Description of atomic friction as forced Brownian motion

Peter Reimann and Mykhaylo Evstigneev
Universität Bielefeld, Fakultät für Physik, 33615 Bielefeld, Germany
E-mail: reimann@physik.uni-bielefeld.de

New Journal of Physics 7 (2005) 25
Received 31 August 2004
Published 31 January 2005
Online at http://www.njp.org/
doi:10.1088/1367-2630/7/1/025

Abstract. A theoretical description of friction force microscopy experiments in terms of a forced Brownian motion model is derived on the basis of microscopic considerations. Particular emphasis is put on the discussion of the relevant state variables/collective coordinates and on a realistic description of dissipation and inertia effects by means of comparison with experimental findings. The main new prediction of the model is a non-monotonic dependence of the friction force upon the pulling velocity of the AFM-tip relative to an atomically flat surface. The region around the force maximum can be approximately described by a universal scaling law and should be observable under experimentally realistic conditions.
1. Introduction

A first central point of Einstein’s celebrated work on Brownian motion is the insight that within the complicated many-body system at hand, a clear-cut time-scale separation exists between the particles’ ‘slow’ centre of mass and the huge number of remaining fast molecular degrees of freedom. A second main achievement represents the mathematical framework of how to exploit this physical insight for a simplified theoretical modelling of the system. Subsequently, the formal extensions and refinements of those basic concepts together with their successful application to a wide variety of different problems in physics, chemistry, and other natural sciences have given rise to an entire new research area of Brownian motion theory. The intention of our present contribution is to exemplify this point in the context of friction phenomena on the nanometre scale. In doing so, our main emphasis will be put on the adequate interpretation and modelling of recent experimental observations in terms of forced Brownian motion theory, while the actual solution of the model will be kept comparatively simple and brief by exploiting previously established concepts and results for related model systems.

Macroscopic friction between solids is well known to be both of paramount practical importance and of notorious difficulty regarding its theoretical understanding. Here, we restrict ourselves to the simpler case of a microscopic contact in the form of a single asperity. Recent progress in the field of nano-technology has also made possible a breakthrough in the exploration of this kind of friction on the atomic scale (nano-tribology). Specifically, we consider experiments by means of an atomic force microscope (AFM), adapted for the special purpose of so-called friction force microscopy [1]–[16], whose basic principle is sketched in figure 1. The tip of
Figure 1. Schematic illustration of friction force microscopy. Typical spatial extensions of the AFM-cantilever are $200 \mu m \times 50 \mu m \times 1 \mu m$. The tip has a typical height $5 \mu m$, basis radius $1 \mu m$, and apex radius $10 nm$. The lattice constant of the substrate is typically about $0.5 nm$. $v$ is the pulling velocity of the substrate, $F(t)$ the instantaneous friction or pulling force (cf footnote 1), $F_N$ the normal force and $x(t)$ the displacement of the tip apex from its rest position.

an AFM-cantilever is brought into contact with an atomically flat surface, which moves with a constant velocity $v$ along the $x$-axis. Via the deflection of a laser beam at the upper side of the cantilever, the resulting torsion is observable and can be immediately translated by means of the known elastic properties of the AFM into the corresponding time-resolved lateral ‘friction force’ $F(t)$ between the cantilever tip and substrate.¹

A representative example for such a friction force $F(t)$ is depicted in figure 2. Though this plot has been obtained by means of numerical simulations, it looks practically the same as in a real experimental measurement [1, 2], [6]–[9], [13]–[15]. The characteristic, saw-tooth like back-and-forth motion of the tip apex $x(t) = F(t)/\kappa$ in figure 2 essentially consists of time segments during which $x(t)$ appears to be ‘stuck’ by the moving substrate in figure 1, and very quick ‘slips’ from one ‘sticky site’ to the next, hence the name ‘stick–slip motion’. By closer inspection of figure 2 one readily sees that the distances between adjacent ‘sticky sites’ reproduce the substrate lattice constant $L$ along the $x$-axis. Hence, atomic resolution is achieved in spite of the fact that the tip apex is at least an order of magnitude ‘bigger’ than a single substrate molecule,² see caption to figure 1. More precisely, the stick-slip motion apparently has its origin in an $L$-periodic interaction potential between substrate and tip, and a ‘slip’ event is nothing else than a transition

¹ The most natural observable would be the pulling force applied to the substrate in figure 1. But in a real experiment, the pulling speed $v$ is imposed, while the concomitant pulling force is not directly accessible. However, that force is by Newton’s third law equal to (the negative of) the reaction force $F(t)$ of the cantilever, apart from inertia effects which drop out in the time average (1), see also sections 2.3.4 and 3.4.

² Estimates for the number of atoms which form the relevant effective ‘contact area’ between tip apex range from ‘a few’ [7], over $5–30$ [13] to 100 [32]. Though the self-consistency of continuum mechanics [30, 31] may be questionable on the scale of a few atoms [6, 13], it may consistently be employed to derive an upper estimate for the volume of the substrate which exhibits significant elastic deformations under the action for the tip apex. With 100 contact atoms and a lattice constant of $1 nm$ one obtains $0.01 \mu m^3$ as a safe upper estimate for the volume within which the contact strain is concentrated [31].
Figure 2. Instantaneous friction force $F(t)$ versus substrate position $s(t) = vt$ for $v = 3 \mu$m s$^{-1}$. Depicted is a small but representative time-slot after transients have died out, obtained by numerical simulations of the model dynamics (4), (7)–(9) with our standard parameters (19)–(22), (38), (39), $m = 5 \times 10^{-13}$ kg, and $\vartheta = 0.75$, see also (26), (36) and (42).

from one local potential minimum to the next, forced by the elastic deformation forces of the set-up. A second basic feature of figure 2 are the random fluctuations and the not completely regular time intervals between subsequent slip-events. These are immediate experimental indications that thermal fluctuations play a significant role.

A second important experimental finding arises if the substrate does not move with the same constant velocity $v$ for all times, but if one periodically switches between $v$ and $-v$. The result is a so-called friction loop of positive area [1, 2], [6]–[9], [13]–[15], implying that the position of the substrate does not yet uniquely fix the global geometrical/mechanical configuration of the set-up in figure 1. Hence, at least one additional relevant state variable/slow collective degree of freedom is needed for an adequate theoretical description of the system [17, 18], see section 2.

As far as quantitative measurements are concerned, even more important than the instantaneous friction force $F(t)$ is its time average

$$\bar{F} := \lim_{t \to \infty} \frac{1}{t} \int_{0}^{t} F(t') dt',$$  

see figure 3 for an example. From the dependence of this force on the pulling velocity $v$, the normal force $F_N$, or the temperature $T$, one tries to gain insight into the specific molecular properties of the probed surface and into the general microscopic principles of surface friction within an exceptionally simple ‘minimal’ system. The interpretation of the measurements along these lines represents a formidable theoretical challenge with respect to simplified model descriptions [4, 5], [7]–[21] and numerical molecular dynamics simulations [22, 23].

Specifically, thermal fluctuations have not been taken into account in some of the earlier theoretical works in this field, see e.g. [5, 17, 18]. A first direct experimental evidence of their significance has been mentioned above. A second direct experimental indication in the same direction is the temperature dependence of the average friction forces reported in [12]. A less direct proof represents the increasing behaviour of the average pulling force (1) as a function.
Figure 3. Experimentally measured average friction force $\bar{F}$ versus pulling velocity $v$ for normal forces $F_N = 12$ nN (dots), $F_N = 10$ nN (squares), and $F_N = 8$ nN (triangles), see also equation (1) and figure 1. The depicted data points and error bars are adopted from figure 1 in [11]. The measurements were performed on a mica surface. For further parameter values see (20)–(24).

of the velocity $v$, which has been observed in numerous experiments, see e.g. figure 3 and [3]–[14], [16]. Usually, this behaviour is theoretically explained with the help of thermal noise effects in models of the type we will study in this paper. While this behaviour can in principle also be explained with the help of somewhat different, purely deterministic models [14], the same is no longer possible for the above-mentioned more direct signatures of thermal-noise effects. Hence, thermal noise is now quite generally appreciated as an indispensable ingredient of an adequate theoretical description of friction force microscopy [4], [7]–[9], [11]–[13], [16], [19]–[21]. In other words, we are indeed faced with a problem of forced Brownian motion.

In this paper, we refine and extend existing theories by way of a more realistic modelling of the dissipation, noise, and inertia effects. In this manner, we are for the first time able to explain not only the well-known initial increasing behaviour of $\bar{F}$ in figure 3, but also the apparent ‘plateaux’ at the high-$v$ end, as reported in the experimental studies [3, 5, 11, 16]. The main consequence is the prediction of a non-monotonic velocity dependence of the ‘effective’ friction force (1) upon the pulling velocity $v$. As a specific example we will focus on the experimental set-up in [11] (see also [16]) since the predicted non-monotonicity should be observable by increasing the velocities $v$ by a rather moderate amount beyond those achieved in [11], see figure 3.

2. Model

The system from figure 1 gives rise to a paradigmatic applications of the well-known general principles of stochastic modelling. Starting from microscopic considerations, basically two relevant (slow) state variables (collective degrees of freedom) can be identified, while the effect
of all the remaining (fast) degrees of freedom essentially boils down to friction, thermal noise, and a renormalization of the relevant potential and inertia within an effective model dynamics for the slow variables alone. While the general mathematical framework for eliminating the fast ‘thermal bath’ variables along these lines is described in detail, e.g. in [24]–[28], here we mainly focus on the physical picture behind those calculations and the very interesting particularities of the specific set-up at hand.

2.1. Basic modelling steps

To model the system from figure 1, we first neglect (average out) the ‘fast’ thermal fluctuations of all participating molecular degrees of freedom of the AFM and the substrate. The remaining possibilities (state space) of the ‘slowly’ varying, ‘global’ mechanical/geometrical configurations in figure 1 are described by the instantaneous substrate position \( s(t) = vt \) and the instantaneous displacement \( x(t) \) of the tip apex from its rest position in the \( x \)-direction.\(^3\) Thus, all the elastic deformations of the cantilever and the tip–substrate contact region are assumed to be uniquely fixed for any given value of \( x \) and \( s \). Furthermore, these two ‘relevant’ collective coordinates are assumed to only vary so slowly that the ‘fast’ molecular degrees of freedom always remain close to the instantaneous accompanying equilibrium [24]–[28]. The justification of these assumption will be addressed in more detail in sections 2.3.2 and 2.3.3. For the time being, we content ourselves with their immediate intuitive plausibility.

By averaging out the thermal fluctuations of the fast degrees of freedom for any given values of the slow ones, \( x \) and \( s \), a free energy type ‘potential of mean force’ \( U_{\text{tot}}(x, s) \) is obtained. Essentially, this potential consists of three contributions. The first accounts for the elastic deformations of AFM and substrate, the second for the molecular interaction between tip and substrate, and the third for the entropy of the fast variables. Since the elastic deformations are typically small [29], we may neglect anharmonic effects and assume that interaction and entropy only depend on the relative tip–substrate position \( x - s \). In other words, we arrive at the approximation

\[
U_{\text{tot}}(x, s) = \frac{\kappa}{2} x^2 + U(x - s),
\]

where \( \kappa \) is the combined elasticity coefficient of cantilever, tip apex, and substrate [18], [29]–[32]. Moreover, focusing on an ideally flat atomic surface with lattice constant \( L \) in the \( x \)-direction, we can conclude that \( U_{\text{tot}}(x, s) \) is invariant under a displacement of the substrate by one period, \( s \mapsto s + L \), implying that

\[
U(x + L) = U(x).
\]

Finally, the experimentally observable lateral force \( F(t) \) (see figure 1) can be identified, according to Newton’s third law actio = reactio, with the negative of the force caused by the elastic deformations, i.e.

\[
F(t) = \kappa x(t).
\]

\(^3\) Note that unlike the majority of previous theoretical investigations in this field, we work in the frame of reference of the AFM rather than in the rest coordinates of the substrate.
Figure 4. An adiabatically slow change from some initial (solid) to some final (dashed) cantilever deformation is a reversible process, but dissipative/damping/friction forces arise as soon as the change takes place at a finite speed.

The so far neglected ‘fast’ thermal fluctuations of the single atoms may be viewed as a ‘thermal bath’ to which the ‘slow’ collective coordinates $x$ and $s$ are coupled, with dissipation and noise as its main effects [24]–[28]. We remark that within our model, $s$ is not an independent dynamic variable but rather externally imposed via the ‘trivial equation of motion’ $s(t) = vt$. Hence only the bath effects on $x$ are of interest.

Our starting point are the elastic deformations of the cantilever, and, in particular, those of the tip apex (see figure 1). If these deformations, or equivalently, the state variable $x$, are changing adiabatically slowly, then the system is at every instance of time in a thermal equilibrium state, i.e., we are dealing with a reversible process. Next we consider the case that these changes are taking place at a finite speed, but still slowly enough that the ‘cantilever bath’, consisting of the ‘fast’ molecular degrees of freedom, always remains close to the instantaneous accompanying thermal equilibrium [24]–[28], see figure 4. The remaining ‘small amount of disequilibrium’ renders the process under consideration ‘slightly irreversible’ and hence gives rise to a linear-response-type dissipative force

$$ F_c(t) = -\eta_c \dot{x}(t) \quad (5) $$

with an effective coupling strength $\eta_c > 0$ between the collective coordinate $x$ and the close to equilibrium ‘cantilever bath’ (subscript ‘c’). In fact, this is nothing else than Onsager’s theory for close to equilibrium processes, associating currents $\dot{x}(t)$ via linear response/Onsager coefficients $[\eta_c]$ with corresponding dissipative forces $[F_c(t)]$. In particular, since the cantilever deformations are typically small [29], the implicitly assumed independence of the Onsager coefficient $\eta_c$ on the state $x$ of the system is well justified.

We emphasize that only the environment (bath) of the actual system of interest (described by $x$) is assumed to remain close to thermal equilibrium, not the system itself!
Next we turn to the elastic deformations of the substrate (see figure 1). To understand the essential point, it is convenient to switch for a moment to the rest frame of the substrate. That is, the substrate is now fixed and the tip–substrate contact travels with velocity $-\dot{s}(t) + \dot{x}(t)$ along the $x$-direction (see figure 5). Hence, also the substrate deformations are moving with velocity $-\dot{s}(t) + \dot{x}(t)$, giving rise, along the same line of reasoning as in (5), to a linear-response-type dissipative force

$$F_s(t) = -\eta_s(\dot{x}(t) - v),$$

where we have returned to the rest frame of the AFM and where we have exploited that $\dot{s}(t) = v$.

Finally, we come to the randomly fluctuating forces acting on the ‘slow’ state variable $x$. They have the same origin as the dissipative forces, namely the large number of ‘fast’ degrees of freedom of the cantilever and substrate baths. Due to this common origin and the fact that the baths always remain close to thermal equilibrium, one can show [33] that those randomly fluctuating forces are already completely fixed (in the statistical sense) by the functional forms (5) and (6) of the dissipative forces. Namely, the thermal ‘cantilever-noise’ acts on $x(t)$ in the usual form [24]–[28] of a fluctuating force $\sqrt{2\eta_c kT} \xi_c(t)$ with temperature $T$, Boltzmann constant $k$, and unbiased, $\delta$-correlated Gaussian noise $\xi_c(t)$. Similarly, the substrate gives rise to thermal fluctuations of the form $\sqrt{2\eta_s kT} \xi_s(t)$ with an unbiased, $\delta$-correlated Gaussian noise $\xi_s(t)$ independent of $\xi_c(t)$. Essentially, the uniqueness of these thermal noises follows from the fact that any deviation from the above-specified statistical properties could be exploited to construct a perpetuum mobile of the second kind [33].

Collecting all acting forces, we arrive at the following equation of motion:

$$m \ddot{x}(t) = -U'(x(t) - vt) - kx(t) - \eta_c \dot{x}(t) + \sqrt{2\eta_c kT} \xi_c(t) - \eta_s(\dot{x}(t) - v) + \sqrt{2\eta_s kT} \xi_s(t),$$

where $m$ is the relevant effective mass associated with inertia effects of cantilever, tip apex, and substrate in figure 1. To better understand its meaning, we consider an arbitrary point within the cantilever in figure 4. Then a small tip apex dislocation $\Delta x$ yields a displacement of our reference
point proportional to $\Delta x$ with some proportionality factor $q$ (directions do not matter). Hence, inertia forces against an acceleration $\ddot{x}(t)$ acquire a local weighting factor $q$, and $m$ follows by integrating over the local (mass-) density times $q$. Note that in full generality, some parts of the cantilever may be accelerated even if the tip apex moves with a constant velocity, and $m$ will itself depend on the instantaneous deformation $x$. However, since the elastic deformations are typically small [29], both these effects are negligible. Similar considerations apply for the substrate, as long as it does not move, i.e. $v = 0$. For $v \neq 0$ it is clear from figure 5, that even for $\ddot{x}(t) = 0$ there are locally accelerated parts of the substrate in the elastically deformed contact region since those deformations are ‘propagating through the substrate’. However, it should be kept in mind that the resulting integrated inertia force equals the change of the total momentum of the substrate material. Since such changes of momenta (accelerations) are invariant under Galilei transformations it follows that the corresponding inertia effects (and hence $m$) are in fact independent of $v$ and it is sufficient to determine them for $v = 0$.

2.2. Comparison with previous models

The equation of motion (7) is the first main result of our paper. Once its is solved, the main experimentally observable quantity $F(t)$ (see figure 1) immediately follows from (4). In order to rewrite (7) in a more convenient form, we first introduce the total and relative friction coefficients

$$\eta := \eta_c + \eta_s, \quad (8)$$

$$\vartheta := \eta_c / \eta \quad (9)$$

with the obvious properties $\eta > 0$ and $\vartheta \in (0, 1)$. Further, we can lump the two independent Gaussian white noise terms in (7) into a single fluctuating force $\sqrt{2\eta kT}\xi(t)$ with an unbiased, $\delta$-correlated Gaussian noise $\xi(t)$. Finally, we introduce the ‘co-moving’ variable

$$X(t) := x(t) - vt, \quad (10)$$

i.e. we go over from the reference frame of the AFM into that of the moving substrate. Along these lines, we can rewrite (7) in the equivalent form

$$m\ddot{X}(t) = -U'(X(t)) - F(t) - \vartheta \eta v - \eta \dot{X}(t) + \sqrt{2\eta kT}\xi(t), \quad (11)$$

where $F(t) = \kappa[X(t) + vt]$. Models reminiscent of equation (11) are long known under the labels ‘Tomlinson model’, ‘Prandtl model’, ‘independent oscillator model’, etc. In spite of the extensive related literature (reviewed, e.g., in [8, 34]) the above section 2.1 represents to our knowledge the first detailed derivation on the basis of microscopic considerations of each single term in equation (11) in the specific context of friction force microscopy experiments conducted by means of an AFM.

In particular, in all previous comparable models in this context [5, 12], [19]–[21], [32], dissipative effects due to deformations of the tip and the concomitant thermal fluctuations have been ignored. One of the main points of our present work is to demonstrate that these effects are in fact quite significant under realistic experimental conditions. In most cases [5, 12], [19]–[21], [32], only substrate dissipation and noise have been taken into account, corresponding to $\eta_c = 0$ and hence $\vartheta = 0$. In the work [32], dissipation and noise originating in the tip–substrate
contact region have been ignored, thus keeping only the remaining cantilever bath effects. This corresponds to $\eta_s = 0$ and hence $\vartheta = 1$ and a much too small value of $\eta_c$, see also section 3 below. It is only in the quite different context of friction between macroscopic solids [34, 35] that models analogous to (11) with $0 < \vartheta < 1$ have been considered previously, however with the conclusion of no significant difference compared to $\vartheta = 0$.

A further problematic point in all previous comparable models is the determination of the relevant effective mass $m$ in (11), as will be discussed in detail in section 3.2 below.

2.3. Additional modelling steps

In this section we discuss in more detail some implicit assumptions and generalizations of the model from section 2.1.

2.3.1. Thermal environment.

In deriving the cantilever and substrate friction effects (5) and (6) we have made the implicit assumption that no irreversible damage of tip or surface occurs. In the opposite case of non-negligible dissipative wear effects, a description of the environment in terms of a bath close to thermal equilibrium would break down. This assumption of wearless friction is justified experimentally, e.g. in [2, 4, 9, 13, 22, 23].

Besides the ‘cantilever and substrate thermal baths’, giving rise to friction terms (5) and (6) and concomitant thermal fluctuations, yet another part of the thermal environment may be an ambient fluid (liquid or gas). Even if the microfluidics of such a set-up may in general be quite complicated [39], with respect to our specific problem at hand we expect that an analogous linear-response-type reasoning as in the derivation of (5) and (6) remains valid. As a result, $\eta_c$ (and possibly also $\eta_s$) will be renormalized to a somewhat increased effective value. However, while friction coefficients associated with dissipation in a solid are (practically) temperature independent, the same is no longer true for friction caused by a fluid.

Yet a fourth relevant part of the environment may be an ‘electronic thermal bath’, giving rise to electronic friction and noise effects [8]. Unless there arise additional slow collective variables in the form of macroscopic electric currents, the result will be just another renormalization of $\eta_c$ and $\eta_s$.

2.3.2. Relevant collective coordinates.

Next, we address the assumption at the beginning of section 2.1 that the two collective coordinates $x$ and $s$ are the only relevant slow variables.

A first set of collective coordinates, which we have omitted in our model, but which may be ‘dangerously slow’ are the amplitudes and phases of the vibration eigenmodes of the cantilever [36]–[38] in contact with the substrate. A typical value of the smallest such eigenfrequency is $v_0 = 200$ kHz [11, 36, 37]. On the other hand, the relative substrate–tip velocity $v$ and the substrate lattice constant $L$ give rise to a characteristic frequency $v/L$ of the tip apex motion $x(t)$. A first guess is therefore that vibration modes will be negligible if they are sufficiently ‘fast’ compared to $x(t)$, i.e. if $v$ is significantly smaller than $L v_0$. For the typical values $L = 0.5$ nm and $v_0 = 200$ kHz [11], this yields the condition $v < L v_0 = 100 \mu$m s$^{-1}$.

4 Taking into account that the tip apex $x(t)$ typically exhibits not a sinusoidal but rather a saw-tooth (‘stick–slip’) motion (see section 1), higher harmonics of the characteristic frequency $v/L$ will also contribute with significant amplitudes to $x(t)$, yielding in fact an even much lower bound for $v$. 

New Journal of Physics 7 (2005) 25 (http://www.njp.org/)
Figure 6. Cartoon of the same set-up as in figure 1, but in a still more schematic representation. The ‘stiff’ cantilever is represented by a ‘rigid plate’, the ‘soft’ tip–substrate contact region by a ‘flexible filament’, see also figure 7. (a) Two different instantaneous conformations (solid and dashed) during a typical friction force microscopy experiment, cf figure 1. The main changes take place in the tip–substrate contact region, while the rest of the cantilever only undergoes very small deformations. (b) Schematic sketch of two different instantaneous conformations (solid and dashed) for a torsional vibration eigenmode of the cantilever in contact with the substrate. The cantilever motion dominates, while the much ‘softer’ tip–substrate contact region adapts itself without much resistance. Conclusion 1: ‘local’ conformation changes like in (a) excite the ‘global’ vibration eigenmode in (b) only very weakly and their ‘back-reaction’ effects to the process (a) are negligible. Conclusion 2: Since the instantaneous substrate position $s(t) = vt$ and the instantaneous displacement $x(t)$ of the tip apex from its rest position are assumed to uniquely fix the entire configuration of cantilever and tip–substrate contact region within the model (7) and since this model is supposed to describe the dynamical processes in (a), it cannot describe at the same time those in (b). Conclusion 3: in (b), the tip–substrate contact affects the cantilever vibrations very little (like a ‘dangling end’). Hence the torsional vibration frequencies in and out of contact with the substrate are similar, in agreement with the experiments from [36, 37].

However, the above viewpoint is too pessimistic. Besides the time-scale separation, a second sufficient condition for neglecting a potentially slow collective coordinate is its weak coupling to $x(t)$. This is indeed the case for vibrational eigenmodes for the following reason. As detailed in section 3.1, under typical experimental conditions the ‘small’ tip–substrate contact region is much ‘softer’ than the ‘large’ rest of the cantilever. Hence, the ‘local’ movements $x(t)$ may excite the ‘global’ torsional cantilever eigenmodes only quite weakly and their ‘back-reaction’ on $x(t)$ is completely negligible compared to the original ‘action’ of $x(t)$ itself, see figure 6. (However, they may have a notable effect on the experimentally observed deflection of the laser beam at the upper side of the cantilever, cf section 2.3.4 below.) The coupling to other than torsional eigenmodes of the cantilever is obviously even weaker.
A second type of ‘potentially dangerous’ slow variables arises from the fact that $x$ and $s$ are assumed (see the beginning of section 2.1) to uniquely fix all the elastic deformations of the cantilever and the tip–substrate contact region, in particular the $y$- and $z$-coordinate of the tip apex. Irrespective of these additional collective coordinates $y(t)$ and $z(t)$, the full ‘potential of mean force’ $U_{tot}(x, y, z, s)$ is of the form $U_1(x, s) + U_2(y, z, s)$, seems to be an acceptable approximation in many cases. In particular, a potential of exactly this type may very well arise in some specific real system and there is no reason to expect that this will lead to a behaviour fundamentally different from that for more general potentials.\footnote{5}

While we are not aware of good arguments for a clear-cut time-scale separation between the stochastic dynamics of $x(t)$ from those of $y(t)$ and $z(t)$, the assumption that $x(t)$ decouples from $y(t)$ and $z(t)$, i.e. the full ‘potential of mean force’ $U_{tot}(x, y, z, s)$ is of the form $U_1(x, s) + U_2(y, z, s)$, seems to be an acceptable approximation in many cases. In particular, a potential of exactly this type may very well arise in some specific real system and there is no reason to expect that this will lead to a behaviour fundamentally different from that for more general potentials.\footnote{5}

As we will see later, the model dynamics (7) turns out to be often dominated by rare, thermally activated escape processes over some potential barriers and the subsequent essentially deterministic relaxation towards some metastable states. For such processes, the effect of additional state variables besides the main reaction coordinates has been studied in detail in the context of reaction rate theory \cite{27,40} and one typically finds rather moderate corrections of the pre-exponential contribution to the rate (entropic factor), while the exponentially dominating Arrhenius-factor (energy barrier) is not affected at all.\footnote{6}

A third type of additional relevant collective variables are discussed in the next subsection.

2.3.3. A refined model. Throughout section 2.1 we have tacitly assumed that the instantaneous displacement $x(t)$ of the tip apex from its rest position in $x$-direction is much larger than the corresponding elastic deformations of the substrate in the $x$-direction (the reasons will become clear below). As will be discussed in section 3.1, this assumption may not be fulfilled very well, see also \cite{30}. Giving up this assumption, we rename the former displacement, $x(t)$, as $x_1(t)$ and the latter is denoted as $x_2(t)$. More precisely, $x_2$ is the deformation of the substrate in the $x$-direction at the point of contact with the tip, see also figure 7 of this paper, figures 2(a) and 3 in \cite{30}, and figure 3 in \cite{32}. Since $x_2(t)$ is not a simple function of $x_1(t)$ we have to consider it as an additional relevant state variable. Similarly as in (2), for any fixed values of $x_1$, $x_2$, and $s$, the potential of mean force $U_{tot}(x_1, x_2, s)$ is expected to be approximately given as the sum of harmonic terms due to elastic cantilever and substrate deformations and an $L$-periodic tip–substrate interaction potential which depends only on the relative positions of the tip $x_1$ and of the substrate at the contact point $s + x_2$.

\footnote{5} The only theoretical study of a model analogous to (7) but with $y(t)$ as a second state variable known to us is \cite{14}. The viewpoint therein is, at first glance, opposite to ours, namely that without this additional variable $y(t)$ significant experimentally observed effects are missed. However, considering $y(t)$ as a substitute for the thermal fluctuations omitted in \cite{14}, the viewpoints are actually quite close.

\footnote{6} In particular, the so-called detailed balance symmetry (microreversibility) guarantees that forward and backward reactions between different metastable states follow essentially the same reaction pathways within the high-dimensional state space.
Figure 7. Elastic deformations of the tip–substrate contact region along the x-direction (magnified detail of figure 1). \( x_1 \) denotes the displacement of the AFM tip apex from its rest position and \( x_2 \) denotes the corresponding deformation of the substrate at the point of contact. Hence, \( x_1 \) and \( x_2 \) have opposite signs and \( x_1 - x_2 \) quantifies the total deformation of cantilever, tip, and substrate.

\[
U_{tot}(x_1, x_2, s) = \frac{\kappa_1}{2} x_1^2 + \frac{\kappa_2}{2} x_2^2 + U(x_1 - (s + x_2)).
\]  

(12)

Along the same line of reasoning as in (4)–(6) one finds for each of the two state variables \( x_i(t), i = 1, 2 \), an equation of motion analogous to (7), i.e.

\[
m_i \ddot{x}_i(t) = -\kappa_i x_i(t) - \eta_c \dot{x}_i(t) + \sqrt{2 \eta_c k T} \xi_i(t) + F_i(t),
\]

(13)

\[
F_1(t) = -F_2(t) = -U'(x_1(t) - x_2(t) - vt) - \eta_{int}(\dot{x}_1(t) - \dot{x}_2(t) - v) + \sqrt{2 \eta_{int} k T} \xi_{int}(t),
\]

(14)

where \( \xi_1(t), \xi_2(t) \) and \( \xi_{int}(t) \) are independent, \( \delta \)-correlated Gaussian noises. The first three terms in the right-hand side of (13) are elastic, dissipative and fluctuating forces caused by the cantilever \((i = 1)\) or the substrate \((i = 2)\). The last term \( F_i(t) \) stands for the respective three types of forces caused by tip–substrate interaction effects (see equation (14)), hence \( F_1 = -F_2 \) according to Newton’s third law. Finally, the two effective masses \( m_i \) in (13) are defined in analogy to the discussion of \( m \) below (7). The mass \( m_1 \) encapsulates inertia effects caused by the cantilever and the tip apex in response to the acceleration of matter in cantilever and tip apex when the displacement of the tip apex \( x_1(t) \) in figure 7 experiences an acceleration \( \ddot{x}_1(t) \). Similarly, \( m_2 \) accounts for inertia effects of the substrate material in response to accelerations of the substrate deformation \( \ddot{x}_2(t) \), which, for the same reasons as explained below (7), are independent of the pulling velocity \( v \) and hence most conveniently determined for \( v = 0 \).

If the substrate is very stiff, i.e. \( \kappa_2 \gg \kappa_1 \), then the substrate deformations \( x_2(t) \) will be negligibly small, i.e. we can make the approximation \( x_2(t) \equiv 0 \). As a consequence, one recovers (7) from (13) for \( i = 1 \) after making the obvious identifications \( x := x_1, \kappa := \kappa_1, \eta_c := \eta_1, \eta := \eta_1 \).
\[ \eta_s := \eta_{\text{int}}, \text{ etc.} \]

In the opposite case that the cantilever (including the tip) is very stiff, one recovers (7) from (13) for \( i = 2 \) with \( x := -x_2, \kappa := \kappa_2, \eta_c := \eta_2, \eta_s := \eta_{\text{int}}, \text{ etc.} \) A third special case of interest arises if \( \kappa_1 = \kappa_2, \eta_1 = \eta_2, \text{ and } m_1 = m_2. \) By means of the following identifications

\[ x := x_1 - x_2, \quad \frac{1}{\kappa} := \frac{1}{\kappa_1} + \frac{1}{\kappa_2}, \quad \eta_c := \eta_{\text{int}}, \]

\[ \eta_c := \frac{\eta_1 + \eta_2}{4}, \quad m := \frac{m_1 + m_2}{4}, \]

one recovers once again (7) by subtracting the two equations (13) for \( i = 1 \) and \( i = 2 \) from each other.

Note that the three relations in (15) are common to all three above-discussed special cases. The first relation accounts for the relevant total deformation of cantilever and substrate, \( x(t) \) (see figure 7), the second for the corresponding total elasticity, \( \kappa \) (see section 3.1), and the third for the dissipative coupling between tip apex and substrate, \( \eta_s. \) The relations in (16) are quite suggestive as well, accounting for the relevant intrinsic dissipation, \( \eta_c \) and effective mass, \( m, \) of the cantilever–substrate entity in the last of the above-discussed special cases, while slightly modified relations apply to the other two cases.

For more general cases it is not possible to recover the simplified model (7) exactly from (13), but it is quite convincing that it will still represent a valid approximate description.

2.3.4. Friction forces. Naively speaking, the experimentally observable quantity \( F(t) \) in figure 1 is the lateral ‘reaction’-force of the cantilever in response to the tip apex displacement (see footnote 1). Within the refined model from the previous subsection, this displacement equals \( x_1(t), \) and according to Newton’s third law the lateral force would then be equal to the negative of the first three terms in the right-hand side of (13) for \( i = 1. \) However, more precisely speaking, it is the deflection of a laser beam at the upper surface of the cantilever in figure 1 which is the primary experimental observable. This is actually not a force but rather a geometrical, coordinate-like quantity. Within the present model, this deflection is directly proportional to the displacement \( x_1(t) \) of the tip apex. This geometrical quantity is only pro forma expressed as the equivalent (negative or reaction) force which one would have to apply at the tip apex in order to achieve exactly the same deflection of the laser beam in the absence of the substrate and of any cantilever intrinsic dissipative and thermal noise effects. Hence, \( F(t) \) equals the negative of only the very first term in the right-hand side of (13) for \( i = 1, \)

\[ F(t) = \kappa_1 x_1(t). \]

We remark that with respect to the average force (1) one does not have to care about these subtleties since the time average of the second and third terms in the right-hand side of (13) vanish. Similarly, by time averaging (13), one sees that (17) leads to the same average force (1) as if one had used the relation

\[ F(t) = -\kappa_2 x_2(t). \]
In other words, apart from unbiased fluctuations, equations (17) and (18) are equivalent and rigorous.\footnote{By adding (13) for \(i = 1\) and \(i = 2\) in the special case that \(\kappa_1 = \kappa_2, \eta_1 = \eta_2,\) and \(m_1 = m_2,\) one finds that \(x_1(t) + x_2(t)\) satisfies a linear stochastic dynamics of the form (13) with \(F_i(t) \equiv 0,\) independent of the process \(x(t) = x_1(t) - x_2(t).\) Hence, in this simple example, the difference between (17) and (18) amounts to a simple Gaussian random process (so-called harmonic noise \cite{41}).}

In fact, our above model assumption that the laser deflection is proportional to the displacement \(x_1(t)\) of the tip apex is itself an approximation. In reality, the upper surface of the cantilever, where the deflection takes place, is itself subjected to additional thermal fluctuations, which are not included in those of \(x_1(t).\) Rather, they are connected with additional slow collective coordinates such as vibration eigenmodes of the cantilever, see section 2.3.2 below. Hence, the short-term fluctuations of the experimentally observed \(F(t)\) cannot be described by our model, and within this restriction, (17) and (18) are once again seen to be equivalent approximations.

Finally, the most direct argument for this equivalence is as follows. Neglecting once again fluctuation and inertia effects, \(\kappa_1 x_1(t)\) and \(\kappa_2 x_2(t)\) are the forces which one has to externally apply to the cantilever and substrate, respectively, in order to maintain the given elastic deformations. Due to Newton’s third law they have equal magnitudes but opposite directions, implying (17) and (18).

In the previous subsection, the refined model (13) was shown to be exactly equivalent to the simplified model (7) in three special cases, with the properties \(\kappa_1 \to \infty, \kappa_2 \to \infty,\) and \(\kappa_1 = \kappa_2.\) In all three cases one also readily recovers the corresponding simplified friction force relation (4) by adding (17) and (18) and exploiting (15).

Finally, we note that we should carefully distinguish between the experimentally observable ‘macroscopic friction force’ \(F(t)\) and ‘microscopic friction forces’ due to dissipation of the system into the various thermal baths, appearing in the form of friction coefficients times velocities in our model equations (7), (11) and (13). On the one hand, it is clear that the few relevant collective coordinates cannot store a substantial amount of energy in the long run and hence, on the average, input of macroscopic friction energy via \(F(t)\) equals the microscopically dissipated energy. On the other hand, this balance of friction energies does not imply any obvious relation between the respective friction forces.

3. Estimating model parameters

Having established the general form of the equation of motion (7) or, equivalently, (11) by means of microscopic considerations, we are left with estimating the model function \(U(x)\) and the various model parameters appearing in those equations.

The temperature \(T,\) the effective coefficient of elasticity \(\kappa\) and the spatial period \(L\) can be readily determined experimentally. Typically, they are of the order of 300 K, 1 N m\(^{-1}\) and 1 nm, respectively \([7, 8, 11, 15, 20]\).

Similar estimates from experiments or \textit{ab initio} molecular dynamics simulations are not available for the functional form of the periodic interaction potential \(U(x).\) Hence it is for the time being considered as a phenomenological model function that has to be adapted later so as to optimally fit the experimentally measured friction forces (1). The simplest and most frequently
used ansatz is

\[ U(x) := \frac{E_0}{2} \sin(2\pi x/L) \]  

(19)

with typical amplitudes \( E_0 \) of the order of \( 10^{-19} \) J. In our numerical examples below, we will employ the specific choice (19), while the analytics will be valid for a quite general \( U(x) \).

Hence, we are left with estimating the effective mass \( m \) and two of the friction quantifiers in (8) and (9). As a typical example, we will focus on the experimental set-up in [11] (see also figure 3 and [16]) when concrete quantitative values will be needed. In particular, for this specific experiment the more precise estimates

\[ T = 293 \text{ K (room temperature)}, \]  

(20)

\[ \kappa = 1.2 \text{ N m}^{-1}, \]  

(21)

\[ L = 0.52 \text{ nm} \]  

(22)

are available.

Our first remark is that in the absence of the tip–substrate contact, the torsional ‘bare elasticity’ of the cantilever is [11]

\[ \kappa_0 = 75 \text{ N m}^{-1}. \]  

(23)

Further, the ‘real mass’ \( m_0 \) of the V-shaped cantilever (Ultralever Type B) used in [11] can be readily determined from its approximate effective spatial extensions \( 180 \mu m \times 72 \mu m \times 1 \mu m \) and its density \( 2330 \text{ kg m}^{-3} \) (silicon), yielding the estimate

\[ m_0 = 3 \times 10^{-11} \text{ kg}. \]  

(24)

The tip consists of the same material as the cantilever but is much smaller (cf caption to figure 1), hence its contribution to \( m_0 \) is negligible. Basically, the same value for \( m_0 \) is recovered from the measured torsional ‘free cantilever’ vibration frequency \( f = 250 \text{ kHz} \) [11] according to the well-known approximation \( m_0 = \kappa_0/(2\pi f)^2 \) for a weakly damped harmonic oscillator. The corresponding quality factor \( Q = \sqrt{\kappa_0 m_0/\eta_0} \) of those oscillations is typically at least 100 (in ambient air) [36], i.e. they are indeed weakly damped, implying for the torsional ‘bare cantilever damping coefficient’ that \( \eta_0 < 5 \times 10^{-7} \text{ kg s}^{-1} \). However, these bare quantities \( \kappa_0, m_0, \eta_0 \) have little to do with their effective counterparts, \( \kappa, m, \eta \) appearing in (7) and (11), since all tip–substrate contact effects are still missing.

3.1. Contact elasticities

As detailed in [30] (see also [18, 29, 31, 32]), the relevant effective elasticity coefficient \( \kappa \) can be viewed as the result of three coupled springs in series, associated with the elastic deformations of cantilever (\( \kappa_0 \)), tip apex (\( \kappa_t \)), and substrate (\( \kappa_s \)) according to

\[ \kappa^{-1} = \kappa_0^{-1} + \kappa_t^{-1} + \kappa_s^{-1}. \]  

(25)
Deviating from the main discussion, we note that in the same spirit \([\kappa_0^{-1} + \kappa_t^{-1}]^{-1}\) and \(\kappa_s\) can be identified with \(\kappa_1\) and \(\kappa_2\) from section 2.3.3, respectively, and (25) then goes over into the equation for \(\kappa\) in (15).

Taking into account the quantitative values (21) and (23) for the experiment from [11], equation (25) implies that the elasticity of the tip–substrate contact region, quantified by \([\kappa_0^{-1} + \kappa_t^{-1}]^{-1}\), must be smaller than 1.2 N m\(^{-1}\). Hence, we can conclude—in accordance with [15]—that ‘in comparison to the elasticity of the tip–substrate contact, the rest of the cantilever is extremely rigid’. Moreover, for analogous reasons, \(\kappa_t\) or \(\kappa_s\) or both must be smaller than 1.2 N m\(^{-1}\) as well. In the experimental work [30] it has been shown that \(\kappa_t\) and \(\kappa_s\) are comparable for the case of a cantilever and substrate for which also \(\kappa_0\) was comparable to \(\kappa_t\) and \(\kappa_s\), i.e. a quite different situation than in our present reference experiment [11]. In the absence of experimental data, a comparable elasticity of substrate and tip apex may be a reasonable working hypothesis for our present system (see also [11]), but substantially differing values cannot be excluded. In particular, the statement from [15] that \(\kappa\) was in the range of 1 N m\(^{-1}\) for various different substrate materials in combination with (23) and (25) indicates that the tip apex may be the most flexible piece. This is also in agreement with what one intuitively expects from a sharp, conically shaped tip, scratching over a flat surface, unless the latter consists of a much softer material.

### 3.2. Effective mass

As detailed in section 2.3.3, the vibration eigenmodes of the cantilever in contact with the substrate give rise to additional slow collective coordinates besides \(x(t)\) and \(s(t) = vt\), which, by construction, are omitted in the model dynamics (7) or (11). An alternative way to see this is provided by conclusion 2 in figure 6. Since our model does not include cantilever vibration modes, it is clear that none of the model parameters in (7) or (11) can be related to those vibrations. In other words, not only properties of the ‘bare cantilever’ are of little use, as discussed below equation (24), but also experimental findings regarding vibration modes of the cantilever in contact with the substrate [36, 37] are useless for estimating model parameters in (7) or (11).

In all previous theoretical works known so far, the effective mass \(m\) in models comparable to (7) or (11) has been set equal to the bare cantilever mass \(m_0\) or it has been estimated by means of cantilever eigenvibrations. According to the reasoning in the preceding paragraph and below equation (24), both these methods are not admissible.

Our starting point to estimate the effective mass \(m\) is its definition as given below equation (7). Essentially, one has to estimate proportionality factors between small displacements \(\Delta x\) of the tip apex and the concomitant displacement of an arbitrary but fixed reference point of the cantilever or the tip–substrate contact region. Due to the conclusion in the preceding section that in comparison to the elasticity of the tip–substrate contact, the rest of the cantilever is extremely rigid, it follows that small displacements \(\Delta x\) of the tip apex will lead to comparable displacements of the material within the elastically deformed tip–substrate contact region, while the displacements within the rest of the cantilever are reduced by roughly a factor \(\kappa/\kappa_0\). Hence, the effective mass \(m\) consists of a summand of the order of \(m_0 \kappa/\kappa_0\) and a summand accounting for the mass of the elastically deformed tip–substrate contact region. Considering (23), (25) and the fact that the mass of the entire tip is much less than \(m_0/10\) (cf caption to figure 1 and discussion below (24)) and similarly for the notably deformed substrate material (cf footnote 2)
we arrive at the following quite conservative estimate

\[ m \leq m_0/10 = 3 \times 10^{-12} \text{kg}, \]

(26)

where we have exploited (24) in the last relation.

### 3.3. General framework for fitting

Since theoretical estimates or experimental data comparable to those for \( \kappa \) and \( m \) are not available for the dissipation coefficients appearing in (7) and (11), we consider them as fit parameters. We recall that only two of them are independent according to (8), (9) and for the sake of convenience we will focus on \( \eta \) and \( \vartheta \).

In order to fit \( \eta \), we introduce the following transformed quantities:

\[ \hat{X}(t) := X(\alpha t), \]

(27)

\[ \hat{F}(t) := F(\alpha t), \]

(28)

\[ \hat{v} := \alpha v, \]

(29)

\[ \hat{\eta} := \eta/\alpha, \]

(30)

\[ \hat{m} := m/\alpha^2, \]

(31)

\[ \hat{\xi}(t) := \sqrt{\alpha} \xi(\alpha t) \]

(32)

with an arbitrary \( \alpha > 0 \). All other quantities appearing in (11) are understood to transform trivially. Then one readily sees that the transformed noise (32) has the same statistical properties as \( \xi(t) \) and that the dynamics (11) remains invariant under this transformation. So, if the model parameters \( \nu, \eta, \) and \( m \) are simultaneously changed in such a way that the combinations \( \nu \eta \) and \( m/\eta^2 \) remain invariant, then the transformed friction force \( \hat{F}(t) \) has exactly the same statistical properties as \( F(t\hat{\eta}/\eta) \) (provided also the initial conditions complementing (11) are transformed accordingly). Moreover, the time-averaged friction force \( \bar{F} \) from (1) remains invariant under this transformation (and is independent of the initial conditions). Hence, in a semi-logarithmic plot of \( \bar{F}(v) \) like that in figure 3, \( \eta \) can be estimated by simply shifting the theoretical curve horizontally until it optimally fits the experimental data.

In order to fit \( \vartheta \), we introduce the following transformed quantities:

\[ \tilde{X}(t) := X(t - \vartheta \eta/\kappa), \]

(33)

\[ \tilde{F}(t) := F(t - \vartheta \eta/\kappa) + \nu \vartheta \eta, \]

(34)

\[ \tilde{\vartheta} := 0. \]

(35)
All other quantities appearing in (11) are understood to transform trivially. Then one readily sees that the dynamics (11) remains invariant under this transformation (i.e., all statistical properties remain the same provided the initial conditions are also transformed appropriately). Considering \( \vartheta \) as a control parameter and indicating parametric \( \vartheta \)-dependences by means of indices \( \vartheta \), we can conclude that the friction force \( F_\vartheta(t - \vartheta \eta / \kappa) + v \vartheta \eta \) has exactly the same statistical properties as \( F_0(t) \), or equivalently,

\[
F_\vartheta(t) = F_0(t + \vartheta \eta / \kappa) - \vartheta \eta v, \tag{36}
\]

\[
\bar{F}_\vartheta = \bar{F}_0 - \vartheta \eta v, \tag{37}
\]

where the second equation follows from the first by means of (1). Hence, as far as friction forces \( F(t) \) and \( \bar{F} \) are concerned, it is sufficient to solve the dynamics (11) for \( \vartheta = 0 \) and then exploit (36) and (37), respectively. Moreover, according to (37), the value of \( \vartheta \) has practically no significance for small velocities \( v \), but is crucial for the large-\( v \) behaviour.

The above findings can be combined into the following recipe for estimating the unknown model parameters \( \eta, \vartheta, m \), and the model function \( U(x) \) in (11) by means of experimentally observed friction forces \( \bar{F}(v) \) for various pulling velocities \( v \), such as e.g. those in figure 3. First, the known parameters in (11) are fixed, e.g. according to (20)–(22). Second, one has to make some choice/guess for \( \eta, m/\eta^2 \) and \( U(x) \), e.g. according to (19). The choice of \( \vartheta \) does not matter, e.g. one may take \( \vartheta = 0 \). Finally, one numerically solves (11) for a variety of \( v \)-values and determines \( \bar{F}(v) \) according to (1). Since \( \vartheta \) has practically no significance for small \( v \)-values, \( \eta \) can be estimated by shifting the theoretical \( \bar{F}(v) \)-curve horizontally in a semi-logarithmic plot until it optimally fits the experimental data in the small \( v \) regime. In doing so, \( m \) is also implicitly subject to such variations such that the combination \( m/\eta^2 \) remains fixed to its initial value. Subsequently, (37) can be exploited to fit \( \vartheta \) in the large \( v \) regime. Finally, one has to make sure that the final \( m \)-value still satisfies (26). If the agreement between numerics and simulations is not satisfactory, one has to repeat the procedure for a different choice of \( m/\eta^2 \) and/or \( U(x) \). If no such choice leads to a satisfactory agreement then this indicates a basic problem with the model (11) or the experimental measurements.

### 3.4. Estimating inertia and friction effects

The general fitting procedure described in the previous section is illustrated in figure 8 for the specific experimental system from [11]. Hence, three parameters of the model dynamics (11) are fixed according to (20)–(22). Regarding the potential \( U(x) \), we focus within the present work on the simplest ansatz (19) with \( E_0 \) as the only remaining fit parameter.

Following the general recipe from the preceding section, figure 8 shows the result of fitting numerical solutions of the model dynamics (11) to the experimental data from [11] for \( F_N = 4 \, nN \) (omitted in figure 3). The corresponding parameter estimates are

\[
\eta = 10^{-5} \, \text{kg s}^{-1}, \tag{38}
\]

\[
E_0 = 2.5 \times 10^{-19} \, \text{J}. \tag{39}
\]

The first result, (38), is quite close to the best fit \( \eta \approx 0.8 \times 10^{-5} \, \text{kg s}^{-1} \) obtained in [20] for a similar experiment [7]. The second estimate, (39), is comparable to those from [11, 14].
Figure 8. The average lateral force (1) versus pulling velocity \(v\) for the model (11) with potential (19) and parameters (20)–(22), (38), (39). Symbols: numerical simulations for \(m = 3 \times 10^{-12}\) kg (cf equation (26)) and \(\vartheta = 0\); lines: numerical solutions of the Fokker–Planck equation [26] equivalent to (11) for \(m = 0\) and (from top to bottom) \(\vartheta = 0, 0.25, 0.5, 0.75, 1\). The numerical uncertainty is comparable to the symbol size and line width, respectively.

A further basic conclusion from the numerical results in figure 8 is that the average force \(\bar{F}\) is almost independent of the mass \(m\) below the upper bound from (26) for \(\vartheta = 0\), and hence for any \(\vartheta\) according to (37). In other words, inertia effects in (7) and (11) are very small, i.e. the stochastic dynamics (11) is very well approximated by its overdamped limit

\[
\eta \dot{X}(t) = -U'(X(t)) - F(t) - \vartheta \eta v + \sqrt{2\eta kT} \xi(t),
\]

where \(F(t) = \kappa [X(t) + vt]\). The same conclusion has also been reached very recently in the experimental work [15]. In all other previous papers addressing this issue, the opposite result of an underdamped dynamics has been found due to inadequate estimates of the relevant effective mass \(m\), as detailed in section 3.2. Nota bene, no direct experimental evidence of such an underdamped dynamics has ever been reported.

In the case of a damped harmonic oscillator, the relative importance of inertia and dissipation effects is quantified by the dimensionless quality factor \(Q = \sqrt{m U''/\eta}\), where \(m, \eta\) and \(U''\) are mass, friction, and curvature of the harmonic potential, respectively. The overdamped limit (negligible inertia) corresponds to \(Q \rightarrow 0\). The generalization of this quantifier for more general (anharmonic) potentials is not obvious. The simplest guess for the relevant potential \(U(x) + \kappa (x + vt)^2/2\) of the dynamics (11) is

\[
Q(x) = \sqrt{m [U''(x) + \kappa]/\eta}.
\]

Using (19), (21), (22), (38), (39) and the upper limit for \(m\) in (26) yields the bound \(Q(x) \leq 0.76\). However, both for deterministic relaxation processes and for thermally activated escape processes [26, 27] it is known that not the maximal curvatures of the relevant potential but rather the
curvatures at its extrema (local maxima and minima) are decisive for the relative importance of inertia and dissipation effects. Evaluating $Q(x)$ at those dominating ‘dynamical bottlenecks’ of the stochastic dynamics (11) (cf section 5) indeed leads to an estimate in agreement with our above conclusion that inertia effects are very small.

Between $v \simeq 10 \mu m s^{-1}$ and the maximal experimental velocity $v \simeq 20 \mu m s^{-1}$, a practically horizontal ‘plateau’ of $\bar{F}(v)$ has been observed in [11], see figure 3. From figure 8 we can infer that this experimental finding is only compatible with

$$\vartheta \geq 0.5.$$  \hspace{1cm} (42)

According to (8) and (9) this means that

$$\eta_c \geq \eta_s,$$  \hspace{1cm} (43)

i.e., dissipation and noise effects of the tip apex are comparable or even more important than those of the substrate, cf (7). A more precise determination of $\vartheta$ would require experimental data beyond $v \simeq 20 \mu m s^{-1}$, but in the generic case we can predict that there will not be just a ‘horizontal plateau’ but that actually $\bar{F}(v)$ exhibits a non-monotonicity.

We finally remark that similar to the variations of $F_N$ in figure 3, variations of the potential amplitude $E_0$ in (19) mainly have the effect of a ‘vertical motion’ of the $\bar{F}(v)$ curves (not shown). On the other hand, we know that varying $\eta$ amounts to a purely ‘horizontal motion’ of the $\bar{F}(v)$ curves in a semi-logarithmic plot like in figures 3 and 8. Hence it is only due to the occurrence of the plateau/maximum that both $\eta$ and $E_0$ can be estimated independently with not too large uncertainties, namely by optimally fitting both the ‘horizontal’ and ‘vertical’ positions of the plateau/maximum. While these two positions in figure 8 indeed agree quite well with the experimental results from [11], the ‘slopes’ of the $\bar{F}(v)$-curves for small $v$ still notably deviate. In order to achieve a better agreement in this respect, one has to go beyond the simple ansatz (19) for the potential $U(x)$. This issue will be addressed in detail elsewhere.

4. Non-monotonic velocity dependence of atomic friction

An increasing behaviour of the average force $\bar{F}$ as a function of the pulling velocity $v$ has been observed in numerous experimental studies, see e.g. [3]–[14], [16]. A subsequent transition into a ‘plateau’ at high velocities $v$ like in figure 3 has only been seen so far in [3, 15, 11, 16] and a consistent theoretical explanation of both features has never been given before. In particular, quite similar qualitative features have been reported in the theoretical study [21] (see figure 1 therein) in basically the same model as in (11) with $\vartheta = 0$. However, in order to achieve at least a semi-quantitative agreement with the data from [11] (cf figure 3) an unrealistically large mass $m \geq 6 \times 10^{-9}$ kg is needed (cf section 3.2). Further, an apparent quantitative agreement of the very same model with the experimental data has been reported in [11] itself. However, the plateaux in the theoretical curves from [11] are an artifact of the employed approximations in solving the dynamics (11), as a comparison with the very accurate numerical solutions in figure 8 of this paper or figure 1(d) in [20] demonstrates. Finally, the above-mentioned plateaux have been explained in [5] again by means of the same model (11) with $\vartheta = 0$ but now by additionally neglecting thermal noise effects (i.e. for $T = 0$ in (11)) in the spirit of the earlier theoretical works in this field, e.g. [17, 18]. However, such a model is unable to describe...
the increasing part of the experimental curves in figure 3. In other words, we recover once again the main conclusion of the previous section, namely that the experimental findings are only compatible with a model of the type (11) with friction coefficients satisfying (42) or, equivalently, (43).

A decreasing behaviour of $\bar{F}$ as a function of $v$ has been observed in [10, 16], however in a quite different system with significant wetting effects due to the humidity of the ambient air and a quite large contact area due to a rather large tip radius of 100 nm. Further, such a decreasing behaviour has also been seen in molecular dynamics simulations [22], however for extremely large $v$ in the range of 1–20 m s$^{-1}$ (as usual in such simulations [23]) and independent of the temperature $T$, in agreement with what one would expect in a model of the type (11) for such high velocities $v$. In other words, a non-monotonic behaviour as we predict it here (see previous section) seems to be a new effect. Comparable non-monotonicities are well-known for systems with ‘mesoscopic’ or even ‘macroscopic’ contact areas [19, 34, 42, 43], but unknown in the context of friction force microscopy.

5. Analytical approximations

In order to gain further insight into the physical mechanism responsible for the non-monotonic behaviour of $\bar{F}(v)$ in figure 8, we employ the following three steps. First, we focus on the approximative overdamped dynamics (40) as justified in the discussion of this equation in section 3.4. Second, in view of (36) we can temporarily focus on the case $\vartheta = 0$. Third, we approximate $F(t) = F_0(t)$ in (40) by its time average,

$$F_0(t) \simeq \bar{F}_0 =: f$$

due to the following argument [19]. As discussed in section 1, the motion of the tip apex $x(t)$ essentially consists of roughly periodic oscillations of $F(t) = \kappa x(t)$ about $\bar{F}$ with an amplitude of the order of $\kappa L/2$, see figure 2. For very small $\kappa$, closer inspection shows that this rough approximation in fact underestimates the amplitude but that the amplitude still approaches zero for $\kappa \to 0$. Hence, the approximation (44) becomes arbitrarily good for sufficiently small $\kappa L/\bar{F}$.

Considering that the deviation of the tip apex from its rest position $x(t)$ remains bounded for all times $t$, it follows that the average velocity, defined in accordance with (1) as $\langle \dot{x} \rangle =: \lim_{t \to \infty} x(t)/t$, vanishes and hence we can conclude from (10) that

$$\langle \dot{X} \rangle = -v.$$  

(45)

By means of the above-mentioned three steps we arrive at the following approximative dynamics (cf (40))

$$\eta \dot{X}(t) = -U'(X(t)) - f + \sqrt{2\eta kT}\xi(t).$$  

(46)

This is nothing else than an overdamped Brownian motion in the tilted periodic potential $U(X) + Xf$, whose well-known exact analytical solution for the average velocity $\langle \dot{X} \rangle$ [26, 44] in combination with (45) yields the result

$$v(f) = \frac{LkT[1 - e^{-Lf/kT}]}{\eta \int_0^L dx \int_x^{x+L} dy e^{U(x) - U(y) + (x - y)f}/kT}. $$  

(47)
Figure 9. The average lateral force (1) versus pulling velocity \( v \) for different values of \( \kappa \) in the overdamped limit \( m \to 0 \). Lower thick line, same as solid line for \( \vartheta = 0.75 \) in figure 8; dotted line, same but for the \( \kappa \)-values (from below) 2.4, 0.6, and 0.3 N m\(^{-1}\); upper solid line, analytical \( \kappa \to 0 \) asymptotics (47); dashed line, analytical approximation (50); and dashed-dotted line, analytical approximation (53), valid for \( R := (f_c - f)/\mu^{1/3}(kT)^{2/3} \gg 1 \) (see main text). Plotted is the regime \( R > 2.5 \) (i.e. \( v < 10 \mu \text{m s}^{-1} \)).

Moreover, from equations (37) and (44) one finds the following behaviour of the average force \( \bar{F}_\vartheta \), considered as a function of \( f \):

\[
\bar{F}_\vartheta(f) = f - \vartheta \eta v(f). \tag{48}
\]

Equations (47) and (48) are the main results of this section. They amount to a parametric representation of \( \bar{F}_\vartheta(v) \), valid under the two assumptions that inertia effects in (11) are negligible (cf section 3.4) and that the approximation (44) is applicable. Our first observation is that this approximative solution of \( \bar{F}_\vartheta(v) \) does not depend on \( \kappa \). According to the discussion below (44), we thus can conclude that in the overdamped limit, \( \bar{F}_\vartheta(v) \) converges for small \( \kappa \) towards the \( \kappa \)-independent limiting behaviour (47), (48). Figure 9 shows that already for experimentally realistic \( \kappa \)-values the convergence is rather good. In the opposite limit \( \kappa \to \infty \) one can infer from (7) that \( \bar{F} \to 0 \) [17]. Hence, for any fixed \( v > 0 \), a monotonically decreasing behaviour of \( \bar{F} \) as a function of \( \kappa \) is expected and indeed numerically observed in figure 9.

In the absence of noise (\( T = 0 \)), it is obvious that if the tilted periodic potential \( U(X) + Xf \) exhibits local minima then the overdamped deterministic dynamics (46) gets permanently stuck in one of these minima and hence \( v = -\langle \dot{X} \rangle = 0 \). This is the case whenever \( f \) is smaller than a certain ‘critical tilt’ \( f_c \). Beyond this threshold, deterministically running solutions with \( v = -\langle \dot{X} \rangle > 0 \) set in, converging for asymptotically large \( f \) towards \( v = -\langle \dot{X} \rangle = f/\eta \) [26] since in this limit the effect of the periodic potential \( U(X) \) becomes negligible. If now a small amount of noise is added then for \( f < f_c \) rare escapes out of the local minima of \( U(X) + Xf \) are possible,
yielding a small but finite velocity \( v = -(\dot{X}) > 0 \), which increases with increasing \( f \). For \( f > f_c \) the weak noise hardly changes the deterministically running solutions.

With this intuitive picture in mind, one readily understands that in figure 8, \( \tilde{F}_{\vartheta=0} \) is a monotonically increasing function of \( v \), approaching zero for \( v \to 0 \) and \( \eta v \) for \( v \to \infty \), and \( \tilde{F}_{\vartheta=1} = F_0 - \eta v \) is a non-monotonic function of \( v \), exhibiting a maximum near \( v(f_c) \) and approaching zero both for \( v \to 0 \) and \( v \to \infty \). The general case follows by interpolating according to (37) between \( \vartheta = 0 \) and \( \vartheta = 1 \). Hence, for small \( \vartheta \), a strictly monotonically increasing \( \tilde{F}_{\vartheta}(v) \) is obtained. Beyond a certain critical threshold of \( \vartheta \), a maximum–minimum pair of \( \tilde{F}_{\vartheta}(v) \) appears. For \( \vartheta \to 1 \) the maximum remains finite, while the location of the minimum tends towards infinity and its value towards zero, see figure 8.

A second conclusion from the above qualitative considerations is that for \( f < f_c \), corresponding to small velocities \( v \), the time-dependent friction force (4) exhibits a pronounced ‘stick–slip’ behaviour, see section 1 and figure 2. In the high-velocity regime, corresponding to \( f > f_c \), the behaviour of \( F(t) \) gets increasingly more smooth and the amplitude of the temporal variations decreases, see figures 10 and 11. Closely related experimental findings have recently been reported in [15].

Denoting by \( X_c \) one of the inflection points of the critically tilted periodic potential \( U(X) + Xf_c \), i.e.

\[
U'(X_c) = -f_c, \quad U''(X_c) = 0,
\]

our above intuitive picture implies that for weak noise and \( f \) near \( f_c \), the dynamics is dominated by the passage through the ‘dynamical bottleneck’ regions in the close vicinity of \( X_c \) modulo \( L \). As a formal counterpart of this intuitive argument one finds that the integral in (47) is dominated by \( x \)- and \( y \)-values close to \( X_c \). In this region, \( U(y) + yf \) can furthermore be approximated by \( U''(X_c)(y - X_c)^3/6 + y(f - f_c) \), yielding [44]

\[
v(f) = \frac{L\mu^{2/3}(kT)^{1/3}}{\eta} G\left(\frac{f - f_c}{\mu^{1/3}(kT)^{2/3}}\right),
\]

**Figure 10.** Same as in figure 2 but for \( v = 30 \mu m s^{-1} \).
Figure 11. Same as in figure 2 but for $v = 300 \mu m s^{-1}$.

where

$$\mu := U'''(X_c)/6,$$  \hspace{1cm} (51)

$$G(\gamma) := \sqrt{\frac{3}{4\pi}} \left[ \int_0^\infty dx e^{-x^2-x^4/4} \right]^{-1}.$$

(52)

For example, $X_c = L/2$, $f_c = \pi E_0/L$, $\mu = (\pi/L^3)(2E_0/3)$ for the potential (11). Figure 9 demonstrates that (50) indeed approximates (47) very well for all experimentally realistic velocities $v$ [11]. Since (50) has the structure of a universal scaling law, the function $\bar{F}(v)$ according to (48) will always have essentially the same ‘shape’ within a rather extended neighbourhood $f_c$.

In the subcritical regime with $f_c - f \gg \mu^{1/3}(kT)^{2/3}$ one can further simplify (50) by means of a saddle-point approximation in (52) with the result

$$v(f) = \frac{L}{\pi \eta} \sqrt{\frac{3}{\mu}} \left( f_c - f \right) \exp \left\{ -\frac{4(f_c - f)^{3/2}}{3\sqrt{3} \mu kT} \right\}.$$

(53)

Essentially, this represents a Kramers rate [27] which can also be derived by means of a saddle-point approximation directly from (47). It is reminiscent of previous predictions [11, 20, 21] of a linear dependence of $(f_c - f)$ upon $(\ln v)^{2/3}$, obtained by determining $v = -\langle \dot{X} \rangle$ from (47) with the help of a different approximation than in (47). If these predictions are adapted to our present notation and rewritten in the form (53) then an exponential factor identical to that in (53) is recovered from [8] [20] and [11] but not from [21] (possibly due to a misprint). However, the pre-exponential factors in [11, 20, 21] all differ from each other and from (53). In particular, all the prefactors in [11, 20, 21] are proportional to $\kappa^{-1}$, implying that $f$ and thus $\bar{F}$ is an increasing

\footnote{The exponential factor from [20] is $\kappa$-dependent, approaching the one in (53) for small $\kappa$, but predicting an increasing force $f$ when $\kappa$ increases and $v$ is kept fixed.}
function of $\kappa$ (with $v$ kept fixed) without a well-defined limiting behaviour for $\kappa \to 0$. Such a behaviour is in disagreement with our numerics in figure 9 and with our asymptotically exact analytical results for small and large $\kappa$ as discussed below equation (47). Figure 9 demonstrates that equation (53) approximates (47) and (50) very well within its regime of validity, which, however, ends before a non-monotonicity of $\bar{F}(v)$ may appear.

6. Outlook

The main points of our present work can be summarized as follows:

1. We provide a quantitative theory for the ‘plateaux’ of the average friction force (1) as a function of the pulling velocity $v$ as observed experimentally in [11], see figure 3.

2. We predict that upon further increasing the velocity $v$, those ‘plateaux’ will turn out to be actually maxima, see figure 8. In other words, a quite surprising decreasing behaviour of $\bar{F}$ for sufficiently large $v$ is predicted, hopefully stimulating future experiments.

3. We provide a careful discussion of inertia effects. Previously, the effective mass governing those effects was always set equal to the cantilever mass or the effective mass of cantilever eigenvibrations, leading to the erroneous conclusion that inertia effects play an important role. We show that the effective mass is actually much smaller and the dynamics is overdamped.

4. In previous theoretical models, only damping and fluctuation effects due to the substrate have been taken into account. We demonstrate that damping and fluctuation effects due to the cantilever and especially the tip apex are at least as important.

5. We put forward an analytical approximation which overcomes several of the shortcomings and limitations of previous theories.

The resulting intuitive picture is as follows. The main elastic deformations in figure 1 occur in the tip apex and in the substrate contact region. Both are accompanied by significant dissipation effects, but only quite small inertia effects. Deformations, dissipation, and inertia of the rest of the cantilever are of little importance either. For small-to-moderate velocities $v$, the tip-apex deformations are of a pronounced ‘stick-slip’ type, see figure 2. With increasing $v$, they get smoother and their amplitude decreases [15], see figures 10 and 11. The result is a reduction of the dissipation within the tip apex and—if its relative importance $\vartheta$ is sufficiently large—also of the average friction force $\bar{F}$.

Important issues which go beyond the scope of our present work include the detailed understanding of more general interaction potentials between substrate and cantilever tip than in (19). Furthermore, an analytical theory going beyond the limit of small $\kappa$ from section 5 represents a challenge of considerable practical importance and fundamental interest. The first issue is the subject of our ongoing investigations, the second has been addressed in [45].

Acknowledgments

We thank Enrico Gnecco, René Overney, and Robert Ros for enlightening discussions. This work was supported by the Alexander von Humboldt-Stiftung, the Deutsche Forschungsgemeinschaft under SFB 613 and RE 1344/3-1, and the ESF-program STOCHDYN.
References

[1] Mate C, McClelland G, Erlandsson R and Chiang S 1987 Phys. Rev. Lett. 59 1942
[2] Overney R M, Takano H, Fujihira M, Paulus W and Ringsdorf H 1994 Phys. Rev. Lett. 72 3546
[3] Gourdon D et al 1997 Tribol. Lett. 3 317
[4] Bouhacina T, Aimé J P, Gauthier S and Michel D 1997 Phys. Rev. B 56 7694
[5] Zwörner O, Hölscher H, Schwarz U D and Wiesendanger R 1998 Appl. Phys. A 66 S263
[6] Bennewitz R, Gyalog T, Guggisberg G, Bamberlin M, Meyer E and Güntherodt H-J 1999 Phys. Rev. B 60 R11301
[7] Gnecco E, Bennewitz R, Gyalog T, Loppacher Ch, Bamberlin M, Meyer E and Güntherodt H-J 2000 Phys. Rev. Lett. 84 1172
[8] Gnecco E, Bennewitz R, Gyalog T and Meyer E 2001 J. Phys: Condens. Matter 13 R619
[9] Bennewitz R, Gnecco E, Gyalog T and Meyer E 2001 Tribol. Lett. 10 51
[10] Riedo E, Lévy F and Brune H 2002 Phys. Rev. Lett. 88 185505
[11] Riedo E, Gnecco E, Bennewitz R, Meyer E and Brune H 2003 Phys. Rev. Lett. 91 084502
[12] Sills S and Overney R 2003 Phys. Rev. Lett. 91 095501
[13] Gnecco E, Bennewitz R, Socoliu C and Meyer E 2003 Wear 254 859
[14] Prioli R, Rivas A M F, Freire F L Jr and Caride A O 2003 Appl. Phys. A 76 565
[15] Socoliu C, Bennewitz R, Gnecco E and Meyer E 2004 Phys. Rev. Lett. 92 134301
[16] Riedo E and Gnecco E 2004 Nanotechnology 15 S288
[17] Tománek D, Zhong W and Thomas H 1991 Europhys. Lett. 15 887
[18] Colchero J, Baró A M and Marti O 1996 Tribol. Lett. 2 327
[19] Heslot F, Baumberger T, Perrin B, Caroli B and Caroli C 1994 Phys. Rev. E 49 4973
[20] Sang Y, Dubé M and Grant M 2001 Phys. Rev. Lett. 87 174301
[21] Dudko O K, Filippov A E, Klafter J and Urbakh M 2002 Chem. Phys. Lett. 352 499
[22] Sørensen M R, Jacobsen K W and Stoltze P 1996 Phys. Rev. B 53 2101
[23] Livihsits A I and Shluger A L 1997 Phys. Rev. B 56 12482
[24] Grabert H 1982 Projection Operator Techniques in Nonequilibrium Statistical Mechanics (Berlin: Springer)
[25] Hänggi P and Thomas H 1982 Phys. Rep. 88 207
[26] Risken H 1984 The Fokker–Planck Equation (Berlin: Springer)
[27] Hänggi P, Talkner P and Borkovec M 1990 Rev. Mod. Phys. 62 251
[28] Reimann P 2002 Phys. Rep. 361 57
[29] Lantz M A, O’Shea S J, Hoole A C F and Welland M E 1997 Appl. Phys. Lett. 70 970
[30] Lantz M A, O’Shea S J, Welland M E and Johnson K L 1997 Phys. Rev. B 55 10776
[31] Carpick R W, Ogletree D F and Salmeron M 1997 Appl. Phys. Lett. 70 1548
[32] Johnson K L and Woodhouse J 1998 Tribol. Lett. 5 155
[33] Reimann P 2001 Chem. Phys. 268 337
[34] Persson B N J, Alborn O, Mancuso F, Peveri V, Samoilov V N and Sivabaek I M 2003 Wear 254 835
[35] Müsser M 2002 Phys. Rev. Lett. 89 224301
[36] Drobek T, Stark R W and Heckl W M 2001 Phys. Rev. B 64 045401
[37] Drobek T, Stark R W, Gräber M and Heckl W M 1999 New J. Phys. 1 15
[38] Butt H-J and Jaschke M 1995 Nanotechnology 6 1
[39] Gad-el-Hak M 1999 J. Fluids Eng. 121 5
[40] Fukui K 1981 Acc. Chem. Res. 14 363
[41] Michl J 1992 Organic chemical systems, theory Encyclopedia of Physical Science and Technology vol 12 (San Diego, CA: Academic)
[42] Daudel R Quantum chemistry Encyclopedia of Physical Science and Technology vol 13 (San Diego: Academic)
[43] Schimskey-Geier L and Züllike C 1990 Z. Phys. B: Condens. Matter 79 451
[44] Liu H, Ahmed S and Scherge M 2001 Thin Solid Films 381 135
[45] Pöschel T, Schwager T and Brilliantov N V 1999 Eur. Phys. J. B 10 169
[46] Reimann P, Van den Broeck C, Linke H, Hänggi P, Rubi J M and Pérez-Madrid A 2002 Phys. Rev. E 65 031104
[47] Evstigneev M and Reimann P 2004 Europhys. Lett. 67 907

New Journal of Physics 7 (2005) 25 (http://www.njp.org/)