Understanding and illustrating the atomic origins of friction

James Ringlein
Lancaster Country Day School, 725 Hamilton Road, Lancaster, Pennsylvania 17603

Mark O. Robbins[a)]
Department of Physics and Astronomy, The Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218

(Received 10 September 2003; accepted 5 March 2004)

Recent advances in the understanding of the atomic origins of friction are described and illustrated with simple simulations. Examples of macroscopic and nanometer scale systems that violate Amontons' laws of friction are discussed. A more general friction relation is motivated and shown to fit data from simple atomic simulations that can be downloaded and modified. The simulations illustrate the fundamental relation between static friction and potential energy, and between kinetic friction and energy dissipation. Conceptual difficulties in understanding how almost all pairs of surfaces lock together in a potential energy minimum are described, and possible resolutions are discussed. We conclude with an explanation of why Amontons' laws work so well in many macroscopic systems. © *2004 American Association of Physics Teachers.*
[DOI: 10.1119/1.1715107]

I. INTRODUCTION

The discussions of sliding and static friction in a typical high school physics course and in popular introductory college physics texts might lead one to believe that friction is a rather straightforward phenomenon. However, some everyday experiences can leave students wondering if the laws they are taught are valid. For example, sticky objects like tape exhibit friction without an applied load and this friction increases with area. Most physics texts do not address such exceptions to the ''laws'' of friction. They also offer little insight into the developments in friction research that have occurred in recent years.

The study of friction, lubrication, and wear is called tribology, or the science of rubbing, and has been at the center of technology since prehistoric times.[1] In the last 15 years, new experimental methods that probe friction with atomic resolution in one or more directions have revealed many details about the molecular origins of friction.[2] Atomic force microscopes measure the friction between a substrate and a few to several hundred atoms on a sharp tip.[3] The lateral resolution can be less than an atomic spacing. The surface force apparatus measures the forces between atomically flat surfaces as their separation is varied with angstrom level resolution.[4] The friction and adhesion are studied as a function of the chemistry and thickness of the material between the surfaces. The quartz crystal microbalance detects friction forces between islands of atoms that are one or two layers thick and the substrate on which they slide.[5] Simulations using new generations of computers have played an important role in interpreting and explaining the findings from these new experimental methods. The study of atomic scale manifestations of friction has been dubbed nanotribology,[2] and an accessible introduction can be found in a recent Resource Letter.[6] More specialized reviews of experimental and theoretical work are also available.[7–11]

The goal of this article is to discuss some simple demonstrations and simulations designed to illustrate important principles associated with the modern view of friction. The purpose of these simulations is to enrich students' conceptions of friction at an early stage in their education. Key concepts are illustrated with simple experiments and with simulations developed using *Interactive Physics 3.0*.[12] All the simulations are available through EPAPS,[13] and make it easy for students to explore the effect of varying geometry, forces, velocities, and other parameters. Movies that illustrate the dynamics for examples discussed in the text are available at the same location, and can be used in lecture or laboratory settings.

We begin with a review of the friction laws developed by Amontons over 300 years ago.[1] Then, simple examples that violate these laws are discussed. The first simulations emphasize the connection between static friction and potential energy.[14] They are followed by simulations that illustrate the connection between kinetic friction and energy dissipation, and show that static friction is surprisingly rare in simple atomistic models. Some explanations for the omnipresence of static friction in macroscopic systems are described, and a mechanism based on the debris present between nearly all surfaces is illustrated using simulations. The article concludes with a brief explanation of why Amontons' laws often work in macroscopic experiments.

II. AMONTONS' LAWS OF FRICTION

The two basic laws of friction that are taught today were first published by Amontons, but known to da Vinci many years earlier.[1] The laws state that the frictional force F is proportional to the normal force or load L holding the two surfaces together, and that F is independent of the area of the surfaces. These rules, particularly the first one, are used by high school and college students in many problems.

In most of these problems we consider the load to be the component of the weight of an object that acts perpendicular to the surface with which it is in contact, plus any other external forces that might be acting on the object in that direction. Amontons' first law is often expressed by saying that the ratio F/L has a constant value known as the coefficient of friction. Euler introduced the symbol μ for the friction coefficient.[1] In general, μ depends on whether one measures the static friction force F_s needed to initiate sliding, or

the kinetic friction force F_k needed to maintain sliding. Many texts discuss a third law, due to Coulomb,[1] which says F_k is independent of the sliding velocity.

A common way of measuring the coefficient of static friction μ_s is to place an object on an inclined plane and increase the slope until it begins to slide. If θ is the angle relative to the horizontal and M is the mass of the object, then the normal force due to gravity or the load is $L = Mg\cos\theta$. The object begins to slide when the tangential force $Mg\sin\theta$ exceeds the static friction F_s. If the angle at which sliding begins is θ_s, then $F_s = Mg\sin\theta_s = L\tan\theta_s$ and $\mu_s = \tan\theta_s$. We can add masses to make the object heavier, and in many cases we find that θ_s does not change, which is consistent with Amontons' first law. To test the second law, we repeat da Vinci's classic experiments[1] by taking an object like a brick and placing surfaces with different areas in contact with the inclined plane. For many common objects the value of θ_s is independent of the macroscopic area of the contacting surfaces, thereby establishing Amontons' second law.

III. FRICTION WITH $L \leqslant 0$?

Students may wonder if Amontons' laws are unbreakable physical laws or simply rules that work well for many common situations. Consider what they predict for two surfaces pulled across each other in the weightlessness of space with no external force. Is there really no friction in the absence of a normal force? It turns out that friction can be even more difficult to deal with in space than on earth.[15] Can students think of an object that will not slide on a vertical plane, like the wall of a room? Most will think of tape, putty, or other sticky objects that adhere to surfaces. These objects don't slide even when the plane is tilted past vertical, so that gravity produces a negative load that seeks to separate the surfaces.

Sticky and compliant objects also violate Amontons' second law because the frictional force increases with the area of contact. This dependence on area can be checked by a simple experiment, where part of the bottom surface of a block is covered with a sheet of compliant material (for example, latex, putty, double-sided tape, caulk, or material from stretchable gloves or toys designed to stick to walls) and placed on a smooth glass substrate so that the area of contact can be observed from below. For stability we may coat strips on two parallel edges of a heavy block and vary the width of the strips. The static friction can be determined by pulling the block with a spring balance along the length of the strips and measuring the force needed to initiate sliding. For such surfaces, the measured friction rises with the area of contact even when the load is fixed, violating Amontons' second law.

The static friction also increases with the weight of the block at a fixed area, as would be expected from the first law. This type of behavior was observed in some systems by Coulomb,[1] who fitted F to a linear function of both load and area A:

$$F = \mu L + cA, \tag{1}$$

where c is a constant coefficient. As illustrated below, Eq. (1) applies to many atomic scale studies of friction.

Our first simulation illustrates how a friction force that violates Amontons' laws can arise in a simple atomic-scale system. The key element is the attraction between atoms at the interface between the two objects. As in many computer simulations of friction,[10] we model the interactions between two atoms separated by distance r with a Lennard-Jones potential[16]

$$U(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]. \tag{2}$$

The parameter ϵ gives the binding energy of two atoms at the separation, $r = 2^{1/6}\sigma$, which minimizes their potential energy. The attractive term in Eq. (2) has the r^{-6} dependence of the attractive van der Waals or fluctuating dipole interaction between all atoms.[4] The r^{-12} term represents the Pauli exclusion repulsion when the electrons around the atoms begin to overlap. The mass m of the atoms is assumed to be large enough that they can be treated as classical objects, which is a good approximation for most atoms at room temperature.[16] All quantities are expressed in units of ϵ, σ, m and the characteristic time $\tau \equiv \sqrt{m\sigma^2/\epsilon}$ (Appendix).

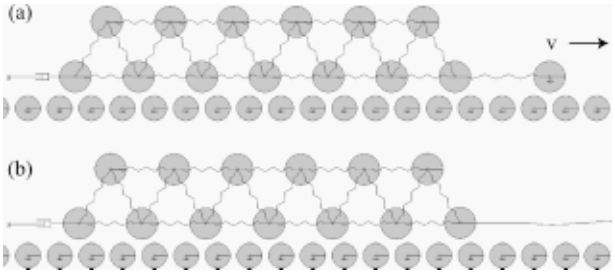

Fig. 1. Geometry of the first simulation [sim1 on EPAPS (Ref. 13)]. Atoms in the bottom wall are fixed, and the top wall is pulled by a spring whose other end (far right) advances at constant speed v. The strong bonds within the top wall are represented by springs ($k = 500\epsilon/\sigma^2$). The atoms are shown as circles of radius 0.4 σ (bottom) or 0.5 σ (top), but their interaction [Eq. (2)] is minimized at the larger separation of $2^{1/6}\sigma$. In the initial state (a) all atoms are at this ideal separation. In (b) the top atoms have been displaced to the right. The external force from the spring is balanced by the force from the bottom atoms, which pushes the top atoms back to the left. The damper stabilizes the system when it starts to slide or the load is suddenly increased.

Figure 1 shows the simple two-dimensional geometry that is used in the first simulation of friction between two crystalline walls (sim1).[13] The bottom wall is represented by a line of fixed atoms. The top wall contains 13 atoms. The strong interactions binding this wall together are modeled by springs between neighboring atoms. Only the Lennard-Jones interaction couples atoms on different walls.

The top wall is pulled by a spring attached to a particle that advances at a constant speed v parallel to the bottom wall. The particle and spring could represent the motion and stiffness of an external measuring device, like an atomic force microscope. We may also think of the particle and spring as modeling a much thicker top wall than would fit in the simulation. The particle then corresponds to the position of the top of the wall and the spring to the elastic stiffness of the wall. A stabilizing damper has been added to keep the top surface from being thrown free of the bottom wall when motion starts. It would not be necessary in a more realistic model, and students can check that setting the damping constant to zero does not affect the static friction.

The surfaces in Fig. 1 are ''commensurate,'' because the ratio of the spacing between atoms on the top surface to that on the bottom wall is a rational number, in this case 2. This ratio allows all top atoms to simultaneously minimize their potential energy. In the lowest energy configuration [Fig.

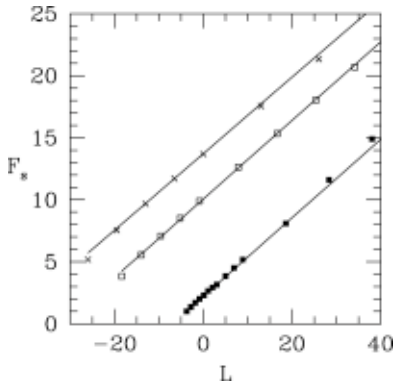

Fig. 2. Static friction as a function of load in simulations with the larger (pluses) and smaller (open squares) commensurate walls (sim1 and sim2), and for a single atom (filled squares) (sim3). The straight lines give a fit to Eq. (1) with the same values of $\mu = 0.308$ and $c = 1.96$.

1(a)], each top atom fits neatly between two bottom atoms and obtains the full binding energy ϵ from each. (It looks like the atoms don't touch only because the circles have a smaller diameter than the spacing that minimizes the energy, $2^{1/6}\sigma$.) Any lateral movement stretches the bond to one of the two bottom atoms and increases the potential energy. The derivative of the potential energy with respect to the lateral displacement represents an internal force, friction, that counteracts the external lateral force.[14]

In the initial phase of the simulation (sim1.avi),[13] the top wall is nearly stationary. The moving particle gradually stretches the spring, and the force it applies on the wall increases linearly with time. A close examination of a snapshot of the simulation [Fig. 1(b)] reveals that there is a small forward motion of the top wall. It advances until the lateral force from the bottom surface matches the external force exactly. This motion is too small to detect in a macroscopic friction experiment, but can be seen using an atomic force microscope.[3] At some point, the external force reaches the largest value that the internal forces can resist. This force corresponds to the static friction force, and is about 14 ϵ/σ for this simulation even though $L=0$. Once the external force exceeds F_s, the top wall begins to accelerate rapidly, lowering the extension of the spring, and thus the external force. Because the load is zero and the friction and adhesion are very large, the top wall rips free before settling back on the bottom wall.

To illustrate the effect of the area, the rear four atoms of the top wall were removed in the second simulation (sim2).[13] This removal reduces the area of contact by a factor of 5/7. Figure 2 shows that the static friction decreases by the same ratio, from about 14 ϵ/σ to 10 ϵ/σ. Thus, this simple simulation gives a friction force at zero load that is proportional to the contact area, violating both of Amontons' laws. Similar violations of Amontons' second law are observed in many nanoscale experiments.[6]

Students can apply a load to the top wall of these two simulations by adding a gravitational acceleration. As long as the magnitude of the load is not too large, the static friction rises linearly with load. Figure 2 shows that the results for both simulations can be fit to Eq. (1) with $\mu = 0.308$, $c = 1.96$, and the ''area'' measured in terms of the number of atoms on the bottom surface of the top wall. This fit even extends to small negative loads, where the static friction is decreased but still present. Of course, when too large a negative load is applied, the surfaces are ripped apart and there is no longer any friction.

The third simulation illustrates why Eq. (1) describes the above results (sim3).[13] Here, a single atom is pulled by the same spring used in the first two simulations. The dashpot is replaced by a drag force proportional to the velocity, which allows the system to equilibrate after sudden changes in load. Figure 2 shows that the force needed to start this atom moving lies on a line with the same μ and c as the first two simulations. For all of these simulations, c represents the force needed to dislodge each atom from its potential energy minimum in the absence of an external load. The third simulation shows that adding a normal force on an atom produces a linear increase in the force needed to dislodge it. Because the increase is linear, the total increase in friction is independent of the number of atoms sharing the load, and all simulations have the same μ. Note that this explanation for Eq. (1) assumes that the forces from each atom add coherently. Section IV discusses problems with this assumption.

A simple geometrical argument explains why the friction on a single atom rises linearly with the normal force.[17,18] The repulsive r^{-12} term in Eq. (2) prevents the atom from approaching too close to atoms on the bottom wall. As a result, the atom must move upward as it moves from a point midway between two atoms to a point directly above an atom. When a normal force is added, an extra lateral force is needed to pull the atom up over atoms in the bottom wall. If the atom moves up at an angle ϕ, then the extra friction is just the load times $\tan\phi$. As expected from this argument, plotting the trajectory of the atom shows that $\tan\phi$ approaches the measured value of μ as the force approaches F_s.

Students can check this simple argument by calculating the force exerted on an atom as it moves over two atoms that form part of the lower surface. A Lennard-Jones interaction, or a simpler purely repulsive power law (r^{-n}) or exponential (e^{-kr}) potential, can be used. For each value of the lateral position, x, the height of the atom must be adjusted until the vertical force equals the constant load. The lateral force at this varying height will vanish at the midpoint between the two lower atoms and rise to a maximum as x increases. This maximum corresponds to the static friction force. Students can verify that the static friction force increases linearly with load over a wide range of loads, as illustrated in Fig. 2.

IV. HOW ARE FRICTION AND ADHESION RELATED?

One common way of explaining Eq. (1) is to argue that the adhesive forces between surfaces act like an effective load that must be added to the external load. If the adhesive load is proportional to the area, then Eq. (1) reverts to Amontons' first law and only the second law is violated. The idea that adhesion between materials is important to friction has historical roots that date back to Desaguliers in 1734.[1] He experimented with smooth metal surfaces and found that both adhesion and frictional forces could be greatly increased by polishing the two surfaces to be placed in contact. The reason for this increase is described in Sec. VI.

Although there is often a correlation between adhesion and friction, adhesion alone is not enough to produce friction. The easiest way to understand this conclusion is to

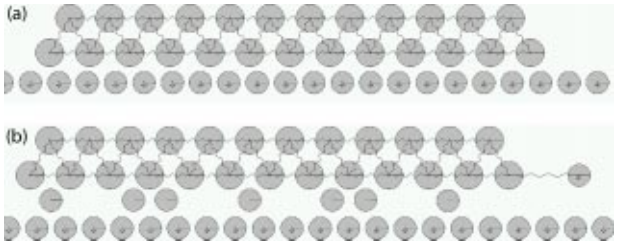

Fig. 3. Initial geometry for simulations with incommensurate walls (a) without (sim4) and (b) with (sim5) debris atoms in between. Atoms in the bottom wall are fixed at a separation of σ and the spacing in the top wall is $\sqrt{2}\sigma$. In (a) a constant force is applied to the rightmost atom on the top wall. In (b) this atom is attached to a spring whose other end moves to the right at constant speed ($v=0.25\sigma/\tau$). The springs within the top wall have $k=500\epsilon/\sigma^2$ in (b) and are reduced to $100\epsilon/\sigma^2$ in (a) to make the vibrations of atoms visible. A small damping is added in (b) to allow the system to settle into an energy minimum.

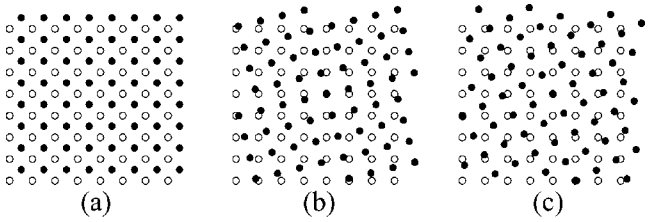

Fig. 4. When two identical walls are in perfect alignment (a), they are commensurate and all atoms on the top surface (closed circles) can nestle between atoms in the surface below (open circles). Rotating the top surface produces incommensurate surfaces [(b), (c)], where each top atom sits at a different position relative to the surface below.

imagine replacing the bottom wall in Fig. 1 by a perfectly smooth wall. The top wall will adhere to the bottom wall as long as there are attractive interactions that pull the walls together. However, there will be no static friction because the potential energy from this attraction is not affected by lateral translations—the system is invariant under sliding. One might argue that real walls are made of discrete atoms that break this translational symmetry. However, it turns out to be remarkably easy to make models with zero static friction.

Figure 3(a) shows the geometry for a simulation where the spacing between atoms on the top and bottom surfaces is an irrational number, $\sqrt{2}$ (sim4).[13] Such surfaces are called incommensurate.[19] Unlike the commensurate surface in Fig. 1, there is no way for all the atoms to simultaneously fall into potential energy minima between atoms on the bottom wall. Indeed, every atom on the top wall lies at a different position relative to the nearest atoms on the bottom wall. The interaction between walls pushes some top atoms to the right and others to the left. If the walls are long enough, the lateral forces cancel exactly, while the adhesive forces grow linearly with area. For short walls, like that in Fig. 3(a), the cancellation is not complete because of edge effects. The simulation shows the top wall begins to advance at forces greater than about $\sim\epsilon/\sigma$. This force is an order of magnitude smaller than the static friction for the first simulation, even though the area is twice as large. Indeed, the friction for the entire incommensurate wall is much smaller than the contribution of each atom on the commensurate wall at zero load, $\sim 2\epsilon/\sigma$. Studies of larger incommensurate systems show that the ratio of friction to adhesion vanishes as the wall area grows.[11,18]

Students can obtain a flavor for incommensurate systems by considering a string of n atoms moving along the x axis in a sinusoidal potential. If the spacing between adjacent atoms is a, and the period of the potential is b, then the maximum lateral force depends on n and a/b. This maximum force corresponds to the static friction, and students will find that it only scales linearly with n when a/b is an integer. How does it scale when a/b is another rational number, or an irrational number like $\sqrt{2}$? Some answers can be found in Refs. 11 and 18.

An examination of the atomic motions in the incommensurate simulation (see sim4.avi[13]) illustrates the connection between kinetic friction and energy dissipation. Unlike static friction, the kinetic friction F_k cannot reflect an increase in the interfacial Lennard-Jones potential energy. Moving the top wall forward by one lattice constant of the bottom wall produces a system with the same geometry, so the average lateral force from the change in interfacial potential energy vanishes.[11,20,21] In the simulation, the top atoms move up and down over bottom atoms at different times. This motion leads to a cancellation of the lateral forces, and the center of mass moves forward smoothly. However, the relative motion of the top atoms excites sound waves that propagate in the top wall. These vibrations are called phonons, and the amount of energy stored in them grows with time during the simulation. The energy comes from the external work being done on the system; that is, the kinetic friction. The rate of external work is $F_k v$, where v is the speed of the top wall. The first law of thermodynamics implies that this work must either increase the internal energy of the system or produce a heat flow out of the system. The top wall in the simulation is isolated so the energy can only go into internal vibrations. These phonons would propagate away from the surface as heat if the coupling to the environment were treated more realistically.[10,11]

Experiments using a quartz crystal microbalance allow these ideas to be tested in a real system. Atoms like krypton or xenon adsorb onto a metal surface forming small incommensurate islands much like that in Fig. 3. The metal is then shaken laterally and the frictional work is measured. For some systems the friction can be quantitatively described in terms of energy transfer to phonons that are generated as the islands slide.[22–24] Energy can also be transported away from the surface by exciting electrons into higher energy levels.[7,22]

V. WHY IS THERE STATIC FRICTION?

The simulation just described (sim4) suggests that static friction should be rare, because two contacting surfaces will generally be incommensurate. Even two identical surfaces become incommensurate when they are rotated about an axis perpendicular to their interface. This is easily illustrated by making two transparencies with a regular lattice of dots and rotating the top one on an overhead projector. Figure 4 illustrates some of the resulting patterns for a square lattice. When the lattices are aligned, each dot has the same position relative to the other lattice. Rotating the lattices out of alignment produces a fascinating variety of Moire patterns, where each dot is at a different position relative to the other lattice. If each dot were replaced by an atom, the lateral forces would be in all different directions and the static friction would vanish. Random surfaces also sample all possible positions, leading to a cancellation of the lateral force.[18]

Figure 1 is unusual even for commensurate systems, because all the top atoms can be moved to a local energy minimum by a rigid translation. A lateral displacement of the wall increases the energy of every top atom at the same time, giving a large static friction. If the ratio of spacings were 3/2, the surfaces could not be brought into perfect registry. We could place half the atoms at local minima, but the other half would be at local maxima. As a result, the energy that can be gained by rigidly translating the wall, and the corresponding friction, is small.[18]

To obtain a static friction force that rises linearly with area, the forces on the atoms at the interface must be large enough to produce large rearrangements of the atoms into new local energy minima. This rearrangement also produces something called adhesion hysteresis: The energy gained by bringing two surfaces together is less than the energy required to break them apart. Recent work suggests that the strength of friction should be correlated with the amount of adhesion hysteresis, rather than the strength of adhesion itself.[25]

A wide variety of surface rearrangements have been proposed to explain experiments, and each may be valid in specific cases. Most rely on the forces at the interface being comparable to the forces holding the contacting solids together. There are some obvious cases where this condition applies. For example, machining, polishing, or sanding generate local pressures strong enough to break internal bonds, removing material and leaving behind grooves. The simplest example is plowing friction where a sharp tip makes a scratch in a surface.[26] The static friction corresponds to the force required to break bonds and push atoms out of the path of the advancing tip. The local pressure required to produce a rearrangement of internal bonds is called the hardness. Plowing friction occurs when the tip is harder than the substrate, otherwise the tip will flatten. Diamond, carborundum, and tungsten carbide are commonly used as abrasives because of their high hardness.

Many textbooks describe another mechanism for surface rearrangement that is observed in contacts between clean metal surfaces in ultrahigh vacuum. The surface atoms have a very high energy that can be lowered by contact with almost any atom. When placed in contact with another metal, the energy of interaction with the other surface is of the same order as the energy holding the atom onto its own surface. Atoms diffuse to maximize metal/metal contact and minimize metal/vacuum interfaces. A beautiful illustration of such atomic rearrangements can be seen in simulations of a nickel tip contacting a gold substrate.[27] The process of forming these metal junctions is called cold welding, and the junctions resist lateral motion much like a bulk solid. The static friction corresponds to the force needed to break these bonds, and experiments show that fracture may occur away from the actual interface, leaving behind patches of one metal on the other's surface.[26]

Although cold-welded junctions provide one explanation for static friction, there are several reasons why this explanation cannot be general. One problem is that breaking the junctions would wear away material at a much higher rate than observed.[26] A second problem is that when surfaces are exposed to air, they react to lower their energy, reducing the energy they can gain from rearranging in response to another surface. Most metals are quickly covered with many atomic layers of an oxide, and all surfaces are rapidly coated with a thin layer of molecules that are absorbed from the air. These contaminant surface layers are difficult to push out unless the pressure is high enough to break bonds in the material below.

The most important difficulty with the cold-welding model is that friction measurements are very sensitive to any surface contamination.[26] As emphasized by Feynman,[28] "friction is never due to copper on copper, but due to the impurities clinging to copper." These impurities are traditionally called "third bodies," because they separate the two macroscopic objects to which the external forces are applied. Since prehistoric times, surfaces have been modified to reduce friction by adding thin layers of animal fat or liquids.[1] Today, single layers of molecules can be precisely deposited to lower friction on hard disks or microelectromechanical machines.[8] Such layers replace the native contamination on surfaces and can reduce the friction by an order of magnitude. They also prevent welding of surfaces.

The layers of molecules between surfaces, whether intentionally applied or native, provide another mechanism for rearrangements that produce static friction. The interactions between surfaces may be too small to rearrange atoms within the contacting solids, but strong enough to rearrange the much weaker bonds within the layer. Unlike the bounding solids considered in many models (for example, Fig. 4),[11] surface layers are generally not in an ordered crystalline state. Like glasses or other amorphous solids, they have many arrangements that are local energy minima. One of these local minima will be able to conform to any given geometry of the bounding surfaces. Once locked in this minimum, the film will resist lateral motion.

The final simulation illustrates how debris atoms between surfaces can lock them together and produce static friction. The bounding solids are the incommensurate walls used in the fourth simulation. Seven spherical atoms of the same size are free to move in between them. These atoms interact with each other and with wall atoms through the same Lennard-Jones potential. When the simulation starts, debris atoms rapidly move to local minima where they are between two atoms on each surface. They then resist any lateral motion, producing friction even though the direct interactions between walls are negligible. The top wall remains pinned until a static friction of about 11 ϵ/σ. This force is more than an order of magnitude larger than the static friction for the same incommensurate walls with no debris, and comparable to the force between commensurate walls without debris in Fig. 1. As in the first three simulations, students can vary the load by changing the gravity. As a project they can verify that results for loads between $-12\epsilon/\sigma$ and more than 100 ϵ/σ can be fit to Eq. (1) with $\mu = 0.20$ and $c = 0.85$. The area can also be changed by deleting or adding wall atoms and debris.

More detailed simulations show that the same mechanism occurs when debris is placed between commensurate or disordered walls.[17,18,29] In all cases, the friction increases with area and load according to Eq. (1). Experiments with the surface force apparatus also show that films of almost any molecule placed between atomically flat surfaces lead to static friction.[30–34]

VI. WHY DO AMONTONS' LAWS OFTEN WORK?

The above simulations and many nanoscale experiments[6,8] indicate that Eq. (1) provides a more accurate description of friction at nanometer scales than Amontons' laws. Students may be left wondering why Amontons' laws work so well in many macroscopic experiments. The answer is that the area

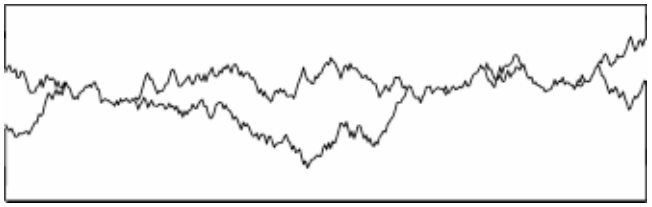

Fig. 5. Illustration of contact between two rough surfaces. The real contact area is limited to the small regions where peaks on the two surfaces overlap.

that enters Eq. (1) in the above examples is the area where atoms on opposing surfaces are close enough that they interact strongly (within ~ 1 nm). For most macroscopic surfaces, this real area of contact is much smaller than the apparent geometrical area A_{app} obtained from the macroscopic dimensions.

Figure 5 illustrates the typical geometry of contacts between macroscopic surfaces. Contact is limited to places where peaks called ''asperities'' on one surface, coincide with peaks on the opposing surface. These contact regions flatten under the applied load and in response to adhesion between the surfaces. The actual area has been measured using conductivity,[26] interfacial stiffness,[35] and optical imaging.[36,37] Under most conditions, the area in these contact regions is small and rises linearly with the applied load. This proportionality can be understood from theoretical models that neglect adhesion and assume the solids respond elastically,[38,39] or deform plastically.[26]

Amontons' laws hold automatically if the real contact area is proportional to the load. In this case, the second term in Eq. (1) is also proportional to the load, and Amontons' first law holds. The second law also holds, because the real area of contact is independent of the apparent geometric area. The proportionality of A and L implies that when surfaces of a brick with different apparent area are placed on an inclined plane, the deformation in the contacts leads to exactly the same true area of contact, and thus the same force.

Amontons' laws fail when adhesive forces are important in determining the flattening of contacts. Adhesion is always important at zero load, because no other force acts to increase the size of contacts. Adhesion also becomes more important when the materials are easy to deform, as in tape, putty, and many children's toys that are designed to stick to walls, ceilings, or any other surface. For these materials the real and apparent areas of contact may be nearly equal. Geckos use a different strategy to achieve a large contact area.[40] Their feet contain many pads that can adjust independently to ensure full contact with a surface. Even though the interactions at each point are weak, they add up to a force that allows geckos to climb walls and ceilings.

Smoothing the surface also makes it easier for adhesive forces to pull opposing surfaces into contact. The resulting increase in contact area is the key to the observations of Desaguliers mentioned previously.[1] By polishing the two contacting surfaces, he caused them to conform to each other. This increased the effective area of contact and the resulting friction force. In general, the second term in Eq. (1) grows as the surfaces are made smoother, while the first term grows as they are roughened. There is an optimum level of roughness that minimizes friction.[26] The surfaces available to Amontons and Coulomb were generally much rougher than this optimum value, and adhesive effects could be ignored in most cases.[1]

VII. SUMMARY AND CONCLUSIONS

The demonstration using blocks coated with compliant material (Sec. III) and the simulations with commensurate walls emphasize the importance of adhesion in producing friction, particularly in the limit of smooth or compliant surfaces and small, or even negative, loads. Adhesion leads to violations of Amontons' laws that are often surprising and have been rediscovered with excitement in many publications. Both laws are violated in the first three simulations (sim1–3),[13] where the friction force is finite at zero load and proportional to the area (Fig. 2). This friction force is explained in terms of the attractive interactions across the interface between the surfaces. The potential energy of each interfacial atom is minimized by the configuration shown in Fig. 1(a). Any lateral motion increases the energy, producing a lateral force that can counteract an external force [Fig. 1(b)]. The maximum derivative of the potential energy corresponds to the largest external force that can be balanced, and thus is the static friction.[14]

The fourth simulation [Fig. 3(a)] shows that adhesion alone is not enough to yield static friction. In general, surfaces will have different periods that prevent the neat interlocking in Fig. 1. For every atom that is at a potential energy minimum, there is another at a maximum. The net change in energy with position, and thus F_s, vanishes for large surface areas. However, motion is still resisted by kinetic friction. The simulation (sim4.avi[13]) illustrates how work done by F_k is converted into lattice vibrations. The latter represent heat that would flow into the surrounding solids. One of the earliest applications of friction was the generation of fire by rubbing sticks together.[1]

Static friction implies that the atoms at the interface have rearranged to achieve a local energy minimum. Different proposals for the types of rearrangement that might occur were discussed. One of these is based on rearrangements of the contaminant layers that are present on all of the surfaces around us. The atoms in these layers can rearrange to produce an energy minimum for any configuration of the bounding surfaces. The final simulation [Fig. 3(b)] illustrates how this rearrangement can produce a static friction that rises linearly with load and area as in Eq. (1).

Although the simulations presented with this article are limited, they provide an interesting visual and conceptual look at the role of surface geometry, interactions, and atomic rearrangements in producing friction. If students were asked to predict the results, many would be mistaken. We hope that the simulations will deepen their understanding of the complexity of friction in real systems, as well as the fundamental connections between static friction and potential energy, and between kinetic friction and dissipation. We also hope that the simulations will spur interest in more advanced questions regarding debris, lubricants, and practical engineering applications in automobiles and other machinery that can be pursued in more extensive reviews.[6–11]

ACKNOWLEDGMENTS

We thank the Materials Research Science and Engineering Center of Johns Hopkins University for their assistance

through the Research Experience for Teachers Program (⟨http://pha.jhu.edu/groups/mrsec/⟩). This material is based upon work supported by the National Science Foundation under Grant Nos. DMR0080031 and DMR0083286. J.R. thanks the administration of the Lancaster Country Day School for their support.

APPENDIX: SIMULATION DETAILS AND PROJECTS

Interactive Physics[12] uses mks units that are much larger than the atomic diameters, weights, and times of the Lennard-Jones potential. Rather than using lengths like 10^{-10} m, we work in Lennard-Jones units of atomic diameter σ, binding energy ϵ, and mass m. This choice gives a characteristic time $\tau \equiv \sqrt{m\sigma^2/\epsilon}$. The unit of force is ϵ/σ and the spring constants have units of ϵ/σ^2. The output from Interactive Physics is reported in mks units, but the user should keep in mind that the numbers are really in Lennard-Jones units. Typical values for models of real atoms are of the order of $\sigma = 0.3$ nm, $\epsilon \sim 0.3$ eV$\sim 5 \cdot 10^{-20}$ J, and $m \sim 5 \cdot 10^{-26}$ kg. These values give $\tau \sim 0.3$ ps, $\epsilon/\sigma \sim 0.2$ nN, and $\epsilon/\sigma^2 \sim 0.5$ J/m^2.

Atoms on the top and bottom walls are drawn with radii of $0.4\ \sigma$ and $0.5\ \sigma$, respectively. If the atomic separation were to decrease to the sum of these values, Interactive Physics would change their velocities as if a collision with perfect elasticity and zero friction had occurred. The sum of the radii is chosen to be small enough that such unphysical, billiard ball-like collisions do not occur in the range of loads considered.

The springs between atoms in the top wall have spring constant $k = 500$ (listed in N/m, but actually in ϵ/σ^2). This value of k is about an order of magnitude larger than the stiffness of the Lennard-Jones interactions between the two surfaces. The damper at the back of the first two simulations produces a drag force of $-10\vec{v}\sqrt{m\epsilon/\sigma^2}$, where $\vec{v}$ is the instantaneous velocity of the back atom. The pulling spring ($k = 1\ \epsilon/\sigma^2$) is attached to a sphere moving with a fixed velocity appropriate to each case.

The interaction between all the particles of the simulation is the Lennard-Jones potential. It is implemented in Interactive Physics by selecting the ''World'' menu, then ''Force Field,'' and ''Pair-Wise.'' The formula: 48*((sqr(self.p-other.p))∧-6.5), -24*((sqr(self.p-other.p))∧-3.5), is then placed in the top slot available. This formula is obtained by differentiating the Lennard-Jones potential between particles centered at positions denoted by ''self.p'' and ''other.p.'' Interactive Physics sums all forcefield interactions and all external forces to determine the net force on a particle at each time step. The time step, desired accuracy, and method used to integrate Newton's equations of motion are adjustable under the ''Accuracy'' item within the ''World'' menu.

For some of the simulations it may be beneficial to add a small gravitational attraction and increase the velocity-dependent damping at the rear of the top wall to stabilize the motion and decrease the ''bouncing'' of the wall that may be observed. Air resistance (proportional to speed) can also be included to mimic heat flow into the surroundings. It is added to the third and fifth simulations to damp sudden changes in load, allowing the system to reach a local energy minimum.

The air resistance or damper prevent the simulations from crashing when the static friction is exceeded. However, only the fourth simulation (sim4) is designed to be meaningful after sliding starts. The other simulations should be stopped once the peak force or static friction is determined. Finite temperature, thicker walls, higher loads, and other factors would generally prevent the top wall from flying off the bottom wall in a real experiment.

As projects, students can vary the gravitational attraction and determine the relation between friction and load for any of the simulations. They can also add or remove particles to explore the effect of contact area, or change the spacing between atoms along the walls to explore the difference between different commensurate and incommensurate cases. They can also explore the effect of spacing when debris particles are present. They will find that decreasing the spacing decreases μ in Eq. (1) and increasing the spacing increases μ. This effect is explained in Ref. 17.

[a]Electronic mail: mr@pha.jhu.edu

[1]D. Dowson, *History of Tribology* (Longman, New York, 1979).
[2]J. Krim, ''Friction at the atomic scale,'' Sci. Am. **275**, 74–80 (1996).
[3]R. W. Carpick and M. Salmeron, ''Scratching the surface: Fundamental investigations of tribology with atomic force microscopy,'' Chem. Rev. (Washington, D.C.) **97**, 1163–1194 (1997).
[4]J. N. Israelachvili, *Intermolecular and Surface Forces* (Academic, London, 1991), 2nd ed.
[5]J. Krim, D. H. Solina, and R. Chiarello, ''Nanotribology of a Kr monolayer: A quartz-crystal microbalance study of atomic-scale friction,'' Phys. Rev. Lett. **66**, 181–184 (1991).
[6]J. Krim, ''Resource Letter: FMMLS-1: Friction at macroscopic and microscopic length scales,'' Am. J. Phys. **70**, 890–897 (2002).
[7]B. N. J. Persson, *Sliding Friction: Physical Principles and Applications* (Springer-Verlag, Berlin, 1998).
[8]B. Bhushan, ed., *Modern Tribology Handbook* (CRC, Boca Raton, FL, 2001).
[9]J. A. Harrison, S. J. Stuart, and D. W. Brenner, ''Atomic-scale simulation of tribological and related phenomena,'' in *Handbook of Micro/Nanotribology*, edited by B. Bhushan (CRC, Boca Raton, FL, 1999), pp. 525–594.
[10]M. O. Robbins and M. H. Müser, ''Computer simulations of friction, lubrication and wear,'' in *Modern Tribology Handbook*, edited by B. Bhushan (CRC, Boca Raton, FL, 2001), pp. 717–765.
[11]M. H. Müser, M. Urbakh, and M. O. Robbins, ''Statistical mechanics of static and low-velocity kinetic friction,'' Adv. Chem. Phys. **126**, 187–272 (2003).
[12]A demo version of Interactive Physics is available at ⟨www.interactivephysics.com⟩.
[13]See EPAPS Document No. E-AJPIAS-72-010406 for Interactive Physics models (simn.ip) and movies (simn.avi) for the five simulations discussed in the text (sim1, ..., sim5). A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps. See the EPAPS homepage for more information.
[14]At finite temperature the static friction is related to the derivative of the free energy rather the potential energy. We refer to the potential energy throughout the text because our simulations are effectively at zero temperature, and because most students do not learn about free energy until their junior or senior year in college.
[15]J. W. R. Jones and M. J. Jansen, ''Space tribology,'' in *Modern Tribology Handbook*, edited by B. Bhushan (CRC, Boca Raton, FL, 2001), pp. 1159–1186.
[16]M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[17]G. He, M. H. Müser, and M. O. Robbins, ''Adsorbed layers and the origin of static friction,'' Science **284**, 1650–1652 (1999).
[18]M. H. Müser, L. Wenning, and M. O. Robbins, ''Simple microscopic theory of Amontons's laws for static friction,'' Phys. Rev. Lett. **86**, 1295–1298 (2001).

[19]Due to roundoff the actual ratio is not truly irrational, but the finite system size has a more pronounced effect than this.
[20]J. Leslie, *An Experimental Inquiry Into the Nature and Propagation of Heat* (J. Newman, 1804).
[21]M. H. Müser, ''Nature of mechanical instabilities and their effect on kinetic friction,'' Phys. Rev. Lett. **89**, 224301 (2002).
[22]M. O. Robbins and J. Krim, ''Energy dissipation in interfacial friction,'' MRS Bull. **23** (6), 23–26 (1998).
[23]M. Cieplak, E. D. Smith, and M. O. Robbins, ''Molecular origins of friction: The force on adsorbed layers,'' Science **265**, 1209–1212 (1994).
[24]M. S. Tomassone, J. B. Sokoloff, A. Widom, and J. Krim, ''Dominance of phonon friction for a xenon film on a silver (111) surface,'' Phys. Rev. Lett. **79**, 4798–4801 (1997).
[25]H. Yoshizawa, Y.-L. Chen, and J. N. Israelachvili, ''Fundamental mechanisms of interfacial friction. 1. Relation between adhesion and friction,'' J. Phys. Chem. **97**, 4128–4140 (1993).
[26]F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids* (Clarendon, Oxford, 1986).
[27]U. Landman, W. D. Luedtke, N. A. Burnham, and R. J. Colton, ''Atomistic mechanisms and dynamics of adhesion, nanoindentation, and fracture,'' Science **248**, 454–461 (1990).
[28]R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, New York, 1963), Vol. 1, p. 12-5.
[29]G. He and M. O. Robbins, ''Simulations of the static friction due to adsorbed molecules,'' Phys. Rev. B **64**, 035413 (2001).
[30]J. N. Israelachvili, P. M. McGuiggan, and A. M. Homola, ''Dynamic properties of molecularly thin liquid films,'' Science **240**, 189–191 (1988).
[31]M. L. Gee, P. M. McGuiggan, J. N. Israelachvili, and A. M. Homola, ''Liquid to solid transitions of molecularly thin films under shear,'' J. Chem. Phys. **93**, 1895–1906 (1990).
[32]S. Granick, ''Motions and relaxations of confined liquids,'' Science **253**, 1374–1379 (1992).
[33]A. L. Demirel and S. Granick, ''Glasslike transition of a confined simple fluid,'' Phys. Rev. Lett. **77**, 2261–2264 (1996).
[34]J. Klein and E. Kumacheva, ''Confinement-induced phase transitions in simple liquids,'' Science **269**, 816–819 (1995).
[35]P. Berthoud and T. Baumberger, ''Shear stiffness of a solid-solid multicontact interface,'' Proc. R. Soc. London, Ser. A **454**, 1615–1634 (1998).
[36]J. H. Dieterich and B. D. Kilgore, ''Direct observation of frictional contacts: New insights for state-dependent properties,'' Pure Appl. Geophys. **143**, 283–302 (1994).
[37]J. H. Dieterich and B. D. Kilgore, ''Imaging surface contacts: Power law contact distributions and contact stresses in quartz, calcite, glass and acrylic plastic,'' Tectonophysics **256**, 219–239 (1996).
[38]J. A. Greenwood and J. B. P. Williamson, ''Contact of nominally flat surfaces,'' Proc. R. Soc. London, Ser. A **295**, 300–319 (1966).
[39]B. N. J. Persson, ''Elastoplastic contact between randomly rough surfaces,'' Phys. Rev. Lett. **87**, 116101 (2001).
[40]K. Autumn, Y. A. Liang, S. T. Hsieh, W. Zesch, W. P. Chan, T. W. Kenny, R. Fearing, and R. J. Full, ''Adhesive force of a single gecko foot-hair,'' Nature (London) **405**, 681–685 (2000).