Atomistic simulations of solid friction

M. H. Müser

Institut für Physik, Johannes Gutenberg-Universität, Mainz, Germany

Abstract. Friction between two solid bodies in relative 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 the effects of friction, fundamental questions remain unanswered. The atomistic origins of well-established phenomenological friction laws are controversial. Many explanations, seemingly well-established, have recently 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 chapter, selected computer simulation studies of friction and their implications will be discussed. Emphasis will be given to the question what one can learn from a friction simulation and how it is possible to avoid effects that merely arise due to poor models. Moreover, it will be outlined 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 nanosecond regime.

1 Introduction

The possibility of developing an atomistic theory of the friction between solid bodies has increased significantly over the last decade. New experimental techniques have made it possible to study well-defined mechanical single-asperity contacts, typically a few nanometers broad in an atomic force microscope [1, 2] (AFM) experiment and a few micrometers broad in a surface force apparatus (SFA) experiment [3, 4]. The physical laws observed in such nano and microscale contacts often deviate qualitatively from those observed in macroscopic systems. For instance the friction-load dependence in a single asperity contact [5, 6] usually deviates strongly from the linear relationship that is almost commonly observed in macroscopic multi-asperity contacts [7]. Another example for scale effects is the onset of oscillatory depletion forces between approaching surfaces when the confinement of a lubricant is reduced to a few nanometers and the frequently observed concurrent dramatic increase in shear forces that oppose relative lateral sliding of the two solids in contact [8, 9]. It is obvious that the fundamental understanding of tribology can only be achieved by a combination of experimental, theoretical, and computational efforts. This is not only an in-

teresting, scientific endeavor, but improving the understanding and ultimately controlling tribology (science of friction, lubrication, and wear) has been and will remain useful for the development of new technologies. One example is the design of novel micro mechanical devices which one may expect to be facilitated through a better atomistic theory of friction. From a computational point of view, even the small scale is not trivial to model realistically, because the chemically detailed, atomistic simulation of an AFM tip scratching on a surface requires the simultaneous use of techniques that are usually employed to address different regimes in space and time. Yet, despite these difficulties, atomistic simulations will not only yield insight into nanotribology but moreover become increasingly important in explaining macrotribological phenomena, among other reasons because they can provide constitutive equations for the use in finite element methods.

Another reason for the importance of simulations in tribological contexts with respect to purely analytical approaches is that there is no principle like minimization of free energy that determines the steady state of non-equilibrium systems. But even if there was, simulations would be needed to address the complex systems of interest, just as in many equilibrium problems.

This chapter is meant as a help to conduct atomistic simulations of friction between solids in a meaningful and efficient way, rather than to give a broad overview of the field. A far more comprehensive review on results of friction simulations was given recently [10]. Here, technical issues will be emphasized. In particular, I will try to point out the traps (of which there are many) that can significantly depreciate the scientific relevance of friction simulations. Their results depend to a large degree on the boundary conditions, the choice of the initial geometry, and the way in which the system is driven. This statement might sound trivial, but the idealized framework or artificial features encountered in simulations often make it difficult to compare the results of the simulations to experiments. Hence it is necessary to know *prior* to the simulation which interactions of the system are relevant and need to be included into the model and which features of the model are irrelevant. In the present context, I call a feature irrelevant if it does not change the tribological behavior in a qualitative way. The relevant features of a model may include the degree of correlation in the surface corrugations between the two sliding objects, the degree and chemical nature of surface contamination, the dimensionalities of the interface and that of the slider, the surface roughness on mesoscopic scales, the elasticity of the solids in contact, thermal and quantum fluctuations of the atomic motion, as well as the age of the contacts, to name a few. Naturally, the relevance of one feature is most always intertwined with that of another feature. Moreover the chemical composition of the interface, which is reflected in the choice of the force field, can also play an important role. In order to understand specific behavior, it may be insufficient to analyze models that merely employ simple two-body potentials.

A typical contact between two real solid bodies is such that prior to con-

tact the bodies are three-dimensional, their two-dimensional surfaces are locally curved, contaminated with dirt or intentionally lubricated, and the surface corrugations of the bodies are uncorrelated. The solids can yield plastically and long-range elastic deformations are possible, which however, are strongly suppressed by long-range elastic interactions. In simulations one often encounters the opposite situation. The surfaces are already flat before contact takes place, there is no contamination, and by some magical force, the corrugation profiles of both bodies are not only identical, but also perfectly aligned. The atoms are sometimes pinned to ideal lattice sites, disabling long-range elastic deformations, or alternatively, they are coupled as a two-dimensional harmonic sheet, thus neglecting long-range elastic interactions. Whether these discrepancies of the real world and the simulation is relevant depends on the problem under investigation. In many cases, however, the tribological properties are qualitatively altered by the simplifications.

Not only the geometry and the interactions of the model are important, but in addition, the results of tribological simulations often depend sensitively on the way in which the system is driven [11]. Driving the slider in an unrealistic way poses the second class of potential traps. One can get qualitatively different behavior if one assumes constant sliding velocity or if the slider is pulled with a weak spring. Furthermore the equivalence of different ensembles that are valid for large systems in equilibrium thermodynamics (constant separation vs. constant load, constant temperature vs. constant energy) usually breaks down in non-equilibrium situations.

The reminder of this chapter will start with the presentation of a simple but rather generic case study emphasizing the role of boundary conditions and other details. In section 2, a simple model for friction will be introduced, originally proposed by Prandtl but commonly referred to as Tomlinson model. The important lesson to be learned is that we need to identify relevant mechanical instabilities in order to understand solid friction. A discussion of dry friction will be given in section 3, including the analysis of friction forces as a function of disorder and dimensionality of interface and slider. Selected studies of simulations incorporating lubricants will be presented in section 4. section 5 focuses on technical issues, in particular on how to model the confining walls and the way in which they are driven. section 6 will contain the conclusions.

1.1 The relevance of details: A simple case study

Let us have a look at the relevance of 'details' in a simple model system. Consider two rigid, impenetrable, identical walls separated by a boundary lubricant, like a quarter layer of simple, non-reactive spherical molecules. A schematic one-dimensional representation is given in Fig. 1. The lower wall, which will be called the substrate, is supposed to be fixed, while the upper wall is pulled via a spring of varying stiffness k at a small, fixed velocity v. For large k, one expects smooth sliding of the upper wall similar to fixed sliding velocity, while at small

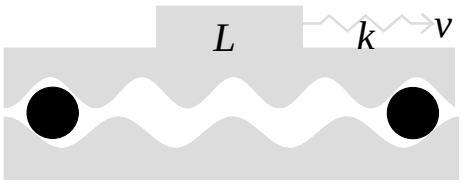

Figure 1: Schematic representation of boundary lubrication. The slider is moved with respect to a fixed substrate at velocity v via a spring of stiffness k. The two solids do not interact directly but only via the boundary lubricant (full circles).

k one expects stick-slip motion, which may best be characterized as a jerky motion of the top wall: The slider is stuck for a long time and then suddenly pops forward before it is stuck again. Stick-slip motion occurs when mechanical stress builds up sufficiently slowly in a contact. After a certain threshold force is reached, namely the static friction force, the two solids start sliding and the stored elastic energy is quickly released as kinetic energy, ultimately leading to generation of heat and/or plastic deformation. From a macroscopic description, unstable, stick-slip trajectories are obtained if the friction force between slider and substrate decreases with increasing velocities. Stick-slip motion is observed from nanoscale junctions up to tectonic plates.

The idealized situation of Fig. 1 can be analyzed in terms of a numerical simulation. This is done by first defining a suitable Hamiltonian for the system. One can then integrate the resulting equations of motion numerically. The friction force can be calculated by averaging the force in the driving spring or alternatively from the force that is exerted from the boundary lubricant and the substrate on the upper wall. Once steady state is reached, the two methods must give the same expectation value since otherwise the top wall would be accelerated. See section 5 for a more thorough description of technical issues. The results of such a molecular dynamics simulation for a two-dimensional interface are shown in Fig. 2. Both surfaces are identical, namely triangular lattices with identical lattice constants. In one simulation, the surfaces are oriented perfectly, in another simulation, one surface is rotated by 90^o, which makes the two surfaces essentially incommensurate. Two crystals are called incommensurate if they do not share a common periodicity. Of course, in a computer simulation two solids cannot be perfectly incommensurate due to the finite system size. But if the smallest common periodicity coincides with the linear length of the simulation cell, one can usually call the surfaces incommensurate for all practical purposes.

On the left-hand side of Fig. 2, one can see that the friction forces depend sensitively on the relative orientation of the two surfaces, even though they are not in direct contact. In particular, the commensurate system shows large

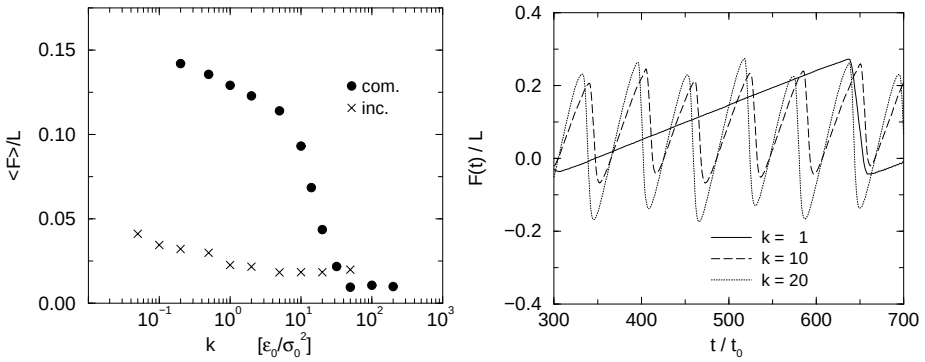

Figure 2: Left: Average force per load $\langle F \rangle / L$ acting on a block pulled at small velocity with spring constant k. Block and substrate are separated by a boundary lubricant. Commensurate (com) and incommensurate (inc) orientations between the identical, confining walls are considered. Right: Instantaneous force acting on the spring of the commensurate system. From Ref. [12].

friction forces in the stick-slip regime at small pulling spring constants k and small forces in the smooth-sliding regime at large k. In the commensurate case, the instantaneous force behaves in a very periodic way. In particular, at those values of k, where the friction forces start to decrease dramatically, one can observe spikes of two different heights in the force spikes, see the right-hand side of Fig. 2. The different spikes can be related to hcp and fcc like configurations of the boundary lubricant (ABA and ABC type layering for lower wall, lubricant, upper wall). However, one needs to be aware that insights into such detailed mechanisms are only useful, if we try to understand experiments in which the effort has been made to orient two surfaces with respect to one another.

Unlike the perfectly aligned walls, the incommensurate system does not show the dramatic drop in $\langle F \rangle$ as it crosses over from stick slip to smooth sliding. Hence, in the present study, we may identify the crossover regime for the commensurate surfaces as an artifact of the commensurability. Moreover the spikes in the instantaneous friction force (not shown here) are rather erratic [12]. To some degree, the resulting trajectory of the incommensurate case is even reminiscent of earthquake dynamics as discussed in Ref. [12]. This can be seen as unexpected, because the microscopic origins of friction between tectonic plates certainly involve much more complicated processes than those considered in the simple molecular dynamics (MD) study of a confined boundary lubricant. So even if we reproduce the effects observed in nature or experiment satisfactorily, it does not necessarily mean that we have identified the relevant atomistic processes leading to friction. The case study also reveals that the friction force derived from the simulations depends strongly on the driving device, namely the

harder the spring the smaller the friction force. This effect is particularly strong for the commensurate surfaces.

2 Solid friction versus Stokes friction

The classical laws of friction go back to Coulomb and Amontons [13]. Static friction F_s, the force needed to initiate relative sliding between two solid bodies, is proportional to the load L (first law) and independent of the (apparent) area of contact (second law), thus

$$F_s = \mu_s L. \tag{1}$$

μ_s is called the static friction coefficient that (to a good approximation) only depends on the chemo-physical properties of the interface and the two solids in contact. The third friction law says that the kinetic friction force F_k, the force needed to keep two solids in relative sliding motion, is (rather) independent of the sliding velocity v. F_k also satisfies the first two laws and in most cases does not vanish in the limit small of sliding velocities v. Of course, in true thermodynamic equilibrium, or alternatively in the mathematical limit $v \to 0$, one would expect kinetic friction to vanish. For most practical and experimental applications, however, one is far away from this limit. The situation is similar to that of the static shear modulus C_{44} of window glass at room temperature. While the true equilibrium value of C_{44} is commonly believed to be zero, the experimentally measured result is in the order of a few dozen Gigapascals.

The solid friction laws are different from those for Stokes or viscous friction which are valid for the motion of a (Brownian) particle in a fluid or a gas. In the case of Stokes friction, there is a linear relationship between driving force and average velocity v, provided that F is sufficiently small. The proportionality coefficient, which is related to the viscosity, can be calculated (at least in principle) from equilibrium statistical mechanics in terms of linear response theory [14]. The viscous force turns out to be the natural consequence of the interaction of one particle with many other particles. The (linear) response of one particle or excitation to an external force can be described as if the other particles were acting like a heat bath composed by a friction term linear in velocity plus random forces. This concept cannot only be applied to the famous Brownian motion, but it can be extended to many different cases such as the damping of phonons or other elementary excitations in solids. Note that the fluctuation-dissipation theorem also predicts the response to a small time-dependent force.

If the coupling of a degree of freedom to a heat bath is a condition for Stokes friction or viscous damping, what is the requirement for F_s and F_k to be different from zero? If two solids interlock geometrically, there will of course be a finite F_s, see section 3.1. However, finite F_s does not imply that $F_k(v \to 0)$ remains finite in the limit $v \to 0$. In particular if we disregard the internal degrees of freedom of the two solids, there is no way to dissipate energy and $F_k = 0$ simply owing to energy conservation. In a first approximation, one may reflect the internal

degrees of freedom in terms of a viscous force, but then F_k would still vanish linearly with v in contradiction to most experimental observations.

In 1928, Prandtl suggested that elastic instabilities change the picture qualitatively [15]. Usually this insight is attributed to Tomlinson, who published similar ideas in 1929 [16]. In his model, Prandtl describes the substrate as completely rigid. The slider is moved at constant velocity v with respect to the substrate. The slider's surface atom feel a force from the substrate that is periodic with the substrate's lattice constant b. Furthermore the surface atoms are supposed to be coupled harmonically to their lattice sites by a spring constant k and there is dissipation linear in the atom's velocity v_n. If we assume that the microscopic origin of the dissipation is the consequence of the interaction of atom n with a heat bath, i.e. the substrate's phonons, then its equation of motion reads:

$$m_n\ddot{x}_n + m\gamma\dot{x}_n = -k(x_n - x_n^0) + f_0 \sin(2\pi x_n/b), \tag{2}$$

where we assume that the equilibrium position x_n^0 of atom n in the slider moves at constant velocity, for example $x_n^0 = v\,t$.

Let us analyze the motion of atom n qualitatively for large values of k and small values of k. If k is larger than $2\pi f_0/b$, then it is easy to show that each atom has only one well-defined mechanical equilibrium site, irrespective of the value of x_n^0. When the upper solid is moved at a constant, small v, each atom will always be close to its unique equilibrium position. This equilibrium position moves with a velocity that is in the order of v. Hence the dissipated friction force is of the order of $m\gamma v$ and consequently F_k vanishes linearly with v as v tends to zero. The situation becomes different for $k < 2\pi f_0/b$. Atoms with more than one stable equilibrium position will now pop from one stable position to another one when the slider is moved laterally. For small pulling velocities v, such a process occurs when an atom does not have a mechanically stable position at time $t + \delta t$ in the vicinity of the old stable mechanical position near which it was located at time t. Such a situation is discussed in Fig. 3 in terms of the time-dependent potential energy $V(x)$ associated with the conservative forces, namely

$$V(x) = \frac{1}{2\pi} f_0 b \cos(2\pi x/b) + \frac{1}{2}k(x_n - x_n^0)^2 \tag{3}$$

with $\dot{x}_0 = v_0 > 0$. In the 'popping' processes (indicated by the thick solid line in Fig. 3), the velocities $\dot{x}_n$ will exceed v by orders of magnitudes for $v \to 0$. At small v, the dynamics along most of the sinusoidal line is rather independent of the precise value of v_0 and the dissipated energy $\int dx\gamma mv$ has a well-defined positive limit $F_k(v_0 = 0)$. Hence in the absence of thermal fluctuations that have been disregarded in this discussion, F_k remains finite even in the limit of infinitely small v.

Based on his model, Prandtl formulated the condition for finite F_k in the limit of small sliding velocities: *If the (elastic) coupling of the mass points is chosen*

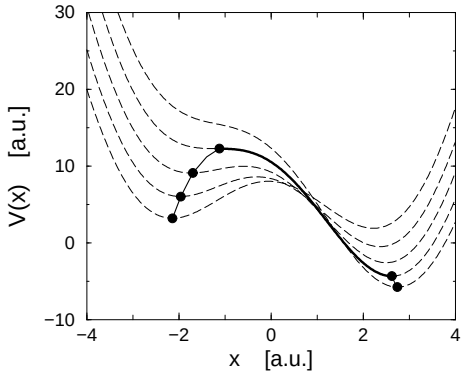

Figure 3: Schematic representation of the time evolution of the potential energy in the Prandtl model (dashed lines), see Eq. (3). All curves are equidistant in time, separated by a time interval Δt. The circles denote mechanically stable positions and the solid line connects mechanically stable points, indicating the motion of an overdamped point particle.

such that at every instance of time a fraction of the mass points possesses several stable equilibrium positions, then the system shows hysteresis.... In the context of friction, hysteresis translates to finite static friction or to a finite kinetic friction that does not vanish in the limit of small sliding velocities. The word 'elastic' in Prandtl's statement is put into parenthesis, in order to make the statement more general, for example, applicable to the case of boundary lubrication. From this analysis, we see that solid friction arises from instabilities. Hence an important lesson to be learned from computer simulations is what these instabilities on a microscopic level are and how these instabilities affect the tribological behavior of a junction. A separate issue is how the heat generated in the pops is transported away from the interface.

Incorporating thermal fluctuations changes the picture qualitatively. In the strict limit $v \to 0$ when all atoms have sufficiently much time to find their true thermal equilibrium for all relative wall positions, friction will vanish linearly with v according to linear response theory. However, in a large regime of small, but finite v, Prandtl predicted that owing to thermal fluctuations, the friction force in his model should only have small, logarithmic corrections in the order of $\ln v$. Prandtl's hypothesis was rederived many times and an analysis of experimental AFM data seemed to confirm that picture [17]. However, a more refined analysis of the thermal fluctuations in the Prandtl-Tomlinson model making use of the theory of fluctuation induced spinodal decompositions [18] suggests corrections in the order of $(k_B T \ln v)^{2/3}$, which indeed fits the experimental results distinctly better than the simple $\ln v$ corrections.

3 Dry friction

The term dry friction obtains a novel meaning in computer simulation, because one can easily prepare absolutely non-contaminated surfaces. Experimentalists often use a less strict definition and mean to express that no lubricant has been added intentionally.

3.1 Rigid walls and geometric interlocking

Early theories of friction were based on the purely geometric argument that friction is caused by interlocking of impenetrable and rigid surface asperities [19, 20]. The idea (Fig. 4) is that the top solid must be lifted up a typical slope $\tan\alpha$ determined by roughness on the bottom surface. If there is no microscopic friction between the surfaces, then the minimum force to initiate sliding is

$$F_s = \mu_s L \tag{4}$$

with $\mu_s = \tan\alpha$. This result satisfies Amontons's laws with a constant coefficient of friction $\mu_s \equiv F_s/L = \tan\alpha$. In 1737, Bélidor obtained a typical experimental value of $\mu_s \approx 0.35$ by modeling rough surfaces as spherical asperities arranged to form commensurate crystalline walls [20]. However, asperities on real surfaces do not match as well as envisioned in these models or sketched in Fig. 4. On average, for every asperity or atom going up a ramp, there is another going down. One concludes that the mean friction between rigid surfaces vanishes unless they happen to have the same periodicity and alignment.

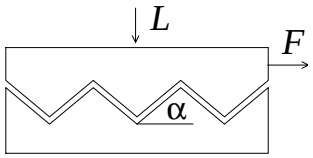

Figure 4: Sketch of two surfaces with interlocking asperities. The top surface experiences a normal load L and a lateral force F, which attempts to pull the top surface up the slope $\tan\alpha$. The bottom wall is fixed.

A cancellation between 'up' and 'down' on an atomic scale happens particularly easily between two flat, incommensurate surfaces. There have been a significant number of computer simulations showing that the *wearless* static friction becomes extremely small, except in the artificial case of identical and perfectly aligned walls. Atomistic computer simulations of iron on iron [21], a blunt pyramidal diamond tip on a metal surface [22], a blunt pyramidal copper tip on incommensurate copper substrates [23], interlayer sliding in multiwalled carbon nanotubes [24], and Lennard Jonesium on Lennard Jonesium [25] all confirm the

simple prediction that lateral forces cancel out to a significant degree. These simulations have in common that the solids were treated truly three-dimensionally and that the atoms (interacting via microscopic interaction potentials) were allowed to relax thus making long-range elastic deformations possible. Significant lateral forces were observed only in combination with strongly irreversible processes such as plastic deformation [22], the production of wear [23, 26], material mixing, or cold welding [25], but *not* due to geometric interlocking.

When two rigid solids in contact are disordered, the cancellation of lateral forces is less systematic than for an incommensurate contact between rigid bodies. For flat solids with atomic scale roughness only, one can expect that the lateral forces F_s grow proportional to the square root of the intimate area of contact A for a given, constant normal pressure σ_n. Since the load L is given by $A\sigma_n$, the ratio F_s/L vanishes with $1/\sqrt{A}$ as A goes to infinity [27]. But geometric interlocking only explains static friction. Kinetic friction can still vanish, because the energy required to lift up the slider in Fig. 4 to the top of the hill can be regained in principle by moving it downhill in a controlled way. As discussed in section 2, the situation can change dramatically in principle if the elastic interactions within the bulk are sufficiently weak to cause elastic or other instabilities.

3.2 Elastic deformations: Role of disorder and dimensions

Although we are concerned with computer simulations of friction rather than with theoretical arguments, it is instructive to analyze the interplay of disorder, dimensionality, and elastic deformations. To do this, let us consider a $d_{\rm obj}$-dimensional elastic solid, in which neighbored atoms are coupled via simple springs. We may safely assume the free elastic solid to be mechanically stable, meaning that the tensor of the elastic constants is positive definite. The dimension of the interface between the slider and disordered substrate is denoted by $d_{\rm int}$.

In such a situation, there will be a competition between the random substrate-slider interactions and the elastic coupling within the solids. An important question to ask is how the interactions change when we change the scale of the system, for example, how strong are the random and the elastic interactions on a scale $2L$ if we know their respective strengths on a scale L. Here L gives the linear dimension of the solids in all directions, that is to say parallel and normal to the interface. As discussed above, the random forces between substrate and slider will scale with the square root of the interface's size, hence the random forces scale with $L^{d_{\rm int}/2}$. The elastic forces on the other hand scale[1] with $L^{d_{\rm obj}-2}$. In the thermodynamic limit $L \to \infty$, the effect of disorder will always dominate

[1] A linear chain can be more easily compressed if we replace one spring by two springs coupled in series. In two dimensions, springs are not only coupled in series but also in parallel, so that the elastic coupling remains invariant to a 'block transformation'. Each additional dimension strengthens the effect of 'parallel' coupling.

the elastic interactions or vice-versa, *unless*

$$L^{d_{\rm int}/2} \propto L^{d_{\rm obj}-2}. \tag{5}$$

For $\lim_{L\to\infty} L^{d_{\rm int}/2}/L^{d_{\rm obj}-2} >> 1$, the random interactions will dominate and hence pinning via elastic instabilities cannot be avoided. This disorder-induced elastic pinning is then similar to that of compliant, ordered systems as discussed above within the Prandtl-Tomlinson model. For $\lim_{L\to\infty} L^{d_{\rm int}/2}/L^{d_{\rm obj}-2} << 1$, the long-range elastic forces dominate the long-range random forces. The slider's motion can only be opposed by elastic instabilities if the elastic coupling is sufficiently weak at finite L in order to make local pinning possible, again akin of the case $\lambda > 1$ in the Prandtl-Tomlinson model.

The so-called *marginal* situation, in which both contributions scale with the same exponent, $d_{\rm int}/2 = d_{\rm obj} - 2$, occurs in the important case of 3-dimensional solid bodies with 2-dimensional surfaces. In the marginal situation, the friction force can stay finite, however, one may expect the friction force per unit area and hence the friction coefficient to be exponentially small [28]. The *marginal* dimension in the case of $d_{\rm obj} = d_{\rm int}$ (adsorbed monoatomic layers, charge density waves, etc.) is $d_{\rm mar} = 4$ [29]. But even in dimensions smaller than the marginal dimension, friction forces may turn out to be small. One example is an experimental quartz crystal microbalance study [30] of solid and liquid krypton films on disordered gold surfaces ($d_{\rm obj} = d_{\rm int} = 2$), for which the pinning forces were undetectably small.

3.3 Extreme conditions and non-elastic deformations

In many cases, the atomistic topology of chemical bonds changes when two solids come into intimate contact and are start to slide. As a consequence, the surface will be altered dramatically when the two solids are removed from one another after the sliding process. A description in terms of *elastic* deformations is not applicable any longer in such a situation. There can be many reasons for non-elastic deformations: (i) plastic flow during contact formation [26, 31] or other thermodynamically driven cold welding [25], (ii) plastic deformations due to large normal pressures [22], (iii) large sliding velocities [32], and (iv) sliding induced generation of dislocations [23], to name a few.

The processes that occur in these strongly non-equilibrium situations may become rather complex. This makes it even more important to set up the simulations and to analyze the processes in a meaningful way. For example, it is important to choose the initial conditions such that the bulks have room to yield and that the generated debris does not necessarily remain within the contact. A simulation methodology which allows debris to be transported away from the interface is shown in Fig. 6 in the context of sliding of lubricated surfaces.

One particularly nice study of strongly irreversible processes consists of large scale molecular dynamics simulations of the indentation and scraping of metal by

Belak and Stowers [22]. Their simulations show that tribological properties are strongly affected by wear or the generation of debris, which in turn may again depend strongly on such 'details' as the system's and the interface's dimensionality. Belak and Stowers considered a blunted carbon tip that was first indented into a copper surface and then pulled over the surface. Since diamond is a hard material, the tip was treated as a rigid unit. Interactions within the metal were modeled with an embedded atom potential and Lennard-Jones potentials were used between Si and Cu atoms.

In the two-dimensional (2D) simulation, indentation was performed at a constant velocity of about 1 m/s. The contact followed Hertzian behavior up to a load $L \approx 2.7$ nN and an indentation of about 3.5 Cu layers. The surface then yielded on one side of the tip, through the creation of a single dislocation edge on one of the easy slip planes. The load needed to continue indenting decreased slightly until an indentation of about five layers. Then the load began to rise again as stress built up on the side that had not yet yielded. After an indentation of about six layers, this side yielded, and further indentation could be achieved without a considerable increase in load. The hardness, defined as the ratio of load to contact length (area), slightly decreased with increasing load once plastic deformation had occurred.

After indentation was completed, the carbon tip was slid parallel to the original Cu surface. The work to scrape off material was determined as a function of the tip radius. A power law dependence was found at small tip radii that did not correspond to experimental findings for micro-scraping. However, at large tip radii, the power law approached the experimental value. Belak and Stowers found that this change in power law was due to a change in the mechanism of plastic deformation from intragranular to intergranular plastic deformation.

In the three-dimensional (3D) simulations, the substrate contained as many as 36 layers or 72,576 atoms. Hence long-range elastic deformations were included. The surface yielded plastically after an indentation of only 1.5 layers, through the creation of a small dislocation loop. The accompanying release of load was much bigger than in 2D. Further indentation to about 6.5 layers produced several of these loading-unloading events. When the tip was pulled out of the substrate, both elastic and plastic recovery was observed. Surprisingly, the plastic deformation in the 3D studies was confined to a region within a few lattice spacings of the tip, while dislocations spread several hundred lattice spacings in the 2D simulations. Belak and Stowers concluded that dislocations were not a very efficient mechanism for accommodating strain at the nanometer length scale in 3D.

When the tip was slid laterally at $v = 100$m/s during indentation, the friction or 'cutting' force fluctuated around zero as long as the substrate did not yield (Fig. 5). This nearly frictionless sliding can be attributed to the fact that the surfaces were incommensurate and the adhesive force was too small to induce locking. Once plastic deformation occurred, the cutting force increased dramatically. Fig. 5 shows that the lateral and normal forces are comparable, implying

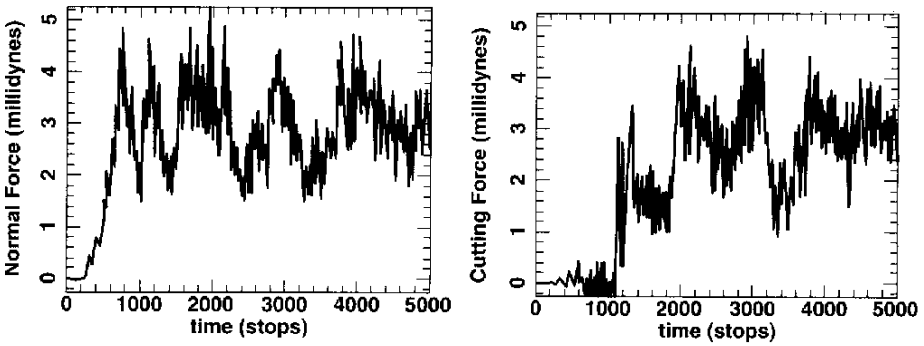

Figure 5: Normal (left) and lateral (right) force on a three dimensional, pyramidal diamond tip on a copper surface as a function of time. No plastic flow was reported up to 1,000 time steps. The indentation stopped at about 5 layers after 2,000 time steps. From Ref. [22].

a friction coefficient of about one. This large value was expected for cutting by a conical asperity with small adhesive forces [37].

4 Lubrication

Any atom or molecule that is bonded weakly to a surface can be considered to perform as a lubricant as long as they do not get squeezed out of the microscopic points of contact. Examples are not only synthetic and mineral oils, but also molecules from the atmosphere physisorbed between the surfaces such as hydrocarbon chains or sometimes even simple nitrogen molecules. Lubricants hinder two surfaces to come into intimate contact. Their presence inhibits or at least reduces the generation of wear and debris. For example, an appropriate additive in a lubricant reacts with a fresh metal surface to form protective surface films. In many cases, a lubricant film glassifies in a point of large normal stress, i.e., in a microscopic point of contact. When the junction breaks, the contact breaks within the lubricant and not within one of the contacting asperities.

The traditional view is that lubricants do not only reduce wear but generally reduce friction between two solids. This is certainly true for most macrotribological processes. However, as we have seen in many examples discussed above, we do not expect any significant solid friction in a microscopic contact as long as no plastic deformation or other strongly irreversible processes occur. In the absence of such processes, the presence of a few adsorbed lubricant particles can increase solid friction, as it is able to accommodate the surface corrugation of both walls simultaneously and thus lock the confining walls together.

The overwhelming majority of experimental and technological sliding systems

incorporates lubricants. It is important to include the effect of these lubricants in simulations if one wants to compare simulations to the bulk of tribological experiments. It is then desirable to mimic the experimental, mechanical and thermodynamic boundary conditions such as temperature T, pressure p, chemical potentials μ_i, etc. as closely as possible. One should of course keep in mind the Gibbs-Duhem relation, which states that it is not possible to specify all intensive thermodynamic variables (such as T, p, the various μ_i) independently from one another.

In a molecular dynamics (MD) or in a Monte Carlo (MC) simulation it is rather simple to keep T and p constant. It is more difficult to keep μ constant, because this requires the use of grand canonical moves. These moves typically equilibrate rather slowly and are likely to interfere seriously with the dynamics in MD or the pseudo-dynamics in MC simulations[2]. Since most tribological effects are non-equilibrium effects and therefore intimately linked to the dynamics of the system rather than to their thermodynamics, it is important not to alter the dynamics in an artificial way.

In some cases, it is nevertheless desirable to change the number of atoms in the contact. It was suggested to include this possibility with the help of reservoirs [33] as shown on the left-hand side of Fig. 6. While the total particle number N is kept fixed in such a simulation, the number of atoms in the contact can change and one may refer to such a simulation as pseudo grand canonical. The externally applied pressure (tensor) and temperature can then be imposed like in an equilibrium simulation, for instance in such a way that the lubricant remains fluid in the reservoir.

The simulation of the reservoir necessitates simulating many bulk-liquid molecules outside the real contact, producing a certain computational overhead. It was suggested to replace the need for a reservoir by fixing T and the parallel pressure $p_\parallel$ to bulk values [34]. It was furthermore suggested to control $p_\parallel$ by adjusting the normal separation D between the solids and to leave the contact area A constant. The strategy would be to repeat simulations with different values of N, all under the same fixed A, $p_\parallel$, and T. One of the results would be the depletion force, which is the average normal pressure $\langle p_\perp \rangle$ times A, as a function of N similar to that shown on the left-hand side of Fig. 6. Note that the differentiation between $p_\parallel$ and $p_\perp$ implies that the lubricant does not correspond to an isotropic fluid any longer. Indeed, the oscillations in the depletion forces are accompanied by strong layering in the lubricant. In order to understand the concurrent dramatic increase in the resistance to sliding observed in experiments [8, 9], one also has to investigate the effect of confinement on the viscosity [35] and the way in which the (atomic-scale) corrugations imprint into the lubricant [36].

[2]If the MC simulation consists of small local moves only, then the generated trajectories correspond to overdamped dynamics that may provide valuable information on solid friction. However, there is no general principle for determining the appropriate probability distribution of steps in a non-equilibrium MC simulation.

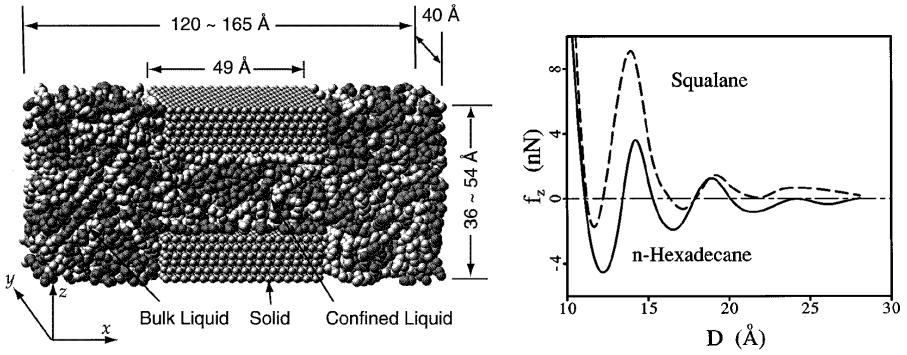

Figure 6: Left: Geometry of a pseudo grand canonical simulation box. Small, regularly arranged spheres represent gold atoms from the solid bulks, while larger circles are associated with lubricant molecules. The lubricant can go back and forth between contact and reservoir. Right: Normal force as a function of separation for two different lubricants as a function of surface separation. From Ref. [33].

4.1 Boundary lubrication

Boundary lubrication refers to a situation in which most of the lubricant has been squeezed out of the contact and the remaining lubricant glassifies. It usually occurs in a mechanical contact under high load and low speed conditions. In the extreme limit, only a few atoms remain in the contact. But even a few atoms alter dramatically the friction between two surfaces.

Since in most cases lubricants are only weakly bonded to surfaces, the most commonly used form for the lubricant-lubricant and lubricant-wall interactions are Lennard Jones (LJ) potentials. LJ potentials reflect the important effect that atoms attract when separated by a sufficiently large distance, but repel and behave like hard disks in the presence of a large external pressure.

It has been shown in a series of molecular dynamics simulations [27, 38, 39, 40] that the presence of a small contamination layer even as thin as a sub-monolayer leads to static and kinetic friction between two incommensurate surfaces. Within the model calculations the two walls would have had zero static friction if no weakly bonded molecules had been introduced into the interface. As argued above, the weakly bonded atoms are able to accommodate the surface roughness of both confining walls simultaneously, which makes the walls lock together. Of course, in order to obtain not only static friction but also non-zero kinetic friction, it is necessary for the atoms to exhibit mechanical instabilities, see section 2. Whether or not the atoms experience instabilities may depend on the special properties of the system such as dimensionality and symmetry of the

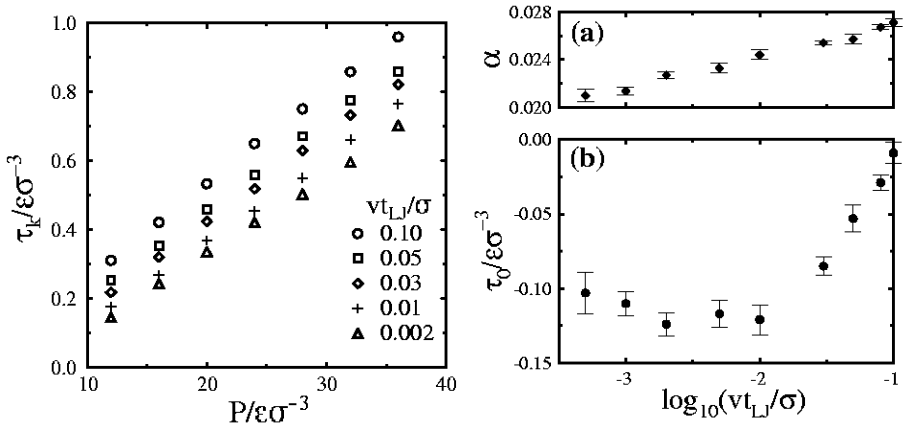

Figure 7: Left: Kinetic shear stress τ_k as a function of normal pressure for different velocities. The data of data like that on the left was fitted to Eq. (6). Right: The resulting fit parameters α (top) τ_0 (bottom). From Ref. [40]

confining walls and their relative orientation, the details of the lubricant wall interactions such as the sign of the first higher harmonic, and further details. Fig. 2 suggests that there are more instabilities present for incommensurate surfaces than for commensurate surfaces in the smooth sliding regime where the pulling spring is hard, while in the stick-slip regime (small values of k) the upper wall in the commensurate case experiences instabilities as a whole owing to geometric interlocking.

The microscopic friction-load relation for boundary-lubricated flat surfaces turns out to be similar to that of Amontons's macroscopic friction law $F = \mu L$, or a simple generalization thereof, namely

$$\tau_k = \tau_0 + \alpha p_\perp, \tag{6}$$

where $\tau_k = F_k/A$ is the (kinetic) shear pressure, τ_0 an adhesive offset, and α a constant that can be associated with the friction coefficient μ if the externally imposed normal pressure $p_\perp = L/A$ is distinctly larger than τ_0. A nearly linear relation between τ_s and p can be observed over a wide range of pressures, which is shown exemplarily on the left-hand side of Fig. 7 for the case of smooth sliding.

It is interesting to analyze the velocity dependence of the coefficients τ_0 and α, which is shown on the right-hand side of Fig. 7. In agreement with Coulomb's observation, the kinetic friction force F_k is barely sensitive to the sliding velocity v: The parameters α and τ_0 only vary logarithmically with v. The overall decrease of τ_k with v can be associated with thermal activation and diffusion of the lubricant atoms out of their metastable traps [40], confirming previous interpretations of simple rate-state models of friction [41].

4.2 Hydrodynamic lubrication and its breakdown

When the contact geometry and the operating conditions are such that the load is fully supported by a fluid film, the surfaces are completely separated. This is generally referred to as the hydrodynamic lubrication and can be well described in most cases by Reynold's equation. When the surfaces are easily deformable, as in rolling contact bearings or human and animal joints, the equations of elasticity and the pressure dependence of lubricant dependence must also be included in the solution of the problem [42]. The hydrodynamic or elastohydrodynamic continuum theory begins to break down as atomic structure becomes important.

A general assumption of continuum theories is the stick condition of the lubricant near a solid wall, that is to say, the tangential velocity of the fluid at the fluid wall interface is set equal to that of the wall. When two surfaces come closer and the confinement is increased, slip can occur at the interface. It is then convenient to introduce a slip length $\mathcal{S}$ into elastohydrodynamic continuum calculations. $\mathcal{S}$ represents the distance into the wall at which the velocity gradient would extrapolate to zero. The calculation of slip length from velocity profiles has some ambiguity. The least ambiguous resolution to the definition of slip length may be given in Ref. [43]. Additional effects due to increasing confinement have been discussed at the beginning of this section 4. A more thorough discussion of the literature on atomistic simulations in the hydroelastic lubrication regime is given in Ref. [10]. Some technical issues relevant to the thermostatization of lubricants will also be given below in section 5.5.

5 Setting up a tribological simulation

5.1 The essential ingredients

When designing a computer simulation of tribological phenomena, one needs to model (i) the physo-chemical properties of the two materials in contact and the lubricant or the atmosphere involved, (ii) the initial conditions/geometry and (iii) in addition to what needs to be specified in an equilibrium simulation, the driving device. Let us assume that we do have a model for the interactions between the atoms involved in the simulation. We then need to set up Newton's equations of motion for the various degrees of freedom and integrate these equations of motion just like in a computer simulation of an equilibrium system [14, 44, 45].

The relevant degrees of freedom are: The center of mass of the confining top and bottom solid $\mathbf{R}_{t,b}$, the coordinates $\mathbf{r}$ of those atoms which are coupled directly to the confining solids (typically the atoms in the outermost layers), and the coordinates $\mathbf{x}$ of all other atoms, which may include lubricant atoms and additional wall atoms that do not belong to the outermost layers.

A quite general form to treat the driving device is to keep the substrate fixed ($\mathbf{R}_b = \text{const}$) and to couple the center-of-mass coordinate $\mathbf{R}_t$ of the upper solid

to an externally driven harmonic spring as shown schematically in Fig. 1. The equation of motion for the top wall then reads:

$$M_{\mathrm{t}}\ddot{\mathbf{R}}_t = K\left(\mathbf{R}_t^0 - \mathbf{R}_t\right) + \mathbf{F}_t\left(\mathbf{R}_t,\,\{\mathbf{r}\}\right). \tag{7}$$

Here M_{t} denotes the inertia of top solid, K is the stiffness of the spring that couples the solid to the driving device $\mathbf{R}_t^0$. In most cases, K is a diagonal matrix, with three independent components that reflect the normal coupling, and the coupling parallel to the two interfacial directions. Two popular choices for $\mathbf{R}_t^0$ are

$$\mathbf{R}_t^0 \;=\; R_{t,y}^0\mathbf{e}_y + R_{t,z}^0\mathbf{e}_z \;+\; \begin{cases} v_x\mathbf{e}_x\,t & \text{'tribological' driving} \\ \Delta R_{t,x}^0\mathbf{e}_x\cos(\omega t) & \text{'rheological' driving} \end{cases} \tag{8}$$

with $\mathbf{e}_\alpha$ a vector of dimension unity pointing in α direction. Here, v_x corresponds to the average sliding velocity of the upper wall if the spring (see Fig. 1) is pulled at constant velocity, and $\Delta R_{t,x}^0$ and ω are the amplitude and frequency of the driving device if the response of the system to oscillations is probed. Despite the simplicity of equations (7) and (8), it is important to realize the implications that various choices can have. This issue is discussed further below in section 5.4.

We still need to specify the coupling $\mathbf{F}_t\left(\mathbf{R}_t, \{\mathbf{r}\}\right)$ between the coordinates of the top solid $\mathbf{R}_t$ and the atoms belonging to the uppermost layer. The discussion is of course equivalent for the bottom wall, even though one usually does not move its position $\mathbf{R}_b$. It is convenient to define equilibrium positions $\mathbf{r}_n^0$ of atom n in the uppermost layer relative to the top solid under the condition that no other atoms are present in the simulation. The equilibrium positions $\mathbf{r}_n^0$ will then always have the same relative displacement $\Delta\mathbf{r}_n^0$ with respect to the top solid, thus

$$\mathbf{r}_n^0(t) = \mathbf{R}_t(t) + \Delta\mathbf{r}_n^0, \tag{9}$$

where periodic boundary conditions are usually employed only in directions parallel to the interface. The real position of atom n will then couple to its equilibrium position. To lowest order this coupling can be considered harmonic and the force that the *ideal* lattice point exerts on an atom of the outermost layer can be written as

$$\mathbf{f}_{t,n} = -k\left[\mathbf{r}_n(t) - \mathbf{r}_n^0(t)\right]. \tag{10}$$

k is again a matrix in the most general case and the actual choice of k is in principle highly non-trivial. To a large extent, k reflects the elastic properties of the top solid, see the discussion of the choice for k in section 5.6. Since $\mathbf{f}_{t,n}$ is the force that the equilibrium site and hence the top solid exerts on atom n, we can finally write down the expression for the last term on the right-hand side of Eq. (7):

$$\mathbf{F}_t\left(\mathbf{R}_t, \{\mathbf{r}\}\right) = -\sum_n \mathbf{f}_{t,n}. \tag{11}$$

As a next step, one has to specify the interactions between atoms within the uppermost layer. A reasonable choice is to do this is again within the harmonic approximation, thus

$$V_{nn'} = \frac{1}{2} \sum_{n<n'} \sum_{\alpha,\beta} k_{nn'}^{\alpha\beta} \left(r_n^\alpha - r_n^{\prime\alpha}\right) \left(r_n^\beta - r_n^{\prime\beta}\right), \tag{12}$$

with r_n^α being the α's component of the vector $\mathbf{r}_n$. Note that the force constants $k_{nn'}^{\alpha\beta}$ do not only approximate the direct interactions between the atoms in the uppermost layer, but they are effective quantities which may also contain information of the elastic and geometric properties of the top solid. A more detailed discussion is given in section 5.6.

Lastly, the particles in the outermost layer must also be coupled to the remaining atoms in the simulation cell, for instance to lubricant atoms or those wall atoms that are not coupled to specific lattice sites. If the forces between wall atoms with coordinate $\mathbf{r}_n$ and remaining atoms with coordinates $\mathbf{x}_k$ are given by $\mathbf{f}_{kn}(\mathbf{r}_k, \mathbf{x}_n)$, then the net force $\mathbf{f}_n$ on atom n in an outermost layer is given by

$$\mathbf{f}_n = \mathbf{f}_{t,n} + \sum_{n' \neq n} \sum_{\alpha,\beta} \mathbf{e}^\alpha k_{nn'}^{\alpha\beta} \left(r_n^\beta - r_{n'}^\beta\right) + \sum_k \mathbf{f}_{nk}(\mathbf{r}_n, \mathbf{x}_k), \tag{13}$$

where the sum in the last term on the right-hand-side of Eq. (13) includes all atoms that do not belong to the uppermost layer. All other interactions, i.e., those in between lubricant atoms or between slider and substrate atoms, should be treated just like in equilibrium simulations of materials [44].

5.2 Physo-chemical properties

The physo-chemical properties of the interface are reflected through the choice of the atomic model potentials. One needs to specify the *intra*bulk and the *inter*bulk atomic interaction potentials as well as the interaction of a lubricant or another adsorbed atom with all other atoms. *Intra*bulk properties are often modeled as entirely elastic, i.e., atoms belonging to one solid are connected to each other and/or to their ideal lattice site via simple harmonic springs. Such simplifying modeling will be sufficient if plastic deformation, cold welding, etc. do not play a significant role in the processes of interest.

Many studies use simple Lennard Jones (LJ) potentials for the interbulk interactions and the interactions between lubricant and walls. LJ interactions are given by

$$V(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right], \tag{14}$$

where ϵ_{ij} and σ_{ij} have units of energy and length, respectively. Typically, the interactions are truncated at a distance $r_{ij}^{(c)}$ and shifted in such a way that $V(r_{ij})$

is continuous at $r_{ij}^{(c)}$. There is a lot one can learn from systems modeled in this way, since LJ potentials are sufficient to describe the generic feature that two atoms attract each other when they are separated by a long distance, while they repel upon close approach. Moreover, there is a lot of flexibility in the choice of the parameters ϵ_{ij}, σ_{ij}, and also the radius $r_{ij}^{(c)}$. The effect of adhesive interactions can be switched on or switched off depending on the choice of $r_{ij}^{(c)}$. A typical choice for $r_{ij}^{(c)}$ for which adhesive effects are eliminated is the cutoff in the minimum of $V(r_{ij})$, hence $r_{ij}^{(c)} = 2^{1/6}\,\sigma_{ij}$.

The use of LJ potentials is also widespread in coarse grained models of polymers. Simulations of friction between polymer bearing surfaces, i.e., simulations of shear forces between polymer brushes, are done in terms of bead-spring models [46, 47, 48] In bead-spring models, all interactions between polymer segments consist of LJ interactions and additional non-linear potentials keep the segments bound together [49]. Note that the effect of a good or a bad solvent can be included into such a coarse grained model via the choice of cutoff radius $r^{(c)}$ and temperature T, making it possible to access distinctly larger length and time scales than if all atoms had been included explicitly into the simulation.

Often, one tries to understand the friction between two specific solids. In such a case, LJ potentials are usually not sufficient and more realistic descriptions are required just like in an equilibrium simulation. Many materials show complicated surface reconstruction after cleavage, which often alters surface properties and hence friction dramatically. This concerns metals, whose surface cannot be described accurately in terms of simple two-body potentials. A popular choice for the simulation of metals are so-called embedded atom potentials. Usually, layered materials and carbon nanotubes cannot be modeled with simple two-body potentials either and more realistic description are required, see Refs. [44, 50] and references therein for a more detailed overview of empirical many-body potentials.

5.3 Initial geometry

The frictional properties of a slider-substrate system do not only depend on the chemical nature of the two solids, the lubricant, and the thermodynamic conditions such as temperature, normal pressure, etc., but also on the way in which the system is initially set up. Most simulations of friction take place between two flat surfaces, as shown on the left-hand side of Fig. 8, while most well-controlled experimental single asperity contacts employ curved tips, either an AFM tip scratching over a smooth substrate or two crossed mica cylinders in SFA experiments. A geometry with similar contact mechanics as those occurring in AFM or SFA experiments is shown on the right-hand side of Fig. 8.

There are many processes that can differ qualitatively as a function of the initial geometry, because of the differences in the distribution of the normal stress in the contact. One obvious example is the squeeze-out dynamics of a

Figure 8: Schematic representation of initial geometries. Left: Flat surfaces. Middle: Blunt tip on substrate. Right: Curved tip on substrate.

fluid lubricant when the asperity approaches the substrate [52]. But also the dry friction depends strongly on the geometry. For flat, disordered surfaces, there is a well-defined friction coefficient μ_s, which however, decreases with the area of contact A. This is shown on the left-hand side of Fig. 9. In the case of a dry, curved tip, F_s depends only linearly on L if tip and substrate are commensurate, which is shown on the right-hand side of Fig. 9. The presence of a few contaminating or lubricating atoms changes these dependencies again.

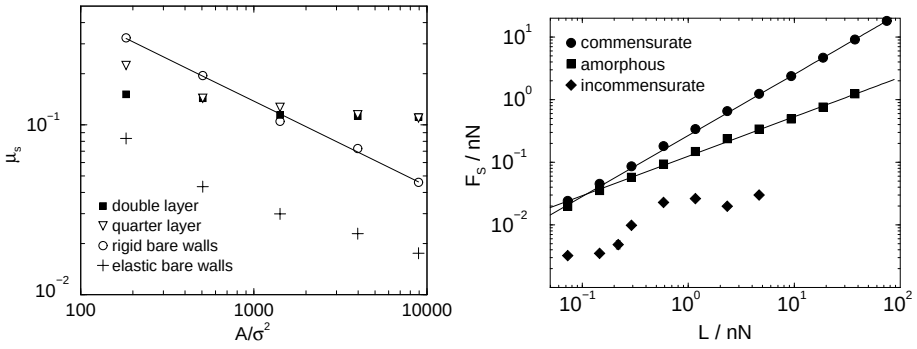

Figure 9: Left: Static friction coefficient $\mu_s = F_s/L$ as a function of interface size A for flat, amorphous walls with different degrees of contamination. From Ref. [27] Right: F_s as a function of load L for a curved tip (commensurate, amorphous, and incommensurate) on a crystalline substrate. Right: From Ref. [51].

Only the quartz crystal microbalance (QCM) [53], which enables one to measure the viscous friction between an adsorbate layer and a smooth crystalline surface does not incorporate curved surfaces. However the simple scaling arguments presented in section 3.2 show that the processes and the potential instabilities occurring in QCM experiments are likely to be different from those occurring between two three-dimensional solids.

Before setting up the simulation it is advisable to consider whether surface curvature plays an important role in the process of interest. However, one needs to be aware that the computationally feasible radii of curvature R_c are distinctly

smaller than experimental R_c. Employing curved surfaces also reduces the real contact and thus produces computational overhead outside of the contact. Moreover it would be advisable to do simulations for different R_c and to test whether scaling arguments to extrapolate to larger R_c can be employed.

5.4 Driving device

There are many ways in which the solids can be coupled to a driving device. A rather generic way is to couple the upper wall to a spring that is moved according to a well-described procedure, i.e. rheologically or tribologically. The scenario is visualized in Fig. 1 and described in more detail in between Eqs. (7) and (11). The use of one simple spring K is a simplistic way to model not only the experimental driving device, but to some degree also the elasticity of the slider.

There are two important limits for the value of K. One limit is to use an infinitely weak spring $K \to 0$ and to set the (hypothetical) position of the driving device a distance ΔZ away from the actual position of the top wall such that $K\Delta Z = L$. This mode is equivalent to impose a constant force and is frequently used to drive the system in the direction normal to the interface, in which case L corresponds to the load. It was shown that small normal spring constants K result in smaller friction forces than if the system was driven with large K, at a given average load L [54]. The reason is that the slider has more possibilities to cross energy-barriers if the spring is weak and hence the slider chooses the path of minimum resistance. Of course, in terms of implementing the condition normal load, one would replace the first term of the right-hand side of Eq. 7 simply with the externally imposed load L. The other limit is $K \to \infty$, in which case the position of the top wall $\mathbf{R}_t$ is identical with that of the driving device $\mathbf{R}_t^0$. For tribological driving, this typically implies a constant sliding velocity mode for the lateral motion and in case of zero normal velocity a constant separation constraint. The effects of the lateral spring constant on the average friction has already been discussed in the simple case study of section 1.1.

The driving device can have different modes in different directions. A natural choice would be to apply a constant load mode in the normal z-direction, while the slider is pulled with a weak spring parallel to the x-axis, and coupled to another non-moving spring in y direction.

5.5 Thermostating

Most experiments are conducted in such a way that, far away from the sliding surfaces, constant temperature is maintained. The heat produced in the sliding process is conducted away from the interface via phonons and in the case of metals heat dissipation also occurs via electrons. One can mimic the effect of heat conduction in a simulation by employing thermostats or heat baths similar to those used in equilibrium simulations. However, there are a few additional difficulties in a non-equilibrium simulation involving sliding surfaces.

(i) There is no well-defined frame of reference. (ii) The driving device does work on the system. The final thermal energy or temperature and thus the system's properties will depend on the rate with which energy is pulled out of the system. In equilibrium simulations, static properties including thermal energy do not depend at all on the strength of that coupling and dynamical properties should depend only little on the coupling to the heat bath.

These three difficulties can be easily overcome if certain rules are respected. (i) Only the outermost layers should be thermostated and thermostatization should take place within the frame of reference defined by the motion of $\mathbf{R}_t$. Alternatively, dissipative particle dynamics (DPD) [55] may be used to thermostat locally in the center-of-mass system of two neighbored wall atoms. DPD ensures that the Gibbs distribution is recovered as the stationary solution to the Fokker-Planck equation. As compared to regular Langevin-type thermostats, DPD is more difficult to implement and slightly more CPU time intensive. However, it has the distinct advantage to act only on a local scale. In some cases, one may also want to use a DPD thermostats within the lubricant, for example to mimic the effect of collisions with solvent atoms that are not simulated explicitly, but only taken into account via effective interactions. If implemented correctly, Navier-Stokes coefficients are recovered in the hydrodynamic limit [56]. (ii) In extreme conditions, the effects of heating cannot be neglected any longer. One possibility is to define the thermostat's coupling strength such that the heat flow into the thermostat corresponds approximately to that which one would have if the system was made infinitely large normal to the interface (size consistency).

5.6 Methods to treat the wall's elasticity

From a computational point of view, we do not want to spend most of the CPU time with the simulation of the bulk in order to reproduce the proper elastic behavior. However, the discussion above shows that a proper description of elastic effects may be crucial. If we only couple the atoms to the equilibrium sites, we suppress elastic deformations and as a consequence unrealistic pressure profiles in the contact may be obtained. On the other hand, we do not want to neglect the effect of long-range elastic interactions by simply connecting the surface atoms with effective springs. This would favor elastic long-range deformations in an unrealistic way and thus make possible elastic instabilities.

It is possible to reproduce the bulk's proper static, elastic response by a suitable choice of the coefficients k in Eq. (7) and $k_{nn'}^{\alpha\beta}$ in Eq. (12), because all harmonic modes can be integrated out in principle. The treatment can even be generalized to a dynamic response in terms of Green's functions [57]. Such sophisticated treatments, however, are very demanding in several aspects, and simplifying procedures are pursued in practice. The easiest approach is to treat the bulk's elasticity in a mean-field approximation, in which case one only needs to define the coupling strength k of a wall atom to its ideal lattice site. From the scaling arguments in section 3.2, one may withdraw that k would only have

a well-defined value in the limit of the slider's linear system size L, if the spatial dimension $d_{\rm obj}$ of the slider is larger than two. In $d = 2$, the proper mean-field choice for k would be $k \propto \ln L$.

A good compromise is to use elastic coupling to the ideal lattice site and next neighbor coupling. Such a model is not only an interesting generalization of the Prandtl-Tomlinson model for the analysis of dry friction [58], but also useful for computer simulations involving lubricants, because the lubricant gets squeezed less easily through the confining walls. For flat surfaces, one may use simple harmonic springs. If one of the two surfaces is curved as shown in Fig. 10, then the harmonic springs have to be replaced with anharmonic interactions: Within the radius of contact r_c, the normal deflection $\delta z(r)$ of the tip atoms with respect to an uncompressed tip is proportional to r^2, where r represents the lateral distance from the center of the tip. In order to reproduce the proper Hertzian contact profile $p_\perp \propto (r - r_c)^{3/2}$ for a tip pressed against a hard but non-adhesive substrate, the normal restoring force $f_\perp$ to the lattice site must be chosen according to [51]

$$f_\perp(\delta z) = \sqrt{\delta z / R_{\rm c}} A_a / K, \tag{15}$$

where $R_{\rm c}$ is the tip's radius of curvature and A_a the surface area per atom represented in the surface. Note that this procedure does not impose Hertzian contact mechanics, but the final pressure distribution depends among other things on the adhesive interaction between tip and substrate.

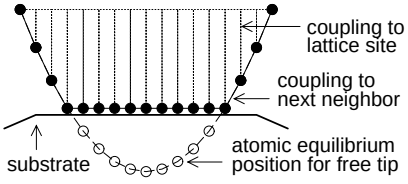

Figure 10: Schematic representation of elastic interactions within a soft tip pressed on a hard surface. Both couplings can have one component normal to the interface and two transversal components.

5.7 Calculation of the friction force

There are various ways to calculate the kinetic friction force. Using the notation introduced in section 5.1, one can measure the force that the external driving device exerts on the wall atoms, i.e., $K(R^0_{t,x} - R_{t,0})$ with x the sliding direction, or alternatively the total force that the top-wall atoms exert on their equilibrium positions, i.e. the projection of $\sum_n \mathbf{f}_{t,n}$, see Eq. (10), onto the sliding direction. We may also sum up the force between top-wall atoms and all other atoms to obtain the friction force. The different ways should be identical since otherwise the

top wall would be accelerated. The argument for the equivalence of the different ways to calculate the friction force breaks down when the surface experiences a time-dependent external force as is the case in the rheological driving mode. This issue has been discussed in detail in Ref. [11] within the Prandtl-Tomlinson model.

In the absence of thermal fluctuations, static friction F_s is defined as the maximum external shear force necessary to invoke lateral sliding. In order to calculate F_s, one can ramp up an external force arbitrarily slowly $F_{\rm ext}(t) = \dot{F}_{\rm ext} t$ and identify the time $t_{\rm dep}$ when the system depins and starts sliding. F_s can then be associated with $F_{\rm ext}(t_{\rm dep})$. At finite temperature, the precise value of F_s will depend on the rate $\dot{F}_{\rm ext}$. Near the depinning threshold, thermal fluctuations will assist the system to overcome the barrier and small corrections to F_s, approximately in the order of $\ln \dot{F}_{\rm ext}$, will presumably apply, similar to the kinetic friction in the Prandtl-Tomlinson model or boundary lubrication [40].

5.8 Interpretation of time scales and velocities

The interpretation of the length and time scales in a computer simulation becomes important when comparison is made to experiment. The length scale is defined in a computer simulation through the size of an atom with a typical radius of about 1.5 Å. The definition of the time scale is much less clear cut. According to the international system of units, a second is defined as the duration of $9.192\,10^9$ and a few periods of a particular vibration of the isotope ^{133}Cs. However, the physics of this vibration is completely irrelevant in a tribological context and a different definition of time scales will be more appropriate. This definition will depend on the particular problem. In the following two exemplary discussions, it will be assumed implicitly that there is a time-scale separation between the vibronic motion of the atoms and all processes requiring thermal activation.

Let us first consider the case of dry friction, i.e., an AFM tip in contact with a substrate. Eq. (2) is then frequently used to model phenomenologically the results of a real experiment or of a computer simulation. If we carry out a simulation of a specific experiment and use realistic model potentials, then we will withdraw a similar value for the tip-substrate interaction strength f_0 from the simulation as from experiment. We might also be successful in reproducing the rate γ with which energy is dissipated and the (effective) coupling strength k between tip and driving device. However, in almost all cases, we will not assign the experimental device/tip mass $m_{\rm exp}$ to our top wall, but a mass $m_{\rm sim}$ that is many orders of magnitude smaller. This means that the relevant frequencies defined either via $\sqrt{k/m}$ or via $\sqrt{f_0/m}$ are also many orders of magnitude smaller in experiment than in the simulations. Thus, in order to compare to experiment our velocities need to be scaled by a factor $\sqrt{m_{\rm exp}/m_{\rm sim}}$ and comparing absolute velocities would be meaningless.

Let us next consider the case of two boundary lubricated surfaces under

shear. The film confined between two surfaces is interpreted as a glassy state, which is able to accommodate the atomistic surface corrugation of both walls simultaneously. After a typical relaxation time τ, a lubricant atom will undergo thermally activated motion, which eventually leads to lateral creep motion of the slider. In correspondence to the concept of the Deborah number D, which is defined as time of relaxation τ divided by the time of observation Δt, one may expect the friction force to depend on the ratio of τ and the time Δt necessary for the slider to move laterally a typical atomic length scale $\Delta l = \Delta t/v$. A distinct advantage of computer simulation is that one can alter the Deborah number at will either by varying τ via the normal pressure or via a change in sliding velocity v.

6 Conclusions

This contribution is meant as a guideline for setting up a meaningful simulation of frictional processes. To do this successfully, it is helpful to have both a good understanding of the underlying theoretical approaches and the knowledge of recent advances in numerical algorithms.

Understanding the underlying theoretical ideas is important in order to prevent the simulations to be meaningless when we want to compare our results to experiments. After all, the goal is to interprete and to understand experiments. Since it is unfeasible to treat all relevant degrees of freedom in a simulation, we are forced to make a model before we carry out the simulation. In my opinion, there are two classes of mistakes that one can make, one class is due to reductional modeling while the other class is due to fictitious modeling. Reductional modeling means that we have not captured all aspects of the system, but what we simulate is at least part of the fuller picture. Fictitious modeling includes a (friction) mechanism that is not relevant in experiment but arises as a consequence of our model. For instance if the walls are described as elastic media although plastic deformation can occur and be a relevant contribution to the net friction, then we see a reduced part of the full picture. An example for fictitious modeling would be a simulation in which long-range elastic interactions are neglected and elastic instabilities occur, although they might only play a minor role in the real system and not contribute to friction. The following remaining compilation of recommendations summarize many aspects of this chapter.

The cardinal and frequently encountered mistake is the use of commensurate surfaces. Even if there is 'stuff' between the surfaces, there will always remain some finite (artificial) resistance to the initiation of sliding. The reason is that commensurability breaks translational invariance in a very specific way. It is certainly true that the effect often becomes negligibly small if the distance between the walls is sufficiently large. It has yet been shown that not only fluids but even a gas can pin two commensurate solids even if they do not interact directly with each other [39]. Similar comments apply to the simulation of friction between

walls that have dense Langmuir-Blodgett type layers grafted onto them. The atomistic processes occurring between commensurate layers in simulations will be strikingly different from those between non-matching surfaces in experiment.

The presence of a few adsorbed atoms on surfaces most always also alters the tribological behavior qualitatively. This concerns in particular reactive and reconstructing surfaces that have much higher friction in UHV than in ambient conditions. In the other extreme, when two (flat) surfaces in contact are believed not to deform irreversibly at all, then sliding in UHV will mainly be opposed by a drag force. In this case, an adsorbed layer will be responsible for a dramatic increase in solid friction.

Other issues discussed in this chapter include the effect of wall curvature and initial geometry, the effects of long-range elasticity, the importance of properly implementing the driving device (artifacts due to constant separation and constant, sliding velocity constraints), and thermostats that allow to transport heat away from the interface in a non-equilibrium situation. Concerning all these aspects, one should of course attempt to mimic the experimental situation as closely as computationally feasible. In some cases, however, it is advantageous not to mimic experiment. For instance, if the mass of the slider is small in the simulation, the gap between the macroscopic processes and the microscopic motion is reduced. This makes it sometimes possible to simulate processes on rather small time scales that occur only on macroscopic time scales in experiment.

Multi-scale techniques that have been used in simulations of fracture like in Ref. [59] will certainly prove valuable in friction simulations as well. These techniques combine ab-initio, atomistic, and coarse-grained modeling within one simulation. In particular, the simulation of an AFM tip substrate interaction seems to be a well-suited problem: The intimate contact can be modeled in terms of ab-initio, the area further outside with an atomistic description, and the proper contact mechanics can be guaranteed with continuum methods for the areas even further away from the intimate contact.

Acknowledgments

The author thanks K. Binder, C. Brangian, M. O. Robbins, and M. Urbakh for useful discussions. This work was supported through the Israeli-German D.I.P.-Project No 352-101 and the Materialwissenschaftliche Forschungszentrum (MWFZ) Rheinland-Pfalz.

References

[1] G. Binning, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 930 (1986).

[2] C. M. Mate, G. M. McClelland, R. Erlandsson, and S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987).

[3] J. N. Israelachvili, P. M. McGuiggan, and A. M. Homola Science, **240**, 189 (1988).

[4] J. N. Israelachvili, Surf. Sci. Rpt., **14**, 109 (1992).

[5] U.D. Schwarz, O. Zwörner, P. Köster, and R. Wiesendanger, Phys. Rev. B **56**, 6987 (1997); *in ibit.* 6997.

[6] A. Berman, C. Drummond, and J. N. Israelachvili, Tribol. Lett. **4**, 95 (1998)

[7] F. P. Bowden and D. Tabor, *The Friction and Lubrication of Solids* (Clarendon Press, Oxford, 1958).

[8] J. Klein and E. Kumacheva, Science **269**, 816 (1995); J. Chem. Phys. **108**, 6996.

[9] A. L. Demirel and S. Granick, Phys. Rev. Lett. **77**, 2261 (1996).

[10] M. O. Robbins and M. H. Müser, in *Modern Tribology Handbook I*, pp. 717-765, Ed. B. Bhushan (CRC Press, Boca Raton, FL, 2001); cond-mat/0001056.

[11] V. Zaloj, M. Urbakh, and J. Klafter, Phys. Rev. Lett. **81**, 1227 (1998); J. Chem. Phys. **110**, 1263 (1999).

[12] M. H. Müser, in *Fundamentals of Tribology and Bridging the Gap between Macro- and Micro/Nanoscales*, pp. 235-240, Ed. B. Bhushan (Kluwer Academic Publishers, Netherlands, 2001); cond-mat/0012100.

[13] B.N.J. Persson, *Sliding Friction: Physical Principles and Applications* (Springer, Berlin, 1998).

[14] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*, (Academic Press, San Diego, 1996).

[15] L. Prandtl, ZS. f. angew. Math. u. Mech., **8**, 85 (1928).

[16] G. A. Tomlinson, Phil. Mag. Series, **7**, 905 (1929).

[17] E. Gnecco et al., Phys. Rev. Lett. **84**, 1172 (2000).

[18] Y. Sang, M. Dubé and M. Grant, Phys. Rev. Lett. **87**, 174301 (2001).

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

[20] D. Dowson, *History of Tribology* (Longman Inc., New York, 1979) pp. 153-67.

[21] M. Hirano and K. Shinjo, Phys. Rev. B **41**, 11837 (1990).

[22] J. Belak and I. F. Stowers, in *Fundamentals of Friction*, ed. I. L. Singer and H. M. Pollock (Springer-Verlag, Berlin, 1990).

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

[24] A. N. Kolmogorov and V. H. Crespi, Phys. Rev. Lett. **85**, 4727 (2000).

[25] M. H. Müser, Tribol. Lett. **10**, 15 (2001); cond-mat/0004494.

[26] A. Buldum and S. Ciraci, Phys. Rev. B **55**, 12892 (1997).

[27] M. H. Müser, L. Wenning, and M. O. Robbins, Phys. Rev. Lett. **86**, 1295 (2001).

[28] B. N. J. Persson and E. Tosatti, Solid State Comm. **109**, 739 (1999).

[29] D. S. Fisher, Phys. Rev B, **31**, 1396 (1985).

[30] C. Mak and J. Krim, Faraday Disc. **107**, 389 (1997).

[31] U. Landman and W. D. Luedtke, J. Vac. Sci. Technol. B, **9**, 414 (1991).

[32] J. E. Hammerberg, B. L. Holian, J. Röder, A. R. Bishop, and J. J. Zhou, Physica D, **123**, 330 (1998).

[33] J. Gao, W. D. Luedtke, and U. Landman, J. Chem. Phys. **106**, 4309 (1997).

[34] J.-C. Wang and K. A. Fichthorn, J. Chem. Phys. **112**, 8252 (2000).

[35] I. Bitsanis, S. A. Somers, H. T. Davis, and M. Tirrell J. Chem. Phys. **93**, 3427 (1990).

[36] P. A. Thompson and M. O. Robbins, Phys. Rev. A **41**, 6830 (1990).

[37] N. P. Suh, *Tribophysics*, (Prentice-Hall, Englewood Cliffs, 1986).

[38] G. He, M. H. Müser, and M. O. Robbins, Science **284**, 1650 (1999).

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

[40] G. He and M. O. Robbins, Tribol. Lett. **10**, 7 (2001).

[41] J. H. Dieterich, J. Geophys. Res. **81**, 2169 (1979).

[42] A. Z. Szeri, in *Modern Tribology Handbook I*, pp. 383-453, Ed. B. Bhushan (CRC Press, Boca Raton, FL, 2001).

[43] L. Bocquet and J.-L. Barrat, Phys. Rev. E **49**, 3079 (1994).

[44] K. Ohno, K. Esfarjani, and Y. Kawazoe, *Computational Materials Science*, (Springer, Berlin, 1999).

[45] J. M. Thijssen, *Computational Physics* (Cambridge University Press, Cambridge, 1999).

[46] For reviews, see e.g., G.S. Grest, in *Advances in Polymer Science 138*, p. 149, ed. S. Granick (Springer, Berlin, 1999); G. S. Grest and M. Murat, in *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*, p. 476, ed. K. Binder (Oxford University Press, New York, 1995).

[47] P. S. Doyle, E. S. G. Shaqfeh, and A. P. Gast, Macromolecules **31**, 5474 (1998).

[48] T. Kreer, M. H. Müser, K. Binder, and J. Klein, Langmuir **17**, 7804 (2001).

[49] K. Kremer and G. S. Grest, J. Chem. Phys. **92**, 5057 (1990).

[50] S. Erkoc, Phys. Rep. **278**, 79 (1997).

[51] L. Wenning and M. H. Müser, Europhys. Lett. **54**, 693 (2001).

[52] B. N. J. Persson, J. Chem. Phys. **113**, 5477 (2000).

[53] J. Krim and A. Widom, Phys. Rev. B **38**, 12184 (1988).

[54] V. Zaloj, M. Urbakh and J. Klafter, Phys. Rev. Lett. **82**, 4823 (1999).

[55] P. Espanol and P. Warren, Europhys. Lett. **30**, 191 (1995).

[56] M. Ripoll, M. H. Ernst, and P. Espanol, J. Chem. Phys. **115**, 7271 (2001).

[57] A list of references for treating elastic solids in terms of Green's function can be found at: http://www.ctcms.nist.gov/gf/references.php3.

[58] M. Weiss and F.-J. Elmer, Phys. Rev. B, **53**, 7539 (1996).

[59] F. F. Abraham, J. Q. Broughton, N. Bernstein, and E. Kaxiras, Europhys. Lett. **44**, 783 (1998); J. Q. Broughton, F. F. Abraham, N. Bernstein, and E. Kaxiras, Phys. Rev. B **60** 2391 (1999).