LUBRICATION: FACTS AND FRICTION
Embraced are familiar and useful, yet their underlying physics is surprisingly subtle and complex.

Steve Granick

Just as new tools—such as the surface forces apparatus, atomic force microscopy (AFM), and the supercomputer (see figure 1)—have pushed research activity forward, so have technological needs pulled progress in the same direction. For instance, ordinary though the need to lubricate a car engine may seem, it is less obvious how to lubricate the hard disks in our computers. To store and read magnetic information at a density that we have come to expect, the lubricating films must have the thickness of only a few molecules.

Other scientists have studied friction because their imaginations have been excited and engaged by problems of geological instabilities, such as the goal of predicting, or even controlling, the onset of earthquakes. And in biology, understanding how the moving parts in our bodies age, wear, and tear through overuse has constituted another emerging frontier (see the box on page 28.)

You may wonder why friction presents so many open questions after centuries of study, and why the mechanisms of energy loss during sliding contact remain understood so incompletely. Explaining what makes these problems so recalcitrant is another aim of this article.

Confined fluids

Imagine you are holding a billiard ball just above a droplet of water on a kitchen countertop. You then let the ball fall. Of course, the liquid squirts out—rapidly at first, then more and more slowly as the liquid thickness becomes less than the radius of the ball.

The macroscopic aspects of the phenomenon, which is depicted in figure 2, can readily be explained in terms of the continuum hydrodynamic equations that Osborne Reynolds developed during the 19th century. Because they give useful predictions of fluid flow, Reynolds's equations are widely used today in engineering calculations of complex fluid flow—provided that the fluid is homogeneous and its viscosity is low.

But a century later, experiments by Jacob Israelachvili (now at the University of California, Santa Barbara) and his coworkers showed that on the microscopic scale the film eventually stabilizes at a finite thickness of a few molecular diameters. The liquid film supports the weight of the ball! Of course, the heavier the ball, the less the ultimate thickness of the liquid film, but the result is quite general. An extraordinarily large pres-


STEVE GRANICK is in the department of materials science and engineering at the University of Illinois at Urbana-Champaign.
sure is needed to squeeze out the final few layers of liquid between two solid surfaces. The experiment's simplicity belies its far-reaching implications: When the thickness of a liquid film becomes comparable to molecular dimensions, classical intuition based on continuum properties no longer applies.

Molecular dynamics simulations of a confined fluid of chain molecules (figure 1) demonstrate the tendency of a fluid to form layers of variable density in the direction normal to the surface. High- and low-density layers alternate with the period of the thickness of the molecule. They form because of the difficulty of packing molecules in any other way. Based, as it is, on quite general considerations, this pioneering work by Uzi Landman and his group at Georgia Institute of Technology demonstrates that layering is ubiquitous.

It had long been known that layering describes fluids confined at rest between parallel surfaces of large dimension relative to their separation, but figure 1 illustrates that it persists for surfaces as small as an AFM tip (top panel) and even when bumpy solid surfaces slide past one another at high velocity (bottom panel). The bottom panel also shows that the fluid layers wrap themselves around the solid surfaces and that the number of molecular layers adjusts dynamically to the spacing between the bumps. Close inspection of the panel reveals the dislocations that occur when the number of layers changes. Simulations of this kind help us to understand how confined fluids form protective layers that keep solids separated. And they illustrate that inhomogeneity is an intrinsic property of confined fluids.

Solidification and melting

Force measurements comprise the most common experimental approach to investigating friction in situ. One of the central experimental findings of the past ten years is
that the mechanical responses of molecularly thin films to forced shear perturbations slow down spectacularly as the films become progressively thinner.4,5 As they are squeezed, the films seem to become progressively more solid.

It is interesting to reflect on what we mean by solidity. It seems to be a matter of timescale. Water is lightly viscous in everyday life, but feels solid when we belly flop into a swimming pool. Glaciers are crystalline, but their flow can be measured over a few years. That gooey plaything called Silly Putty snaps if pulled too hard (reminiscent of static-to-kinetic friction), but flows like a viscous fluid if it sits on a table. Solidity and fluidity, in the mechanical senses of these words, depend on the time of observation.

Fluid films become progressively more sluggish as they approach a thickness of molecular dimensions. It is convenient to divide the resistance to sliding into three regimes of thickness. A sufficiently thick film flows, of course, like a bulk fluid, but below a thickness of 5-10 molecular dimensions (the exact point depending on the system in question), the effective viscosity rises, the longest structural relaxation time increases, and the system increasingly tends to respond non-linearly at a given excitation rate. These three tendencies grow with diminishing film thickness and diverge below a thickness of usually 2-4 molecular dimensions (again, depending on the system). At this point, the system has become solid-like in the sense that sliding over a macroscopic distance cannot be accomplished without exceeding a yield strength.

Dynamical experiments may be performed at vanishingly low rates of shear, where the observed linear response gives information about the dynamical structure of the fluid at rest. Or they may be performed at large shear rates, such that the experiment alters the structure of the film. Either way, the resulting patterns of dynamical response in the three thickness regimes are generic, and because these generic patterns have been observed for fluids of many kinds—linear and branched alkanes, polymers, glass formers, liquid crystals, mixtures of polymers with small molecules, and even aqueous solutions—they seem to depend on confinement to a molecularly thin film, as opposed to the chemical composition of the system at hand. In other words, knowledge of bulk viscosity gives little guidance in predicting the viscosity of the same fluid in a lubricated contact. Already enhanced at low shear rates, viscosity is further modified by rapid shear.

Solidification appears to be a universal aspect of confined fluids—regardless of whether the fluid-surface attraction is strong, weak, or absent, or whether the surface is periodic (commensurate or incommensurate with the confined fluid) or even aperiodic. Even so, quantitative details of the surface-induced structure and mechanical response do seem to depend on these variables. Being able to cleanly separate system specificity from generic effects of geometric confinement clearly has a quantitative, though not qualitative, effect and remains a problem to be solved.

**FIGURE 2. EXPERIMENT** showing that a liquid can support a normal force. A liquid droplet is placed between a ball and a flat surface. The graph, in which liquid thickness is plotted schematically against time after the ball has begun to fall, shows that the film thickness remains finite (a few molecular dimensions), even at equilibrium. The side panel shows the local liquid density plotted against the distance between the solid boundaries. A confined fluid's tendency to order in layers parallel to the boundaries is indicated by decaying oscillations with a period of about a molecular dimension. This behavior differs from the radial distribution function of a bulk liquid in the sense that interesting degrees of order and disorder parallel to the boundary are also possible. When the density waves shown on the right come sufficiently close to interfere with one another, the liquid can support a normal force at equilibrium.

---

**Tribology in the Biological Environment**

Oil and lubrication are so synonymous that it is easy to discount the importance of aqueous-based lubrication—not least in our own bodies, which are full of surfaces in sliding contact. The current poor understanding of how these surfaces are lubricated has medical ramifications that affect everyone's quality of life.

For example, prosthetic hips and knees are implanted into over 100,000 people in the US each year. And even now, when the incidence of tooth decay has declined so much because of the fluoridation of drinking water, mitigating the long-term wear of tooth surfaces remains a challenging problem. The effective lubrication of the eye is also important. Many people over 70 suffer from dry eye, which is a painful and debilitating syndrome. A fourth important example concerns the need to have slippery surfaces in biomedical devices. If catheters, contact lenses, vascular grafts, heart valves, and so forth are insufficiently slippery, tissue damage and pain can ensue.

Without wishing to minimize the differences among these examples, it is certainly fair to say that fundamental issues come into play. Motion in joints takes place in the presence of biologically based lubricating fluids that occupy the joint cavity (long-chain water-soluble polymers, water electrolytes, and proteins). Blinking is mediated by a tear layer containing mucous proteins that seem to produce a continuous tear film while preserving a viscosity close to that of water. The drag of shear forces is believed, in the normal eye, to remove dying or dead cells from the eye surface, but this mechanism breaks down when the eye is dry.

Engineers debate the criteria to produce biomedical surfaces of low friction, but the underlying physics has seen so little study that scientists must resort to speculation about the mechanisms.
Rapid deformation

What happens if a very thin film is deformed more rapidly than it can keep up? Figure 3 shows the transition from rest to sliding of atomically smooth solids separated by a molecularly thin film of a simple liquid and sheared with sinusoidal oscillations of increasing amplitude at 256 Hz. The resulting motion is decomposed into one part in phase with the sinusoidal drive and another part in phase with its rate of change.

Responding like springs, the in-phase forces measured the energy stored during each oscillation cycle. The out-of-phase forces are dissipative, and followed Newton's law of viscosity, which states that force is proportional to the rate of deformation and to the energy dissipated. The figure shows that elastic forces dominate for small deflections, increasing in proportion to the amplitude of deformation—that is, the sliding surfaces are stuck.

The deformation at around 3 angstroms in figure 3 can be identified with the end of the stick region—the point when resistance gave way; the confined film snapped, and the sliding surfaces careened forward, racing to catch up with the applied oscillation. In this sort of slip regime, the forces were predominantly dissipative and could be identified with kinetic friction.

The attribution of this dramatic behavior to the interfacial fluid film, rather than to some residual surface-to-surface contact between asperities on rough surfaces, is secure because the experiments involve sliding of parallel, step-free single crystals whose surface separation can be measured with confidence to within an experimental resolution of less than 0.1 µg/cm² of fluid in a monolayer-thick film. With current techniques, the structure of fluid film in intimate contact with a static surface can just be measured—but only under restricted conditions. To be able to do so under sliding is presently just a dream.

With porous media, however, structural measurements of confined fluids can be made routinely. This method of confinement presents the advantage of a vastly greater surface area. Unfortunately, porous media are not uniform, nor can they be subjected to external shear fields. Seeking to get around these obstacles, some scientists have turned to computer simulations. But simulations are limited by the time and length scales accessible to the largest and most efficient computers, which may be inadequate. For example, the rate of a typical experiment (256 Hz as in figure 3) is too slow to be simulated by a computer at the molecular level. Mindful of these difficulties, other scientists have turned to phenomenological theory. Arguing that the microscopic details of the system at rest may not be necessary to describe the system when it is strongly driven, they have gone on to develop suggestive phenomenological models of the dynamics of driven systems.

Although many mechanisms of instability are clearly possible, they all serve, in essence, to introduce heterogeneity into a previously homogeneous system. Consequently, forces that would otherwise be distributed uniformly can be concentrated at localized spots. The mechanisms that are responsible for breakup instabilities in lubricants include slip, shear bands, cavitation, dilation of the film thickness, and, if the films are sufficiently thick, various types of secondary flow.

Instabilities occur when the sliding force is relatively low—much lower than the yield strength of a hard solid. This condition helps to explain how lubrication acts as the weak link between two solid surfaces. Instabilities were initially regarded as anomalies, but it is now clear that they are essential aspects of strongly driven fluids. As such, they will provide a rich field of study for years to come.

Controlling friction

The friction force in dry sliding is usually proportional to the force pressing the surfaces together, with a constant...
of proportionality (the coefficient of friction) on the order of 0.5 for most material pairings. Of course, lubrication reduces friction, but, except for a few rare cases, we do not understand how it does so. Commercial lubricants, such as engine oil, have so many components that it is impossible to know how any single component acts or how the many components interact with one another. The many components interact with one another. Recent work reveals fine structure during wet sliding in friction traces, as plots of frictional force against time are known. From experience, it had long been known that irregularity is endemic whenever friction is measured. But this was interpreted as artifact (from some kind of instrumental resonance or sample inhomogeneity), so that all observed irregularity in friction measurements was dismissed as having no fundamental origin. These difficulties disappear when dealing appropriately with single crystals. Fluctuations remain. Accordingly, the forces of friction should not be thought of as single entities, as textbooks imply with their force-velocity models. Rather, shear causes a state of sliding in which the time-averaged friction is essentially independent of velocity, but the instantaneous value fluctuates. Kinetic friction is familiar, but the fluctuations, which can be huge, are a new discovery.

As depicted in figure 4, friction traces show the dissipative forces building smoothly, but dropping abruptly and breaking time-reversal symmetry. Small-scale structure is found nested within larger structure, and it appears to be self-similar. The dissipation mechanism must involve coherent dissipative structures that extend over distances that are huge compared with molecular dimensions because, if they did not, they would average out over the vastly larger sample dimension. The avalanche-like events shown in the figure are reminiscent of earthquake models, granular flow, and contact angle depinning (the flow in fits and starts of droplets on a surface)—to pick just a few examples.

In these and other cases, the mechanical response of a fluid confined between sliding surfaces can be thought of in terms of interacting springs. The sliding object, having its own intrinsic stiffness, is one spring; the confined fluid, as illustrated by the elastic forces (black squares) in figure 3, can also be represented by springs. During sliding, these springs interact. Since coupled oscillators easily generate chaotic motion, it is not surprising, perhaps, to find chaos in friction. Though unnoticed, chaos may be found very generally in the world of sliding motion about us.

The growing appreciation of irregularity in friction traces has sparked research that seeks to control friction by applying notions of limiting chaotic motion. Similar ideas have already proven fruitful in other fields, such as laser technology, and are certain to assume increasing prominence in friction, too.

Friction can also be controlled, at least in principle, by frustrating the development of friction-increasing structure in the confined lubricant. In this approach, low-amplitude modulation is applied externally to the lubricating film at a rate comparable to, or higher than, the rate characteristic of confinement-induced solidification. Recent, though still tentative, experimental results support this idea.

In seeking to control friction, we must recognize that friction force and normal load can be utterly decoupled. That is, they are not, in essence, proportional to each other. A thought experiment illustrates this. Imagine that the repulsive forces between two magnets are balanced by a force that pushes them together, and then they are impelled to slide past one another while separated by a simple Newtonian fluid. The friction forces depend trivially on the fluid's viscosity and the shear rate, yet forces pushing the magnets together have nothing to do with these variables. The naive interpretation would be to say that the friction coefficient is astonishingly low, but this would miss the point that the origins of sliding and normal forces are unrelated in this example.

Five years ago, Jacob Klein and his coworkers at the Weizmann Institute of Science in Israel elegantly demonstrated decoupling in a system that may find practical application. The group found that although end-attached polymer chains in solvent (polymer brushes) strongly resist being pressed together for thermodynamic reasons, their interfacial sliding over one another may be domi-
nated by an entirely different physical mechanism.\textsuperscript{14} Thus, decoupling, which can cause friction forces to remain small even as normal forces become large, constitutes another approach to minimizing friction.

**Looking forward**

Friction is so commonplace that its connotations of impediment, obstruction, and waste crop up in our language far outside physics. But in academia, studies of friction are rare and are scattered among different disciplines. The physics of friction has a short pedigree, and the investigation of how energy is dissipated when a system is forced far from equilibrium has not been fashionable until recently.

By contrast, the applied side of friction studies has deep roots in mechanics (control of friction and wear) and in applied chemistry (lubrication). Not surprisingly, one reason why progress in understanding friction has been slow is the minimal interaction among physicists, engineers, and chemists. Publishing in journals appropriate to their disciplines, these Balkanized researchers have tended to focus on problems in their own specialties and to ignore the larger common problems. Fortunately, tribology, which bridges the gaps between the traditional disciplines, is now benefiting from the push toward complex interdisciplinary research.

We are living in a period of exciting advances. The traditional interest in dry sliding is being augmented by promising efforts to understand the molecular mechanisms of wet sliding. Although difficulties in defining the system under study plagued earlier attempts to understand molecular mechanisms, the use of atomically smooth single crystals makes it possible for experimenters to precisely determine the area of contact of sliding surfaces and the thickness of an intervening fluid layer. At the same time, ever more powerful computers are becoming available, and simulations of increasing molecular realism are emerging.

Identifying the molecular origins of confinement-induced solidification remains an active problem, but it has become clear that confinement may introduce a new length scale that, depending on geometry and the strength of fluid–surface interactions, might modify the glass transition. Indeed, in some circumstances, the glass transition temperature of polymer liquids in thin films at rest can be tens of degrees elevated from its value in the bulk (under other circumstances it appears to be depressed). This unexpected observation has triggered considerable activity in polymer physics,\textsuperscript{15} but it remains to be embraced by tribology.

Confined fluids also belong in the realm of liquid physics. The surface-to-surface spacing of a confined fluid may be comparable to the correlation length of short-range packing, the size of supramolecular structures, the size of molecule, or all three. The connection to friction is that high-density, short-range packing and dynamical rearrangements of structure are also interrelated when dealing with the liquid state. Unlike the solid and gas states, many of whose basic properties are understood in principle, the liquid state remains imperfectly understood at the most fundamental level. The more deeply the physics of liquids—confined or free—is understood, the more effectively will technology be able to exploit and control friction.

I conclude by mentioning three interesting questions which I do not have room here to explore in detail.

First, when rapidly driven far from equilibrium, how does a fluid avoid the development of forces so large that they would tear apart its container? The problem does not arise when pouring water from a pitcher because the shear rates are so slow, but it is endemic when dealing with the very high shear rates characteristic of confined fluids (the natural length scale by which to normalize sliding velocity is the film thickness). It seems clear that there is no single mechanism. Solutions to this problem will bring the field into closer contact with emerging issues of polymer processing and, more generally, the boundary conditions for fluid flow over a surface.\textsuperscript{16}

Second, how do fluids respond to a severe and sudden change of environment? Until now, the field of friction studies has concentrated on systems that are subject to steady external conditions. But it often happens that fluids sustain ultrafast and transient alteration of pressure, deformation rate, or film thickness. We do not yet have good ways of thinking about how a fluid responds to a severe change of environment that might last for a tiny fraction of a second.

Third, is it realistic to expect a general theory of surfaces in sliding contact? For a long time, it has been clear that—intriguingly—much of the flow behavior of confined fluids so resembles granular materials such as sand, powder, and colloidal particles\textsuperscript{11} that an observer could not distinguish between the two if the experiment were performed in a black box. Too often, the models that one proposes for a given system are specific to that system—but common responses strongly suggest more universality, which could reflect the fact that high-density, short-range packing, and dynamical rearrangements of structure by instability are interrelated. The challenge, then, is to predict from theory rather than empiricism, what makes lubricant molecules of one chemical structure more effective in lowering friction forces than those of another.

I have written this article under the auspices of the National Science Foundation and the US Department of Energy’s division of materials science, under award no. DEFG02-ER9645439, through the Frederick Seitz Materials Research Laboratory at the University of Illinois at Urbana-Champaign. I am indebted to many colleagues, especially Uzi Landman, J. Klaver, Nenad Kostic, and A. Levent Demirel, for reading the manuscript critically.

**References**