Refs. [8] and [14]. The two walls

contain discrete atoms tied to their equilibrium sites by springs of stiffness κ. Periodic boundary conditions are applied in the plane of the walls. Mobile molecules between the surfaces contain n spherical monomers bound into a chain. All monomers and wall atoms also interact with a Lennard-Jones potential. For the results presented below, $n = 1$ or 6, and the energy and length scales (ϵ and σ, respectively) of the Lennard-Jones potential are the same for monomers and wall atoms. Other parameters produced equivalent results. Disordered walls were made by quenching bulk fluid states and then removing all atoms above or below some height.

Figure 3 shows representative results for the scaling of friction with contact area. The top wall was coupled to a constant normal load and a slowly increasing lateral force. The static friction was determined from the threshold force needed to initiate lateral motion of the top wall. For all cases considered, the coefficient of friction between bare disordered surfaces vanished as $A^{-1/2}$ (solid line) in agreement with prediction (iv). In contrast, when molecules were inserted into the interface, the static friction rapidly approached a constant as A was increased. It is interesting to note that the bare and disordered surfaces have equal friction at an area corresponding to about 1000 atoms. This is slightly larger than contacts in atomic force microscopy but much smaller than the micron size contacts between typical surfaces. In agreement with prediction (vi), decreasing κ to the smallest reasonable value [6] ($\kappa = 400\epsilon\sigma^{-2}$) reduced the friction between bare walls considerably (crosses). Much smaller reductions (~20%) were observed when mobile atoms were present between elastic surfaces. In both cases the reductions reflect the ability for atoms to move vertically to minimize the steric
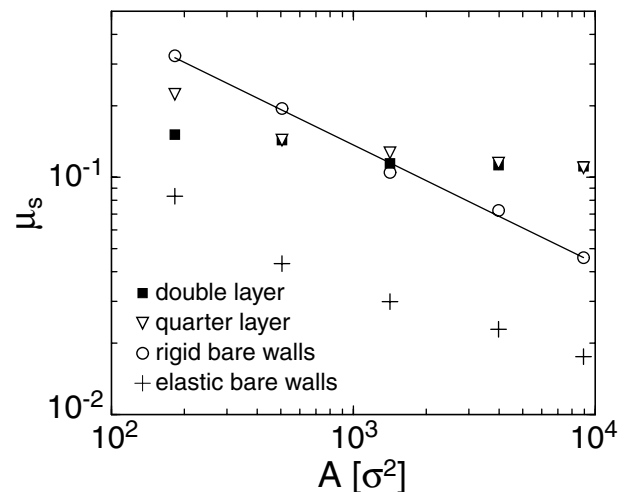

FIG. 3. Static friction coefficient as a function of contact area A. The friction coefficient for rigid or elastic ($\kappa = 400\epsilon\sigma^{-2}$) bare walls drops as $A^{-1/2}$ (solid line). Inserting enough mobile molecules ($n = 6$) to make either a quarter or two monolayers yields nearly the same area independent value of μ_s. Statistical error bars are comparable to the symbol size.
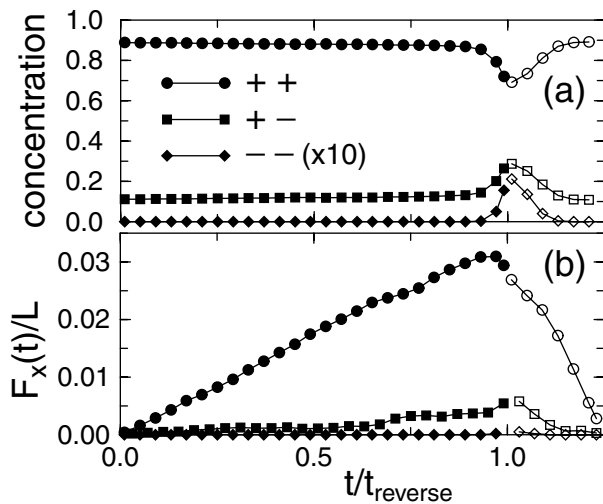
FIG. 4.   (a) Fraction of atoms in $(++)$, $(+-)$, and $(--)$ sites (Fig. 2) and (b) their contribution to the restoring friction force as a function of time. For these runs $t_{reverse}$ was about 300 in Lennard-Jones units.

overlap with the opposing surface. When atomic displacements were constrained to the horizontal plane, we observed no reduction in F_s.

The simulations also allowed us to test the assumption that monomers in $(++)$ configurations (Fig. 2) dominate the pinning. Figure 4 shows results for a quarter monolayer of spherical molecules $(n = 1)$ as the force was ramped up to a value slightly above F_s over a time $t_{reverse}$ and then returned 4 times as rapidly to zero. About 90% of mobile atoms sit at $(++)$ positions until the wall begins to slide $(t/t_{reverse} \approx 0.95)$, and they contribute an even larger fraction of F_s. Atoms that sit at convex positions relative to both walls $(--)$, were seen only during sliding. Even in this dynamic state, almost 70% of the atoms are in $(++)$ sites and they continue to provide most of the lateral force. A detailed analysis of these runs shows that the mechanisms of kinetic and static friction are closely related in this model. At any instant in time, most of the mobile atoms are in local energy minima and resist lateral motion. When these sites become unstable $(+-)$ or $(--)$ sites, the atoms pop rapidly to a new site. Energy is dissipated during these rapid pops and flows into the walls as heat.

In conclusion, we have presented a simple analytic model for the molecular origins of friction. Although the model assumes exponential repulsion between surfaces, we presented simulations with more realistic potentials that show that the predictions are more generally applicable. Specifically the model provides a microscopic foundation for Amontons's law $F_{lateral} = \mu L$ but shows that μ vanishes for bare incommensurate or disordered surfaces as the size of contacts increases. Introducing mobile atoms between the surfaces yields Amontons's laws: μ is independent of surface area and load for

any contact geometry. The simulations also reveal deep connections between static and kinetic friction. Of course, experimental systems contain additional features that have not been treated. These include long-range elastic and plastic deformations of the walls, generation of wear debris, and chemical reactions [1]. All of these may be important in particular systems.

[1] F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids* (Clarendon Press, Oxford, 1986).

[2] D. Dowson, *History of Tribology* (Longman Inc., New York, 1979), pp. 153–167.

[3] J. N. Israelachvili, S. Giasson, T. Kuhl, C. Drummond, A. Berman, G. Luengo, J.-M. Pan, M. Heuberger, W. Ducker, and M. Alcantar, in *Thinning Films and Tribological Interfaces,* Tribology Series Vol. 38 (Elsevier, New York, 2000), pp. 3–12; A. Berman, C. Drummond, and J. N. Israelachvili, Tribol. Lett. **4**, 95 (1998).

[4] M. Hirano and K. Shinjo, Phys. Rev. B **41**, 11 837 (1990); Wear **168**, 121 (1993).

[5] M. R. Sørensen, K. W. Jacobsen, and P. Stoltze, Phys. Rev. B **53**, 2101 (1996).

[6] M. H. Müser and M. O. Robbins, Phys. Rev. B **61**, 2335 (2000).

[7] M. O. Robbins and M. H. Müser, in *Modern Tribology Handbook,* edited by B. Bhushan (CRC Press, Boca Raton, FL, 2001), Vol. 1, pp. 717–765.

[8] G. He, M. H. Müser, and M. O. Robbins, Science **284**, 1650 (1999); G. He and M. O. Robbins, Tribol. Lett. (to be published).

[9] Previously, exponentially repulsive forces between surfaces had only been shown to result in Amontons's laws for Gaussian height probability distributions of contacting asperities: A. Volmer and T. Nattermann, Z. Phys. B **104**, 363 (1997).

[10] Summing Lennard-Jones interactions from each atom also leads to exponentially decaying Fourier content in the wall potential: W. Steele, Surf. Sci. **36**, 317 (1973). Higher order contributions to the cumulant expansion from smaller $\vec{k}$ are also exponentially small.

[11] L. Wenning and M. H. Müser, cond-mat/0010396.

[12] This trend is reversed in the limit where the solids become extremely soft, allowing them to deform into locally commensurate regions [4,6]. A detailed analysis shows that this limit corresponds to having stronger interactions between atoms on opposing surfaces than within each surface [6,7].

[13] P. Berthoud and T. Baumberger, Proc. R. Soc. London A **454**, 1615 (1998).

[14] P. A. Thompson, M. O. Robbins, and G. S. Grest, Isr. J. Chem. **35**, 93 (1995).