Towards an atomistic understanding of solid friction by computer simulations

Martin H. Müser

Inst. f. Physik, WA 331, Johannes Gutenberg Universität, 55099 Mainz, Germany

Abstract

Friction between two solid bodies in sliding motion takes place on a large spectrum of length and time scales: From the nanometer/second scale in an atomic force microscope up to the extremely macroscopic scales of tectonic motion. Despite our familiarity with friction, fundamental questions about its atomistic origins remain unanswered. Phenomenological laws that describe the friction in many systems were published more than 300 years ago by Amontons: The frictional force is proportional to the applied load and independent of the apparent area of contact. The atomistic origins of this simple law is still controversial. Many explanations, which seemed to be well-established until recently, have been called into question by new experimental results. Computer simulations have also revealed flaws in previous theoretical approaches and led to new insights into the atomistic processes responsible for friction. In this paper, selected computer simulation studies of friction will be discussed. Special attention will be given to how it is possible to gain insight into tribological processes that take place on macroscopic time scales with the help of atomistic computer simulations which are typically constrained to the nanometer and nanosecond regime.

Keywords: Friction; Lubrication; Computer simulation; Material transfer

1. Introduction

Friction is not always an unwanted phenomenon. It makes our jam glasses stay on the breakfast table and allows us to walk up a hill and to come back down, usually in a well controlled way. The first technological revolution involving friction dates back to about 200,000 BC and consists of the invention of fire rods. This technology was supposedly brought to us (homo sapiens) by the Neanderthal and not by Prometheus [1]. As a matter of fact, this myth on friction is not the only one that had to be dismissed. In recent years, computer simulations (see Ref. [2] for an extensive review on recent computer simulations of friction, lubrication and wear) have revealed flaws in the atomistic picture of the origin of friction between solid bodies. This concerns in particular the issue why there should be wearless friction at all. A deeper understanding of friction might also prove useful in the developments of new technologies. What can we learn from biological systems in order to reduce friction and wear? How can we adopt biological concepts and maybe even improve them for specific applications? The experimental [3] and computational [4] study of friction between surfaces bearing end-anchored
polymers in good solvents is partially motivated by such questions.

It is a challenge for computer simulations to contribute to a more detailed but also to a more coherent picture of tribological phenomena. There are many pitfalls that can render a well-designed computer simulation study to be applicable to very specific, usually unencountered circumstances only, e.g., the use of commensurate walls or one-dimensional solids, absence of surface curvature or contamination, neglect of bulk dislocations and long-range elasticity, constant separation and constant velocity constraints, to name a few. Not all of these points are relevant for every sliding system, but at least one of them will matter in a qualitative way. It is surprising how large the percentage of studies is that disregard all these points at once, albeit the use of chemically realistic interaction potentials is emphasized. A more careful modeling of boundary conditions and other “details” may therefore prove to be as important if not more important than the invention of new algorithms that allow us to access larger length scales and longer times than presently possible.

Some tribological behavior seems to be surprisingly universal, e.g., the stick-slip phenomenon appears to be similar on the micro-scale and on the macro-scale, see Fig. 1. Stick-slip motion occurs when shear forces build up with time in a mechanical contact (stick phase) and the stored potential energy is released abruptly (slip phase). In a computer simulation this can be mimicked by pulling the slider with a (weak) spring over a substrate whose center of mass is fixed.

Fig. 1. Comparison of macroscopic and nanoscale stick-slip motion. Left-hand side: History of uplift at Muroto point (adopted from Ref. [5]). Right-hand side: MD simulation of two incommensurate crystals separated by a submonolayer layer of simple atoms. The time scale \( t_0 \) and the length scale \( \sigma_0 \) are in the order of 3.5 Å and 1 ps, respectively.

Some other tribological properties show much stronger scale effects. This concerns in particular the relationship between (static) friction \( F_s \) and load \( L \) which is nearly always linear for macroscopic solids and independent of the apparent contact area (Amontons’s law), while surface force apparatus (SFA) [6] and atomic friction microscope (AFM) [7] experiments often show a sublinear dependence between friction and load that resemble a \( F_s \propto (L - L_0)^{2/3} \) relationship [8,9]. One may ask whether this discrepancy is due to a scale effect or whether it is due to other conditions, e.g., Berman et al. [9] showed that the \( F_s \) dependence on \( L \) becomes linear in an SFA contact when the adhesive interaction between the contacting surfaces are shielded. Also computer simulations of flat crystalline surfaces separated by thin lubricating films confirm the picture of a linear change of \( F_s \) with \( L \) [10].

This discussions above left out chemical details. This invokes the question whether we can look for “universality”¹ in tribological behavior. Surprisingly little effort has been made in order to classify friction depending on its atomic process. Experimental studies are specific per definition. Computer simulations, however, enable us to study generic systems. They allow us to change boundary conditions, contamination, surface curvature, etc. at will to explore

¹ Here universality is not meant in the sense of the renormalization group theory.
their effects on friction, lubrication, and wear. In this article, various computer simulation studies are discussed that aim to understand the implications of such conditions on friction in a systematic way. The model underlying the simulations is presented in Section 2. Results are divided into those for dry surfaces (Section 3) and those for lubricated surfaces (Section 4). Section 5 discusses in how far the length-scale and the time-scale gap between simulations and experiments allows a comparison between simulation and experiment. Conclusions are drawn in Section 6.

2. Model

In the first chapter of his famous lecture notes, Feynman imagined a situation in which all scientific knowledge were to be destroyed and only one sentence could be passed on to the next generation containing the most information in the fewest words. He would have picked: All things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling each other upon being squeezed into one another. Of course, in order to understand details of concrete tribological systems, it is crucial to reflect the interactions sufficiently well, but a computer simulation that is based on Feynman’s seemingly simple statement presumably allows us to gain deeper insight into the fundamental mechanisms than studies of high chemical accuracy. Thompson, Robbins, and Grest suggested the use of such a simplistic model for the study of tribological systems [11]. Most simulations discussed in the remainder of this paper are based upon this model. It allows us to change physical properties (degree of adhesion, lattice constants, etc.) in a continuous way. This flexibility can be used to better highlight relevant mechanisms. Simple models also enable us to access longer time scales and larger system sizes than those feasible in studies of high chemical accuracy.

The system consists of two walls composed of discrete atoms. The walls are coupled to their equilibrium lattice sites by springs of stiffness $\kappa$. In the limiting case of rigid walls, the coupling is considered infinitely strong, $\kappa = \infty$, and the atoms are constrained to their equilibrium positions. The periodic image convention is applied in the plane of the walls, which is always chosen to coincide with the $xy$ plane. The particles between the walls (if present) are either atoms or short chains of varying length. All monomers interact with each other and with wall atoms via a truncated Lennard–Jones (LJ) potential,

$$V(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right] + V_c,$$

where $r$ is the separation, and $\epsilon$ and $\sigma$ are characteristic energy and length scales, respectively. Wall atoms from opposing solids interact via the same LJ potential. The potential is cut off at $r_c$ and shifted by $V_c$ so that $V(r_c) = 0$. Adjacent monomers on a chain interact via an additional FENE potential [12]

$$V_{\text{CH}}(r) = -(1/2)kR_0^2 \ln \left[ 1 - \left( r/R_0 \right)^2 \right],$$

where $R_0 = 1.5\sigma$ and $k = 30\epsilon/\sigma^2$. Frequently, quantities are expressed in units of $\sigma$, $\epsilon$, and the mass $m$ of one monomer. The characteristic time is $t_{\text{LJ}} = \sqrt{m\sigma^2/\epsilon}$.

In order to treat tip-sample systems, the coupling of atoms in the upper wall to their equilibrium positions was altered such that a Hertzian pressure profile is generated if the tip (upper wall) is pressed on a perfectly flat, infinitely hard, and non-adhering surface [13]. This can be achieved by using a normal restoring force

$$f(\delta z) = \sqrt{\delta z}/(K\sqrt{R_c}),$$

where $\delta z$ is the normal deflection of an atom from its ideal lattice position, $R_c$ is the radius of curvature of the tip, and $K$ the bulk modulus of the tip.

3. Analysis of dry friction by computer simulation

Most contacts are contaminated with atoms, even if novel sliding systems such as carbon nanotubes nested in each other are an exception to this rule [14]. However, the study of dry friction is of fundamental relevance, e.g., it allows us to test whether concepts such as the Frenkel–Kontorova or the Tomlinson model [2,15] capture the relevant atomistic friction processes or how these models have to be altered in order to achieve this.

3.1. Geometric interlocking

Early theories of friction were based on the purely geometric argument that friction is caused by inter-
locking of surface asperities [1,16]. The idea is that asperities of the top surface must be lifted up a typical slope \( \tan \alpha \) determined by the roughness on the bottom surface. If there is no microscopic friction between the surfaces, then the minimum force to initiate sliding is \( F_s = L \tan \alpha \). However, asperities on real surfaces usually do not match well and on average, for every asperity or atom going up a ramp, there is another going down. Such considerations were used to predict the following dependence of the (static) friction coefficient \( \mu_s \) on the area of contact \( A \) (of a periodically repeated flat surface) [17]: For commensurate walls, \( \mu_s \) is independent of \( A \), while for incommensurate walls \( \mu_s \) is exactly zero. If at least one of the surfaces is disordered (where we assume that the roughness does not increase with \( A \)), then the law of large numbers applies and \( \mu_s \propto 1/\sqrt{A} \) follows.

Real contacts, however, are neither flat nor periodically repeated. Yet, the arguments stressed above can be generalized to a tip-substrate system such as a disordered AFM tip moving on a crystalline substrate [13]: Unlike SFA contacts, the contact mechanics of AFM tips can be reasonably well described by the so-called Hertz-plus-offset model [8,18]. We may therefore subsume adhesive effects by defining \( L \) with respect to the pulloff force so that \( A \propto L^{2\beta/3} \) applies. If we insert this into the relation \( F \propto \mu_s(A)L \) with \( \mu_s \propto 1/\sqrt{A} \), we obtain the friction-load relation frequently observed experimentally, namely \( F \propto L^{2\beta/3} \). These relations are also seen in the simulations, which were based on the model described in the previous section. In the simulations adhesion was suppressed by cutting off the interactions such that atoms from opposing solids only repel. The results for the amorphous tip are shown in Fig. 2. For comparison, tips are included that are either commensurate or incommensurate to the substrate. Such small crystalline tips can be analyzed in computer simulations, however, not yet produced experimentally. They also confirm the picture that \( \mu_s \) is independent of \( A \) for the commensurate case. Since the contact is finite, the incommensurate configurations do not show zero friction, but the value is strongly reduced with respect to the amorphous case.

It is certainly important to test whether this geometric interpenetration of the \( F_s \propto L^{2\beta/3} \) law can also be seen in chemically detailed simulations. The lack of appropriate model potentials for the experimentally relevant tip sample systems and the large number of atoms to be taken into account will make this a difficult task.

### 3.2. Elastic and plastic instabilities

Atoms in real solids have the possibility to deflect from their ideal positions. Under certain circumstances this flexibility may lead to finite friction even if contacts are flat and incommensurate: If the elastic coupling of atoms in the contact is sufficiently weak such that at every instance of time a fraction of the atoms possesses several stable equilibrium positions, then the system shows static friction, i.e. a finite force is needed in order to initiate sliding. This requirement is the quintessence of any elastic theory of solid friction and was first formulated by Prandtl in 1928 [19]. As a matter of fact, Prandtl’s model is isomorphic to Tomlinson’s model which was published one year later [20] and achieved much more recognition. While today’s model of elastic friction have reached a high level of sophistication [15], there remains a major concern: Can there be elastic instability on an atomic length scale without plastic instability? Computer simulations of incommensurate solids suggest that finite (static) shear forces should hardly ever occur without plastic flow, e.g., flat pyramidal diamond [21] and flat pyramidal copper tips [22] move.
essentially frictionless on copper surfaces before plastic deformation takes place.

In order to address this issue in more detail, it is convenient to study the generic model presented in Section 2. It has two free parameters, namely the coupling strength of atoms to their lattice sites $\kappa$ as well as the distance $a$ between neighbors in the solid. The parameters $\epsilon$ and $\sigma$ describing the LJ potential between the atoms is used as usual to define the units of energy and length scale, respectively. If one assumes that the intrabulk interactions between atoms were as large as the atomic interbulk interactions, one would find $a \approx a' = 2^{1/6} \sigma$ and $\kappa \approx \kappa' = 140 \epsilon / \sigma^2$ [23]. If one keeps $a$ fixed at $a'$ and varies $\kappa$ as a free model parameter, it turns out that finite values of $F_s$ can only be found for incommensurate surfaces if $\kappa$ is distinctly smaller than the intrinsic value $\kappa'$ [23]. This results suggests that the atomistic forces between the solids need to be larger than those within the solids in order to see elastic instabilities/friction forces. For such large interbulk interactions, however, one would expect not only elastic instabilities but also plastic instabilities including material transfer.

This expectation is supported by more detailed simulations, in which atoms are not simply pinned to their lattice sites by harmonic springs but where each solid consists of $l = 2, \ldots, 6$ layers of explicitly treated atoms and only the outermost layer of each wall is treated according to the model described in Section 2 [24]. The free parameter now is the LJ interaction strength $\epsilon_1$ that acts between atoms originating from different solids, while $\epsilon_0$ is the interaction strength for atoms originally located in the same solid. Hence for $\epsilon_1 / \epsilon_0 > 1$, the ground state of the system will be a state in which atoms of the top solid have diffused into the bottom solid and vice versa. Simulations show that due to large energy activation barriers, mixing is suppressed kinetically for a long amount of time (a few 100,000 molecular dynamics time steps) even if $\epsilon_1 / \epsilon_0 \approx 6$. Under the influence of a thermal force only, the top solid diffuses freely on the bottom solid as long as mixing does not yet take place, see Fig. 3.

The sharp drop of the diffusion constant, indicative of the onset of finite static friction forces, is accompanied by a plastic/mixing instability that leads to cold welding of the junction. A configuration for a mixed system with $l = 2$ is shown in Fig. 4. Larger systems $l > 2$ show the same trends, namely cold welding of the junction occurs within about 10,000 MD steps near values $\epsilon_1 / \epsilon_0 \approx 8$ independent of the linear dimensions within the plane and $l$.

Simulations in which mixing and cold-welding was suppressed by only using repulsive forces between atoms originally belonging to different solids does not change the picture: Only viscous drag forces oppose the relative lateral motion unless load induced plastic deformation took place. The results presented in this section thus strongly anticipate that wearless dry friction between incommensurate (monoatomic) solids should in almost all cases only be due to finite contact areas but not due to elastic instabilities.
4. Analysis of friction between lubricated surfaces

The lateral forces between incommensurate solids that are only allowed to deform elastically but not plastically changes dramatically if a thin film of physisorbed atoms or molecules is placed between the solid surfaces [10,17]. The presence of such atoms automatically leads to finite friction forces provided that the interface is sufficiently large [23]. This is achieved through a nearly infinite compliance of the film in the interface, which hence satisfy Prandtl’s condition for finite friction formulated in the beginning of Section 3.2. Simple geometric arguments and simulations even show that this friction mechanism satisfies Amontons’s law of a linear $F_s(L)$ dependence reasonably well for flat nanoscale mechanical contacts [10, 17]. In agreement with experiment, $F_s$ would be expected to be independent of the contact area if the adhesion between the contacting solids could be suppressed [9].

It has yet to be emphasized that even in the presence of a lubricating film, commensurability still influences the tribological properties strongly [25]: The static friction force can be expected to be much larger for commensurate (lubricated) systems, while the kinetic friction force is larger for an incommensurate system. Another difference is that the stick-slip motion is more erratic for lubricated incommensurate solids than for a commensurate alignment.

4.1. Adsorbed layers in curved contacts

So far, we have seen that finite friction forces can be due to the finiteness of a curved (Hertzian) contact and/or due to the presence of adsorbed layers in a flat contact. It is thus normal to ask what happens when both things come together. This question was addressed among other questions in Ref. [13]. The model consisted of a tip which allowed for long-range elastic deformations normal to the interface, see the description at the end of Section 2. The interactions were cutoff in such a way that the lubricating atoms adhered to the substrate but not to the tip. The interactions between tip and surface were also treated as non-adhesive. A typical cross section through the system is shown in Fig. 5.

It turns out that the friction-load law of this particular system is a power law $F_s \propto L^\alpha$ with an exponent $\alpha \approx 0.8$. A detailed analysis of normal and shear forces in the tip reveals that the normal forces are large at the center of the contact as well as at certain rings that can be associated with one layer lubrication and two layer lubrication. The contribution to the net shear force coming from the center of the tip is negligible. The bulk of the shear forces come from those lubricated parts where the normal forces are large as well. For this model the largest shear pressures were found at the entrance of the tip. Hence we can understand the distribution of normal and shear forces in the tip based on the simulations of dry finite tips and lubricated flat contacts, however, the value of $\alpha \approx 0.8$ is not easy to predict based on the previous studies.

4.2. Friction between polymer bearing surfaces

Wear and friction in biological systems are believed to be minimized through anchoring polymers to surfaces [3]. Grafting density, degree of polymerization, and solvent quality are such that the polymers form brushes at small external loads that occupy only relatively small volume fractions. This brings up the questions why it is efficient to graft polymers as opposed to have them adsorbed like in a typical industrial application and what tribological effects can we expect by merely imposing the boundary condition ‘end-grafted’.

These questions were addressed in a recent comparative computer simulation study of frictional drag mechanisms between polymer bearing surfaces [26]. Adsorbed and grafted polymers were considered in good and bad solution The solvent was not treated explicitly but indirectly in terms of a Langevin thermostat. It was found that the shear forces are strongly suppressed for polymers that are tilted parallel to the shear direction. This tilting is much stronger for grafted polymers than for adsorbed polymers, see Fig. 6. The tilting explains the observation that a differential friction coefficient $\mu = \partial F_{\text{shear}} / \partial L$ decreases with sliding velocity for grafted polymers but increases with sliding velocities for adsorbed polymers.
Another outcome of the simulations was that the net brush-brush interactions are mainly determined by the interpenetration of the brushes, see Fig. 7. This came as a surprise, because the simulations are in a small load regime where the shear forces are viscous, e.g., the static friction force is zero. However, the trend outlined in Fig. 7 can be understood if we keep in mind that large sliding velocities invoke strong tilting plus a reduction of the overlap and thus small friction forces.

A shortcoming of the simulations certainly is the limitation to relatively short chains (about 30 Kuhn’s segments) and the simplistic treatment of the solvent. However, the results suggest interesting applications for future research for both simulation and experiment: Does a certain degree of chemical crosslinking between the monomers in an uncompressed brush lead to a reduced interpenetration of the brushes and thus to smaller friction and more resistance to wear?

5. The time-scale and length-scale gap

One important motivation of tribological computer simulations is to find the relevant atomistic processes that occur in a single well-defined sliding contact such as in the SFA or in the AFM. While it has become computationally feasible to study atomistically the core of a nanometer scale AFM contact, the sliding speeds \( v_{sl} \) employed in the simulations are still about 10 decades larger than in experiment. The use of new, clever algorithms such as kinetic Monte Carlo might prove fruitful to overcome this time-scale gap in a brute-force way, however, such an approach is not necessarily required in order to make a meaningful comparison between experiment and simulation. Instead, one can try to find out what the relevant time scales are and what parameters define them, e.g., by studying a phenomenological description of an embedded system under shear. In such an analysis [27], it turned out that the motion of the top plate/tip is not sensitive to the atomic characteristic frequencies \( \omega_{mic} \)—as long as there is a reasonable time-scale separation between \( \omega_{mic} \) and the characteristic frequency of the tip \( \Omega_{tip} \). The experimentally relevant time scale \( \Omega_{tip} \) is given by \( \sqrt{K/M} \) where \( K \) is a restoring spring constant (defined by the tip’s elasticity and/or the tip substrate interactions) and \( M \) is the tip’s inertia, eventually including...
the cantilever’s mass. The experimental value of $M$ will be about 20 decades larger than in the simulation so that the relevant parameter $v_s\Delta \Omega_{\text{tip}}$ are similar in simulation and experiment. However, it is discussed controversially whether the treatment of the tip as a mass point is legitimate or whether long-range elasticity is relevant, see Ref. [28] and Chapter 9.9 in Ref. [5].

Another way to overcome time-scale gaps is to alter interaction strengths, temperature, or other variables in such a way that processes are moved to shorter times that usually would take place at very long times. One such example is the cold welding study of Ref. [24], which is discussed in Section 3.2: It would have been by no means desirable to base the calculations on interfacial strengths that correspond to typical experimental values. E.g., if there had been only a small interaction asymmetry in the real system (see Fig. 4) which favored cold welding, then the mixing process would have taken place on macroscopic time scales. Yet the system would have been driven in a similar direction, i.e. the ground state of the mixed system is not expected to depend sensitively on the magnitude of the interaction asymmetry. Thus, in order to study qualitatively the tribological side effects of cold welding, one has to enhance the mixing process. This can easily be done by magnifying the interaction asymmetry artificially, resulting in a reduced activation barrier for the mixing process. This simple trick thus makes it possible to bridge indirectly several orders of magnitude in time scales between experiment and simulation.

Some open issues, however, supposedly require the explicit treatment of many different length scales. This concerns in particular the question whether the elastic coherence length $\ell_e$ in a ‘typical’ macroscopic solid-solid contact is small enough so that different contacting asperities can move independently from each other. ‘Typical’ values given for $\ell_e$ range from a few $\mu$m [29] to values that are distinctly larger than the linear size of a macroscopic object [30]. This issue is of importance for the question whether clean, corrugated (macroscopic) surfaces have the potential to show elastic pinning and hence wearless friction. In order to address this, it will be necessary to do simulations that incorporate both atomistic and finite element methods.

6. Conclusions

The computer simulations discussed in this paper aimed to assist building a general atomistic understanding of tribological processes rather than to elucidate a specific sliding system; the central question being, how is it possible to obtain finite (static) shear forces in the absence of plastic deformation, wear or other types of strongly irreversible processes? The advantage of studying a simplified model as opposed to a realistic system is that we can change the system’s properties at will and study the relevance of adhesion, presence of adsorbed atoms, surface curvature, etc. in a unique way. For example, if we want to understand the effects due to adhesion we can compare the results of a simulation where all interactions are long-ranged (thus incorporating the attractive tail) and those where we cut off the relevant (intrabulk) atomic interactions such that adhesion is eliminated by leaving everything else unaltered.

Many theories on wearless solid friction rely on the picture of elastic instabilities. The simulations discussed above, however, show that simple solids should hardly ever show elastic instabilities on an atomic scale without concurrent plastic flow or material transfer. Despite large efforts to ‘design’ an atomistic system whose (static) friction mechanism is entirely elastic, no such model could be found as long as the ingredients were: three-dimensional incommensurate solids, flat contacts, and two-body interactions between all atoms. The relative lateral motion is found to be only opposed by a viscous type force that goes to zero for small sliding velocities—or—the onset of elastic instabilities is accompanied by plastic flow.

Wearless (static) friction between solids is only found if the contacts are finite or if adsorbed atoms are present in the interface: Finite contacts—in particular those where at least one of the solids has finite surface curvature—can pin geometrically. Given the validity of Hertzian contact mechanics, this leads to a $F \propto L^{2/3}$ relation for an amorphous tip on a crystalline substrate. Macroscopic solids can not pin geometrically as long as the surface roughness does not increase sufficiently fast with system size. If adsorbed atoms are present, these atoms can lock the surfaces together by adjusting their positions such that interactions are satisfied with both walls simultaneously.
One may conclude that often the shear forces depend more sensitively on the way in which the system is modeled than on the details of the interactions themselves. A popular artifact in simulations is the use of commensurate, perfectly aligned (confining) walls, which effect the simulated system strongly even in the presence of a lubricating film. Also the curvature of surfaces in a contact is frequently not treated with a high degree of care. It will be a challenge in the computer modeling of solid friction to single out all relevant details in order to make a meaningful comparison to experiment for chemically specific applications.

Acknowledgements

The research reported in here was partially supported by the German–Israeli Project DIP # 352-101 and by the Bundesministerium für Bildung und Forschung (BMBF) under grant # 03N6015.

References