NanoScience and Technology

Enrico Gnecco
Ernst Meyer *Editors*

Fundamentals of Friction and Wear on the Nanoscale

Second Edition

Springer

NanoScience and Technology

The series NanoScience and Technology is focused on the fascinating nano-world, mesoscopic physics, analysis with atomic resolution, nano and quantum-effect devices, nanomechanics and atomic-scale processes. All the basic aspects and technology-oriented developments in this emerging discipline are covered by comprehensive and timely books. The series constitutes a survey of the relevant special topics, which are presented by leading experts in the field. These books will appeal to researchers, engineers, and advanced students.

More information about this series at http://www.springer.com/series/3705

Enrico Gnecco · Ernst Meyer
Editors

Fundamentals of Friction and Wear on the Nanoscale

Second Edition

Springer

Editors
Enrico Gnecco
Instituto Madrileño de Estudios Avanzados en Nanociencia
Madrid
Spain

Ernst Meyer
Department of Physics
University of Basel
Basel
Switzerland

This is a second edition. ISBN 1st ed.: 978-3-540-36806-9

ISSN 1434-4904 ISSN 2197-7127 (electronic)
ISBN 978-3-319-10559-8 ISBN 978-3-319-10560-4 (eBook)
DOI 10.1007/978-3-319-10560-4

Library of Congress Control Number: 2014952447

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The second edition of "Fundamentals of Friction and Wear on the Nanoscale" has been motivated by the significant progress made by nano tribology in the last seven years. New chapters on triboluminescence, friction in liquids, nonlinear mechanisms of friction, fractal surfaces, multiscale modeling of contacts, capillary condensation, nano manipulation in SEM, colloidal systems, graphene, nanowear of polymers, Casimir forces, and cell motility have been added. Other key chapters, such as those on atomic-scale friction in ultra-high vacuum and nano manipulation have been completely revised. On the other side, we have omitted some chapters dealing with side aspects of nano tribology which did not undergo significant changes in the last few years. We hope that this new edition will attract the interest of a broad readership of scientists and engineers, and stimulate new experiments and theoretical models in this exciting multidisciplinary research field.

Madrid — Enrico Gnecco
Basel — Ernst Meyer

Contents

Part IV Nanomanipulation

Contributors

Satish Achanta Department of MTM, Katholieke Universiteit Leuven, Leuven, Belgium

Nicolas Agraït Departamento de Física de la Materia Condensada C-III, Universidad Autónoma de Madrid, Madrid, Spain

Guillaume Anciaux Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Itay Barel Department of Chemistry and Biochemistry, University of California, Santa Barbara, USA

Roland Bennewitz INM—Leibniz-Institute for New Materials, Saarbrücken, Germany; Physics Department, Saarland University, Campus D2 2, Saarbrücken, Germany

Oleg M. Braun Institute of Physics, National Academy of Sciences of Ukraine, Kiev, Ukraine

Lorenzo Bruschi CNISM Unitá di Padova, Padova, Italy

Seymur Cahangirov Nano-Bio Spectroscopy Group, Departamento Fisica de Materiales, Centro de Fisica de Materiales CSIC-UPV/EHU-MPC and DIPC, Universidad Del Pais Vasco, San Sebastian, Spain

Rosario Capozza International School for Advanced Studies (SISSA), Via Bonomea 265, Trieste, Italy

Giuseppe Carbone CEMeC Politecnico di Bari, Bari, Italy

Jean-Pierre Celis Department of MTM, Katholieke Universiteit Leuven, Leuven, Belgium

Salim Ciraci Department of Physics, Bilkent University, Ankara, Turkey

Rogerio Colaço Department of BioEngineering and Centro de Química Estrutural, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

Maria Teresa Cuberes Dpto. Mecánica Aplicada, Universidad de Castilla-La Mancha, Almadén, Spain

Mario D'Acunto Istituto di Struttura Della Materia, Consiglio Nazionale Delle Ricerche (ISM-CNR), Roma, Italy; Istituto di Scienza E Tecnologie Dell'Informazione, Consiglio Nazionale Delle Ricerche, ISTI-CNR, Pisa, Italy; NanoICT Laboratory, Area Della Ricerca CNR, Pisa, Italy

Serena Danti Department of Surgical, Medical, Molecular Pathology and Emergency Medicine, University of Pisa, Pisa, Italy

Martin Dienwiebel Karlsruhe Institute of Technology, Institute for Applied Materials—Reliability of Systems and Components, Microtribology Center μTC, Karlsruhe, Germany

Dirk Dietzel Institute of Applied Physics (IAP), Justus-Liebig-Universität, Giessen, Germany

Franco Dinelli Istituto Nazionale di Ottica, INO-CNR, Pisa, Italy

Leonid Dorogin Institute of Physics, University of Tartu, Tartu, Estonia

Carlos Drummond Centre de Recherche Paul Pascal, CNRS-Université Bordeaux 1, Pessac, France

Mykhaylo Evstigneev Faculty of Physics, University of Bielefeld, Bielefeld, Germany

Joost W. M. Frenken Kamerlingh Onnes Laboratory, Leiden University, Leiden, The Netherlands

Enrico Gnecco Instituto Madrileño de Estudios Avanzados en Nanociencia, IMDEA Nanociencia, Campus Universitario de Cantoblanco, Madrid, Spain

Urs Gysin Institute of Physics, University of Basel, Basel, Switzerland

Haidara Hamidou Institut de Science des Matériaux de Mulhouse, Mulhouse, France

Kenichi Hiratsuka Chiba Institute of Technology, Narashino-shi, Chiba, Japan

Ilmar Kink Institute of Physics, University of Tartu, Tartu, Estonia

Marcin Kisiel Institute of Physics, University of Basel, Basel, Switzerland

Markus Langer Institute of Physics, University of Basel, Basel, Switzerland

Dong-Weon Lee MEMS and Nanotechnology Laboratory, Chonnam National University, Gwangju, South Korea

Rünno Lõhmus Institute of Physics, University of Tartu, Tartu, Estonia

Sabine Maier Department of Physics, University of Erlangen-Nürnberg, Erlangen, Germany

Nicola Manini Dipartimento di Fisica, Università degli Studi di Milano, Milano, Italy

Juan J. Mazo Departamento de Física de la Materia Condensada, Instituto de Ciencia de Materiales de Aragón, CSIC-Universidad de Zaragoza, Zaragoza, Spain

Ernst Meyer Institute of Physics, University of Basel, Basel, Switzerland

Giampaolo Mistura CNISM and Dipartimento di Fisica e Astronomia G. Galilei, Universitá di Padova, Padova, Italy

Jean-Francois Molinari Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Karine Mougin Institut de Science des Matériaux de Mulhouse, Mulhouse, France

Martin H. Müser Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany; Universität des Saarlandes, Saarbrücken, Germany

Roman Nevshupa IETCC, CSIC, Madrid, Spain

Bo Persson Peter Grünberg Institut, Forschungszentrum Jülich, Jülich, Germany

Carlos Pimentel Departamento de Cristalografía y Mineralogía, Universidad Complutense de Madrid, Instituto de Geociencias (UCM-CSIC), Madrid, Spain

Carlos M. Pina Departamento de Cristalografía y Mineralogía, Universidad Complutense de Madrid, Instituto de Geociencias (UCM-CSIC), Madrid, Spain

Pasqualantonio Pingue NEST, Scuola Normale Superiore and Istituto Nanoscienze—CNR, Pisa, Italy

Boris Polyakov Institute of Solid State Physics, University of Latvia, Riga, Latvia

Srinivasa Babu Ramisetti University of Edinburgh, Edinburgh, UK

Simon Rast Institute of Physics, University of Basel, Basel, Switzerland

Peter Reimann Faculty of Physics, University of Bielefeld, Bielefeld, Germany

Philippe Richetti Centre de Recherche Paul Pascal, CNRS-Université Bordeaux 1, Pessac, France

Juan J. Riquelme Departamento de Física de la Materia Condensada C-III, Universidad Autónoma de Madrid, Madrid, Spain

Gabino Rubio-Bollinger Departamento de Física de la Materia Condensada C-III, Universidad Autónoma de Madrid, Madrid, Spain

Ovidio Salvetti Istituto di Scienza E Tecnologie Dell'Informazione, Consiglio Nazionale Delle Ricerche, ISTI-CNR, Pisa, Italy

Vladimir N. Samoilov IFF, FZ-Jülich, Jülich, Germany

Matthias Scherge Fraunhofer IWM, Microtribology Center, Pfinztal, Germany

André Schirmeisen Institute of Applied Physics (IAP), Justus-Liebig-Universität, Giessen, Germany

Udo D. Schwarz Department of Mechanical Engineering and Materials Science and Center for Research on Structures and Phenomena (CRISP), Yale University, New Haven, CT, USA

Ion M. Sivebaek Department of Mechanical Engineering, Technical University of Denmark, Lyngby, Denmark

Ugo Tartaglino Pirelli Tires, Milan, Italy

Erio Tosatti CNR-IOM Democritos National Simulation Center, Trieste, Italy; International School for Advanced Studies (SISSA), Trieste, Italy; International Center for Theoretical Physics (ICTP), Trieste, Italy

Michael Urbakh School of Chemistry, Tel Aviv University, Tel Aviv, Israel

Andrea Vanossi CNR-IOM Democritos National Simulation Center, Trieste, Italy; International School for Advanced Studies (SISSA), Trieste, Italy

Sebastian Vieira Departamento de Física de la Materia Condensada C-III, Universidad Autónoma de Madrid, Madrid, Spain

Sergei Vlassov Institute of Physics, University of Tartu, Tartu, Estonia

Aleksandr I. Volokitin Samara State Technical University, Samara, Russia

Chunyan Yang IFF, FZ-Jülich, Jülich, Germany

Part I
Experimental Techniques

Chapter 1
Friction Force Microscopy

Roland Bennewitz

Abstract This chapter introduces Friction Force Microscopy, which is possibly the most important experimental technique in nanotribology. In spite of the apparent simplicity of this technique, a special care is required in the calibration of the force sensors, as discussed in the chapter. We will also present a few key results on the load, material and temperature dependence of friction. The chapter ends with an overview on dynamic measurements of friction, in which the probing tip is oscillated laterally while sliding in contact with the sample surface or even while translating at very close distance from it.

1.1 Introduction

Friction Force Microscopy (FFM) is a sub-field of scanning force microscopy addressing the measurement of lateral forces in small sliding contacts. In line with all scanning probe methods, the basic idea is to exploit the local interactions with a very sharp probe for obtaining microscopic information on surfaces in lateral resolution. In FFM, the apex of a sharp tip is brought into contact with a sample surface, and the lateral forces are recorded while tip and sample slide relative to each other. There are several areas of motivation to study FFM. First, the understanding of friction between sliding surfaces in general is a very complex problem due to multiple points of contact between surfaces and the importance of lubricants and third bodies in the sliding process. By reducing one surface to a single asperity, preparing a well-defined structure of the sample surface, and controlling the normal load on the contact the complexity of friction studies is greatly reduced and basic insights into the relevant processes can be obtained. Furthermore, with the decrease of the size of mechanical devices (MEMS) the friction and adhesion of small contacts becomes a technological issue. Finally, the lateral resolution allows to reveal tribological contrasts caused by material differences on heterogenous surfaces.

R. Bennewitz (✉)
INM–Leibniz Institute for New Materials, Saarbrücken, Germany
e-mail: roland.bennewitz@inm-gmbh.de

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_1

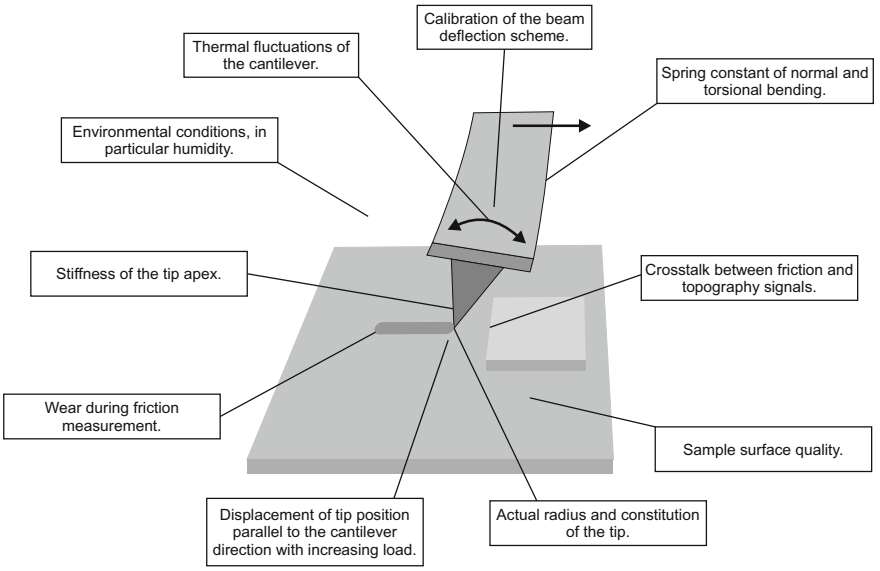


Fig. 1.1 Critical issues in experimental friction force microscopy which are discussed in this chapter

The experimental field of FFM has been pioneered by Mate et al. [1]. The group built a scanning force microscope where the lateral deflection of a tungsten wire could be measured through optical interferometry. When the etched tip of the tungsten wire slid over a graphite surface, lateral forces exhibited a modulation with the atomic periodicity of the graphite lattice. Furthermore, a essentially linear load dependence of the lateral force could be established.

In this chapter we will describe aspects of instrumentation and measurement procedures. In the course of this description, a series of critical issues in FFM will bee discussed which are summarized in Fig. 1.1.

1.2 Instrumentation

1.2.1 Force Sensors

The force sensor in the original presentation of FFM by Mate et al. was a tungsten wire [1]. Its deflection was detected by an interferometric scheme where the wire constituted one mirror of the interferometer. A similar concept was later implemented by Hirano et al., who optically detected the deflection of the tungsten wire in a Scanning Tunneling Microscope when scanning the tip in close proximity to the surface [2]. Mate and Hirano report lateral spring constants from 1.5 to 2,500 N/m, depending on the wire thickness and length. Etching the wire to form a tip at its end,

mounting the wire, aligning of the light beam, and determination of the spring constant comprise some experimental difficulties. These difficulties are greatly reduced by the use of dedicated micro-fabricated force sensors. A very sophisticated instrumental approach to the solution of those problems has been realized by Dienwiebel et al. [3]. The group has attached a stiff tungsten wire to a micro-fabricated force sensor made of silicon. The central part of the sensor is a pyramid holding the tip. The position of the pyramid is detected in all three dimensions by means of four optical interferometers directed towards the faces of the pyramid. It is suspended in four symmetric high-aspect ratio legs which serve as springs with isotropic spring constant in both lateral directions and a higher spring constant in normal direction. The symmetric design of the instrument allows for determination of normal and lateral forces acting on the tip with minimal cross talk. An overview over different experimental realizations of FFM is given in Fig. 1.2.

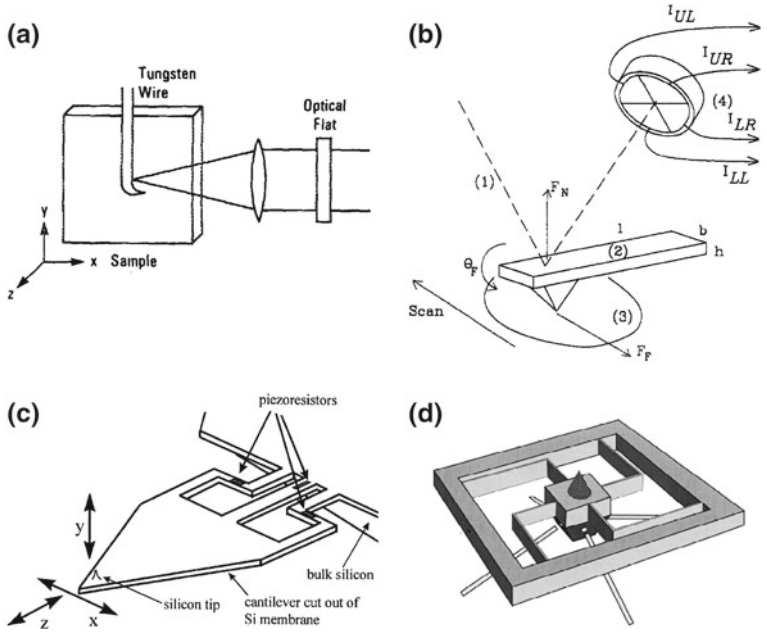


Fig. 1.2 Four design options for Friction Force Microscopy. **a** Concept of the original instrument used by Mate et al. for their pioneering experiments [1] The deflection of a tungsten wire is detected by optical interferometry. The bent end of the wire is etched into a sharp tip. **b** Beam-deflection scheme as devised by Marti et al. [5]. Normal force F_N and friction force F_F cause bending and twisting of the cantilever. The deflection of a reflected light beam is recorded by comparing currents from four sections of a photodiode. **c** Cantilever device for the measurement of lateral forces with piezoresistive detection [8]. Lateral forces acting on the tip cause a difference in stress across the piezoresistors. **d** Micro-fabricated force detector for isotropic measurements of friction forces. The block in the center holds a tungsten tip, pointing upwards in this figure. The position of the block in all three dimensions is recorded by four interferometric distance sensors which are indicated by the four light beams below the devices [9]

The most widely used form of micro-fabricated force sensors for FFM is the micro-fabricated cantilever with integrated tip. The cantilever can be either a rectangular beam or a triangular design based on two beams. The lateral force acting on the tip is detected as torsional deflection of the cantilever. This scheme has been implemented in 1990 by Meyer et al. [4] and Marti et al. [5]. It is interesting to note that the triangular design is more susceptible to deflection by lateral forces than the rectangular beam, contrary to common belief and intuition [6]. However, triangular cantilevers are less prone to the highly unwanted in-plane bending [7].

The deflection of cantilever-type force sensors is usually detected by means of a light beam reflected from the back side of the cantilever at the position of the tip. The reflected light beam is directed towards a position-sensitive photodiode which detects normal and torsional bending of the cantilever as a shift in the position of the light beam in orthogonal directions. Realistically, there is always some cross-talk between the signals for normal and torsional bending. It can be detected by exciting the cantilever to oscillate at the fundamental normal and torsional resonance and measuring the oscillation amplitude in the orthogonal channels. The cross-talk can be minimized by rotation of the position-sensitive photodiode or accounted for in the detection electronics or software. Cross-talk can transfer topographic features into the lateral force signal and create topographic artifacts from friction contrast, the latter even amplified by the feedback circuit acting on the sample height.

Calibration of the beam-deflection scheme is not a simple task, however very important in order to compare FFM results from different sources. Many publications in the past have reported on relative changes in frictional properties, without providing any calibration at all. While such relative changes certainly represent important physical findings, it is nevertheless of utmost importance to provide all experimental information available, often allowing for a rough quantitative estimate of the lateral forces. Lateral forces in FFM can easily range from piconewton to micronewton, spanning a range of very different situations in contact mechanics, and knowing at least the order of magnitude of forces helps to sort the results qualitatively into different regimes.

The calibration comprises two steps. First, the spring constant has to be determined for the force sensor. Note that the beam-deflection scheme actually determines the angular deflection of the cantilever. Nevertheless it has become custom to quantify the force constant in N/m, where the length scale refers to the lateral displacement of the tip apex relative to the unbent cantilever. Second, a relation between the deflection of the cantilever and the voltage readout of the instrument has to be established.

For the determination of the spring constant, several methods have been suggested. The easiest to calculate it from the dimensions of the cantilever. While width and thickness are easily determined by optical or electron microscopy, thickness is better deduced from the cantilever's resonance frequency. Alternatively, the spring constant can be determined from changes in the resonances caused by the addition of masses to the free end of the cantilever. Also, the analysis of a cantilever's resonance structure in air can provide the required quantities. The latter two methods have recently be described and compared by Green et al. [10]. The relation between tip displacement

and voltage readout can be established by trapping the tip in a surface structure and displacing the sample laterally by small distances. For a rough estimate one can also assume that the sensitivity of the position-sensitive photodiode is the same for normal and torsional deflection. Taking into account the geometry of the beam-deflection scheme, the torsional deflection sensitivity can be deduced from the normal deflection sensitivity (See [11] and page 352 of [12]). Since the quantification of the thermal noise driven torsional resonance can be difficult, a combination of thermal noise and beam geometry methods can be useful for the calibration of FFM [13].

A method which provides a direct calibration of the lateral force with respect to the readout voltage is the comparison with a calibrated spring standard. Recent implementations of this approach suggest as calibrated standards optical fibers [14] or micro-fabricated spring-suspended stages with spring constants that can be traced to international standards [15]. Similarly, the lateral stiffness of a magnetically levitated graphite sheet can be used as [16]. A particularly elegant method to calibrate FFM experiments is the analysis of friction loops, i.e. lateral force curves from forward and backward scans, recorded across surfaces with well-defined wedges [11, 17, 18]. Dedicated micro-fabrication design in form of a hammer-shaped cantilever can also help to calibrate the torsional bending [19].

The torsional deflection of a cantilever can in principle be detected also by optical interferometry, provided that the beam diameter is smaller than the cantilever and the point of reflection is shifted off the torsional axis [20]. However, FFM results including normal and lateral force measurements require the differential reading of multiple interferometers [3, 21].

An alternative to the detection of the cantilever bending via the beam-deflection scheme is the implementation of piezoresistive strain sensors into the cantilever. In order to measure both lateral and normal forces acting on the tip in FFM, two such strain sensors need to be realized on one sensor. Chui et al. have created a piezoresistive sensor which decouples the two degrees of freedom by attaching a normal triangular cantilever to a series of vertical ribs sensing lateral forces [22]. Gotszalk et al. have constructed a U-shaped cantilever with one piezoresistive sensor in each arm, allowing for the detection of lateral forces at the tip [23]. While the publications presenting these novel instrumental approaches contain experimental proofs of concept, no further use of piezoresistive sensors in FFM experiments has been reported. This is certainly due to a lack of commercial availability. Furthermore, the signal-to-noise ratio in static force measurements using piezoresistive cantilevers seems not to reach that of optical detection schemes.

1.2.2 Control Over the Contact

The exact knowledge of the atomic configuration in the contact between tip apex and surface is prerequisite for a complete understanding of the results in Friction Force Microscopy. It is the most severe drawback in FFM that this knowledge is not available in most cases. While sample surfaces can often be prepared with atomic

precision and cleanliness, the atomic constitution of the tip apex is usually less controlled. Friction signals vary with tip shape, as has been investigated for steps on graphite [24]. Furthermore, in the course of sliding atoms may be transferred from the tip to the surface or vice versa. Such transfer processes occur even for very gentle contact formation, as shown in experiments combining Scanning Probe Microscopy with a mass spectrometry analysis of the tip apex [25–27]. The transfer of atoms may quite often not only quantitatively but also qualitatively change the lateral forces encountered. Chemical reactions between surface and tip have been found to significantly increase friction between a Pt(111) surface for silicon but not for diamond tips [28]. The occurrence of atomic stick-slip motion can depend on the establishment of a certain degree of structural commensurability between tip and surface in the course of scanning [29, 30]. For atomic stick-slip measurements on graphite surfaces, the role of small graphite flakes attached to the tip has long been discussed and recently confirmed experimentally [1, 31].

The best control over the atomic structure of the tip apex has been achieved for metal tips in vacuum environments. By applying the established procedures of Field Ion Microscopy (FIM), the tip structure can not only be imaged but also conditioned on the atomic scale. Cross et al. have characterized the adhesion between a tungsten tip and a gold surface and proved the conservation of the atomic tip structure by means of FIM [32]. Even with instruments of lower resolution, FIM can at least be used for cleaning procedures and for a determination of the crystalline orientation of the apex cluster [2].

The integrated tips at the end of micro-fabricated silicon cantilevers have a well-defined crystalline orientation, usually pointing with the (100) direction along the tip. However, the tip surface and with it the whole tip apex are at least oxidized and possibly contaminated through packaging, transport, and handling. Furthermore, many tips are sharpened in a oxidation process which introduces large stresses at the apex. While etching in hydrofluoric acid can remove the oxide and for some time passivate silicon surface bonds by hydrogen, a stable formation and reproducible characterization comparable with FIM of metal tips has not yet been reported. Tips integrated into silicon nitride cantilevers are amorphous due to the chemical vapor deposition process and may exhibit an ever more complex structure and chemistry at the tip apex.

One way of overcoming the uncertainty of the tip constitution is to use methods of surface chemistry to functionalize the tip [33]. Specific interactions between molecules attached to the tip and molecules on the surface can be sensed by means of FFM [34]. At the same time, very strong adhesion has been reduced by covering the tip with a passivating layer to allow for lateral force imaging for example on silicon [35]. Numerous studies using this method have been published, mainly concentrating on organic monolayers on tip and surface. A review of the field has been given by Leggett et al. [36]. While most tip functionalization relies on thiol bonding to gold-coated tips, carbon bonding to nanocrystalline diamond tips has also been realized [37]. Schwarz et al. have prepared well-defined tips for FFM by deposition of carbon from residual gas molecules in a Transmission Electron Microscope, keeping control of the tip radius for a quantitative analysis of a contact mechanics

study [38]. Force measurements explicitly aiming at interactions between colloidal particles and a surface have been performed by gluing micrometer-sized spheres of the desired size to the cantilever [39, 40]. As a final note, one should always be aware of the possible occurrence of major tip wear which has been observed to happen in a concerted action of mechanical and chemical polishing [41].

1.3 Measurement Procedures

The standard measurement in FFM is the so-called friction loop: The lateral force acting on the tip is recorded for a certain distance of scanning in the direction perpendicular to the long cantilever axis and for the reverse direction. The area in the loop represents the dissipated energy, and the area divided by twice the distance is the mean lateral force. It is always very instructive to record the topography signal of forward and backward scan at the same time, as differences will reveal cross-talk between normal and torsional bending of the cantilever.

Whenever lateral forces are measured as a function of some experimental parameter, the influence of that parameter on adhesion should be studied simultaneously. In order to interpret the experimental results in terms of contact sizes versus dissipation channels the knowledge of adhesion is essential. An excellent example is the jump in lateral forces observed on a C_{60} crystal when cooling to the orientational order-disorder phase transition, which was fully explained by a change in adhesion [42]. For experiments carried out in ambient environment, the dominant contribution to adhesion are usually capillary forces which dependent greatly on the humidity and on the hydrophobicity of the surface [43]. The humidity dependence of FFM results itself can depend again on the temperature [44–46]. Consequently, an enclosure of FFM experiments for humidity control greatly enhances the reproducibility of results.

1.3.1 Friction as a Function of Load

One of the central experiments in tribology is the quantification of friction, i.e. the change of lateral force with increasing normal load on the sliding contact. One of the questions to be addressed is whether the relation between lateral and normal force is linear for FFM experiments, i.e. whether Amontons' law extends to the nanometer scale [47]. The number of FFM studies reporting lateral force as a function of load is very large, and the overall physical picture is multifaceted, to express it in a positive way. A collection of results is shown in Fig. 1.3. From a procedural point of view it is extremely important to measure the lateral forces for the full range of small normal forces until the tip jumps out of contact, usually at a negative normal force. In this way the adhesion in the system can be categorized and even maps of adhesion can be produced from friction versus load experiments [48]. Furthermore,

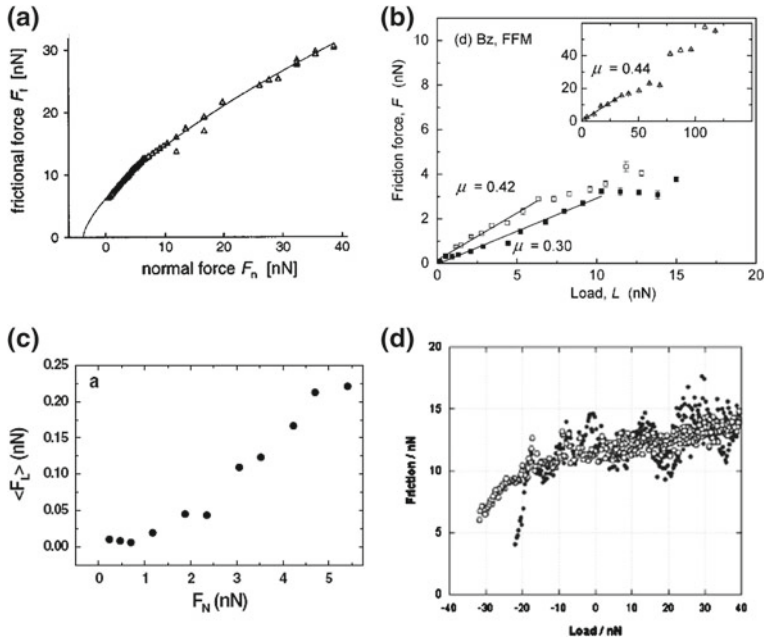


Fig. 1.3 Examples for the diversity of friction versus load curves measured by FFM. **a** Amorphous carbon measured in an argon atmosphere [38]. The sub-linear characteristic resembles the results of contact mechanics models. **b** Phenyltrichlorosilane monolayer studied in ethanol [50]. A linear dependence is found until the monolayer collapses under the tip pressure. **c** Atomic friction on NaCl(100) recorded in ultra-high vacuum [51]. A regime of vanishing friction is found for low loads. **d** Friction measurement on a hydrogen-terminated diamond surface with nanometer-scale roughness [52]. The closed circles represent the erratic load dependence of FFM results when the lateral displacement of the tip for increasing load is not compensated. The open circles show the expected sub-linear characteristic after activating the compensation

possible nonlinear characteristics at minimal loads are not overlooked. A useful way of analyzing load dependence data from FFM experiments is the representation in lateral force histograms, where for example friction on terraces and friction at steps could automatically be distinguished [49].

When the normal load on the tip is varied the position of the contact may be displaced along the long axis of the cantilever. This effect is caused by the tilt of the cantilever with respect to the surface. On heterogeneous surfaces such displacement may distort the friction measurement and, therefore, has to be compensated [52]. Another effect that can seriously disturb friction experiments is the onset of wear and the concomitant increase of lateral forces. Wear thresholds in FFM can be as low as a few nanonewton normal load, and wear at a constant low load may suddenly start after repeatedly scanning the same area [53].

1.3.2 Friction as a Function of Material

On inhomogeneous surfaces Friction Force Microscopy can image contrasts between different materials with high lateral resolution. Such contrast has been found to arise from a difference in chemical interactions between different molecular patches at the surface and the tip [54]. FFM can thus serve to identify partial coverage of a surface, for example by graphene patches [55]. As mentioned above, it is crucial to complement lateral friction contrast with local measurements of adhesion in order to elucidate whether adhesion and contact size or different channels of dissipation are dominating the contrast. Care has to be taken regarding topographical artifacts, as different materials on heterogeneous surfaces are often found at different topographic heights. Interestingly, friction contrast is also found between domains of identical molecular layers with anisotropic lateral orientation [56–58]. Friction anisotropy on a given surface has to be clearly distinguished from friction anisotropy for different azimuthal orientations between the tip and the surface. In order to measure the latter, the sample has to be rotated with respect to the tip [31, 59].

1.3.3 Friction Effects in Normal Force Measurements

When the sample is approached towards the tip, the normal force can be determined as a function of distance by measuring the normal bending of the cantilever. In all beam-deflection type FFM the cantilever is tilted with respect to the sample surface to make sure that the tip is the foremost protrusion of the force sensor. Once the tip is in contact, the tilt causes a lateral displacement of the tip position upon further approach. The friction forces arising from this lateral displacement influence the normal force measurement [40]. A detailed analysis of the process proves that one can actually perform a calibrated friction experiment through normal force versus distance curves, in particular when using extended tips like colloid probes [60]. Even when probing the surface in a dynamic intermittent contact mode these frictional contributions can be detected as a phase shift between excitation and cantilever oscillation [57].

1.3.4 Fluctuations in Friction Force Microscopy

Friction Force Microscopy is naturally subject to thermal fluctuations. Such thermal fluctuations can influence the frictional behavior of sliding contacts, as evident in the logarithmic dependence of friction on velocity at low scanning velocities [61, 62] which has been linked to thermal fluctuations via its temperature dependence [63]. Cantilever-type force sensors have a distinct resonance structure which dominates the thermal noise spectrum. A full treatment of thermal noise and mechanical vibrations and their influence on FFM have been provided in [64]. Typically, oscillations

at resonances with frequencies of several kHz are averaged out in FFM experiments. However, these resonances influence the experimental result and it is therefore very instructive to study the lateral force signal with high bandwidth [65, 66]. The statistical distribution of lateral forces in atomic stick-slip experiments can be analyzed to reveal the role of thermal fluctuations [67]. The limited scanning velocity of FFM normally separates the frequency regimes of fast fluctuations and of slower occurrence of topographic or even atomic features. The velocity limitations of FFM have been addressed by new designs combining the force sensor of an FFM with a dedicated sample stage [68, 69].

1.3.5 Friction as a Function of Temperature

The study of friction as a temperature is an obvious field of great interest. However, the number of groups including a temperature dependence into FFM studies is increasing recently [42, 44, 63, 70–74]. Thermal drift is a severe problem in the design of Friction Force Microscopes working at variable temperature, since the optical lever of the beam-deflection scheme needs to have a certain length for sensitivity. Variable-temperature instruments with thermal-expansion compensated design comparable to dedicated Scanning Tunneling Microscopes [75] have not been reported so far. One interesting approach to circumvent drift problems is the local heating of the very tip [46, 76].

1.3.6 Dynamic Lateral Force Measurements

1.3.6.1 Dynamic Friction Force Microscopy

When the sample is periodically displaced in lateral direction, the lateral force acting on the tip and detected by the cantilever will be modulated with the same periodicity. An early application of such a lateral modulation by Maivald et al. was the enhancement of contrast at step edges [77]. Dynamic Friction Force Microscopy detects the periodic lateral force signal by means of a lock-in amplifier. This idea was implemented by Göddenhenrich et al., who applied the periodic sample displacement along the long axis of the cantilever and detected the lateral force as periodic buckling of the cantilever [78]. Simultaneously, their fiber-interferometric setup could statically measure the deflection of the cantilever caused by normal forces. The same technique was implemented by Colchero et al. for a beam-deflection instrument. The authors provided a detailed analysis for the evaluation of the lateral forces when the sample is displaced in a sinusoidal movement [79]. They also pointed to the fact that using their method of Dynamic Friction Force Microscopy one will obtain quantitative results when taking data, while static experiments need subtraction of forward and backward scan before numbers can be obtained. Carpick et al. have used a similar

technique with very small sample displacement amplitudes to avoid any slip of the tip over the surface [80]. In such experiments, the amplitude of the lateral force provides a measure for the contact stiffness. Dynamic friction force microscopy has been combined with sophisticated versions of the pulsed-force mode for a simultaneous measurement of all relevant properties of mechanical contacts [81]. In a recently published study, Haugstad has analyzed the non-linear response of the lateral force to the sinusoidal sample displacement in a Fourier analysis [82]. Using this technique he was able to gain new insights into the transition from static to kinetic sliding on a polymer blend.

Dynamic Friction Force Microscopy can gain sensitivity by tuning the periodic excitation to resonances of the cantilever [83, 84]. However, the coupling between the mechanical properties of the contact and the flexural modes of the cantilever require a complex analysis, as provided in a recent review which also references previous work in the field of ultra-sonic force microscopy [85].

1.3.6.2 Dynamic Non-contact Lateral Force Experiments

The success of dynamic non-contact force microscopy in atomic resolution imaging of insulating surfaces and its prospect of measuring dissipation phenomena with the same resolution [86] has initiated projects which aim at a dynamic non-contact microscopy using lateral oscillation of the tip. Jarvis et al. have constructed a novel force sensor which allows to excite and detect oscillations of the tip in normal as well as in lateral direction [87]. The independent oscillations were achieved by suspending the tip holder in hinges at the end of two normally oscillating cantilevers. The group has controlled the tip-sample distance by changes in the normal oscillation frequency, and simultaneously recorded changes in the amplitude of the lateral oscillation pointing to frictional tip-sample interactions.

A standard rectangular cantilever has been employed by Pfeiffer et al. for the dynamic detection of interactions between a laterally oscillating tip and a surface close to but not in contact [88]. In this study, the cantilever was excited to oscillate at its first torsional resonance, making the tip oscillate laterally. The distance between tip and a copper surface was controlled using the tunneling current as feedback quantity. The lateral interaction between tip and monatomic steps or single impurities could be detected as frequency shift in the torsional oscillation. Giessibl et al. attached a tungsten tip to a quartz tuning fork such that it would oscillate laterally over the surface. Again using tunneling as feedback, they were able to study dissipation in the lateral movement with atomic resolution on a Si(111)7×7 surface, thereby tracing friction to a single atom [89]. The damping of the lateral oscillation has been explained in terms of a fast stick-slip process involving one adatom. The same surface has recently been studied in dynamic lateral force microscopy using a standard rectangular cantilever by Kawai et al. [90]. In this study a small frequency shift in the torsional resonance frequency upon approach was used to control the tip-sample distance. The torsional resonance was detected using a heterodyne interferometer scheme, where the focus of the light beam was positioned on one side of

the cantilever in order to be sensitive to the torsional bending. This is actually a very informative method to study the resonance structure of cantilevers which can show significant deviations from ideal modeling due to extra masses and asymmetries [20].

The dynamic non-contact experiments introduced in this section are very interesting tools to study conservative and dissipative interactions in lateral motion even before a repulsive contact is established. Their full strength has recently demonstrated by determination of the lateral force needed to move an atom on a surface [91] and by relating atomic structure to the anisotropy of lateral forces [92].

1.4 Outlook

Friction Force Microscopy is now a widely distributed experimental method. The experimental procedures and the calibration have been established to allow for reproducible studies of frictional properties in single-asperity contacts. The biggest drawback within the method is the lack of methods for a reproducible preparation and characterization of tips on atomic scale, as compared to the surface preparation by means of methods of Surface Science. Such control over the atomic constitution of the contact area would greatly advance our understanding of tribological processes on the nanometer scale. Other instrumental challenges in the field include the further improvement of FFM experiments at variable temperatures and in liquid environments, where atomic friction phenomena have been observed with a resolution similar to vacuum experiments [93].

References

1. C. Mate, G. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**(17), 1942 (1987)
2. M. Hirano, K. Shinjo, R. Kaneko, Y. Murata, Phys. Rev. Lett. **78**(8), 1448 (1997)
3. M. Dienwiebel, E. de Kuyper, L. Crama, J. Frenken, J. Heimberg, D.J. Spaanderman, D. van Loon, T. Zijlstra, E. van der Drift, Rev. Sci. Instr. **76**(4), 43704 (2005)
4. G. Meyer, N. Amer, Appl. Phys. Lett. **57**(20), 2089 (1990)
5. O. Marti, J. Colchero, J. Mlynek, Nanotecnology **1**(2), 141 (1990)
6. J. Sader, R. Sader, Appl. Phys. Lett. **83**(15), 3195 (2003)
7. J. Sader, C. Green, Rev. Sci. Instrum. **75**(4), 878 (2004)
8. T. Gotszalk, P. Grabiec, I. Rangelow, Ultramicroscopy **82**, 39 (2000)
9. T. Zijlstra, J. Heimberg, E. van der Drift, D. van Loon, M. Dienwiebel, L. de Groot, J. Frenken, Sens. Actuators a-Phys. **84**(1–2), 18 (2000)
10. C. Green, H. Lioe, J. Cleveland, R. Proksch, P. Mulvaney, J. Sader, Rev. Sci. Instrum. **75**(6), 1988 (2004)
11. D. Ogletree, R. Carpick, M. Salmeron, Rev. Sci. Instr. **67**(9), 3298 (1996)
12. E. Meyer, R. Overney, K. Dransfeld, T. Gyalog, *Nanoscience: Friction and Rheology on the Nanometer Scale* (World Scientific, Singapore, 1998)
13. R. Alvarez-Asencio, E. Thormann, M.W. Rutland, Rev. Sci. Instr. **84**(9), 096102 (2013)
14. N. Morel, M. Ramonda, P. Tordjeman, Appl. Phys. Lett. **86**(16), 163103 (2005)
15. P. Cumpson, J. Hedley, C. Clifford, J. Vacuum Sci. Technol. B (Microelectron. Nanometer Struct.) **23**(5), 1992 (2005)

16. Q. Li, K. Kim, A. Rydberg, Rev. Sci. Instr. **77**(6), 065105 (2006)
17. M. Varenberg, I. Etsion, G. Halperin, Rev. Sci. Instrum. **74**(7), 3362 (2003)
18. H. Wang, M.L. Gee, Ultramicroscopy **136**, 193 (2014)
19. M.G. Reitsma, R.S. Gates, L.H. Friedman, R.F. Cook, Rev. Sci. Instr. **82**(9), 093706 (2011)
20. M. Reinstaedtler, U. Rabe, V. Scherer, J.A. Turner, W. Arnold, Surf. Sci. **532–535**, 1152 (2003)
21. G. Germann, S. Cohen, G. Neubauer, G. McClelland, H. Seki, J. Appl. Phys. **73**, 163 (1993)
22. B. Chui, T. Kenny, H. Mamin, B. Terris, D. Rugar, Appl. Phys. Lett. **72**(11), 1388 (1998)
23. T. Gotszalk, P. Grabiec, I. Rangelow, Sens. Actuators, A **123–124**, 370 (2005)
24. Y. Dong, X.Z. Liu, P. Egberts, Z. Ye, R.W. Carpick, A. Martini, Tribol. Lett. **50**(1), 49 (2013)
25. U. Weierstall, J. Spence, Surf. Sci. **398**, 267 (1998)
26. T. Shimizu, J.T. Kim, H. Tokumoto, Appl. Phys. A **66**, S771 (1998)
27. A. Wetzel, A. Socoliuc, E. Meyer, R. Bennewitz, E. Gnecco, C. Gerber, Rev. Sci. Instrum. **76**(10), 103701 (2005)
28. A. Caron, D.V. Louzguine-Luzguin, R. Bennewitz, Acs Appl. Mater. Interfaces **5**(21), 11341 (2013)
29. A. Livshits, A. Shluger, Phys. Rev. B **56**, 12482 (1997)
30. R. Bennewitz, M. Bammerlin, M. Guggisberg, C. Loppacher, A. Baratoff, E. Meyer, H.J. Güntherodt, Surf. Interface Anal. **27**, 462 (1999)
31. M. Dienwiebel, G. Verhoeven, N. Pradeep, J. Frenken, J. Heimberg, H. Zandbergen, Phys. Rev. Lett. **92**(12), 126101 (2004)
32. G. Cross, A. Schirmeisen, A. Stalder, P. Grütter, M. Tschedy, U. Dürig, Phys. Rev. Lett. **80**, 4685 (1998)
33. T. Nakagawa, K. Ogawa, T. Kurumizawa, J. Vacuum Sci. Technol. B (Microelectron. Nanometer Struct.) **12**(3), 2215 (1994)
34. C. Frisbie, L. Rozsnyai, A. Noy, M. Wrighton, C. Lieber, Science **265**(5181), 2071 (1994)
35. L. Howald, R. Lüthi, E. Meyer, P. Güthner, H.J. Güntherodt, Z. Phys, B **93**, 267 (1994)
36. G. Leggett, N.B. NJ, K. Chonga. Phys. Chem. Chem. Phys. **7**, 1107 (2005)
37. M.E. Drew, A.R. Konicek, P. Jaroenapibal, R.W. Carpick, Y. Yamakoshi, J. Mater. Chem. **22**(25), 12682 (2012)
38. U. Schwarz, O. Zwörner, P. Köster, R. Wiesendanger, Phys. Rev. B **56**, 6987 (1997)
39. W. Ducker, T. Senden, R. Pashley, Nature **353**, 239 (1991)
40. J. Hoh, A. Engel, Langmuir **9**, 3310 (1993)
41. W. Maw, F. Stevens, S. Langford, J. Dickinson, J. Appl. Phys. **92**(9), 5103 (2002)
42. Q. Liang, O. Tsui, Y. Xu, H. Li, X. Xiao, Phys. Rev. Lett. **90**(14), 146102 (2003)
43. E. Riedo, F. Levy, H. Brune, Phys. Rev. Lett. **88**, 185505 (2002)
44. F. Tian, X. Xiao, M. Loy, C. Wang, C. Bai, Langmuir **15**(1), 244 (1999)
45. R. Szoszkiewicz, E. Riedo, Phys. Rev. Lett. **95**(13), 135502 (23 Sept. 2005)
46. C. Greiner, J.R. Felts, Z. Dai, W.P. King, R.W. Carpick, ACS NANO **6**(5), 4305 (2012)
47. J. Gao, W. Luedtke, D. Gourdon, M. Ruths, J. Israelachvili, U. Landman, J. Phys. Chem. B **108**(11), 3410 (2004)
48. R. Alvarez-Asencio, J. Pan, E. Thormann, M.W. Rutland, Tribol. Lett. **50**(3), 387 (2013)
49. E. Meyer, R. Lüthi, L. Howald, M. Bammerlin, M. Guggisberg, H.J. Güntherodt, J. Vac. Sci. Technol. B **14**, 1285 (1996)
50. M. Ruths, N. Alcantar, J. Israelachvili, J. Phys. Chem. B **107**(40), 11149 (2003)
51. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**(13), 134301 (2004)
52. R. Cannara, M. Brukman, R. Carpick, Rev. Sci. Instr. **76**, 53706 (2005)
53. A. Socoliuc, E. Gnecco, R. Bennewitz, E. Meyer, Phys. Rev. B (Condens. Matter Mater. Phys.) **68**(11), 115416 (2003)
54. R. Overney, E. Meyer, J. Frommer, D. Brodbeck, R. Luethi, L. Howald, H.J. Guentherodt, M. Fujihira, H. Takano, Y. Gotoh, Nature **359**(6391), 133 (1992)
55. A.J. Marsden, M. Phillips, N.R. Wilson, Nanaotechnology **24**(25), 255704 (2013)
56. M. Liley, D. Gourdon, D. Stamou, U. Meseth, T. Fischer, C. Lautz, H. Stahlberg, H. Vogel, N. Burnham, C. Duschl, Science **280**(5361), 273 (1998)
57. M. Marcus, R. Carpick, D. Sasaki, M. Eriksson, Phys. Rev. Lett. **88**(22), 226103 (2002)

58. M. Kwak, H. Shindo, Phys. Chem. Chem. Phy. **6**(1), 129 (2004)
59. H.S. Liao, B.J. Juang, W.C. Chang, W.C. Lai, K.Y. Huang, C.S. Chang, Rev. Sci. Instr. **82**(11), 113710 (2011)
60. J. Stiernstedt, M. Rutland, P. Attard, Rev. Sci. Instrum. **76**(8), 83710 (2005)
61. T. Bouhacina, J. Aime, S. Gauthier, D. Michel, Phys. Rev. B **56**, 7694 (1997)
62. E. Gnecco, R. Bennewitz, T. Gyalog, C. Loppacher, M. Bammerlin, E. Meyer, H. Güntherodt, Phys. Rev. Lett. **84**, 1172 (2000)
63. S. Sills, R. Overney, Phys. Rev. Lett. **91**, 095501 (2003)
64. A. Labuda, M. Lysy, W. Paul, Y. Miyahara, P. Gruetter, R. Bennewitz, M. Sutton, Phys. Rev. E **86**(3), 031104 (2012)
65. T. Kawagishi, A. Kato, Y. Hoshi, H. Kawakatsu, Ultramicroscopy **91**, 37 (2002)
66. S. Maier, Y. Sang, T. Filleter, M. Grant, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. B **72**, 245418 (2005)
67. A. Schirmeisen, L. Jansen, H. Fuchs, Phys. Rev. B **71**, 245403 (2005)
68. N. Tambe, B. Bhushan, Nanotechnology **16**(10), 2309 (2005)
69. E. Tocha, T. Stefanski, H. Schonherr, G. Vancso, Rev. Sci. Instrum. **76**(8), 83704 (2005)
70. X. Yang, S.S. Perry, Langmuir **19**(15), 6135 (2003)
71. R.H. Schmidt, G. Haugstad, W.L. Gladfelter, Langmuir **19**(24), 10390 (2003)
72. L. Jansen, A. Schirmeisen, J.L. Hedrick, M.A. Lantz, A. Knoll, R. Cannara, B. Gotsmann, Phys. Rev. Lett. **102**(23), 236101 (2009)
73. L. Jansen, H. Hoelscher, H. Fuchs, A. Schirmeisen, Phys. Rev. Lett. **104**(25), 256101 (2010)
74. X. Zhao, S.R. Phillpot, W.G. Sawyer, S.B. Sinnott, S.S. Perry, Phys. Rev. Lett. **102**(18), 186102 (2009)
75. M. Hoogeman, D. van Loon, R. Loos, H. Ficke, E. de Haas, J. van der Linden, H. Zeijlemaker, L. Kuipers, M. Chang, M. Klik, J. Frenken, Rev. Sci. Instrum. **69**(5), 2072 (1998)
76. B. Gotsmann, U. Durig, Langmuir **20**(4), 1495 (2004)
77. P. Maivald, H. Butt, S. Gould, C. Prater, B. Drake, J. Gurley, P. Hansma, Nanotecnology **2**, 103 (1991)
78. T. Göddenhenrich, S. Müller, C. Heiden, Rev. Sci. Instr. **65**, 2870 (1994)
79. J. Colchero, M. Luna, A. Baro, Appl. Phys. Lett. **68**, 2896 (1996)
80. R. Carpick, D. Ogletree, M. Salmeron, Appl. Phys. Lett. **70**, 1548 (1997)
81. H.U. Krotil, T. Stifter, O. Marti, Appl. Phys. Lett. **77**(23), 3857 (2000)
82. G. Haugstad, Tribol. Lett. **19**(1), 49 (2005)
83. M. Reinstadtler, U. Rabe, V. Scherer, U. Hartmann, A. Goldade, B. Bhushan, W. Arnold, Appl. Phys. Lett. **82**(16), 2604 (2003)
84. L. Huang, C. Su, Ultramicroscopy **100**(3–4), 277 (2004)
85. M. Reinstadtler, T. Kasai, U. Rabe, B. Bhushan, W. Arnold, J. Phys. D: Appl. Phys. **38**(18), 269 (2005)
86. S. Morita, R. Wiesendanger, E. Meyer, *Noncontact Atomic Force Microscopy, NanoScience And Technology* (Springer, Berlin, 2002)
87. S. Jarvis, H. Yamada, K. Kobayashi, A. Toda, H. Tokumoto, Appl. Surf. Sci. **157**, 314 (2000)
88. O. Pfeiffer, R. Bennewitz, A. Baratoff, E. Meyer, P. Gruetter, Phys. Rev. B **65**, 161403 (2002)
89. F. Giessibl, M. Herz, J. Mannhart, Proc. Natl. Acad. Sci. USA **99**(19), 12006 (2002)
90. S. Kawai, S.I. Kitamura, D. Kobayashi, H. Kawakatsu, Appl. Phys. Lett. **87**(19), 173105 (2005)
91. M. Ternes, C.P. Lutz, C.F. Hirjibehedin, F.J. Giessibl, A.J. Heinrich, Science **319**(5866), 1066 (2008)
92. A.J. Weymouth, D. Meuer, P. Mutombo, T. Wutscher, M. Ondracek, P. Jelinek, F.J. Giessibl, Phys. Rev. Lett. **111**(12), 126103 (2013)
93. A. Labuda, W. Paul, B. Pietrobon, R.B. Lennox, P.H. Gruetter, R. Bennewitz, Rev. Sci. Instr. **81**(8), 083701 (2010)

Chapter 2
Surface Forces Apparatus in Nanotribology

Carlos Drummond and Philippe Richetti

Abstract The Surface Forces Apparatus (*SFA*) has proven to be an excellent tool for research in nanotribology. It allows the study of single or multiple asperity contacts lubricated or not. The normal load, the contact area and the sliding velocity between the surfaces can be controlled and unambiguously measured with higher accuracy than in any conventional tribometer. Furthermore, an image of the surfaces in contact can be obtained as the surfaces are slid, allowing the monitoring of the real size and shape of the contact area and the distance or film thickness profile between the surfaces when atomically smooth surfaces are used. It is relatively simple to perform a comprehensive exploration of the full space of parameters to determine the important variables in the frictional behavior of the system. In this chapter the principles of operation and some experimental details of the Surface Forces Apparatus nanotribometer are described.

2.1 Introduction

The measurement of normal interaction forces between solids dates back to the 1920s, when Tomlinson investigated the interaction between crossed filaments of different metals [1]. Later, research groups in the Netherlands and Russia led by Overbeek and Derjaguin developed different techniques for measuring the force between surfaces of quartz or glass as a function of their separation [2, 3]. The example of these seminal pieces of work was promptly followed by many other groups. Particularly in Cambridge a remarkable body of work was accomplished, leading to the development of the Surface Forces Apparatus, *SFA*, by Tabor, Winterton and Israelachvili [4, 5].

The study of lateral forces between surfaces has a longer history. The problem of friction between surfaces attracted great thinkers as Da Vinci, Coulomb, Euler, Amontons and many others. A fascinating historical account of the history of

C. Drummond (✉) · P. Richetti
Centre de Recherche Paul Pascal, CNRS-Université Bordeaux 1,
Avenue Albert Schweitzer, 33600 Pessac, France
e-mail: drummond@crpp-bordeaux.cnrs.fr

P. Richetti
e-mail: richetti@crpp-bordeaux.cnrs.fr

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_2

tribology was compiled by Dowson [6]. Friction has also been investigated with *SFAs* modified for that purpose; the first friction measurements using this technique date back to the 70's [7]. A decade later Briscoe and Evans reported extensive results on the study of friction of adsorbed monolayers in air [8]. Nevertheless, it was not until the late 80's that nanotribology studies with the *SFA* became a very active field of research. Since then, various modifications to the technique have been introduced. In the present chapter we describe the principles of operations and some experimental details of the SFA-nanotribometer.

2.2 Surface Forces Apparatus Technique: Generalities

In a typical *SFA*-nanotribometry experiment molecularly smooth mica surfaces are glued to cylindrically curved silica lenses, and used to confine thin films. The use of mica as a substrate for surface force experiments was originally proposed by Bailey and Courtney-Pratt [9]. The cylindrically-shaped silica disks are placed with their axes perpendicular to each other, a configuration that presents several advantages. First, it circumvents the difficult—if not impossible—task of accurately aligning two parallel plates. Unwanted edge effects are easily avoided by this approach. Second, it allows the investigation of different contact spots on the same pair of surfaces, simply by laterally displacing the crossed cylinders. If wear or contamination of the surfaces appears during the experiment, a fresh contact zone can be readily found. Finally, this geometry is convenient for comparing the results of the measurements with theoretical descriptions, typically sketched for flat surfaces. If the separation between the curved surfaces is much smaller than their radii of curvature, R, the *SFA* cross-cylinder configuration is equivalent to a sphere-on-plate contact. The force between two such surfaces, F, can be related to the energy of interaction between flat surfaces per unit area, E, by using the so-called Derjaguin approximation [10], $E = F/2\pi R$. This provides a normalization method in order to quantitatively compare data from different experiments. The question of the normalization of the measured interaction forces is more involved in friction experiments, as will be discussed below.

One of the major strengths of the *SFA* technique rests on the possibility of imaging the area of contact to determine the distance between the surfaces, the refractive index of the film confined between them and the geometry of the contact region. The *SFA* is one of few techniques in the field of tribology that allows to image in situ and in real time the geometry of the contact area, and probably the only one with subnanometric resolution. Multiple Beam Interferometry (*MBI*) is used for this purpose [11]. A highly reflective layer is deposited on the back side of the mica surfaces, and white light is passed through this built-in Fabry-Perot interferometer. The intensity of the light transmitted through the stratified media between the two mirrors depends on the optical thickness in a nontrivial way: only wavelengths that interfere constructively after the multiple reflections in the cavity traverse the multilayer system. The emerging beam of light can then be focused on a spectrometer. The resulting constructive interference fringes (Fringes of Equal Chromatic Order, *FECO*) carry with them the

information about the thickness and the refractive index of the different layers in the path of the light. Particularly, the thickness and the refractive index of the film confined between the mica surfaces can be determined with an accuracy of 0.1 nm and 0.01 respectively. Israelachvili developed simple explicit expressions to calculate these quantities from the wavelength of the *FECO* for a film confined between mica surfaces of identical thickness [12]. Later, the analysis has been extended to asymmetric, adsorbing, anisotropic or more complicated multilayer systems [13–15]. The potential of extending the analysis to obtain information about the roughness of the surfaces has also been demonstrated [16].

Many different experimental setups for the measurement of the surface forces have been reported. One of the oldest versions, the Mk I, was designed by Israelachvili and Adams for the measurement of forces between liquids and vapours [17]. It was based on the earlier designs of Tabor, Winterton and Israelachvili [4, 5], and was later followed by greatly improved and modified versions, the Mk II and III [18, 19]. Parker and co-workers developed later a circular steel apparatus (Mk IV) which is simpler to clean and assemble than Mk I or II [20]. The stability and reliability of the apparatus, as well as the simplicity of handling, have been progressively improved on each design. Recently, Israelachvili et al. have developed the SFA 2000, a device which is simpler to assemble and operate than earlier models and integrates a number of new functionalities [21]. The interested reader is referred to the original publications for the particular details of each apparatus.

For measuring the normal force of interaction in a typical *SFA* experiment, one of the surfaces is displaced using a combination of motors and piezoelectric elements, while the other surface is coupled to a calibrated spring with a fix end. Double cantilever springs are typically used in order to minimize the tilting and/or sliding between the surfaces when the spring is deflected. The interaction force between the surfaces is measured by progressively changing the distance between the fix end of the double cantilever spring and the second surface, and allowing the separation between the surfaces to come to an equilibrium situation where the surface forces are balanced by the elastic force of the spring. The difference between the displacement carried out and the actual change in the separation between the surfaces, Δx (measured by *MBI*) will correspond to a deflection of the spring. It will be used to calculate the difference in interaction force between the two equilibrium positions (before and after the motion), ΔF, by using Hooke's law, $\Delta F = k\Delta x$, where k is the elastic constant of the double cantilever spring. Albeit being conceptually simple, measuring the forces by this procedure is limited by spring instabilities. Quickly decaying forces, with a force-distance gradient larger than the spring constant are inaccessible because of the mechanical instability of the system under such conditions. Derjaguin et al. proposed the use of a force feedback technique to overcome this problem [3]. The idea is to control the force applied to the surfaces independently of the displacement: an external force is applied to the surfaces to maintain the spring undeflected. Effectively, this translates into a continuously changing spring constant, which eliminates the mechanical instability above mentioned. Several implementations of this idea have been reported [22–24]. Steward and Parker modified a Mark IV by incorporating a magnetic force transducer and a bimorph displacement sensor. Tonck et al. introduced

a feedback apparatus with capacitive displacement transducers [25]. An interesting description of the different techniques used for the measurement of the normal force between surfaces was presented by Lodge [26].

2.3 Surface Forces Apparatus Nanotribometer

In a nanotribology experiment with the *SFA*, the mica surfaces are brought to a certain separation, T. By using motors or electromechanical transducers a lateral displacement between the surfaces is imposed, and the force induced by this displacement is measured. Usually a certain normal load is applied, L. If the load is high enough, the glue layer under the surfaces undergoes elastic deformation, and a thin film is confined to a flat circular region of uniform thickness T and area of contact A, as illustrated in the Fig. 2.1. By using *MBI* an image of the surfaces in contact can be obtained as the surfaces are slid, allowing monitoring of the size and the profile of the contact area and the distance between the surfaces. Shear-induced elastohydrodynamic deformation can also be distinguished. In addition, damage of the surfaces can be easily detected as soon as it occurs, allowing to discriminate between undamaged sliding and friction with wear, and to independently study the two scenarios.

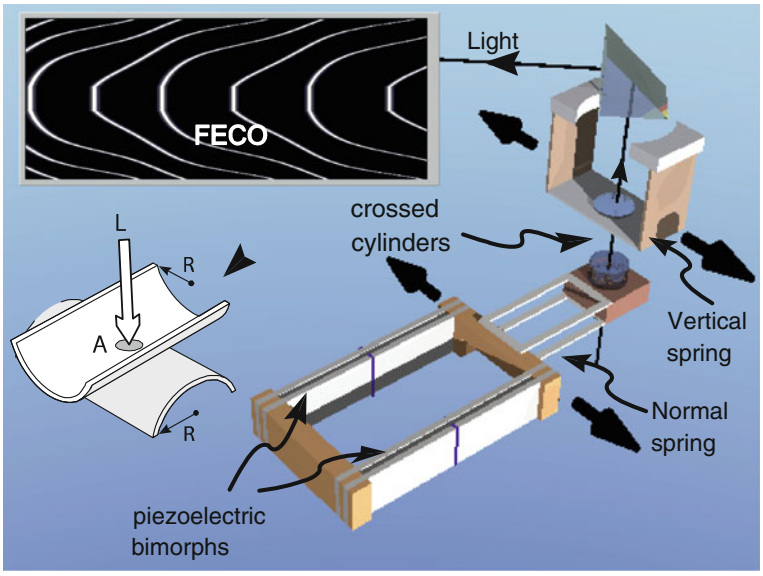


Fig. 2.1 Functional scheme of the *SFA* designed by Israelachvili et al. configured for friction experiments. The mica sheets are mounted in *crossed-cylinder* geometry, and their back surfaces are coated with reflective silver layers to allow for multiple beam interference. The *upper* and *lower* surfaces are mounted on cylindrically curved silica discs which are attached to the friction sensing device and the piezoelectric bimorph slider, respectively

A subject of major importance in the analysis of a *SFA*-nanotribology experiment is to identify the area over which the frictional force takes action. Often the friction force between sliding surfaces will be dominated by the flatten area. In that case, the sharp edge of the *FECO* allows recognizing the "area of contact", used to normalize the measured force and to calculate the shear stress. This operation is necessary to quantitatively compare the results of different experiments. From this point of view, the customary used friction coefficient is a less fundamental parameter than the shear stress. There is, however, an important caveat to this operation: very often the measured shear stress depends on the applied pressure. Given that curved surfaces are used in a *SFA* experiment, the normal pressure is not constant over the flatten area. Its value is given by a nonlinear function of the position in the contact area, a problem that has been extensively treated by the contact mechanics community [27]. It is clear then that the shear stress calculated in a *SFA* experiment is an average quantity, to be treated with caution.

An even more complicated scenario is found when there is a significant contribution to the friction force by regions of the surfaces outside the flatten area. This situation can be envisaged, for example, if there is a contribution to the friction coming from breaking and reforming bonds of long molecules that are able to bridge the two surfaces together. In that case, there is not an obvious way to identify the effective contact area. One possibility is to adopt a cut-off length, and to assume that the contribution to the frictional force is negligible at larger separations. However, at least two problems persist: the choice of the characteristic length rests somehow arbitrary and the contribution of a given region to the total force will most likely be a function of the local surface separation. It is important to emphasize at this point that the experimental difficulties just outlined are shared by most—if not all—the experimental techniques in nanotribology. Besides, the *SFA*-nanotribometer in its interferometric version is possibly the only technique in nanotribology that allows the observation of the contact geometry while rubbing the surfaces.

2.3.1 Experimental Setup

Several *SFA* experimental setups have been proposed during the last three decades, each with its own capabilities and limitations. In the following we will briefly describe few systems which are broadly used in the field. The reader interested in more complete information is referred to the original papers.

A number of experimental designs have been proposed by the group of Israelachvili [21, 28, 29]. The main features of the most recent version are illustrated in Fig. 2.1. The lower surface is mounted on a bimorph-driven slider [29], which moves laterally in a linear fashion when a constant slope voltage ramp is applied between the two electrodes of sectored piezoelectric bimorphs (electromechanical transducers). Alternatively, a constant frequency sinusoidal input can be imposed to the slider to perform nanorheological experiments. The upper surface is itself attached to a vertical double cantilever spring, whose deflection is monitored using strain gauges

connected to form the arms of a Wheatstone bridge. If the displacement of the lower surface induces a viscous or friction force on the upper surface, the vertical spring will deflect. From the deflection of this spring of known spring constant K, the friction force between the surfaces F can be calculated, simply by using Hooke's law of elasticity [28]. The mechanical properties of the measurement system (e.g., compliance and inertial mass) will influence the results; these factors have to be taken into account in order to obtain meaningful information from the signal measured. This can be done in a straightforward fashion in the *SFA* because of its mechanical simplicity and easy-to-characterize mechanical properties.

The maximal distance that can be slid with this setup depends on the characteristics of the bimorph strips used, being typically of the order of several tens of micrometers. A slider with a larger displacement range has been designed for the SFA 2000. A larger displacement can also be achieved by mechanically driving the upper surface using a reversible, variable speed motor-driver micrometer shaft that displaces the translation stage holding the vertical double cantilever spring. The detection limit for the friction force of this setup is typically of the order of several μN. By changing the frequency and the amplitude of the input signal to the bimorph slider, the driving speed can be typically varied between several Å/s to 0.1 mm/s. This device has been used to study a large number of systems. Some examples included confined simple liquids [30–32], polymer melt and solutions [28, 29, 33–35], self-assembled surfactant and polymer layers [36–38]. A number of systems in biotribology have also been extensively investigated [39, 40].

An alternative design conceived for the study of smaller deformations was introduced by Granick et al. [41, 42]. The goal of these low amplitude studies is to focus the investigation on the linear response of the confined films. By applying small deformations, the flow of fresh liquid in the contact zone is avoided. This allows the study of long time relaxation process that may be occurring in the contact region. A schematic of this device is illustrated in Fig. 2.2. In this design, the bottom surface remains stationary, while the upper surface is mounted on a holder attached to a double cantilever. Compared with the design of Israelachvili, they replaced the vertical metallic cantilevers by two piezoelectric bimorph strips. One of the bimorphs is used as an actuator and the other as a sensor. In the experiment, a voltage difference is applied to one of the bimorphs (actuator). Typically, a constant frequency sinusoidal signal is used, inducing an oscillating force on it. Simultaneously, the deformation-induced voltage of the second bimorph (sensor) is measured. This data is used to determine the actual displacement of the surface. By comparing this response with the one observed when no interaction between the two surfaces is presented, the influence of the confined film on the movement can be extracted. The electromechanical characteristics of the system are model as a series of effective masses, springs and dashpots representing the different components of the apparatus. The friction appears as a force acting on the holder of the lower surface, from which an effective viscosity can be extracted [42]. Although mainly conceived for the study of small deformations (of the order of the film thickness) typical displacements range from few nm to few

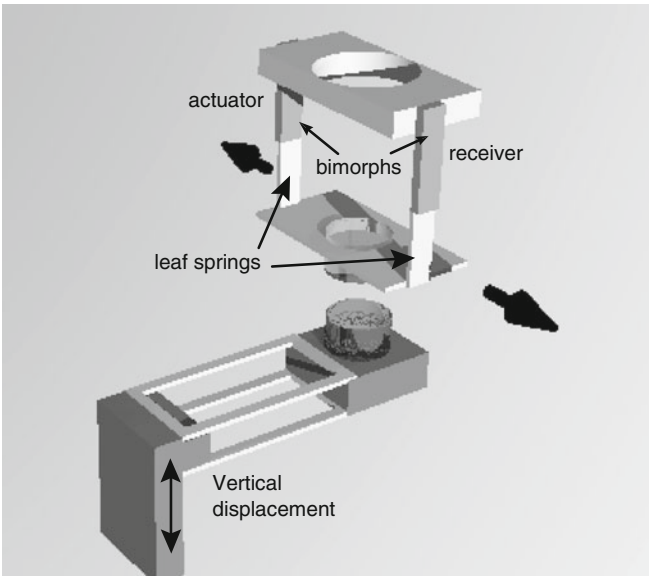


Fig. 2.2 Schematic illustration of the *SFA* designed by Granick et al. The shear force is generated by one bimorph (*actuator*) and the response of the device induces a voltage across the other bimorph (*receiver*)

μm. The reported sensitivity on the friction force is around 5 μN. Many different systems have been explored with this device, including simple liquids [41, 43–45] polymer melts [46–48] and solutions [49, 50].

A third experimental setup widely used in the literature has been developed by Klein et al. [51, 52]. A schematic of this device is presented in Fig. 2.3. In this device the sensitivity to the measured friction forces is greatly improved with respect to the previous designs. Inversely to the configuration adopted by Israelachvili et al. in the most recent version of their design the upper surface is driven and the effect on the lower surface is measured. A sectored piezoelectric tube is use to produce a normal or lateral displacement of the upper surface. An air-gap capacitor is used to measure the lateral displacement of the lower surface, which is coupled to a calibrated double cantilever spring with a fix end. The shear induced frictional force is then directly extracted from this displacement (e.g., the deflection of the spring) by using Hooke's law. The reported sensitivity of the friction force is 50 nN and the maximal displacement of the upper surface is few tens of μm. The improved sensitivity of this device comes from the detection method used. This has proven to be very valuable for the study of polymer melts and solutions [53–56] where small forces are typically observed. Research on water and other simple liquids has also been conducted with this experimental setup [57, 58]. Studies of several biotribological related systems have been recently reported [59, 60].

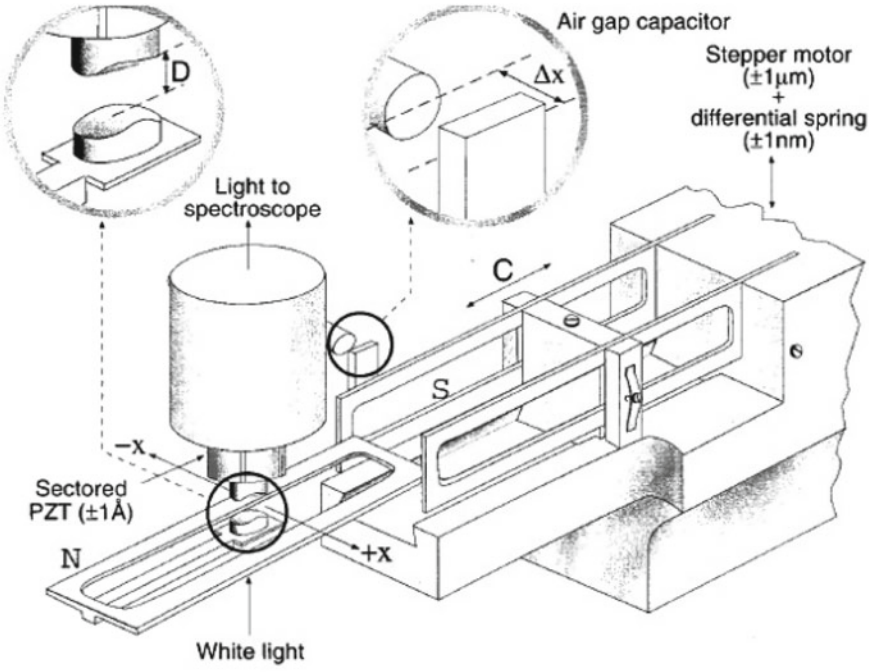


Fig. 2.3 Schematic illustration of the surface force balance (*SFB*) designed by Klein et al. The separation between the surfaces is controlled via a three-stage mechanism with a sectored piezoelectric tube on which the top surface is mounted. The piezoelectric element produces both normal and lateral displacement. The bending of the shear force spring is detected by an air-gap capacitor. Reprinted with permission from The *Journal of Physical Chemistry B* Volume **105(34)**, 8125–8134 (2001). Uri Raviv, Rafael Tadmor and Jacob Klein

As the SFA technique involves the displacement of curved surfaces, it is not a trivial matter to achieve large displacements under a constant applied load. In addition, achieving velocities greater than 100 μm/s can prove technically difficult and may involve large accelerations, due to the limited range of displacement. Bureau [61], and Israelachvili et al. [62] have addressed some of these issues. Qian et al. extended the capabilities of the apparatus to include the movement of the surfaces and the measurement of the friction force between them in two orthogonal directions in the plane of contact [63]. This new apparatus should prove to be useful in investigating shear-induced effects (e.g. shear alignment) on the confined thin films. Efforts in the same direction have also been pursued by Israelachvili et al. [21, 64].

Further improvements on different aspects of the experimental technique have been proposed during the last decade. First, substantial efforts have been devoted to automate the procedure of measurement, in order to improve the accuracy and simplicity of the technique. Second, the strategy used to determine the separation between the surfaces has been extended by using non-interferometric techniques. The automatic detection of the *FECO* have posed some challenges in the past, but increasingly

accurate and affordable charge couple devices in the market are currently used in several laboratories in the world for this purpose. Different strategies for the automation of the measurement process have been extensively described by Quon et al. [65], Grunewald and Helm [66], and more recently by Heuberger et al. [67].

As mentioned before, some efforts have also been devoted to determine the surface displacement by noninterferometric techniques. Several groups have proposed to use piezoelectric bimorphs for this purpose [68–70]. This solution is inadequate for long or quasi-static measurements due to the intrinsic drift and leakage of bimorph sensors (electrical drift and decay). A partial solution to these problems was presented by Parker [71], who suggested the use of an ultra-high impedance amplifier to lengthen the decay time of the bimorph sensor. A different method was presented by Frantz et al. [72]. They proposed to monitor the capacitance between the silver layers deposited on the back surface of the mica sheets for a fast detection of the surface separation, and described the use of this setup for the study of contact mechanics. Tonck et al. [25], and later Stewart [73], proposed to use capacitance dilatometry to measure the separation between the surfaces: one plate of a parallel-plate capacitor was attached to the moving surface, and the other to the chamber of the apparatus. This technique allows for a fast and accurate measurement of the displacement of the surface and eliminates the constraint of having to use transparent surfaces. A similar configuration, replacing one of the capacitors for a sensitive interferometric sensor to measure the deflection of the normal spring, has been used for very precise nanorheological studies [74]. Nevertheless, when opaque surfaces are used it is impossible to obtain an image of the contact region while shearing, which is one of the major strength of the SFA technique.

2.3.2 Local Structural Information: Combination of the SFA with Other Techniques

The information gathered in a conventional *SFA* experiment is limited to the average response of the confined film under shear. For achieving a better understanding of its behavior it is desirable to obtain structural information at the molecular level. Obtaining this information implies a colossal experimental challenge. On one hand the number of molecules involved in a thin film is relatively small, particularly when a localized area is explored, which inevitably reduces the intensity of any measured signal. On the other hand the investigated thin film is surrounded by layers of different materials that are susceptible to interact with the used probe (e.g. light, x-ray or neutrons) increasing the level of noise of the measured signal. Despite of these difficulties, several experimental groups have reported encouraging results of experiments combining the capabilities of the *SFA* with other techniques in-situ. It is reasonable to expect that techniques revealing the local molecular properties of the confined films will improve our understanding about the friction phenomena in the years to come.

The earliest effort in this direction aimed to combine *SFA* with x-ray diffraction (*XSFA*) [75, 76]. The second generation of *XSFA* combines the force measurement capabilities of the *SFA* with in-situ small angle x-ray scattering experiments [77]. Films of several liquid crystals have been studied with this technique, and the effect of shear on the molecular alignment in confined films has been directly evidenced. The application of this technique has so far been limited to films thicker than 500 nm, mainly because of the poor signal to noise ratio obtained otherwise. Obviously, the research in nanotribology calls for much thinner films. Moreover, the results reported with this method have been limited to an average investigation of the contact area, because of the size of the x-ray probe used. The possibility of investigating small regions of the contact area by using a micro focused x-ray beam has been suggested, but no results in this direction has been reported so far. Despite of the difficulties encountered, several research groups are actively working in this technique. It has been shown that X-ray reflectivity can be used to obtain structural information of ultra confined molecular films of OMCTS only few layers thick [78]. A recent publication describes an analogous study for water films [79]. Nevertheless, simultaneous determination of interaction forces and structural information in thin films has not been reported.

Helm et al. [80] showed that *MBI* can be used to obtain structural information of the confined thin films without any modification to the original *SFA* technique. Information about orientation and intermolecular interactions can be extracted from the *FECO* if optically active molecules are investigated. They were able to study ultra thin films, given that the light absorption by the confined molecules is enhanced by the multiple reflections in the optical cavity. Nevertheless, this technique limits the molecules that can be studied to large dye molecules. In addition, for best determination of the adsorption spectra of the confined dyes, relatively thick mica has to be used, reducing the accuracy of the film thickness determination.

In other order of ideas, Salmeron et al. suggested to couple second harmonic and sum-frequency generation to the *SFA* to study alignment and relaxation of confined ultra thin films, and showed the potential of the application by investigating self-assembled and Langmuir-Blodget monolayers of several surfactants [81]. However, results with other experimental systems have not been reported.

The combination of the *SFA* with other optical techniques has been limited by the reflective silver layer used to determine the surface separation by *MBI*. This layer strongly reduces the intensity of the illumination of the confined films, seriously limiting the in-situ performance of other optical methods. Granick et al. have overcome this limitation by replacing the reflective silver layer by multilayer dielectric coatings, which are transparent in different regions of the optical spectrum. In that way, they have been able to apply different spectroscopic tools to obtain in situ structural information of ultra thin films under shear [82, 83].

By combining the *SFA* nanotribometer with fluorescence correlation spectroscopy they measured the molecular diffusion coefficient in thin films within spots of submicron size, obtaining spatially resolved measurements [70]. This method has the drawback that fluorophore molecules have to be added to the liquid investigated in order to have a fluorescence signal. However, the authors have shown that the small

amount of fluorophores added didn't modify they behavior under shear and compression of the fluids investigated. Thus, diffusion coefficient of rhodamine in 1,2 propane diol was found to decrease by 2 orders of magnitude under confinement. Similar results were observed for the diffusion of cumarin 153 in OMCTS. They also found that the diffusion coefficient decreases from the edges towards the center of the contact region. Their results seem to suggest a heterogeneous dynamic in the confined thin films, where the diffusion appears to involve cooperative rearrangements of many molecules.

They have also reported results on the combination of *SFA* with Confocal Raman Spectroscopy [84]. The confocal geometry was used to reduced the bulk contribution to the measured signal. By using a multilayer reflective coating transparent to the argon laser and to the scattered Raman signal, they were able to monitor the geometry of the contact area simultaneously with the Raman scattering signal. They reported spatially resolved Raman scattering before and after shear, evidencing the influence of shear on the orientation of the molecules inside the confined film.

A third technique developed in Granick's group is the combination of photoluminescence and absorption dichroism with the *SFA* [85]. The shear-induced alignment of pre-adsorbed polymer molecules on mica was quantified both by photoluminescence and spectral absorption. They found molecular alignment parallel and perpendicular to the shear direction. which seems to be extremely sensitive to small changes in the initial conditions of the test. Although this technique is limited to the investigation of optical active substances, the information obtained can help to understand the behavior of lubricants with similar molecular structure.

In a different direction, Berg et al. recently suggested incorporating a Quartz Crystal Resonator in the *SFA* [86]. Because of the high oscillation frequency of the Quartz Crystal, this configuration allows the study of sliding velocities much higher than typically investigated in a conventional *SFA* nanotribology experiment. Nevertheless, in order to obtain meaningful results extremely thin mica surfaces need to be used, complicating its implementation as a routine technique. In addition, a sphere-on-plate geometry is required, which complicates the procedure of preparation of the mica surfaces.

2.3.3 Beyond Mica: Alternative Substrates

As mentioned previously, mica surfaces are the most popular substrates for *SFA* experiments. It gathers a set of properties seldom observed in other materials. It is transparent and can be prepared in the form of thin sheets of molecularly smooth surfaces over large areas by successive cleaving. The combination of these properties is at the heart of the *SFA* technique: transparent surfaces are required to determine the geometry of the contact by *MBI*. Besides, mica is a fairly incompressible material, so the forces measured are not flawed by the deformation of the surfaces. In addition it is inert to chemical reaction, so it is hardly modified during experiments. As a drawback, the process of producing mica surfaces thin enough to fulfill the requirements of the

SFA technique calls for a skillful experimentalist. This constraint is greatly relaxed by carefully implementation of automatic thickness measurement. On the other hand, thanks to the smoothness of the surfaces the geometry of the contact between the two surfaces can be easily described, simplifying the description and interpretation of the results.

The investigation of substrates other than mica is of interest for obvious reasons. The substrate plays a major role in most of the phenomena investigated by *SFA*, and particularly in tribology. It acts not only as a geometrical barrier, but as a major player: the interaction between the surfaces and with the confined films determines the general frictional behavior. For these reasons considerable efforts have been devoted to integrate different substrates in SFA experiments, to expand the range of applications of the technique. Mica surfaces can be modified by deposition or adsorption of different materials. By properly controlling the modification process, the smoothness of the surfaces can be preserved. In addition, modified surfaces may be more prone to chemical modifications.

Several groups have investigated the behavior of mica surfaces modified by self-assembly [37, 38, 87, 88] or deposition of Langmuir-blodgett films [37, 89, 90] of different substances. Mica acquires a negative surface charge when immersed in water, so positively charged species (e.g. cationic surfactants) spontaneously adsorbed on it; the structure of the adsorbed layer and its relationship to the molecular structure of the adsorbed material has been a very active area of research during the last 25 years [91]. The frictional behavior of the modified surfaces depends strongly on the characteristic of the adsorbed layers: surface properties like the adhesion energy and the morphology of the adsorbed layer will ultimately determine their behavior under shear.

As mentioned before, mica is an inert material. Nevertheless, it can be chemically modified by water vapor plasma treatment, increasing their reactivity to different species, e.g. chlorosilanes, as suggested by Parker et al. [92]. In this way, molecularly smooth hydrophobic surfaces can be prepared, given that the chemical structure is modified without increasing the roughness of the substrates. Mica surfaces treated by this procedure have been used in SFA studies [93]. Kessel and Granick modified this procedure to be able to induce the self-assembly of alkoxysilanes on mica, showing that strongly bound monolayers were formed [94].

Several groups have proposed to modify the mica surfaces simply by depositing on them thin films of different materials, including metals and dielectrics. In order to be able to monitor the geometry of the contact region by *MBI*, it is important for the deposited layers not to be completely opaque. This does not impose a serious limitation for sufficiently thin films. The interpretation of the *FECO* becomes more involved because of the larger number of optical layers in the optical path of the white light, but the information about the thickness and optical properties of the confined film can nonetheless be extracted. Different algorithms which are adequate for the modified experimental conditions have been described in the literature [95, 96].

Studies of mica modification by deposition of many different materials have been reported in the past. Silver [95, 97], gold [95, 98], platinum [98, 99], silica [99, 100], are only a few of a long list of materials investigated. Horn et al. grown single crystals

of aluminum oxide [101] by vapor phase condensation. The tribological behavior of these surfaces was later investigated by Berman et al. [102]. Vigil et al. deposited smooth layers of amorphous silica on mica, and study the behavior under compression and shear of the resulting surfaces [100]. They found that oscillatory structural forces were absent of the interaction between the surfaces due to the increased roughness. In addition, they observed long time-dependent adhesion and friction of the surfaces in the presence of water. Mc Guiggan et al. deposited amorphous carbon by magnetron sputtering on mica, and used these surfaces in the *SFA*-nanotribometer [103]. They found the friction force to be proportional to the area of contact between the surfaces, and the measured shear stress to decrease strongly with increasing relative humidity. Hirz et al. sputtered thin films of zirconia and alumina on mica, and investigated the behavior of these surfaces when lubricated with a linear perfluoropolyether [104]. They showed that these metal oxide formed smooth films on mica susceptible of being used as alternative substrates in *SFA* experiments.

Other groups have proposed to simplify the method of substrate preparation by eliminating the use of mica all together. A method of preparing silica surfaces for use in the *SFA* was proposed by Horn et al. although its use has not became widespread [105]. Golan et al. proposed to deposit a thin layer of silicon nitride on rigid silica disks previously coated with a reflective layer to replace the mica substrates [106]. They also reported a succinct tribological study of this generic substrate. Chain and Klein proposed to use mica as a template to produce extremely smooth gold surfaces [107, 108]. This method has been adapted by other groups to study electrochemical processes using the SFA [109, 110].

In general, the surface modification processes abovementioned may alter the smoothness of the surfaces at some degree complicating the geometry of the system, changing it from a single-asperity to a multiple-asperity contact. However, they allow the investigation of surfaces of interest in many different fields, extending the range of applications of the *SFA*. In addition, in most of the cases the roughness of the deposited layers can be controlled and/or modified to certain extent, allowing the investigation of the effect of surface roughness on friction, an important field of research on its own. *SFA* studies involving controlled roughness are in progress in several laboratories in the world [109].

2.4 Case Study: Weakly Adhesive Surfaces Under Shear

To illustrate the potential of the *SFA* technique for nanotribology studies, some experimental results obtained with self-assembled surfactant layers are described in this section.

We investigated the following system: the mica surfaces in the *SFA* are immersed in bulk aqueous surfactant solutions. Cationic surfactants are chosen, so that self assembled layers are formed on the mica surfaces. For surfactant concentrations above the critical micelle concentration (cmc) the adsorbed films show different morphology depending on the surfactant. Some surfactants adsorb as flat bilayers, while others

form rather modulated layers, suggesting the adsorption of globular or cylindrical micelles [91]. If two flat bilayers are compressed, eventually the hemifusion of the layers can be induced. In the hemifused region the mica surfaces end up cover by a monolayer of surfactant, and the surfaces are held together by an adhesive interaction, because of the hydrophobic attraction between the hydrophobic chains of the surfactant molecules. The precise measurement of the thickness of the trapped layer allows the clear identification of the hemifusion; an abrupt change of the confined film, corresponding to the expulsion of two monolayers from the contact region, is induced by compression and/or shear [38, 87, 111].

The behavior under shear of these systems is very complex. In general, when two intact bilayers are sheared, we do not detect any frictional resistance at any applied velocity or normal load: the friction force is below the detection limit of our experimental setup, which is similar to the one designed by Israelachvili [28, 29]. On the contrary, after the hemifusion of the layers is induced, a higher friction force can be observed. A typical friction trace measured during the hemifusion process is presented in Fig. 2.4, together with friction traces measured at different driving velocities after the hemifusion has taken place. The general behavior of the measured friction force with the driving velocity after hemifusion is illustrated in Fig. 2.5. At least 5 different regimes can be identified. At low velocities smooth sliding is observed. The force increases first linearly and then logarithmically with the driving velocity, before reaching a plateau. Above a certain critical velocity the movement becomes unstable and stick-slip is observed. At even higher velocities the movement becomes again stable and a second smooth sliding regime is observed, when the frictional resistance increases linearly with the driving velocity. As can be observed in the Fig. 2.5, an extensive dynamic regime is necessary to be able to observe the five regimes just described. This exploration of the space of parameters can be readily performed with the *SFA* nanotribometer.

Within the experimental accuracy, the shear stress, defined as $\sigma = F/A$, appears to be independent of the normal load L over the range of load investigated, both along the plateau regime preceding the stick-slip instability, and for the high velocity smooth-sliding regime. This implies that the friction force is proportional to the contact area A, rather than to the contact diameter or the load L. The load independence of the shear stress is no longer verified along the logarithmic regime. It is only due to the possibility of monitoring the real area of contact with the *SFA* (from the flat region on the *FECO*) that the shear stress can be univocally calculated at all times during the experiment.

All the trends observed in the sliding curve can be described by a model originally proposed by Schallamach [112] and that we have extensively discussed in the past [38]. The general behavior of the friction force can then be interpreted in terms of a model based on the kinetics of formation and rupture of small adhesive links (bonds) between the two shearing surfaces. Under this scenario, the observed stick-slip regime is just a manifestation of the mechanical instability due to the negative slope of the force versus velocity curve in a certain range of speeds. This adhesive model is insufficient to account for the steady smooth sliding regime observed at

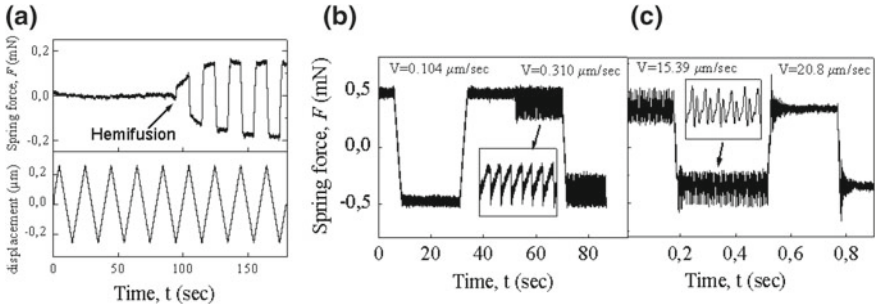


Fig. 2.4 **a** Friction signal recorded when the lower surface is displaced at constant velocity in a reciprocate mode, at the moment of the shear-induced hemifusion. A dramatic increase in friction force is accompanied by a film thickness reduction from 6.5 to 3.5 nm, indicating the hemifusion of the adsorbed bilayers. The normal load remains practically constant. **b** A smooth sliding regime is observed at low velocities, $V < V_c$. In the stick-slip regime the friction force oscillates between the kinetic value F_k and a *lower* kinetic value F_{sk}. **c** Increasing the driving velocity the measured spring force changes from an oscillatory state to a smooth steady state. Every time the driving velocity is reversed there is a transient response of few hundredth of a second before the system reaches steady-state sliding

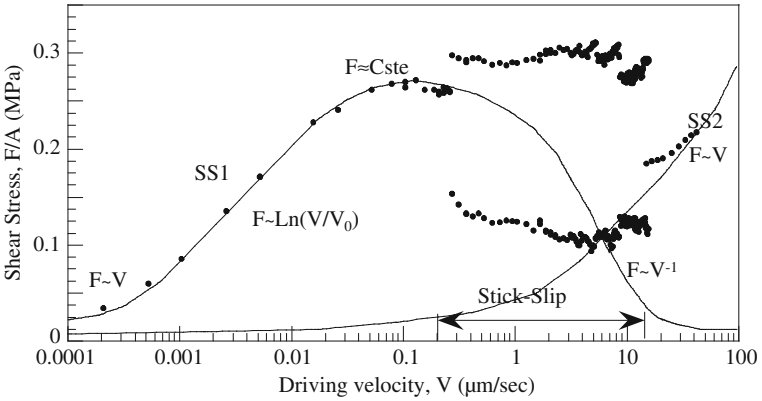


Fig. 2.5 Driving-velocity dependence of the shear stress measured while shearing two adsorbed monolayers of the 12-3-12-3-12 surfactant under a load of $L = 4.51$ mN at $T = 20$ C. The smooth sliding to stick-slip transition occurs at $V_c \approx 0.3$ μm/s. Prior to the transition, the kinetic stress σ_k levels off at V_1 after a logarithmic σ-V dependence. The quasi-smooth regime persists up to the transition at V_c. At high driving velocities a new transition to a smooth-sliding regime is observed

high velocities. A second contribution to the friction force (other than the elastic contribution) must be considered in order to re-stabilize the mechanical system in a kinetic state with finite friction. This extra contribution may be, for instance, the viscous dissipation in the trapped layer. The linear increase of the force at high sliding velocity seems to support this idea [38, 111].

This example illustrates the capabilities of the *SFA* nanotribometer. It shows how the measurement of the parameters pertinent to the friction problem, in particular the capability to monitor the geometry of the rubbing surfaces, greatly improves the understanding of the phenomena involved, allowing a quantitative comparison of the behavior of the system with theoretical models.

References

1. G.A. Tomlinson, Philos. Mag. **6**, 695–712 (1928)
2. J.T.G. Overbeek, M.J. Sparnaay, Discuss. Faraday Soc. **18**, 12 (1954)
3. Y.I. Rabinovich, B.V. Derjaguin, N.V. Churaev, Adv. Colloid Interface Sci. **16**, 63–78 (1982)
4. D. Tabor, R.H.S. Winterton, Proc. R. Soc. A **312**, 435–450 (1969)
5. J.N. Israelachvili, D. Tabor, Proc. R. Soc. A **331**, 19–38 (1972)
6. D. Dowson, *History of Tribology*, 2nd edn. (Professional Engineering Publishing Limited, London, 1998)
7. J.N. Israelachvili, D. Tabor, Wear **24**, 386–390 (1973)
8. B.J. Briscoe, D.C.B. Evans, Proc. R. Soc. A **380**, 389–407 (1982)
9. A.I. Bailey, J.S. Courtney-Pratt, Proc. R. Soc. A **227**, 500–515 (1955)
10. B. Derjaguin, Kolloid-Zeitschrift **69**, 155–164 (1934)
11. S. Tolansky, *Multiple Beam Interferometry of Surfaces and Films* (University Press, Oxford, 1948)
12. J.N. Israelachvili, J. Colloid Interface Sci. **44**, 259–272 (1973)
13. R.G. Horn, D.T. Smith, Appl. Opt. **30**, 59–65 (1991)
14. C. Mueller, P. Maechtle, C.A. Helm, J. Phys. Chem. **98**, 11119–11125 (1994)
15. M. Heuberger, Rev. Sci. Instrum. **72**, 1700–1707 (2001)
16. M. Heuberger, G. Luengo, J. Israelachvili, Langmuir **13**, 3839–3848 (1997)
17. J.N. Israelachvili, G.E. Adams, J. Chem. Soc., Faraday Trans. **1**(74), 975–1001 (1978)
18. J. Israelachvili, Proc. Natl. Acad. Sci. U. S. A. **84**, 4722–4724 (1987)
19. J.N. Israelachvili, P.M. McGuiggan, J. Mater. Res. **5**, 2223–2231 (1990)
20. J.L. Parker, H.K. Christenson, B.W. Ninham, Rev. Sci. Instrum. **60**, 3135–3138 (1989)
21. J. Israelachvili, Y. Min, M. Akbulut, A. Alig, G. Carver, W. Greene, K. Kristiansen, E. Meyer, N. Pesika, K. Rosenberg, H. Zeng, Rep. Prog. Phys. **73**, 036601 (2010)
22. J. Parker, A. Stewart, Prog. Colloid Polym. Sci. **88**, 162–168 (1992)
23. A.M. Stewart, J.L. Parker, Rev. Sci. Instrum. **63**, 5626 (1992)
24. W.H. Briscoe, R.G. Horn, Langmuir **18**, 3945–3956 (2002)
25. A. Tonck, J.M. Georges, J.L. Loubet, J. Colloid Interface Sci. **126**, 150–163 (1988)
26. K.G. Lodge, Adv. Colloid Interface Sci. **19**, 27–73 (1983)
27. K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985)
28. A.M. Homola, J.N. Israelachvili, M.L. Gee, P.M. McGuiggan, J. Tribol. **111**, 675 (1989)
29. G. Luengo, F.-J. Schmitt, R. Hill, J. Israelachvili, Macromolecules **30**, 2482–2494 (1997)
30. M.L. Gee, P.M. McGuiggan, J.N. Israelachvili, A.M. Homola, J. Chem. Phys. **93**, 1895–1906 (1990)
31. J. Israelachvili, P. McGuiggan, M. Gee, A. Homola, M. Robbins, P. Thompson, J. Phys.: Condens. Matter **2**, SA89–SA98 (1990)
32. C. Drummond, J. Israelachvili, Phys. Rev. E **63**, 041506 (2001)
33. G. Luengo, J. Israelachvili, A. Dhinojwala, S. Granick, Wear **200**, 328–335 (1996)
34. P.A. Schorr, T.C.B. Kwan, S.M. Kilbey, E.S.G. Shaqfeh, M. Tirrell, Macromolecules **36**, 389–398 (2003)
35. G. Luengo, M. Heuberger, J. Israelachvili, J. Phys. Chem. B **104**, 7944–7950 (2000)
36. H. Yoshizawa, C. You-Lung, J. Israelachvili, Wear **168**, 161–166 (1993)

37. H. Yoshizawa, Y.L. Chen, J. Israelachvili, J. Phys. Chem. **97**, 4128–4140 (1993)
38. C. Drummond, J. Israelachvili, P. Richetti, Phys. Rev. E **67**, 066110 (2003)
39. B. Zappone, M. Ruths, G.W. Greene, G.D. Jay, J.N. Israelachvili, Biophys. J. **92**, 1693–708 (2007)
40. G.W. Greene, X. Banquy, D.W. Lee, D.D. Lowrey, J. Yu, J.N. Israelachvili, Proc. Natl. Acad. Sci. U. S. A. **108**, 5255–5259 (2011)
41. J. Van Alsten, S. Granick, Phys. Rev. Lett. **61**, 2570–2573 (1988)
42. J. Peachey, J. Van Alsten, S. Granick, Rev. Sci. Instrum. **62**, 463–473 (1991)
43. H. Hu, G. Carson, S. Granick, Phys. Rev. Lett. **66**, 2758–2761 (1991)
44. Y. Zhu, S. Granick, Phys. Rev. Lett. **87**, 096104 (2001)
45. A.L. Demirel, S. Granick, J. Chem. Phys. **115**, 1498 (2001)
46. S. Granick, H.-W. Hu, Langmuir **10**, 3857–3866 (1994)
47. S. Granick, H.-W. Hu, G.A. Carson, Langmuir **10**, 3867–3873 (1994)
48. J. Van Alsten, S. Granick, Macromolecules **23**, 4856–4862 (1990)
49. Y. Zhu, S. Granick, Macromolecules **36**, 973–976 (2003)
50. M. Ruths, S.A. Sukhishvili, S. Granick, J. Phys. Chem. B **105**, 6202–6210 (2001)
51. J. Klein, D. Perahia, S. Warburg, Nature **352**, 143–145 (1991)
52. U. Raviv, R. Tadmor, J. Klein, J. Phys. Chem. B **105**, 8125–8134 (2001)
53. R. Tadmor, J. Janik, J. Klein, Phys. Rev. Lett. **91**, 115503 (2003)
54. U. Raviv, S. Giasson, N. Kampf, J.-F. Gohy, R. Jérôme, J. Klein, Nature **425**, 163–165 (2003)
55. E. Eiser, J. Klein, T. Witten, L. Fetters, Phys. Rev. Lett. **82**, 5076–5079 (1999)
56. J. Klein, E. Kumacheva, D. Perahia, D. Mahalu, S. Warburg, Faraday Discuss. **98**, 173–188 (1994)
57. J. Klein, E. Kumacheva, J. Chem. Phys. **108**, 6996–7009 (1998)
58. U. Raviv, P. Laurat, J. Klein, Nature **413**, 51–54 (2001)
59. J. Seror, Y. Merkher, N. Kampf, L. Collinson, A.J. Day, A. Maroudas, J. Klein, Biomacromolecules **12**, 3432–3443 (2011)
60. J. Seror, Y. Merkher, N. Kampf, L. Collinson, A.J. Day, A. Maroudas, J. Klein, Biomacromolecules **13**, 3823–3832 (2012)
61. L. Bureau, Rev. Sci. Instrum. **78**, 065110 (2007)
62. D.D. Lowrey, K. Tasaka, J.H. Kindt, X. Banquy, N. Belman, Y. Min, N.S. Pesika, G. Mordukhovich, J.N. Israelachvili, Tribol. Lett. **42**, 117–127 (2011)
63. L. Qian, G. Luengo, D. Douillet, M. Charlot, X. Dollat, E. Perez, Rev. Sci. Instrum. **72**, 4171–4177 (2001)
64. E. Charrault, X. Banquy, K. Kristiansen, J. Israelachvili, S. Giasson, Tribol. Lett. **50**, 421–430 (2013)
65. R.A. Quon, J.M. Levins, T.K. Vanderlick, J. Colloid Interface Sci. **171**, 474–482 (1995)
66. T. Grünewald, C.A. Helm, Langmuir **12**, 3885–3890 (1996)
67. M. Zaüch, J. Vanicek, M. Heuberger, Rev. Sci. Instrum. **74**, 260–266 (2003)
68. J. Van Alsten, S. Granick, Tribol. Trans. **33**, 436–446 (1990)
69. J.N. Israelachvili, S.J. Kott, L.J. Fetters, J. Polym. Sci., Part B: Polym. Phys. **27**, 489–502 (1989)
70. A. Dhinojwala, S. Granick, J. Chem. Soc., Faraday Trans. **92**, 619 (1996)
71. J.L. Parker, Langmuir **8**, 551–556 (1992)
72. P. Frantz, N. Agrait, M. Salmeron, Langmuir **12**, 3289–3294 (1996)
73. A.M. Stewart, Meas. Sci. Technol. **11**, 298–304 (2000)
74. F. Restagno, J. Crassous, E. Charlaix, M. Monchanin, Meas. Sci. Technol. **12**, 16–22 (2001)
75. S. Idziak, I. Koltover, J. Israelachvili, C. Safinya, Phys. Rev. Lett. **76**, 1477–1480 (1996)
76. S.H. Idziak, C.R. Safinya, R.S. Hill, K.E. Kraiser, M. Ruths, H.E. Warriner, S. Steinberg, K.S. Liang, J.N. Israelachvili, Science **264**, 1915–1918 (1994)
77. Y. Golan, A. Martin-Herranz, Y. Li, C. Safinya, J. Israelachvili, Phys. Rev. Lett. **86**, 1263–1266 (2001)
78. O.H. Seeck, H. Kim, D.R. Lee, D. Shu, I.D. Kaendler, J.K. Basu, S.K. Sinha, Europhys. Lett. **60**, 376–382 (2002)

79. S. Chodankar, E. Perret, K. Nygård, O. Bunk, D.K. Satapathy, R.M. Espinosa Marzal, T.E. Balmer, M. Heuberger, J.F. van der Veen, Europhys. Lett. **99**, 26001 (2012)
80. P. Mächtle, C. Müller, C.A. Helm, J. Phys. II **4**, 481–500 (1994)
81. P. Frantz, F. Wolf, X.-D. Xiao, Y. Chen, S. Bosch, M. Salmeron, Rev. Sci. Instrum. **68**, 2499–2504 (1997)
82. A. Mukhopadhyay, J. Zhao, S.C. Bae, S. Granick, Rev. Sci. Instrum. **74**, 3067–3072 (2003)
83. S.C. Bae, J.S. Wong, M. Kim, S. Jiang, L. Hong, S. Granick, Philos. Trans. R. Soc., A **366**, 1443–154 (2008)
84. S.C. Bae, H. Lee, Z. Lin, S. Granick, Langmuir **21**, 5685–5688 (2005)
85. S.C. Bae, Z. Lin, S. Granick, Macromolecules **38**, 9275–9279 (2005)
86. S. Berg, M. Ruths, D. Johannsmann, Phys. Rev. E **65**, 026119 (2002)
87. P. Richetti, C. Drummond, J. Israelachvili, M. In, R. Zana, Europhys. Lett. **55**, 653–659 (2001)
88. M. Ruths, N.A. Alcantar, J.N. Israelachvili, J. Phys. Chem. B **107**, 11149–11157 (2003)
89. C.A. Helm, J.N. Israelachvili, P.M. McGuiggan, Biochemistry **31**, 1794–805 (1992)
90. S. Yamada, J. Israelachvili, J. Phys. Chem. B **102**, 234–244 (1998)
91. G.G. Warr, Curr. Opin. Colloid Interface Sci. **5**, 88–94 (2000)
92. J.L. Parker, D.L. Cho, P.M. Claesson, J. Phys. Chem. **93**, 6121–6125 (1989)
93. J.L. Parker, P.M. Claesson, D.L. Cho, A. Ahlberg, J. Tidblad, E. Blomberg, J. Colloid Interface Sci. **134**, 449–458 (1990)
94. C.R. Kessel, S. Granick, Langmuir **7**, 532–538 (1991)
95. J.M. Levins, T.K. Vanderlick, Langmuir **10**, 2389–2394 (1994)
96. M.T. Clarkson, J. Phys. D: Appl. Phys. **22**, 475–482 (1989)
97. J.L. Parker, H.K. Christenson, J. Chem. Phys. **88**, 8013–8014 (1988)
98. C.P. Smith, M. Maeda, L. Atanasoska, H.S. White, D.J. McClure, J. Phys. Chem. **92**, 199–205 (1988)
99. N.A. Alcantar, C. Park, J.-M. Pan, J.N. Israelachvili, Acta Mater. **51**, 31–47 (2003)
100. G. Vigil, Z. Xu, S. Steinberg, J. Israelachvili, J. Colloid Interface Sci. **165**, 367–385 (1994)
101. R.G. Horn, D.R. Clarke, M.T. Clarkson, J. Mater. Res. **3**, 413–416 (1988)
102. A. Berman, S. Steinberg, S. Campbell, A. Ulman, J. Israelachvili, Tribol. Lett. **4**, 43–48 (1998)
103. P.M. McGuiggan, S.M. Hsu, W. Fong, D. Bogy, C.S. Bhatia, J. Tribol. **124**, 239–244 (2002)
104. S.J. Hirz, A.M. Homola, G. Hadziioannou, C.W. Frank, Langmuir **8**, 328–333 (1992)
105. R.G. Horn, D.T. Smith, W. Haller, Chem. Phys. Lett. **162**, 404–408 (1989)
106. Y. Golan, N.A. Alcantar, T.L. Kuhl, J. Israelachvili, Langmuir **16**, 6955–6960 (2000)
107. L. Chai, J. Klein, Langmuir **23**, 7777–7783 (2007)
108. L. Chai, J. Klein, Langmuir **25**, 11533–11540 (2009)
109. M. Valtiner, X. Banquy, K. Kristiansen, G.W. Greene, J.N. Israelachvili, Langmuir **28**, 13080–13093 (2012)
110. T. Kamijo, M. Kasuya, M. Mizukami, K. Kurihara, Chem. Lett. **40**, 674–675 (2011)
111. C. Drummond, J. Elezgaray, P. Richetti, Europhys. Lett. **58**, 503–509 (2002)
112. A. Schallamach, Wear **17**, 301–312 (1971)

Chapter 3
Nanoscale Friction and Ultrasonics

Maria Teresa Cuberes

Abstract The chapter describes different procedures to monitor ultrasonic vibration at a sample surface using an AFM cantilever tip. Both the excitation of normal and shear surface ultrasonic vibration are considered. The possibility to reduce and eliminate friction at nanometer-sized contacts by means of ultrasonic vibration is discussed. Experiments that provide information about nanoscale adhesion hysteresis, and its relationship to friction, are described in detail. The ability of Phase—Heterodyne Force Microscopy to resolve tiny differences in adhesion hysteresis with high sensitivity is remarked.

3.1 Introduction

Ultrasonic technology finds many applications in our society. It is used in chemistry, biology and medicine, i.e. for preparation of colloids or emulsions, the pregermination of seeds, for imaging of biological tissues, etc. Also, in non-destructive testing (NDT), for measurement of materials properties, in metrology, etc. Ultrasonic vibrations are commonly employed in mechanical machining of materials [1]. Procedures such as ultrasonic cutting of metals, ultrasonically-assisted wire-drawing, ultrasonically-assisted drilling, etc. take advantage of a modification of friction by ultrasonic vibration. Macroscopically, it is well-known that friction and acoustics are very much related [2]. The development of *nanoscale ultrasonics* can be of interest in nanotechnology. Nevertheless, studies related to the emission of ultrasound from nanoscale contacts or to the influence of ultrasonic vibrations on nanofriction are still scarce [3].

The investigation of friction at the nanometer scale can be realized with an Atomic Force Microscope (AFM). A specific AFM-mode, Friction Force Microscopy (FFM), has been developed to this purpose [4]. FFM monitors the torsion of a microcantilever as a sample is laterally displaced by means of piezoelectric actuators, being

M.T. Cuberes (✉)
Dpto. Mecánica Aplicada, Universidad de Castilla-La Mancha,
Plaza Manuel de Meca 1, 13400 Almadén, Spain
e-mail: Teresa.Cuberes@uclm.es

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_3

the cantilever tip in contact with the sample surface. Typically, the deformation of the cantilever is sensed by optical beam deflection, and both bending in normal direction and torsion are simultaneously recorded with a four-quadrant photodiode detector [5]. The measurement of the lateral forces that act upon the tip-sample contact during forward and backward scans allows us to distinguish frictional forces, which reverse when reversing the scanning direction, from the lateral forces that stem from topographical features. The lateral resolution in FFM depends on the tip-sample contact area, which is typically of 10–100 nm in diameter, in ambient conditions.

Ultrasound refers to mechanical vibrations of frequencies ranging from 20 KHz up to GHz. Typical ultrasound propagation velocities in solid materials are of the order of 10^3 ms^{-1}. Hence, ultrasonic wavelengths in solid materials are of the order of mm, much larger than the diameter of the mean tip-sample contact area. Actuation of ultrasonic vibration at a nanocontact will always be accomplished in the "near-field" regime. Understanding of whether it is possible to detect ultrasonic vibration at the contact of an AFM cantilever tip and a sample surface is not trivial at first sight. A cantilever tip in contact with a surface will certainly be subjected to forces when the surface atoms displace due to ultrasound excitation, but if the ultrasonic frequency is sufficiently high, considering the cantilever tip as a point mass, it is clear that it will not be able to follow the surface motion due to its inertia.

Starting from 1992, different procedures to monitor ultrasonic vibrations at a sample surface using an AFM cantilever tip have been explored, which will be described in this chapter [6–23]. A first motivation for most of those studies was to implement a near-field approach that provided the kind of information that is obtained with the Acoustic Microscope, i.e. information about the elasticity and viscoelasticity of materials, but with a lateral resolution on the nanometer scale. To this aim, different AFM-based techniques such as Ultrasonic Force Microscopy (UFM) [7, 9], Atomic Force Acoustic Microscopy (AFAM) [10], and Heterodyne Force Microscopy (HFM) [21] have been quite successfully implemented. The different methods and their main opportunities for the characterization of nanoscale materials properties will be briefly outlined in Sect. 3.2.

Shear ultrasonic vibration excited at a sample surface can also be detected with the tip of an AFM cantilever [24–36]. Experiments that monitor the cantilever response to shear ultrasonic vibration excited at the tip-sample interface, being the tip in contact with the sample surface, provide novel methods to study nanoscale friction. Some interesting results concerning the response of nanocontacts to shear ultrasonic vibration will be introduced in Sect. 3.3.

In Sect. 3.4, experimental evidence of the reduction and/or elimination of friction at nanometer-sized contacts by means of ultrasonic vibration will be considered. The opportunity to control friction at a nanometer scale is of tremendous significance in nanotechnology. By now, it has been unambiguously demonstrated that ultrasound of sufficiently high amplitude can act as a lubricant in nanoscale contacts [37–40]. Nevertheless, only a few experiments that address this topic have been performed up to date, and hence the opportunities of ultrasonic vibration to modify the mechanisms of friction at a nanometer scale are still an open question.

In Sect. 3.5, some attempts to obtain information about adhesion and/or the adhesion hysteresis using ultrasonic AFM techniques will be summarized [21, 41–47]. Procedures for the measurement of adhesion hysteresis from UFM have been investigated, and a relationship between adhesion hysteresis and friction has been formally established [44]. Phase-HFM provides information about dynamic relaxation processes related to adhesion hysteresis nanoscale contacts with an extremely high time sensitivity, superior to any other ultrasonic-AFM procedure [21]. In view of a comparison of phase-HFM and friction data, the opportunities to take advantage of the time resolution of HFM for the study of nanoscale friction processes will be discussed

3.2 Normal Ultrasonic Vibration at Nanocontacts

In the following, we will consider the nanocontact formed by the tip of an AFM cantilever in contact with a sample surface. Normal ultrasonic vibrations at the tip-sample interface can be excited using, for instance, an appropriate piezoelectric element attached to the back of the sample; longitudinal acoustic waves originated by mechanical vibrations of the piezo will propagate through the sample, and reach the surface-tip contact area.

As indicated in the introduction, in the limit of high ultrasonic frequencies (100 MHz for instance), it is not expected that the cantilever tip in contact with the sample surface can move fast enough to keep up with surface atomic vibrations at ultrasonic frequencies, due to its inertia. Nevertheless, the displacement of the surface atoms will lead to modification of the tip-sample interaction forces. In the absence of ultrasound, being the tip in contact with the sample surface, in the repulsive interaction force regime, the cantilever is bent to compensate for the sample surface repulsive interactions, so that the net force at the tip-sample interface is zero, and the tip is indented into the sample to a certain extent, which depends on both the cantilever and the tip-sample contact stiffness. In the presence of normal ultrasonic vibration the tip-sample distance is varied at ultrasonic frequencies between minimum and maximum values, which depend upon the amplitude of ultrasound excitation and the initial set-point force (see Fig. 3.1a). If the amplitude of ultrasound is small, the tip-sample distance sweeps a linear part of the tip-sample interaction force curve. The net average force that acts upon the cantilever during an ultrasonic time period will be in this case the initial set-point force. However, if the amplitude of ultrasound is increased, and the tip-sample distance is swept over the nonlinear part of the force curve, the average force will then include an additional force. If the ultrasonic amplitude is sufficiently high, the cantilever experiences an additional displacement due to force, which can be easily detected with the optical lever technique [7]. This additional force constitutes the so-called*ultrasonic force* and it is the physical parameter evaluated in *Ultrasonic Force Microscopy* (UFM) [7, 9]. The ultrasonic force induces a “static” cantilever displacement (UFM signal) as long as vertical ultrasonic vibration of sufficiently high amplitude is present at the

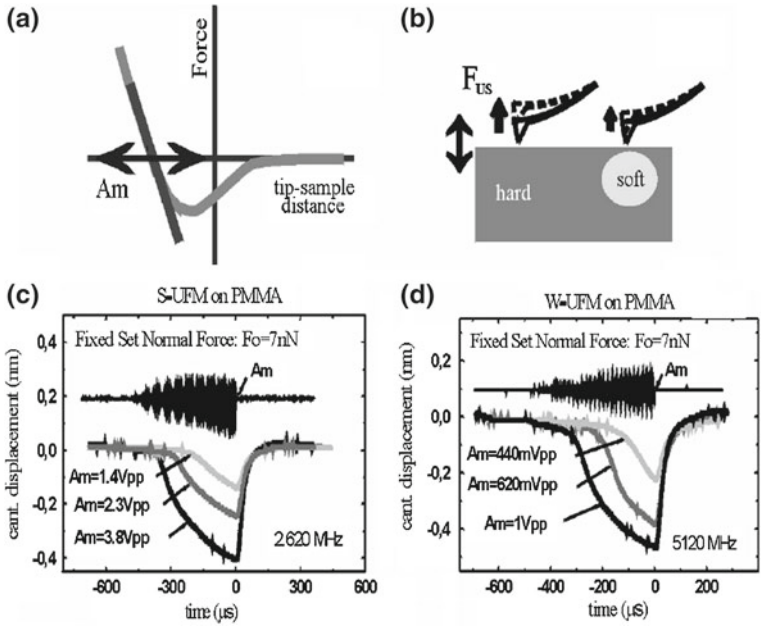


Fig. 3.1 **a**, **b** The physical principle of UFM measurements (see text). The ultrasonic excitation may be introduced through the sample (S-UFM), **c** or through the tip using the cantilever as a waveguide (W-UFM) (**d**). The piezo excitation is give a triangular modulation, with maximum amplitude Am. The effect of varying the static force Fo (set-point force) is similar for S-UFM and W-UFM (from [22])

tip-sample contact. In this sense, the cantilever behaves a "mechanical diode", and UFM has also received the name of mechanical diode mode.

The ultrasonic force is hence understood as the averaged force experienced by the tip during each ultrasonic period. Its magnitude depends upon the part of the tip-sample force regime over which the tip-sample distance varies while being modulated at ultrasonic frequencies, i.e. on the initial tip-sample distance (the initial indentation or set-point force) and on the ultrasonic amplitude. The ultrasonic response will be dependent on the details of the tip-sample interaction force, and hence on sample materials properties such as local elasticity and adhesion. Figure 3.1a, b illustrate the physical principle of the UFM measurements. Softer surface or near-surface regions of nanoscale dimensions at the sample under consideration will be easily distinguished from harder regions because of a smaller UFM signal at the former (Fig. 3.1b). In Fig. 3.1c, d, UFM responses of a sample of poly (methylmethacrylate) about 3 mm thick are shown ([see [22] for more details about these measurements]. As shown in the Fig., the piezo excitation is given a triangular modulation, with maximum amplitude Am. In (c), the piezo is located at the back of the sample, and works at a frequency of 2.620 MHz (the way ultrasound is excited at the tip-sample contact in (d) will be discussed below). The set-point force is kept constant

at 7 nN. UFM responses for different maximum ultrasonic amplitudes are shown. As it is noticeable from the figure, the UFM response is zero until the amplitude of ultrasound excitation reaches a threshold value, and it then increases as the ultrasonic amplitude is increased. If the ultrasonic excitation amplitude is periodically varied at some low KHz frequencies, the UFM response will change accordingly, and by monitoring its magnitude at every surface point by means of a lock-in amplifier, UFM images can be measured. To date, it has already been demonstrated that UFM is a useful technique to map the nanoscale elasticity and adhesive properties of surface and subsurface regions in a variety of both stiff and compliant samples [9, 19].

When working in the UFM mode, the high-frequency cantilever vibration is not directly monitored. If the cantilever is regarded as a simple point-mass, the amplitude of vibration at the driving frequency should vanish in the limit of very high frequencies [7]. Nevertheless, the cantilever is not a point mass, but a tiny elastic beam that can support high-frequency resonant modes. Atomic Acoustic Force Microscopy (AFAM) [10, 13] monitors the resonance frequencies of the high-order bending modes of the cantilever, being the tip of an AFM cantilever in contact with the sample surface, in the presence of normal ultrasonic vibration at the tip-surface interface. According to the wave theory of elastic beams, the flexural resonance frequencies of a rectangular cantilever are the solutions of a fourth-order differential equation, which can be analytically solved for a clamped-free cantilever, and for a clamped spring-coupled cantilever with the tip in contact in contact with a sample surface [13]. In the latter case, the resonances are shifted in frequency and the vibration amplitudes along the cantilever changes. Using a linear approximation for the tip-sample interaction forces, the frequency shift can be calculated. Figure 3.2 shows the resonance frequencies of the clamped spring-coupled cantilever as a function of

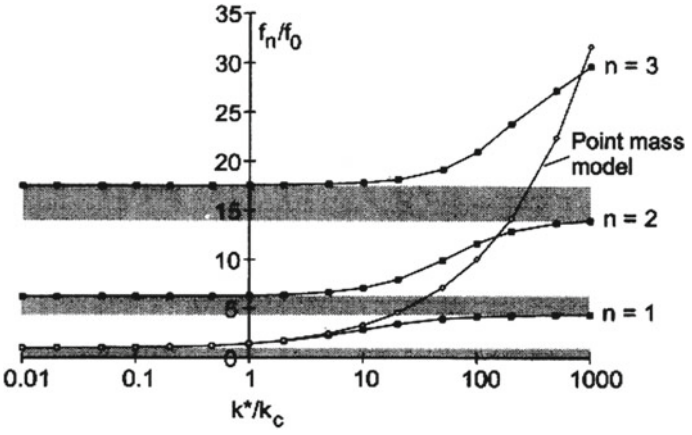


Fig. 3.2 Resonance frequencies f_n of the clamped spring-coupled cantilever with the tip in contact with a sample surface (*black squares*) normalized to the first resonance frequency of the clamped-free cantilever f_o. K^* and K_c are the tip-sample contact stiffness and the cantilever stiffness, respectively. A comparison with the point-mass model for the cantilever (*open circles*) shows that this model predicts too large frequency shifts for $K^*/K_c > 1$ (from [13])

the stiffness of the tip-sample contact normalized to the cantilever stiffness for the first three modes. The experimental determination of the shift of the resonance frequencies of the high-order flexural cantilever modes provides a measurement of the tip-sample contact stiffness, with lateral resolution in the nanometer scale. From the contact stiffness, the sample indentation modulus can be derived using, for instance, Hertz contact theory [13].

In UFM, it is assumed that the cantilever is *dynamically frozen*, and does not vibrate at ultrasonic frequencies [7]. Even though cantilever resonant modes can certainly be excited at a microcantilever, the point-mass picture for the AFM cantilever tip allows us to understand certain peculiarities of its high-frequency dynamic behaviour. Thus, the inertia of the cantilever "explains" that in ultrasonic-AFM techniques soft cantilevers can indent hard samples, and yield information about surface and subsurface elastic inhomogeneities. In the limit of high ultrasonic frequencies, the amplitude of vibration at the crests of the resonant modes of a clamped spring-coupled cantilever is expected to be very small, and extremely difficult, if possible, to detect. Hence, UFM appears as the most appropriate technique for measurements at higher ultrasonic frequencies. Typically, in AFAM, the tip-sample distance is kept sufficiently small that the tip-sample interactions remain in the linear regime. In contrast, UFM relies on the nonlinearity of the tip-sample interaction force; if the tip-sample interactions are in the linear regime, no ultrasonic force is expected to set off at the tip-sample contact.

The detection of surface ultrasonic vibration with the tip of an AFM cantilever was first demonstrated in [6] by exciting Surface Acoustic Waves (SAWs) at slightly different frequencies, and using a cantilever tip in contact with the sample surface to detect the surface vibration at the difference frequency. SAWs are acoustic modes that are confined within a wavelength to the surface of a solid, and propagate along specific crystalline directions. They can be excited using interdigital transducers (IDTs) on appropriate substrates. Scanning Acoustic Force Microscopy (SAFM) is particularly implemented for the characterization of SAWs field amplitudes [11] and phase velocities [18]. The procedure in SAFM is actually equivalent to this in UFM: the superposition of two SAWs of slightly different frequencies leads to surface high frequency vibration that is modulated in amplitude at the lower difference frequency. When the surface vibration amplitude is sufficiently high, a cantilever tip detects the rectified signal via the mechanical diode effect, due to the nonlinearity of the tip-sample force curve.

In *Scanning Local Acceleration Microscopy* (SLAM) [14], the cantilever tip is considered a point mass. Three different working modes are distinguished: the "contact mode", the "mechanical diode" mode and the "subharmonic" mode. In "contact mode" SLAM, the sample is vibrated at high frequency, being the tip in contact with the sample surface, and the tip displacement, which yields the contact stiffness, is monitored at the excitation frequency; the high-frequency surface vibration amplitude is kept sufficiently low that the tip-sample interaction remains in the linear regime. The "mechanical diode" SLAM mode is equivalent to UFM. The "subharmonic" SLAM mode proposes that the sample surface is excited at very high ultrasonic vibration amplitudes; according to interesting reported data [12], the

analysis of the generation of subharmonics and chaos may provide information about the local coefficient of restitution of a tip bouncing on a sample surface.

Scanning Microdeformation Microscopy (SMM) [8] uses a piezoelectric element to both excite ultrasonic vibration at a sample, and detect the acoustic wave generated by the microdeformations caused by a tip in contact with a sample surface. The technique can operate in "transmission mode", with the piezo located at the back of the sample. In this way, contrast of local elastic constants, inhomogeneities and/or subsurface features is obtained with a lateral resolution essentially related to the tip diameter.

It is worth to remark at this stage that most of the different ultrasonic-AFM approaches discussed so far have capabilities of *subsurface imaging* [8, 9, 14]. Nevertheless, so far the resolved buried feature sizes are typically much smaller than the used acoustic waves, the sensitivity to subsurface features does not appear "straightforwardly" related to acoustic wave propagation, but rather to a "near-field" effect.

The development of AFAM has proved that in the presence of ultrasound, being the tip in contact with a sample surface, flexural resonant modes are excited at typical AFM cantilevers at frequencies of some MHz. Nevertheless, UFM usually also works quite well in the frequency range of some MHz. In principle, the ultrasonic frequency selected for UFM measurements should not be coincident with the cantilever contact resonances in order that the high-frequency displacements of the tip are as small as possible. However, it has additionally been demonstrated that ultrasound can be excited at a sample surface from a piezoelement located at the cantilever base. In this case, the cantilever acts as an acoustic waveguide that propagates the ultrasonic signal to the sample. As in AFAM, the measurement of the amplitude and resonant frequency of the high-order resonances of a cantilever in contact with the sample surface when ultrasound is excited from the cantilever base provides information of the sample elasticity with nanoscale resolution [15, 16]. SMM has also been implemented in the so-called "reflexion mode", with a piezoelement located at the cantilever base which is used for both the excitation and the detection of ultrasound [17]. And even though the propagation of ultrasound from the cantilever base to the sample surface necessarily requires that the cantilever tip vibrates at the excitation frequency, it has been experimentally demonstrated that UFM works in this configuration, renamed as Waveguide-UFM (W-UFM) for distinction. similarly as in the case that ultrasound is excited at the tip-sample contact from the back of the sample (Sample-UFM, S-UFM) [22, 23]. In W-UFM, the ultrasonic excitation is input at the tip-sample contact via tip displacements. W-UFM and S-UFM signals recorded on PMMA can be compared in Fig. 3.1c, d. In Fig. 3.1d, a piezo located at the cantilever base is excited at 5.120 MHz. As it is apparent from the Fig., both procedures lead to remarkably similar qualitative responses. In principle, excitation of ultrasound from the cantilever base in ultrasonic-AFM techniques is potentially advantageous as there are many fewer restrictions on the sample shape or its internal structure (e.g. porous or hollow samples can be studied). In addition, the use of same piezo-cantilever-tip assembly for different samples simplifies a quantitative comparison of nanoscale mechanical data.

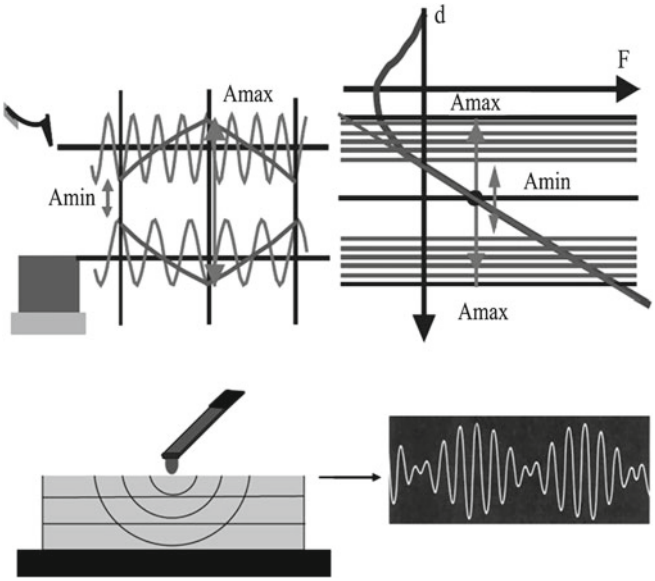


Fig. 3.3 A schematic diagram illustrating HFM. Small phase-delays between tip and sample vibration (at ω_1 and ω_2 respectively) will cause a phase variation of the cantilever vibration at the difference frequency $\omega_1 - \omega_2$. This is detected as the HFM response

In *Heterodyne Force Microscopy* (HFM) [21], ultrasound is excited both at the tip (from a transducer at the cantilever base) and at the sample surface (from a transducer at the back of the sample) at adjacent frequencies, and mixed at the tip-sample gap (see Fig. 3.3). The physical principle of HFM is described in Fig. 3.3. As the sample vibrates at a frequency ω_1 and the tip at a frequency ω_2, the maximum tip-sample distance, is modulated at $\omega_1 - \omega_2$ (beat frequency). Provided that the total amplitude is large enough to cover the nonlinear range of the tip-sample interaction force, an ultrasonic force (stronger for larger amplitudes) will act upon the cantilever and displace it from its initial position. Owing to the varying ultrasonic force, the cantilever vibrates at the difference mixed frequency. In HFM, this vibration is monitored in amplitude and phase with a lock-in amplifier, using the (externally) electronically mixed signal as a reference. The information provided by the Amplitude-HFM (A-HFM) response is very similar to that obtained by UFM. Nanoscale lateral variations in sample elasticity and/or adhesive properties will give rise to A-HFM contrast. A unique feature of HFM is its ability to monitor phase shifts between tip and sample ultrasonic vibrations with an extremely high temporal sensitivity, i.e. fractions of an ultrasonic time period. Small differences in the sample dynamic viscoelastic and/or adhesive response to the tip interaction result in a shift in phase of the beat signal that is easily monitored in phase-HFM (ph-HFM). In this way, HFM makes it possible to study dynamic relaxation processes in nanometre volumes with a time-sensitivity of nanoseconds.

Recently, *Scanning Near-Field Ultrasound Holography* (SNFUH) [23] has been proposed as a non-destructive imaging method. The technique is implemented in a similar way as HFM, save that here the difference frequency is chosen in the range of hundreds of KHz whereas in [21] difference frequencies of some KHz are used. The experimental data obtained by SNFUH demonstrate its capability to provide elastic information of buried features with great sensitivity. Interestingly, in Phase-HFM most of the contrast apparently stems from surface effects, as will be discussed in Sect. 3.5 of this chapter.

3.3 Shear Ultrasonic Vibration at Nanocontacts

If we consider the nanocontact formed by the tip of an AFM cantilever in contact with a sample surface, shear ultrasonic vibrations at the tip-sample interface can be excited using, for instance, a shear piezoelectric element attached to the back of the sample; shear acoustic waves originated by mechanical vibrations of the piezo will propagate through the sample, and reach the surface-tip contact area.

With a shear wave transducer oriented in such a way that the surface in-plane vibrations are polarized perpendicular to the long axis of the cantilever, torsional resonant modes of a cantilever with the tip in contact with the sample surface are excited. *Lateral-Acoustic Friction Force Microscopy* (L-AFAM)*(or Resonant Friction Force Microscopy* (R-FFM)*)* [24–27] monitors the vibration amplitudes of the cantilever torsional resonant modes at different surface points. In this technique, the sample is typically laterally vibrated at MHz frequencies, and the torsional vibration amplitudes provide information about the lateral forces between tip and sample. Apparently, L-AFAM images are independent of the scanning direction, i.e. not influenced by topography-induced lateral forces [25]. When scanning in the presence of shear ultrasonic vibration at the tip-sample contact, the relative tip-sample velocities are of the order of $1\,\mathrm{mm\,s^{-1}}$, much larger than those in conventional FFM (about $100–250\,\mu\mathrm{m\,s^{-1}}$), and nearer to the sliding operating velocities in MEMs and NEMs (in the range of tens of $\mathrm{mm\,s^{-1}}$ to few $\mathrm{m\,s^{-1}}$) [48].

The analysis of the torsional contact resonances of AFM cantilevers in contact with a sample surface provides a novel means to study friction and stick-slip phenomena at the nanometer scale [26, 27]. At low shear excitation voltages, the resonance curve torsional cantilever vibration amplitude versus excitation frequency is a Lorentzian with a well-defined maximum; the cantilever with the AFM tip stuck to the sample surface following the surface motion, behaves like a linear oscillator with viscous damping. Above a critical shear excitation amplitude, which depends on the static cantilever load, and is of the order of 0.2 nm for bare and lubricated silicon samples [26], the shape of the resonance curve exhibits a characteristic flattening, attributable to the onset of sliding friction at the tip-sample contact. Experimental evidence of energy dissipation before sliding friction sets in has been related to microslip, i.e. slipping of an annulus at the tip-sample contact before the whole contact starts to slide (see [26] for further details).

The local vibration amplitudes and phases of the torsional resonances of clamped-free AFM cantilevers have been studied using optical interferometry [28]. The finite size of the cantilever beam and asymmetries in its shape leads to coupling between flexural and torsional vibrations. Lateral resonant modes of AFM cantilevers, which consist in flexural vibration modes in the cantilever width direction parallel to the sample surface, have also been experimentally observed [29]; asymmetries in the cantilever thickness lead to a z component of the displacement that can be monitored by optical beam deflection with an AFM.

The torsional resonant modes of a cantilever tip in contact with a sample surface have also been excited using a shear piezo located at the cantilever base [30, 31]. In the *Torsional Resonance Dynamic-AFM mode* (TRmode) [32] torsional vibrations of the cantilever are excited via two piezoelectric elements mounted beneath the holder of the chip, which vibrate out-of-phase, in such a way that they generate a rotation at the length axis of the cantilever. Using this procedure, the torsional resonances of the cantilever can be monitored in both near-contact and contact modes. In ultra-high vacuum (UHV), torsional cantilever resonances can be excited via vertical vibrations, due to their high quality factors. Lateral forces between a cantilever tip and objects on surfaces have been measured in UHV by monitoring the induced change of the frequency of the fundamental cantilever torsional resonant mode [33]. In the *Torsional Overtone Microscopy* [34], torsional cantilever resonances excited by thermal noise are used to obtain information about the shear stiffness of the tip-sample contact.

In the limit of high ultrasonic frequencies, it is questionable if high-order torsional resonances will be excited at the cantilever. Nevertheless, in *Lateral Scanning Acoustic Force Microscopy* (LFM-SAFM) [35, 36] SAWs with in plane oscillations components such as Love waves have been detected by modulating the rf signal's amplitude at some KHz. When the tip is in contact with the sample surface, in the presence of shear ultrasonic vibration at the tip-sample contact, the cantilever experiences an additional amplitude-dependent torsion or lateral-mechanical diode effect. From the ultrasound-induced additional torsion, information about the amplitude and phase velocity of in-plane polarized SAWs can be obtained.

In *Lateral Ultrasonic Force Microscopy* (L-UFM) [9] lateral vibrations of the sample surface at a relatively low frequency of some KHz, polarized perpendicular to the length axis of the cantilever, are superimposed on a continuous vertical ultrasonic surface vibration. The measurement of the amplitude of torsion of the cantilever at the lateral low-frequency surface vibration provides information about the sample shear elastic properties with subsurface sensitivity.

3.4 Reduction of Friction by Ultrasonic Vibration

The reduction of friction by ultrasound is a well-known macroscopic effect [1, 2]. Its occurrence at the nanometer scale is only recently being investigated.

Dinelli et al. [37] studied the influence of out-of-plane ultrasonic vibration on the frictional response of a Si sample in ambient conditions, using FFM and UFM.

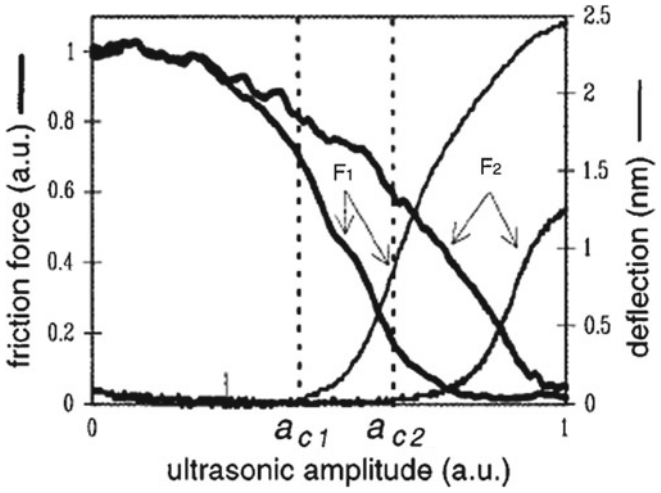


Fig. 3.4 Experimental measurements of dynamic friction (*thick line*) and cantilever deflection (*thin line*) dependencies on the ultrasonic amplitude, for two different applied loads $F_1 = 0$ N and $F_2 = 2$ nN on a Si sample (from [37])

Their results clearly demonstrated that dynamic friction vanishes in the presence of ultrasound when the tip-surface contact breaks for part of the out-of-plane vibration cycle (see Fig. 3.4). Figure 3.4 shows the friction force and the cantilever deflection measured at different surface ultrasonic vibration excitation amplitudes. The friction force in Fig. 3.4 was independently determined for each of the different amplitudes of surface ultrasonic vibrations by laterally scanning the sample back and forth in the direction perpendicular to the cantilever axis, using a lock-in amplifier (see [37] for further details). The cantilever deflection signal in Fig. 3.4 corresponds to the cantilever response to the ultrasonic force, i.e. the UFM signal, which depends on the ultrasonic amplitude (see Fig. 3.1). The onset of an UFM response for a given set point force roughly indicates the ultrasonic amplitude needed for the tip to detach from the sample surface at part of the surface ultrasonic vibration cycle.

The breaking of the tip-sample contact at each ultrasonic cycle explains the reduction or elimination of friction, because of a reduction of slippage during sliding. Interestingly, it is apparent in Fig. 3.4 that, for a given applied load, the friction force considerably reduces well before the onset of the UFM response, i.e. while the tip remains in "linear contact" with the sample surface during the ultrasonic vibration cycle. For the case of F_2 in Fig. 3.4, the reduction of friction already amounts to about 60 % when the UFM cantilever response sets off.

A study of influence of out-of-plane ultrasonic vibration on the static friction force, keeping the amplitude of the lateral displacement small enough that the tip sticks to a surface point without sliding, (see [37] for details) demonstrated that this begins to decrease at very low ultrasonic amplitudes, and that the onset of friction reduction does not depend on the applied shear force. Evidence on this latter point

ruled out the possibility that the reduction of friction is due to slippage during the part of the period that the tip-sample forces are the lowest.

In order to explain a reduction of friction at low ultrasonic amplitudes, the presence of a surface layer at the tip-sample gap, i.e. a liquid layer formed by water and possibly organic contaminants, has been considered [37]. In the absence of ultrasonic vibration, such a layer might organize in a solid-like configuration between the tip and the sample and partially sustain the load. As the tip-sample distance is varied at ultrasonic frequencies, the viscosity of the layer would hinder its re-arrangement, thereby reducing the probability of tip stick-slip processes, and hence friction.

Using molecular dynamics (MD) simulations, Gao et al. [49] demonstrated that small amplitude (of the order of 0.1 nm) oscillatory motion of two confining interfaces in the normal direction to the shear plane can lead to transitions of a lubricant from a high-friction stick-slip shear dynamics to an ultralow kinetic friction state (superkinetic friction regime), provided that the characteristic relaxation time for molecular flow and ordering processes in the confined region be larger than the time constant of the out-of-plane mechanical oscillations.

Heuberger et al. [50] observed load- and frequency- dependent transitions between a number of dynamic friction states of a lubricant using a surface forces apparatus, modified for measuring friction forces while simultaneously inducing normal (out-of-plane) vibrations between two boundary-lubricated sliding surfaces. In particular, they found regimes of vanishingly small friction at interfacial oscillation amplitudes below 0.1 nm, and demonstrated that they originate due to the dynamics of the relaxation processes of the lubricant at the molecular level.

Recently, Socoliuc et al. [51] have demonstrated that mechanical vibrations normal to the plane of sliding at cantilever resonance frequencies in a range of hundreds of KHz in ultra-high-vacuum (UHV) conditions lead to an ultra-low friction regime in atomic scale friction even in the case that the amplitude is not sufficiently high that the tip detaches from the sample during the vibration cycle. Previously [52], the authors had reported on the observation of an ultralow dissipation state in atomic friction related to the absence of mechanical instabilities, attained by varying the normal force. Such a state may exist because a modification of the tip-sample normal load leads to changes in the lateral surface corrugation felt by the tip without significantly altering the stiffness of the tip-sample contact. In the case that the tip-sample force is periodically varied at high frequencies, it is feasible that the tip slides through ultralow dissipation atomic friction states when being laterally displaced.

The effect of in-plane ultrasonic vibration in nanoscale friction has also been considered. Scherer et al. [25] observed that when lateral ultrasonic vibrations are excited at a sample surface at ambient conditions using a shear piezo bonded to the back of the sample, friction nearly vanishes at certain frequency bands, whereas remains as high as on a non-vibrating surface at other frequencies. However, they verified that the near-zero friction bands coincided with frequencies at which a lift-off (vertical displacement) of the AFM cantilever occurred. As discussed by the authors [25] such “lift-off” is likely attributable to the set off of a vertical ultrasonic force due to parasitic out-of-plane motions of the sample surface or to mode coupling in the cantilever. Nevertheless, the build up of an elastohydrodynamic lubrication film

whose viscosity and hence thickness is dependent on the lateral tip-sample relative velocity was proposed as a reasonable hypothesis, that could account for a vertical cantilever displacement in the absence or in the case of low-amplitude out-of-plane surface vibrations.

Behme et al. [38–40] studied the influence of Surface Acoustic Waves (SAWs) on nanoscale friction. SAWs constitute a precise source of acoustic vibration, with well-defined surface oscillations in a perfectly determined polarization, whereas when working with bulk shear wave transducers parasitic surface displacements due to the existence of boundaries, etc. can hardly be avoided. LFM and multimode SAFM were used to measure and distinguish the influence of in-plane and vertical surface oscillations components on the cantilever torsion and bending. To this aim, the authors [38–40] excited a standing Rayleigh wave field, and considered the dependence of friction on the acoustic excitation amplitude. In Rayleigh waves, the atoms oscillate on elliptical trajectories with a large vertical and a smaller lateral oscillation component. The experiments showed that by increasing the rf. amplitude, friction is locally reduced an eventually suppressed. In addition, it was clearly demonstrated that at the point in which friction disappears, the lateral-SAFM signal breaks down. Hence, it was concluded that the effect of friction reduction is essentially due to the vertical "mechanical diode effect" that leads to an effective shift of the cantilever, whereas in-plane oscillations do not play a significant role. This hypothesis is further reinforced by the fact that apparently in-plain polarized Love-type SAWs do not significantly alter the frictional behaviour. In these experiments, no cantilever lift-off induced by a lateral-oscillation of the sample [25] was observed. At very high Rayleigh wave amplitudes a lateral force rectification of the longitudinal component of the standing wave field is apparent, which results in a scan-direction independent appearance of the LFM traces.

Ultrasonic vibration covers a broad range of frequencies, and the processes involved in a reduction of friction by ultrasound can vary at different relative tip-sample velocities. Kessermakers et al. [53] studied the influence on nanoscale friction of lateral high frequency vibration of the cantilever, up to frequencies of 1 MHz, on a NbS_2 sample at ambient conditions, and observed gaps of lowered or eliminated friction at specific frequencies, presumed to be around torsional and/or lateral cantilever resonances. In these experiments a Au-coated cantilever was used, and the oscillating lateral cantilever vibration was applied by means of an electrostatic field. At a particular friction gap frequency, a slow increase in driving field amplitude caused a gradual increase in friction, and above a certain threshold level of driving amplitude, a partial stick/slip behaviour with the tip periodically alternating between a zero friction an a non-zero-friction state was apparent.

Riedo et al. [54] also reported about a reduction of friction when lateral oscillations around a frequency of 19.5 KHz were applied to an AFM cantilever sliding on mica. In the range of scanning velocities they used, the thermally activated hopping of contact atoms over the effective lateral interatomic potential led to increased energy dissipation when increasing the sliding velocity. By superimposing a lateral oscillation on the cantilever and sweeping its frequency between 15–100 KHz, and a clear peak of friction reduction was observed around 19.5 KHz, independently of

the applied load. This friction reduction peak was attributed to the excitation of a cantilever torsional contact resonance, which increased the attempt frequency for thermally activated jumps during sliding. The effect did not occur above a certain critical value of the sliding velocity.

In recent experiments performed by Gnecco et al. [51] on KBr samples in UHV no reduction-of-friction effect was apparent upon the excitation of torsional cantilever contact resonances in the frequency range from 40 up to 200 KHz, even though friction was strongly reduced when the excitation frequency matched one of the normal resonance frequencies of the pinned lever or half its value.

Other works that have considered the possibility to control nanoscale friction by mechanical action at high frequencies on the system motion are described in [55, 56] and therein.

3.5 Adhesion Hysteresis at Ultrasonic Frequencies

On the nanoscale, adhesion phenomena become decisive to the performance of nanodevices, and surface properties acquire a particular relevance. Usually, the work of adhesion is defined as the energy needed to separate two surfaces, assuming that this is reversible [57]. The adhesion hysteresis is defined as the difference between the work needed to separate two surfaces and that gained when bringing them together. The fact that those two works are different in magnitude, i.e. the adhesion hysteresis is different from zero, can be attributed to elastic, viscoelastic and plastic deformations in the contact zone, reconfiguration of surface molecules during contact, chemical reactions, etc.

Recently, novel methods to obtain information about the work of adhesion and the adhesion hysteresis at the tip-sample contact using UFM have been proposed [41–45]. Essentially, they take advantage of the fact that the ultrasonic excitation amplitude at which an UFM response sets off when increasing the excitation is different from this at which it falls down when decreasing the excitation. This is illustrated in Fig. 3.5 [41], in which both experimental and simulated UFM signal versus ultrasonic excitation amplitude curves have been drawn. In UFM, being the tip in contact with the sample, when increasing the out-of plane ultrasonic amplitude at the tip-sample contact, at certain amplitude the tip detaches from the surface at part of the ultrasonic period, and the ultrasonic force (see Sect. 3.2 of this chapter) experiences a sudden increase that give rise to a “jump out” of the cantilever (see Fig. 3.5). When decreasing the ultrasonic amplitude, at certain amplitude the tip cannot separate anymore from the surface, and the ultrasonic force experiences a sudden decrease, that gives rise to a “jump in” of the cantilever (see Fig. 3.5). For the evaluation of the ultrasonic force, it is considered that mechanical hystereses i.e. snap-in and -out of the cantilever when approaching or separating from the sample surface do not occur. In the absence of ultrasound, compliant cantilevers are subjected to large mechanical hysteresis when approaching or separating from a sample surface due to the force gradient being larger than the cantilever spring constant. However, at ultrasonic frequencies, the

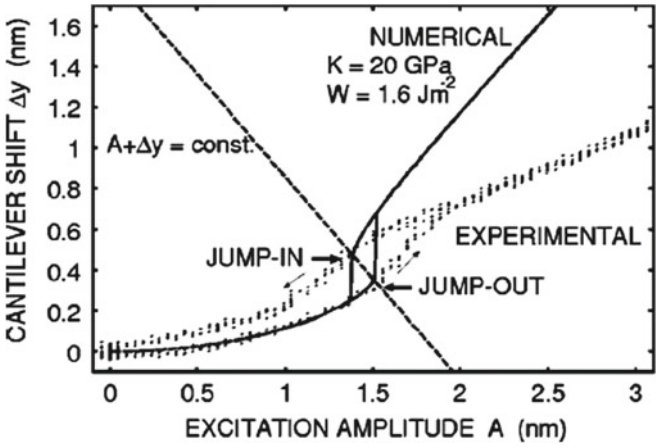


Fig. 3.5 UFM signals recorded when increasing and decreasing the ultrasonic excitation amplitude (see *arrows* to distinguish each case) on an aluminium thin film. The *continuous lines* correspond to a numerical evaluation of the UFM responses according to the model detailed in [41] (from [41])

inertia of the cantilever leads to an effective much larger cantilever stiffness, which can hence probe the hysteretic cycle of tip-sample in and out interactions, without a decrease of its sensitivity for force field detection.

In [41] a method for quantitative analysis of the UFM signal is proposed in order to determine both the sample elastic modulus and the work of adhesion by monitoring the cantilever jumps such as those in Fig. 3.5. In UFM, both elasticity and adhesion contribute to the ultrasonic force. Dinelli et al. [46] evaluated the contact stiffness by comparing the jump-in positions in ultrasonic amplitude for different applied loads. Using the Johnson-Kendall-Roberts-Sperling (JKRS) model to account for both elastic and adhesive forces between tip and sample, the authors in [41] evaluate both the stiffness and the work of adhesion as defined in JKRS by calculating the jump-in and jump-out cantilever shifts. According to their modelling, the normalized cantilever jump-in shift turns out to be constant and effectively independent on the set point force, the stiffness and the work of adhesion. Hence, they derived a universal relation between the work of adhesion, the stiffness and the cantilever shift at jump in, the latter being easily measured from the experimental data (see [41] for further details).

In [42] the area between experimental curves such as those in Fig. 3.5 is measured and defined as the UFM hysteresis area (UH), and it is assumed that UH scales with the local adhesion hysteresis. A detail procedure to obtain quantitative information about the adhesion hysteresis from UFM signal versus ultrasonic excitation amplitude curves is discussed in [45]. The correlations between adhesion hysteresis and local friction have been theoretically and experimentally investigated [44]. According to a model based on the classical theory of adhesional friction and contact mechanics which includes the effects of capillary hysteresis and nanoscale roughness and assumes an adhesive, elastic and weariless tip-sample contact, a relationship

between adhesion hysteresis and friction has been derived, which depends on the varying ratio of the tip-sample work of adhesion over the reduced Young modulus (see [44] for further details). In the model, the adhesion hysteresis is estimated as the pull-off force times the critical separation at which the tip-sample contact is about to be broken. Measurements on a wide range of engineering samples with varying adhesive and elastic properties have confirmed the model [42, 44]. The aforementioned ratio does not vary much between typical metallic samples, and for a limited number of specimen's adhesion hysteresis and friction the experimental relationship may appear lineal. In addition, it is found that capillary hysteresis offsets the measured adhesion hysteresis from the friction force, and that roughness reduces both friction and adhesion hysteresis: friction decreases because of a smaller area of a real contact, and adhesion hysteresis drops due to a smaller pull-off force at rough surfaces. Recently, it has been demonstrated that the study of the dependence of local adhesion hysteresis on relative humidity using UFM may provide information about protein-water binding capacity with molecular scale resolution [43].

Procedures to obtain information about the work of adhesion using AFAM are also being considered [44]. In AFAM, monitoring of the resonance frequency of an AFM cantilever with the tip in contact with the sample surface allows us to determine the tip-sample contact stiffness (see Sect. 3.2 of this chapter). Strictly, the contact stiffness is influenced by both tip-sample elastic properties and the work of adhesion. Typically, the tip-sample distance in AFAM is kept sufficiently small that the tip-sample interactions remain in the linear regime. Recently, a method has been proposed to evaluate both these properties quantitatively from the analysis of the nonlinear AFAM cantilever response excited when the tip-sample distance sweeps the nonlinear part of the tip-sample interaction, but in such a way that the tip always remains in contact with the sample surface, considering the case of a perfect contact. To this aim, the dependence of the resonance frequency on the vibration amplitude is studied; the elastic properties and the work of adhesion are separately determined by finding the optimal set of values which minimizes the difference between the theoretical and empirical relationship of cantilever resonance frequency versus ultrasonic excitation amplitude (see [46] for further details).

In HFM, the phase signal provides information of the adhesion hysteresis related to the formation and breaking of the tip-surface contact [21]. Contrast in Phase-HFM mostly stems from dissipative processes; an exceptional feature of the technique being its ability to probe a local response in extremely short times, i.e. HFM may test effects that take place at nanoseconds in nanometer scale volumes. Hence, Phase-HFM can reveal dissipation due to extremely quick transitions that otherwise remains unresolved from other dissipative effects occurring at larger time scales. For instance, using Phase-HFM, it has been possible to distinguish differences in contrast at identical thin polymer layers with different boundary constraints on the nanometer scale. Those layers however exhibited a same FFM contrast, which confirms the ability of Phase-HFM to resolve dynamic dissipative processes in a much shorter time scale than conventional FFM. In the following, the results presented in [21] relative to those experiments will be summarized here, with a main focus in understanding the

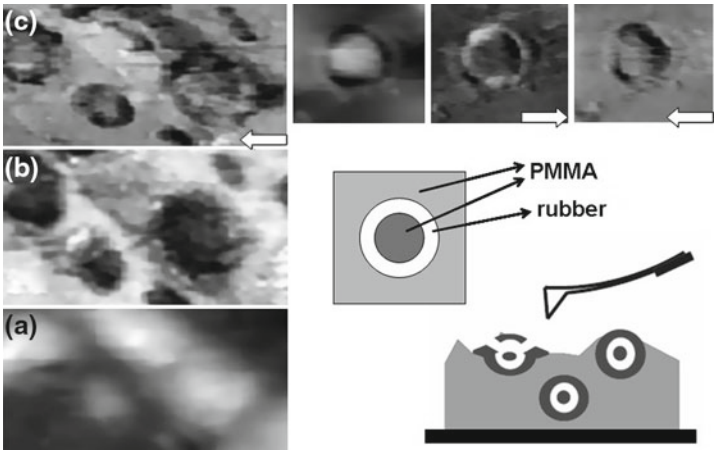


Fig. 3.6 (**a**–**c**) AFM contact-mode topography (**a**), Phase-AFM (**b**) and LFM images recorded over a same surface region of a PMMA/rubber sample. The images at the *top right-hand* side correspond to AFM contact-mode topography, and LFM images recorded scanning from *left* to *right*, and viceversa (see *arrows*) respectively, recorded over a same surface region of the sample, different from this in (**a**–**c**). Below, schematic drawings illustrate the apparent structure at the PMMA/rubber sample surface

opportunities of phase-HFM to provide information about adhesion hysteresis with extremely high time resolution, in the nanosecond time scale.

In metals, anelastic or viscoelastic contributions are expected to be small. On the contrary, in polymeric materials intra- or inter-molecular perturbations upon tip actuation and/or dissipative effects of the molecules due to adhesion to the tip or to other neighbouring molecules will play a significant role in the Phase-HFM contrast. Phase-HFM has been applied to PMMA/rubber nanocomposites that consist in an acrylic matrix, a copolymer based upon PMMA, and toughening particles, composed of a core of acrylic enclosed with rubber with a bonded acrylic outer shell to ensure good bonding to the matrix (see Fig. 3.6).

Figure 3.6a–c shows contact-mode AFM (a), Phase-AFM (b) and LFM images recorded over a same surface region of a PMMA/rubber sample. The topographic protrusions in Fig. 3.6a indicate the presence of core-shell PMMA particles in the surface and/or near surface region. Two different kinds of topographic protrusions may be distinguish from those and other images recorded on the PMMA/rubber sample surface: (i) some that give rise to a lower Ph-HFM contrast than the PMMA matrix, and (ii) others that show a Ph-HFM contrast similar to that of the PMMA matrix. Such different protrusions are apparent from the comparison of Fig. 3.6a, b. The drawings in Fig. 3.6 illustrate a model for the two different protrusions: at some of particles, the PMMA particle shell is well-bonded and indistinguishable from the PMMA matrix, whereas in others the rubber particle is still capped with the PMMA layer, but this is detached from the matrix material. Such a picture is corroborated when considering FFM images (see Fig. 3.6c) as well as UFM and A-HFM images (not shown here, see [21]) recorded in the same surface region. Both UFM and

A-HFM reveal the presence of the toughening particles by a clear darker contrast, indicative of the presence of a softer material in the surface or near-surface region; the aforementioned different particles cannot be distinguished from the UFM and A-HFM measurements [21]. However, they are clearly differentiated in Ph-HFM, and discernible by the presence or absence of kind of halo-contrast in FFM.

At the top right-hand side of Fig. 3.6, contact-mode AFM and FFM images recorded over a particular PMMA/rubber particle scanning from left to right (forward scan), and viceversa (backward scan, see arrows in the figure) are shown. This particle is representative of those that typically give rise to Ph-HFM contrast, and the image quality is a bit better than this in Fig. 3.6c. From those images it is apparent that the particle is characterized by a halo-shaped frictional contrast, in both forward (bright halo) and reversed (dark halo) FFM scans, which can be attributed to the presence of rubber directly exposed at the sample surface. Notice that *the PMMA layer on top of the rubber exhibits the same frictional contrast than the PMMA matrix, being indistinguishable from that in both forwards and backwards FFM scans*. In contrast, Ph-HFM resolves small differences in viscoelastic and/or adhesion hysteresis response time of the PMMA on top of the rubber that is not linked to the PMMA rubber matrix. Relaxation processes of polymeric materials are strongly dependent on the constraints for molecular movement. A different molecular density, entanglement density and/or molecular weight in the PMMA layer on top of rubber that is detached from the PMMA matrix may lead to differences in the PMMA viscoelastic and/or adhesion hysteresis response. In addition differences in interfacial bonding between the rubber and the PMMA on top depending on whether the PMMA is well adhered to the PMMA matrix or not, may also modify the PMMA dynamic behaviour. According to the obtained experimental results, the contrast provided by Ph-HFM allows us to distinguish differences in the locally-probed dynamical response of PMMA on top of rubber depending on whether the PMMA is well adhered to the matrix or not, in spite of the fact that no difference between can be resolved in conventional FFM. Hence, Ph-HFM allows us to study quick dissipative transitions not resolved by FFM which will however surely play an important role in MEM/NEMs devices working at much higher sliding velocities than those typically used in AFM/FFM measurements.

It is also worth to point out that, when probed with extreme sensitivity, a locally measured response might be strongly affected by small dissipative effects induced by long-range interactions (via molecular entanglements) at molecules outside the immediate contact region. The possibility that those kinds of interactions might be detected in an extremely short time scale can be of interest in the implementation of dynamic mechanical procedures for communications in nanodevices.

E. Gnecco is gratefully acknowledged for scientific discussions, and careful reading and critic of the manuscript. Financial support from the Spanish Ministerio de Educación y Ciencia (MEC) under project MAT2002-00076), the Junta de Comunidades de Castilla-La Mancha (JCCM) under projects PBI-02-003 and PBI-05-018, and the European Science Foundation (ESRF), under the ESRF Scientific Programme NATRIBO, is also gratefully acknowledged.

References

1. T.B. Thoe, D.K. Aspinwall, M.L.H. Wise, Review on ultrasonic machining. Int. J. Mach. Tools Manufact. **38**, 239 (1998)
2. A. Akay, Acoustics of friction. J. Acoust. Soc. Am. **111**, 1525 (2002)
3. K. Dransfeld, Generation of ultrasonic waves in sliding friction, Chap. 7, in *Nanoscience: Friction and Rheology on the Nanometer Scale*, ed. by E. Meyer, R.M. Overney, K. Dransfeld, T. Gyalong (World Scientific, Singapore, 1998)
4. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Atomic scale friction of a tungsten tip on a graphite surface. Phys. Rev. Lett. **59**, 942 (1987)
5. G. Meyer, N. Amer, Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. Appl. Phys. Lett. **57**, 2089 (1990)
6. W. Rohrbeck, E. Chilla, Detection of surface acoustic waves by scanning force microscopy. Phys. Stat. Sol. (a) **131**, 69 (1992)
7. O. Kolosov, K. Yamanaka, Nonlinear detection of ultrasonic vibrations in an Atomic, Force microscope. Jpn. J. Appl. Phys. **32**, L1095 (1993)
8. B. Cretin, F. Sthal, Scanning microdeformation microscopy. Appl. Phys. Lett. **62**, 829 (1993)
9. K. Yamanaka, H. Ogiso, O. Kolosov, Ultrasonic force microscopy for nanometer resolution subsurface imaging. Appl. Phys. Lett. **64**(2) (1994)
10. U. Rabe, W. Arnold, Acoustic microscopy by atomic force microscopy. Appl. Phys. Lett. **64**, 1493 (1994)
11. T. Hesjedal, E. Chilla, H.-J. Froehlich, Scanning acoustic force microscopy measurements in grating-like electrodes. Appl. Phys. A **61**, 237 (1995)
12. N.A. Burnham, A.J. Kulik, G. Gremaud, G.A.D. Briggs, Nanosubharmonics: the dynamics of small nonlinear contacts. Phys. Rev. Lett. **74**, 5092 (1995)
13. U. Rabe, K. Janser, W. Arnold, Vibrations of free and surface-coupled atomic force microscope cantilevers: theory and experiments. Rev. Sci. Instr. **67**, 3281 (1996)
14. K. Yamanaka, S. Nakano, Ultrasonic atomic force microscopy with overtone excitation of the cantilever. Jpn. J. Appl. Phys. **35**, 3787 (1996)
15. K. Yamanaka, S. Nakano, Ultrasonic atomic force microscopy with overtone excitation of the cantilever. Jpn. J. Appl. Phys. **35**, 3787 (1996)
16. S.C. Minne, S.R. Manalis, A. Atalar, C.F. Quate, Contact imaging in the AFM using a high order flexural mode combined with a new sensor. Appl. Phys. Lett. **68**, 1427 (1996)
17. P. Variac, B. Cretin, Scanning microdeformation microscopy in reflexion mode. Appl. Phys. Lett. **68**, 461 (1996)
18. E. Chilla, T. Hesjedal, H.-J. Fröhlich, Nanoscale determination of phase velocity by scanning acoustic force microscopy. Phys. Rev. B **55**, 15852 (1997)
19. F. Dinelli, M.R. Castell, D.A. Ritchie, N.J. Mason, G.A.D. Briggs, O.V. Kolosov, Mapping surface elastic properties of stiff and compliant materials on the nanoscale using ultrasonic force microscopy. Phil. Mag. A **80**, 2299 (2000)
20. K. Inagaki, O. Kolosov, A. Briggs, O. Wright, Waveguide ultrasonic force microscopy at, 60 MHz. Appl. Phys. Lett. **76**, 1836 (2000)
21. M.T. Cuberes, H.E. Assender, G.A.D. Briggs, O.V. Kolosov, Heterodyne force microscopy of PMMA/rubber nanocomposites: nanomapping of viscoelastic response at ultrasonic frequencies. J. Phys. D.: Appl. Phys. **33**, 2347 (2000)
22. M.T. Cuberes, G.A.D. Briggs, O. Kolosov, Nonlinear detection of ultrasonic vibration of AFM cantilevers in and out of contact with the sample. Nanotechnology **12**, 53 (2001)
23. G.S. Shekhawat, V.P. Dravid, Nanoscale imaging of buried structures via scanning near-field, ultrasound holography. Science **310**, 89 (2005)
24. V. Scherer, B. Bhushan, U. Rabe, W. Arnold, Local elasticity and lubrication measurements using atomic force and friction force microscopy at ultrasonic frequencies. IEEE Trans. Magn. **33**, 4077 (1997)
25. V. Scherer, W. Arnold, B. Bhushan, Lateral force microscopy using acoustic force microscopy. Surf. Interface Anal. **27**, 578 (1999)

26. M. Reinstädtler, U. Rabe, V. Scherer, U. Hartnann, A. Goldade, B. Bhushan, W. Arnold, On the nanoscale measurement of friction using atomic-force microscopy cantilever torsional resonances. Appl. Phys. Lett. **82**, 2604 (2003)
27. M. Reinstädtler, U. Rabe, A. Goldade, B. Bhushan, W. Arnold, Investigating ultra-thin lubricant layers using resonant friction force microscopy. Tribol. Int. **38**, 533 (2005)
28. M. Reinstädtler, U. Rabe, V. Scherer, J.A. Turner, W. Arnold, Imaging of flexural and torsional resonance modes of atomic force microscopy cantilevers using optical interferometry. Surf. Sci. **532**, 1152 (2003)
29. A. Caron, U. Rabe, M. Reinstätler, J.A. Turner, W. Arnold, Imaging using lateral bending modes of atomic force microscopy cantilevers. Appl. Phys. Lett. **85**, 6398 (2004)
30. K. Yamanaka, S. Nakano, Quantitative elasticity evaluation by contact resonance in an atomic force microscope. Appl. Phys. A **66**, S313 (1998)
31. T. Kawagishi, A. Kato, U. Hoshi, H. Kawakatsu, Mapping of lateral vibration of the tip in atomic force microscopy at the torsional resonance of the cantilever. Ultramicroscopy **91**, 37 (2002)
32. M. Reinstädtler, T. Kasai, U. Rabe, B. Bhushan, W. Arnold, Imaging and measurement of elasticity and friction using the TRmode. J. Phys. D: Appl. Phys. **38**, R269 (2005)
33. O. Pfeiffer, R. Bennewitz, A. Baratoff, E. Meyer, P. Grütter, Lateral-force measurements in dynamic force microscopy. Phys. Rev. B **65**, 161403 (2002)
34. T. Drobek, R.W. Stark, W.M. Heck, Determination of shear stiffness based on thermal noise analysis in atomic force microscopy: passive overtone microscopy. Phys. Rev. B **64**, 0454001 (2001)
35. G. Bheme, T. Hesjedal, E. Chilla, H.-J. Fröhlich, Transverse surface acoustic wave detection by scanning acoustic force microscopy. Appl. Phys. Lett. **73**, 882 (1998)
36. G. Behme, T. Hesjedal, Simultaneous bimodal surface acoustic-wave velocity measurements by scanning acoustic force microscopy. Appl. Phys. Lett. **77**, 759 (2000)
37. F. Dinelli, S.K. Biswas, G.A.D. Briggs, O.V. Kolosov, Ultrasound induced lubricity in microscopic contact. Appl. Phys. Lett. **71**, 1177 (1997)
38. G. Behme, T. Hesjedal, Influence of ultrasonic surface acoustic waves on local friction studied by lateral force microscopy. Appl. Phys. A **70**, 361 (2000)
39. G. Behme, T. Hesjedal, Influence of surface acoustic waves on lateral forces in scanning force microscopies. J. Appl. Phys. **89**, 4850 (2001)
40. T. Hesjedal, G. Behme, The origin of ultrasound-induced friction reduction in microscopic mechanical contacts. IEEE Trans. Ultrason. Ferroelectr. Freq. Control **49**, 356 (2002)
41. K. Inagaki, O. Matsuda, O.B. Wright, Hysteresis of the cantilever shift in ultrasonic force microscopy. Appl. Phys. Lett. **80**, 2386 (2002)
42. R. Szoszkiewicz, B.D. Huey, O.V.Kolosov, G.A.D. Briggs, G. Gremaud, A.J. Kulik, Tribology and ultrasonic hysteresis at local scales. Appl. Surf. Sci. **219**, 54 (2003)
43. R. Szoszkiewicz, A.J. Kulik, G. Gremaud, M. Lekka, Probing local water contents of in vitro protein films by ultrasonic force microscopy. Appl. Phys. Lett. **86**, 123901 (2005)
44. R. Szoszkiewicz, B. Bhushan, B.D. Huey, A.J. Kulik, G. Gremaud, Correlations between adhesion hysteresis and friction at molecular scales. J. Chem. Phys. **122**, 144708 (2005)
45. R. Szoszkiewicz, A.J. Kulik, G. Gremaud, Quantitative measure of nanoscale adhesion hysteresis by ultrasonic force microscopy. J. Chem. Phys. **122**, 134706 (2005)
46. F. Dinelli, S.K. Biswas, G.A.D. Briggs, O.V. Kolosov, Measurements of stiff-material compliance on the nanoscale using ultrasonic force microscopy. Phys. Rev. B **61**, 13995 (2000)
47. M. Muraoka, W. Arnold, A method of evaluating local elasticity and adhesion energy from the nonlinear response of AFM cantilever vibrations. JSME Int. J. Ser. A **44**, 396 (2001)
48. N.S. Tambe, B. Bhushan, Recently, a novel AFM-based technique for studying nanoscale friction at velocities near to 10 nm s-1 has been implemented; see A new atomic force microscopy based technique for studying nanoscale friction at high sliding velocities. J. Phys. D: Appl. Phys. **38**, 764 (2005)
49. J. Gao, W.D. Luedtke, U. Landman, Friction control in thin-film lubrication. J. Phys. Chem. B **102**, 5033 (1998)

50. M. Heuberger, C. Drummond, J. Israelachvili, Coupling of normal and transverse motions during frictional sliding. J. Phys. Chem. B **102**, 5038 (1998)
51. A. Socoliuc, E. Gnecco et al., submitted.
52. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Transition from stick-slip to continous sliding in atomic friction: entering a new regime of ultralow friction. Phys. Rev. Lett. **92**, 134301 (2004)
53. J. Kerssemakers, J.T M. De Hosson, Probing the interface potential in stick/slip friction by a lateral force modulation technique. Surf. Sci. **417**, 281 (1998)
54. E. Riedo, E. Gnecco, R. Bennewitz, E. Meyer, H. Brune, Interaction potential and hopping dynamics governing sliding friction. Phys. Rev. Lett. **91** (2003)
55. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, The nonlinear nature of friction. Nature **430**, 523 (2004)
56. Z. Tshiprut, A.E. Filippov, M. Urbakh, Tuning diffusion and friction in microscopic contacts by mechanical excitations. Phys. Rev. Lett. **95**, 016101 (2005)
57. J. Israelachvili, *Chapter 15 in Intermolecular and Surface Forces*, 2nd Ed (Academic Press Elsevier Ltd, New york, 1992)

Chapter 4
Triboluminescence

Roman Nevshupa and Kenichi Hiratsuka

Abstract This works seeks to summarize recent advances in experimental studying of triboluminescence—defined here as the light emission when a material is subjected to rubbing, scratching, rolling, impacting or other mechanical agitation—and elucidate the basic mechanisms whereby triboluminescence is excited.

4.1 Introduction and Brief Historical Survey

Triboluminescence (TL) is defined by the Oxford English Dictionary as the emission of light from a substance caused by rubbing, scratching, or similar frictional contact [1] or, more generally, as the quality of emitting light at interacting surfaces in relative motion. It should be noted that the substance referred to in this definition may include both the materials of the contacting bodies and the surrounding gas or liquid, e.g., atmospheric air and lubricating oil.

The art of making fire by rubbing two wooden sticks together is a prehistoric example of how light emission can be produced by friction first due to incandescent radiation and then due to chemical reaction of wood oxidation [2]. Also triboluminescence can naturally occur when frictional contact is composed of dielectric materials susceptible of intensive frictional electrification. In this case triboluminescence is usually manifested by a shine or lightning around the frictional contact and is related to electrical discharge with gas or liquid. This phenomenon can be observed during precise turning of glassy polymers by a diamond tool in dry air [3], in which process weak glowing occurs around the cutting tool near the cutting edge as well as at the rake surface of the tool (see Fig. 4.1).

R. Nevshupa (✉)
IETCC, CSIC, C/Serrano Galvache 4, 28033 Madrid, Spain
e-mail: r.nevshupa@csic.es

K. Hiratsuka
Chiba Institute of Technology, 2-17-1 Tsudanuma,
Narashino-shi, Chiba 275-8588, Japan
e-mail: hiratsuka@sea.it-chiba.ac.jp

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_4

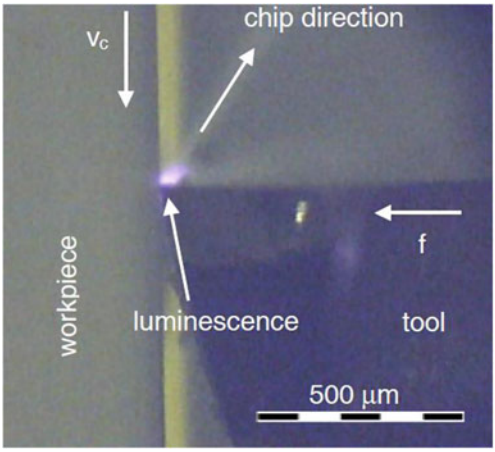


Fig. 4.1 *Side view* of the turning process of polycarbonate, showing light emission ($v_c = 1\,\mathrm{m/s}$, $f = 5\,\mu\mathrm{m/rev}$) (reprinted from [3] © 2004, with permission from CIRP)

Luminous signals in the atmosphere such as glowing, light flashes and lightning, which result from fracture and slippage of rocks in earthquake-prone regions, are also attributed by many researchers to triboelectricity and triboluminescence [4–7]. These signals, which occur prior to, concurrently or after severe earthquakes, are a supreme manifestation of triboluminescence which can be observed on the Earth. Even more tremendous are the lightnings and sprites in volcanic plumes and dust devils in planets of our solar system and in the atmospheres of brown dwarfs and gas giant planets outside of our solar system [8]. Triboelectric charging of dust particles is frequently invoked as a mechanism for spectacular lightning discharges in dust clouds of silicate planetary regolith analogues [9].

Though the oldest records on systematic studies of triboluminescence comes from the beginning of the seventeenth century, the TL must surely be observed before [10]. Bearing in mind that the triboluminescence frequently accompanies triboelectrification, this phenomenon could probably be noticed since the remote antiquity, from the works on electrostatics of Thales of Miletus dated back to the sixth century BC. The knowledge on the triboluminescence developed in the period from the 17th century to the third quarter of the 20th century was reviewed from various perspectives by A. Walton. A comprehensive literature survey on mechanically induced luminescence for the period 1968–1998 can be found in [11].

The word *triboluminescence* was coined by Wiedeman in 1888 to denote the light emission resulting on fracturing materials [10]. In its long history *triboluminescence* was applied to various luminescent processes, some of which, at first sight, have little in common, e.g., emission during breaking of adhesive bonds [12, 13]; shaking of mercury in a glass vessel [14]; rapid crystallization of certain substances [15]; collapse of small gas bubbles in a liquid [16]; excitation of a laser-induced shock wave in solids [17]; elastic and plastic deformation of solids, scratching, milling and fracture [10, 18]. Though the forms of mechanical action significantly differ between these processes, many of them rely on quite similar physico-chemical reactions

(see Sect. 4.2). As a natural consequence of the diversity of the forms of mechanical action, in literature, the name *triboluminescence* is concurrently used as a synonym of *fractoluminescence*, *crystalloluminescence*, *sonoluminescence* and *piezoluminescence,* all of which are different forms of *mechanoluminescence* [19]. In this work we endeavour to adhere to the original signification of *triboluminescence* as defined by the Oxford Dictionary, therefore, this chapter is focused principally on those triboluminescence phenomena which are directly related with tribological interaction between solids.

Even though the triboluminescence has been studied for centuries, this phenomenon still remains something of an enigma at the conceptual and theoretical levels. Triboluminescence is intimately related to both friction and wear and all three require an understanding of highly nonequilibrium processes occurring at the molecular level [20]. These processes significantly differ depending on the tribological conditions, environment and materials. Furthermore, it is important to understand how all these elements come together producing synergetic coupling on different length and time scales.

Triboluminescence has great potential for the development of advanced materials, for example, the materials capable of sensing and imaging the stress intensity and distribution under applied load [21, 22]. The TL can be used, as well, for remote diagnosing of failures of space and vacuum mechanisms [23]. In medicine, the TL and triboemission from biological tissues have been applied for early diagnosing of cancer and other diseases [24, 25].

4.2 Basic Processes and Activation Mechanisms

During long time it was believed that frictional heating is the dominant mechanism for many tribophysical and tribochemical phenomena including triboluminescence. This belief was induced by the empirical evidence coming from the observation of heat generation in heavily loaded joints and brakes. Nevertheless, as early as 1792 Saussure demonstrated that in certain cases light could be obtained under such conditions that a pure temperature radiation was altogether unlikely [15].

The obvious contradiction related to the fact that frictional heating is weak or negligible under mild and light frictional conditions, and that the net temperature increase under these conditions is not sufficient to initiate TL was overcame by introducing the flash temperature [2] and magma–plasma concepts [26]. The main idea behind these concepts is that heat generation occurs on sub-micrometric zones of real contact, called hot spots, under almost adiabatic conditions. It was suggested that, while the total generated heat is quite low, the instant temperature increase on the minute hot spots can be sufficiently high to produce melting, thermionic electron emission, triboplasma and triboluminescence [27]. Though these hypothesis have been proven for high speed and relatively heavy-loaded contacts, there is still a lot of debates about the existence of hot spots and the extent of the temperature increase on them in lightly loaded and microscopic tribocontacts as well as under low sliding velocities [28].

In the studies of energy dissipation at sliding of various amorphous carbon coatings using molecular dynamic simulation it was found that the temperature increase on the hot spots was less than 300 °C even under severe sliding conditions: contact pressure in the range 7.6–56 GPa and sliding velocity 90.2 ms^{-1} [29, 30]. The temperature flashes were extremely short-time, in picosecond range, and subnanometer size. These findings corroborated that the mechanism which can confidently be ruled out is that the average mechanical energy fed in per atom of a material under such sliding conditions is sufficient to excite an atom directly into emitting visible or UV light [10, 15]. Different excitation mechanisms of triboluminescence must exist.

Various authors suggested that at a tribocontact mechanical energy is not converted instantaneously and fully into Joule heat. Though the major part of mechanical energy is converted into heat, there exist de-excitation pathways in which energy is transferred to different forms [31]. Deformation, bonds breaking and stress-induced chemical transformations resulting from quasi-adiabatic energy release excite various energy-rich states [18, 32–36], whereas subsequent de-excitation of these excited states promotes cascade of multiple emission processes. Radiation and radiationless de-excitation can lead to luminescence, emission of electrons and ions, chemical transformations, generation of electric and magnetic fields and so on. Collision of the emitted particles and photons with atoms and molecules can lead to the secondary excitation and decay processes and so on [37]. Thus, energy dissipation at the sliding contact is an extremely complex process involving a large variety of elementary physical processes and chemical reactions. These processes include incandescent radiation [36, 38], molecular deformation and direct mechanical excitation of electronic levels in mechanically affected zones [18, 36]; space charge generation due to cleavage, contact electrification [12, 36, 39, 40]; piezoelectric effect [21]; recombination of charged particles, holes and radicals [18, 21]; electric discharge through surrounding gas or liquid [10, 15, 39]; electroluminescence [10, 26]; photoluminescence [41]; cathode- and ionoluminescence [42]; resonance radiation from excited particles [10]; phase transformation [36]; gas adsorption and chemical reactions [26, 36]; unpinning and annihilation of dislocations, defect stripping, motion of charged dislocations [39]; and others.

In an exhaustive study involving 23 different materials and aimed at distinguishing the major possibilities for excitation of mechanoluminescence obtained by grinding the crystals in a glass tube, Sweeting [43] found two clearly distinct mechanisms of mechanoluminescence: (i) luminescence that requires a discharge through air or other surrounding gas and (ii) luminescence which comes from the processes within the crystal, possibly the release of energy from recombination of energetic defects. Some of the materials presented parallel action of both mechanisms. In addition, three separate modes were observed for the first excitation mechanism: excitation of the surrounding gas by electron bombardment; excitation of the photoluminescence of the material by electron bombardment; and excitation of the photoluminescence of the material by absorption of the ultraviolet (UV) emission of the excited gas.

Recently, rare-earth complexes have drawn the attention of many researchers since mechanoluminescence of these materials can be induced by mechanical deformation without fracture and without electric discharge. During deformation the strain

energy excites the filled traps to release holes to the valence band. The holes then excite rare-earth ions and return to the ground state by emitting light [44]. Another material that can repeatedly emit light without destruction is zinc sulphide doped with transition metal ions. Non-destructive mechanoluminescence materials are very promising candidates for the development of smart damage sensors capable of real-time detection of the magnitude and location of deformation within the material.

A number of experiments coupled to theoretical modelling have made great advances in unifying apparently diverse phenomena relative to mechanical activation of physical and chemical reactions and provided a crucial benchmark for mechanisms and behaviour of mechanoluminescence under elemental mechanical actions such as fracture, grinding or cohesive detachment [10, 15, 18, 21, 36, 40, 43]. Varentsov and Khrustalev [45] summarized these findings in a generalized scheme of the emission phenomena and chemical reactions occurring when solids are subjected to an external mechanical influence, mainly in form of deformation or fracture. This scheme is shown in Fig. 4.2 by black boxes and lines. The boxes highlighted in yellow designate various luminescence phenomena. This scheme was further developed by introducing other non-intuitive processes involved in the mechanical energy dissipation at frictional contact (blue boxes and lines). Excitation of metastable states just by contacting without deformation and fracture, e.g., due to contact electrification, is one of these processes. These new processes are specifically relative to frictional mechanical interaction and have not been considered within the framework of mechanoemission and mechanochemistry.

For tribo- and mechanoluminescence, a common basis underpins the behaviour on atomic and nanometer scale (Fig. 4.3b). However, as a natural consequence of a complex tribological behaviour characterized by synergism of processes and self-organization of structures across time and length-scales of several orders of magnitude, TL is a much more complex phenomenon than elementary mechanoluminescence. In the mechanoluminescence studies, in which the materials are usually milled, ground, strained or fractured, the geometry of interfaces is highly irregular, while the contact interactions and interface processes have a random character. In contrast, when the wear rate is low, the macroscopic geometry of tribological contacts in machines and in the nature is usually well-defined and stable or quasi-stable in time (see Fig. 4.3a). The overall geometry of a tribocontact keeps stable despite the fact that in each instant new surfaces come into contact [37]. Quasi-stable distribution of physical parameters (pressure, electric field, temperature, and so on) around a frictional contact associated to its temporal and geometric stability creates favourable conditions for spontaneous formation of complex, self-organized and thermodynamically steady dissipative structures, which can occur at significant distances from the real contact zone. This explains why the triboluminescence phenomena can span the length scales from the quantum to the astronomical one, while the mechanoluminescence related to fracture or milling usually does not occur beyond microscopic scale.

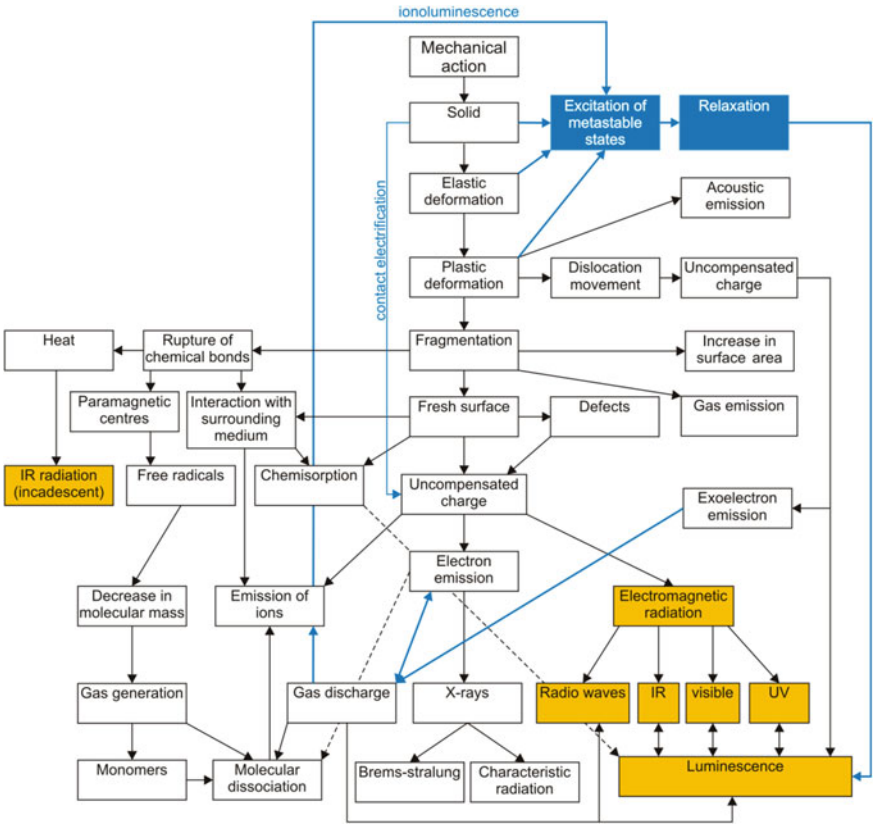


Fig. 4.2 Schematic drawing of some physical processes and chemical reactions involved in the mechanical energy dissipation at frictional contact and relevant triboemission and triboluminescence phenomena. *Black boxes* and *lines* correspond to the generalized scheme of the emission phenomena and chemical reactions occurring when solids are subjected to an external mechanical influence (reprinted from [45] © 1995, with permission from the Russian Academy of Sciences and Turpion Ltd). The *boxes* highlighted in *yellow* designate various luminescence phenomena. The *boxes* and *lines* highlighted in *blue* indicate the processes which are specifically related to frictional mechanical interaction and which have not been considered within the framework of mechanoemission and mechanochemistry

4.3 Experimental Techniques for Studying Triboluminescence

Triboluminescence occurs as a response of a system to the introduction of mechanical energy at a frictional contact. Thus, an experimental set-up for measuring the TL should necessarily have two main subsystems: (i) the tribological device for producing relative motion of two solids under controlled velocity and loading conditions and (ii) the optical system for measuring the intensity and spectral characteristics of the emitted light in different spectral ranges and on various time scales.

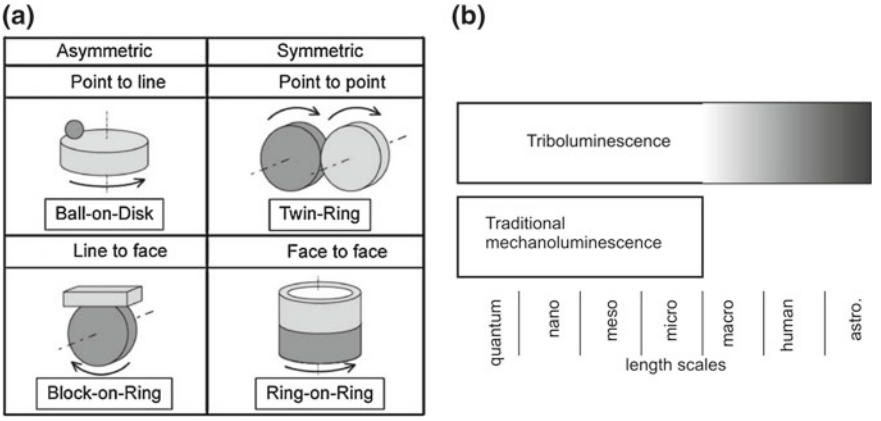


Fig. 4.3 **a** Classifications of contact configurations of four types of sliding-motion tribometers (reprinted from [46] © 2011, with permission from Elsevier). **b** Length scales of traditional mechanoluminescence and triboluminescence processes

Though in the past several experimental techniques for measuring the TL have been developed, only few of them can be considered as true tribological systems, i.e., the systems where two surfaces slide and/or roll past each other [39, 44, 47, 48], while most of other devices are aimed at grinding, milling, fracturing or straining the material being studied [18, 49, 50]. The reason for that relies on the misleading use of the term *triboluminescence* instead of *deformational luminescence* or *fracto-luminescence*.

Simultaneous scanning of both the friction force being a measure of the rate of mechanical energy dissipation and the TL is essential to reveal common mechanisms lying behind these processes. However, it is not an easy task because in the laboratory tests the TL occurs directly in the buried interface or in the close proximity to the contact zone, typically at the distance below 1 mm from it. If both solids being in frictional contact are non-transparent in the spectral range of interest for the TL study, the emitted photons can only be collected from the sides of the contact zone. In case one or both solids are transparent in this spectral range, the TL can be recorded either from the side or through the transparent solid.

The necessity to focus the optical system on a tiny luminescent zone buried in the interface rules out the tribometer configurations in which the contact zone is not stationary. This is why a pin-on-disk configuration with a stationary pin and rotating disk has been normally used in most of the TL studies. Schematic drawing of a typical experimental system with a pin-on-disk configuration aimed at the TL measurement through a transparent disk is shown in Fig. 4.4a. Light from the luminescent zone at the frictional contact passes through the quartz disk (Fig. 4.4b), a sapphire viewport and is condensed by lenses of an optical microscope. The light beam can be directed to a high-sensitive camera (in this case CCD) or to a spectrometer for the analysis of the spectral characteristics. Quartz and sapphire have been chosen because of the

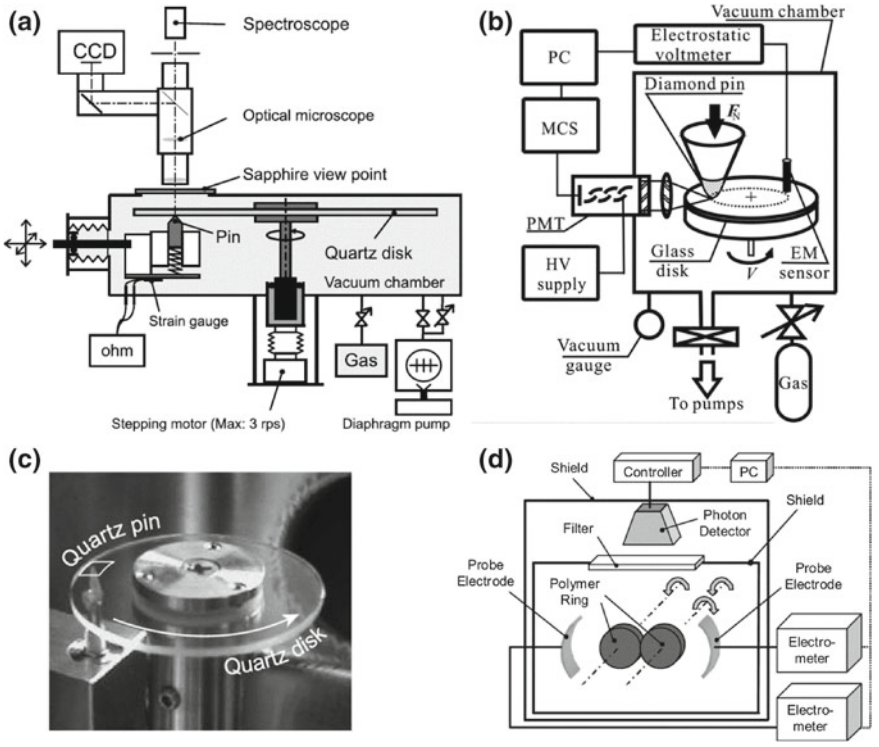


Fig. 4.4 **a** Schematic drawing of a pin-on-disk set-up in which spatial distributions and spectroscopic characteristics of the TL can be studied. Light passing through the quartz disk is measured from the opposite side of the contact by an optical microscope combined with a CCD camera and a spectrometer. **b** *Photograph* of the quartz pin and quartz disk. **c** Schematic drawing of the experimental vacuum tribometer with a pin-on-disk configuration in which the TL is measured from the side using a photomultiplier tube (PMT). Electric impulses at the anode of a PMT were acquired by a multichannel scaler (MCS). Surface potential was measured by an electrostatic voltmeter using an electromechanical (EM) sensor. **d** The twin-ring tribometer with a chemiluminescence analyzer. Photons passing through a filter are detected by a photomultiplier. Charges induced on the surfaces by rubbing are measured by an electrometer through an air probe electrodes which were set at the farthest end of the specimens without contacting them. (**a**) and (**b**) reprinted from [47] © 2006, with permission from Elsevier; (**c**) reprinted from [51] © 2003, with permission from American Institute of Physics; (**d**) reprinted from [52] © 2012, with permission from Elsevier

transmission in the ultraviolet (UV) and visible (vis) ranges. The disk and the pin are placed inside a vacuum chamber, so both the gas pressure and the composition can be controlled. By introducing specific optical filters in the optical path of the microscope the TL can be studied in different spectral ranges.

Another example of a TL pin-on-disk system in which the emitted photons are measured from the side is shown in Fig. 4.4c. In this system light is condensed by a lens which is used also as a viewport of a vacuum chamber and guided to a photomultiplier tube (PMT). The output signal of the PMT is acquired by a multichannel

scaler (MCS). The PMT can operate in either current mode or pulse count mode. In the pulse count mode, nanosecond resolution of the TL signal can be achieved.

The disadvantages of the pin-on-disk configuration are related to asymmetric sliding conditions: different wear rate of the disk and the pin, extremely large difference in the area of the contact zone on the pin and on the disk, dissimilar adsorption conditions and so on [52]. These shortcomings have been removed in the twin-ring tribometer (Fig. 4.4d) [46]. In addition to the sliding motion this tribometer allows to do pure rolling or rolling/sliding with controlled slip ratio. Due to symmetrical configuration the surface charge can be measured on both surfaces and simultaneously with the TL.

4.4 Characteristics of the TL

4.4.1 Spatial Distribution of the TL at a Tribological Contact

K. Nakayama and his research group were probably the first who succeeded to measure two-dimensional spatial distribution of the TL at a sliding contact. They used a pin-on-disk tribometer situated in the ambient air [39, 53, 54] or in a vacuum chamber [54–56] and intensified CCD camera equipped with CaF_2 lenses transparent in UV, vis and IR ranges.

Figure 4.5a shows the unfiltered pseudo-colour image of in plane spatial intensity distribution of the TL taken through the disk. The images represent integrated TL intensity in approx. 12 disk revolutions. TL completely surrounded the contact zone and spread out to the distances up to 200 μm from the centre of the contact in the direction of the relative disk motion (shown by the arrow). The total TL was the most intensive in the zone situated behind the contact at a distance of approx. 50 μm from its centre. Figure 4.5b shows the intensity distribution of the TL intensity in UV range. For the total TL and UV TL the shape of the intensity distributions were

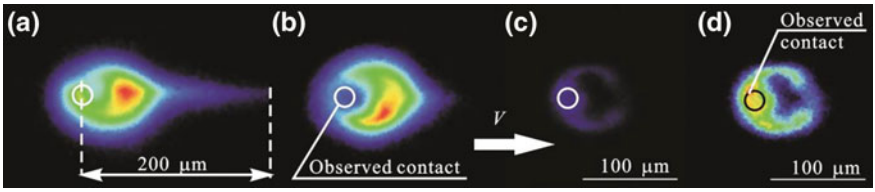


Fig. 4.5 Two-dimensional spatial distribution of the TL at the sliding contact of a diamond pin (radius of curvature 300 μm) and a sapphire disk in ambient air. **a** unfiltered pseudo-colour *image*; **b** band-filtered pseudo-colour *image* of the TL in the range 290–420 nm; **c** band-filtered pseudo-colour *image* of the TL in the range 720–2800 nm; **d** enhanced *image* from (**c**). A *circle* in each *figure* denotes the size and position of the nominal contact area determined visually using optical microscopy (reprinted from [54] © 2003, with permission from ASME)

quite similar, although small variation in shape can be ascribed to the fluctuations of the TL in time [39] since the images in Fig. 4.5 were acquired sequentially. The total luminescent power was also quite similar in both cases. That was taken as indicating that the TL arose mainly from the excitation of nitrogen from the surrounding air. Spectroscopic measurements generally corroborated this hypothesis (see Sect. 4.4.2).

The TL intensity in IR range was low (Fig. 4.5c) and exhibited unexpected distribution. Figure 4.5d presents the enhanced image of the IR TL from Fig. 4.5c. Apart from IR radiation at the contact area that can be ascribed to the frictional heating there were two lateral branches aside from the friction track protruding from the contact zone in the direction of disk rotation. The shape of the IR TL intensity distribution was defined as an open circle or a horseshoe. The most surprising was the fact that IR luminescence on the zones of the branches was much more intense than on the friction track behind the contact although these branch zones had not been in contact with the pin. These findings suggested that the gas discharge at a friction contact can be an important channel of energy dissipation and a source of the surface/gas heating inside and outside the friction track.

Miura et al. [53] found that the TL can also occur in front of the sliding contact. This luminescent zone, designated γ in Fig. 4.6, situated at a distance from the centre of the contact almost twice larger than the luminescent zone β behind the contact. Nitrogen excitation in electrical gas discharge was identified as the main mechanisms of the TL in the zones β and γ since most of the light was emitted in the UV range. Other regular and irregular patterns of the TL intensity distribution were described by various authors, e.g., four strong emission points connected each other and surrounding a dark zone just behind the contact to form a ring, elongated tails, irregular strips aside the friction track which spanned for more than hundred micrometres behind the contact and so on [55, 57, 58]. The characteristics of some of these TL features and the reasons of their occurrence are discussed in the following sections.

The size and the shape of the luminescent zone strongly depended upon the surrounding gas pressure. Figure 4.7 shows a series of pseudo-colour images of the TL under various pressures of dry air [55]. While the gas pressure decreased from the

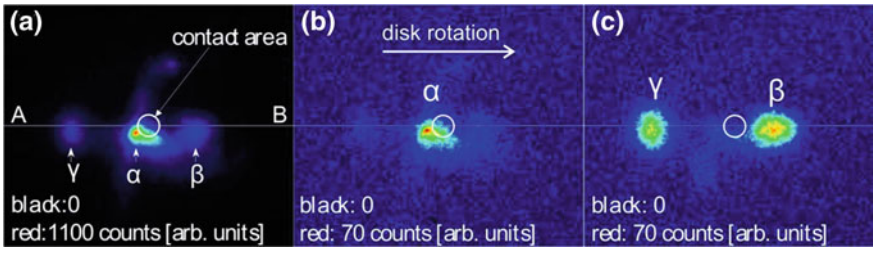


Fig. 4.6 **a** unfiltered pseudo-colour *image* of the TL during sliding in air; **b** band-filtered pseudo-colour *image* of the TL in the range 610–780 nm; **c** band-filtered pseudo-colour *image* of the TL in the range 300–420 nm. The size of all images is 250 × 300 μm^2. *Red colour* corresponds to the maximum intensity indicated in each *figure*. A *circle* shows the position and the size of the contact area (reprinted from [53] © 2001, with permission from American Institute of Physics)

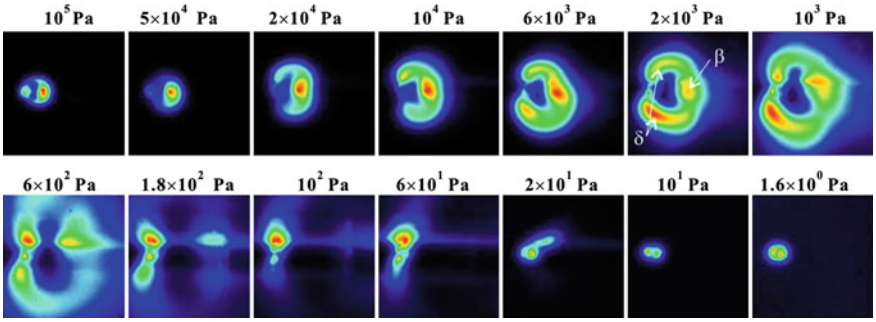


Fig. 4.7 Unfiltered pseudo-colour images of the TL obtained during sliding of a diamond pin on a sapphire disk under various pressures of dry air (reprinted from [55] © 2004, with permission from Elsevier)

atmospheric to 600 Pa the size of the luminescent zone expanded more than six-fold. It deserves special attention the fact that when the gas pressure decreased below 10 kPa the zone of the maximum TL intensity shifted from the position behind the contact (β) to lateral positions aside the contact zone designated hereinafter δ. When the gas pressure further decreased below 600 Pa the TL zone rapidly shrank. The luminescent zones β, γ and δ disappeared below 20 Pa, whereas weak light emission could be observed solely at the contact zone (zone α in Fig. 4.6). Though the expansion of the luminescent zone with the decreasing gas pressure followed by its collapse resembles typical behaviour of gas discharge, e.g., glowing discharge in a tube [59] or corona [60], the behaviour of the TL is much more complex because of the complex geometry and the configuration of the electric field.

4.4.2 Effect of the Ambient Gas and the Material of the Counterbodies on Spectral Characteristics and Intensity Distribution of the TL

When the TL of insulating materials occurs in air, important part of light emission comes from lightning as indicated by the characteristics lines of N_2 discharge in the TL spectrum [54, 57, 61]. Nakayama and Nevshupa [54] studied the TL spectra for diamond pin—sapphire disk configuration in various gas and compared the obtained TL spectra with the glowing discharge spectra of the same gases. They found that the TL emission spectra in air, N_2, O_2, He and CH_4 were very similar to the spectra of the glowing discharge, although the proportion between various peaks was slightly different (compare, for example, Figs. 4.8a and b, 4.9a and d). In H_2 the glowing discharge spectrum had one strong line $H_\alpha = 656.279$ nm, which, however, was not found in the TL emission spectrum. In ethane and propane neither the TL nor the glowing discharge spectra had any noticeable feature.

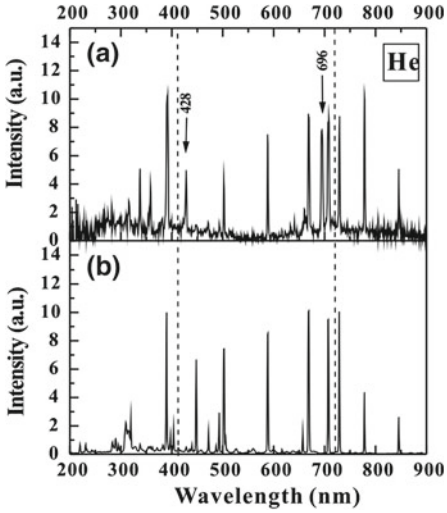


Fig. 4.8 **a** The TL spectrum for diamond-sapphire sliding in He atmosphere. **b** The luminescence emission spectrum of glowing discharge in He (reprinted from [54] © 2003, with permission from ASME)

Apart from the emission corresponding to the gas discharge a line at approximately 696 nm was observed in all TL spectra with exception for ethane and propane, but not in the glowing discharge spectra of the same gases (Figs. 4.8 and 4.9a). This line was only found when sapphire disk was used in the test (Fig. 4.9a–c). In accurate study Nakayama [57] confidently demonstrated that this line should be attributed to the ^{2}E metastable state luminescence of Cr impurities in sapphire (694.5 nm). Also he showed that the irregular stripes on the sides of the friction track in the images of the TL spatial distribution corresponded to this luminescence component. These stripes began at the lateral branches of the TL and spanned more than 100 μm behind the contact zone [55, 57]. While the intensity of the gas discharge luminescence stabilized after 2–3 min, the intensity of the induced luminescence from sapphire slowly increased in time and reached a steady value after 7 min that corresponded to approximately 1,100 disk revolutions [54]. Though the excitation mechanism for the secondary luminescence from sapphire is not clear yet, both the gas discharge luminescence and mechanical activation should be ruled out since the luminescence occurred outside the friction track and the origin of the luminescent zones did not coincide with the zone of the maximal discharge luminescence. This luminescence can probably be excited by electron [43, 57] or ion [42, 62] bombardment.

The ease of gas discharge increases as the pressure decreases, but it decreased again when the density becomes too low and the mean free path of electrons becomes too large for sustaining the discharge. In high vacuum the gas discharge is suppressed and the light emission occurred mainly from the contact zone. This explains the observed behaviour of the spatial distribution of the TL (Fig. 4.7). Miura et al. [62] studied spectral characteristics of light emitted from the contact zone in high vacuum for different pin materials. For diamond pin and quartz disk configuration they observed three broad bands at 400, 630 and 900 nm. The measured bands

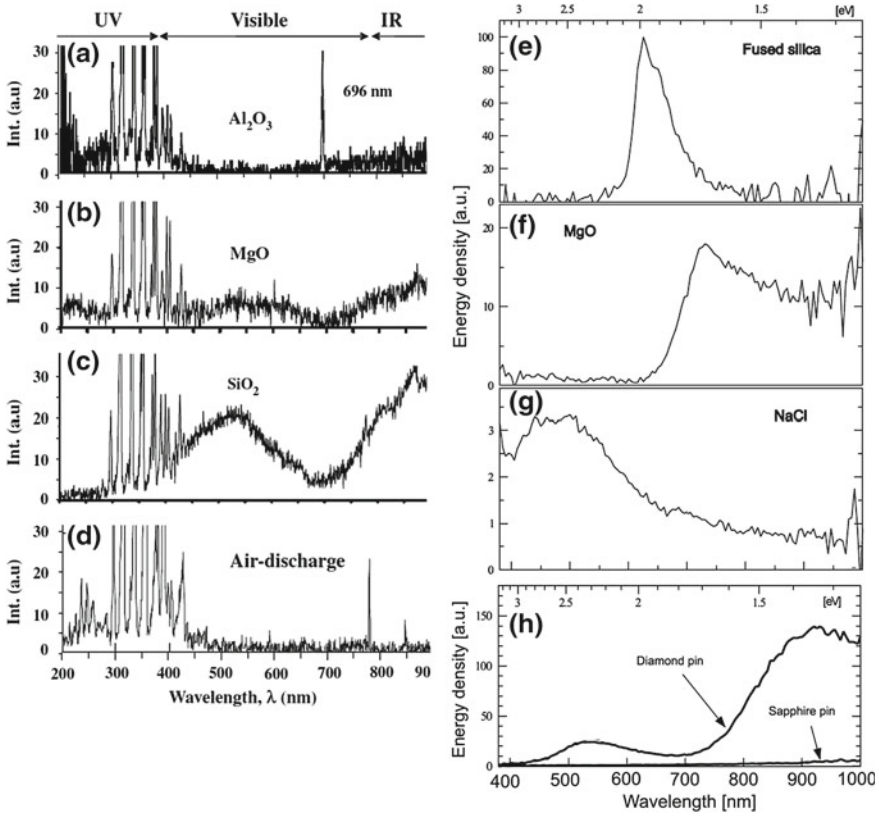


Fig. 4.9 Effect of material of the disk (**a**)–(**c**) and the pin (**e**)–(**h**) on the spectral characteristics of the TL. **a** A natural diamond pin and a synthetic sapphire disk; **b** a natural diamond pin and a MgO disk; **c** a natural diamond pin and a crystal SiO_2 disk; **d** spectrum of the luminescence of glowing discharge in air; **e** a synthetic fused silica pin and a synthetic sapphire disk; **f** a MgO single crystal pin and a synthetic sapphire disk; **g** a NaCl pin and a synthetic sapphire disk; **h** a natural diamond pin and a synthetic sapphire disk. **a**–**c** measured in ambient air; **e**–**h** measured in vacuum (10^{-3} Pa). (**a**)–(**d**) reprinted from [57] © 2010, with permission from Springer Science and Business Media. (**e**)–(**h**) reprinted from [62] © 2009, with permission from Elsevier

significantly differed from the theoretical Planck's distribution and did not depend on the sliding conditions. These findings led the authors to the conclusion that thermal radiation was not the dominant mechanism of the luminescence excitation. The band at 900 nm (Fig. 4.9h) was caused by fractoluminescence of diamond. In addition to the broad band at 900 nm another weak band between 350 and 650 nm was observed during diamond abrasion (Fig. 4.9h). For the ionic crystals studied in [62] the TL emission spectra were very similar to the corresponding fractoluminescence spectra and were related to electronic excitation of vacancies or impurities during abrasion. For quartz, the band at 630 nm (Fig. 4.9e) was ascribed to fractoluminescence of non-bridging oxygen hole centres. Broad band in the TL emission spectrum of MgO

(Fig. 4.9f) was attributed to the emission of Mn and Cr^{3+} impurities and was similar to thermoluminescence and photoluminescence spectra of MgO.

4.4.3 Effect of Friction Type and Humidity on the TL and Triboelectrification of Polymers

Despite a number of works on triboluminescence of ceramics, only limited researches have been carried out in the field of polymers. Nowadays, polymers are used for numerous machine parts and the TL phenomena can have undesired effect on their operation life, safety and reliability. As polymers can be easily charged during friction, triboluminescence is likely to occur wherever polymers are in rubbing contact. However, not much attention has been paid on the TL of polymers since the TL usually occurs in the UV range, while if the TL is visible, its effect is often neglected. The gas discharge associated with the TL from polymers releases charged particles and photons of extremely high specific energy.

The mechanism of Triboelectrification (TE) is essential for the process of the TL. However, it is difficult to measure how much charge has been generated. This is because the measured charge on each surface reflects the resultant charge after the recombination of charges as a result of the discharge between them as well as charge leakage through the bulk or on the surface. However, charge can be measured when the gas discharge is suppressed and the materials of high electrical resistance are used. The twin-ring tribometer (Figs. 4.3a and 4.4d) is the most appropriate for this purpose, because it enables the sequential measurements of both the TE and the TL. Additionally, this test rig has the advantage of realizing the rolling and sliding contacts by just reversing the rotation direction of one specimen. The following results are obtained from this apparatus.

Figure 4.10 compares the charging processes of polyamide (PA66) and polytetrafluoroethylene (PTFE) in sliding and rolling contacts together with the associated photon emission. Sliding or rolling occurred in the period between 25 and 75 s of the data records. During sliding the TL intensity was two orders of magnitude higher than during rolling, whereas the saturation charge in both experiments was almost the same. It should be mentioned that during the first revolution of the cylinders in the sliding test the discharge did not occur, therefore the measured surface charge reached the maximum. In rolling friction, the surface charges gradually increased with time. These clearly indicated that the charge accumulation rate during sliding was higher than during rolling. This fact can be attributed to the multiple real contacting of the same surface zones during sliding. The decrease in the TL intensity during sliding which is seen in Fig. 4.10a can be related to the gradual transfer of PTFE onto PA66, which would transform the rubbing pair from PA66/PTFE to PTFE/PTFE. Therefore, the charging rate per contact/separation on similar material decreased leading to the corresponding decrease in the TL intensity.

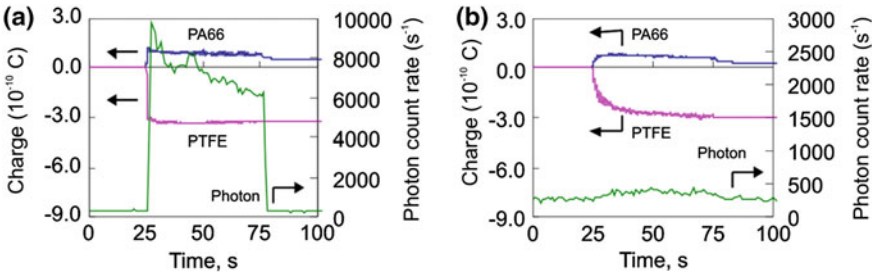


Fig. 4.10 TE and TL processes during sliding (**a**) and rolling (**b**) between PA66 and PTFE. (reprinted from [52] © 2012, with permission from Elsevier)

The fact that the TL from the sliding contact of polymers reached its maximum soon after the onset of rubbing indicated that the luminescence could not be attributed to the frictional heating. Rather, photon emission was triggered by the electron emission which could be caused by the following two mechanisms: fracto-emission and field emission. In contrast to rolling friction, sliding friction produces continuous fracture of the material in the contact zone. Therefore, electrons can be emitted from the fractured surface. Frictional electrification contributes to the build-up of a strong electric field which can induce field electron emission.

Environmental humidity is one of the most influential factors in triboelectrification [63], which would then give rise to the change in TL. Figure 4.11 shows a typical result of the effect of humidity on the TE and TL from the rubbing between acrylonitrile butadiene styrene (ABS) and PTFE. When relative humidity (RH) increased from 0 to 58 % the TL intensity also increased. This process was associated with the increase in the surface charge that was especially intensive between 10 and 58 %. However, the saturation charge on both polymers at 10 % RH was higher than at both 0 and 58 %. The relationship between the charging rate/saturation charge and photon amount led to the conclusion that TL rate is determined by the charging rate per one frictional contact and not by the saturation charge on the surface. It is worth mentioning that almost no TL could be observed at 0 % RH, although the leakage rate is expected to be the lowest. This finding can be due to the limited number of charge carrier which

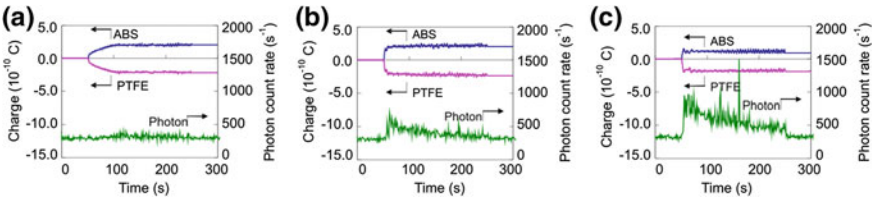


Fig. 4.11 Effects of humidity on the triboelectrification and triboluminescence in the rubbing between ABS and PTFE at various relative humidity: **a** 0 %, **b** 10 % and **c** 58 % (reprinted from [52] © 2012, with permission from Elsevier)

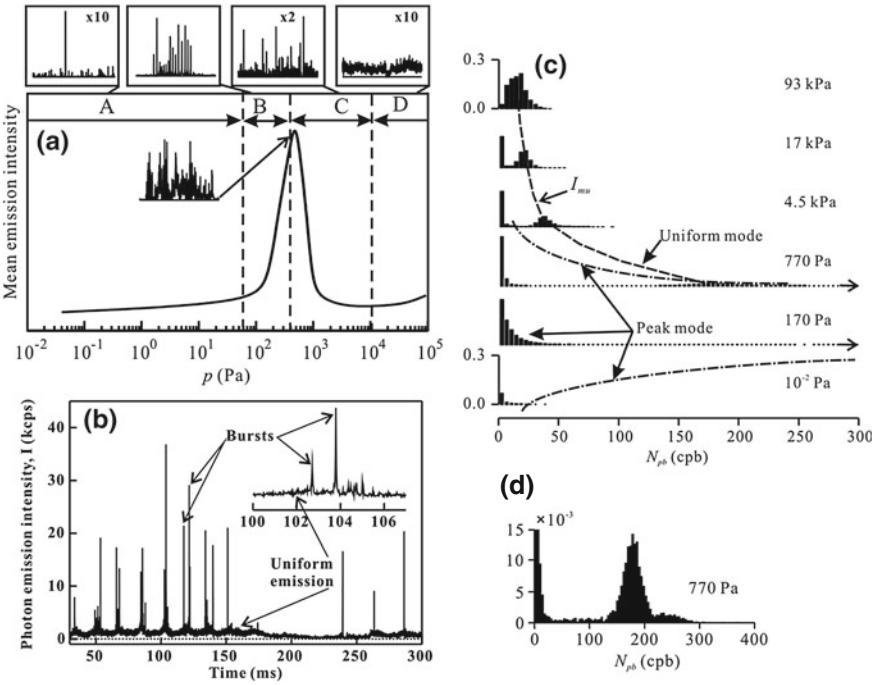


Fig. 4.12 **a** The effect of the Ar pressure on the total TL intensity and behaviour patterns (reprinted from [42] © 2013, with permission from IOP Publishing Ltd); **b** fragment of the TL time series measured with a resolution of 20.48 μs (reprinted from [51] © 2003, with permission from American Institute of Physics); **c** histograms of the TL time series distributions on intensity at various pressures of Ar atmosphere; **d** a histogram of the TL time series distributions at gas pressure 770 Pa showing two components: the *bell-shape* and the exponential-like

inhibits a subsequent discharge between surfaces. This means that charges have to be carried by ions (H_3O^+ and OH^-) generated from the adsorbed water [64]. On the other hand, when the relative humidity increased to a certain value, the surface charge decreased due to discharge and recombination.

4.4.4 Behaviour of the TL on Different Time Scales

Gas pressure is an important external parameter not only for the intensity and spatial distribution of the TL related to the gas discharge, but also for its behaviour on various time scales [42, 51]. In the experiments with a diamond pin—sapphire disk configuration four behaviour patterns of the TL time series depending on the Ar gas pressure were distinguished (Fig. 4.12a) [42]: at $p < 20$ Pa the TL was of low intensity with sparse irregular bursts. In the range $20\,\text{Pa} < p < 300\,\text{Pa}$ the TL had a regular pattern consisting of a series of groups of almost equally spaced bursts (Fig. 4.12b).

The period between the bursts varied from 20 to 40 ms, whereas the groups of the bursts appeared synchronously with the disk rotation. In the range 300 Pa $< p <$ 10 kPa, uniform emission with irregular bursts occurred. At the pressure higher than 10 kPa only weak uniform emission was observed. The appearance of burst with the frequency of 33 Hz is puzzling since the typical frequencies of the gas discharge oscillation is three to four orders of magnitude higher.

Density distribution functions of the PMT output time series on the number of photon counts per time bin are shown in Fig. 4.12c. The distributions have either exponential-like or the bell-shaped distributions, or the combination of both (see, for example, Fig. 4.11d). The exponential-like distribution component was ascribed to the burst TL mode, while the bell-shaped one was attributed to the uniform mode. From the analysis of the density distribution function it was found that the uniform mode increased in intensity when the gas pressure decreased from 100 kPa to 300 Pa and vanished below 300 Pa. The burst mode also increased with decreasing gas pressure and reached maximum at 170 Pa. Then it decreased, but not totally vanished.

4.5 Modelling Approach

Electric gas discharge being an important excitation mechanism for the TL still remains poorly understood. From the point of view of electrostatics, pin-on-disc configuration can be considered as an extreme case of a point-to-plane gap with a point being in contact with a plane. In the case of insulating electrodes the gas discharge, referred to as barrier or silent discharge, is likely to be a very complex process because of the displacement field and very low mobility of charges in the electrodes [65]. Furthermore, in contrast to traditional barrier discharge being driven by alternating feeding potential, the polarities of the pin and the disk are rather stable [42, 51].

When embarking on a study of electric gas discharge at a sliding contact the first question which should be addressed is related to the configuration and the strength of electric field. Bearing in mind the constrain geometry and possible alteration of the electric field by introducing conductive measuring electrodes, direct measuring of the electric field in a gap around the contact zone is not a trivial task. The electric field can be roughly estimated if the rate of the charge generation is known. The latter was determined from the surface potential on the disk measured in high vacuum when the gas discharge is hindered (Fig. 4.13a) [66]. For the sliding contact composed of a diamond pin (r 1000 μm) and a sapphire disk under the normal load 0.96 N and sliding velocity 4.4 cm s^{-1} (1 rps) the rate of the charge generation was 3.39×10^{-11} A. Then, taking into account the geometry of the friction zones on the pin and the disk and assuming that, in this rough and ready approach, the charge leakage through the solid and on the surfaces can be neglected, the configuration of the electric field was modelled using finite elements method (Fig. 4.13b).

The negative charge on the disk had no effect on the radial component of the electric field, which was dominant, and only minor effect on the axial component. It was concluded that concentrated positive charge on the pin controlled the overall

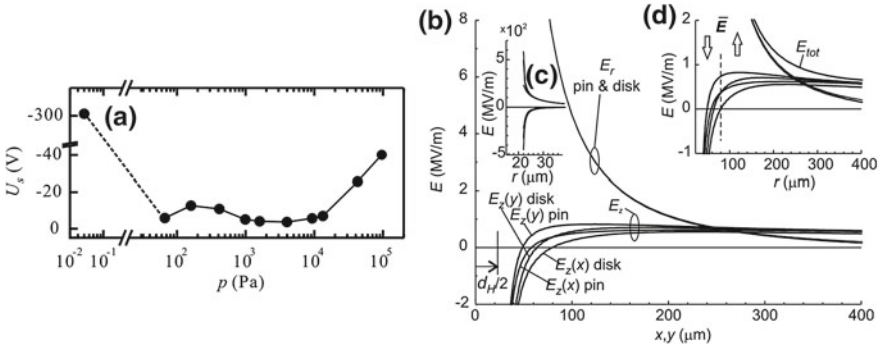


Fig. 4.13 **a** Mean electrostatic potential measured 2 mm above the friction zone on the disk as function of gas pressure. **b** Radial, E_r, and axial, $E_x(x)$ and $E_y(y)$, components of the electric field along the *x*- and *y*-*axes* on the disk and pin surfaces near the contact zone. *Insets* **c** and **d** show the *graphs* on different scales (reprinted from [66] © 2013, with permission from IOP Publishing Ltd)

configuration of the electric field. This can explain the experimental finding that the gas discharge occurred even from the disk which had a 5 nm thick conductive Al coating [67]. The electric field was highly non-uniform that contrasted with the common believe and other models based on the assumption of near symmetric bipolar configuration of the electric field in the gap [57, 68]. Such configuration of the electric field with a highly stressed anode and a low-field cathode is typical for positive corona discharges.

The strong and negative axial component of the electric field near the contact zone provided necessary conditions for the emission of triggering electrons through a field emission mechanism. An analysis of impact ionization and secondary processes at various gas pressures [42] showed that the corona model provided plausible explanations for the experimentally observed TL behaviour (Sect. 4.4.4) consisting of various combinations of two main modes. It was shown that when the region of electron collision ionization coincides with both the regions of the secondary processes and triggering electron emission, the TL emission is uniform. At lower pressures the first two regions decouple leading to the instability of the gas discharge and appearance of bursts.

Acknowledgments This work was supported through the grants of the Ministry of Economy and Competitiveness of Spain RYC-2009-0412, BIA2011-25653 and the project IPT-2012-1167-120000 with the participation of European Regional Development Fund (FEDER). One of the authors (R.N.) also acknowledges the contribution of the COST action TD1208.

References

1. "Triboluminescence", *Oxford Dictionaries*. (Oxford University Press, Oxford, 2010)
2. H. Blok, The flash temperature concept. Wear **6**, 483–94 (1963)
3. G.P.H. Gubbels, G.J.F.T. van der Beek, A.L. Hoep, F.L.M. Delbressine, H. van Halewijn, Diamond tool wear when cutting amorphous polymers. CIRP Annal. Manuf. Technol. **53**, 447–450 (2004)
4. F. Freund, Charge generation and propagation in igneous rocks. J. Geodyn. **33**, 543–570 (2002)
5. F. St-Laurent, J.S. Derr, F.T. Freund, Earthquake lights and the stress-activation of positive hole charge carriers in rocks. Phys. Chem. Earth Parts A/B/C **31**, 305–312 (2006)
6. T.V. Losseva, I.V. Nemchinov, Earthquake lights and rupture processes. Nat. Hazards Earth Syst. Sci. **5**, 649–656 (2005)
7. J. Muto, H. Nagahama, T. Miura, I. Arakawa, Frictional discharge at fault asperities: origin of fractal seismo-electromagnetic radiation. Tectonophysics **431**, 113–122 (2007)
8. C. Helling, M. Jardine, D. Diver, S. Witte, Dust cloud lightning in extraterrestrial atmospheres. Planet. Space Sci. **77**, 152–157 (2013)
9. A.A. Sickafoose, J.E. Colwell, M. Horányi, S. Robertson, Experimental investigations on photoelectric and triboelectric charging of dust. J. Geophys. Res. Space Phys. **106**, 8343–8356 (2001)
10. A.J. Walton, Triboluminescence. Adv. Phys. **26**, 887–948 (1977)
11. L.J. Kricka, J. Stroebel, P.E. Stanley, Triboluminescence: 1968–1998. Luminescence **14**, 215–220 (1999)
12. B.V. Derjaguin, N.A. Krotova, Y.P. Toporov, in *Emission of High-Speed Electrons and Other Phenomena Accompanying the Process of Breaking Adhesion Bonds*, ed. by J.M. Georges. Tribology Series (Elsevier, Amsterdam, 1981), pp. 471–87
13. T. Miura, M. Chini, R. Bennewitz, Forces, charges, and light emission during the rupture of adhesive contacts. J. Appl. Phys. **102**, 103509–103516 (2007)
14. Experience faite a l'observatoire sur le barometre simple touchant un nouveau phenomene qu'on ya découvert. Le Journal des sçavans (Académie des inscriptions et belles-lettres, Paris, France, 1676)
15. H.B. Weiser, I.I. Crystalloluminescence, J. Phys. Chem. **22**, 576–595 (1917)
16. M.P. Brenner, S. Hilgenfeldt, D. Lohse, Single-bubble sonoluminescence. Rev. Mod. Phys. **74**, 425–484 (2002)
17. Y. Tsuboi, T. Seto, N. Kitamura, Laser-induced shock wave can spark triboluminescence of amorphous sugars. J. Phys. Chem. A **112**, 6517–6521 (2008)
18. P.Y. Butyagin, The luminescence accompanying mechanical deformation and rupture of polymers. Vysokomol soyed **A12**, 290–299 (1970)
19. B.P. Chandra, K.K. Shrivastava, Dependence of mechanoluminescence in rochelle-salt crystals on the charge-produced during their fracture. J. Phys. Chem. Solids **39**, 939–940 (1978)
20. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, The nonlinear nature of friction. Nature **430**, 525–528 (2004)
21. B.P. Chandra, Mechanoluminescence of nanoparticles. Open Nanosci. J. **5**, 45–58 (2011)
22. J.-C. Zhang, C.-N. Xu, S. Kamimura, Y. Terasawa, H. Yamada, X. Wang, An intense elastico-mechanoluminescence material CaZnOS:Mn^{2+} for sensing and imaging multiple mechanical stresses. Opt Express **21**, 12976–12986 (2013)
23. E.A. Deulin, V.P. Mikhailov, Y.V. Panfilov, R.A. Nevshupa, *Mechanics and Physics of Precise Vacuum Mechanisms* (Springer, Dordrecht, 2010)
24. V.E. Orel, I.N. Kadyuk, N.N. Dzyatkovskaya, M.I. Danko, Y.I. Mel'nic, Mechanoluminescence: lymphocyte analysis after exposure to ionizing radiation. Luminescence **15**, 29–36 (2000)
25. V.E. Orel, A.V. Romanov, N.N. Dzyatkovskaya, I. Mel'nik, The device and algorithm for estimation of the mechanoemisson chaos in blood of patients with gastric cancer. Med. Eng. Phys. **24**, 365–371 (2002)

26. G. Heinike, *Tribochemistry* (Carl Hanser Verlag, Munchen, 1984)
27. B. Vick, M.J. Furey, An investigation into the influence of frictionally generated surface temperatures on thermionic emission. Wear **254**, 1155–1161 (2003)
28. R. Nevshupa, The role of athermal mechanisms in the activation of tribodesorption and triboluminisence in miniature and lightly loaded friction units. J. Frict. Wear **30**, 118–126 (2009)
29. J.D. Schall, G. Gao, J.A. Harrison, Effects of adhesion and transfer film formation on the tribology of self-mated DLC contacts. J. Phys. Chem. C **114**, 5321–5330 (2010)
30. T.-B. Ma, Y.-Z. Hu, H. Wang, Molecular dynamics simulation of shear-induced graphitization of amorphous carbon films. Carbon **47**, 1953–1957 (2009)
31. A.I. Rusanov, Thermal effects in mechanochemistry. Russ. J. Gen. Chem. **72**, 327–344 (2002)
32. P.Y. Butyagin, Problems in mechanochemistry and prospects for its development. Russ. Chem. Rev. **63**, 965 (1994)
33. P.Y. Butyagin, Structural disorder and mechanochemical reactions in solids. Russ. Chem. Rev. **53**, 1025–1038 (1984)
34. T.E. Fischer, Tribochemistry. Annual review material. Science **18**, 303–323 (1988)
35. S.L. Craig, Mechanochemistry: a tour of force. Nature **487**, 176–177 (2012)
36. B.P. Chandra, S. Tiwari, M. Ramrakhiani, M.H. Ansari, Mechanoluminescence in centrosymmetric crystals. Cryst. Res. Technol. **26**, 767–781 (1991)
37. R.A. Nevshupa, Triboemission: an attempt of generalized classification, in *Tribology: Science Applications*, ed. by C. Kajdas (PAS, Vienna, 2004), pp. 11–25
38. J. Thevenet, M. Siroux, B. Desmet, Measurements of brake disc surface temperature and emissivity by two-color pyrometry. Appl. Therm. Eng. **30**, 753–759 (2010)
39. K. Nakayama, R.A. Nevshupa, Plasma generation in a gap around a sliding contact. J. Phys. D Appl. Phys. **35**, L53–L56 (2002)
40. V.A. Kluev, T.N. Vladikina, Y.P. Toporov, V.J. Anisimova, B.V. Derjaguin, Emission phenomena accompanying the triboelectrification process in vacuum. IEEE Trans. Ind. Appl. **IA-14**, 544–546 (1978)
41. J.P. Duignan, I.D.H. Oswald, I.C. Sage, L.M. Sweeting, K. Tanaka, T. Ishihara et al., Do triboluminescence spectra really show a spectral shift relative to photoluminescence spectra? J. Lumin. **97**, 115–126 (2002)
42. R.A. Nevshupa, Effect of gas pressure on the triboluminescence and contact electrification under mutual sliding of insulating materials. J. Phys. D Appl. Phys. **46**, 185501 (2013)
43. L.M. Sweeting, Triboluminescence with and without Air. Chem. Mater. **13**, 854–870 (2001)
44. C.N. Xu, Coatings, in *Encyclopedia of Smart Materials*, ed. by M. Schwartz (Willey, New York, 2002), pp. 190–201
45. E.A. Varentsov, Y.A. Khrustalev, Mechanoemission and mechanochemistry of molecular organic crystals. Russ. Chem. Rev. **64**, 783–797 (1995)
46. K. Hiratsuka, T. Yoshida, The twin-ring tribometer–characterizing sliding wear of metals excluding the effect of contact configurations. Wear **270**, 742–750 (2011)
47. J. Muto, H. Nagahama, T. Miura, I. Arakawa, Frictional discharge plasma from natural semiconductor/insulator junctions: origin of seismo-electromagnetic radiation. Phys. Chem. Earth Parts A/B/C **31**, 346–351 (2006)
48. K. Nakayama, H. Hashimoto, Effect of surrounding gas pressure on Triboemission of charged particles and photons from wearing ceramic surfaces. Tribol. Trans. **38**, 35–42 (1995)
49. G.E. Hardy, J.I. Zink, Triboluminescence and pressure dependence of the photoluminescence of tetrahedral manganese(II) complexes. Inorg. Chem. **15**, 3061–3065 (1976)
50. L.M. Sweeting, J.L. Guido, An improved method for determining triboluminescence spectra. J. Lumin. **33**, 167–173 (1985)
51. R.A. Nevshupa, K. Nakayama, Triboemission behavior of photons at dielectric/dielectric sliding: time dependence nature at 10^{-4}-10^{4} s. J. Appl. Phys. **93**, 9321–9328 (2003)
52. K. Hiratsuka, K. Hosotani, Effects of friction type and humidity on triboelectrification and triboluminescence among eight kinds of polymers. Tribol. Int. **55**, 87–99 (2012)
53. T. Miura, K. Nakayama, Two-dimensional spatial distribution of electric-discharge plasma around a frictional interface between dielectric surfaces. Appl. Phys. Lett. **78**, 2979–2981 (2001)

54. K. Nakayama, R.A. Nevshupa, Characteristics and pattern of plasma generated at sliding contact. J. Tribol-T Asme **125**, 780–787 (2003)
55. K. Nakayama, R.A. Nevshupa, Effect of dry air pressure on characteristics and patterns of tribomicroplasma. Vacuum **74**, 11–17 (2004)
56. K. Nakayama, The plasma generated and photons emitted in an oil-lubricated sliding contact. J. Phys. D Appl. Phys. **40**, 1103–1107 (2007)
57. K. Nakayama, Triboplasma generation and triboluminescence: influence of stationary sliding partner. Tribol. Lett. **37**, 215–228 (2010)
58. K. Nakayama, Mechanism of triboplasma generation in oil. Tribol. Lett. **41**, 345–351 (2011)
59. Y.P. Raizer, *Gas Discharge Physics* (Springer, New York, 1991)
60. L.B. Loeb, *Electrical Coronas, Their Basic Physical Mechanisms* (University of California Press, Berkley, 1965)
61. T. Miura, K. Nakayama, Spectral analysis of photons emitted during scratching of an insulator surface by a diamond in air. J. Appl. Phys. **88**, 5444–5447 (2000)
62. T. Miura, E. Hosobuchi, I. Arakawa, Spectroscopic studies of triboluminescence from a sliding contact between diamond, SiO_2, MgO, NaCl, and Al_2O_3 (0001). Vacuum **84**, 573–577 (2009)
63. E. Németh, V. Albrecht, G. Schubert, F. Simon, Polymer tribo-electric charging: dependence on thermodynamic surface properties and relative humidity. J. Electrost. **58**, 3–16 (2003)
64. L.S. McCarty, G.M. Whitesides, Electrostatic charging due to separation of ions at interfaces: contact electrification of ionic electrets. Angewandte Chemie Int. Ed. **47**, 2188–2207 (2008)
65. H.E. Wagner, R. Brandenburg, K.V. Kozlov, A. Sonnenfeld, P. Michel, J.F. Behnke, The barrier discharge: basic properties and applications to surface treatment. Vacuum **71**, 417–436 (2003)
66. R. Nevshupa, Effect of gas pressure on the triboluminescence and contact electrification under mutual sliding of insulating materials. J. Phys. D Appl. Phys. **46**, 185501 (2013)
67. R.A. Nevshupa, K. Nakayama, Effect of nanometer thin metal film on triboemission of negatively charged particles from dielectric solids. Vacuum **67**, 485–490 (2002)
68. K. Nakayama, M. Tanaka, Simulation analysis of triboplasma generation using the particle-in-cell/Monte Carlo collision (PIC/MCC) method. J. Phys. D Appl. Phys. **45**, 495203 (2012)

Chapter 5
The Quartz Crystal Microbalance as a Nanotribology Technique

Lorenzo Bruschi and Giampaolo Mistura

Abstract The quartz crystal microbalance (QCM) technique is a powerful probe of interfacial phenomena that has been successfully employed to investigate the sliding friction of objects of nanoscopic size subject to lateral speeds as large as a few m/s. After a description of the quartz acoustics, the chapter presents the more common circuits used to drive the QCM and discusses the main problems in the application of such a technique to the study of nanotribolgoy; the quality of the surface electrodes and surface contamination.

5.1 Introduction

The quartz crystal microbalance (QCM) technique is a powerful probe of interfacial phenomena that has been successfully employed to investigate the sliding friction of objects of nanoscopic size subject to lateral speeds as large as a few m/s [1, 2]. The microbalance is a small quartz disk whose principal faces are optically polished and covered by two metal films, which are used both as electrodes and as adsorption surfaces. By applying an AC voltage across the two electrodes, it is possible to drive the crystal to its own mechanical resonance with the two parallel faces oscillating in a transverse shear motion. The quality factor of these resonances is usually very high ($\gtrsim 10^5$) and this explains why the QCM is quite sensitive to interfacial phenomena. A change in the disk inertia, as caused, for example, by the adsorption of a film on the metal electrodes, is signalled by a shift in the resonant frequency. Similarly, any dissipation taking place in the system determines a decrease in the resonance amplitude.

L. Bruschi (✉)
CNISM Unitá di Padova, via Marzolo 8, 35131 Padova, Italy
e-mail: lorenzo.bruschi@unipd.it

G. Mistura
CNISM and Dipartimento di Fisica e Astronomia G. Galilei, Universitá di Padova, via Marzolo 8, 35131 Padova, Italy
e-mail: giampaolo.mistura@unipd.it

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_5

Thanks to the pioneer work of Krim and coworkers [3], it was found that molecules and atoms weakly bound to the surface of a quartz crystal can slip relative to the oscillating substrate. The slip occurs as a result of the force of inertia F acting on the adsorbates during the vibrational motion of the crystal. Actually, because of its extremely small value, the force F induces a slow, thermally activated motion of the adsorbate along its direction, with a drift velocity proportional to F. Most of the friction experiments carried out so far with the QCM technique have dealt with molecularly thin films of simple gases adsorbed at low temperatures on the metal electrodes, generally gold or silver, evaporated over the faces of a quartz crystal. For such studies, the QCM was implemented in standard cryostats that guarantee very good temperature controls, of the order of a few mK or better [4–7]. In this way, it was possible to achieve very stable quartz resonance curves, an essential requirement to detect the tiny dissipation associated with the friction of very thin films. However, more systematic and quantitative investigations now require the use of very uniform and clean surfaces, well characterized at the microscopic level by techniques like STM or AFM. Furthermore, it is also important to change in-situ and in a controlled way the morphology of the surface. In other words, it is necessary to design a new generation of QCM experiments that combine cryogenics with surface science [8].

In the literature, there are several general reviews on the application of the QCM to nanotribology. However, none of them has covered in detail the experimental aspects of such a technique. Therefore, in this chapter, after a brief description of the acoustics of the QCM, we present the main electronic circuits used to drive the QCM, emphasizing their pro and con. We then discuss the surface quality of the quartz electrodes and conclude with a short summary of the main features of a new apparatus that we have expressly built to fulfil the requirements of low temperatures and an ultra-high-vacuum environment. Finally, we present some preliminary data acquired with set-up that seem to suggest structural depinning of Ne films adsorbed on Pb(111) at a coverage above 0.4–0.5 layers.

5.2 The Acoustics of Quartz Crystal

The more common quartz crystals used in interfacial physics are of the so called AT-cut. It has been extensively used in the electronics because the temperature dependence of its resonance frequency is very small around room temperature. Other types are the SC-cut (stress-compensated) quartz crystal, which is insensitive to radial stresses, although the minimum of the quartz resonance versus temperature curve occurs close to 200 °C. Furthermore, the SC-cut crystals are much more expensive and require a complicated driving circuit with respect to the more common AT-cut crystals. In practice, QCMs employing SC-cut crystals have been used so far in adsorption studies but not in the field of nanotribology [9].

Let us now consider an AT-cut quartz plate, which is characterized by a shear motion of its two parallel faces. The AT crystal has a natural mechanical resonance when the plate thickness h is half of the transverse mode wavelength λ, or an odd

multiple of $\lambda/2$, e.g. $h = n\frac{\lambda}{2}$, where n is called the overtone number ($n = 1$ is the fundamental mode, $n = 3$ is the third overtone…). At room temperature, the resonance frequency of such a plate oscillating in vacuum is related to its thickness h by the simple relation:

$$f_{0,n} = \frac{1.75n}{h} - C \tag{5.1}$$

where $f_{0,n}$ is measured in MHz, h in mm and C is a small correction factor which increases with electrode thickness [10] . (Typical values of $f_{0,1}$ for AT plates employed as QCM sensors lie in the range 1–10 MHz).

At a certain frequency f*, its behavior in vacuum can be described by a complex acoustical impedance:

$$Z_0 \equiv R_0 - jX_0 = R_0 - j\pi\, n\, A\, Z_q \frac{f^* - f_{0,n}}{f_{0,n}} \tag{5.2}$$

where A is the area of one electrode, $Z_q = 8.862 \times 10^5 g/cm^2$ sec is the quartz acoustic impedance and the dissipative term R_0, which accounts for all the losses in the plate, is related to the quality factor Q_0 via

$$\frac{1}{Q_0} = \frac{2R_0}{\pi\, n\, A\, Z_q} \tag{5.3}$$

When the quartz plate is immersed in a fluid, its impedance will change because of the adsorption of a film onto the quartz surfaces and of the viscous coupling with the surrounding vapor. The global contribution per unit area can be expressed in terms of a complex impedance $R_{sfv} - jX_{sfv}$, in series with Z_0. If both faces of the quartz plate are exposed to the fluid, the total dissipative and inertial terms become, respectively, $R_0 + 2AR_{sfv}$ and $X_0 + 2AX_{sfv}$. The quality factor will then decrease by an amount $\Delta\frac{1}{Q}$ equal to:

$$\Delta\frac{1}{Q} \equiv \frac{1}{Q} - \frac{1}{Q_0} = \frac{4R_{sfv}}{\pi\, n\, Z_q} \tag{5.4}$$

and the resonance frequency f will also be diminished by:

$$\Delta f \equiv f - f_{0,n} = -2X_{sfv}\frac{f_{0,n}}{\pi\, n\, Z_q} \tag{5.5}$$

The exact shifts will obviously depend on the explicit forms of R_{sfv} and X_{sfv}. In order to determine them, we have applied the linearized Navier–Stokes equation to the combined system quartz crystal-adsorbed film-bulk vapor [11]. Let d be the thickness of the adsorbed film and ρ_f and η_f its bulk mass density and viscosity, respectively, while η_v and ρ_v represent the viscosity and the density of the bulk vapor. If we assume, as customary, that the transverse velocity field depends only on the vertical distance z from the electrode surface, the general stationary solutions to

the Navier–Stokes equations in the vapor and film regions are determined apart from four integration constants.

These can be univocally determined by imposing the following boundary conditions on the velocity fields v_f and v_v: (a) $v_v = 0$ very far from the film; (b) $v_v = v_f$ at the film-vapor interface (i.e. no slippage at this boundary); (c) at this interface, the force exerted by the vapor on the film must be equal to that caused by the film on the vapor, that is $\eta_v \left(\frac{dv_v}{dz}\right)_{z=d} = -\eta_f \left(\frac{dv_f}{dz}\right)_{z=d}$; (d) we assume, in general, that there may be slippage at the solid-film interface. Because of this, there will be a frictional force F_{sf} at this boundary. As the last condition, we then impose that the force F_{sf} must be equal and opposite to that due to the film, that is: $F_{sf} = +\eta_f \left(\frac{dv_f}{dz}\right)_{z=0}$. Finally, we make the plausible assumption that F_{sf} depend linearly on the relative velocity between the quartz plate and the film:

$$F_{sf} = -\eta_2 \left[v_0 - v_f(0)\right] \tag{5.6}$$

where η_2 is called coefficient of sliding friction or interfacial viscosity, v_0 is the velocity of the electrode and $v_f(0)$ that of the film at the electrode surface. This condition is consistent with recent QCM studies of the velocity dependence of the interfacial friction [12]. If there is no slippage at the solid-fluid interface, $\eta_2 = \infty$. The opposite limit, $\eta_2 = 0$, corresponds instead to a superfluid whose motion is totally decoupled from that of the oscillating substrate.

By carrying out the necessary algebra, one finds that the reciprocal of Z_{sfv} can be easily rewritten as:

$$\frac{1}{Z_{sfv}} = \frac{1}{Z_v + Z_{fd}} + \frac{1}{\eta_2} \tag{5.7}$$

which says that the total acoustic impedance Z_{sfv} of the combined system substrate-film-vapor can be considered as the parallel between the series of the vapor i mpedance, Z_v, and that of the film Z_{fd}, and the impedance η_2 due to the slippage of the film at the solid boundary.

The formula 5.7 means that it is possible, at least in principle, to measure the friction force of a film adsorbed on a solid surface with a quartz microbalance. In nanotribology one is interested in studying the friction of an adsorbed monolayer. This implies that the acoustic impedance of the film can be simplified as

$$Z_{fd} \simeq -j\,\omega\,\rho_f\,d \tag{5.8}$$

where $\omega = 2\pi f$.

If we solve the 5.7 in terms of R_{sfv} and X_{sfv} we get:

$$\frac{X_{sfv}}{R_{sfv}^2 + X_{sfv}^2} = \frac{\omega\rho_f d + X_v}{R_v^2 + (\omega\rho_f d + X_v)^2} \tag{5.9}$$

and

$$\frac{R_{sfv}}{R_{sfv}^2 + X_{sfv}^2} = \frac{R_v}{R_v^2 + (\omega\rho_f d + X_v)^2} + \frac{1}{\eta_2} \tag{5.10}$$

The first equation yields the film thickness d as

$$d = \frac{1}{2\omega\rho_f}\left[R_{sfv}^2 + X_{sfv}^2 + \sqrt{\left(R_{sfv}^2 + X_{sfv}^2\right)^2 - 4R_v^2\, X_{sfv}^2}\right] - \frac{X_v}{\omega\rho_f} \tag{5.11}$$

which can be substituted in the second one in order to calculate the interfacial viscosity η_2.

Finally, the slip time τ_s, which represents the time required for the adsorbed film speed to decay to 1/e of its initial value after that the oscillating substrate has been put to rest in the absence of a bulk vapor, can be calculated from the ratio [3]:

$$\tau_s = \frac{\rho_f\, d}{\eta_2} \tag{5.12}$$

In the limit of very low vapor density, this approach yields identical results with the formulas introduced heuristically by other authors [13], according to which the vapor impedance is in series with the parallel of the film impedance and the interfacial viscosity. As an example, for Kr vapor pressures below 1 Torr, the two approaches yield essentially the same results, while for pressures above 10 Torr this discrepancy can be as high as 15 % or more, depending on the amount of sliding observed.

5.3 QCM Driving Circuits

As we have already mentioned, in order to use the quartz crystal as a nanotribology sensor it is necessary to measure its resonance frequency and its quality factor very accurately. Close to resonance, an AT-plate can in fact be accurately described [10] by the equivalent circuit shown in the enlargement of Fig. 5.1. The capacitor C_0 (of the order of a few pF) represents the static capacitance of the crystal between the two electrodes. The values of L_q and C_q are related to the kinetic and potential energies of the plate. The resistor R_q (of the order of a few tens of Ω) accounts for all the losses of the crystal and thus determines the intrinsic Q_q of the crystal. A quartz plate exhibits a series resonance at $f_s = 1/2\pi\sqrt{L_q C_q}$ and a parallel resonance at f_p where $f_p – f_s \approx f_s C_q / 2C_0$. The detailed values of these characteristic parameters are provided by the manufacturer.

Various techniques have been devised to measure f_{res}. The simplest one is that of the oscillator, which uses the quartz crystal (either at its series or parallel resonance) in a positive feedback network in order to obtain an oscillating circuit. Many different

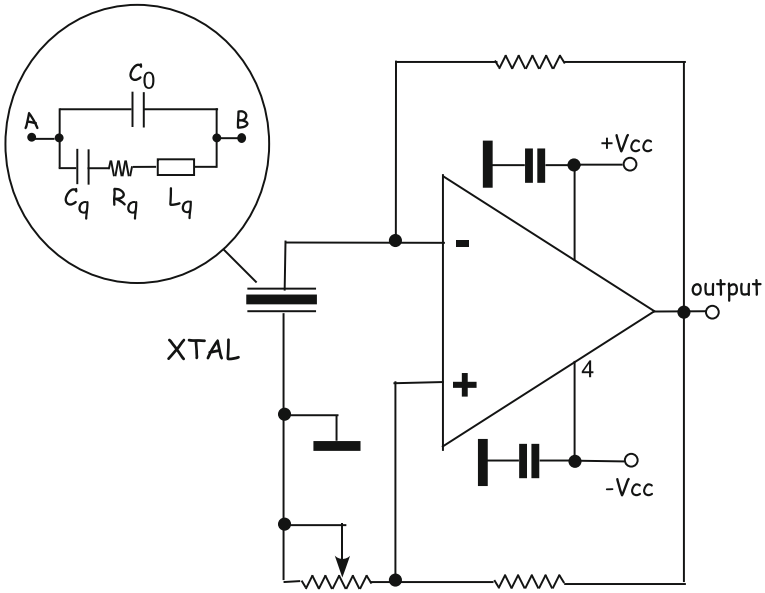


Fig. 5.1 Schematic circuit of a quartz oscillator used in our laboratory

configurations of oscillators exist in the literature. For example, Fig. 5.1 shows a very schematic diagram of an oscillator we have built in which the quartz crystal is connected to the inverting channel of a wide-band operational amplifier. By adjusting the resistance of the trimmer close to R_q, the circuit will auto-oscillate to the series resonance frequency of the quartz.

In practice, because of the unavoidable phase shifts introduced by the electronic components employed, f_{osc} is shifted with respect to f_{res} and this difference depends on the width of the resonance curve, if the overall phase shift does not change. For example, a typical phase error of 10° implies an error in the frequency $\left|\frac{f_{osc}-f_{res}}{f_{res}}\right| \approx 2 \times 10^{-6}$ with a quality factor of 40,000. If Q decreases during the measurements by 30 %, (a value observed in our studies with multilayers and which may become much bigger if one studies systems like heavy, organic fluids characterized by a large damping), it determines a change in the frequency of the oscillator of about 6×10^{-7} parts, e.g. an error of about 3 Hz for a crystal of 5 MHz. Another drawback of the oscillator is that it does not allow an easy variation and monitoring of the excitation power of the crystal. Furthermore, the oscillator selects by itself the resonance mode, typically it is the fundamental series, and it is not easy to switch to other modes. In conclusion, such a technique, although it is the simplest one, is not very flexible and might cause serious measurement errors particularly in situations where the Q of the quartz microbalance is small.

Another simple way to measure both the frequency shift and the dissipation taking place in a QCM experiment is the ring-down QCM, ring-down technique method

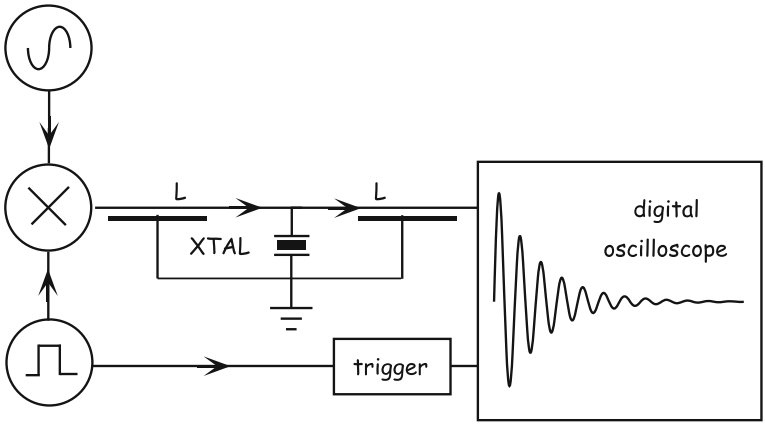


Fig. 5.2 Block diagram of the ring-down technique

shown in Fig. 5.2. The quartz crystal is excited at its resonance frequency either by an high-stability radio frequency generator [14] or by an oscillator [15], which are connected to the electrodes through an electronic switch. This latter one is controlled by a pulse generator that also triggers a digital oscilloscope. In this way, it is possible to turn the excitation on and off at a fixed rate, typically comprised between 10 and 100 Hz, and store the free decaying voltage after the excitation has stopped in the memory of the oscilloscope. The data can then be analyzed and fit according to the function $\mathrm{Aexp}(-\mathrm{t}/\tau)\cos(2\pi\,\mathrm{f}\,\mathrm{t}+\phi)$, where f represents the resonance frequency while the decay time τ is related to the quality factor of the crystal through the simple relation $\mathrm{Q}=\pi\,\mathrm{f}\,\tau$.

In our view, the technique best suited to drive the QCM in nanotribology applications is certainly that of the QCM, frequency modulation technique frequency modulation. In such a technique, $\mathrm{f_{osc}}$ corresponds to the maximum (or to the minimum) of the amplitude of the quartz electrical impedance while the quality factor is deduced from the amplitude of the detected signal [16]. Its main advantages can be summarized as: (i) it is possible to lock on any resonance mode of the QCM (series or parallel, fundamental or overtone) in a very simple and fast way; (ii) the excitation power can be easily varied from a few nW to several μW and, more important for nanotribology studies, its precise value can be accurately determined from the analysis of the crystal circuitry [17]; (iii) it is possible to achieve very high sensitivities and time stabilities; (iv) most of its main components can be easily found in any laboratory.

Figure 5.3 shows the block diagram of the electronics used in the FM technique. The output of an high stability radio-frequency generator equipped with the external frequency-modulation option (FM) drives the quartz crystal (XTAL) with a frequency $\mathrm{f}=\mathrm{f_{car}}+\Delta\mathrm{f}\sin(2\pi\,\mathrm{f_{mod}}\mathrm{t})$, where $\mathrm{f_{car}}$ is the so-called frequency carrier, set by the operator sufficiently close, but not necessarily equal, to the resonance frequency of the quartz mode one wants to lock-on. This $\mathrm{f_{car}}$ is modulated at a low-frequency $\mathrm{f_{mod}}$

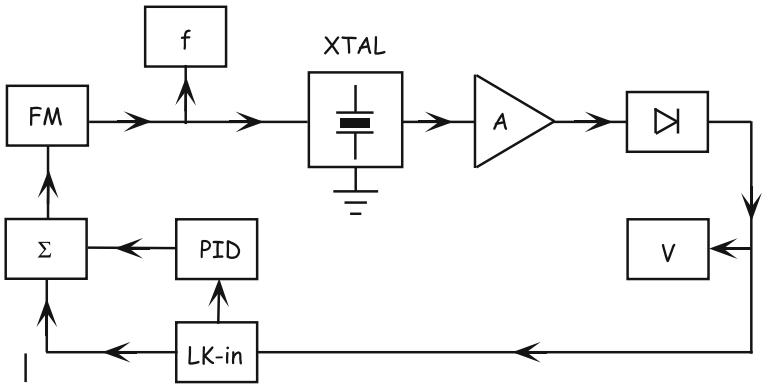


Fig. 5.3 Block diagram of the frequency modulation technique. See text for further details

between the extremes $f_{car} - \Delta f$ and $f_{car} + \Delta f$. The frequency of this modulation has to be smaller than the inverse decay time of the crystal f_{res}/Q_q [16]. The quartz crystal may either be inserted in a transmission line or have one electrode grounded, depending on the experimental set-up. As a matter of course, the actual configuration affects the choice of the amplifier A. If $\Delta f \ll f_{car}$, the amplified voltage can be well approximated by $A(f) \sim A(f_{car}) + \frac{dA(f_{car})}{df} \Delta f \sin(2\pi f_{mod} t)$. The high-frequency component of this signal can then be rectified by a diode detector. In our case, to bypass the problems connected with the use of the diode, e.g. periodic calibrations and thermal and time drifts, we have realized a multiplier whose output yields the square of the signal followed by a low-pass filter [17]. The DC amplitude of this signal is read by a high-precision voltmeter. The small component at f_{mod} is instead detected by a lock-in amplifier (Lk-in). The DC output of Lk-in, which changes sign as the frequency passes through quartz resonance, is used to control the value of f_{car} and thus locks the circuit onto the resonance frequency of the quartz. The block Σ adds up the low frequency modulation signal and the output of the Lk-in. The PID sums up the Lk-in output, its integral and its derivative. The integral is needed to perfectly locate the resonance frequency, the derivative guarantees stability to the feedback loop.

Finally, we mention another simple and sensitive circuit to drive a QCM. A generator is locked to the series resonance of the crystal by an high-frequency lock-in amplifier, which detects the reactive component of the transmission signal of the crystal and shifts the synthesizer frequency to null this component, after an offset is applied to cancel the contributions of shunt capacitance [18].

5.4 Quality of the Surface Electrodes

For a he quantitative use of the QCM in nanotribology, it is of paramount importance to have very homogeneous and well characterized surfaces. In fact, surface roughness not only complicates the comparison with theoretical models but can also prevent

the sliding of the adsorbed film. Although quartz resonators consist of thin disks of single crystalline silicon dioxide which are optically polished on both sides to roughness of about 1 nm rms, the thermally deposited metal electrodes usually exhibit a roughness of a few nm rms. Unfortunately, this roughness cannot be significantly reduced. To anneal a thermally deposited gold electrode, heating above 600 °C would be required. However, at 573 °C a phase transition occurs from α- to β-quartz. This phase transition is reversible upon cooling, but may introduce regions where the direction of the electrical axis in the quartz is reversed and thus remove or greatly reduce the piezoelectricity.

In a systematic study [19] of the surface morphology of Cu films on quartz in a UHV chamber, it was observed that annealing to 340 °C of the films deposited onto quartz preplated with titanium reduces the rms roughness to about 3 nm over a scan size of 500×500 nm^2, although they maintain a self-affine fractal scaling behavior over the length scale $10 \sim 500$ nm.

More recently, Pb films have been grown by physical deposition using an e-beam heated evaporation source at a rate of 0.5 nm/s [20]. The substrate was a quartz blank polished down to an RMS roughness of about 0.3 nm. Prior to Pb evaporation the quartz substrate was annealed under UHV conditions up to 140 °C in order to remove condensed surface impurities. When Pb deposition is performed at or above room temperature, the thermally activated diffusion of Pb atoms is so high that a non connected percolated network of Pb clusters is formed [20]. Therefore, a substrate temperature of 150 K was chosen to hinder adatom mobility and a connected film is formed. In Fig. 5.4 we can see a large scale STM image (2×2 μm) showing the morphology of a 150 nm thick Pb film deposited at 150 K, followed by annealing at

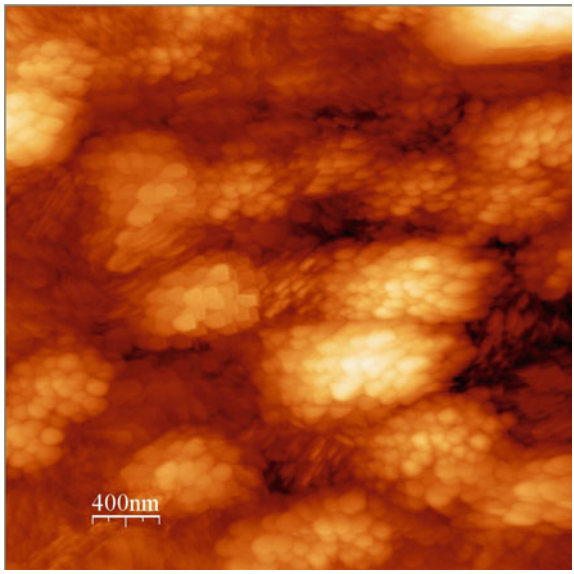


Fig. 5.4 STM topography of a $2{,}000 \times 2{,}000$ nm ^{2} area representative of a 150 nm thick Pb film deposited at 150 K on the surface of an AT-cut quartz. The film is formed by the assembly of 0.5 μm sized domains which are formed by a stacking of platelets. (*Photo* courtesy of F. Buatier de Mongeot)

room temperature. The image shows a distribution of domains with lateral dimensions around 0.5 μm . The various domains can be identified by the different orientation of the platelets. The majority of them are stacked parallel to the quartz surface, with an in-plane rotational mismatch, while a minor fraction is stacked with a tilt angle with respect to the substrate.

Another method to obtain very homogenous surfaces consists in glueing very thin layers of mica to the metal electrodes of a QCM [21]. In this way, it is possible to have a quartz sensor with a macroscopic atomically flat area without a significant loss of its sensitivity. These mica covered crystals have been used in studies of contact mechanics [22]. So far, however, no QCM measurements on the sliding friction of monolayers adsorbed on mica have been reported. Following a similar approach, a piece of Grafoil, a type of graphite, was glued uniformly on the Ag electrodes of a quartz crystal in order to study the slippage of ^{4}He films below 3 K [23].

5.5 UHV Apparatus

Another problem that affects QCM measurements of the sliding friction of adsorbed monolayers is the contamination of the active surface of the quartz sensor [8]. To overcome these problems, we have recently assembled a new apparatus specifically designed to perform friction experiments on molecularly thin films carried out with the QCM technique in ultra-high-vacuum and at a temperature as low as 4 K [24]. The main chamber is provided with a stainless steel jacket that allows a quick change of the various temperature inserts (cryocooler head, liquid nitrogen insert, Peltier-cell stage…) that span the working range 4–400 K, without ever breaking the vacuum. The crystal mounting is compatible with UHV conditions and with the Omicron standard and guarantees a good thermal and electrical contact over the entire temperature range. The quartz is inserted in a specially designed copper sample holder, housing the wiring for the temperature control and signal detection systems in a volume separated by that of the chamber. In this way, we can control the QCM temperature within 5 mK or less at any temperature below 10 K. Furthermore, the quartz frequency and amplitude stabilities are found to be as good as the best recorded values achieved with QCMs mounted in standard high-vacuum, liquid bath cryostats. The system is provided with a sputtering ion gun mounted in the fast-entry load section to clean the QCM electrodes from surface contaminants. The movements of the quartz sample inside the vacuum volume of the system are performed through the combination of a magnetic translator and a wobble stick.

Facing the circular hole in the crystal mounting, there is a thin-wall stainless steel tube whose end is attached to a sapphire variable leak valve. The high pressure side of the valve is connected to a high-purity gas cylinder and another port allows to purge the system effectively. A film is condensed onto the QCM, kept at low temperature, by slowly leaking gas through this nozzle. Depending onto the vapor pressure, the film may cover either only one or both electrodes. In the former case, which typically occurs at very low temperatures, the deposition of a Ne film can be controlled within

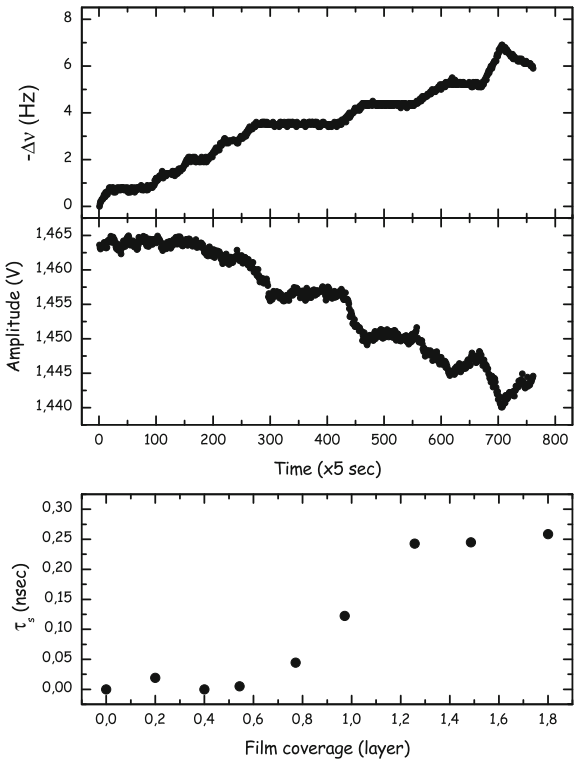


Fig. 5.5 Raw data of the resonance frequency shift (*top*) and amplitude (*middle*) during an adsorption isotherm of Ne on lead at 6.5 K. *Bottom* calculated slip time as a function of Ne film coverage

0.1 layers working at a fundamental frequency of 5 MHz. Obviously, this figure improves significantly using an heavier adsorbate and/or an higher overtone. For instance, the dosing of a Kr film onto a QCM running at 15 MHz (and third overtone) is done with a resolution better than 0.05 layers. If necessary, the adsobed film can be annealed to higher temperatures by simply turning off the cryocooler.

In the following, we present some preliminary data acquired with this setup at temperatures below 10 K. They refer to Ne deposited on a Pb(111) electrode grown and characterized in the group of Prof. Valbusa following the procedure described in the previous section. The data have been acquired at the third overtone of a 5 MHz quartz plate characterized by a quality factor of 380,000 at low temperatures. The two top graphs in Fig. 5.5 show the variations in the measured QCM resonance parameters during Ne evaporation. By acting on the leak valve, the film was slowly grown in steps of about 0.2 layers and then we waited for the system to equilibrate. At low coverages, there is no change in the quartz amplitude and accordingly the slip time is zero. Close to 0.4–0.5 layers, dissipation starts to appear and the slip time reaches values close to 0.3 ns, which are typical for rare gases sliding on metal surfaces. The slip times described in the bottom of Fig. 5.5 have been normalized with coverage according to the formulas reported in [8].

Although there is no data available in the literature on the 2D phase diagram of Ne adsorbed on Pb(111), it is tempting to interpret our data in terms of a structural depinning of the film. At low coverages, the Ne film is in a fluid phase that at such low temperatures is locked to the substrate. Close to about 0.4 layers, the film enters an incommensurate solid phase which is weakly bound to the substrate and can easily slides. This interpretation is consistent with the structural phase diagrams of heavy rare gases adsorbed on Ag(111) [25], systems that are very similar to Ne/Pb(111). Our measurements are also in very good qualitative agreement with the results of extensive computer simulations of a model system carried out by Persson [26] in the case of a low-corrugated substrate.

A systematic study of the dependence of the slip time on film coverage, driving amplitude and temperature is currently under way in order to establish the phase diagram of this intrinsic depinning [27]. In a previous QCM study of the sliding friction of Kr monolayers adsorbed on Au(111) at 85 K we have also observed a nonlinear behavior, but in that case the depinning was induced by increasing the driving amplitude above a certain threshold [7]. More recently, highly sensitive AFM experiments have found changes in the friction behavior that are based on similar structural effects. Dienwiebel et al. observed that friction between graphite surfaces is significantly reduced when the surfaces are rotated out of the commensurate locking angle [28]. Socoliuc et al. by varying the normal load on the contact between tip and substrate, have also observed a new regime of very low friction in which negative and positive lateral forces sum up to a vanishing average force in the time average instead of the spatial average [29].

It is a pleasure to acknowledge several clarifying explanations of the theoretical aspects of nanofriction we have had over the past few years with Francesco Ancilotto, Bo Persson and Erio Tosatti. We also wish to thank our experimental partners Francesco Buatier de Mongeot, Renato Buzio, Bruno Torre, Corrado Boragno and Ugo Valbusa for many interesting discussions and suggestions. We have greatly benefitted from daily interactions with our students who have been involved with these studies: Alessandro Carlin, Moira Ferrari, Luca Stringher, Francesco dalla Longa, Giovanni Fois and Alberto Pontarollo, and we must also thank Giorgio Delfitto for his technical mastery. Finally, funding from INFM, PRA *Nanorub*, and MIUR, FIRB *Carbon based micro and nanostructures* and PRIN *Nanotribologia,* is kindly acknowledged.

References

1. B.N.J. Persson, *Sliding Friction* (Springer, Berlin, 1998), Chap. 1
2. J. Krim, Sci. Am. **275**, 74 (1996)
3. J. Krim, A. Widom, Phys. Rev. B **38**, 12184 (1988)
4. J. Krim, D.H. Solina, R. Chiarello, Phys. Rev. Lett. **66**, 181 (1991)
5. C. Daly, J. Krim, Phys. Rev. Lett. **76**, 803 (1996)
6. R.L. Renner, P. Taborek, J.E. Rutledge, Phys. Rev. B **63**, 233405 (2001)
7. L. Bruschi, A. Carlin, G. Mistura, Phys. Rev. Lett. **88**, 046105 (2002)

8. A. Carlin, L. Bruschi, G. Mistura, Phys. Rev. B **68**, 045420 (2003)
9. G. Hayderer, M. Schmid, P. Varga, H.P. Winter, F. Aumayr, Rev. Sci. Instrum. **70**, 3696 (1999)
10. J.C. Brice, Rev. Mod. Phys. **57**, 105 (1985)
11. L. Bruschi, G. Mistura, Phys. Rev. B **63**, 235411 (2001)
12. C. Mak, J. Krim, Phys. Rev. B **58**, 5157 (1998)
13. E.T. Watts, J. Krim, A. Widom, Phys. Rev. B **41**, 3466 (1990)
14. S. Berg, T. Prellberg, D. Johannsmann, Rev. Sci. Instrum. **74**, 118 (2003)
15. M. Rodahl, F. Höök, A. Krozer, P. Brzezinski, B. Kasemo, Rev. Sci. Instrum. **66**, 3924 (1995)
16. M.J. Lea, P. Fozooni, P.W. Retz, J. Low Temp. Phys. **54**, 303 (1984)
17. L. Bruschi, G. Delfitto, G. Mistura, Rev. Sci. Instrum. **70**, 153 (1999)
18. G.B. Hess, M.J. Sabatini, M.H.W. Chan, Phys. Rev. Lett. **78**, 1739 (1997)
19. S.M. Lee, J. Krim, Thin Solid Films **489**, 325 (2005)
20. B. Torre, F. Buatier de Mongeot, F. Krok, R. Alessio, C. Boragno, U. Valbusa, to be published
21. S. Berg, M. Ruths, D. Johannsmann, Rev. Sci. Instrum. **74**, 3845 (2003)
22. S. Berg, D. Johannsmann, Phys. Rev. Lett. **91**, 145505 (2003)
23. N. Hosomi, A. Tanabe, M. Hieda, M. Suzuki, J. Low Temp. Phys. **138**, 361 (2005)
24. L. Bruschi, A. Carlin, F. Buatier de Mongeot, F. dalla Longa, L. Stringher, G. Mistura, Rev. Sci. Instrum. **76**, 023904 (2005)
25. J. Unguris, L.W. Bruch, M.B. Webb, J.M. Phillips, Surf. Sci. **114**, 219 (1982)
26. B. Persson, Phys. Rev. B **48**, 18140 (1993)
27. L. Bruschi et al., to be published
28. M. Dienwiebel, G. Verhoeven, N. Pradeep, J. Frenken, J. Heimberg, H. Zandbergen, Phys. Rev. Lett. **92**, 126101 (2004)
29. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)

Part II
Atomic-Scale Friction

Chapter 6
Atomic-Scale Friction Measurements in Ultra-High Vacuum

Sabine Maier, Enrico Gnecco and Ernst Meyer

Abstract This chapter reviews friction force microscopy experiments that reveal atomic-scale processes in single asperity contacts during sliding. Different regimes of atomic friction will be discussed including smooth sliding with low dissipation and the dissipative *atomic stick–slip* motion, where the tip jumps from one lattice site to the next. Furthermore, effects owing to finite temperatures, varying scan velocity and the influence of surface structures on the atomic friction are presented. Finally, the empirical Prandtl-Tomlinson model is introduced, which explains well the main observations in atomic-scale friction force microscopy experiments.

6.1 Introduction

Apart from playing a vital role in everyday life, friction between two sliding surfaces is a phenomenon of key importance to a variety of advanced technological applications and fundamental scientific studies. Understanding friction at the atomic scale is essential in order to find ways to minimize wear and energy losses in micro- and nano-scale devices involving sliding contacts. Sliding friction without wear, as it is discussed in this chapter, corresponds to the energy transfer from macroscopic degrees of freedom arising from the relative motion of the bodies in contact to microscopic degrees of freedom such as phonons or electronic excitations. The adaptation of the atomic force microscope to study the friction in single asperity contacts led to a

S. Maier (✉)
Department of Physics, University of Erlangen-Nürnberg, 91058 Erlangen, Germany
e-mail: sabine.maier@physik.uni-erlangen.de

E. Gnecco
IMDEA Nanociencia, 28049 Madrid, Spain
e-mail: enrico.gnecco@imdea.org

E. Meyer
University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland
e-mail: ernst.meyer@unibas.ch

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_6

profound knowledge of the underlying microscopic mechanisms on various surfaces and conditions.

The first atomic-scale friction measurements were performed by Mate et al. using a tungsten wire tip sliding on the basal plane of a graphite surface [1]. The frictional force on the tip, that was measured by optical interferometry, showed a saw-tooth-like modulation as a function of support position (or time). This modulation with the periodicity of the graphite lattice is due to elastic instabilities accompanying the sliding motion of the tip, causing it to hop between neighboring lattice sites. This characteristic motion is referred to as *atomic-scale stick-slip* and has been observed on a wide variety of materials. If the lateral force signal is measured in forward and backward directions a hysteresis loop occurs (so-called friction loop). The area embraced by the two curves gives a direct measure of the amount of energy dissipated while scanning over the surface. The friction force is defined as the average value of the lateral force signal. An example of a friction map and the corresponding friction loop obtained on a KBr(100) surface in ultra-high vacuum (UHV) is shown in Fig. 6.1. The lateral force in this measurement varies in the nano-newton range. In contrast, the first measurement on graphite revealed lateral forces in the micro-newton range [1]. These comparatively large forces for an atomic-scale process possibly occurred owing to a graphite flake which might have detached from the surface at high loads thus providing a multi-atom contact.

In this chapter we will describe different aspects of the atomic stick-slip motion on flat surfaces and on long-range ordered superstructures induced by surface reconstructions. In Sect. 6.2 we introduce the Prandtl-Tomlinson model, which explains the main features observed in atomic-scale friction experiments. In the rest of the chapter some significant experiments are described in detail to discuss effects arising at different loads and finite temperature.

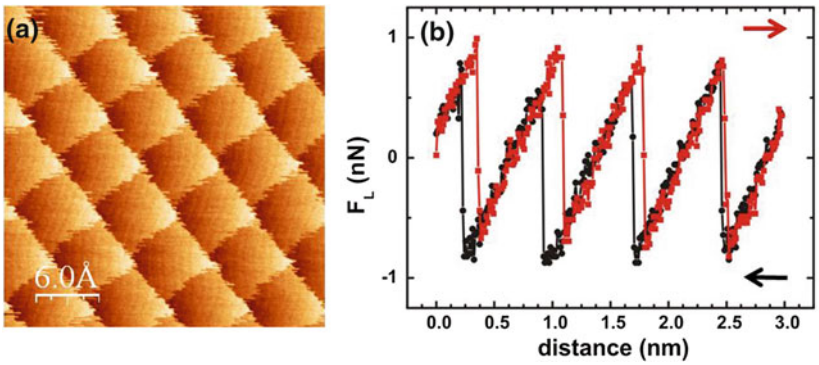


Fig. 6.1 **a** Friction force map and **b** the corresponding friction loop along the (001)-direction on a KBr(100) surface. The periodic instabilities represent the atomic-scale stick-slip motion of the tip, where it hops from one atomic site to the next. Adapted from [2]

6.2 The Prandtl-Tomlinson Model

Friction experiments on atomically flat terraces are well interpreted by the Prandtl-Tomlinson (PT) model [3, 4], proposed in the early 1920s. In spite of its simplicity, the PT model explains the main features in atomic friction experiments such as the periodic modulation of the lateral force. The PT model describes a point mass (the tip apex) being dragged along the surface by a rigid support via a spring k. Neglecting inertia, the total energy of the system is expressed by the sum of U_{eff}, that reflects the periodic tip–surface interaction, and the elastic energy stored in the spring:

$$U_{tot}(\mathbf{r}, t) = U_{eff}(\mathbf{r}) + \frac{1}{2}k(\mathbf{r} - \mathbf{v}t)^2. \tag{6.1}$$

In (6.1) the two-dimensional vector $\mathbf{r} \equiv (x, y)$ is the position of the tip and $\boldsymbol{X} = \mathbf{v}t$ the position of the support, with $\mathbf{v}$ the scan velocity.

6.2.1 One-dimensional Prandtl-Tomlinson Model

In one dimension, the total energy stored in the system simplifies to

$$U_{tot}(x, t) = -U_0 \cos\frac{2\pi x}{a} + \frac{1}{2}k(x - vt)^2, \tag{6.2}$$

assuming a sinusoidal interaction potential with the atomic lattice periodicity a and an amplitude U_0. $U_{tot}(x, t)$ has the form of a corrugated parabola, whose shape depends on the tip and support position, see Fig. 6.2a. The tip is located initially in a minimum of U_{tot}, described by the solution of $\partial U_{tot}/\partial x = 0$ and is subject to the stability condition $\partial^2 U_{tot}/\partial x^2 > 0$. When the support starts moving over the surface, the spring extends which results in a growing lateral force and a flattening of the local minima, see Fig. 6.2a. The tip becomes unstable and experiences an elastic instability at the position x^*. It initiates a jump towards the next stable position. With the energy profile given by (6.2), the instability condition translates into

$$x^* = \frac{a}{2\pi}\arccos\left(-\frac{1}{\eta}\right) \quad \text{with} \quad \eta = \frac{4\pi^2 U_0}{ka^2}. \tag{6.3}$$

The dimensionless parameter η reflects the ratio between the strength of the surface potential and the spring energy. It is evident from (6.3), that instabilities of the tip only occur if $\eta > 1$, which resembles the atomic stick-slip motion in the experiment. In this case, a hysteretic friction loop is observed while scanning forth and back along the same line. The area of the friction loop represents the energy dissipated in this process. If the tip oscillations are overdamped, hopping occurs from one equilibrium position to the next. In an underdamped system, where the PT model

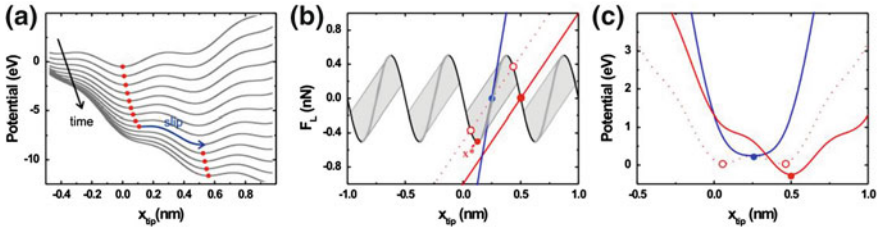


Fig. 6.2 **a** Temporal sequence of U_{tot} and the corresponding position of the tip (*red dot*) capturing a stick-slip instability ($\eta = 3.2$). **b** Graphical solution of $\partial U_{tot}/\partial x = 0$ to illustrate the instabilities: The *straight lines* represent the spring term and the *curved lines* the periodic surface force. The *blue line* has a single intersection with the surface force for each value of x_{tip}, and thus a smooth sliding from one minimum to the next is possible ($\eta < 1$). The *red line* has multiple intersections in the *grey shaded* areas indicating instable regions causing a hysteresis loops between forward and backward scans ($\eta > 1$). **c** The corresponding U_{tot} for the different regimes in (**b**)

is not applicable, multiple slips are possible, as will be discussed in Sect. 6.3. For $\eta < 1$ the instabilities and dissipation are suppressed and the tip smoothly follows the surface potential in an oscillatory way.

In the PT model, the lateral force $F_L = k(x_{tip} - vt)$ is represented by

$$F_L = -\frac{2\pi U_0}{a} \sin \frac{2\pi x_{tip}}{a}, \tag{6.4}$$

using the aforementioned condition for the tip being in a potential minima $\partial U_{tot}/\partial x = 0$. The maximal absolute value of the lateral force $F_L^{\max}$ in the stick-slip motion is found at $x_{tip} = a/4$ and we obtain

$$U_0 = \frac{aF_L^{\max}}{2\pi}. \tag{6.5}$$

Hence the corrugation of the surface potential is linearly related to the maximum lateral force and can be experimentally determined from the measured maxima of the lateral force in friction loops.

A second important parameter to consider in the framework of the PT model is the stiffness. The effective spring constant k reflects the respective compliance of tip and surface and can be modeled by a series of springs [5]:

$$\frac{1}{k} = \frac{1}{k_{contact}} + \frac{1}{k_{tip}} + \frac{1}{c_L}, \tag{6.6}$$

where c_L is the lateral spring constant of the cantilever and $k_{contact}$ and k_{tip} the lateral stiffness of the contact and tip, respectively. There are different methods to experimentally determine the lateral spring constant of the cantilever c_L, e.g. the added mass method of Cleveland [6, 7], the unloaded resonance technique of Sader [7, 8] or from the geometry and the elastic modulus [9]. Lantz et al. found by

scanning transmission electron microscopy and finite element analysis that the lateral stiffness of high aspect ratio Si tips is comparable with the lateral spring constant of the cantilever, i.e. $k_{tip} \approx 85$ N/m [5]. Finally, the lateral contact stiffness $k_{contact}$ for a macroscopic contact between a sphere and a flat surface is $k_{contact} = 8aG^*$ using continuum mechanics, [10] where a is the contact radius and the effective shear modulus G^* is defined by

$$\frac{1}{G^*} = \frac{2-\nu_1^2}{G_1} + \frac{2-\nu_2^2}{G_2} \tag{6.7}$$

($G_{1,2}$ are the shear moduli of the sphere and the surface, respectively, and $\nu_{1,2}$ the corresponding Poisson numbers). According to the Hertzian model the contact radius and hence the contact stiffness follows a power-law in dependence of the applied load. In atomic stick-slip measurements on NaCl, Socoliuc et al. [11] showed however, that the measured effective stiffness k and the extracted $k_{contact}$ are nearly constant, see Fig. 6.3. Instead, the corrugation of the surface potential, $E_0 = 2U_0$, increases comparatively much stronger with load than the effective stiffness. Experimentally, the effective spring constant is determined by the sticking slope k_{exp} at the beginning of the friction loop (F_x vs. x curve) provided η is large [12]. For η approaching 1, the relation has to be corrected to [13]

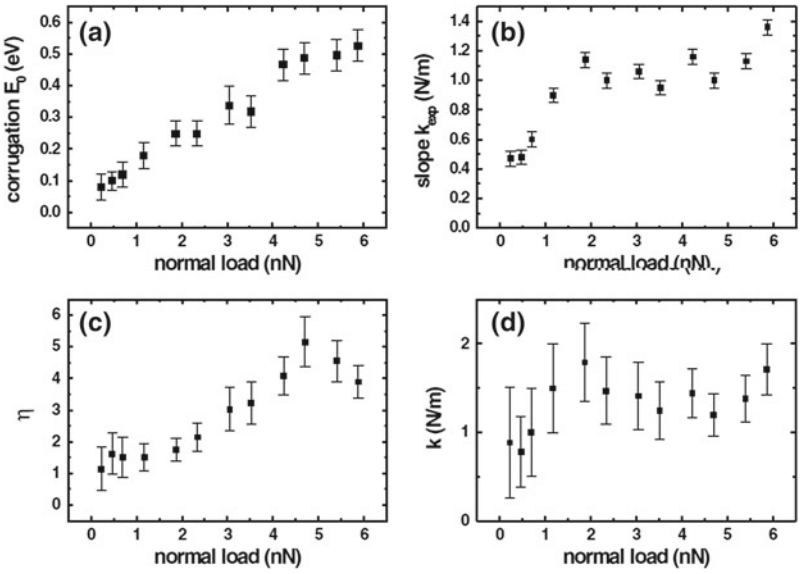


Fig. 6.3 **a** Energy corrugation E_0, **b** slope k_{exp} of the lateral force versus distance curve in the sticking part, **c** parameter η, and **d** effective lateral stiffness k of the contact, as a function of the normal load F_N acting on the tip. The quantities were determined from atomic stick-slip experiments on NaCl(001). From [11]

$$k = \left(1 + \frac{1}{\eta}\right) k_{\text{exp}}, \tag{6.8}$$

which results from a combination of (6.3) and (6.5). The breakdown of the continuum models at the nanometer-scale has been discussed with respect to the contact stiffness by Luan and Robbins, who found that the applicability of continuum mechanics depends strongly on the atomic structure of the tip apex [14]. Tips which have a spherical shape even on the nanometer-scale are well described by continuum models, while tips with a stepped atomic structure exhibit discontinuities in the contact area with increasing load.

6.2.2 Extensions of the Prandtl-Tomlinson Model

There are several extensions of the one-dimensional PT model to account for the complexity of certain experimentally observed phenomena. These specific extensions of the PT model will be discussed to some extent along this chapter in conjunction with the corresponding experimental validations.

Generally applicable is the extension of the PT model to two dimensions to reproduce the basic features of an experimental friction force map, whereas the one-dimensional model is limited to friction loops only [15]. The interaction potential $U_{eff}(x, y)$ in (6.1) depends on the properties of tip and sample and is generally unknown. However, in many cases it is sufficient to consider only the periodicity and symmetry of the sample surface, neglecting the fine structure of the contact within the unit cell. To represent the square lattice of NaCl(001) [16] and the hexagonal structure of graphite(0001) [17], the following effective potentials can be used:

$$U_{square}(x, y) = -U_0 \cos\frac{2\pi x}{a} \cos\frac{2\pi y}{a} \text{ and} \tag{6.9}$$

$$U_{hex}(x, y) = -\frac{U_0}{2.25}\left[2\cos\frac{2\pi x}{a}\cos\frac{2\pi y}{a\sqrt{3}} + \cos\frac{4\pi y}{a\sqrt{3}}\right]. \tag{6.10}$$

Both interaction potentials reproduce the experimentally observed features in friction maps well. The two-dimensional model allows to locate the tip position on the surface and to determine how the tip moves in a two-dimensional 'zig-zag' motion across the surface lattice [18].

Although the conventional one-dimensional PT model depicts the basic mechanism of the stick-slip motion, it cannot explain effects such as the variation of the maximal lateral force in a friction loop [19] or the increase of the mean friction force with the logarithm of the scan velocity v [20, 21]. Including thermal activation processes in the PT model accounts for both the velocity and temperature dependence of friction. The thermally activated PT model will be discussed in detail in Sect. 6.3. The dynamics of the tip on the surface at finite temperatures can be described more

detailed by numeric simulations based on the Langevin equation [22]

$$m\frac{d^2x}{dt^2} + m\gamma\frac{dx}{dt} + \frac{dU_{tot}}{dx} = \xi(t), \tag{6.11}$$

where m *is* an effective mass and γ is a damping coefficient describing the coupling with phonon and possible electron oscillations in the substrate. The quantity $\xi(t)$ is a random noise term satisfying the fluctuation-dissipation relation $\langle\xi(t)\xi(t')\rangle = 2m\gamma k_B T\delta(t-t')$, in which k_B is the Boltzmann's constant.

Further, the PT model should be extended to a two-mass-two-spring model, if both the resonance frequency of the cantilever and the experimentally found lateral stiffness, which is dominated by the contact, are well reproduced [2, 23]. In the two-mass-two-spring model, one spring represents the cantilever and the other spring the microscopic contact.

Taking into account that several atoms are usually in contact with the sample, requires the usage of a Frenkel-Kontorova (FK) or Frenkel-Kontorova-Tomlinson (FKT) model. This model addresses the periodic sliding of two surfaces in contact, whereas the atoms of one surface are harmonically coupled with their nearest neighbors [24, 25]. The FK model was first proposed for the description of dislocation in solids, and then used to address effects related to the commensurability and incommensurability of contacts as well as adsorbed films on surfaces [26]. In one dimension one can think of a chain of particles of mass m connected by elastic springs and interacting with a periodic potential. In this case the total potential takes the form

$$U_{tot} = \sum_n \left(-U_0 \cos\frac{2\pi x_n}{a} + \frac{1}{2}k(x_{n+1} - x_n - b)^2\right),$$

where b is the distance between two particles in the chain. Static friction can be probed by adding an external force F_{ext} adiabatically increasing till sliding initiates. The ratio *a/b* is very important. For any (irrational) value of *a/b* there is a critical value η_c of the parameter η defined by (6.3) such that the static friction force vanishes when $\eta < \eta_c$. The parameter η_c takes a maximum value (=1) when the ratio a/b is equal to the golden mean $(1+\sqrt{5})/2 = 1.618$ [27].

6.3 Experimental Observations of Atomic Stick-slip

Since the invention of friction force microscopy (FFM) by Mate et al. [1], atomic-scale stick-slip has been experimentally observed on a variety of materials including insulators, metals, and semiconductors. Here, we will focus on FFM measurements in ultra-high vacuum. Other environments, especially liquid cells, are described elsewhere in this book. Ionic crystals are commonly used in atomic-scale friction studies in UHV to study the characteristics of the stick-slip motion owing to their simple

structure and comparatively easy and reproducible preparation. The periodicity of the stick-slip motion and instabilities, respectively, corresponds on ionic crystals to the unit-cell given that the contact is underdamped. Atomic stick-slip was observed for example on KBr [28], NaCl [11, 20], and KF [29]. Measurements on NaF(100) by Ishikawa et al. are the exception, as they could resolve both components depending on the applied load [30].

Friction force experiments on clean metal surfaces prepared in UHV are often more difficult to carry out due to the high surface energy of metals under clean conditions, which leads to a high reactivity of surface atoms resulting in a junction formation under the tip apex [31]. For example, the formation of a neck of copper atoms at the tip is proposed to result in the observation of stick-slip type sliding over Cu(111) surfaces [32]. Friction force microscopy experiments at low loads performed on Cu(100) and Au(111) surfaces revealed a clear stick-slip modulation in the lateral force with almost zero dissipation [33]. For higher loads, significant friction occurred together with the onset of wear. The Cu(100) surface is more easy worn than Cu(111), because it is not a crystallographic glide plane [34]. Even on the more ductile Pt(111) surface atomic stick-slip was observed using passivated tips [35].

Recently, two-dimensional materials have been of wide interest due to their extraordinary electronic, mechanical, optical, and chemical properties with respect to their bulk counterparts, making them suitable for applications in electronic and NEMS devices [36]. The most widely studied two-dimensional material is graphene, whose nano- and microtribology properties are outlined in detail elsewhere in this book. Using FFM, the nano-scale frictional response of atomically thin sheets of supported graphene, molybdenum disulfide (MoS_2), niobium diselenide ($NbSe_2$), and hexagonal boron nitride have been investigated [37, 38]. Both graphene and MoS_2 exhibited clear atomic stick-slip friction for exfoliated films down to a single atomic sheet. The friction monotonically increased as the number of layers decreased for all four materials [38].

6.3.1 Load Dependence: From Smooth Sliding and Stick-slip to Wear

Classically, dry friction is governed by the Coulomb (or Amontons) equation $F_L = \mu F_N$, where the friction coefficient μ mainly depends on the material of the two bodies in motion. In single asperity contacts the dependence of the friction on applied load is usually non-linear. In nano-scale contacts, the load dependence is often described by means of contact mechanics continuum theories [39] which however do not give information about the atomic interactions in the sliding contact region as they assume a spherical tip. Fusco et al. suggested a power law dependence of atomic friction by simulating the dynamics of a tip scanning a rigid monolayer graphite surface in a PT-like model [40].

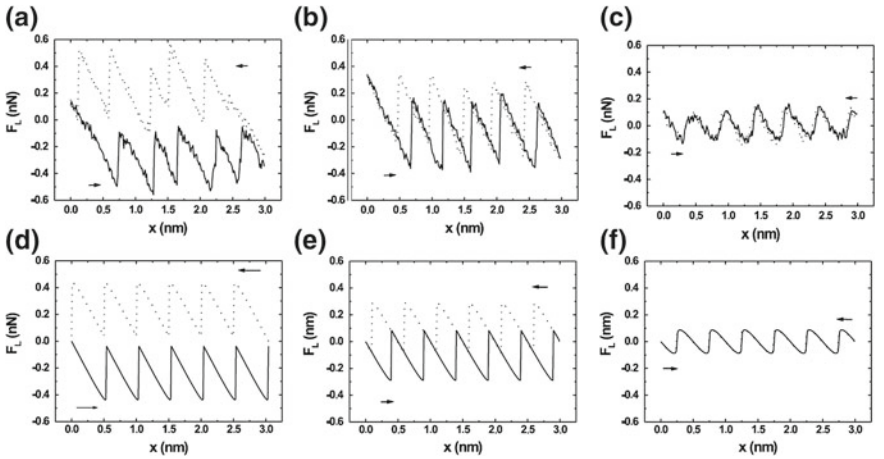


Fig. 6.4 Measurement of the lateral force sliding the tip forward and backward in (100) direction over a NaCl(001) surface, where the externally applied load was **a** $F_N = 4.7\,nN$, **b** $F_N = 3.3\,nN$, and **c** $F_N = -0.47\,nN$. **d–f** Corresponding numerical results from the Tomlinson model for **d** $\eta = 5$, **e** $\eta = 3$, and **f** $\eta = 1$. The stiffness was chosen as $k = 1\,\mathrm{N/m}$ and the lattice constant as $a = 0.5\,\mathrm{nm}$. From [11]

Experimentally, the amount of energy dissipated while sliding over an atomically flat surface can be controlled by the variation of load on the contact [11]. The mean friction force decreases with decreasing normal load and reaches zero before the probing tip jumps out of contact. The normal force in the experiment is represented in the PT model by the parameter η. If $\eta < 1$ elastic instabilities are suppressed and the tip smoothly follows the surface corrugation with negligible friction. This transition from stick-slip to smooth sliding was observed experimentally for friction loops recorded on NaCl(001) by Socoliuc et al. when a normal loading of about 1 nN was applied, see Fig. 6.4.

There are other ways to achieve a state of ultra-low friction at the atomic scale. Dienwiebel et al. [41] observed vanishing friction while dragging a graphite flake out of registry over a graphite surface. The corrugation of the interaction potential between the graphite substrate and the graphite flake could be continuously varied by simply rotating the substrate with respect to the flake. This state of ultra-low friction is only observed in incommensurable contacts. It is referred to as *superlubricity*, which goes back to a concept discussed by Hirano [26]. Friction can also be decreased by reducing the scan speed to a few nm/s or less, so that thermally activated jumps occur [42]. The thermally induced suppression of friction is called *thermolubricity*. Another efficient way to switch friction on and off at the atomic scale is found by exciting the mechanical resonances of the sliding system perpendicular to the contact plane [43]. In this way, a similar transition from stick-slip to smooth sliding as in Fig. 6.4 was observed on NaCl, however by changing the excitation amplitude instead of the load. The state of *dynamic superlubricity* so-achieved could be also exploited to acquire lattice-resolved lateral force maps of crystal surfaces with no occurrence of abrasive wear [44]. No evidence of wear was also proven by SEM images of a silicon

tip sliding on a polymer surface and vibrating at the contact resonance over several hundred meters [45]. A significant decrease of friction is also observed if a nanotip is shaken laterally instead than normally [46]. In this case, an accurate analytical relation between the shaking amplitude and the parameters of the PT model can be derived [47].

In contrast, wear may occur in a sliding contact, if the normal forces applied exceed a critical point. The initial stage of abrasive wear on alkali halides was investigated by comparing friction loops and topography images before and after wearing off a KBr(001) surface [48]. From this experiment it was concluded that on KBr only a minor part (30 %) of the total energy dissipation during scratching the surface went into wear. The formation of wavy patterns ('ripples') with periodicities in the order of 100 nm is also an interesting effect accompanying abrasive wear processes on the nanoscale [49].

6.3.2 The Slip

The slip instability is accompanied by the dissipation of mechanical energy into phonon and electron excitations. In conventional FFM experiments the slip occurs so fast that the relaxation time is not measurable. However, recording the atomic-scale stick-slip movement with higher bandwidth revealed a wide variation of slip durations up to several milliseconds, by far longer than expected for a relaxation process on the atomic scale [2]. These long slip events are possibly resulting from a multiple contact established between tip and surface, as suggested by the correlation between irregular features in lateral force maps and long slip durations. This conclusion was supported also by a comparison of the experimental results with a multi-tip (multi-spring) simulation using a PT model including thermal activation.

6.3.3 Thermal Effects and Velocity Dependence

Coulomb's law of friction states that kinetic friction is independent of the sliding velocity for macroscopic contacts. In contrary, most experiments report a logarithmic dependence of the mean friction force in atomic-scale contacts at low velocities, e.g. mica [21], Cu(111) [32], NaCl [20], graphite [50], and silicon oxide [51].

The origin of the logarithmic velocity dependence can be understood with a thermally activated PT model based on reaction rate theory. At zero temperature the tip does not jump to the next equilibrium position until an energy barrier $\Delta E(t)$, that depends on the interaction potential, becomes zero. However, at finite temperatures T, the tip can hop even if $\Delta E(t) \neq 0$. The probability that the tip remains pinned at a given location, $p(t)$, changes with time according to the master equation [20]

$$\frac{dp(t)}{dt} = -f_0 \exp\left(-\frac{\Delta E(t)}{k_B T}\right) p(t), \tag{6.12}$$

In (6.12) f_0 denotes the resonance frequency of the tip in its actual potential minimum. Assuming a linear dependence of the energy barrier ΔE with increasing lateral force F, a logarithmic dependence of the friction is obtained [20]

$$F(v) = \text{const.} + \frac{k_B T}{\lambda} \log \frac{v}{v_0}, \tag{6.13}$$

where λ is in the order of the lattice constant. Equation 6.13 suggests that the slower the tip is scanning the higher the probability of an early thermally activated jump and hence a lower friction force results. Sang et al. [22] refined the model above and proposed that barrier-hopping fluctuations occur preferentially when the tip is very close to athermal slipping at the top of the barrier. In this case, primarily at high velocities $\Delta E \propto (\text{const.} - F_L)^{3/2}$ has to be used. The master equation leads then to

$$F(v) = \text{const.} - \left(\frac{k_B T}{\lambda}\right)^{2/3} \left(\log \frac{v_c}{v}\right)^{2/3} \tag{6.14}$$

for velocities below the critical velocity

$$v_c = \frac{\pi\sqrt{2}}{2} \frac{f_0 k_B T}{ka}. \tag{6.15}$$

At high velocities the thermal vibrations have no time to contribute, and a transition from the logarithmic increase in friction to a constant plateau occurs. The transition to a constant plateau has been observed in FFM experiments on mica in ambient conditions [21]. More recently, Reimann et al. [52] proposed a model with special emphasis on the realistic description of dissipation and inertia effects. The model predicts a non-monotonic dependence of the friction force upon the sliding velocity of the tip over an atomically flat surface.

Thermal effects do not only play an important role in the velocity dependence of nano-scale friction. Krylov and Frenken [42] found that the thermal fluctuations of the localized contact can also result in a variety of friction regimes not predicted by the athermal PT model. In particular, the range of ultra-low friction at low loads observed by Socoliuc et al. on NaCl could be extended toward higher loads by thermal effects. Their calculations suggest also that thermal delocalization of the contact could lead to extremely low friction independent of normal load, despite the observation of atomic stick-slip in the lateral force signal. As mentioned above, this concept of low friction owing to thermal fluctuation is referred to as thermolubricity.

With the recent instrumental development of FFMs being operational with high resolution at cryogenic temperatures, the predictions of the thermally activated PT model were experimentally verified. Atomic stick-slip measurements between 100 and 300 K have been performed on graphite [50] and MoS_2 [53]. Both experiments revealed an increase in average friction with decreasing temperature in agreement with the PT model.

6.3.4 Maximal Lateral Force

Another interesting feature of the stick-slip motion is the maximum lateral force, as it is a direct measure of the surface-tip potential strength (see 6.5) and hence varies much stronger with load than the stiffness of the contact [11]. The values of the maximal lateral force are spread by thermal activation. Schirmeisen et al. extracted the maximum lateral force from friction force maps on HOPG and performed a statistical analysis of them [22]. A site-dependent energy barrier of the surface potential from the side to the center of the hollow-site rows was found from fits to the experimental distribution functions using a thermally activated PT model. A velocity dependent statistical analysis of the jump heights on HOPG by Evstigneev et al. [54] showed that the single-step reaction scenario assumed by Schirmeisen et al. is valid only at relatively high velocities, while at slower pulling speeds, a more complicated hopping mechanism must be at work. As for the slip duration analysis, the latter mechanism might be associated to multiple bond formation of the tip-substrate contact, where additional bonds take finite time to establish, so that at fast pulling, new bonds do not have sufficient time to develop during a single stick phase.

6.3.5 Multiple Slips

Most of the early atomic-slip measurements revealed jump periodicities of one lattice spacing. Johnson and Woodhouse have predicted that under certain conditions, slips are not restricted to a single lattice constant and hopping over multiple spacings may occur (i.e. multiple slips) [55]. The transition between single and double slip mode and its dependence on contact damping, sliding velocity, and finite temperature was studied in numeric simulations employing a generalized PT model [56, 57]. Medyanik et al. derived the threshold values of the parameter η (see 6.3) for the transition between different slip regimes from a simple analytic criterion based on the one-dimensional PT model in the quasi-static limit [58]. They found that the tip slides smoothly if $\eta < 1$, executes stick-slip motion with single slips for $1 < \eta < 4.6$ or, possibly exhibits double slips for $4.6 < \eta < 7.79$ or even bigger multiple slips for $\eta > 7.79$. However, the threshold values are significantly modified by the damping coefficient γ and the temperature. A chaotic regime (with alternate jump lengths) is always expected if γ is low enough [59].

Experimentally, multiple slips can occur only with sufficiently low energy dissipation during the slip. Therefore the transition from single to multiple slips occurs at high loads. Figure 6.5 shows the experimental dependence of stick–slip behavior with load on an HOPG surface in ambient condition. The system exhibited superlubricity at the lowest applied load whereas at higher loads, stick–slip instabilities occurred with periodicity of the HOPG lattice and multiples of it. Roth et al. [60] verified the occurrence of multiple slips also in UHV and used multiple slips to quantify the viscous damping accompanying the stick-slip motion on NaCl(001). They found from

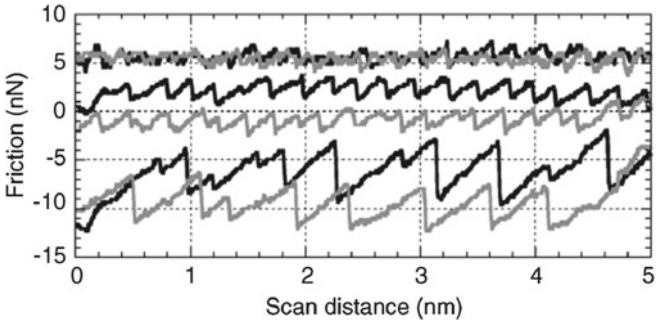


Fig. 6.5 Representative experimental friction loops with increasing load (offset for clarity) demonstrating the transitions from smooth sliding (*top*) to single (*middle*) and mostly double slips (*bottom*). From [58]

a comparison between simulations based on an extended PT model and experiments that nanometer-sized contacts are underdamped at intermediate loads (13–26 nN) and become slightly overdamped at higher loads.

6.4 Atomic-Scale Friction Beyond Flat Terraces

In this section we will change from atomically flat to corrugated substrates and discuss the influence of surface step edges and of periodic height variations induced by superstructures and surface reconstructions on the atomic-scale friction.

6.4.1 Atomic-Scale Friction at Step Edges

A common observation in several friction experiments on the micrometer-scale is the enhancement of friction at step edges related to an increased energy barrier at the step, known as Schwöbel barrier [61–63]. The enhanced frictional forces at step edges are strongly direction dependent. While the friction increases linearly with the applied load in upward scans, it is load-independent for downward scans. This direction dependence of the friction has been experimentally observed on HOPG, MoS_2, and NaCl in ambient and UHV environment for monatomic up to about tenfold steps [61, 62].

Recently, Steiner et al. [64] performed atomic-scale studies on NaCl(001) and Ge(001) revealing the atomic stick-slip motion along a step. They found that with blunt tips the lateral force was increased while scanning both up and down an edge, consistent with the experiments at the micro-scale. However, with atomically sharp tips, the lateral force still increased upwards but decreased and even changed sign

downwards leading to overlapping lateral force profiles for both directions. Hence, the sharpness of the tip facilitates scanning the step edge without occurrence of a Schwöbel barrier at small enough loads. The experimental results could be matched to calculated friction loops based on a modified PT model using a second contribution to the tip-surface potential to account for the step edge, which was modeled by an asymmetric well simulating a sharp drop followed by a slow recovery.

6.4.2 Atomic-Scale Friction on Ordered Superstructures and Reconstructions

Atomic-scale friction measurements featuring stick-slip have recently been performed on a variety of ordered superstructures, including ultra-thin KBr films on NaCl(001) [65] and Cu(001) [66] as well as single and bilayer graphene on SiC(0001) [37], all exhibiting interesting modulation effects. KBr thin film on NaCl(100) have a periodic (6×6)-superstructure due to the lattice mismatch between the two materials [67]. This long-range modulation of the surface-potentials has strong effects on the atomic stick-slip motion. The lateral force signal shows a variation of the local amplitude of the trace-retrace friction loops along with a clear transition from a dissipative stick-slip to smooth sliding regime in concert with the underlying superstructure (see Fig. 6.6) [65]. This transition was attributed to a variation of the potential-energy corrugation, which is also supported by numerical simulations using a one-dimensional PT model with a spatially modulated tip-sample interaction. Specifically, the constant surface potential amplitude U_0 in the conventional PT model (6.2) is replaced with

$$U_0 \left(1 + \alpha \cos \frac{2\pi x}{b}\right), \tag{6.16}$$

where α defines the normalized modulation amplitude and b the periodicity of the superstructure. In contrast for graphene, the local center position of the trace-retrace friction loops varies in concert with the superstructure [37]. The centerline modulation does not affect the width of the friction loop, and thus there is no change in the energy dissipated during frictional sliding. The origin of a centerline modulation is not entirely clear, as it can result from a geometrical effect, e.g. a local slope of the surface [68], or can be the result of a friction asymmetry. For graphene, the experimental results are better reproduced using a superposition of the potentials $U = U_{eff} + \beta U_{\text{sup}}$ where U_{eff} describes the interaction with the atomic lattice, as in (6.1), and U_{sup} the interaction with the superlattice [69].

Friction force microscopy in a low load regime has been also successfully applied to image surface reconstructions on metals, e.g. the Au(111) herringbone [33, 70], as well as on single-crystal semiconductors [71]. In order to resolve these surface modulations, the contact size in those experiments was clearly below the characteristic length of the superstructure and reconstruction, respectively (i.e. a few nm).

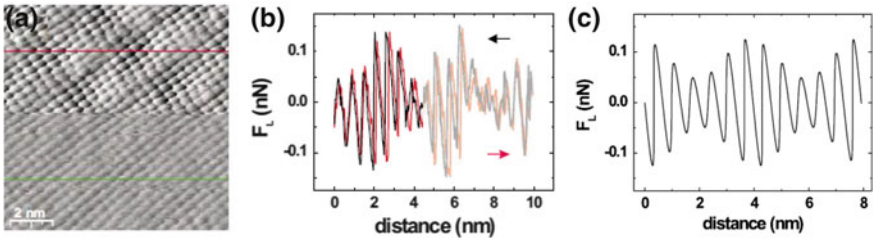


Fig. 6.6 **a** Lateral force image of a KBr double layer on NaCl(100) with **b** the corresponding forward and backward traces along the red line in (**a**). The average normal force was reduced from −0.01 to −0.32 nN in the lower half of the measurement. **c** Simulated lateral force trace assuming a modulated energy corrugation (see 6.16) with a period $b = 6a$ and a constant stiffness of $k = 1\,\mathrm{N/m}$ in a one-dimensional PT model. From [65]

6.5 Anisotropy Effects

There are various mechanisms in friction force microscopy experiments which lead to an asymmetry in frictional forces. For instance anisotropy effects are related to the transition of commensurate and incommensurate contacts, respectively, or to a structural anisotropy of the sample surface.

As discussed already in Sect. 6.3, the importance of the relative orientation of two lattices in sliding contact has been first experimentally demonstrated by Hirano et al. [26]. The friction increased when the two surfaces formed a commensurate contact and superlubricity was obtained in the incommensurate case. This anisotropy was also observed in stick-slip friction, when measuring the lateral forces between a graphite flake and the HOPG surface at rotation angles between 0° and 60° [41].

There are several examples of friction anisotropy owing to a structural anisotropy of the surface. For instance using a UHV-FFM, Park et al. [72] reported a large directional anisotropy of the friction force on Al–Ni–Co decagonal quasicrystals. They found friction forces to be eight times larger when sliding along the periodic direction of the surface than when sliding along the aperiodic one. The anisotropy in the friction force was attributed to the intrinsic structural anisotropy of the surface. On the quasicrystal surface the friction anisotropy is manifested in both nanometer-sized contacts obtained with sharp AFM tips and macroscopic contacts in pin-on-disk tribometers [73].

Friction anisotropy was also observed in molecular single crystals, although not in UHV. For instance the kinetic friction was found to peak up along well-defined crystallographic directions on a pentacene single crystal [74] and organic crystal potassium hydrogen phthalate (KAP) [75]. In these experiments, the frictional anisotropy was measured by recording the transverse component of the friction forces acting on the AFM tip while scanning. Friction anisotropy is also often found between individual domains of molecules, which are chemically homogenous but structurally rotated. For instance, in an early FFM study on freshly cleaved (010) surfaces of the ferroelectric triglycine sulfate frictional anisotropy could be observed under ambient

condition between terraces inside individual domains, which are structurally rotated by 180° [76]. Another example is the flower-shaped islands of lipid monolayers on mica, consisting of domains with different molecular orientations. In these experiments, the angular dependence of friction directly reflected the tilt direction of the alkyl chains in the monolayer [77].

An asymmetry can also occur in the lateral force of atomically and molecularly resolved friction force maps revealing stick-slip characteristics. For instance lateral force maps on layered organic compounds such as transition-metal oxalate complexes showed a clear contrast between molecules with different orientation revealing a friction asymmetry on a molecular scale [78]. The static and kinetic friction experienced by a point mass elastically driven at different angles on surface lattices with square, hexagonal, and honeycomb symmetries were estimated by analytical and numeric calculations based on the PT model. The anisotropy of static friction was found to be strongly dependent on the density of the surface atoms packing [79]. These results await experimental verification, however.

6.6 Mechanical Properties of Molecular Chains

The FK model introduced in Sect. 6.2 has been recently applied to an interesting experiment, in which molecular chains lying on a flat surface were pulled up at one of its end by an AFM tip. Kawai et al. succeeded in lifting up single polymer chains from a reconstructed gold surface using AFM in a dynamic mode at low temperature (4 K) in UHV [80]. Under those conditions, it was possible to measure the force gradients accompanying the detachment of individual polymer chains with extremely high accuracy (cf. Fig. 6.7). The primary observation was the modulation of the force during detachment of fluorene groups, which could be precisely related to the adhesion energy of these submolecular groups by a theoretical analysis based on an extended FK model and realistic interaction potentials. A small modulation of the force gradient due to the sliding on the gold surface was also observed. This modulation indicates that the polymer chain was sliding in a superlubric way, as expected from the incommensurability between the lattice constants of the gold substrate and the equilibrium distance between consecutive fluorine units.

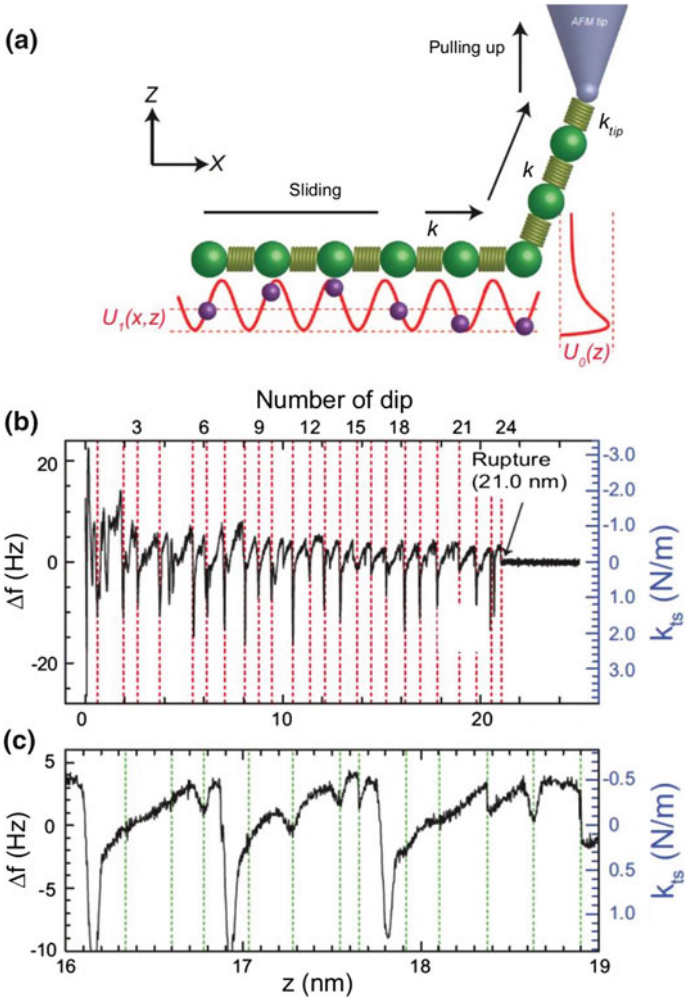


Fig. 6.7 **a** Sketch (*side view*) of a polymer chain consisting of polymerized polyfluorene initially lying on a Au(111) surface and pulled up by an AFM tip. **b** The force gradient shows periodic variations due to the detachment of the fluorene units. **c** Zoom on the from [80]

These measurements also showed that the intermolecular stiffness is large (around 200 N/m), which makes elastic deformation to accommodate the molecules on the substrate quite unfavorable.

6.7 Conclusions

In conclusion, we have shown that the PT model reproduces the main characteristics of the atomic-slip motion observed in friction force microscopy experiments quite well. The experimental observations include the load dependence, specifically the transition from smooth sliding to atomic stick-slip and multi-slips, as well as the velocity and temperature dependence of atomic-scale friction. Atomic stick-slip experiments on reconstructed surfaces revealed a strong influence of the long-ranged height modulations on the lateral force. Very recent AFM investigations on molecular chains lifted up by the probing tip open interesting possibilities in coupling nanotribology to the realm of molecular electronics. Remarkably, all these experiments were realized in ultra-high vacuum conditions, where environmental contaminants could be removed from the nanocontacts.

References

1. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987)
2. S. Maier, Y. Sang, T. Filleter, M. Grant, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. B **72**, 245418 (2005)
3. L. Prandtl, ZAMM—J. Appl. Math. Mech./Zeitschrift für Angewandte Mathematik und Mechanik **8**, 85 (1928)
4. G.A. Tomlinson, Philos. Mag. Ser. **7**(7), 905 (1929)
5. M.A. Lantz, S.J. O'Shea, A.C.F. Hoole, M.E. Welland, Appl. Phys. Lett. **70**, 970 (1997)
6. J.P. Cleveland, S. Manne, D. Bocek, P.K. Hansma, Rev. Sci. Instrum. **64**, 403 (1993)
7. C.P. Green, H. Lioe, J.P. Cleveland, R. Proksch, P. Mulvaney, J.E. Sader, Rev. Sci. Instrum. **75**, 1988 (2004)
8. J.E. Sader, I. Larson, P. Mulvaney, L.R. White, Rev. Sci. Instrum. **66**, 3789 (1995)
9. T.R. Albrecht, S. Akamine, T.E. Carver, C.F. Quate, J. Vac. Sci. Technol. A **8**, 3386 (1990)
10. K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985)
11. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)
12. R.W. Carpick, D.F. Ogletree, M. Salmeron, Appl. Phys. Lett. **70**, 1548 (1997)
13. E. Gnecco, R. Bennewitz, T. Gyalog, E. Meyer, J. Phys.-Condens. Matter **13**, R619 (2001)
14. B. Luan, M.O. Robbins, Nature **435**, 929 (2005)
15. T. Gyalog, M. Bammerlin, R. Lüthi, E. Meyer, H. Thomas, Europhys. Lett. (EPL) **31**, 269 (1995)
16. P. Steiner, R. Roth, E. Gnecco, A. Baratoff, S. Maier, T. Glatzel, E. Meyer, Phys. Rev. B **79**, 045414 (2009)
17. H. Hölscher, U.D. Schwarz, O. Zwörner, R. Wiesendanger, Phys. Rev. B **57**, 2477 (1998)
18. E. Gnecco, Europhys. Lett. **91**, 66008 (2010)
19. A. Schirmeisen, L. Jansen, H. Fuchs, Phys. Rev. B **71**, 245403 (2005)
20. E. Gnecco, R. Bennewitz, T. Gyalog, C. Loppacher, M. Bammerlin, E. Meyer, H.J. Guntherodt, Phys. Rev. Lett. **84**, 1172 (2000)
21. E. Riedo, E. Gnecco, R. Bennewitz, E. Meyer, H. Brune, Phys. Rev. Lett. **91**, 084502 (2003)
22. Y. Sang, M. Dubé, M. Grant, Phys. Rev. Lett. **87**, 174301 (2001)
23. D.G. Abel, S.Y. Krylov, J.W.M. Frenken, Phys. Rev. Lett. **99**, 166102 (2007)
24. T. Gyalog, H. Thomas, Europhys. Lett. **37**, 195 (1997)
25. M. Weiss, F.-J. Elmer, Phys. Rev. B **53**, 7539 (1996)
26. M. Hirano, Wear **254**, 932 (2003)

27. O.M. Braun, Y.S. Kivshar, *The Frenkel-Kontorova Model: Concepts, Methods, and Applications* (Springer-Verlag, Berlin, 2004)
28. R. Luthi et al., J. Vac. Sci. Technol. B **14**, 1280 (1996)
29. R.W. Carpick, Q. Dai, D.F. Ogletree, M. Salmeron, Tribol. Lett. **5**, 91 (1998)
30. M. Ishikawa, S. Okita, N. Minami, K. Miura, Surf. Sci. **445**, 488 (2000)
31. U. Landman, W.D. Luedtke, N.A. Burnham, R.J. Colton, Science **248**, 454 (1990)
32. R. Bennewitz, T. Gyalog, M. Guggisberg, M. Bammerlin, E. Meyer, H.J. Guntherodt, Phys. Rev. B **60**, R11301 (1999)
33. N. Gosvami, T. Filleter, P. Egberts, R. Bennewitz, Tribol. Lett. **39**, 19 (2010)
34. R. Bennewitz, E. Gnecco, T. Gyalog, E. Meyer, Tribol. Lett. **10**, 51 (2001)
35. M. Enachescu, R.W. Carpick, D.F. Ogletree, M. Salmeron, J. Appl. Phys. **95**, 7694 (2004)
36. A.K. Geim, Science **324**, 1530 (2009)
37. T. Filleter, R. Bennewitz, Phys. Rev. B **81**, 155412 (2010)
38. C. Lee, Q. Li, W. Kalb, X.-Z. Liu, H. Berger, R.W. Carpick, J. Hone, Science **328**, 76 (2010)
39. M. Enachescu, R.J.A. van den Oetelaar, R.W. Carpick, D.F. Ogletree, C.F.J. Flipse, M. Salmeron, Phys. Rev. Lett. **81**, 1877 (1998)
40. C. Fusco, A. Fasolino, Appl. Phys. Lett. **84**, 699 (2004)
41. M. Dienwiebel, G.S. Verhoeven, N. Pradeep, J.W.M. Frenken, J.A. Heimberg, H.W. Zandbergen, Phys. Rev. Lett. **92**, 126101 (2004)
42. S.Y. Krylov, K.B. Jinesh, H. Valk, M. Dienwiebel, J.W.M. Frenken, Phys. Rev. E **71**, 065101 (2005)
43. A. Socoliuc, E. Gnecco, S. Maier, O. Pfeiffer, A. Baratoff, R. Bennewitz, E. Meyer, Science **313**, 207 (2006)
44. E. Gnecco, A. Socoliuc, S. Maier, J. Gessler, T. Glatzel, A. Baratoff, E. Meyer, Nanotechnology **20**, 025501 (2009)
45. M.A. Lantz, D. Wiesmann, B. Gotsmann, Nat. Nano. **4**, 586 (2009)
46. R. Roth, O.Y. Fajardo, J.J. Mazo, E. Meyer, E. Gnecco, Appl. Phys. Lett. **104**, 083103 (2014)
47. O.Y. Fajardo, E. Gnecco, J.J. Mazo, Phys. Rev. B **89**, 075423 (2014)
48. E. Gnecco, R. Bennewitz, E. Meyer, Phys. Rev. Lett. **88**, 215501 (2002)
49. A. Socoliuc, E. Gnecco, R. Bennewitz, E. Meyer, Phys. Rev. B **68**, 115416 (2003)
50. L. Jansen, H. Hölscher, H. Fuchs, A. Schirmeisen, Phys. Rev. Lett. **104**, 256101 (2010)
51. A. Schirmeisen, L. Jansen, H. Holscher, H. Fuchs, Appl. Phys. Lett. **88**, 123108 (2006)
52. P. Reimann, M. Evstigneev, Phys. Rev. Lett. **93**, 230802 (2004)
53. X. Zhao, S.R. Phillpot, W.G. Sawyer, S.B. Sinnott, S.S. Perry, Phys. Rev. Lett. **102**, 186102 (2009)
54. M. Evstigneev, A. Schirmeisen, L. Jansen, H. Fuchs, P. Reimann, Phys. Rev. Lett. **97**, 240601 (2006)
55. K.L. Johnson, J. Woodhouse, Tribol. Lett. **5**, 155 (1998)
56. J. Nakamura, S. Wakunami, A. Natori, Phys. Rev. B **72**, 235415 (2005)
57. Z. Tshiprut, S. Zelner, M. Urbakh, Phys. Rev. Lett. **102**, 136102 (2009)
58. S.N. Medyanik, W.K. Liu, I.-H. Sung, R.W. Carpick, Phys. Rev. Lett. **97**, 136106 (2006)
59. E. Gnecco, R. Roth, A. Baratoff, Phys. Rev. B **86**, 035443 (2012)
60. R. Roth, T. Glatzel, P. Steiner, E. Gnecco, A. Baratoff, E. Meyer, Tribol. Lett. **39**, 63 (2010)
61. H. Hölscher, D. Ebeling, U.D. Schwarz, Phys. Rev. Lett. **101**, 246105 (2008)
62. T. Müller, M. Lohrmann, T. Kässer, O. Marti, J. Mlynek, G. Krausch, Phys. Rev. Lett. **79**, 5066 (1997)
63. E. Meyer, R. Luthi, L. Howald, M. Bammerlin, M. Guggisberg, H.J. Guntherodt, J. Vac. Sci. Technol. B **14**, 1285 (1996)
64. P. Steiner, E. Gnecco, F. Krok, J. Budzioch, L. Walczak, J. Konior, M. Szymonski, E. Meyer, Phys. Rev. Lett. **106**, 186104 (2011)
65. S. Maier, E. Gnecco, A. Baratoff, R. Bennewitz, E. Meyer, Phys. Rev. B **78**, 045432 (2008)
66. T. Filleter, W. Paul, R. Bennewitz, Phys. Rev. B **77**, 035430 (2008)
67. S. Maier, O. Pfeiffer, T. Glatzel, E. Meyer, T. Filleter, R. Bennewitz, Phys. Rev. B **75**, 195408 (2007)

68. D.F. Ogletree, R.W. Carpick, M. Salmeron, Rev. Sci. Instrum. **67**, 3298 (1996)
69. P. Steiner, E. Gnecco, T. Filleter, N. Gosvami, S. Maier, E. Meyer, R. Bennewitz, Tribol. Lett. **39**, 321 (2010)
70. Q. Li, Y. Dong, A. Martini, R. Carpick, Tribol. Lett. **43**, 369 (2011)
71. M. Goryl, J. Budzioch, F. Krok, M. Wojtaszek, M. Kolmer, L. Walczak, J. Konior, E. Gnecco, M. Szymonski, Phys. Rev. B **85**, 085308 (2012)
72. J.Y. Park, D.F. Ogletree, M. Salmeron, R.A. Ribeiro, P.C. Canfield, C.J. Jenks, P.A. Thiel, Science **309**, 1354 (2005)
73. J.Y. Park, D.F. Ogletree, M. Salmeron, C.J. Jenks, P.A. Thiel, J. Brenner, J.M. Dubois, J. Mater. Res. **23**, 1488 (2008)
74. V. Kalihari, G. Haugstad, C.D. Frisbie, Phys. Rev. Lett. **104**, 086102 (2010)
75. M. Campione, E. Fumagalli, Phys. Rev. Lett. **105**, 166103 (2010)
76. H. Bluhm, U.D. Schwarz, K.P. Meyer, Appl. Phys. A-Mater. Sci. Process. **61**, 525 (1995)
77. M. Liley et al., Science **280**, 273 (1998)
78. G. Fessler et al., Appl. Phys. Lett. **98**, 083119 (2011)
79. E. Gnecco, O. Fajardo, C. Pina, J. Mazo, Tribol. Lett. **48**, 33 (2012)
80. S. Kawai et al., Proc. Natl. Acad. Sci. **111**, 3968 (2014).

Chapter 7
Stochastic Modeling and Rate Theory of Atomic Friction

Mykhaylo Evstigneev, Juan J. Mazo and Peter Reimann

Abstract Atomic friction involves objects whose dynamics is strongly influenced by thermal fluctuations. In stochastic modeling, one focuses on a few relevant degrees of freedom, whereas the atomistic ones are taken into account by introducing dissipation and noise. We review applications of this approach to atomic friction, namely, the basic Prandtl-Thomlinson model, some of its multidimensional generalizations, and the rate approximation, which allows one to obtain analytical results not easily accessible by other methods.

7.1 Introduction

Macroscopic friction between solids is well known to be both of paramount practical importance and of notorious difficulty regarding its theoretical understanding [1–4]. While macroscopic friction involves interactions between numerous asperities of the two contacting surfaces, employing an atomic force microscope (AFM) offers a unique opportunity to probe the frictional forces between a single asperity—the tip of an AFM cantilever—and an atomically flat surface. Therefore the research direction of friction force microscopy (FFM) [5] had been initiated only a year after the invention of the AFM in 1986 [6] and is being intensively pursued since then (see the reviews [7–13] and references therein).

M. Evstigneev (✉) · P. Reimann
Faculty of Physics, University of Bielefeld, Bielefeld, Germany
e-mail: mevstigneev@mun.ca

P. Reimann
e-mail: reimann@Physik.Uni-Bielefeld.de

Juan J. Mazo
Departamento de Física de la Materia Condensada,
Instituto de Ciencia de Materiales de Aragón, CSIC-Universidad de Zaragoza,
50009 Zaragoza, Spain
e-mail: juanjo@unizar.es

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_7

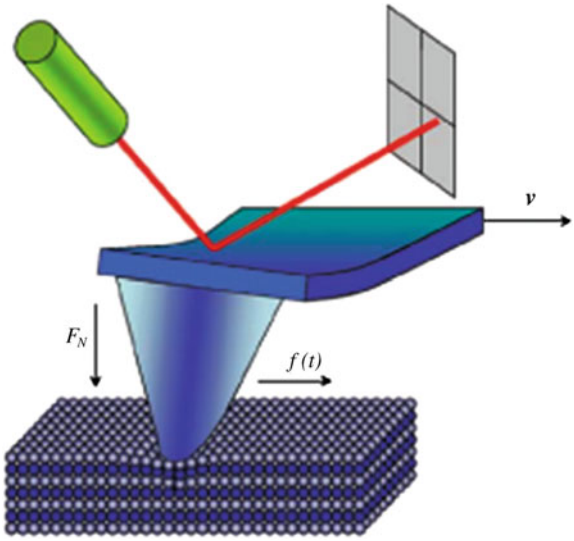


Fig. 7.1 Schematic illustration of an FFM experiment. Note that this picture is severely out of scale: in a real experiment, the tip radius is of the order of 10 nm, and the contact region consists of several hundreds of atoms

In a typical FFM experiment [5], the tip of an AFM cantilever is brought in contact with an atomically clean surface by means of a normal load F_N, while the cantilever base is set in motion at a constant velocity v (see Fig. 7.1). The interaction between the tip and the surface leads to a torsional deformation of the cantilever. One can determine the magnitude of this deformation by optical means and thus deduce the resulting elastic force $f(t)$, which, by Newton's third law, equals the instantaneous force of friction. The central quantity of interest is the time-averaged friction force

$$\bar{f} := \lim_{t\to\infty} \frac{1}{t} \int_0^t dt'\, f(t'). \tag{7.1}$$

While it has been known from the time of Coulomb that the force of friction between two macroscopic bodies in contact is independent of their relative velocity, friction force on the nanoscale exhibits a non-trivial velocity dependence, which will be discussed in this contribution.

Experimental results reveal that the effects of thermal noise play an important role in nanoscale friction. Understanding such a phenomenon is the most important challenge of the stochastic modeling of atomic friction. The specific indications of the importance of thermal effects are thermal fluctuations of the instantaneous friction force $f(t)$ and, in particular, randomness of interstitial jumps of the cantilever in the so-called stick-slip regime of motion, the temperature dependence of atomic friction, and the approximately logarithmic dependence of the friction force, which can be explained using a model based on the assumption that the tip transitions (slips) from one lattice site to the next are due to thermally activated rate processes [10, 14–18].

Though simpler than macroscopic friction, the adequate interpretation and modeling of nanofriction experiments still represents a formidable challenge. In particular, direct molecular dynamics (MD) simulations, initiated in the mid-nineties [19, 20], have only recently started to approach experimentally realistic pulling velocities. Without resorting to special techniques, the velocities accessible in MD studies are in the $1–10^3$ m/s range [21], orders of magnitude too fast in comparison with the experimentally relevant values. Smaller velocities can be probed with the help of methods that accelerate the algorithm performance, such as parallel replica dynamics [22, 23] (velocity of a few mm/s), or its combination with hyperdynamics [24, 25] that has further reduced the pulling velocities to the experimental microns per second. At the same time, the number of atoms that can be simultaneously accounted for in MD is still several orders of magnitude smaller than in the experiment. Last but not least, MD simulations may take up to several weeks of computational time. The reason for these limitations is the enormous time scale separation between atomic vibrations proceeding on the subpicosecond time scale, and sliding motion of the tip, which covers only a few lattice constants in a millisecond.

Hence, non-trivial theoretical modeling steps are indispensable, in particular the concepts of non-linear stochastic processes [26–31]; the above-mentioned time-scale separation justifies and greatly facilitates the calculations within such models. In stochastic modeling, one focuses on just a few relevant degrees of freedom, which, in the case of nanofriction, describe the tip geometric configuration. The huge number of the remaining atomistic degrees of freedom are accounted for by introducing the effects of randomness and dissipation in the tip equations of motion. In this contribution, we review two different types of stochastic approaches to nanofriction modeling, both stemming from the early works due to Prandtl [32] and Tomlinson [33]: one is based on the Langevin equation, and the other on the theory of thermally activated rate processes.

7.2 Langevin Modeling

7.2.1 Langevin Equation

The system from Fig. 7.1, though small, still involves a huge number of atomic degrees of freedom. In a one-dimensional model, the experimentally observable lateral force $f(t)$ can be deduced from the torsional deformation of the cantilever and is directly related to the displacement $x(t) - vt$ of the AFM tip from its equilibrium position vt at a moment of time t, cf. (7.5) below. To obtain the evolution equation for this relevant collective degree of freedom $x(t)$, one writes down the equation of motion for all coordinates of the system, and then projects the system's microstate onto the subspace characterized by a given value of $x(t)$ [26]. As a result of this procedure, an equation of motion for $x(t)$ is obtained, in which the effect of atomic degrees of freedom is accounted for by introduction of the following objects: (i) a

free-energy type potential $U(x,t)$ of mean force, (ii) memory-dependent dissipative force, and (iii) a random force (noise) of finite correlation time. In view of the fact that the characteristic frequency associated with the collective variable $x(t)$ is of the order of 10^5 Hz, i.e. many orders of magnitude lower than the Debye frequencies describing the time-scale of atomic motion, both the memory effects in dissipative force and the finite noise correlation time can be neglected.

The potential of mean force $U(x,t)$ consists of two contributions, the first one accounting for the elastic deformations of AFM and substrate, and the second for the tip-substrate interaction. Since the elastic deformations are typically small [34], we may neglect anharmonic effects in the elastic energy. Furthermore, we can assume that interaction only depends on the relative tip-substrate position x. We thus arrive at the approximation

$$U(x,t) = \frac{\kappa_L (x - vt)^2}{2} + U(x). \tag{7.2}$$

The argument in the first term indicates that the cantilever moves at a constant velocity $v > 0$ to the right (cf. Fig. 7.1). Furthermore, focusing on an ideally flat atomic surface with lattice constant a in x-direction, we conclude that $U(x)$ is invariant under a displacement by one period,

$$U(x + a) = U(x). \tag{7.3}$$

The lateral spring constant κ_L describes the combined effect of the elastic deformation of the cantilever, the tip and the elastically deformed surface in the contact region [34–39]:

$$\frac{1}{\kappa_L} = \frac{1}{\kappa_{cantilever}} + \frac{1}{\kappa_{tip}} + \frac{1}{\kappa_{surface}}. \tag{7.4}$$

The experimentally observable lateral force $f(t)$ can be identified, according to Newton's third law, with the negative of the force caused by the elastic deformations, i.e.

$$f(t) = -\kappa_L \left(x(t) - vt\right). \tag{7.5}$$

Next, let us consider the elastic deformations of the cantilever, and, in particular, those of the tip apex (see Fig. 7.1). If these deformations, or equivalently, the state variable $x(t)$, 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. If these changes are taking place at a finite speed, but still slowly enough that the thermal bath of the cantilever's atoms always remains close to the instantaneous accompanying thermal equilibrium, the remaining "small amount of disequilibrium" renders the process "slightly irreversible" and hence gives rise to a linear-response type dissipative force which, in the frame of reference of the cantilever, will be proportional to the velocity of the tip relative to the cantilever base. In the laboratory frame, this dissipative force assumes the form

$$F_c(t) = -\eta_c(\dot{x}(t) - v) \tag{7.6}$$

with an effective coupling strength $\eta_c > 0$ between the collective coordinate x and the close to equilibrium "cantilever and tip bath" (subscript "c"). In particular, because of the smallness of the tip deformations [34], the implicitly assumed independence of η_c on the state $x(t)$ of the system is well justified. In a similar manner, the influence of the microscopic degrees of freedom of the substrate will result in a dissipative force $F_s(t)$, which is proportional to the tip velocity with respect to the substrate with the proportionality coefficient η_s:

$$F_s(t) = -\eta_s \dot{x}(t). \tag{7.7}$$

Finally, we come to the randomly fluctuating forces acting on the slow state variable $x(t)$. They have the same origin as the dissipative forces, namely, the large number of fast degrees of freedom of the cantilever, tip and substrate baths. Due to this common origin and the fact that the baths always remain close to thermal equilibrium, one can show that those randomly fluctuating forces are completely fixed (in the statistical sense) by the functional form of the dissipative forces via the fluctuation-dissipation theorem [31, 40, 41]. Namely, the thermal "cantilever-and-tip-noise" acts on $x(t)$ in the usual form [26–30] of a fluctuating force $\sqrt{2\eta_c k_B T}\xi_c(t)$ with temperature T, Boltzmann constant k_B, and unbiased δ-correlated Gaussian noise $\xi_c(t)$ of unit strength. Similarly, the substrate gives rise to thermal fluctuations of the form $\sqrt{2\eta_s k_B T}\xi_s(t)$ with an unbiased δ-correlated Gaussian noise $\xi_s(t)$ independent of $\xi_c(t)$:

$$\langle \xi_c(t)\, \xi_c(t')\rangle = \langle \xi_s(t)\, \xi_s(t')\rangle = \delta(t'-t), \quad \langle \xi_c(t)\, \xi_s(t')\rangle = 0. \tag{7.8}$$

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 [41]. Their independence is an approximation which is well justified by the fact that the contact between the two baths consists of comparatively few atoms.

Collecting all acting forces, we arrive at the following equation of motion [42, 43]:

$$m\,\ddot{x}(t) + \eta\dot{x}(t) = -U'(x(t)) - \kappa_L(x(t) - vt) + \eta_c v + \sqrt{2\eta k_B T}\xi(t). \tag{7.9}$$

where m is the relevant effective mass associated with inertia effects of cantilever, tip, and substrate, and

$$\eta := \eta_s + \eta_c \tag{7.10}$$

is the total damping coefficient. The zero-noise limiting case of this equation of motion is essentially equivalent to the early model of friction due to Prandtl [32] and Tomlinson [33], whereas the thermal noise term was introduced [15, 16] about 70 years after Prandtl and Tomlinson's publications. Such an equation

(or its noise-free version) has been considered in a number of papers, the important difference being that either η_c or (much less often) η_s was assumed to vanish, whereas, in general, there is no reason to expect that any of these coefficients is zero. One can, however, straightforwardly relate the results obtained for arbitrary η_s, η_c to those where either of them is set to zero. For instance, one can introduce a time translation $\tilde{t} = t + \eta_c/\kappa_L$ in (7.9), which eliminates the $\eta_c v$ term from the equation and allows one to express $\bar{f}(\eta_s, \eta_c) = \bar{f}(\eta_s + \eta_c, 0) - \eta_c v$.

It is instructive to rewrite (7.9) in the co-moving reference frame defined by

$$z(t) = x(t) - vt, \quad \dot{z}(t) = \dot{x}(t) - v, \tag{7.11}$$

in which the equation of motion assumes the form

$$m\,\ddot{z}(t) + \eta\dot{z}(t) = -U'(z(t) + vt) - \kappa_L z(t) - \eta_s v + \sqrt{2\eta k_B T}\xi(t) \tag{7.12}$$

and the instantaneous friction force from (7.5) becomes

$$f(t) = -\kappa_L z(t). \tag{7.13}$$

Then, the term $\eta_s v$ can be eliminated by a change of variables $\tilde{z} = z - \eta_s v/\kappa_L$, $\tilde{t} = t + \eta_s/\kappa_L$, so that $\bar{f}(\eta_s, \eta_c) = \bar{f}(0, \eta_s + \eta_c) + \eta_s v$. One of the consequences of (7.12) is that at high velocities the third term in the right-hand side, which describes the viscous drag of the substrate, exceeds all the other acting forces, and the force of friction (7.13) behaves as

$$\bar{f} \to \eta_s v \quad \text{for } v \to \infty, \tag{7.14}$$

allowing us, at least in principle, to experimentally determine the coefficient η_s associated with the substrate from the slope of the force-velocity plot at high v.

7.2.2 Parameter Values

In the simplest version, the potential $U(x)$ is taken to be sinusoidal,

$$U(x) = -\frac{\Delta U}{2}\cos\frac{2\pi x}{a} \tag{7.15}$$

with a lattice constant $a \cong 0.2 \ldots 0.5$ nm and a corrugation depth ΔU, which varies from values close to zero to ca. 0.1 nN·nm, depending on the conditions of the experiment one wishes to model [44]. Note that, although a single-harmonic potential (7.15) is the one that is used most often in modeling, other possibilities have also been considered in the literature, e.g. a potential with sharp minima and flat maxima [45], quasiperiodic potentials [46, 47], potentials with localized Gaussian perturbations [48], fractal potentials [49], etc.

The value of the effective stiffness, κ_L, can be experimentally established from the slope of the force-distance curve in the regime where the cantilever is "stuck" to some lattice site of the substrate while its base moves at a constant velocity; then, the force evolves according to $f(t) \cong \kappa_L t$, up to an additive constant and small fluctuations [see (7.17), (7.18) below for a more precise estimate]. This procedure yields, for various experimental conditions, the value of κ_L between 0.1 N/m and 10 N/m [14, 17, 50].

With respect to the mass parameter, m, in the Langevin equation (7.9), its naive identification with the total mass of the cantilever is unjustified, which already becomes obvious from the fact that the effective stiffness κ_L is at least one order of magnitude smaller than the "bare" torsional stiffness of the cantilever. This fact suggests that only a very small portion of the tip at the apex experiences a significant deformation, while most of the cantilever is relatively rigid during its motion. The mass parameter usually employed in the simulations is of the order of $m \cong 10^{-12}$ kg: this estimate follows from the fundamental torsional resonance frequency of the cantilever in contact with the sample, $(2\pi)^{-1}\sqrt{\kappa_L/m}$, which has a typical value of ca. 400 kHz [51]. It has also been proposed [42, 43] that in many experiments, the relevant mass parameter is so small that an overdamped limit ($m \to 0$) is an adequate approximation.

Finally, the damping coefficient η is difficult to measure directly, because the damping force $-\eta v$ is typically very small. Various estimates [15, 42, 43, 52, 53], however, agree within an order of magnitude and yield $\eta \cong 10^{-6} \ldots 10^{-5}$ kg/s. Note that this value is close to the critical damping of the cantilever, $2\sqrt{\kappa_L m}$; thus, depending on the experimental conditions, the motion of the cantilever may be either slightly underdamped or overdamped.

7.2.3 Regimes of Motion

Depending on the relative importance of cantilever mass, damping, spring constant, pulling velocity, potential corrugation amplitude, and temperature, different dynamical regimes can be achieved: continuous sliding [44], thermolubricity [54], regular stick-slip motion [15, 16], stick-multislip [52, 55–57], chaotic motion [58, 59], etc. Here, we will briefly describe the regimes of continuous sliding and stick-(multi)slip, as they have attracted most of the experimental interest so far.

Whether the cantilever will slide or perform stick-slip motion depends on the ratio of the maximal potential curvature, $\max_x |U_0''(x)|$, to the stiffness κ_L, also known as the Prandtl-Tomlinson parameter [44, 56]. For a sinusoidal potential (7.15), it is given by

$$\gamma_{PT} = \frac{2\pi^2 \Delta U}{a^2 \kappa_L}. \tag{7.16}$$

For $\gamma_{PT} < 1$, the potential (7.2) has a single minimum located roughly at vt, implying smooth sliding of the tip. In this regime, the force fluctuates around the value $\eta_s v$, see

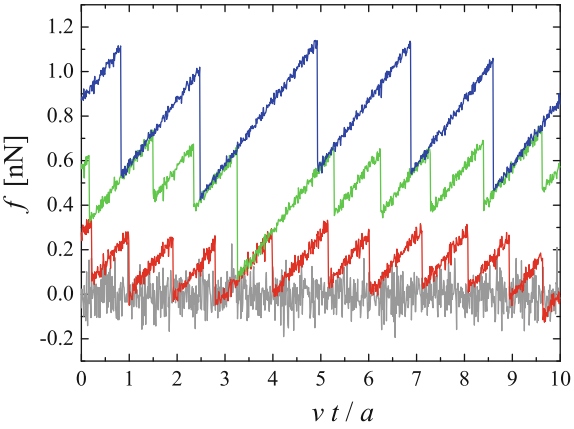


Fig. 7.2 Evolution of the friction force obtained from numerical simulation of the Langevin equation (7.9) with the following parameters: $a = 0.3$ nm, $\kappa_L = 1$ N/m, $m = 10^{-12}$ kg, $\eta_s = \eta_c = 10^{-6}$ kg/s, $v = 0.1\,\mu$m/s, $T = 300$ K. The *curves* are obtained for different values of the potential corrugation depth, $\Delta U = 2$, 50, 100, and 150 pN·nm (from *bottom* to *top*) and feature smooth sliding (*gray*), stick-single slip motion (*red*), a mixture of single and double slip events (*green*), and stick-double slip motion (*blue*)

Fig. 7.2, gray curve. On the other hand, if $1 < \gamma_{PT} < 4.604\ldots$, the potential (7.2) becomes bistable. Further increase of γ_{PT} to a value between ca. 4.604 and 7.788 ... results in a potential (7.2) having three minima; for $7.788 < \gamma_{PT} < 10.95\ldots$, the potential has four minima, etc. [56].

The multistability of the potential (7.2) implies the possibility of stick-slip motion, whose physical picture is as follows. In a stick phase, the tip apex is confined to the nth lattice site, while the cantilever base moves at a velocity v, leading to an approximately linear (up to thermal and instrumental noise) increase of the elastic force and a reduction of the energy barrier separating the tip from the next lattice site. At some point, thermal noise drives the tip over that energy barrier into the next, $(n + 1)$st lattice site in a single slip event (see Fig. 7.2, red curve), whereas the kinetic energy of the tip is dissipated into the atomistic degrees of freedom and the force drops by a fixed amount. Then, a new stick phase begins. If inertia of the tip is large, then the tip may not be able to dissipate all of its kinetic energy in a single slip event and, as a result, the tip will perform a jump over two or even more lattice constants [38]. In general, the multiplicity of slips for a given value of γ_{PT} is smaller than the number of minima of the total potential (7.2). The force evolution in the stick-multislip regime is exemplified in Fig. 7.2, showing a mixture of single and double slips (green curve) and pure double slips (blue curve).

In the stick phases, the force increases according to

$$f_n(t) = \kappa(vt - na), \tag{7.17}$$

with the rate of force increase characterized by a renormalized stiffness κ given by [44, 60]

$$\frac{1}{\kappa} = \frac{1}{\kappa_L} + \frac{1}{U''(b)}. \tag{7.18}$$

The parameter b can be taken as the position of the minimum of the corrugation potential (7.15), e.g., $b = 0$ [44]. A better accuracy is achieved if one identifies it with the position of the minimum of the total potential (7.2) corresponding to the mean force value $\bar{f}$ from (7.1), i.e. finds b from the relation $U'(b) = \bar{f}$ [50]. Calculations show that the renormalized stiffness κ weakly depends on velocity v and may be smaller than the "bare" counterpart κ_L by at most 10 % [50].

In the stick-slip regime, the friction force decreases with temperature due to a reduction of the cantilever force at the onset of the slip event. However, in general, the situation can be more complex and in some cases a temperature-induced enhancement of nanoscale friction was predicted [46, 48, 61] due to the effect of temperature on the slip length. Friction reduction is more significant at low velocities and moderate damping values. A change in the slip length is relevant for small enough spring constants (large γ_{PT}), where the total potential profile shows multiple accessible metastable states [57].

Multiple slip events have been observed experimentally [5, 52, 56]; it has also been suggested that the statistics of multislip events can be used to estimate the damping coefficient η in the Langevin equation [52]. On the other hand, the majority of experimental studies that we are aware of focus on the stick-single slip regime of motion; moreover, some publications [44, 62] state explicitly that only single-slip, but no multiple slip events were detected in the measurements. This suggests that, in those studies, the tip dynamics is overdamped or the effective potential (7.2) is bistable.

7.2.4 Some Generalizations of the Standard PT Model

7.2.4.1 Disordered Potential

One of the biggest advantages of the PT model is its flexibility. With suitable modifications, it can be applied to study many variants of the nanofriction set-up. For instance, the standard PT model is characterized by a sinusoidal tip-substrate interaction potential (7.15) and corresponds to a perfectly periodic substrate lattice. However, other cases are worth to be analyzed, such as quasiperiodic lattices [46, 47] and lattices including defects [48, 61]. The presence of disorder or defects changes the local potential profile by modifying the potential barriers to overcome by the tip in every stick-slip cycle, and, on the other hand, it can also change the length of different slip events. Depending on parameter values, both effects cooperate or compete to change the friction force. This is especially noticeable at low temperatures, while thermal fluctuations at high enough temperatures (close to room temperature) can screen other effects.

7.2.4.2 Additional Slow Degrees of Freedom

As discussed in Sect. 7.2.2, the cantilever in contact with the substrate is extremely soft in the apex region. Correspondingly, it is a natural extension of the PT model to consider the apex and the tip as two separate objects connected by a spring. This consideration leads to the so-called two-mass-two-spring models, which have been introduced and analyzed recently [63–67]. A multitude of friction regimes have been discovered within such a two-dimensional extension of the PT model, see [64] for a comprehensive review.

The assumption implicit in the Langevin equation (7.9) is that, during a nanofriction experiment, the tip moves along a one-dimensional manifold, whereas the substrate is, obviously, a two-dimensional object. Therefore, friction effects have been studied in two-dimensional geometries, with the main issue being friction force as a function of the tip-motion angle [52, 57, 68–71]. The comparison of experimental and theoretical models at finite temperatures is an active current research topic.

7.2.4.3 AC Actuation Effects

Ultrasonic vibrations have been proposed as a valid method for reducing friction at the nanoscale [72–77]. The PT model can be useful to study the dynamics of the system in the presence of in-plane and out-of-plane actuation fields. In the context of the model, in the first case the support position is affected by a shaking term. For out-of-plane actuation, the tip-substrate potential amplitude is modulated by an oscillating term [76], or the tip-sample distance is taken explicitly into account with a corresponding modification of the tip potential [72]. As in the regular case, thermal fluctuations are incorporated in the model as an additive Gaussian noise.

The most important finding is the existence of a wide medium-frequency range ($\sim$kHz), where friction force is significantly reduced and even almost suppressed for intense enough actuation. The lower bound of this friction-reduced zone is determined by the inverse time to cover one lattice constant, v/a, and the upper bound by the effective damping.

7.2.5 Friction Force-velocity Relations

7.2.5.1 Stratonovich Formula

In the asymptotic case of very low effective stiffness κ_L, the magnitude of force fluctuations, which is of the order of $\kappa_L a$, is also small, so that one can approximately replace the instantaneous elastic force (7.5) in (7.9) with its average value, $\bar{f}$. Furthermore, if one considers the overdamped ($m \to 0$) limit [42, 43], then the Langevin equation (7.9) assumes the form

$$\eta\dot{x}(t) = -U'(x) + \bar{f} + \eta_c v + \sqrt{2\eta k_B T}\xi(t) \tag{7.19}$$

that describes the diffusion of a Brownian particle in a tilted periodic potential $U_{tilted}(x) = U(x) - (\bar{f} + \eta_c v)\, x$. The problem of finding the average velocity of such a particle, $\langle \dot{x} \rangle \equiv v$, has been solved by Stratonovich, who derived the analytic formula [78]

$$v = \frac{a k_B T (1 - e^{-a(\bar{f}+\eta_c v)/k_B T})}{\eta \int_0^a dx_1 \int_{x_1}^{x_1+a} dx_2\, e^{[U(x_1)-U(x_2)+(x_1-x_2)(\bar{f}+\eta_c v)]/k_B T}}. \tag{7.20}$$

The argument of the function in the right-hand side is not the average force $\bar{f}$, but rather the combination $\bar{f} + \eta_c v$. In order to plot the $\bar{f}$-v relation, one can, first, for each given value of the combined force $\bar{f} + \eta_c v$ calculate the corresponding velocity v using the Stratonovich formula (7.20), and then deduce the average friction force $\bar{f}$ corresponding to this velocity value by subtracting the value of $\eta_c v$ from the combined force. Apart from the result (7.20), we are not aware of any exact force-velocity relation applicable to the general case of arbitrary inertia m or effective stiffness κ_L.

7.2.5.2 General Case

In a number of works, the Langevin equation (7.9) was simulated numerically using a random number generator. In the stick-slip regime, the ensuing relation between the friction force (7.1) and pulling velocity was found to be logarithmic at fast pulling, where the cantilever performs only forward slips, namely,

$$\bar{f} \propto |\ln(v/v_0)|^{\alpha}, \tag{7.21}$$

where v_0 is some reference velocity. Fitting the simulation results with an expression of the type (7.15) yielded the exponents α close to unity [68]. From a theoretical perspective, the exponent α is related to the functional form for a potential barrier decrease as the cantilever base moves. The $\alpha = 1$ value [14] is achieved for a barrier decreasing linearly with the force (Bell-type expression [79]), whereas the exponent $\alpha = 2/3$ results from a linear-cubic approximation of the potential at small barriers [15, 16]. It has been recently pointed out [80] that the regimes corresponding to these two values of α can be distinguished only if one can probe a wide range of velocities covering many decades, whereas the typical experimentally accessible velocity range (usually not more than three decades) is too narrow to determine α unambiguously.

In the slow-pulling regime, the back-slips of the cantilever play a significant role, and the linear-response arguments predict that

$$\bar{f} \propto v, \tag{7.22}$$

with the proportionality constant being different from the substrate damping coefficient η_s (unless the potential corrugation is zero). Recently, it has been suggested [80] that the two regimes, logarithmic (7.21) and linear (7.22), can be unified by a phenomenological ansatz of the type

$$\bar{f} = f_0(T)\sinh^{-1}[v/v_0(T)] \tag{7.23}$$

with $f_0(T) \propto T^{2/3}$ and $v_0(T) \propto T$. The accuracy of this ansatz has been demonstrated for a velocity range covering about seven decades [80].

The logarithmic (or $\sinh^{-1}$) force-velocity relation can be obtained within a general framework of the rate theory, which is the subject of the next section.

7.3 Rate Theory

7.3.1 Rate Equation

The Langevin equation (7.9) can be obtained, at least formally, by projecting the microscopic state of the system onto a subspace characterized by given values of the slow collective degrees of freedom. If the total state space consists of many "regions of attraction", such that the system spends most of the time within any such region and only rarely performs transitions between them, then an even more coarse-grained description is possible, namely, the one that uses the language of occupation probabilities of such regions and transition rates between them [29]. This is the case for the cantilever stick-slip motion: the average time spent by the cantilever within a given lattice site, a/v, is of the order of 0.1 ms, whereas the timescale of the tip coordinate fluctuations, $x(t)$, can be estimated as the inverse resonance frequency, $\sqrt{m/\kappa_L}$, or, in the overdamped case, as the tip relaxation time, η/κ, and is at least two orders of magnitude faster for both estimates.

In view of Fig. 7.2, the elastic force (7.5) in the nth stick phase of the stick-slip motion can be naturally separated into two contributions, regular and random:

$$f(t) = f_n(t) + \delta f(t), \tag{7.24}$$

where the regular part, $f_n(t)$, is the force (7.17) corresponding to a minimum of the combined potential (7.2), and the random part $\delta f(t)$ results from the tip fluctuations about that minimum. The rate of interstitial slips, i.e. the slip probability per unit time, depends on the regular part of the force, $f_n(t)$, whereas the fluctuating part, $\delta f(t)$, becomes irrelevant in this coarse-grained picture. We will denote the forward rate out of the nth lattice site as $\omega_+(f_n)$. If the elastic force is not too high, the cantilever can also perform a back-transition into the previous lattice site with the backward rate $\omega_-(f_n)$. For symmetric substrates, the two rates, $\omega_+(f_n)$ and $\omega_-(f_n)$, are related: the rate to jump forward "along the force" equals the rate to jump

backward "against the force", i.e. $\omega_+(f_n) = \omega_-(-f_n)$. Furthermore, both rates are given by the Kramers-Arrhenius law [29]

$$\omega_+(f_n) = \omega_-(-f_n) = \omega_0\, e^{-\Delta U(f_n)/k_B T}, \tag{7.25}$$

where $\Delta U(f_n)$ is the force-dependent height of the energy barrier separating the current minimum from the next one, and the prefactor ω_0 depends usually quite weakly on force and temperature.

Neglecting the possibility of multiple slips, the probability $p_n(t)$ for the tip to find itself in the nth lattice site at the moment of time t obeys the rate (or master) equation [81, 82]:

$$\dot{p}_n(t) = -\left[\omega_+(f_n(t)) + \omega_-(f_n(t))\right] p_n(t) + \omega_+(f_{n-1}(t)) p_{n-1}(t) + \omega_-(f_{n+1}(t)) p_{n+1}(t). \tag{7.26}$$

Here, the first term in the right-hand side describes the transitions out of the nth lattice site into the next $(n+1)$st and the previous $(n-1)$st ones, and the remaining two terms correspond to transitions into the nth lattice site from the $(n-1)$st and the $(n+1)$st ones.

We are interested in the long-time limiting solution of the rate equation (7.26), which is established after the decay of transient processes. In this limit, the probability to find the tip in the nth lattice site at the moment of time t is the same as the probability to find it in the previous lattice site at the earlier time, shifted by an interval a/v necessary for the cantilever base to cover one lattice constant a at the velocity v:

$$p_n(t) = p_{n-1}(t - a/v). \tag{7.27}$$

We define the force probability distribution, $p(f)$, as the probability to find the tip in the nth lattice site at that moment of time t when the corresponding elastic force $f_n(t)$ has the given value f, i.e. $f_n(t) = f$. According to (7.17), this time is given by $t = f/(\kappa v) + na/v$, so that

$$p(f) = p_n\left(f/(\kappa v) + na/v\right), \quad p_n(t) = p(f - n\kappa a). \tag{7.28}$$

In view of the relation (7.27), the so defined $p(f)$ is independent of the index n. It can be shown [82] that the time-averaged value of any function of force, $g(f)$, can be expressed as

$$\langle g(f)\rangle = \frac{1}{\kappa a}\int_{-\infty}^{\infty} df\, g(f)\, p(f), \tag{7.29}$$

i.e. the function $p(f)/(\kappa a)$ has the physical meaning of force probability density.

The normalization condition expresses the fact that the tip finds itself in some lattice site with probability one for all times:

$$\sum_{n=-\infty}^{\infty} p_n(t) = \sum_{n=-\infty}^{\infty} p(f + n\kappa a) = 1. \tag{7.30}$$

In the second part, we used the second relation (7.28) and replaced $n \to -n$.

Using the definition (7.28) and making the change of variables (7.17) in the rate equation (7.26), we find that the force probability distribution satisfies the rate equation in the force domain [82]:

$$\kappa v p'(f) = -\left[\omega_+(f) + \omega_-(f)\right] p(f) + \omega_+(f + \kappa a) p(f + \kappa a) + \omega_-(f - \kappa a) p(f - \kappa a). \tag{7.31}$$

7.3.2 Validity Conditions

The rate approximation is valid if the system possesses two very different time scales: the fast one describing the relaxation of the tip within a given potential well and the much slower one describing the thermally activated interwell transitions of the tip. This is realized if the typical height of the barrier separating two adjacent potential minima is at least a few times larger than the thermal energy [29]:

$$\Delta U(f) \gg k_B T. \tag{7.32}$$

This condition implies that the Prandtl-Tomlinson parameter (7.16) must satisfy $\gamma_{PT} \gg 1$. Furthermore, pulling must proceed sufficiently slowly to allow the slips to occur before the condition (7.32) is violated, which happens at some force f_{max} for which $\Delta U(f_{max})$ equals a few $k_B T$. According to an estimate from [83], the pulling velocity must therefore satisfy the condition

$$v \ll -\frac{k_B T \omega(f_{max})}{\kappa \Delta U'(f_{max})}. \tag{7.33}$$

Pulling velocities v bigger than in (7.33) lead to the onset of the opposite regime of steady sliding [42, 43] characterized by friction forces increasing as $\eta_s v$.

7.3.3 Parameterization

The parameter values used in the rate approach (7.26) can be derived from the Langevin equation (7.9) in the limit of deep corrugation ΔU or low κ_L. In particular, the effective stiffnesses in the two approaches, κ and κ_L, are slightly different, see (7.18). The barrier height in the rate expression (7.25) can be approximated as

$$\Delta U(f) \cong \Delta U_0 \left(1 - f/f_c\right)^{\beta}, \tag{7.34}$$

where ΔU_0 is the barrier height at zero force, which is related to the corrugation depth ΔU of the potential (7.15), f_c is the critical force value at which the barrier vanishes, and the exponent β describing the barrier reduction with force depends on the functional form of the potential $U(x)$ from (7.2). Considering the sinusoidal potential (7.15), it is often assumed that the exponent $\beta = 3/2$, which results from the linear-cubic interpolation of the potential $U(x)$ at forces f slightly below f_c [15, 16]. Note that, in practice, other β-values of the order of unity have also been either assumed to fit experimental or simulation results or deduced from such a fitting [14, 50, 68, 84].

For a sinusoidal potential (7.15), the critical force is given by [15–17]

$$f_c = \pi \Delta U / a \tag{7.35}$$

and usually does not exceed several nanonewtons. The height ΔU_0 of the force-dependent potential barrier (7.34) at zero force was initially taken to be equal the corrugation depth ΔU of the coordinate-dependent potential energy (7.15) [15–17]. Such an approximation is valid, strictly speaking, in the limit of vanishingly small lateral stiffness κ_L, which is inherent in the Stratonovich formula (7.20). For a finite spring constant κ_L, an improved approximation is due to [85], namely

$$\Delta U_0 = \Delta U + \kappa_L a^2/8, \tag{7.36}$$

still leaving ΔU_0 and ΔU of comparable order of magnitude. We note that, while the expression (7.34) is a useful simple ansatz for the barrier height, other approximations also have been introduced, that lead to an almost perfect agreement of the results for the friction force obtained within the rate and Langevin approaches, see [60, 86] for details.

The expression for the rate prefactor ω_0 depends on whether the tip dynamics is underdamped or overdamped [29]. In the practically important overdamped limit, it is [29]

$$\omega_0 = \frac{\sqrt{U''_{min}|U''_{max}|}}{2\pi\eta}, \tag{7.37}$$

where $U''_{min,max}$ denote the curvature at the minimum and maximum of the potential (7.2) when the force (7.5) has the value f. More generally, the rate parameters can be derived from any multidimensional version of the Langevin equation (e.g. the two-mass-two-spring model [66]), and even from the all-atom description used in MD, where the heat-bath degrees of freedom are not "integrated out". This is achieved using the transition state theory, see [23] for details. While the energy parameter ΔU_0 and the force parameter f_c are more or less consistent in different publications, the parameter ω_0 is found to vary in a wide range of values, from tens of kHz [17, 50] to hundreds of MHz [62] and GHz [66].

7.3.4 Types of the Stick-slip Motion

Following [82], we describe here four regimes of the stick-slip motion that are inherent within the rate approach; see Fig. 7.3 for a numerical illustration obtained by simulating the stick-slip process using a Monte Carlo technique [82].

Perhaps, the most widely investigated type is realized for relatively high effective stiffness and relatively fast pulling, so that the back-transitions are negligible, see Fig. 7.3a. The characteristic feature of such curves is the existence of a force interval, where the stick probability is close to 1. In this regime, the dependence of the mean friction force on the pulling velocity is approximately logarithmic (7.21), reflecting the exponential dependence of the transition rate $\omega_+(f)$.

If pulling is slow, Fig. 7.3b, the back-transitions become important and have indeed been observed experimentally [87]. For ultraslow pulling, they occur at almost the same frequency as the forward transitions, resulting in a linear force-velocity relation (7.22).

The velocity value at which the transition between the linear and logarithmic regimes occurs can be estimated from the following reasoning [82]. A typical stick

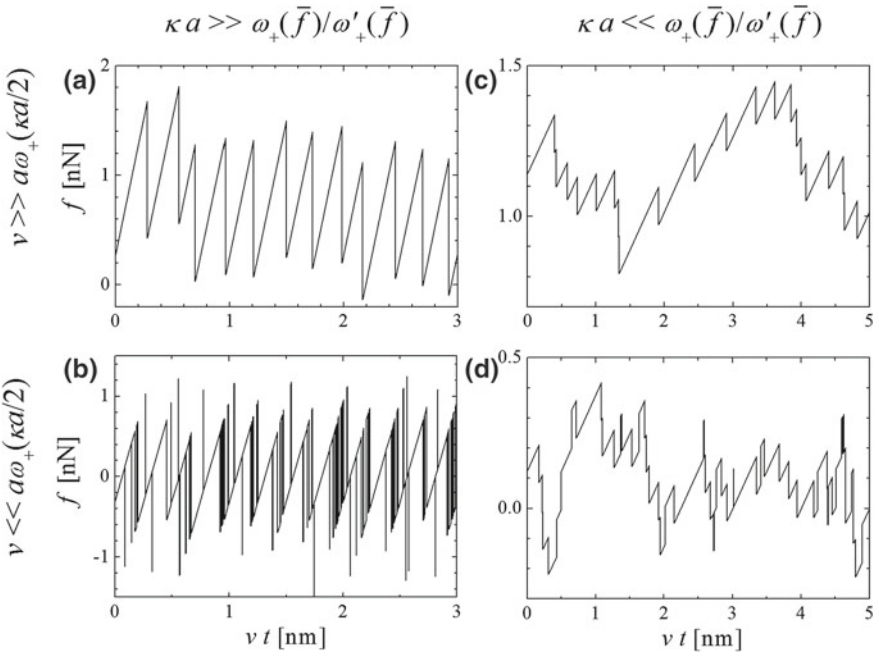


Fig. 7.3 The four basic types of stick-slip motion obtained by a numerical Monte Carlo simulation of the stick-slip process [82]. The rate parameters are as follows: $\omega_0 = 1$ MHz, $\Delta U_0 = 10\,k_BT$, $f_c = 3$ nN, $\beta = 3/2$, $a = 0.25$ nm. The effective stiffness $\kappa = 5$ N/m for curves (**a**) and (**b**), and $\kappa = 0.5$ N/m for (**c**) and (**d**). The pulling velocity $v = 2$ μm/s for curves (**a**) and (**c**), and $v = 0.01$ μm/s for curves (**b**) and (**d**)

phase begins at a force close to $\bar{f} - \kappa a/2$. Therefore, the typical frequency of the back transitions is $\omega_-(\bar{f} - \kappa a/2) = \omega_+(\kappa a/2 - \bar{f})$. Assuming that we are in the linear response regime where the mean force $\bar{f} \propto v$ is small, this can be approximated as $\omega_+(\kappa a/2)$. The back-slips will not be observed if the time to cover one lattice constant is faster than the inverse of the back-slip frequency. Thus, the force-velocity relation will be approximately logarithmic with back-jumps playing practically no role for $v \gg a\omega_+(\kappa a/2)$ and approximately linear for $v \ll a\omega_+(\kappa a/2)$.

The two regimes exemplified by the stick-slip curves from Fig. 7.3a and b are realized if the forward rate is very small below some velocity-dependent force value and high above this force value. Then, once formed, a given stick phase will survive in a rather wide force interval where the rate is small. Once the force increases beyond this interval, the rate will become very large, and the tip will slip into the next lattice site, resulting in the force drop by κa, a strong rate reduction, and a small probability of the next slip event until the force f reaches again the upper limit of that interval. Thus, the regimes shown in Fig. 7.3a, b are realized if the rate varies strongly within the force interval of the order of κa around the mean force value $\bar{f}$, i.e. $\omega'_+(\bar{f})\kappa a \gg \omega_+(\bar{f})$.

In the opposite case, $\kappa a \ll \omega_+(\bar{f})/\omega'_+(\bar{f})$, the rate $\omega_\pm(f)$ depends weakly on the applied force f. Then, right after some slip event, the probability to perform another slip remains significant, so that the short stick phases will be scattered in a rather broad force interval in such a manner that there is no particular force value that would be common to all of them. This situation is illustrated in Fig. 7.3c (fast pulling, no back-slips) and (d) (slow pulling, back-slips present).

7.3.5 Force-velocity Relations

7.3.5.1 Most Probable and Average Slip Force at High κ

In the regime of fast pulling and large stiffness exemplified by Fig. 7.3a, all stick phases cross some common force interval. Any force value f_L from that interval is characterized by the occupation probability $p(f_L)$ very close to one, and thus can be regarded as a "starting point" common to all stick phases. Then, a description even simpler than the rate equation (7.31) is possible [14–17, 50, 83, 88]. Namely, one can ask about the probability $\mathcal{P}(t|t_L)$ of staying within the same lattice site up to the moment of time t, provided that the tip entered this site at the initial time t_L, i.e. $\mathcal{P}(t_L|t_L) = 1$. The time evolution of $\mathcal{P}(t|t_L)$ for $t \geq t_L$ is governed by the rate equation, initially considered within the context of nanofriction in [14]

$$\frac{\partial \mathcal{P}(t|t_L)}{\partial t} = -\omega_+(f(t))\,\mathcal{P}(t|t_L). \tag{7.38}$$

With the help of the transformation of variables (7.17), we find from the rate equation (7.38) the probability that the transition into the next site occurs at a force value

between f and $f + df$, provided that the initial lower force value for a given stick phase was f_L,

$$-\frac{\partial \mathcal{P}(f|f_L)}{\partial f} = \frac{1}{\tilde{\kappa} v}\omega_+(f)\,\mathcal{P}(f|f_L) \tag{7.39}$$

with the initial condition $\mathcal{P}(f_L|f_L) = 1$. The most probable force f_*, at which the transition into the next site occurs, is evaluated by setting the second derivative of $\mathcal{P}(f|f_L)$ to zero. This results in the relation between the pulling velocity and the most probable force at the moment of slip:

$$v = \frac{\omega_+^2(f_*)}{\kappa \omega_+'(f_*)}. \tag{7.40}$$

This equation in various forms has been presented in [14–17, 66], the difference between these works stemming from different assumptions regarding the functional dependence of $\omega_+(f)$ on f. For the rate ansatz (7.25) with a constant prefactor ω_0 and the barrier height (7.34), the ensuing relation between the most probable slip force and the velocity is

$$\ln\frac{v}{v_0} = -\frac{\Delta U_0}{k_B T}\left(1 - \frac{f_*}{f_c}\right)^{\beta} - (\beta - 1)\ln\left(1 - \frac{f_*}{f_c}\right), \quad v_0 = \frac{\omega_0 f_c k_B T}{\beta \kappa \Delta U_0}. \tag{7.41}$$

It can be solved analytically [83]:

$$\frac{f_*}{f_c} = 1 - \left(\frac{\beta - 1}{\beta}\frac{k_B T}{\Delta U_0} W(z)\right)^{1/\beta}, \quad z = \frac{\beta}{\beta - 1}\left(\frac{\omega_0 f_c}{\beta \kappa v}\right)^{\beta/(\beta-1)}\left(\frac{k_B T}{\Delta U_0}\right)^{1/(\beta-1)} \tag{7.42}$$

where $W(z)$ is Lambert function defined implicitly by

$$W(z)\, e^{W(z)} = z. \tag{7.43}$$

Other approximations that have been introduced previously can be derived from the result (7.42). In particular, at large arguments z, the Lambert function behaves as a natural logarithm, thus leading to the asymptotic law (7.21) complemented by a simple relation between the exponents, α and β:

$$\alpha = 1/\beta. \tag{7.44}$$

With respect to the mean force $\bar{f}$, an accurate analytical approximation has been obtained in [50, 86], namely, a relation between force and velocity of the form

$$v = a\omega_+(\bar{f} + \kappa a/2)\, Q\left(\frac{\omega_+'(\bar{f} + \kappa a/2)}{\omega_+(\bar{f} + \kappa a/2)}\kappa a\right), \tag{7.45}$$

where the function $Q(x)$ can be approximated by

$$Q(x) \cong 1/\sqrt{1+(e^{-\gamma}x)^2}, \tag{7.46}$$

with $\gamma = 0.5772156649\ldots$ being Euler's constant. The accuracy of this approximation is better than a few per cent for all values of x. The relation (7.45) is valid for arbitrary stiffness κ, but fast pulling, so that the back-slips are absent. For large κ, the argument of the Q-function becomes large, allowing us to replace $Q(x)$ with e^{γ}/x. Then, the equation for the mean slip force $\bar{f}_{slip} = \bar{f} + \kappa a/2$ becomes almost identical to the one for most probable force, (7.40):

$$v = e^{\gamma}\frac{\omega_+^2(\bar{f}+\kappa a/2)}{\kappa\omega_+'(\bar{f}+\kappa a/2)}. \tag{7.47}$$

The solution is the same as (7.42), but with v replaced by $e^{-\gamma}v$ and f_* replaced by $\bar{f}+\kappa a/2$.

7.3.6 Force Probability Distribution

Because of the "advanced" and "retarded" terms on the right-hand side, an analytical solution of the rate equation in the force domain (7.31) subject to the normalization condition (7.30) is a highly non-trivial task. Analytical results can be obtained for the asymptotic cases of small and large stiffness κ, but for arbitrary pulling velocities. For this, it is convenient to look for the solution of (7.31) in the form that respects the normalization condition (7.30) from the outset, namely,

$$p(f) = P(f) - P(f-\kappa a), \tag{7.48}$$

such that the new unknown function $P(f)$ monotonically increases between two extreme values $P(-\infty)$ and $P(\infty)$ related by

$$P(\infty) - P(-\infty) = 1. \tag{7.49}$$

It has been found [82] that, for large κ, the asymptotic result is

$$P(f) = \frac{1}{\kappa v}\int_{-\infty}^{f} df'\, e^{-\frac{1}{\kappa v}\int_{f'}^{f} df''[\omega_+(f''+\kappa a)+\omega_-(f'')]}\omega_+(f'+\kappa a), \tag{7.50}$$

and for small κ, it is

$$P(f) = \frac{1}{2}\,\mathrm{erf}\left(\frac{f-\bar{f}+\kappa a/2}{\sigma}\right), \tag{7.51}$$

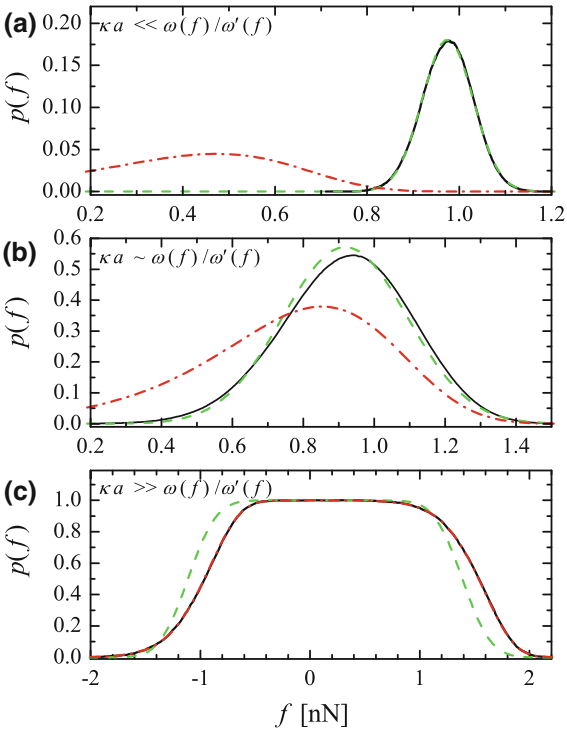


Fig. 7.4 Force probability distribution $p(f)$ for velocity $v = 1$ μm/s and other parameters as in Fig. 7.3. The three sets of data were obtained for different values of the stiffness, namely $\kappa =$ 0.1 N/m (**a**), 1 N/m (**b**), and 10 N/m (**c**), corresponding to small, intermediate, and large κa. *Black solid line*: Monte Carlo simulation of the rate equation [82]. *Red dash-dotted line*: high-κa approximation (7.48), (7.50). *Green dashed line*: low-κa approximation (7.48), (7.51). The *black* and the *green curves* practically coincide in panel **a**. In panel **c**, the *red* and the *black curves* are almost indistinguishable

where the mean force and the force dispersion are given by

$$v = a\omega_+(\bar{f})\frac{\sinh x}{x}e^{\frac{x}{2}[1-L(x)]}\left(1 - e^{-4\bar{f}x/(\kappa a)}\right), \quad \sigma = \kappa a\sqrt{\frac{1-L(x)}{2x}}, \tag{7.52}$$

with $x := \frac{\kappa a}{2}\frac{d\ln\omega_+(\bar{f})}{d\bar{f}}$ and $L(x) := \coth x - 1/x$ (Langevin function). The high accuracy of these expressions is demonstrated in Fig. 7.4.

7.4 Concluding Remarks

The equations used in stochastic modeling, in particular, Langevin equation and rate theory, are not exact laws of nature. Rather, they represent a useful approximation that accounts for the heat-bath effects on a nanoscopic system. In comparison to direct molecular dynamics, they have a number of advantages. Their simplicity often makes them amenable to analytical treatment, whose results can be used to interpret experimental findings. Even when analytical studies of stochastic models are difficult, they still can be easily simulated numerically, with the simulation time being orders

of magnitude faster than in the all-atom molecular dynamics approach. Stochastic models can usually be generalized to include additional experimental factors, such as the multidimensional nature of the problem, additional slow degrees of freedom, time-dependent external forcing, etc. Such modifications can be motivated by new experimental results, or introduced with the purpose of designing future experiments that would allow us to discover new phenomena in the nanofriction research area.

Acknowledgments We are grateful to the Deutsche Forschungsgemeinschaft (RE 1344/7-1) and the ESF programs NATRIBO and FANAS (collaborative research project Nanoparma) for financial support. JJM acknowledges Spanish MINECO through Project N. FIS2011-25167, cofinanced by FEDER funds.

References

1. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Oxford University Press, New York, 2008)
2. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, Nature **430**, 525 (2004)
3. N.J. Mosey, M.H. Müser, T.K. Woo, Science **307**, 1612 (2005)
4. S.Y. Krylov, J.W.M. Frenken, J. Phys.: Condens. Matter **20**, (35) (2008)
5. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987)
6. G. Binning, C.F. Quate, Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986)
7. C.M. Mate, *Trobology on the Small Scale* (Oxford University Press, New York, 2008)
8. G.V. Dedkov, Physics Uspekhi **43**, 541 (2000)
9. B.N.J. Persson, *Sliding Friction* (Springer, Berlin, 2000)
10. E. Gnecco, R. Bennewitz, T. Gyalog, E. Meyer, J. Phys.: Condens. Matter **13**, R619 (2001)
11. I. Szlufarska, M. Chandross, R.W. Carpick, J. Phys. D: Appl. Phys. **41**, 123001 (2008)
12. Y. Dong, A. Vadakkepatt, A. Martini, Tribol. Lett. **44**, 367 (2011)
13. A. Vanossi, N. Manini, M. Urbakh, S. Zapperi, E. Tosatti, Rev. Mod. Phys. **85**, 529 (2013)
14. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, H.-J. Güntherodt, Phys. Rev. Lett. **84**, 1172 (2000)
15. Y. Sang, M. Dubé, M. Grant, Phys. Rev. Lett. **87**, 174301 (2001)
16. O.K. Dudko, A.E. Filippov, J. Klafter, M. Urbakh, Chem. Phys. Lett. **352**, 499 (2002)
17. E. Riedo, E. Gnecco, R. Bennewitz, E. Meyer, H. Brune, Phys. Rev. Lett. **91**, 084502 (2003)
18. S. Sills, R. Overney, Phys. Rev. Lett. **91**, 095501 (2003)
19. M.R. Sorensen, K.W. Jacobsen, P. Stoltze, Phys. Rev. B **53**, 2101 (1996)
20. A.I. Livshits, A.L. Shluger, Phys. Rev. B **56**, 12482 (1997)
21. G. Gao, R.J. Cannara, R.W. Carpick, J.A. Harrison, Langmuir **23**, 5394 (2007)
22. A. Martini, Y. Dong, D. Perez, A. Voter, Tribol. Lett. **36**, 63 (2009)
23. D. Perez, Y. Dong, A. Martini, A.F. Voter, Phys. Rev. B **81**, 245415 (2010)
24. W.K. Kim, M.L. Falk, Modelling Simul. Mater. Sci. Eng. **18**, 034003 (2010)
25. W.K. Kim, M.L. Falk, Phys. Rev. B **84**, 165422 (2011)
26. H. Grabert, *Projection Operator Techniques in Nonequilibrium Statistical Mechanics* (Springer, Berlin, 1982)
27. P. Hänggi, H. Thomas, Phys. Rep. **88**, 207 (1982)
28. H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1984)
29. P. Hänggi, P. Talkner, M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990)
30. P. Reimann, Phys. Rep. **361**, 57 (2002)
31. M. Evstigneev, P. Reimann, Phys. Rev. B **82**, 224303 (2010)
32. L. Prandtl, Zeitschr. f. Angew. Math. u. Mech. **8**, 85 (1928)
33. G.A. Tomlinson, Philos. Mag. **7**, 905 (1929)

34. M.A. Lantz, S.J. O'Shea, M.E. Welland, K.L. Johnson, Phys. Rev. B **55**, 10776 (1997)
35. M.A. Lantz, S.J. O'Shea, A.C.F. Hoole, M.E. Welland, Appl. Phys. Lett. **70**, 970 (1997)
36. J. Colchero, A.M. Baró, O. Marti, Tribology Lett. **2**, 327 (1996)
37. R.W. Carpick, D.F. Ogletree, M. Salmeron, Appl. Phys. Lett. **70**, 1548 (1997)
38. K.L. Johnson, J. Woodhouse, Tribology Lett. **5**, 155 (1998)
39. R. Bennewitz, T. Gyalog, M. Guggisberg, M. Bammerlin, E. Meyer, H.-J. Güntherodt, Phys. Rev. B **60**, R11301 (1999)
40. R. Kubo, Rep. Progr. Phys. **29**, 255 (1966)
41. P. Reimann, Chem. Phys. **268**, 337 (2001)
42. P. Reimann, M. Evstigneev, Phys. Rev. Lett. **93**, 230802 (2004)
43. P. Reimann, M. Evstigneev, New J. Phys. **7**, 25 (2005)
44. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)
45. O.J. Furlong, S.J. Manzi, V.D. Pereyra, V. Bustos, W.T. Tysoe, Tribol. Lett. **39**, 177 (2010)
46. O.Y. Fajardo, J.J. Mazo, Phys. Rev. B **82**, 035435 (2010)
47. A.E. Filippov, A. Vanossi, M. Urbakh, Phys. Rev. Lett. **104**, 074302 (2010)
48. O.Y. Fajardo, J.J. Mazo, J. Phys.: Condens. Mat. **23**, 355008 (2011)
49. A.E. Filippov, V.L. Popov, Phys. Rev. E. **75**, 027103 (2007)
50. M. Evstigneev, A. Schirmeisen, L. Jansen, H. Fuchs, P. Reimann, Phys. Rev. Lett. **97**, 240601 (2006)
51. Y. Hoshi, T. Kawagishi, H. Kawakatsu, Jpn. J. Appl. Phys. Part 1 **39**, 3804 (2000)
52. R. Roth, Th Glatzel, P. Steiner, E. Gnecco, A. Baratoff, E. Meyer, Tribol. Lett. **39**, 63 (2010)
53. M. Reinstädtler, U. Rabe, V. Scherer, U. Harmann, A. Goldade, B. Bhushan, W. Arnold, Appl. Phys. Lett. **82**, 2604 (2003)
54. S.Y. Krylov, K.B. Jinesh, H. Valk, M. Dienwiebel, J.W.M. Frenken, Phys. Rev. E **71**, 065101 (2005)
55. J. Nakamura, S. Wakunami, A. Natori, Phys. Rev. B **72**, 235415 (2005)
56. S.N. Medyanik, W.K. Liu, I.-H. Sung, R.W. Carpick, Phys. Rev. Lett. **97**, 136106 (2006)
57. E. Gnecco, R. Roth, A. Baratoff, Phys. Rev. B **86**, 035443 (2012)
58. J.S. Helman, W. Baltensperger, J.A. Hołyst, Phys. Rev. B **49**, 3831 (1994)
59. W.G. Conley, A. Raman, C.M. Krousgrill, J. Appl. Phys. **98**, 0535519 (2005)
60. M. Evstigneev, P. Reimann, Phys. Rev. B **73**, 113401 (2006)
61. Z. Tshiprut, S. Zelner, M. Urbakh, Phys. Rev. Lett. **102**, 136102 (2009)
62. L. Jansen, H. Hölscher, H. Fuchs, A. Schirmeisen, Phys. Rev. Lett. **104**, 256101 (2010)
63. S.Y. Krylov, J.A. Dijksman, W.A. van Loo, J.W.M. Frenken, Phys. Rev. Lett. **97**, 166103 (2006)
64. S.Y. Krylov, J.W.M. Frenken, New J. Phys. **9**, 398 (2007)
65. S.Y. Krylov, J.W.M. Frenken, Phys. Rev. B **80**, 235435 (2009)
66. Y. Dong, H. Gao, A. Martini, EPL **98**, 16002 (2012)
67. Z. Tshiprut, A.E. Filippov, M. Urbakh, J. Phys.: Condens. Matter. **20**, 354002 (2008)
68. C. Fusco, A. Fasolino, Phys. Rev. B **71**, 045413 (2005)
69. J. Kerssemakers, J.Th.M. De Hosson, Appl. Phys. Lett. **67**, 347 (1995)
70. P. Steiner, R. Roth, E. Gnecco, A. Baratoff, S. Maier, Glatzel, E. Meyer, Phys. Rev. B **79**, 045414 (2009)
71. E. Gnecco, O.Y. Fajardo, C.M. Pina, J.J. Mazo, Trib. Lett. **48**, 33 (2012)
72. Z. Tshiprut, A.E. Filippov, M. Urbakh, Phys. Rev. Lett. **95**, 016101 (2005)
73. A. Socoliuc, E. Gnecco, S. Maier, O. Pfeiffer, A. Baratoff, R. Bennewitz, E. Meyer, Science **313**, 207 (2006)
74. S. Jeon, T. Thundat, Y. Braiman, Appl. Phys. Lett. **88**, 214102 (2006)
75. R. Guerra, A. Vanossi, M. Urbakh, Phys. Rev. E **78**, 036110 (2008)
76. R. Capozza, A. Vanossi, A. Vezzni, S. Zapperi, Phys. Rev. Lett. **103**, 085502 (2009)
77. O.Y. Fajardo, R. Roth, E. Meyer, J.J. Mazo, E. Gnecco, Phys. Rev. B 89, 075423 (2014); R. Roth, O.Y. Fajardo, J.J. Mazo, E. Meyer, and E. Gnecco, Appl. Phys. Lett. **104**, 083103 (2014)
78. R.L. Stratonovich, Radiotekhn. Elektron. (Moscow) **3**, 497 (1958) (English translation ed. by P.I. Kuznetsov, R.L. Stratonovich, and V.I. Tikhonov, Non-linear Transformations of Stochastic Processes (Pergamon, Oxford, 1965))

79. G.I. Bell, Science **200**, 618 (1978)
80. M. Müser, Phys. Rev. B **84**, 125419 (2011)
81. K.B. Jinesh, S.Yu. Krylov, H. Valk, M. Dienwiebel, J.W.M. Frenken, Phys. Rev. B **78**, 155440 (2008)
82. M. Evstigneev, P. Reimann, Phys. Rev. B **87**, 205441 (2013)
83. M. Evstigneev, Phys. Rev. E **78**, 011118 (2008)
84. M. Evstigneev, A. Schirmeisen, L. Jansen, H. Fuchs, P. Reimann, J. Phys.: Condens. Matt. **20**, 354001 (2008)
85. O.J. Furlong, S.J. Manzi, V.D. Pereyra, V. Bustos, W.T. Tysoe, Phys. Rev. B **80**, 153408 (2009)
86. M. Evstigneev, P. Reimann, Phys. Rev. E **71**, 056119 (2005)
87. S. Maier, Y. Sang, T. Filleter, M. Grant, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. B **72**, 245418 (2005)
88. A. Schirmeisen, L. Jansen, H. Fuchs, Phys. Rev. B **71**, 245403 (2005)

Chapter 8
Experimental Observations of Superlubricity and Thermolubricity

Martin Dienwiebel and Joost W.M. Frenken

Abstract The chapter introduces and discusses nanoscale effects called superlubricity and thermolubricity. Superlubricity is the phenomenon in which two surfaces slide over each other in dry contact without the atomic-scale instabilities. Superlubricity can reduce friction forces by orders of magnitude. Thermolubricity is the effect that thermal excitations significantly assist the contact between two bodies in overcoming the energy barriers against sliding, resulting in a reduction of the friction forces of contacts that are not superlubric.

8.1 Introduction

In this chapter we discuss manifestations of two effects, which we shall refer to as *superlubricity* and *thermolubricity*. Superlubricity is the phenomenon in which two surfaces slide over each other in dry contact without the atomic-scale instabilities that are thought to be the main source for energy dissipation. Superlubricity can reduce friction forces by orders of magnitude. Thermolubricity is the effect that thermal excitations significantly assist the contact between two bodies in overcoming the energy barriers against sliding, resulting in a reduction of the friction forces of contacts that are not superlubric. We shall argue that together, the two effects may lead to near-frictionless sliding over a wide range of conditions. Although we demonstrate these special effects for nanoscale contacts, we propose that they play

M. Dienwiebel (✉)
Karlsruhe Institute of Technology, Institute for Applied Materials—Reliability of Systems and Components, Microtribology Center μTC, Kaiserstr. 12, 76131 Karlsruhe, Germany
e-mail: martin.dienwiebel@iwm.fraunhofer.de

J.W.M. Frenken
Kamerlingh Onnes Laboratory, Leiden University, Im Schlehert 32, P.O. Box 9504, 2300 RA Leiden, The Netherlands

J.W.M. Frenken
Advanced Research Center for Nanolithography, Science Park 104, 1098 XG Amsterdam, The Netherlands
e-mail: frenken@physics.leidenuniv.nl

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_8

a key role in the well-known lubricating properties of some layered materials, such as graphite and molybdenum disulfide, and speculate they may hold a promise for further, low-friction applications.

8.1.1 The Transition to Frictionless Sliding in the One-Dimensional Case

In friction force microscopy (FFM) experiments at the atomic scale the lateral force signals often show a sawtooth-like modulation with the periodicity of the lattice of the substrate over which the tip is being moved. Observations of this type have been made on many different materials, such as graphite [1], mica [2], MoS_2 [3], copper [4], diamond [5, 6], and alkali-halides (NaF, NaCl, KF, KCl, KBr) [7–9]. The general characteristics of these observations are described well by a simple model that has been formulated first by Prandtl [10] and by Tomlinson [11]. Applied to the nanoscale geometry of a friction force microscope the model describes the motion of a point-like tip which is coupled by a spring to a moving support. The tip is in contact with a rigid solid, which is treated as a periodic potential energy surface. If this potential energy landscape has only a single Fourier component, with period a and amplitude V_0, we can write the lateral force on the tip as

$$\frac{2\pi}{a} V_0 \sin\left(\frac{2\pi}{a} x_t\right) = k(x_m - x_t) \tag{8.1}$$

where k is the stiffness of the spring and x_t and x_m denote the positions of the tip and the support.

The relative strength of the interaction potential with respect to the stiffness of the spring is often expressed in the form of a dimensionless parameter $\gamma \equiv 4\pi^2 V_0 / k a^2$. When γ exceeds unity, multiple solutions exist to (8.1). The tip remains stuck in a metastable equilibrium position until the spring force is large enough to force the tip to rapidly slip to the next equilibrium position, which may again be merely metastable. These two elements, the sticking and the slipping, represent the stick-slip motion, commonly observed in FFM experiments. A tacit assumption in this description is that the excess potential energy that is released during the slip event is irretrievably lost to other degrees of freedom of the system, such as phonons in the substrate and the tip. As a consequence, stick-slip motion implies energy dissipation. By contrast with this scenario for energy dissipation, (8.1) predicts continuous sliding of the tip over the counter surface for $\gamma < 1$, i.e. when the surface potential is sufficiently weak and the spring is sufficiently stiff. In this case, the lateral force oscillates between negative (against the direction of attempted motion) and positive (forward) values and the average lateral force, i.e. the *friction* force, is zero. The transition from zero dissipation to non-zero friction at $\gamma = 1$, due to the breaking of analyticity, is known as an Aubry transition [12]. The Aubry transition has been investigated primarily in the context of the one-dimensional Frenkel-Kontorova model [12, 13], in which two

one-dimensional lattices are sheared over each other. In this model, static friction and the Aubry transition depend strongly on the ratio p/q of the lattice constants of the top and bottom solid.

8.1.2 Superlubricity

The term *superlubricity* has been introduced by Hirano and Shinjo [14]. Originally, it was defined as the extension of the Aubry transition to a two-dimensional geometry and describes the effect that friction can vanish almost completely when two crystalline surfaces slide over each other in dry contact without wear and plastic deformation. This phenomenon was first demonstrated in a quasistatic calculation for rigid crystals with fcc, bcc and hcp symmetry and for various surface orientations [15]. In the two-dimensional case it was found that the frictionless or superlubric regime can be reached for a much wider range of values of γ and they noted that superlubricity should appear for any combination of flat and clean metals when the interaction potential is weak. Hirano and Shinjo concluded that a way to tune the interaction potential experimentally, is to change the commensurability between two surfaces. Of the theoretical work that has been performed after these first calculations we mention that by Sørensen et al. [16], who used molecular dynamics simulations to investigate friction at $T = 0\,\mathrm{K}$ between flat copper asperities, e.g. 19×19 atoms large, and a copper surface. When the asperity and the surface were both (111) oriented, the sliding did not involve wear. For an aligned contact, regular stick-slip motion was observed with high friction, whereas the friction force vanished when the contact was twisted 16.1° out of registry.

In this chapter, we will use the word superlubricity, even though it suggests an analogy between the *structural lubricity* at an incommensurate interface and the phenomena of superconductivity and superfluidity [17]. Lately, the term superlubricity has been used by several authors to also indicate other situations with extraordinarily low friction forces, not involving a lattice mismatch effect. Here, we will stick to the original meaning of the word and concentrate on experiments that probe the effect of commensurability on friction.

8.1.3 In Search for Superlubricity

In an early experiment, Hirano et al. [18] have employed a surface forces apparatus (SFA) to measure the orientation dependence of friction between mica sheets. They found a friction force of ($8 \times 10^{-4}\,\mathrm{N}$) when the orientations of the mica sheets matched. The friction force was reduced by as much as a factor 4 when the crystallographic directions of the mica sheets were misoriented relative to each other. Since in the incommensurate case friction was still relatively strong, either the superlubricity effect was incomplete or other mechanisms were responsible for additional channels of energy dissipation.

As was shown by Ko and Gellman [19], one such additional type of dissipation can be the internal friction that arises when the contact pressure is high enough to cause plastic deformation. These authors measured the friction force as function of the misfit angle between two Ni(100) crystal surfaces using an ultrahigh vacuum (UHV) tribometer and found a lower friction coefficient for 45° and 135° misfit angles than for 0° and other orientations. Although at first sight, this observation seems consistent with superlubricity, the orientational variations were still observed after adsorption of as much as 20 monolayers of ethanol or sulfur on the nickel surfaces prior to contact formation, which made Ko and Gellman conclude that the low friction in certain directions was caused by easy shearing along the preferred slip planes in the bulk. This explanation is consistent with the computer-simulation result obtained by Sørensen et al. [16] for shearing contacts between clean copper surfaces, which revealed that the shear occurred predominantly along the (111) planes, even for (001) oriented surfaces.

In another macroscopic experiment Martin et al. found a remarkably low friction coefficient between clean MoS_2 surfaces after a short sliding distance using a UHV tribometer [20]. After the experiment MoS_2 flakes were collected and examined with a transmission electron microscope (TEM). The TEM images showed that the flakes were rotated with respect to each other. The authors concluded that the low friction coefficient was due to the incommensurability between the flakes, thus due to superlubricity.

In 1997 Hirano et al. [21] have performed a scanning tunnelling microscope (STM) experiment and claimed the observation of superlubricity in UHV between a tungsten tip and a Si(001) surface. The tungsten tip was first imaged using field electron microscopy, after which it was advanced towards the Si surface until a tunnelling current could be measured. Since an STM is usually not capable of detecting forces, the bending of the tip was monitored optically and translated into a lateral force. When the major crystallographic axes of the two surfaces were aligned, a bending of the tungsten wire over an estimated 100 nm was measured, which was absent when the orientation was rotated over 45°.

8.2 Atomic-Scale Observation of Superlubricity

8.2.1 Commensurability-Dependent Superlubricity Between Finite Graphite Surfaces

In this chapter we concentrate on nanotribological experiments, conducted with a dedicated friction force sensor, the Tribolever [22]. This sensor is part of an unconventional friction force microscope [23] that allows quantitative tracking of the forces on the scanning tip in three directions, with a high resolution in the lateral forces, down to 15 pN. The instrument can rotate the sample to change the relative orientation between the tip and sample lattices. Initial measurements with this

instrument have been performed on low-grade, highly oriented pyrolytic graphite (HOPG) [24, 25].

Although in some cases, these measurements showed traditional friction loops, with stick-slip character and measurable energy dissipation, in many friction loops the average friction force was very low and the tip was sliding over the graphite surface without stick-slip motion. Also it was found that the variation of the friction force with the normal force was rather weak. These observations strongly suggested that the sliding had been taking place between two graphite surfaces, one being the HOPG substrate, and the other being a small piece of graphite, i.e. a graphite flake, that was attached to the tungsten tip. The difference between the high- and low-friction force loops could then be attributed to the difference in commensurability between the flake and the substrate, high friction corresponding to a fully commensurate contact and low friction to an incommensurate contact.

In order to obtain further support for our interpretation of these initial observations, we have repeated the experiment much more carefully with a high-quality HOPG sample that had an average grain size of several millimeters. In the experiment we rotated this sample in small steps with respect to the tip. For each orientation, we performed a complete set of friction force spectroscopy measurements for a range of normal forces between +25 nN and pull-off (−22 nN) and a range of sliding directions. We recorded the lateral forces in the X- and the Y-direction of the sensor, from which we reconstructed the average friction force in the sliding direction (for details see [25]).

Figure 8.1 shows lateral force maps and force loops measured in the X-direction for different rotational orientations. A typical force loop is shown in Fig. 8.1d, which was measured at a normal force of 18 nN. The lateral force in Fig. 8.1d displays clearly resolved atomic-scale stick-slip sliding and the average friction force parallel to the sliding direction is 203.3 ± 20 pN. Figure 8.1b, e and c, f show FFM measurements obtained with the graphite substrate rotated +12° and −22° with respect to Fig. 8.1a, d around an axis normal to the surface and parallel to the tip. The rotation by 12° has caused the average friction force to reduce by more than one order of magnitude, to 15 ± 15 pN. Rotating 22° away from the first measurement in the opposite direction also has caused a reduction, to 8^{+16}_{-8} pN, which is equal to zero friction within the detection limit of our instrument. This variation of the friction force with the rotation angle Φ was completely reversible. Notice that the ultra-low lateral forces in Fig. 8.1e–f still exhibit regular variations with the periodicity of the graphite substrate.

Figure 8.2 displays the average friction forces measured over a 100° range of substrate rotation angles. We recognize two narrow angular regions with high friction, separated by a wide angular interval with nearly zero friction. The distance between the two friction peaks is $61 \pm 2°$, which corresponds well with the 60° symmetry of individual atomic layers in the graphite lattice. After every 60° rotation, the lattices of the substrate and the graphite flake align and the friction is high. For intermediate angles, the lattices are incommensurate and the friction force is close to zero.

The peak width in Fig. 8.2 can be used to estimate the flake diameter. For finite-size contacts, the cancellation of lateral forces, which causes superlubricity, can be

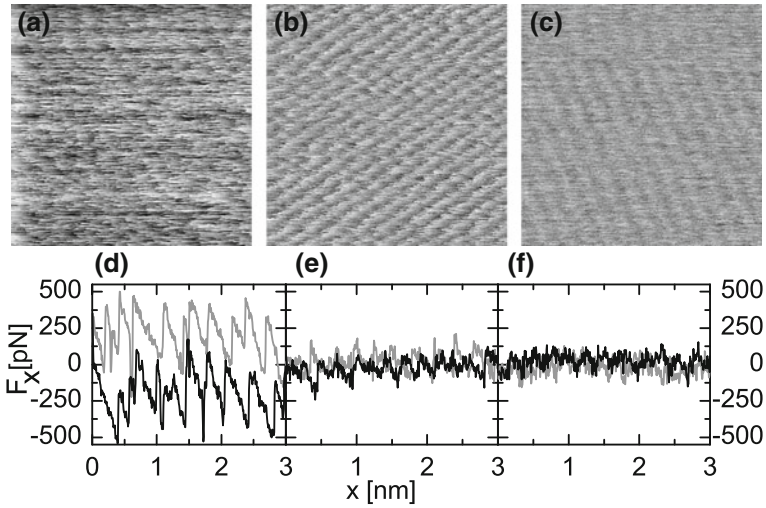


Fig. 8.1 Lateral force images (*forward direction*) and friction loops measured between a tungsten tip and a graphite substrate. The displayed signals correspond to the X-direction of the Tribolever sensor and rotation angles Φ of the graphite sample of 60° (**a**, **d**), 72° (**b**, **e**) and 38° (**c**, **f**). The normal force between the tip and the substrate amounted to $F_N = 18$ nN in (**a**, **d**) and (**c**, **f**) and $F_N = 30.1$ nN in (**b**, **e**). The grey scales in the force images cover force ranges of **a** 590 pN, **b** 270 pN, and **c** 265 pN. The image size is 3 nm × 3 nm. After [25]

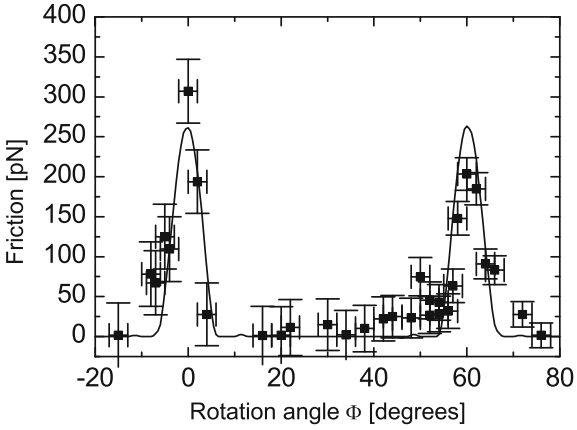


Fig. 8.2 Average friction force between a tungsten tip and a graphite substrate, plotted versus rotation angle Φ of the graphite sample with respect to an axis normal to the sample surface. Two *narrow peaks* of high friction are observed at 0° and 61°, respectively. Between these peaks a wide angular range with ultra-low friction, close to the detection limit of the instrument, is found. The first peak has a maximum friction force of 306 ± 40 pN, and the second peak has a maximum of 203 ± 20 pN. The curve through the data points shows results from a calculation according to the Tomlinson model for a symmetric 96-atom graphite flake sliding over the graphite surface (see text). After [25]

considered complete when the mismatch between the two lattices adds up to one lattice spacing over the diameter of the contact. The mismatch condition provides us with the estimate that $\tan(\Delta\Phi) = 1/D$, where $\Delta\Phi$ is the full width at half maximum of the friction peak, and D is the flake diameter, expressed in lattice spacings. From the widths of the two peaks in Fig. 8.2, of $5.4 \pm 1.0°$ for the first peak and $6.5 \pm 0.8°$ for the second, we estimate that the flake diameter is between 7 and 12 lattice spacings. The contact size was determined more precisely by Verhoeven et al. [26]. He modelled the flake as a rigid, finite lattice, with the hexagonal symmetry of a single layer of graphite. Because the relative positions of the atoms in the N-atom flake $(x_i, y_i, 0)$ with respect to the position (x_t, y_t, z_t) of the center of mass (CM) of the flake were fixed, the flake-surface interaction potential could simply be obtained by the summation over N atomic contributions:

$$V_{int}(x_t + x_i, y_t + y_i, z_t) = -V_0(z_t)[2\cos(b_1(x_t + x_i))\cos(b_2(y_t + y_i)) + \cos(2b_2(y_t + y_i))] + V_1(z_t), \qquad (8.2)$$

with $b_1 = 2\pi/(0.246\,\text{nm})$ and $b_2 = 2\pi/(0.426\,\text{nm})$. The height-dependent corrugation amplitude per flake atom is given by $V_0(z)$, while $V_1(z)$ indicates the overall, i.e. N-atom position-averaged z-dependence of the interaction, expressed per atom. The amplitude of the summed potential depended strongly on the orientation angle Φ of the flake lattice with respect to the substrate lattice. The flake was coupled to a support by springs in the x- and y-directions (see Fig. 8.3), with which it was dragged through this N-atom interaction potential.

Symmetric flakes of various sizes were considered in the calculation. Each flake was a piece of graphene sheet and had a shape with 60° rotational symmetry.

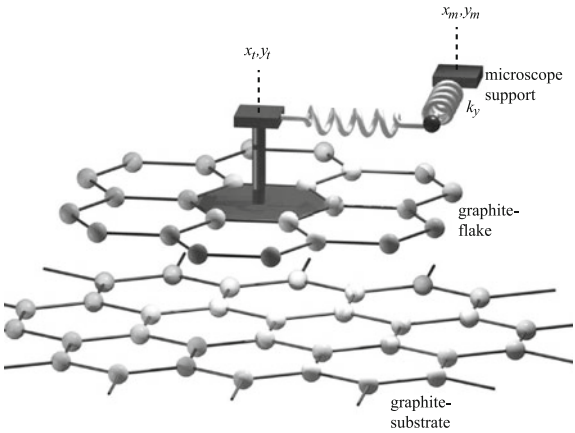


Fig. 8.3 Illustration of the modified Tomlinson model used in our calculations. A rigid graphite flake consisting of N carbon atoms (here $N = 24$) is connected by an x-spring and a y-spring to the support of the microscope. The support is moved in the x-direction. The substrate is modelled as a rigid, infinite, single layer of graphite. From [26]

As expected, the friction force was maximal if the misfit angle Φ was zero (or a multiple of 60°). For these orientations, the friction force increased linearly with the number of atoms N in the flake. In order to compare all different flake sizes for the same fixed total interaction between the flake and the surface, the potential amplitude per atom V_0 was lowered with increasing flake size such that the calculated friction force with the flake and substrate in registry was the same for all flakes, namely 265 pN at 0° pulling direction, between the values measured experimentally at misfit angles of 0° and 60°.

The effective interaction potential energy surface (PES) for the flake as a whole V_{int}^{flake} is shown in Fig. 8.4a for matching lattices ($\Phi = 0°$) for a flake size of $N = 96$. The small grey areas overlayed on the PES are the flake positions recorded in the 'forward' scan direction, during the 3 nm × 3 nm scan, parallel to the x-axis ($\Theta = 0°$). In Fig. 8.4a, the flake is only found in limited regions, slightly displaced to the upper right with respect to the minima of the PES. Also shown are flake pathways for three separate scan lines. During the scanning process the flake moves continuously through the grey 'sticking' regions, while force is built up in the spring. From the end of such a region it jumps (slips) to the beginning of the next sticking region. When the 96-atom flake is misaligned by 7°, the calculated lateral forces become small and for most trajectories the average lateral force, i.e. the friction

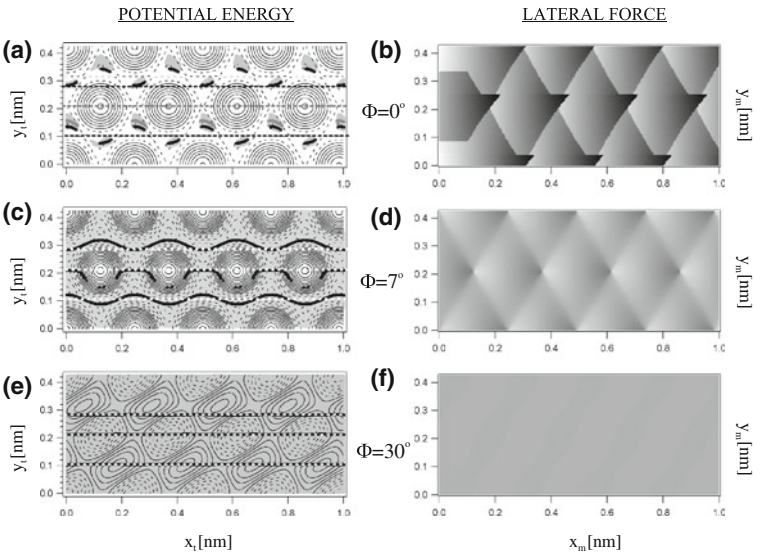


Fig. 8.4 Total potential energy surfaces and lateral force images (1.0 nm × 0.426 nm), calculated in the forward x-direction for a symmetric, 96-atom graphite flake sliding over a graphite substrate, for misfit angles $\Phi = 0°$ (**a**, **b**), $\Phi = 7°$ (**c**, **d**) and $\Phi = 30°$ (**e**, **f**). The *grey* scale in the lateral force images corresponds to the range [−1.04, 0.63] nN. For this range, **b** has maximal contrast. The *grey areas* in the potential energy *contour plots* denote positions that were visited by the flake. The *black lines* denote pathways of the flake during single scan lines of the support. From [26]

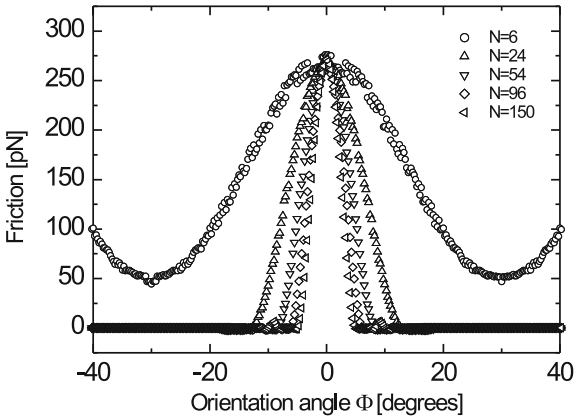


Fig. 8.5 Friction force as a function of the orientation angle for different symmetric graphite flakes ranging in size from 6 to 150 atoms, sliding over a graphite substrate. The potential amplitude per atom has been chosen such that all flakes share the same maximum friction value of 265 pN at $\Phi = 0°$. From [26]

force, vanishes completely within the precision of the calculation. The corrugation of the PES has decreased with respect to the situation at $\Phi = 0°$ and the regions addressed by the flake have merged, indicating that the flake moves continuously through most of the PES.

If the misalignment between the 96-atom flake and the substrate is further increased to 30°, the corrugation of the PES becomes so low that the pathway of the flake through the PES is identical to that of the support, within the precision of the calculation. The flake-graphite contact is now completely superlubric.

Figure 8.5 displays the computed friction force as a function of the misfit angle Φ, for five symmetric flakes with different sizes. We find an angular region with high friction around 0°, repeating every 60° due to the rotational symmetry of the flakes. At intermediate angles, near-zero friction is calculated, except for the 6-atom flake, for which the friction drops to 52 pN. These numerical calculations confirm the simple geometrical estimate, mentioned above, the best fit being produced by a flake with a size of $N = 96$ atoms. Further calculations showed that the shapes of the peaks in Fig. 8.5 also depended on the shapes of the flakes, the best fit to the experimental data being obtained for a symmetric flake, as shown by the curve in Fig. 8.2.

8.2.2 The Role of the Normal Force

As is clear from the description in Sect. 8.1.1, sliding without instabilities is possible also in the case of a single-atom contact or a contact between two commensurate lattices, provided that γ is smaller than 1. The required reduction of V_0 can be achieved by making the normal force sufficiently small, with which the tip is pressed

against the substrate. In case of an attractive tip-substrate interaction, this may even require a negative external force, i.e. pulling the tip. This approach has been demonstrated recently by Socoliuc et al. In their FFM experiment, a silicon tip was scanned over a NaCl surface along the (100) direction in UHV. The observed force loops showed excellent agreement with the transition to frictionless sliding that is predicted by the one-dimensional Prandtl-Tomlinson model. As γ approached unity at a normal force of $F_N = -0.47$ nN, the area enclosed in the friction loop and, thus, the energy dissipated in a cycle reduced to zero and the tip was observed to slide over the NaCl surface without stick-slip motion. At that point the lateral force still showed slightly distorted sinusoidal variations with the periodicity of the surface lattice, but the average force was zero (Fig. 8.6c).

Following the original definition by Hirano and Shinjo [14], we should not refer to this form of near-frictionless sliding at low (or even negative) normal forces as superlubricity, since it does not involve the cancellation between lateral forces on individual atoms in the contact resulting from a mismatch between the two contacting surface lattices. By contrast, in the case of superlubricity, the normal force can be made surprisingly high [25]. We expect that the natural limit in the normal force, or rather contact pressure, will be that a new energy dissipation channel is introduced when the contact pressure is made high enough to induce noticeable lateral elastic deformations in the two contacting surfaces. In the case of such deformations, we can no longer describe the situation as that of two translating, rigid bodies [17]. In addition, the contact will then carry a rapidly shifting deformation pattern that one may view as a lattice of dislocation lines. Moving this pattern involves the combination of breaking atoms out of registry on one side of these lines and the popping of atoms back into registry on the other side. As these changes in atomic

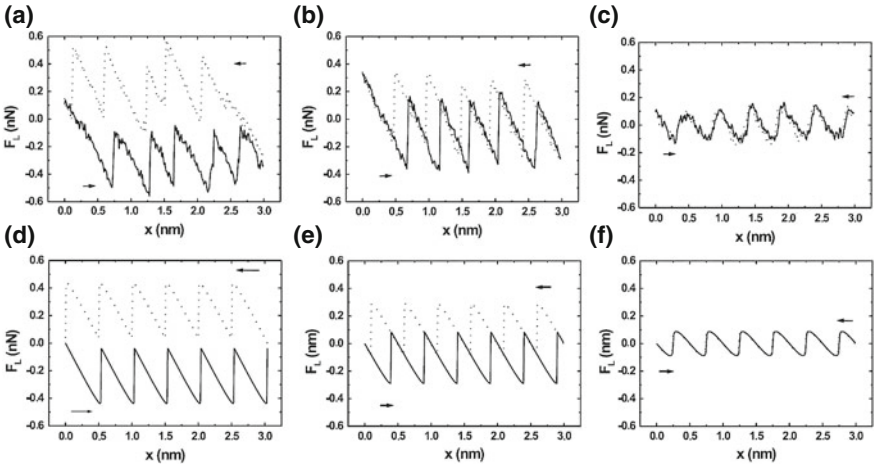


Fig. 8.6 Friction force loops of a Si tip sliding along the (100) direction of NaCl at **a** $F_N = 4.7$ nN, **b** $F_N = 3.3$ nN, and **c**, **a** $F_N = -0,47$ nN. Prandtl-Tomlinson calculation with **d** $\gamma = 5$, **e** $\gamma = 3$, and **f** $\gamma = 1$. Reprinted from [27] with permission by A. Socoliuc and E. Gnecco

positions are no longer rigidly connected, much of the energy released on one side will not be re-invested on the other side and will be lost in the form of heat (phonons).

8.3 The Role of Temperature

Our discussion in the previous sections has been presented in terms of the classical mechanics of the sliding system, combined with the assumption of instantaneous loss of the excess energy during each slip event. These elements form the basis of the Tomlinson model that was used to fit the data in Fig. 8.2. One of the obvious simplifications in the Tomlinson model is the complete absence of effects due to the spontaneous thermal excitations that are present at finite temperatures. Such effects have been anticipated already by Prandtl in 1928 [10]. Here, we summarize how they lower the average friction force and may lead to a strongly modified type of sliding motion.

8.3.1 Weak Thermal Effects

It is easy to see that thermal excitations will assist the tip in overcoming the energy barrier for sliding from one well in the potential energy surface to the next. One might expect this phenomenon to become noticeable only when the amplitude of the potential V_0 (see 8.1 and 8.2) and, thus, the energy barriers would be limited to only a few times the thermal energy $k_B T$. However, one should realize that the tip moves in the *combined* potential of its interaction with the surface and its interaction with the spring that connects it to the moving support. Within the Tomlinson model the tip remains stuck in one well of this combined potential until the spring is sufficiently extended that the energy barrier to the next well vanishes, at which point the system is unstable and the tip necessarily slips into the next well. This means that even when the barrier to the next well starts out at a high value, it decreases continuously to zero while the spring is being stretched. This implies that thermal excitations will *always* play a role, since the stick part of the stick-slip cycle always contains a portion during which the barrier is sufficiently low with respect to $k_B T$ that it can be overcome by a thermally activated jump. This so-called pre-critical jump somewhat lowers the maximum lateral force that is exerted by the spring and thus reduces the friction force. The effect depends weakly on the sliding velocity because at lower velocities there is time for more attempts of the system to thermally overcome each barrier, which should make friction reduce more. It is easy to provide an estimate of this reduction in the friction force. Due to the exponential nature of the thermal excitations there is a range of sliding velocities over which the friction force is expected to increase proportionally with the logarithm of the sliding velocity v [28–30] or with $(\log v)^{2/3}$ [31]. Such a weak dependence of friction on sliding velocity has indeed been observed experimentally [30, 32, 33].

8.3.2 Strong Thermal Effects: Thermolubricity

As the measurements in Fig. 8.2 show, the rotation angle Φ of the graphite flake with respect to the graphite substrate determines the amplitude of the potential V_0 and can tune it anywhere between a high value of γ, at which there is strong friction, and low values $\gamma < 1$ that result in superlubricity. If the spring coefficient of the sensor k is low, as was the case in the measurements of Fig. 8.2, the transition to superlubricity takes place at a comparatively low value of the potential amplitude V_0. For example, at a typical (low) spring coefficient of $k = 1\,\mathrm{N/m}$ and a typical lattice constant of $a = 0.25\,\mathrm{nm}$ the transition to superlubricity is at a potential amplitude of only 10 meV, which is lower than the thermal energy at room temperature of 25 meV. This implies that for a contact that is close to superlubricity the role of thermal excitations will be much more dramatic than that described in the previous section. Rather than to merely facilitate the jumps that were bound to happen anyway (pre-critical jumps), thermal excitations are now sufficient to efficiently promote the system over all barriers, both to the next well(s) and to the previous one(s) [28]. As a result, the tip will conduct a random walk over the surface, its average position following the slow translation of the support. The stochastic nature of this driven diffusion of the tip is characterized by rapid force variations that replace the periodic stick-slip character, typical for higher γ values. The average lateral force, i.e. the observed friction force, is strongly reduced by these thermal jumps. It is this behavior for which we have proposed the term *thermolubricity* [34].

Figure 8.7 shows two selected force loops measured between a graphite substrate and a tungsten tip (dressed with a graphite flake) for two relative orientations, corresponding to two different values of γ. While the force loop for the higher γ still displays recognizable stick-slip behavior, the force variations at the lower γ value are almost completely stochastic. Note that the average friction force of the lower loop is close to zero, although γ is still well above unity, i.e. the system is not yet superlubric. These qualitative features agree well with the thermolubricity scenario. A stronger test is a quantitative confrontation of these observations with numerical calculations of thermolubricity. In these calculations we describe the surface potential as a one-dimensional sine function (8.1) to which we add the potential due to the interaction with the spring, $\frac{1}{2}k(x_t - x_m)^2$. If $\gamma < 1$ this combined potential shows only a single minimum for every position x_m of the support and the sliding is frictionless (superlubricity). When $\gamma > 1$ the combined potential shows several wells. Rather than to calculate individual trajectories or concentrate on average behavior, we describe the process in terms of the probabilities p_i for the tip to reside in each well i at every point in time. These probabilities evolve according to a simple set of continuity equations of the type:

$$v\frac{dp_i}{dx_m} = -(r_i^+ + r_i^-)p_i + r_{i-1}^+ p_{i-1} + r_{i+1}^+ p_{i+1}, \tag{8.3}$$

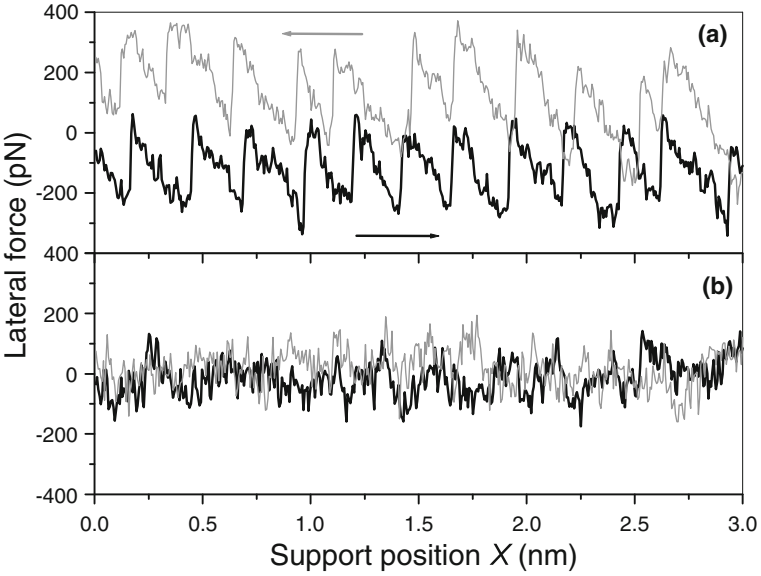


Fig. 8.7 Characteristic lateral force loops measured with a tungsten tip (with a graphite flake) on a graphite substrate at two different relative orientations, corresponding to **a** $\gamma = 5.0$ and **b** $\gamma = 2.5$. After [34]

where the role of time is played by the coordinate of the support $x_m = vt$. The rates r_i^+ and r_i^- of jumps from well i to the right and to the left are calculated according to the Arrhenius law:

$$r_i^{\pm} = r_0 \exp(-\frac{\Delta E_i^{\pm}}{k_B T}). \tag{8.4}$$

Here, $\Delta E_i^{\pm}$ are the energy barriers from well i to the next well and to the preceding one. The locations and heights of these barriers depend on the support position x_m. The prefactor r_0 is the frequency of attempted jumps, which is treated as a fitting parameter (see below). Equation (8.3) can be solved analytically in the limit of very strong thermolubricity, i.e. low velocities v, low amplitudes of the potential V_0 and/or high temperatures T. For other conditions the probabilities need to be evaluated numerically.

Figure 8.8 demonstrates that for $\gamma > 1$ the experiments presented before, probing the lateral forces between a graphite-decorated tungsten tip and a graphite substrate, clearly exhibit thermolubricity. For each value of the relative strength of the potential γ the experimental data fall significantly below the dashed curve, which is the friction force expected according to the Tomlinson model. The full curves show the results obtained from (8.3) for the values of T, v, a, and k taken from the experiment and for four different values of the attempt frequency r_0. Like the experiment, the theoretical curves fall below the curve for the Tomlinson model. The second curve, for $r_0 = 1.6\,\text{kHz}$, provides an excellent fit to the experiment. This frequency is in the

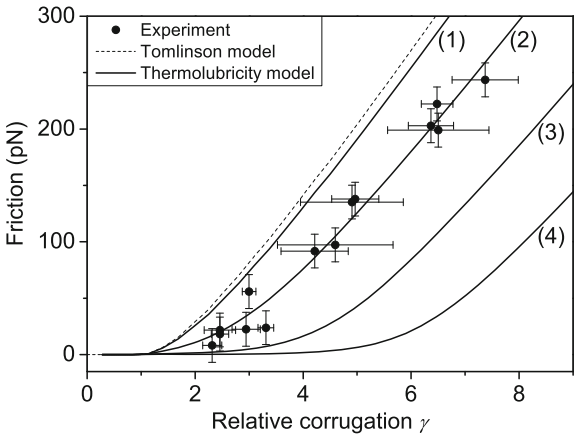


Fig. 8.8 Friction force as a function of relative surface corrugation γ. Experiments for a graphite substrate and a tungsten tip with a graphite flake are compared with numerical solutions of (8.3) for $v = 30\,\text{nm/s}$, $a = 0.25\,\text{nm}$, $k = 1.8\,\text{N/m}$ and $T = 290\,\text{K}$ and for $v/ar_0 = 7.53 \times 10^{-n}$ with (from *left* to *right*) $n = 1, 2, 3, 4$. Experiment and calculations all fall well below the *dashed curve*, which shows the corresponding result from the Tomlinson model, i.e. the friction force in absence of thermal excitations. After [34]

order of the eigenfrequency of the employed Tribolever sensor. Although this may not seem very surprising, further work will be necessary to resolve why the much higher vibration frequencies of the apex of the tip seem not to dominate the value of r_0.

8.4 Other Manifestations of Superlubricity and Thermolubricity

8.4.1 Lubrication by Graphite and Other Lamellar Solids

Graphite is a popular solid lubricant that is usually applied in the form of flaky powder. Several other layered materials also show excellent lubricating properties, for example MoS_2 and Ti_3SiC_2. Traditionally, the good lubrication by these materials is ascribed to the weak interaction between adjacent layers [35]. However, the intuitive idea that this would lead to easy shear cannot be correct, since it would still require the simultaneous rupture of all bonds in a plane, which involves a tremendously high energy, even for the weakly interacting layers in graphite [36]. Based on the lateral force measurements discussed in this chapter it seems natural to add the extra element of superlubricity and possibly also that of thermolubricity to the low-friction scenario of graphite. As argued above, the easy shear cannot take place *within* individual pieces of graphite, but it can occur *between* flakes of graphite since in a lubrication film flakes will be oriented randomly with respect to each other, which introduces the mismatch required for superlubricity and thermolubricity for

almost all graphite-graphite contacts. It is known that when metals are lubricated by graphite under ambient conditions, the oxide layers on the metals tend to fix some of the graphite to the metal, leading to a smooth transfer film [37], which would indeed concentrate most of the shear motion within the film of graphite flakes, rather than between the graphite and the metal surfaces. In a practical sliding geometry, the normal and shear forces will not be carried by a single, macroscopic flake-flake contact but it will be distributed over a large ensemble of simultaneous microcontacts between flakes, most of which will be slippery due to superlubricity or thermolubricity, while only a small fraction will be in temporary registry.

Strong support for the suspicion that it is the slipperiness of misoriented flakes that makes the friction low when macroscopic contacts are lubricated by layered solids comes from transmission electron microscopy observations by Martin et al. on MoS_2 [20]. TEM inspection of wear particles harvested from a MoS_2 lubrication film that had been exposed to sliding friction revealed Moiré patterns characteristic for superimposed flakes that were rotated with respect to each other around their c-axis.

8.4.2 Lubrication by Diamond-Like Carbon and Related Coatings

We also briefly speculate about the extremely good lubricating properties of diamond-like carbon (DLC) coatings and related, carbon containing materials and propose that it is again the superlubricity and thermolubricity of graphite that might be responsible. During the running-in phase, some of the DLC coating may be transformed into graphite and remain loosely bound to the DLC film. A relatively small amount of graphitized material should be sufficient to decorate all asperities and thus dominate the shear response. The main role of the DLC film would thus be to provide the material (carbon) and the conditions (e.g. through its hardness) necessary to produce small amounts of graphite. An essential element of this scenario is that it is self-terminating. The high friction forces at the beginning of run-in provide local pressures and temperatures that should be high enough to shear off carbon from the DLC film and graphitize it. Once it has been formed, the graphite dramatically reduces friction, so that the local shear stresses on the DLC film are too low to continue wearing off the film and graphitizing it. It has indeed been demonstrated that under sliding conditions, a graphitized tribolayer is formed on top of diamond-like carbon (DLC) coatings, which goes hand in hand with the decrease of the friction coefficient during run-in [38].

8.4.3 Lubrication by Fullerenes and Carbon Nanotubes

Fullerenes show very low friction when they are oriented in an incommensurate fashion. Miura et al. [39] have constructed a molecular ball bearing by placing a

monolayer of C_{60} molecules between two graphite sheets of 1 mm^2. When they moved the upper graphite flake with an FFM tip, they observed that the forward and backward traces were identical and no energy dissipation was measurable up to normal loads of more than 100 nN. Surprisingly the lateral force traces still showed sawtooth-like force variations which are typical for instabilities and thus significant friction should be expected. Therefore the Prandtl-Tomlinson model apparently is not able to describe the low friction behavior of this system. Falvo et al. [40, 41] manipulated carbon nanotubes (CNTs) on a graphite surface using the tip of an FFM. They observed that the CNTs changed from sliding to rolling motion, depending on the orientation of the tubes on the substrate. The rolling motion of the CNTs in the case of a commensurate contact was found to require a higher lateral force than the sliding motion of the CNTs in the case of an incommensurate contact. Cumings and Zettl [42] have used a TEM to estimate the friction force between two tubes of a multiwall carbon nanotube (MWNT) in the direction of the long axis. They pulled the core tube out of the outer tube and calculated the friction force from the retraction time, which was below 1.5×10^{-5} nN/atom.

Nanoparticles form another interesting model system to study the friction between two finite surfaces. By deposition of Sb particles on graphite and MoS_2, Ritter et al. [43] have created incommensurate contacts of various sizes. At a certain size the particles undergo a transition from amorphous to crystalline and at that point an increase in friction was observed. Although their friction was low, the amorphous particles were found not to be completely superlubric in air [44]. For future applications in superlubric meso- or macroscopic mechanical systems it is very important to learn why superlubricity is not seen in this case.

8.5 Concluding Remarks

In this chapter we have reviewed friction force microscopy experiments on extremely low friction and their interpretation in terms of lattice mismatch effects (superlubricity) and effects due to thermal excitations (thermolubricity). The prospect of exploiting these effects in practical applications, for example in nano- and micro-electro-mechanical systems (NEMS and MEMS) is exciting. The examples of excellent lubrication by graphite and other layered materials and by modern, carbon containing coatings such as DLC strongly suggest that these effects can indeed manifest themselves on macroscopic length scales and under realistic loading conditions.

It is evident that many questions have not been addressed in the studies reviewed here on nanoscale contacts. Similarly, many aspects remain to be explored on the way from a single, nano-contact under modest loading pressures to a large ensemble of larger contacts with possibly higher loading. An example is the effect of the load-dependent elastic deformations that may be expected to undermine the slipperiness, as mentioned in Sect. 8.2.2.

Another important issue is whether we can evoke superlubricity- and thermolubricity-based slipperiness also on other materials than layered solids such as graphite and related materials. Simulations for copper sliding over copper [16] suggest that there is no fundamental reason against this possibility, which is therefore demanding to be examined.

Acknowledgments The authors are grateful to a large number of people for their valuable contributions to the work reviewed in this chapter. In particular we mention J.A. Heimberg for the design and construction of the friction force microscope, K.B. Jinesh and N. Pradeep for performing part of the experiments and analysis, S. Yu. Krylov for setting up the theory of thermolubricity and G.S. Verhoeven and H. Valk for numerical calculations of superlubricity and thermolubricity. The work presented in Sects. 8.2.1 and 8.3.2 is part of the research program of the "Stichting voor Fundamenteel Onderzoek der Materie (FOM)" and was made possible by financial support of the "Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO)".

References

1. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987)
2. S. Fujisawa et al., J. Vac. Sci. Technol. B **12**, 1635 (1994)
3. S. Morita, S. Fujisawa, Y. Sugawara, Surf. Sci. Rep. **23**, 1 (1996)
4. R. Bennewitz et al., Phys. Rev. B **60**, R11301 (1999)
5. G.J. Germann et al., J. Appl. Phys. **73**, 163 (1993)
6. R.J.A. van den Oetelaar, C.F.J. Flipse, Surf. Sci. Lett. **384**, L828 (1997)
7. R.W. Carpick, Q. Dai, D.F. Ogletree, M. Salmeron, Tribol. Lett. **5**, 91 (1998)
8. L. Howald et al., J. Vac. Sci. Technol. B **12**, 2227 (1994)
9. R. Lüthi et al., J. Vac. Sci. Technol. B **14**, 1280 (1996)
10. L. Prandtl, ZS f. angew. Math. u. Mech. **8**, 85 (1928)
11. G.A. Tomlinson, Phil. Mag. S.7 **7**, 905 (1929)
12. S. Aubry, The New Concept by Breaking of Analyticity in a Crystallographic Model, in *Solitons and Condensed Matter Physics*, ed. by A.R. Bishop, T. Schneider (Springer, Berlin, 1979), p. 264
13. S. Aubry, Phys. D **7**, 240 (1983)
14. K. Shinjo, M. Hirano, Surf. Sci. **283**, 473 (1993)
15. M. Hirano, K. Shinjo, Phys. Rev. B **41**, 11837 (1990)
16. M.R. Sørensen, K.W. Jacobsen, P. Stoltze, Phys. Rev. B **53**, 2101 (1996)
17. M.H. Müser, Europhys. Lett. **66**, 97 (2004)
18. M. Hirano, K. Shinjo, R. Kaneko, Y. Murata, Phys. Rev. Lett. **67**, 2642 (1991)
19. J.S. Ko, A.J. Gellman, Langmuir **16**, 8343 (2000)
20. J.M. Martin, C. Donnet, T. LeMogne, T. Epicier, Phys. Rev. B **48**, 10583 (1993)
21. M. Hirano, K. Shinjo, R. Kaneko, Y. Murata, Phys. Rev. Lett. **78**, 1448 (1997)
22. T. Zijlstra et al., Sens. Actuators A **84**, 18 (2000)
23. M. Dienwiebel et al., Rev. Sci. Instrum. **76**, 043704 (2005)
24. M. Dienwiebel et al., Phys. Rev. Lett. **92**, 126101 (2004)
25. M. Dienwiebel et al., Surf. Sci. **576**, 197 (2005)
26. G.S. Verhoeven, M. Dienwiebel, J.W.M. Frenken, Phys. Rev. B **70**, 165418 (2004)
27. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)
28. T. Baumberger, P. Berthoud, C. Caroli, Phys. Rev. B **60**, 3928 (1999)
29. C. Caroli, P. Nozières, in *Physics of Sliding Friction, Vol. 311 of NATO ASI Series E: Applied Sciences*, ed. by B.N.J. Persson, E. Tosatti (Kluwer, Dordrecht, 1996), p. 27
30. E. Gnecco et al., Phys. Rev. Lett. **84**, 1172 (2000)

31. Y. Sang, M. Dubé, M. Grant, Phys. Rev. Lett. **87**, 174301 (2001)
32. E. Riedo et al., Phys. Rev. Lett. **91**, 084502 (2003)
33. K.B. Jinesh, J.W.M. Frenken, Phys. Rev. Lett. **96**, 166103 (2006)
34. S.Y. Krylov et al., Phys. Rev. E **71**, 065101(R) (2005)
35. W. Bragg, *An Introduction to Crystal Analysis* (G.Bell and Sons Ltd., London, 1928)
36. G.I. Finch, Proc. Phys. Soc. A **63**, 785 (1950)
37. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Clarendon Press, Oxford, 1964)
38. Y. Liu, A. Erdemir, E.I. Meletis, Surf. Coat. Technol. **86–87**, 564 (1996)
39. K. Miura, S. Kamiya, N. Sasaki, Phys. Rev. Lett. **90**, 055509 (2003)
40. M.R. Falvo et al., Nature **397**, 236 (1999)
41. M.R. Falvo, J. Steele, R.M. Taylor II, R. Superfine, Phys. Rev. B **62**, R10665 (2000)
42. J. Cumings, A. Zettl, Science **289**, 602 (2000)
43. C. Ritter, M. Heyde, B. Stegemann, K. Rademann, Phys. Rev. B **71**, 085405 (2005)
44. C. Ritter, Private communication

Chapter 9
Friction and Wear of Mineral Surfaces in Liquid Environments

Carlos M. Pina, Carlos Pimentel and E. Gnecco

Abstract Lateral Force Microscopy (LFM) is a very suitable technique to investigate the structure and reactivity of mineral surfaces in liquids. Studies performed in the last two decades have shown that the dissolution and growth of mineral surfaces immersed in water and aqueous solutions can be monitored by recording friction signals with LFM. Moreover, the sensitivity of lateral forces to both structure and chemistry makes possible to use LFM to obtain information about monolayers formed on mineral faces. Finally, numerous mineral surfaces are excellent substrates on which nanoparticles and complex organic molecules can be deposited and subsequently imaged and manipulated. This opens the way to future applications in molecular electronics. This chapter presents an overview of the recent use of LFM in liquid to investigate mineral surfaces and processes occurring on them.

9.1 Introduction

Minerals have been revealed as an almost inexhaustible source of surfaces which can be studied with atomic force microscopy (AFM). As early as 1992, the (104) surface of the mineral calcite was chosen by several researchers to conduct first investigations of a relatively complex crystal face using AFM [1–3]. Since then, mineral surfaces have been extensively used to investigate the structure, reactivity (e.g. growth and

C.M. Pina (✉) · C. Pimentel
Departamento de Cristalografía y Mineralogía, Universidad Complutense de Madrid, Instituto de Geociencias (UCM-CSIC), c/José Antonio Novais, 2, 28040 Madrid, Spain
e-mail: cmpina@geo.ucm.es

C. Pimentel
e-mail: cpimentelguerra@geo.ucm.es

E. Gnecco
Instituto Madrileño de Estudios Avanzados en Nanociencia, IMDEA Nanociencia, Campus Universitario de Cantoblanco, 28049 Madrid, Spain
e-mail: enrico.gnecco@imdea.org

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_9

dissolution), mechanical properties and wearing of crystalline matter [4–6, and references therein]. Furthermore, numerous mineral surfaces are suitable substrates to perform nanomanipulation experiments in which overgrowths, nanoparticles and complex biomolecules are imaged, disrupted, removed or displaced [7–9]. However, to obtain reliable quantitative information from these experiments it is essential to measure and control both vertical (loading) and lateral (friction) forces between the tip of the AFM and the investigated surfaces. When an AFM is used to record friction forces as a function of loading forces it is usually called lateral force microscope (LFM) or friction force microscope (FFM) [10]. Since recorded friction forces are sensitive to both the composition and structure of the substrates and the bonds formed between them and the scanning tip, LFM can, in principle, provide information about the crystallochemistry of surfaces. This is why LFM is also sometimes named chemical force microscopy (CFM), despite quantitative chemical analysis of surfaces is still beyond the current capabilities of this technique. The main obstacle to extract reliable structural and chemical information of surfaces from recorded lateral forces is the high complexity of the interactions between the tip probe and the surfaces. When an AFM tip slides on a surface, measured friction forces are influenced by a high number of factors, such as the elastic properties of both tip and surface, the tip-surface contact area, the sliding velocity, the formation and breaking of (chemical) bonds between the tip apex and the atoms of the surface, and the capillarity and adhesion forces. To reduce the number of factors affecting frictional forces and to increase the reproducibility of friction data, researchers perform their measurements using AFM tips with well characterised sizes, geometries and elastic properties, and operating under controlled conditions. The optimum controlled conditions are ultra-high vacuum (UHV) and/or low temperatures. This decreases the concentration of impurities, eliminates the adhesion forces between tip and surfaces due to humidity, and reduces the thermal vibration of atoms, which negatively affects the quality of the friction data. These advantages are, however, compromised by the demanding technical requirements of LFM in UHV. Moreover, relevant studies on the reactivity and frictional properties of mineral surfaces usually require observational conditions similar to those found in natural environments and, among them, liquid environments (i.e. aqueous solutions) are doubtless the most important.

The nanotribological investigation of mineral surfaces in liquid environments constitutes a new research field. Both the measurement of friction and nanomanipulation of molecules, particles or overgrowths on minerals immersed in liquids (e.g. water, aqueous solutions, and organic liquids such as ethanol) are not only of interest because natural environments and processes can be reproduced but also because it has been shown that LFM microscopy in liquids is can be an alternative to other AFM techniques in UHV which provide images with comparable resolution.

In this chapter, we review recent nanotribological investigations of mineral surfaces immersed in liquids. This review will show that LFM operating in a liquid environment is a very suitable technique to obtain structural and, to some extent, chemical information of crystal surfaces. Furthermore, it will be shown that the friction maps acquired during the interaction of mineral surfaces with liquids can be used to study mineral reactivity. Finally, recent investigations demonstrate that

mineral surfaces can act as adequate substrates to deposit organic molecules. The subsequent nanomanipulation of such molecules can provide new insights into the processes that control the interaction between inorganic and organic matter.

9.2 Structural Studies of Mineral Surfaces Using Lateral Force Microscopy

Lateral force microscopy can be used to investigate structural details of surfaces which are impossible or extremely difficult to resolve using conventional AFM. For instance, it has been demonstrated that recorded friction signals are sensitive to the orientation of some ionic groups (e.g. sulphate and carbonate groups) protruding from cleavage mineral surfaces [11, 12]. On a number of surfaces of alkaline earth sulphate crystals, i.e. anhydrite (100), barite (001) and celestine (001), Shindo et al. [11] measured an alternation of friction values in successive crystal monolayers. While the friction contrast can be attributed to the relative orientation of the S-O bonds of sulphate groups with respect to the scan direction, its alternation is due to the reversal of such an orientation by the operation of two-fold screw axes perpendicular to the studied mineral surfaces, which rotate 180° the position of the sulphate groups in successive monolayers (see Fig. 9.1). In the case of the three sulphates investigated, the higher friction was found when the monolayers were scanned against the tilt of the sulphate groups.

Kwak and Shindo [12] also studied the frictional asymmetry due to the tilt of carbonate groups on calcite {104} surfaces [12]. To this end, these authors scanned calcite (104) surfaces of two crystals oriented opposite. Although again different friction values were found depending on the scan direction, a lower friction was

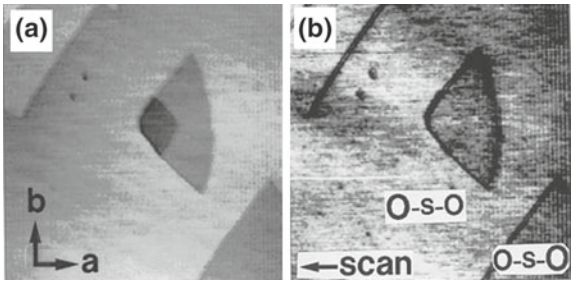


Fig. 9.1 **a** Topography AFM image of a celestine (001) surface showing a triangular etch pit with a rhombus-shaped pit inside. The depth of both etch pits is 0.34 nm, i.e half a unit cell each. **b** LFM image of the same area showing the reversal of friction in successive monolayers. Larger "O" symbols indicate the position of the protruding oxygens belonging to the sulphate groups. With the set up used, dark areas indicate higher friction. Scan area: $1.46 \times 1.46\,\mu m^2$. Both images were taken with a scan direction from the *right* to the *left*. Reprinted (adapted) with permission from Shindo et al. [11]. Evidence of the contribution of molecular orientations on the surface force friction of alkaline earth sulfate crystals. Phys. Chem. Chem. Phys. 1, 1597–1600 Copyright 1999 PCCP

measured when scanning against the tilt of the carbonate groups. This result contrasts with the tribological behaviour previously observed on sulphates. An explanation for such a discrepancy can be found in the different tilt angles of the C-O and S-O bonds with respect to the surfaces, which result in different elastic responses when the tip is pushing down the CO_3 and SO_4 groups during the scan. Although a conclusive explanation for the reversed friction anisotropy observed on faces of alkali earth sulphates and calcite requires further experimental and modelling work, the current observations clearly show that the interpretation of friction contrast in terms of molecular orientations is not straightforward.

The effects of molecular orientations on friction were observed by Shindo and collaborators on mineral surfaces exposed to air. Nanotribological studies of mineral surfaces immersed in water have provided similar results. For example, *in situ* observations of crystal growth on anhydrite (100) surfaces from aqueous solutions have shown that friction contrast reverses in successive monolayers during growth [13] (see Fig. 9.2). Such a friction contrast reversal can be again explained as due to the alternate orientation of sulphate groups within elementary growth layers related by two-fold screw axes.

LFM imaging in water is also useful to obtain information of slight structural modifications of mineral surfaces like calcite (104) and dolomite (104) faces. In first high resolution AFM studies of calcite (104) surfaces conducted in water, a rectangular lattice consistent with the calcite bulk structure was reported [1, 2, 14–16]. In addition, AFM images showed two peculiarities in the termination of the calcite (104) face: (i) the existence of rows of atoms parallel to the [010] direction which alternate in height and they are paired, i.e. the so-called "row-pairing"; and (ii) the height modulation along the direction [010] with a periodicity of about 1 nm which results in a (2×1) reconstruction of the calcite (104) surface [15, 16]. Both the row-pairing and (2×1) reconstructions have been confirmed recently by non-contact AFM in water and in UHV [4, 17, and references therein].

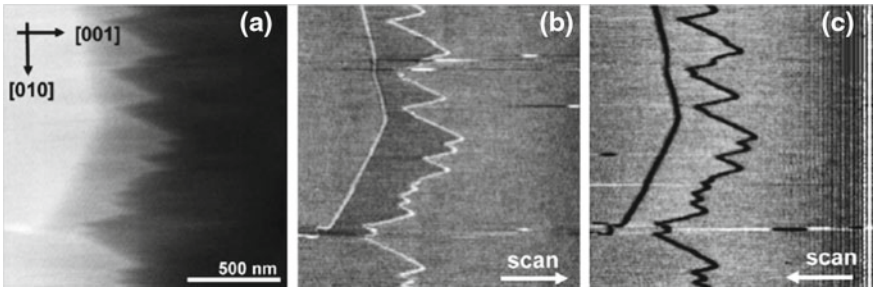


Fig. 9.2 **a** Topography image showing the growth of a monolayer on an anhydrite surface. **b**, **c** LFM images of the same area taken with opposite scan directions and showing the reversal of friction contrast. Reprinted (adapted) with permission from Pina [13]. Nanoscale dissolution and growth on anhydrite cleavage faces. Geochimica et Cosmochimica Acta 73, 7034–7044. Copyright 2009 Elsevier

High resolution LFM images of the (104) faces of both calcite and the structurally-related mineral dolomite, $CaMg(CO_3)_2$, have provided further information about their surface structures and reconstruction phenomena described above. Friction maps of calcite and dolomite (104) surfaces in water show two distinct friction peaks in each rectangular surface unit cell, which can be univocally attributed to the interaction of the AFM tip with the protruding oxygens of the triangular carbonate groups (see Fig. 9.3).

A more detailed analysis of the friction forces during the scan of calcite and dolomite (104) demonstrated that they depend on the scan direction, i.e. there is a clear anisotropy of friction [18]. This is evidenced by the weakening of the intensity of one of the friction peaks within the dolomite and calcite surface unit cells when the scan direction approaches the $[0\bar{1}0]$ crystallographic direction (see Fig. 9.4). In addition, high resolution LFM images of dolomite and calcite (104) surfaces in water confirmed the row-pairing previously observed using other AFM modes (i.e. topography and deflection). By changing the scan direction LFM images also show for the first time an inversion of the row-pairing. However, up to date, high resolution LFM images of both dolomite and calcite (104) surfaces do not show evidences of the (2×1) reconstruction.

The reported atomic scale friction maps of dolomite and calcite (104) surfaces, as well as their modifications with the scan direction, have been adequately reproduced by numeric calculations. Lattice resolved friction force maps like those shown in Figs. 9.3 and 9.4 can be understood using the Prandtl-Tomlinson model [18]. Here, specific chemical groups in the unit cell of the crystal surface act as pinning centers for the AFM tip. In this case, the lateral force exerted on the tip increases linearly

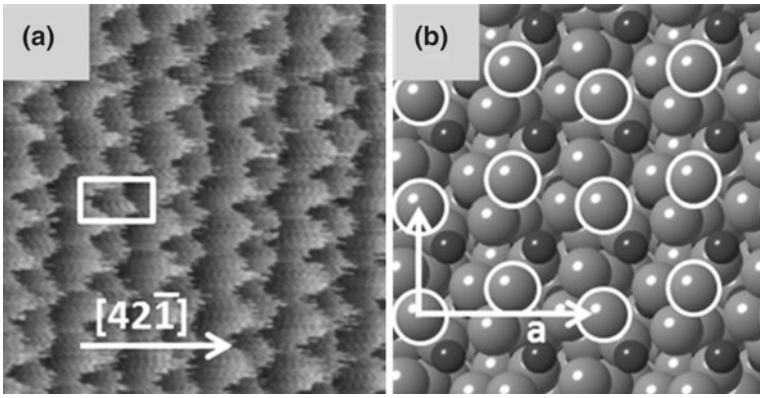


Fig. 9.3 **a** Friction map of a calcite (104) face showing two friction peaks per unit cell (marked by a white rectangle). **b** Projection of the calcite structure on the (104) plane. Oxygen atoms are represented by *light grey spheres*, carbon atoms by *black spheres*, and calcium atoms by *dark grey* spheres. Carbon atoms are not visible in this projection. Protruding oxygens atoms from the surface are enhanced with a *white* rim. In the dolomite structure, half of the Ca atoms are replaced by Mg along the $[42\bar{1}]$ direction, the surface structure being essentially identical. Reprinted (adapted) with permission from Pina et al. [18]. Anisotropic coupling while sliding on dolomite and calcite crystals. Physical Review B, 85, 073402

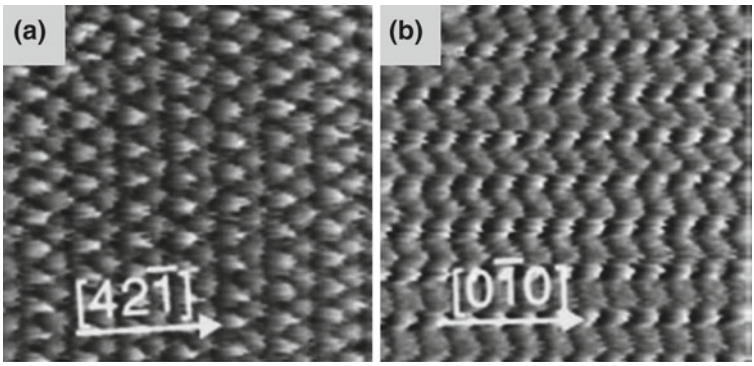


Fig. 9.4 Friction maps of calcite (104) surface when scanning almost parallel to (**a**) the [42$\bar{1}$] direction and the, (**b**) the [0$\bar{1}$0] direction. Reprinted (adapted) with permission from Pina et al. [18]. Anisotropic coupling while sliding on dolomite and calcite crystals. Physical Review B, 85, 073402

when the tip is pulled laterally away from the pinning center, as if an elastic spring was present between tip and surface. When the elongation of the spring (and the corresponding lateral force) reaches a critical value, pinning is no more sustainable and the tip suddenly jumps into a new pinning site defined by the scan direction. In this context, we should notice that, unless isolated defects or consecutive atoms below and above a step edges are imaged, it is not correct to speak about true molecular resolution. The number of atoms forming the contact area can indeed be in the order of some tens, so that the contrast seen in the friction force maps is ultimately due to the convolution of the real crystal structure with the atomic arrangement at the tip apex.

9.3 Obtaining Chemical Information of Surfaces from Frictional Forces

On atomically flat surfaces, frictional contrast can be related to compositional contrast [19]. This enables the *in situ* observation of adsorption and/or solid solution formation on mineral surfaces from multicomponent aqueous solutions, which is fundamental to better understand a number of mechanisms that control the removal of some contaminants from natural and industrial waters. The first systematic nanotribological study of the interaction of dissolved metal ions with a mineral surface was conducted by Hay et al. [19]. These authors observed with LFM the reactions between calcite (104) surfaces and aqueous solutions containing Ca^{2+}, Sr^{2+} and La^{3+}. Such reactions led to the formation of monolayer overgrowths which were identified by clear differences in the recorded friction signals (Fig. 9.5). In the three cases, newly-grown layers display higher friction contrast than the calcite substrate, which was attributed to epitaxial strain. Differently, epitaxial monolayers of celestine and anglesite grown on barite (001) have been found to show lower friction than the substrate [20].

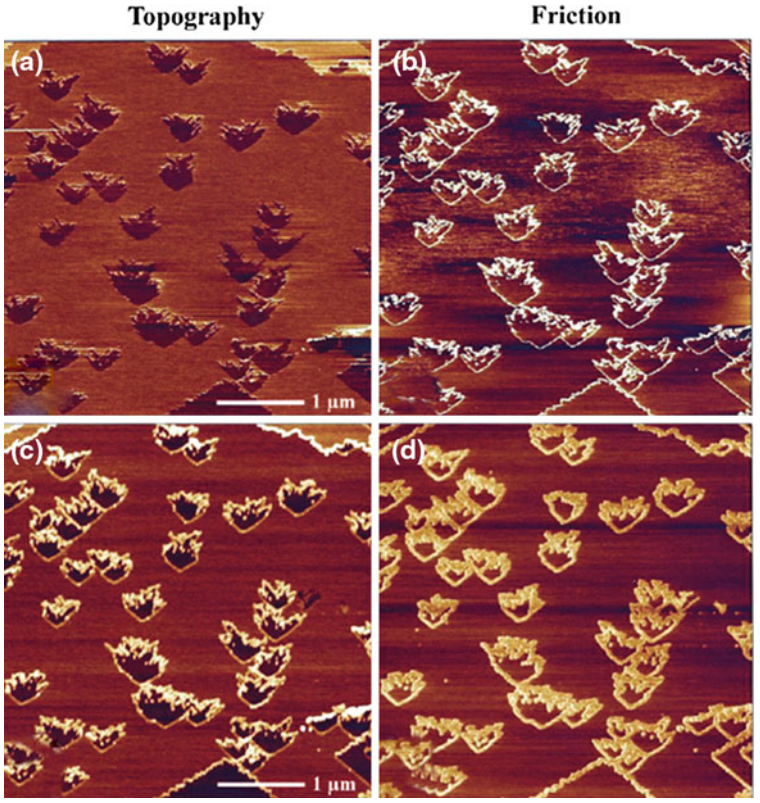


Fig. 9.5 Topography and friction images of a growing calcite (104) surface in the presence of La^{3+}. **a** and **b** after 150 s of reaction. **c** and **d** after 1330 s of reaction. Higher contrast areas in the friction images correspond to the places where a La-carbonate nucleated. Reprinted (adapted) with permission from Hay et al. [19]. Mechanisms of metal ion sorption on calcite: composition mapping by lateral force microscopy. Langmuir, 19, 3727–3740. Copyright 2003 American Chemical Society

These somehow contradictory experimental results indicate that the origin of friction when epitaxial layers are formed is not clear and further experimental and computational work is required.

Despite the current limitations of the use of LFM to detect chemical changes on mineral surfaces and overgrowths, the study of friction have already provided semi-quantitative information about the up-take of dissolved metals by common mineral surfaces and its effect on the nanotribological properties of the overgrowths. An interesting case study, the formation of Cd-bearing carbonate layers on calcite, was recently investigated with LFM [21]. Systematic measurements of friction versus applied loading force were conducted both on calcite and Cd-bearing carbonate layers and successfully fitted to the Johnson-Kendall-Roberts model, which accounts for adhesion in single asperity contacts [22] (see Fig. 9.6a). According to this model, differences in friction between calcite substrate and Cd-bearing layers have a different origin depending on the loading force. At low loading forces, differences in friction

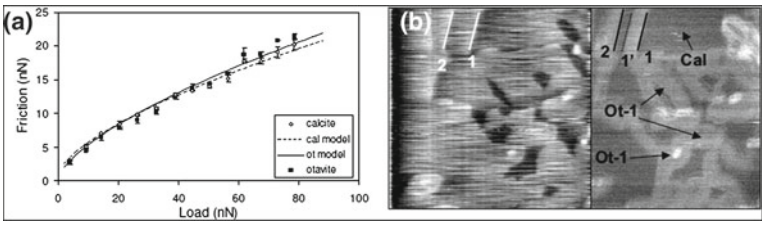


Fig. 9.6 **a** Friction versus loading force measured on calcite substrate and otavite overgrowths. **b** Overgrowth of otavite (Ot-1) on calcite (104) surface (Cal). *Left* topography image; *Right* friction image. Modified from Cubillas and Higgins [21]. Friction characteristics of Cd-rich carbonate films on calcite surfaces: implications for compositional differentiation at the nanometer scale. Geochemical Transactions 10.7. doi:1186/1467-4866-10-7

are determined by adhesion forces. Differently, at high loading forces, friction forces are mainly sensitive to differences in contact shear strengths. As a result, while at low loading forces (up to 40 nN), Cd-bearing layers have lower friction than the calcite substrate, at loading forces higher than approximately 40 nN, friction measured on calcite substrates is lower than that on Cd-bearing layers (see Fig. 9.6b).

The results summarized above, although somehow disappointing, clearly demonstrate that future quantifications of chemical variations on mineral surfaces using LFM will require a precise knowledge of the regime (i.e. adhesion-controlled or shear strength-controlled) under which friction is measured.

9.4 Wear and Nanomanipulation of Mineral Surfaces and Overgrowths

AFM can be used not only to image but also to produce wear and reorganisation of mineral surfaces at the nanoscale under different environmental conditions including aqueous solutions. In addition, overgrowths formed on mineral surfaces can be eroded and/or manipulated using the tip of the AFM. Such wear and nanomanipulation experiments are conducted by controlling the loading forces of the AFM tip and they provide information about the mechanical properties and stability of both mineral surfaces and mineral overgrowths. Furthermore, nanowear experiments allow one to study phenomena such as the propagation of ductile and brittle deformation, stress-enhanced crystal growth and dissolution and surface pattern formation (e.g. ripple initiation and propagation [23]).

Nanowear experiments conducted on calcite ($CaCO_3$) {104} faces and brushite ($CaHPO_4-2H_2O$) {010} faces immersed in undersaturated aqueous solutions have demonstrated that when vertical forces are increased dissolution behaviour is strongly modified [24, 25]. On these faces, it was observed that moderate to high loading forces produce relatively atomically flat surfaces. This surface planarization, also observed on dolomite (104) and gypsum ($CaSO_4 \cdot 2H_2O$) (010) surfaces, can be explained by a combination of enhanced dissolution of pre-existent steps and etch pits edges with a tip assisted redeposition of dissolved material during scan. In contrast, redeposition

outside the scan areas lead to an increase in roughness, further indicating the strong effect of the AFM tip on the surface reorganisation.

When supersaturated solutions with respect to calcite are used, crystal growth is enhanced by the AFM tip along pre-existing steps, even at relatively low loading forces. Such a growth enhancement rapidly increases with supersaturation. The result of this tip-enhanced growth is the formation of defect-free calcite (104) surfaces at high velocities [26]. To promote such a high growth rates in the absence of scanning higher supersaturations with respect to calcite are required, which however do not lead to a rapid layer-by-layer material deposition but to a three-dimensional nucleation on calcite surfaces. As in the case of dissolution, tip-enhanced growth can be explained by the mobilisation of ions and clusters of ions during scan (see Fig. 9.7). In particular, the AFM tip seems to have the ability of moving growth units from calcite terraces to the proximity of step edges where edge diffusion occurs. Then growth units are eventually incorporated into the calcite structure.

The tip of the AFM can be used not only to modify the dissolution and growth behaviour of mineral surfaces but also to remove overgrowths previously formed on them. When such nanomanipulation experiments are conducted with controlled loading forces and by recording friction forces, quantitative information on the stability of overgrowths on mineral substrates can be obtained. Two interesting model examples for studying the stability of overgrowths on mineral surfaces are the formation of calcite islands on dolomite and the isostructural kutnohorite ($MnCa(CO_3)_2$) {104} surfaces [7]. At high supersaturations with respect to calcite, spontaneous growth of calcite three-dimensional islands on dolomite and kutnohorite can be observed. This growth is in both cases epitaxial and calcite {104} rhombohedra are deposited on the dolomite and kutnohorite substrates with an almost perfect parallelism of the $[42\bar{1}]$ and [010] crystallographic directions, i.e. the main directions lying on the (104) epitaxial contact plane (see Fig. 9.8).

Calcite islands grown on dolomite and kutnohorite (104) surfaces can be removed by increasing the loading force of the AFM. Figure 9.9a, b shows a typical event of island detachment in which a single island and an aggregate of two calcite crystals were removed from a dolomite (104) surface by the AFM tip. As can be seen in Fig. 9.9c the removal of each calcite island results in a sharp increase in the recorded friction force. From such frictional peaks, estimations of the shear strengths required to remove epitaxial overgrowths can be obtained using the simple formula $\tau = F_L/A$ where F_L is the lateral force measured at moment and A is the overgrowth-substrate contact area, which can be measured from topography images. First nanomanipulation experiments provided the following shear strengths: $\tau \approx 7$ MPa for calcite islands grown on dolomite (104) surface and $\tau \approx 140$ MPa for calcite islands grown on kutnohorite (104) surface. The significant difference in shear strengths can be partially explained by the differences in lattice misfits between calcite overgrowth and substrates (which are higher in the case of the dolomite (104) substrate).

Recent nanomanipulation experiments have been also addressed to remove nitratine ($NaNO_3$) islands grown on calcite (104) surfaces [8]. Nitratine and calcite are isostructural compounds and as in the case of calcite on dolomite and kutnohorite substrates, epitaxial growth occurs. Nanomanipulation experiments, similar to those

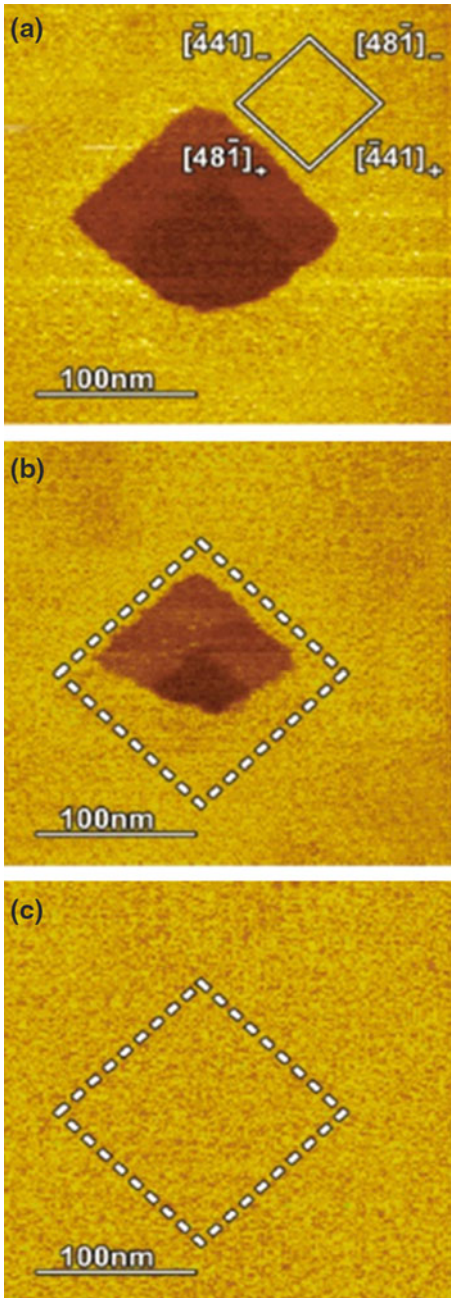


Fig. 9.7 Growth within an etch pit two monolayers deep on a calcite (104) surface. The growth occurred from a slightly supersaturated aqueous solution with respect to calcite and applying a tip loading force $F_N = 7$ nN. The crystallographic directions of the steps edges that define the etch pit are indicated by *solid lines*. The *dashed line* mark the initial shape of the etch pit. Reprinted (adapted) with permission from McEvoy et al. [26]. Scanning-Induced Growth on Single Crystal Calcite with an Atomic Force Microscope. Langmuir 22, 6931–6938. Copyright 2006 American Chemical Society

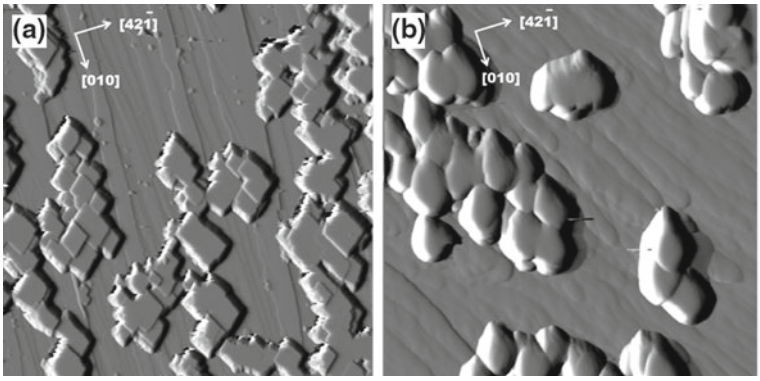


Fig. 9.8 AFM deflection images showing the epitaxial growth of calcite islands on the (10.4) surfaces of (**a**) dolomite (scan area: $14 \times 14\,\mu m^2$) and (**b**) kutnohorite (scan area: $5 \times 5\,\mu m^2$). Main crystallographic directions are indicated by white arrows. Reprinted (adapted) with permission from Pimentel et al. [7]. Epitaxial growth of calcite crystals on dolomite and kutnohorite (104) surfaces. Crystal Growth and Design 13, 2557–2563. Copyright 2013 American Chemical Society

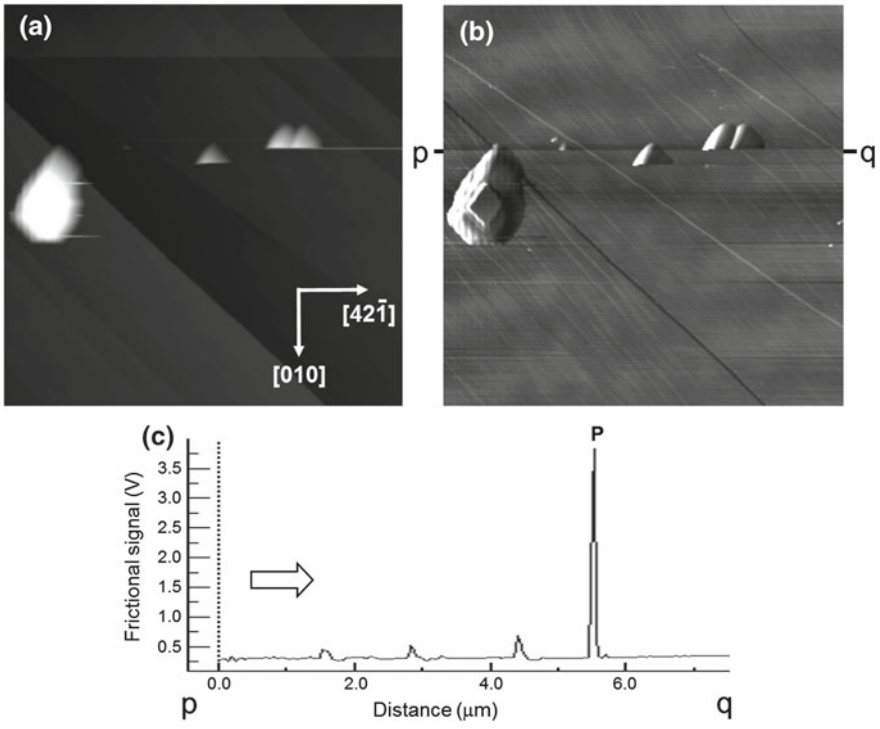


Fig. 9.9 Removal of calcite islands grown on a dolomite (104) surface. **a** Topography AFM image. **b** Friction AFM image of the same area. Both images are $7.5 \times 7.5\,\mu m^2$ in size. **c** Friction profile taken along the p-q line in (**b**). The frictional peak (*P*) corresponds to the removal of the aggregate of two calcite islands in (**b**). The *arrow* indicates the scan direction. Reprinted (adapted) with permission from Pimentel et al. [7]. Epitaxial growth of calcite crystals on dolomite and kutnohorite (104) surfaces. Crystal Growth and Design 13, 2557–2563. Copyright 2013 American Chemical Society

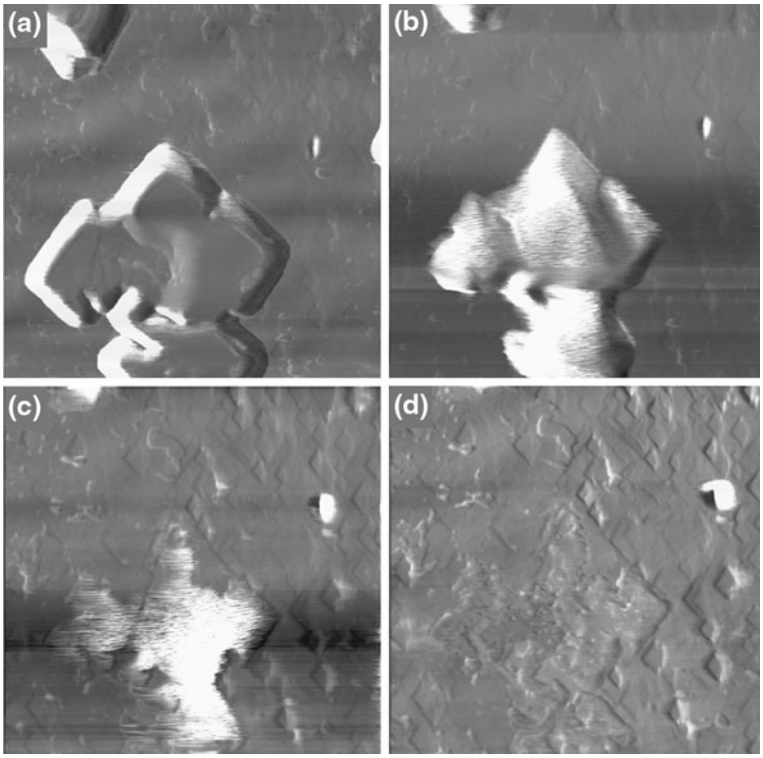


Fig. 9.10 Sequence of AFM friction images showing the nanowear of nitratine islands grown on a calcite (104) face. The applied normal force was $F_N \approx 80\,\mu$N. Reprinted (adapted) with permission from Benages-Vilau et al. [8]. Epitaxial crystal growth of nitratine on calcite (10.4) cleavage faces at nanoscale. Crystal Growth and Design 13, 5397–5403. Copyright 2013 American Chemical Society

conducted in the case of calcite islands on dolomite and kutnohorite (104) surfaces, provided estimates of shear strengths required to remove epitaxial overgrowths of about 30 MPa. However, due to the softness of nitratine islands (1.5–2 in the Mohs scale), the removal of nitratine islands is often accompanied by their erosion, which is also an example of nanowear under controlled conditions (Fig. 9.10).

9.5 Organic Molecules on Mineral Surfaces

Mineral surfaces provide a number of substrates on which both organic molecules can be deposited and subsequently nanomanipulated. This kind of experiments are of interest for various research and industrial fields such as the production of new composite materials, the studies of biomineralisation mechanisms and the design of new insulating and semiconductor materials. Insulating mineral surfaces are very

promising materials in the context of molecular electronics. Here, the minerals can be used as substrates for growing self-assembled structures of organic molecules, which need to be electrically decoupled from the surrounding environment. A thick insulating substrate makes the use of scanning tunneling microscopy as direct imaging technique, clearly impossible. On the other side, the AFM can still provide detailed information on the morphological structures of the organic layers so formed. The organic molecules are usually deposited onto the substrate in ultra-high vacuum, and observed by AFM in the same environment. For some recent examples we refer the reader to a recent review by Rahe et al. [27]. Nevertheless, an accurate resolution is also possible in different contexts, which better reproduce the conditions under which the molecular devices are supposed to be operated.

Figure 9.11 shows a topographic image of dolomite (104) covered by about 2 monolayers of copper phthalocyanine (CuPc) molecules [28]. Elongated molecular stripes with height of about 4.5 nm are seen all over the surface. The width of the stripes varies between 200 and 600 nm, and their length is about 1–3 μm. The stripes do not reveal any preferential orientation, meaning that the intermolecular interactions are much stronger than the molecule-substrate interaction.

Some features in the internal structure of the molecular stripes could be resolved by friction force microcopy, as shown in Fig. 9.12a. Here, a stacked intermolecular arrangement can be recognized. The stacks run along the axis of the corresponding stripe and, consequently, with no preferential orientation with respect to the substrate.

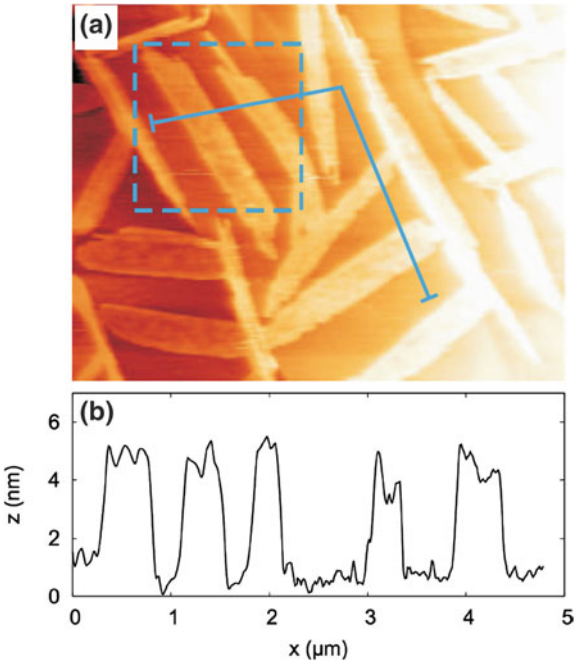


Fig. 9.11 **a** Contact mode AFM topography ($4.6 \times 6\,nm^2$) of the dolomite (104) surface covered by 2.0 monolayers of CuPc. The image was acquired at room temperature in water. **b** Topography profile taken along the *continuous line* in (**a**). Reprinted from Nita et al. [28]. Molecular resolution friction microscopy of Cu phthalocyanine thin films on dolomite (104) in water. Nanoscale 6, 8334–8339. Copyright 2014 The Royal Society of Chemistry

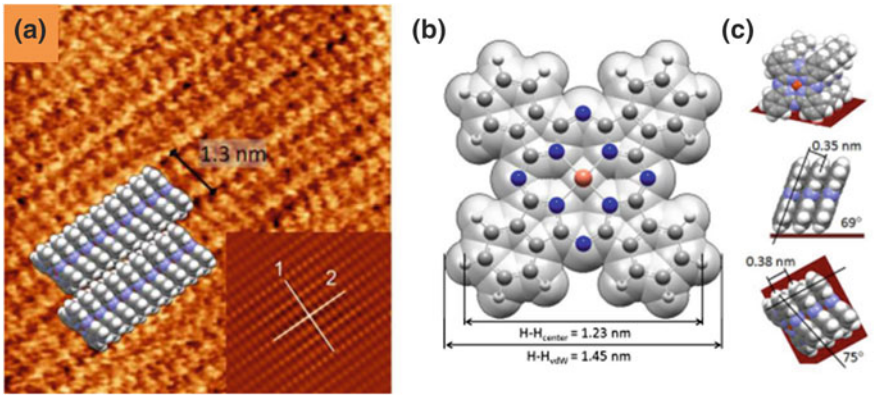


Fig. 9.12 **a** High resolution friction map ($10 \times 10\,\text{nm}^2$) on a CuPc stripe. The periodic arrangement of the molecules in parallel stacks, running along the axis of the stripe, is clearly visible. The average friction force is 5.5 nN (corresponding to a normal force $F_N = 2.7$ nN). Inset shows a 2D self-correlation of the friction map. **b** Chemical structure of a copper (II) phthalocyanine (CuPc) molecule. **c** Molecular stack geometry as determined from single crystal x-ray diffraction. Reprinted from Nita et al. [28]. Molecular resolution friction microscopy of Cu phthalocyanine thin films on dolomite (10.4) in water. Nanoscale 6, 8334–8339. Copyright 2014 The Royal Society of Chemistry

Two elongated spots per molecule are identified. Intermolecular structural parameters can be estimated from the 2D self-correlation analysis shown in the inset, which was performed along and perpendicular to the stacking direction. Along the stack, the periodicity is 0.36 nm, close to the reported periodicity from the single crystal x-ray value (0.38 nm) (Fig. 9.12b). The measured stack width is 1.3 nm, while the van der Waals (vdW) stack width calculated from single crystal data is $1.45\,\text{nm} \times \sin 75^{\text{o}} = 1.40\,\text{nm}$ (Fig. 9.12c). This indicates an interdigitated structure of neighboring CuPc stacks on dolomite just like in the single crystal which should effectively reduce the stack width by roughly the vdW radius of hydrogen (0.11 nm), resulting in a value of 1.29 nm, in perfect agreement with the measured value. Accordingly, the crystal structure stack (projected against the stacking plane) can be perfectly overlaid with the FFM image, as shown in Fig. 9.12a. In this way, the friction spots correspond to standing benzene rings. The good agreement of the stacking structure on dolomite with the single crystal data gives full evidence that substrate induction is weak on dolomite, and the growth of CuPc is essentially driven by self-assembly.

When the normal force increases, so do the pinning effect and the lateral force, and the resolution achieved via the stick-slip mechanism is enhanced. Nevertheless, this can also lead to irreversible damage of the molecular film. The image presented in Fig. 9.12a corresponds to a threshold value of $F_L = 5.5$ nN. If the same area is repeatedly scanned while keeping this value, the molecular layers are gradually worn off. This is seen in Fig. 9.13, where three 'snapshots' of the abrasive process at

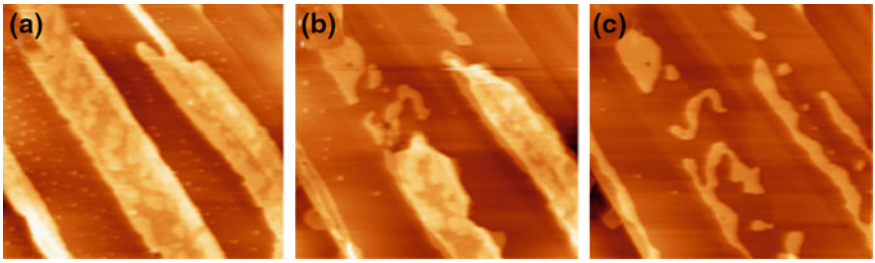


Fig. 9.13 Effect of prolonged scanning on the square region highlighted in Fig. 9.11. The topographic images **a**, **b** and **c** were acquired during the 3rd, 19th and 22nd scan back and forth. Reprinted from Nita et al. [28]. Molecular resolution friction microscopy of Cu phthalocyanine thin films on dolomite (104) in water. Nanoscale 6, 8334–8339. Copyright 2014 The Royal Society of Chemistry

different times are shown. Two layers are progressively removed till the step edges of the underlying dolomite (104) surface become clearly visible in Fig. 9.13c. Note that, before being worn off, the molecular strips were running across the step edges of the substrate in a carpet-like fashion. This gives a third indication of the weak interaction between CuPc and the dolomite surface.

9.6 Conclusions and Outlook

To summarize, LFM in water is a promising new technique for obtaining high (atomic-scale) resolution images of both mineral and synthetic crystalline surfaces. Recent results demonstrate that the resolution and quality of the images achieved is comparable to those recorded using more demanding AFM techniques such as dynamic modes in UHV. This may be attributed to the absence of capillary effects in water, which reduces the adhesion between tip and surface and minimizes the damage while scanning. Analysis of friction maps at the nanoscale allows one to distinguish growth monolayers with different chemical composition. However, quantitative chemical analysis is still beyond the current capabilities of LFM. Nanomanipulation experiments where epitaxial overgrowths are removed from mineral surfaces can provide estimates of shear strengths, related to the adhesion forces between the overgrowths and substrates. Last but not least, the crystal structure of organic molecules self-assembled on the mineral surface can be also identified by FFM in water. The weak interaction between molecule and substrate observed in recent investigations on CuPc grown on dolomite (104) opens the way to interesting applications in molecular electronics.

Acknowledgments This work was supported by the Spanish Ministerio de Economía y Competitividad (MINECO; project MAT2012-34487). Carlos Pimentel is grateful to Spanish Ministry of Education, Culture and Sports for a FPU grant. We thank Dr. Pawel Nita, Dr. Luo, Dr. Milián-Medina and Dr. Johannes Gierschner for their key collaboration on LFM of organic molecules on mineral surfaces.

References

1. P.E. Hillner, S. Manne, A.J. Gratz, P.K. Hansma, Ultramicroscopy **42–44**, 1387 (1992). doi:10.1016/0304-3991(92)90432-J
2. P.E. Hillner, A.J. Gratz, S. Manne, P.K. Hansma, Geology **20**, 359 (1992). doi:10.1130/0091-7613(1992)020(lt)0359:ASIOCG(gt)2.3.CO;2
3. F. Ohnesorge, G. Binnig, Science **260**, 1451 (1993). doi:10.1126/science.260.5113.1451
4. S. Rode, N. Oyabu, K. Kobayashi, H. Yamada, A. Kühnle, Langmuir **25**, 2850 (2009). doi:10.1021/la803448v
5. I. Sethmann, J. Wang, U. Becker, A. Putnis, Cryst. Growth Des. **10**, 4319 (2010). doi:10.1021/cg100202h
6. C.M. Pina, G. Jordan, in *Nanoscopic Approaches in Earth and Planetary Sciences*, ed. by F. Brenker, G. Jordan. EMU—Notes in Mineralogy (2006)
7. C. Pimentel, C.M. Pina, E. Gnecco, Cryst. Growth Des. **13**, 2557 (2013). doi:10.1021/cg400315g
8. R. Benages-Vilau, T. Calvet, M.A. Cuevas-Diarte, C. Pimentel, C.M. Pina, Cryst. Growth Des. **13**, 5397 (2013). doi:10.1021/cg401327h
9. S. Mangold, K. Harneit, T. Rohwerder, G. Claus, W. Sand, Appl. Environ. Microbiol. **74**, 410 (2008). doi:10.1128/AEM.01812-07
10. E. Gnecco, R. Bennewitz, O. Pfeiffer, A. Socoliuc, E. Meyer, in *Springer Handbook of Nanotechnology*, ed. by B. Bhushan (Springer, Berlin, 2007)
11. H. Shindo, K. Shitagami, T. Sugai, S. Kondo, Phys. Chem.. Chem. Phys. **1**, 1597 (1999). doi:10.1039/A808691E
12. M. Kwak, H. Shindo, Phys. Chem. Chem. Phys. **6**, 129 (2004). doi:10.1039/B310991G
13. C.M. Pina, Geochimica et Cosmochimica Acta **73**, 7034 (2009). doi:10.1016/j.gca.2009.08.032
14. A.L. Rachlin, G.S. Henderson, M.C. Goh, Am. Mineral. **77**, 904 (1992)
15. S.L.S. Stipp, C.M. Eggleston, B.S. Nielsen, Geochim. et Cosmochim. Acta **58**, 3023 (1994). doi:10.1016/0016-7037(94)90176-7
16. S.L.S. Stipp, Geochim. et Cosmochim. Acta **63**, 3121 (1999). doi:10.1016/S0016-7037(99)00239-2
17. P. Rahe, J. Schütte, A. Kühnle, J. Phys.: Condens. Matter **24**, 084006 (2012). doi:10.1088/0953-8984/24/8/084006
18. C.M. Pina, R. Miranda, E. Gnecco, Phys. Rev. B **85**, 073402 (2012). doi:10.1103/PhysRevB.85.073402
19. M.B. Hay, R.K. Workman, S. Manne, Langmuir **19**, 3727 (2003). doi:10.1021/la020647s
20. A.E. Murdaugh, M. Liddelow, A.M. Schmidt, S. Manne, Langmuir **23**, 5852 (2007). doi:10.1021/la063548d
21. P. Cubillas, S.R. Higgins, Geochem. Trans. **10**, 7 (2009). doi:10.1186/1467-4866-10-7
22. K.L. Johnson, K. Kendall, A.D. Roberts, Proc. R. Soc. Lond. A **324**, 301 (1971)
23. A. Socoliuc, E. Gnecco, R. Bennewitz, E. Meyer, Phys. Rev. B **68**, 115416 (2003). doi:10.1103/PhysRevB.68.115416
24. N.-S. Park, M.-W. Kim, S.C. Langford, J.T. Dickinson, J. Appl. Phys. **80**, 2680 (1996). doi:10.1063/1.363185

25. L. Scudiero, S.C. Langford, J.T. Dickinson, Tribol. Lett. **6**, 41 (1999). doi:10.1023/A:1019134901387
26. A.L. McEvoy, F. Stevens, S.C. Langford, J.T. Dickinson, Langmuir **22**, 6931 (2006). doi:10.1021/la0608359
27. P. Rahe, M. Kittelmann, J.L. Neff, M. Nimmrich, M. Reichling, P. Maass, A. Kühnle, Adv. Mater. **25**, 3948 (2013). doi:10.1002/adma.201300604
28. P. Nita, C. Pimentel, F. Luo, B. Milián-Medina, J. Gierschner, C.M. Pina, E. Gnecco, Nanoscale **6**, 8334–8339 (2014). doi:10.1039/C4NR01522C

Chapter 10
Nanotribology: Nonlinear Mechanisms of Friction

N. Manini, Oleg M. Braun and A. Vanossi

Abstract Friction with its related nonlinear dynamics is a vast interdisciplinary field, involving complex physical processes over a wide range of length and time scales. The accelerated progress in experimental and computational techniques, often leading to complex detailed dynamical patterns, has vigorously stimulated the search and implementation of idealized experimental frameworks and simpler mathematical models, capable of describing and interpreting, in a more immediate way, the essential physics involved in nonlinear sliding phenomena.

10.1 Introduction

Frictional motion plays a central role in diverse systems and phenomena that span vast ranges of scales, from the nanometer contacts inherent in micro- and nanomachines [1] and biological molecular motors [2] to the geophysical scales characteristic of earthquakes [3]. Due to its enormous practical importance, the problem has stimulated progress over the centuries. Historical figures from Leonardo da Vinci onward have brought friction into the field of physics, with the formulation of time-honored phenomenological frictional laws, which have been referred to as the Coulomb-Amontons laws. These statements can be summarized as follows: (i) frictional force

N. Manini (✉)
Dipartimento di Fisica, Università degli Studi di Milano,
Via Celoria 16, 20133 Milano, Italy
e-mail: nicola.manini@mi.infm.it

O.M. Braun
Institute of Physics, National Academy of Sciences of Ukraine,
46 Science Avenue, Kiev 03028, Ukraine
e-mail: obraun.gm@gmail.com

A.Vanossi
CNR-IOM Democritos National Simulation Center,
Via Bonomea 265, 34136 Trieste, Italy

A.Vanossi
International School for Advanced Studies (SISSA),
Via Bonomea 265, 34136 Trieste, Italy
e-mail: vanossi@sissa.it

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_10

is independent of the apparent area of contact; (ii) frictional force is proportional to the normal load; (iii) kinetic friction (the force to keep relative motion at constant speed) does not depend on the sliding velocity and is smaller than static friction (the force needed to initiate motion between two contacting bodies at rest). Serious attempts were made in the first half of the 20th century toward a microscopic understanding of these laws [4]. Whereas the basic physics underlying sliding friction—non equilibrium statistical mechanics of solids, sheared fluids, and moving surfaces—is in principle quite exciting, the field as a whole has (even if with notable exceptions) failed to attract adequate interest by the physicist until the last few decades, mainly because of a lack of microscopic experimental data.

Three quiet revolutions, of broad nature and unrelated to friction, are radically changing this state of affairs. First, progress in the general area of complexity provided new tools to tackle non-equilibrium disordered systems with interacting degrees of freedom. Second, and crucial, the developments in nanotechnology extended the study of friction and permitted its analysis on well-characterized materials and surfaces at the nano and microscale [5]. Notably the invention of scanning tip instruments of the Atomic Force Microscope (AFM) family [6] has opened *nanofriction* as a brand new avenue, the use of the Surface Force Apparatus (SFA) [7] has led to the systematic studies of confined mesoscopic systems under shear, and the Quartz Crystal Microbalance (QCM) [8, 9] has allowed us to measure the inertial sliding friction of adsorbate submonolayers. Thanks to these methods, a mass of fresh data and information on well defined systems has accumulated in the last two decades. Third, computer simulations have had a strong boost, also allowed by the fantastic growth of computer power. The numerical study of frictional models on one hand, and direct atomistic molecular dynamics (MD) simulations on the other hand, are jointly advancing our theoretical understanding [10]. Invaluable initial reviews of the progress in our understanding of sliding friction can be found in the books [11, 12].

Despite the importance of friction and the growing efforts in the field, many key aspects of friction dynamics are not yet fully understood. Fundamental theory is still difficult in all fields of sliding friction, including nanofriction, since the sliding motion generally involves sudden nonlinear stick-slip events, that cannot be treated within traditional theoretical approaches such as linear-response theory and hydrodynamics. Experiments in tribology have long suffered from the inability to directly observe what takes place at the sliding interface. Although AFM, SFA and QCM techniques have identified many friction phenomena on the nanoscale, many interpretative pitfalls still result from indirect or ex-situ characterization of contact surfaces. In the present chapter, we will briefly cover some aspects, progress, and problems in the current modeling and simulation of sliding friction, from nano to mesoscale.

One of the main difficulties in understanding and predicting frictional response is the intrinsic complexity of highly nonlinear and non-equilibrium processes going on in any tribological contact, which include detachment and reattachment of multiple microscopic junctions (bonds) between the surfaces in relative motion while still in contact [1, 2, 13]. Therefore friction is intimately related to instabilities that occur on a local microscopic scale, inducing an occasional fast motion of the corresponding

degrees of freedom even if the slider's center-of-mass velocity is extremely small. Understanding the physical nature of these instabilities is crucial for the elucidation of the mechanism of friction, as we will emphasize below.

The present chapter covers the following different types of theoretical approach to sliding friction: "minimalistic" models (MMs) described in Sects. 10.2 and 10.3, atomistic MD simulations considered in Sect. 10.4, and mesoscopic earthquake-like (multicontact) models briefly discussed in Sect. 10.5 (phenomenological rate-state models will not be considered here; this topic is covered, e.g., in [14]).

10.2 The Prandtl-Tomlinson Model

The Prandtl-Tomlinson (PT) model [16, 17] is the most successful and influential MM so far suggested for description of nanoscale friction. In particular, it addresses friction force microscopy (FFM) where friction forces are measured by dragging an AFM tip along a surface. Qualitative conclusions drawn with this model provide guidance to understand friction at the nanoscale, that often retain their validity in more advanced models and MD simulations.

PT assumes that a point-like mass m (e.g., mimicking the AFM tip) is dragged over a one-dimensional (1D) sinusoidal potential representing the interaction between the tip and a crystalline substrate. The point-like tip is pulled by a spring of effective elastic constant K, extending between the tip position x and the position of the microscope support stage, that is driven with a constant velocity v relative to the substrate, see Fig. 10.1a. Thus, the total potential experienced by the tip consists of two parts, the tip-substrate interaction and the elastic interaction between the tip and the support, and can be written as

$$U(x,t) = U_0 \cos\left(\frac{2\pi}{a}x\right) + \frac{K}{2}(x - vt)^2 , \tag{10.1}$$

where $2U_0$ is the amplitude and a is the period of the tip-substrate potential. Note that in an AFM experiment the actual "spring constant" mimicked by K in the PT model is not only due to the torsional stiffness of the cantilever but includes also the contribution from the lateral stiffness of the contact. There is no attempt in the model to describe realistically the energy dissipation into the substrate—all dissipation is described by a viscous force $-m\gamma\dot{x}$, where γ is a damping coefficient. The instantaneous lateral friction force measured in FFM experiments reads $F = -K(x - vt)$, and the kinetic friction F_k is the time average of F.

The PT model predicts two different modes for the tip motion, depending on the dimensionless parameter $\eta = 4\pi^2 U_0/(Ka^2)$, which represents the ratio between the stiffness of the tip-substrate potential and that of the pulling spring. When $\eta < 1$, the total potential $U(x)$ exhibits only one minimum and the time-dependent sliding

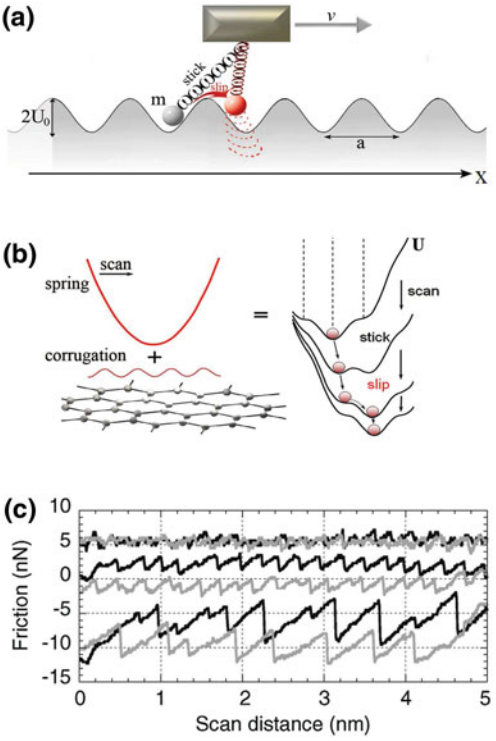


Fig. 10.1 **a** The Prandtl-Tomlinson model; **b** Energy landscape for a soft spring (low K). The total potential (harmonic spring + sinusoidal substrate) exhibits different metastable minima, giving rise to the stick-slip behavior. **c** A representative experimental friction pattern, for increasing load. Lateral AFM force vs position traces demonstrate transitions from smooth sliding (*top*) to single (*middle*) and mostly double slips (*bottom*). (Reproduced from [15]). Similar patterns can be generated within the PT model. (Adapted from [10])

motion is smooth; for $\eta > 1$ two or more minima appear in $U(x)$, and the sliding is discontinuous, characterized by stick-slip, Fig. 10.1b. The value $\eta = 1$ represents the transition from smooth sliding to slips by one lattice site (single-slip regime).

Physically, stick-slip motion corresponds to jumps of the tip between successive minima of $U(x)$, due to elastic instabilities induced by the driving spring ($\partial U/\partial x = 0,\ \partial^2 U/\partial x^2 = 0$). Close to the inflection point the height of the barrier preventing the tip sliding decreases with increasing applied force as $\Delta E \propto (\text{const} - F)^{3/2}$ [18–20]. This type of externally induced topological change in the free energy landscape is known as a fold catastrophe, and it has been found in many driven systems, including superconducting quantum interference devices [21, 22], mechanically deformed glasses [23], and stretched proteins [24, 25]. The simulation results obtained for diverse systems show that the fold catastrophe scaling is accurate not only in the immediate vicinity of the inflection point but over reasonably large intervals of loads.

The possibility of slips of higher multiplicity (multiple-slip regime) occurs for larger values of $\eta > 4.604$ [15]. However, this is the necessary but not sufficient condition to observe multiple slips, since the dynamics depends also on the damping coefficient γ. In particular, one can distinguish between the overdamped regime of motion, $\gamma > (\eta K/m)^{1/2}$, where the tip jumps between nearest-neighbor minima of the potential, and the underdamped regime, $\gamma < (\eta K/m)^{1/2}$, where, for $\eta > 4.604$,

the tip may perform multiple slips over a number of lattice sites and even overshoot the lowest well of the potential $U(x)$. In the latter case the minimal spring force reached during stick-slip oscillations is negative.

The elastic instability occurring for $\eta > 1$ results in a nonzero value of the low-velocity kinetic friction that is given by the energy drop from the point of instability to the next minimum of the potential divided by a [26]. For $\eta < 1$ this instability disappears and the friction is viscous, $F_k \to 0$ for $v \to 0$. The emergence of static friction can be interpreted as the arousal of a saddle-node bifurcation as a function of η, realizing a sort of fold-catastrophe scenario [27].

In experiment, the effective value of the PT parameter η can be controlled by the variation of the normal load on the contact, which changes the potential corrugation U_0 more than the contact stiffness. FFM experiments at low normal loads indeed demonstrated smooth sliding with ultralow friction, connected to the absence of elastic instabilities [15, 28]. At higher loads instead, "atomic" stick-slip took place with the atomic periodicity of the substrate lattice, while increasing load further led to a multiple slip regime as predicted by the PT model, see Fig. 10.1c.

In real systems at *finite temperature*, hysteresis and dissipation must disappear in the zero-speed limit of sliding, where stick-slip instabilities are preempted by thermal fluctuations. This regime is sometimes termed "thermolubricity" [18, 19, 29–32]. The main aspects of thermal effects on friction were considered in the pioneering work of Prandtl [16]. Thermal effects can be incorporated into the PT model (10.1) by adding a thermal random force $\hat{f}(t)$ and the damping term $-m\gamma\dot{x}$ to the conservative force between the slider and substrate, so that the tip motion is described by the Langevin equation

$$m\ddot{x} + m\gamma\dot{x} = -\partial U\,(x,t)/\partial x \;\; + \hat{f}(t)\,. \tag{10.2}$$

The random force should satisfy the fluctuation-dissipation theorem. As usual, it is chosen with zero mean $\left\langle \hat{f}\,(t) \right\rangle = 0$ and δ-correlated:

$$\left\langle \hat{f}(t)\, f(t') \right\rangle = 2m\gamma\, k_B T\, \delta(t - t')\,, \tag{10.3}$$

where k_B denotes the Boltzmann constant and T temperature. The random force and the damping term arise from interactions with phonons and/or other fast excitations that are not treated explicitly.

In the thermal PT model, (10.2) and (10.3), beside the PT-parameter η, thermal fluctuations bring out a new dimensionless parameter δ representing the ratio between the pulling rate v/a and the characteristic rate of thermally activated jumps over the potential barriers, $\omega_0 \exp(-U_0/k_B T)$, where ω_0 is the attempt frequency [32]. As a result, one should distinguish between two regimes of motion: (i) $\delta \ll 1$, regime of very low velocities or high temperatures (typically $v < 1$ nm/s at room temperature), where the tip has enough time to jump back and forth across the barrier, and (ii) $\delta \gg 1$, the stick-slip regime of motion, where thermal fluctuations only

occasionally assist the tip to cross the barrier before the elastic instability is reached. In these two regimes the following expressions for kinetic friction have been suggested [18, 19, 32]:

$$F_k(v,T) = \alpha(T)\, v + O\left(v^3\right), \quad \delta \ll 1, \tag{10.4}$$

$$F_k(v,T) = F_0 - bT^{2/3} \ln^{2/3}\left(B\frac{T}{v}\right), \quad \delta \gg 1 \ \text{ and } \ v < BT. \tag{10.5}$$

Here F_0 is the athermal ($T = 0$) low-velocity limit of friction, $\alpha(T) \propto (K/\omega_0) \times (U_0/k_BT)\exp(U_0/k_BT)$ is the equilibrium damping experienced by the tip (note that α is independent of the ad-hoc damping coefficient γ), and b, B are positive constants which depend on m, K, a, U_0 and γ but not on v and T. Equation (10.4) describes the slow friction regime (thermolubricity) and corresponds to the linear-response regime, while (10.5) has been derived assuming that thermally activated depinning occurs in the vicinity of the athermal instability point. The velocity and temperature dependences of friction force predicted by (10.5) result from the fold catastrophe scaling of the potential barriers, $\Delta E \propto (\text{const} - F)^{3/2}$. In between the regimes described by (10.4) and (10.5) one should observe a logarithmic dependence of F_k on velocity. However, it is very difficult to distinguish between $[\ln(v)]^{2/3}$ and simple $\ln(v)$ behavior in experiments as well as in numerical simulations [33]. The logarithmic (or $[\ln(v)]^{2/3}$) regime can span several decades, until v becomes so large that the inertial or viscous-like effects set in. The $[\ln(v)]^{2/3}$ dependence of the average rupture force has been also found in single-molecule unbinding experiments where the energy landscape of complex biomolecules is probed by applying time-dependent forces [34].

The theoretical framework outlined above explained a number of FFM experimental results on single crystal surfaces [29, 30, 35]. Furthermore, the statistical distribution of friction forces was measured to match predictions from the PT model [36]. These results provide strong evidence that atomic stick-slip in FFM is attributable to thermally activated slip out of a local minimum as described by the PT model. Thermally activated stick-slip friction is seen in MD simulation at sufficiently low speeds only, which are so far achievable through accelerated MD [37]. At higher speeds, friction is mostly determined by dissipative athermal dynamical processes, which correspond to a fundamentally different regime of sliding. This limits severely the regime of validity of comparisons of the PT model with MD simulations.

Equations (10.4) and (10.5) also predict that kinetic friction should decrease with increasing temperature [18, 19, 38]. Thermal excitations help overcome energy barriers and reduce the stick-slip jump magnitude, so that nanofriction should decrease with temperature provided no other surface or material parameters are altered by temperature [39]. Recent experimental results [40–43], however, strongly disagree with the predictions of (10.4) and (10.5). Friction forces exhibit a peak at cryogenic temperatures for different classes of materials, including amorphous, crystalline, and layered surfaces. Instead, the temperature and velocity dependence of the kinetic friction is well described by the multicontact model [42, 43].

Several generalizations of the original 1D PT model include considerations of:

- two-dimensional (2D) structure of surfaces that led to the introduction of frictional imaging of interfaces [44–47];
- coupling between normal and lateral motion of the slider [48, 49] that led to a new approach to control friction and wear by modulating the normal load [50, 51];
- flexibility of the AFM tip apex that led to a predictions of new regimes of motion exhibiting complex stick-slip patterns [52, 53].

10.3 The Frenkel-Kontorova Model

The basic model describing the sliding of crystalline interfaces is the 1D Frenkel-Kontorova (FK) model ([54] and references therein). First analytically treated in [55] and then independently introduced to describe dislocations in solids [56–58], the FK model found subsequently a broad area of application, in particular, in surface physics, where it is used to unravel the behavior of adsorbed monolayers.

The standard FK model Hamiltonian

$$H = \sum_{i=1}^{N} \left[\frac{{p_i}^2}{2m} + \frac{1}{2} K (x_{i+1} - x_i - a_c)^2 + \frac{1}{2} U_0 \cos \frac{2\pi x_i}{a_b} \right] \tag{10.6}$$

describes a 1D chain of N harmonically coupled "atoms" subjected to a sinusoidal potential, see Fig. 10.2. The first term in (10.6) is the kinetic energy of the chain, the second one describes the harmonic interaction of the nearest neighboring atoms with the elastic constant K and equilibrium distance a_c, and the last term is the interaction of the chain atoms with the periodic potential of magnitude U_0 and period a_b. Static friction is probed by driving all atoms with an extra adiabatically increasing force F until sliding initiates.

The success of the FK model is partly due to the relevance of its continuum limit, valid for large K, where the FK equations of motion reduce to the exactly integrable sine-Gordon (SG) equation, the solutions of which, in addition to linear waves (phonons), include the topological solitons (called "kinks") and dynamical solitons ("breathers"). Tribological processes in the FK model are ruled by kinks. Consider the simplest case of the trivial commensurate ground state (GS), when the number of atoms N is equal to the number of minima of the substrate potential

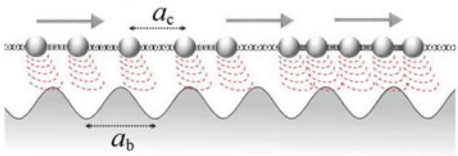


Fig. 10.2 A sketch of the FK model with the two competing lengths: interparticle and substrate periodicities

M, so that the dimensionless concentration $\theta = N/M = a_b/a_c$ is 1. In this case, adding (or subtracting) one atom results in a chain configuration with one kink (or antikink) excitation (more rigorously, kinks may be defined on the background of any *commensurate* GS; in what follows, $\theta = 1$ kinks will be indicated as "trivial" kinks, while for a general case we will use the term "superkinks"). After relaxation, the minimum-energy configuration corresponds to a local compression (or extension in the antikink case) of the chain. Kinks move along the chain far more easily than atoms because the activation energy $\varepsilon_{\rm PN}$ for kink motion [known as the Peierls-Nabarro (PN) barrier] is always smaller (or much smaller) than the amplitude U_0 of the substrate potential.

Because the kinks (antikinks) are associated with extra atoms (vacancies), their motion provides a mechanism for mass transport along the chain so that they are responsible for mobility, conductivity and diffusivity. The higher the concentration of kinks, the higher is the system mobility [60]. When the GS is commensurate (e.g., $\theta = 1$), the first step to initiate motion in the FK model is the (e.g. thermally induced) creation of a kink-antikink pair, see Fig. 10.3.

When the elastic layer is of finite extension, kinks are usually generated at a free end of the chain and then propagate along the chain until disappearing at the other free end [61]. Each run of the kink through the chain results in the shift of the whole chain by one lattice constant a_b. In the case of a finite film confined between two solids, one may similarly expect that the onset of sliding is initiated by the creation of a kink at the boundary of the contact. Subsequent kink motion is the basic mechanism of

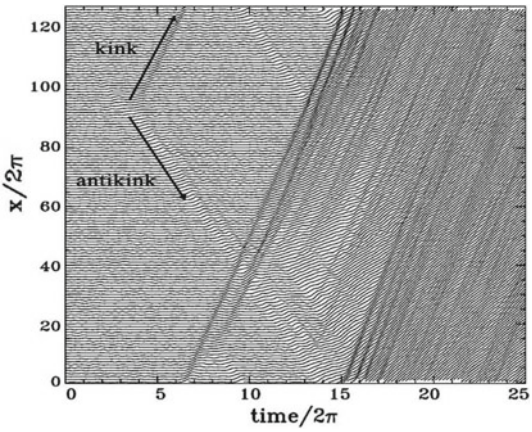


Fig. 10.3 Atomic trajectories as a function of time of the perfectly commensurate ($\theta = 1$) FK chain at the depinning transition, at a small temperature. The onset of motion is marked by the nucleation of a kink-antikink pair. The kink and antikink move in *opposite directions*, collide quasielastically (because of the periodic boundary conditions), and soon a second kink-antikink pair is created in the tail of the primary kink. This process repeats with an exponential (avalanche-like) growth of the kink-antikink concentration, leading to a sliding state. (Adapted from [59])

sliding. In 2D or three-dimensional (3D) systems, concepts of domain walls or misfit dislocations are used instead of kinks, but the physics of these processes remains qualitatively the same.

A crucial role in the FK model is played by incommensurability. Let the substrate period a_b and the natural period of the chain a_c be such that, in the limit of an infinite system's length, their ratio $\theta = a_b/a_c$ is irrational. In this case, under a not too restrictive condition on θ [62], there exists a critical value of the elastic constant K, such that for a higher rigidity the chain becomes effectively free of the substrate, i.e., the static friction F_s drops to zero, and the kinetic friction becomes very small. This phenomenon is known in physics from the beginning of the 1970s as the commensurate-incommensurate transition, or (later on, when S. Aubry developed the rigorous mathematical theory) "the transition by breaking of analyticity", or simply the *Aubry transition* [63–68]. A simple explanation of the $F_s = 0$ sliding state is the following: in this state, for every atom going up over the barrier, there is another atom going down, and the energy costs of these processes compensate exactly each other. Roughly speaking, the incommensurate FK chain acquires a "staircase" deformation, with regions of approximate commensurability separated by regularly spaced superkinks. If there is a nonzero probability to find atoms arbitrarily close to the maximum potential energy U_0 these superkinks are unpinned and mobile, otherwise they are pinned [69]. For a fixed value of U_0, the FK GS undergoes a transition between these two states at a critical value $K = K_c$ of the chain stiffness. The value K_c depends dramatically and discontinuously on the incommensurability ratio a_b/a_c defining the interface; it takes the minimal value $K_c \approx 1.0291926$ [in units of $2U_0(\pi/a_b)^2$] for the irrational golden mean ratio $a_b/a_c = (1+\sqrt{5})/2$ [54]. For $K > K_c$ there is a continuum set of ground states that can be reached adiabatically through nonrigid displacements of chain atoms at no energy cost (*the sliding mode*, or *the Goldstone mode*). On the other hand, for $K < K_c$, the atoms are trapped close to the minima of the substrate potential and thus require a finite energy to move over the corrugated substrate. Thus, for the incommensurate contact above the Aubry transition ($K > K_c$) chain sliding is initiated by even the smallest driving force and, accordingly, the static friction force vanishes, $F_s = 0$. On the contrary, for K below K_c the two incommensurate 1D surfaces are locked together due to pinning of the superkinks that separate local regions of common periodicity, and in this case we expect stick-slip. Note also that a finite-size FK chain is always pinned, even for an irrational value of a_b/a_c because of the locking of the free ends of the chain (although an Aubry-like transition, exhibiting a symmetry-breaking nature, can still be defined [70–72]).

In order to characterize the Aubry transition, it is convenient to introduce a "disorder" parameter ψ defined as the minimum distance of any atom from the nearest top of the substrate potential. Near the critical point the transition from pinned to sliding ground states occurs according to a power law,

$$\psi \propto (K_c - K)^{\chi_\psi}, \qquad F_s \propto \varepsilon_{\mathrm{PN}} \propto (K_c - K)^{\chi_{PN}}, \tag{10.7}$$

where the critical exponents χ depend on the incommensurability ratio, in particular, for the golden-mean case $\chi_\psi \approx 0.7120835$ and $\chi_{\rm PN} \approx 3.0117222$ [68, 73–79]. Notice that the Aubry transition exhibits a scaling behavior as typical for critical phenomena. The exponents in (10.7) are called super-critical because they only apply to the locked side of the transition, $K \le K_c$.

Likewise, one may introduce also sub-critical exponents for the sliding state above the Aubry transition. An important subcritical quantity is the effective viscosity $\Gamma = \lim_{F\to 0} F/(mv)$ which describes the steady-state average velocity v in response to an infinitesimally small dc force F applied to all atoms (to avoid infinite acceleration, an external damping γ should be included in the equation of motion). Γ is zero in the SG limit ($K \to \infty$) and diverges at the Aubry transition. For the golden-mean concentration, Γ scales as

$$\Gamma(K) \propto (K - K_c)^{-\chi_\Gamma} \tag{10.8}$$

with $\chi_\Gamma \approx 0.029500$. The scaling (10.8) is only appropriate immediately above the Aubry transition, while at $K \gg K_c$, Γ decreases toward γ. Accordingly, for any $K > K_c$ the mobility $B = v/F$ remains lower than its maximum value $(m\gamma)^{-1}$ even in the $F \to 0$ limit as illustrated for $K \gtrsim K_c$ in Fig. 10.4. A "frictionless" motion of the truly incommensurate GS exists only in the SG ($K \to \infty$) limit, where the substrate-corrugation U_0 term adds no dissipation to the one brought in by the γ term. Vanishing static friction has been first found within the FK model for mutually incommensurate periodicities and sufficiently hard infinite lattices [68]. Later on [80–82], this effect was predicted for infinite incommensurate contacts, and called *superlubricity*. The term superlubricity has been criticized as misleading, since it might wrongly suggest null friction in the sliding state in analogy to superconductivity and superfluidity. Instead, incommensurability of periodic interfaces cancels only one of the channels

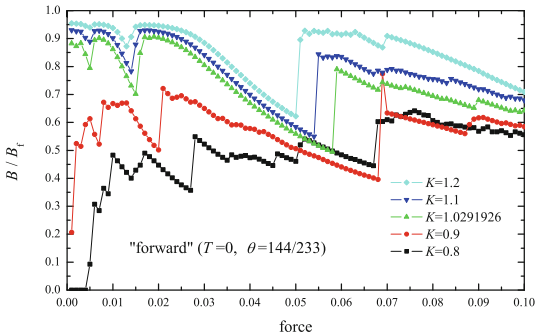


Fig. 10.4 The mobility $B = v/F$ normalized to the free-motion value $B_f = (m\gamma)^{-1}$ as a function of the dc force F for the classical FK model with the "golden-mean" concentration for different values of the elastic constant K below and above the Aubry threshold $K_c \approx 1.0291926$. The equation of motion included an external viscous damping, with a friction coefficient $\gamma = 0.1$. (Adapted from [54])

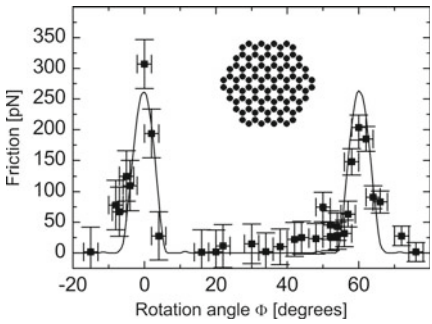


Fig. 10.5 The data points show the average friction force versus the rotation angle measured by [83]. The *curve* through the data points shows the calculated friction force from a PT-like model based on the sketched symmetric 96-atom flake. (Adapted from [83, 84])

of energy dissipation, that originating from the low-speed stick-slip elastic instability. Other dissipative processes, such as the emission of sound waves, still persist, and therefore even in the case of complete incommensurability the net kinetic friction force does not vanish, although in the superlubric regime one expects a substantial reduction of the friction force relative to a similar, but commensurate case.

Detailed experimental studies of superlubricity have been performed recently for friction between a graphite flake attached to the FFM tip and an atomically flat graphite surface [83–85]. Super-low friction forces (<50 pN) were found for most relative orientations of the flake and the substrate, for which the contacting surfaces find themselves in incommensurate states (see Fig. 10.5). For narrow ranges of orientation angles corresponding to commensurate contacts, stick-slip motion was observed and friction was high (typically 250 pN).

The kinetic friction properties of the FK model [86, 87] are probed by adding a (e.g. Langevin) thermostat as described for the PT model above. Even above the Aubry transition, where $F_s = 0$, the kinetic friction force F_k is nonzero, because the dynamics at any finite speed results in the excitation of phonons in the chain. At finite T, pinning can be overcome by thermal fluctuations, which can initiate sliding even in the most-pinned state, the fully commensurate one, see Fig. 10.3. Finally, we remark that friction in the dynamically driven FK model describes fairly the onset of sliding of a crystalline contact [88], but it cannot account for the highly inelastic plastic or quasi-plastic deformations of the surfaces characterizing real-life friction experiments.

The dimensionless atomic concentration $\theta = N/M = a_b/a_c$ in the FK system plays a crucial role since it defines the concentration of "geometrical" superkinks. As mentioned above, these excitations can be defined for any background *commensurate* atomic structure $\theta_0 = p/q$, where p and q are relative prime integers. If the concentration θ slightly deviates from the background value θ_0, the GS of the system corresponds to large domains with background commensurate structure θ_0, separated by localized incommensurate zones of compression (expansion) called superkinks (super-antikinks).

When the external force increases, the FK system with a non-trivial GS exhibits a hierarchy of first-order dynamical phase transitions from the completely immobile state to the totally running state, passing through several intermediate stages

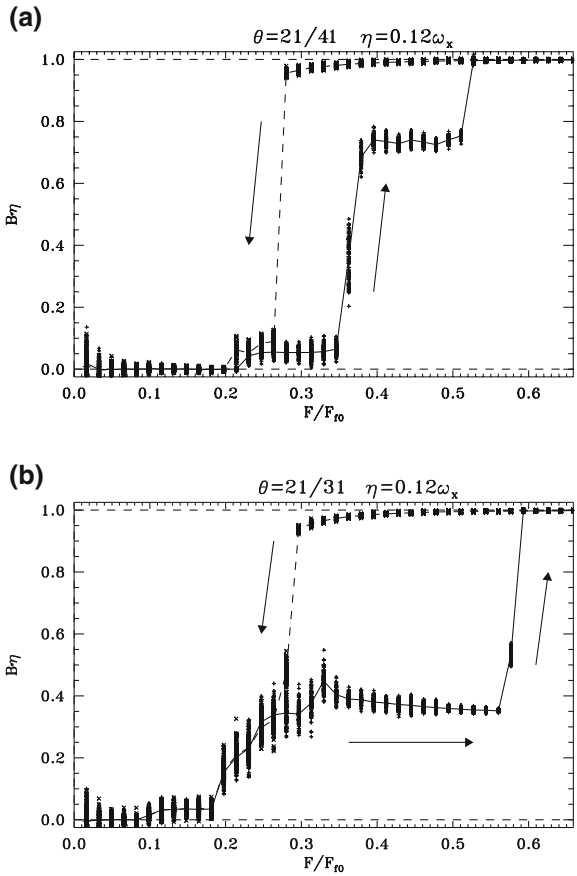


Fig. 10.6 The mobility $B = v/F$ versus the force F for the underdamped ($\gamma = 0.12$) FK model with exponential interaction ($K_{\rm eff} = 0.58$) **a** for $\theta = 21/41$ (superkinks on the background of a $\theta_0 = 1/2$ structure), and **b** for $\theta = 21/31$ (superkinks on the background of the complex $\theta_0 = 2/3$ structure). (Adapted from [89])

characterized by the running state of collective quasiparticle excitations, or kinks of the FK model. As an example, let us consider the $\theta = 21/41$ case when the mass transport along the chain is carried out by trivial kinks constructed on the background of the $\theta_0 = 1/2$ structure. As the average distance between the kinks is large (equal to 41 a_b in the GS), the kink-kink interaction is weak, and the atomic flux is restricted by the overcoming of kinks over the PN barriers (see Fig. 10.6a). When the driving force F increases, the now-tilted effective PN barriers are lowered (simultaneously with the original barriers of the substrate potential), resulting in the increase of the single-kink mobility. Thus, at zero temperature the crossover from the locked $B = 0$ state to the kink-running state takes place at the force $F \approx F_{tk} = C\pi\varepsilon_{\rm PN}/a_b$, where the factor $C \sim 1$ depends on the shape of the PN potential. The mobility in the kink-running state is $B \approx \theta_k B_f$, where $\theta_k = 1/41$ is the dimensionless kink concentration.

The further scenario depends on the value of the damping coefficient γ. At very low damping, $\gamma < 0.05$, there is no intermediate stages, because the running kinks destroy themselves as soon as they start to move: they will cause an avalanche

driving the whole system to the totally running state of atoms similarly to that shown in Fig. 10.3. At larger damping, $\gamma > 0.05$, the above-mentioned intermediate stages with running kinks exist. A mechanism of the second abrupt increase of the mobility depends on γ too (for details see [54, 89, 90]). Between the kink-running stage and the totally running state there may be a specific "traffic-jam" regime [54].

This qualitative picture holds also for a more complex atomic structure like $\theta = 21/31$ [89] (see Fig. 10.6b for $N = 105$ and $M = 155$). In this case the state of running trivial kinks is preceded by the state of running superkinks. The GS in this case corresponds to domains of the complex $\theta_0 = 2/3$ commensurate structure, separated by *superkinks* with an average spacing $30\,a_b$ between them. On the other hand, the $\theta = 2/3$ structure can be viewed as a dense lattice of *trivial* kinks defined on the background of the $\theta_0 = 1/2$ structure. This specificity clearly manifests itself in the $B(F)$ dependence. During the force-increasing process, there are now two sharp steps of increasing of the mobility B. The first one, at $F = F_{sk} \approx 0.08\,F_s$, corresponds to the situation where the superkinks start to slide, whereas the second step, occurring at $F = F'_{tk} \approx 0.18\,F_s$, corresponds to the transition of the trivial kinks to the running state.

10.3.1 Extensions of the Frenkel-Kontorova Model

Many relevant generalizations of the FK model have been proposed so far to cover a large class of frictional relevant phenomena. They mainly consist of modifications of model interactions or of dimensionality. For realistic physical systems, anharmonicity can be introduced in the chain interatomic potential, see [54]. The main novelties here include effects such as a broken kink-antikink symmetry, new types of dynamical solitons (supersonic waves), a possible breakup of the antikink soliton followed by a chain rupture, and a modified kink-kink interaction. Strong anharmonic effects are responsible of a strong kink-antikink asymmetry in recent experiments of friction in repulsive colloids [91, 92], see Chap. 19. Likewise, nonsinusoidal periodic substrates, characterized, e.g., by sharp bottoms and flat barriers [93], have been investigated to address atoms adsorbed on a metal surface. Complex unit cell substrates [60, 94], as well as quasiperiodic [95, 96] and disordered corrugated profiles [97–99] have also been considered. These deviations from the standard FK model may lead to qualitatively different excitations such as different types of kinks, phonon branches, and to changes in the kink-antikink collision scenario. From a tribological point of view, different types of sliding behavior are to be expected at low-driving forces, when the dynamics is mainly governed by the motion of kink-like structures.

A very important generalization of the standard FK chain with relevant consequences for the tribological properties (critical exponents, scaling of friction force with system size, mechanisms of depinning, etc.) involves increasing the dimensionality of the model. Especially the 2D generalized versions of the FK model [11, 54] are naturally applicable to the description of a contact of two crystalline surfaces (i.e., the case of "dry" friction), in particular as is realized in QCM experiments, where 2D monoatomic islands of adsorbate atoms slide over a periodic crys-

talline substrate [8], or, very recently, in 2D colloidal monolayers over laser-generated optical lattices [91, 92, 100, 101]. Among 2D generalized FK models we mention the model consisting of two coupled FK chains [102], the 2D "springs and balls" FK model describing a 2D layer of harmonically interacting atoms in the 2D periodic substrate, the scalar anisotropic 2D FK model treating a system of coupled 1D FK chains, the vector anisotropic 2D FK model (e.g., the zigzag FK model), the vector isotropic 2D FK model [103–106], and the 2D tribology model [88, 107] (see also [54] and references therein).

The approaches based on these models are especially powerful in the investigation of the transient behavior at the onset (or stopping) of sliding, which is quite difficult to study in fully realistic 3D models (e.g., see [108]). As a typical example, let us describe the onset of sliding observed in the vector 2D FK model, where a 2D layer of atoms is subjected to a periodic substrate potential with the triangular symmetry [108]. The transition from the locked to running state is mediated by the formation of an island of moving atoms in a sea of essentially stationary particles. The size of the moving island grows quickly in the direction of the driving force, and somewhat more slowly in the perpendicular direction. Inside the island the atoms largely maintain their triangular structure due to the stiffness of the atomic layer. Hence one sees areas of essentially perfect triangular lattice surrounded by a closed boundary of partial dislocations. In simulation, due to periodic boundary conditions, the island eventually joins up on itself. There forms a strip ("river"), oriented parallel to the driving force and bounded in the direction perpendicular to the driving force, in which particles move along the periodically-continued system. Outside this stripe the particles are immobile. This stripe then broadens perpendicularly to the driving direction until all atoms are moving. The evolution of this scenario is illustrated in a series of snapshots in Fig. 10.7.

Noncontact AFM tips oscillating on top of kink-like adsorbate regions [109] dissipate significantly more than nearly in-registry regions. This mechanism is explained by the higher softness and mobility of solitonic regions [110–113], and it has been

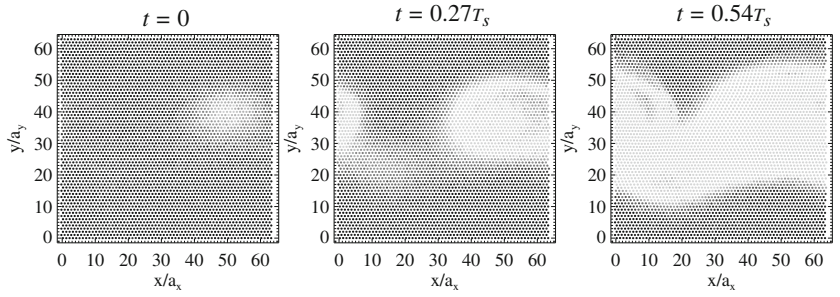


Fig. 10.7 Successive snapshots of the mechanism of the locked-to-running transition in the vector 2D FK model for the LJ interaction with $K_{\text{eff}} = 0.9$, $\gamma = 0.141$, $T = 0.05$, and $F = 0.9933$. The positions of the atoms are indicated by *circles*. The x component of the particle speed is shown in *gray* scale: from zero (*black*) to maximum (*lightest gray*) velocity. (Adapted from [108])

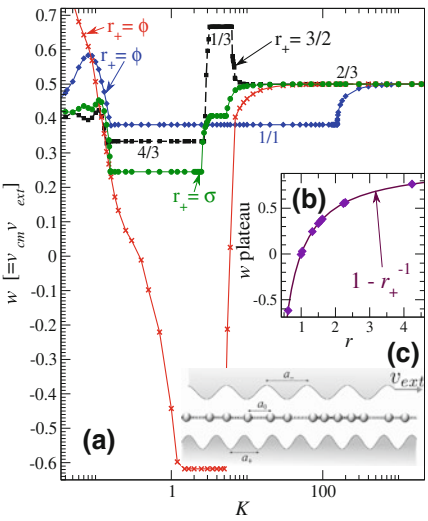


Fig. 10.8 **a** The average drift velocity ratio $w = v_{cm}/v_{ext}$ of the chain as a function of its lubricant stiffness K for different commensurability ratios (r_+, r_-), with $r_\pm = a_\pm/a_0$: commensurate $(3/2, 9/4)$, golden mean (GM) (ϕ, ϕ^2) $(\phi \simeq 1.6180\ldots)$, spiral mean (SM) (σ, σ^2) $(\sigma \simeq 1.3247\ldots)$, and (ϕ^{-1}, ϕ). The (ϕ, ϕ^2) 1/1 plateau value is $w = 0.381966\ldots$, identical to $1 - \phi^{-1}$ to eight decimal places. **b** The main plateau speed w as a function of r_+. **c** A sketch of the model. (Adapted from [121])

demonstrated by the dynamics of an incommensurate FK chain, forced and probed by a locally-acting oscillation [114].

In investigating *confined* systems under shear, FK-like models with just one particle [115–117] or an interacting atomic chain [118–120] embedded between *two* competing substrates have led to uncover peculiar tribological phenomena related to stick-slip dynamics or to the appearance of remarkable sliding regimes of motion. For example, velocity quantization phenomena have been reported [121, 122] in the motion of an idealized 1D solid lubricant.

In slider-lubricant-slider geometry [120] sketched in Fig. 10.8c, the lattice mismatch can give rise to peculiar and robust "quantized" sliding regimes, characterized by a nontrivial geometrically fixed ratio of the mean lubricant drift velocity and the externally imposed translational velocity. In detail, the speed ratio $w = v_{cm}/v_{ext}$ remains pinned to exact "plateau" values over wide ranges of parameters, such as sliders corrugation amplitude, external velocity, chain stiffness and dissipation (see Fig. 10.8a), and is strictly determined by the length ratios alone. The plateau mechanism has been interpreted in terms of solitons, formed by the mismatch of the lubricant periodicity to that of the nearer matching substrate, being rigidly dragged forward by the advancing sinusoid representing the other, more mismatched, slider.

The finding of exact plateaus implies a kind of "dynamical incompressibility", namely identically null response to perturbations or fluctuations trying to deflect the CM velocity away from its quantized value. In order to probe the robustness of the plateau attractors, an additional constant force F_{ext}, acting on all particles in the chain, was introduced. As expected, as long as F_{ext} remains sufficiently small, it has no effect whatsoever on the velocity-plateau attractor. The plateau dynamics is only abandoned above a critical force F_c. The transition, occurring for increasing external driving force F_{ext} acting on the lubricant, displays a large hysteresis, and

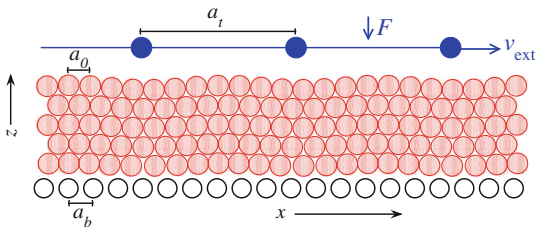


Fig. 10.9 A sketch of the model with rigid *top* (*solid circles*) and *bottom* (open) layers (of lattice spacing a_t and a_b respectively), the former moving at externally imposed x-velocity v_{ext}. One or more lubricant layers (*shadowed*) of rest equilibrium spacing a_0 are confined in between. (Adapted from [126])

has the features of depinning transitions in static friction, only taking place "on the fly" [123, 124]. Although different in nature, this phenomenon appears isomorphic to a static Aubry depinning transition [67, 68], the role of particles now taken by the moving solitons of the lubricant-substrate interface. The confined model was extended beyond the standard sinusoidal corrugation of (10.6): the quantized velocity remains, but a nonsinusoidal corrugation can affect the parametric region where the velocity plateau extends [125].

A quantized sliding state of the same nature has been demonstrated by MD simulations for a substantially less idealized two-dimensional model [126], where atoms of a lubricant multilayer film were also allowed to move perpendicularly to the sliding direction and interact via Lennard-Jones potentials (see Fig. 10.9). This dynamical regime, is shown to be robust against the effects of thermal fluctuations, quenched disorder in the confining substrates, and over a broad range of loading forces. By evaluating its tribological properties in terms of averaged kinetic friction $\overline{F_k}$ exerted on the top slider, this lubricant quantized sliding has been found [127] to be characterized by significantly low values of $\overline{F_k}$, see Fig. 10.10.

While certain of these phenomena, such as chaotic and inverted stick-slip motion, two types of smooth sliding and transitions between them, have already been observed [128, 129], others are still waiting for experimental confirmation. Last but not least, the combined Frenkel-Kontorova-Tomlinson (FKT) model [130, 131] has been introduced including harmonic coupling of the interacting chain atoms to a sliding body. The FKT model introduces more degrees of freedom than the PT model, and it has been used to describe effects of finite size and stiffness of the AFM tip and of normal load on friction [132, 133]. The latter effect has been modeled assuming that the amplitude U_0 of potential corrugation increases proportionally to the applied normal force. The validity of the FKT model has been tested by 3D MD simulations [133], which confirmed the outcome of the model for most investigated regimes except for the limit of very low stiffness and high normal load. Unlike the FKT model where the breakdown of superlubricity coincides with the emergence of the metastable states, in 3D simulations certain metastable states appear to reduce the frictional force leading to nonmonotonic dependence of force on normal load and tip

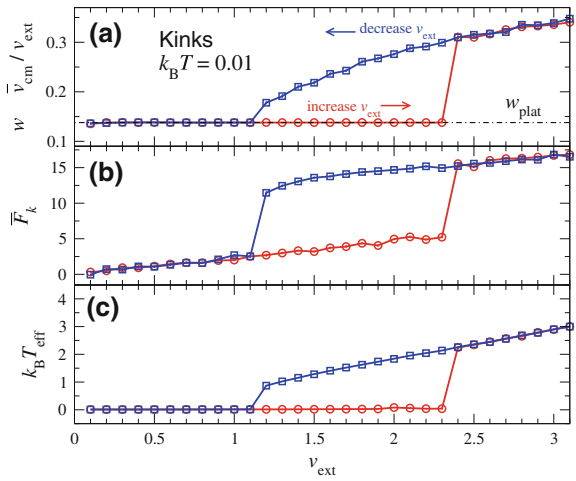


Fig. 10.10 As a function of the adiabatically increased (*circles*) or decreased (*squares*) top-substrate velocity v_{ext}, the three panels report: **a** the average velocity ratio w; **b** the average friction force experienced by the top substrate; **c** the effective lubricant temperature, computed using the average kinetic energy in the frame of reference of the instantaneous lubricant center of mass. (Adapted from [127])

compliance. Increasing dimensionality and adding realistic features to the FK model brings its extensions into closer and closer contact to full-fledged MD simulations.

10.4 Molecular Dynamics Simulations

The simple low-dimensional MMs discussed above are useful for a qualitative understanding of many physical aspects of friction. To address subtler features of a specific interface, one should go beyond MMs including atomistic structural details of the interface. Such an approach is provided by MD simulations.

Advances in computing hardware and methodology have dramatically increased our ability to simulate frictional processes and gather detailed microscopic information for realistic tribological systems. MD simulations represent controlled computational experiments where the dynamics of all atoms is obtained by solving numerically Newton or Langevin equations of motion based on suitable interparticle interaction potentials and the corresponding interatomic forces. The geometry of the sliding interface and the boundary conditions (e.g. as sketched in Figs. 10.11 and 10.12) can be chosen to explore friction, adhesion and wear.

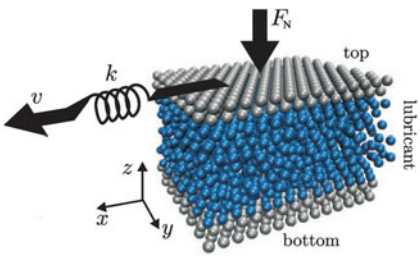


Fig. 10.11 Sketch of a typical MD simulation of a sheared boundary-lubricated interface. Periodic boundary conditions are applied in the lateral (x-y) directions, and a normal load F_N is applied to the top substrate in the z direction.(Adapted from [10])

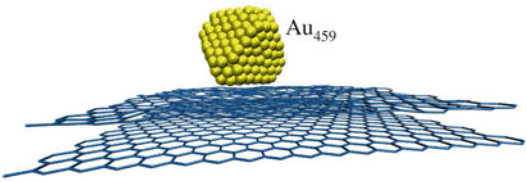


Fig. 10.12 A simulated truncated-octahedron Au_{459} cluster sliding with one of its (111) facets over a mobile graphite substrate. (Adapted from [134]).

A worthwhile guide to atomistic MD simulations of frictional processes focusing on fundamental technical aspects (realistic construction of the interface, appropriate ways to impose load, shear, and the control of temperature) can be found in the review articles by [135, 136]. By following the Newtonian dynamics of a system executing sliding for a significant amount of time, quantities of physical interest such as instantaneous and average friction force, mean (center-of-mass) slider velocity, heat flow, and correlation functions are numerically evaluated. Unlike standard equilibrium MD simulations of bulk systems, frictional modeling inherently involves non-equilibrium conditions and a nonlinear dissipative response to the external driving. A standard practical assumption is to add Langevin terms to Newton's equations, like in (10.2) and (10.3) for the PT model at finite temperature.

The choice of the appropriate interaction forces between atoms represents a major problem. If $U\{R_1, R_2, \ldots R_N\}$ is the total energy of the system as a parametric function of all atomic coordinates $\{R_i\}$, the force on atom i is $F_i = -\nabla_{R_i} U$, perfectly determined once U is known. Unfortunately, this is generally not the case, because U is determined by the quantum mechanics of electrons—a much bigger and unsavory problem to solve. *Ab-initio* MD, e.g. of the Car-Parrinello type [137], has not really been of use so far in sliding friction, mainly because it can handle only rather small systems, typically hundreds of atoms, for relatively short times, typically $\ll$1 ns. Most MD frictional simulations are therefore based on reasonable empirical interatomic forces ("force fields"), ranging from relatively sophisticated energy surfaces accounting for electrons at the density-functional level or at the tight-binding level [138], to angle-dependent many-particle potentials, to simple pairwise potentials (e.g. Lennard-Jones or Morse), to basic simple models of elastic springs which represent extensions of FK-type formulations. In practice, several reliable force fields, parameterized to fit different ranges of experimental data and material combinations, are available in the literature [139–143]. While this allows qualitative atomistic simulations of sliding friction, it is often far from quantitative. For example, during such a violent frictional process as wear, atoms may change substantially their coordination, their chemistry, sometimes their charge. Once a specific system is understood after the elaborate development of satisfactory potentials, the mere change of a single atomic species may require a complete reparameterization of the interatomic forces. As a result, systematic frictional studies may become quite a challenge in the absence of a consistent set of potentials. A promising approach consists in

the use of the so-called reactive potentials [144–146], capable of describing chemical reactions and interface wear, with the advantage, for large-scale atomic simulations, of a good computational efficiency compared to first-principle and semi-empirical approaches.

10.4.1 Thermostats and Joule Heat

In a tribology experiment, mechanical energy is converted to Joule heat which is carried away by phonons (and electron-hole pairs in metals). In a small-size simulation, the excitations generated at the sliding interface propagate and crowd-up into an excessively small region of "bulk" substrate, where they are back-reflected by the cell boundaries, rather than properly dispersed away. To avoid overheating and in order to attain a frictional steady state, the Joule heat must therefore be steadily removed. If this removal is done by means of standard equilibrium thermostats such as velocity rescaling or Nosé-Hoover or even Langevin dynamics, an unphysical dissipation is distributed throughout the simulation cell, so that simulated atoms do not follow their real conservative motion, but rather execute an unrealistic damped dynamics which turns out to affect the overall tribological properties [147]. Similarly in the PT and FK models, the damping parameter γ is known to modify kinetic and frictional properties, but there is no clear way to chose the value of γ.

To solve this problem, one should attempt to modify the equations of motion inside a relatively small simulation cell so that they reproduce the frictional dynamics of a much larger system, once the remaining variables are integrated out. One approach is to use in Langevin equations a damping coefficient which depends on the coordinate and velocity of each lubricant atom; these dependences can be taken to fit the known dissipation of atoms adsorbed on a surface [148]. In turn, this method requires a modification of the standard Langevin technique [149]. A more rigorous approach is a recent implementation of a non-conservative dissipation scheme, based on early formulations by [150–152] and subsequent derivations by [153–155], that has demonstrated the correct disposal of friction-generated phonons, even in the relatively violent stick-slip regime [156, 157]. All atoms near the sliding interface follow plain Newton's equation, while the atoms in the deepest simulated layer, representing the boundary layer in contact with the semi-infinite heat bath, acquire additional non-conservative (and non-Markovian) terms which account for the time history of this layer through a memory kernel [153, 154]. Nanofriction simulations exploiting this dissipative scheme have recently been implemented that conceptually and practically improve over a traditional Langevin simulation.

10.4.2 Size- and Time-scale Issues

Modern CPUs perform of the order of 10^9 floating-point operations per second (FLOPS) per core. Classical MD can take advantage of medium-scale parallelization, with fairly linear scaling to approximately 10^2 cores, thus affording about 10^{11} FLOPS. As the calculation of the force acting on each atom (usually the dominating

step in a MD calculation) can require, depending on the complexity and range of the force field, 10–10^2 operations, the product of the number of simulated particles N times the number of time-integration steps N_{step} per runtime second on a modern medium-size parallel computer is approximately $N\,N_{\text{step}} \simeq 10^{10}$. With a typical time-step of $\sim$1 fs, a medium-size simulation involving $N = 10^5$ atoms can progress at a rate of 10^5 fs per second, i.e. approximately 10^9 fs $= 1$ μs in a simulation day. This total time scales down for larger systems sizes.

These estimates should be compared with the typical times, sizes, and speeds of tribology experiments. If we wish to address macroscopic sliding experiments, the speed would be in the 0.1–10 m/s range: in 1 μs the slider advances by 0.1– 10 μm, i.e. approximately 10^3–10^4 typical lattice spacings, enough for a fair statistics of atomic-scale events (but hardly sufficient to gather significant data about phenomena such as the diffusion of additives or of wear particles within the lubricant, or step- or defect-related phenomena). In a nanoscale FFM experiment, however, the tip advances at a far smaller average speed (i.e. $\simeq$1 μm/s) and we can simulate a miserable $\simeq$1 pm advancement in a typical run, far too short to observe even a single atomic-scale event, let alone reaching a steady state. Therefore, whenever long-distance correlations and/or slow diffusive phenomena and/or long equilibration times are expected, MMs will perform better than fully atomistic MD simulations. There is nevertheless so much physical insight to be extracted from MD simulations that it makes sense to run them even at larger speeds than in AFM or SFA experiments; and indeed, the sliding speed adopted in most current atomistic MD frictional simulations is in the m/s region.

While the high-speed kinetic friction is reproduced adequately in MD simulation, it is not so for the static friction which essentially depend on the system size, usually decreasing with the increase of the interface area. To overcome this problem, one may use scaling arguments which allows to find the large-area static friction from MD simulation for a rather small system [158].

One of the challenging problems for MD simulations is to account for the transition from stick-slip to steady sliding. In SFA and AFM experiments, stick-slip with its associated hysteresis and large friction generally disappears for speeds larger than $\sim$1 μm/s, while in MD simulations the transition takes place in the m/s range. This major discrepancy (up to $\sim$6 orders of magnitude in speed!) between simulations and experiments has been discussed [159–162], and relates to the effective spring-force constants and mass distributions, that are hugely different in the two cases, and much oversimplified in simulations. Several attempts to fill these gaps rely on methods, including hyperdynamics, parallel-replica dynamics, temperature-accelerated dynamics, and on-the-fly kinetic Monte Carlo devised in recent years [163–165].

Another important aspect present in experiments and largely missed by MD simulations is the ageing of contacts due to the interface relaxation. Contact ageing is believed to be responsible for the increase of the static friction force as a function of the contact time. Direct imaging of contact regions in samples under a normal load shows a logarithmic growth with time [166], leading therefore to a slowly increasing static friction. At the phenomenological level, frictional ageing is well described by rate and state friction laws, widely used in geophysics [167], but its microscopic

origin is still debated. The main mechanisms that have been invoked in the past to explain it are plastic creep [168] or chemical strengthening at the interface [169]. In a recent paper [169], AFM was used to explore ageing in nanoscale contact interfaces, finding supporting evidence for the second mechanism, since when the contact surface was passivated it showed no ageing. It is however likely that at larger scales and loads plastic creep would also play an important role. Beyond its direct relevance for friction, the intriguing issue of contact ageing occurs in other non-equilibrium disordered systems such as granular media or glasses.

10.4.3 Multiscale Models

Since it is currently impossible to treat atomistically all the characteristic length scales that mark the dynamical processes entering the friction coefficient of engineering materials, a rising effort is nowadays devoted to develop multiscale approaches. The basic consideration is that unless conditions are very special, all processes far away from the sliding interface can be described approximately by continuum mechanics and simulated using finite elements, allowing for a macroscopic description of elastic and plastic deformation. The advantage of these continuum-theory methods is that it is possible to increasingly coarse-grain the system as one moves away from the sliding contact, thereby highly reducing the computational effort. Several groups [170, 171] combine the atomistic treatment of the interfacial mating region, where displacements occur on an atomic or larger length scale, with a coarse-grained or finite-element continuum description elsewhere, where strains are small and continuous. The main difficulty lies in the appropriate matching conditions between the atomistic and continuum regions [172]. Since the detail of lattice vibrations (the phonons), which are an intrinsic part of any atomistic model, cannot be fully represented at the continuum level, conditions must be met that at least the acoustic phonons should not be reflected excessively at the atomistic-continuum interface. In other words, matching at this interface must be such that sound deformations transmit with reasonable accuracy in both directions, which is vital to a proper account of Joule-heat disposal into the bulk. Chap. 14 covers this issue.

10.4.4 Selected Results of MD Simulations

Here we survey some results from the growing simulation literature, certainly not providing an adequate review of the field. With two sliding surfaces separated by a thick fluid film, friction is mainly determined by the lubricant viscosity. The friction coefficient can be calculated using the Navier-Stokes equations, showing a monotonic increase with the relative sliding velocity [173]. For small driving velocity and/or high load, the lubricant cannot usually keep the surfaces apart and solid-solid contact eventually ensues. But even before full squeezeout, a liquid confined within a

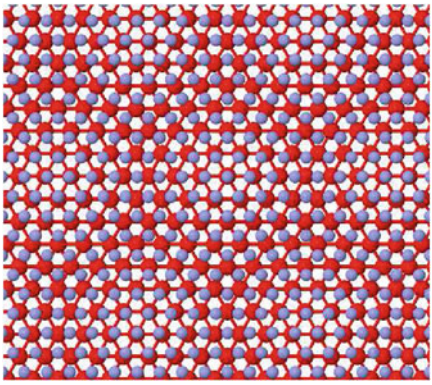


Fig. 10.13 Top view of a snapshot of a 3D MD simulation of the solitonic pattern arising at the boundary layer of a solid lubricant (*light gray*) in contact with a perfect crystalline surface (*dark/red*), induced by a 16 % lattice-constant mismatch. The Lennard-Jones interaction of this simulation favors in-registry hollow sites, while unstable top sites mark solitonic regions. Other layers were omitted for clarity. (Adapted from [10])

nanometer-scale gap ceases to behave as a structureless fluid—it becomes layered and even may solidify when thickness decreases below about five molecular layers. Pioneering studies of confined systems under shear reveal a sequence of drastic changes in the static and dynamic properties of fluid films in this "boundary-lubrication" regime.

SFA experiments [174] and MD simulations [175, 176] have both shown clear upward frictional jumps, in correspondence to squeezeout transitions from N to $N-1$ lubricant layers. The lubricant squeezeout for increasing load becomes harder and harder, corresponding to a (near) crystallization of the initially fluid lubricant [177–179]. But friction would not necessarily always jump upward during the layer-by-layer squeezeout: it could jump downward as well, if lattice mismatch between the compressed boundary lubricant layer and the rigid substrates jumped from commensurate to incommensurate, the latter superlubric with a mobile soliton pattern, as sketched in Fig. 10.13.

MD investigations of a melting-freezing mechanism in the stick-slip phenomenology of boundary-lubricated films were carried out by [180–183]. Various realistic models for lubrication layers in very specific contexts have been investigated with extensive MD simulations [162, 184–187]. During sliding, a thin lubricant film may be solid or liquid, depending on the interplay between the strength of interaction of lubricant molecules with the surfaces and between themselves. In the case of weaker interaction with the surfaces, the sliding takes place at the lubricant-surface interface, and the lubricant film remains in the solid state. And if the substrate and the solid lubricant both have near-ideal crystalline structures and these structures are incommensurate, superlubric sliding with zero static friction and very low kinetic friction ensues: the solid lubricant may provide minimal friction coefficient. But in practice the surfaces, as well as the lubricant, are unlikely to retain undefected crystalline structures. The presence of impurities or defects between the sliding surfaces

can, even in a relatively small concentration, leads to pinning and nonzero static friction [188], thus destroying superlubricity. However, if one carefully chooses the parameters of the lubricant, the perfect sliding could again be achieved, because the lubricant may self-order itself during sliding [189].

In the opposite case of a strong interaction of lubricant molecules with the surfaces, the latter are covered by lubricant monolayers which protect surfaces from wear, and sliding occurs somewhere in the bulk of the lubricant film. In this case the film usually melts during sliding (a thin lubricant film, however, is not completely liquid typically, it has a layered structure imposed by the surfaces, and slips often occur by a layer-over-layer mechanism [162]). The kinetic friction for a liquid lubricant film, even for very thin films (but thicker than two molecular layers), is caused by lubricant viscosity. A lower viscosity implies better lubrication, i.e. lower kinetic friction. Thus, the "best lubricant" is vacuum: the viscosity is zero, so the friction should be zero too (more rigorously, even in the case of a vacuum gap between sliding bodies, due to quantum effects a nonzero friction arises even at zero temperature [190, 191]). Air and water are excellent lubricants (recall how slippery is a thin water film over the ice surface; note also that nature adopted aqueous solutions as a lubricant in the articular joints of animals).

However, a low viscosity easily leads to squeezing out of the lubricant from the sliding interface; then the surfaces come in direct contact and start to be eroded by rubbing wear. That is why in machinery oil-based lubricants are used typically. An oil has large viscosity which leads to high friction, but it is also hard to squeeze out of the contact zone due precisely to its high viscosity. Thus, lubrication engineering is in a permanent search for a compromise: on one hand, the viscosity should be low to provide low friction, but on the other hand, it must be large enough to avoid oil squeezing out and machinery wear. This problem is especially actual for nano-devices, where traditional lubricants often fail to operate.

For a liquid lubricant, the role of the shape of lubricant molecules is also nontrivial. Simulation [192] showed that brush-forming lubricants, e.g. head-glued molecules which work like a hair, provide better lubrication—even if the surfaces are pushed together so strongly that most lubricant is squeezed out leaving fewer than two monolayers, it continues to operate providing smooth sliding.

To some degree, MD can address relevant realistic features, including the roughness of real surfaces. Even if one polishes the surfaces and makes them smooth, they still consist of domains with different orientation because of the mismatch between the crystalline structures. Simulation [193, 194] showed that when two crystalline surfaces (even with a few lubricant layers in between) are rotated relatively each other, the static friction force changes with the misfit angle over more than one order in magnitude.

Finally, in the case of a lubricant which melts during sliding but solidifies at stick, the formation of solid droplets (grains) pin the surfaces by bridges; but these bridges are not formed instantaneously, not all at once. The bridges, grains, domains, asperities, etc. acquire different sizes and therefore different stress values to be broken at the subsequent slip. The interface is always divided into areas characterized by different thresholds for the onset of sliding. Therefore, such a parameter as the static

friction force is in fact not a physical quantity, it describes some average value measured in an experiment (and its value depends on the way the experiment is organized). This fact must be taken into account in description of real tribosystems, and this may be done with the help of the earthquake-like model considered in the next Sect. 10.5.

10.5 Earthquake-Like Models

An earthquake-like (EQ) models, known also as a multi-contact models, assumes that the contact between two surfaces is realized only at certain points. Typically for a dry contact of rough surfaces these contacts are associated with asperities, but they may otherwise represent molecular bonds or capillary bridges, or they may account for patches of solidified lubricant or its domains for the case of lubricated friction. The contacts are characterized by a continuous distribution of the static threshold values $P_c(x_s)$. A contact itself behaves as follows from MD simulation and tip-based experiments—it operates as an elastic spring until the local shear force $f_i = kx_i$ (k is the elastic constant and x_i is the contact stretchings) is below a threshold value $f_{si} = kx_s$, and breaks when the threshold is exceeded. When the upper block moves, the forces on the contacts increase, and at some moment they start to break in sequence, one after another, with weaker contacts breaking earlier and strongest contacts resisting to the last. Once a contact is broken, it slips and then is reformed again. Such a model was used in many studies [42, 43, 159, 195–200] and successfully accounted for friction at the meso- and macroscale. Models of the same class were also used to describe the failure of fibre bundles and faults [201–203].

The master equation. The EQ model, being the cellular automaton model, allows no analytical treatment. Its kinetics, however, may be reduced to the so-called master equation (ME), also known as the kinetic equation, the Boltzmann equation, etc. It reads as follows [204, 205]:

$$\left[\frac{\partial}{\partial X} + \frac{\partial}{\partial x} + P(x)\right] Q(x;X) = \delta(x) \int_{-\infty}^{\infty} d\xi\, P(\xi)\, Q(\xi;X)\,, \tag{10.9}$$

where $Q(x;X)$ is the distribution of the stretching x when the bottom of the sliding plate has advanced at position X, and $P(x)\Delta X$ is the fraction of contacts that break at the stretching x when the plate moves by ΔX. The latter is related to the distribution of the breaking thresholds $P_c(x)$ by

$$P(x) = P_c(x)/J_c(x)\,, \quad J_c(x) = \int_{x}^{\infty} d\xi\, P_c(\xi)\,, \tag{10.10}$$

which simply says that the fraction of the contacts that break when X increases by ΔX are those that have their thresholds between x and $x + \Delta X$ divided by the total of fraction $J_c(x)$ of contacts which are not yet broken at stretching x.

Equations (10.9) and (10.10) can be generalized to incorporate thermal effects, and can also be supplemented by another equation describing the ageing of contacts [205]. The ME can often be solved analytically, thus allowing us to describe the dependence of friction on temperature and velocity, the stick-slip motion and the transition to smooth sliding.

Smooth sliding: friction force versus velocity. The steady-state solution of the ME may be found analytically; it describes the dependence of the kinetic friction force on the sliding velocity in the smooth-sliding regime [206]. According to the second Amontons law (also known as the Coulomb law), the friction force does not depend on the sliding speed; however, this is not true in a general case. The friction force does depend on the speed—$f_k(v)$ increases with v at small velocities, reaches a maximum and then decreases. At low driving velocities the kinetic friction force increases linearly with speed—if the slider moves slowly, all contacts will break sooner or later, purely due to thermal fluctuations. The slower the slider moves, the longer time the contacts have to receive a fluctuation above the threshold, so the smaller is the friction force. The linear $f_k(v)$ dependence sometimes is treated as a (typically very high) "viscosity" of a thin lubricant film [197, 199]. At intermediate speed, the role of thermal fluctuations becomes more and more marginal, and friction is dominated by the so-called ageing effects: when a contact breaks, soon it re-forms and grows in size. This leads to a weak (logarithmic) $f_k(v)$ dependence, which is basically consistent with Amonton-Coulomb's law: the actual $f_k(v)$ dependence is hard to detect experimentally (however, see recent papers [42, 43]). Eventually at high velocities the kinetic friction reaches a maximum and starts to decrease, when sliding is so fast that no time is left for contact re-forming.

Stick-slip: elastic instability. The EQ approach also accounts for the stick-slip motion and the transition to smooth sliding [197, 199, 205]. Roughly speaking, it may be explained as follows: when the slider begins to move, the contacts start to break but they are formed again later. The main question is: do the re-formed contacts produce a force capable to compensate the externally applied driving force? If not, an elastic instability appears, and the slider will slide fast until the pulling spring force decreases enough; then the process repeats itself. This is the stick-slip regime, typical e.g. of creaking doors and squeaking brakes. If, on the other hand, the reformed contacts build up a force strong enough to compensate the driving one, the system proceeds with smooth sliding. When the ageing of contacts is taken into account, such an approach explains the transition from stick-slip to smooth sliding with the increase of driving velocity.

Interaction between the contacts. Above we considered the model with a rigid slider where the contacts do not interact. In reality, elasticity of the sliders leads to contact-contact interaction—when one of contacts breaks, the forces on surrounding contacts should increase by some δf. Numerics [207] shows that $\delta f(r)$ decays with the distance r from the broken contact as $\delta f(r) \propto r^{-1}$ at short distances $r \ll \lambda$, and as $\delta f(r) \propto r^{-3}$ at long distances, where $\lambda \sim a^2 E/k$ is the elastic correlation

length [207, 208] expressed in terms of the slider Young modulus E and the average distance a between the contacts. The model may then be simplified by considering the slider as rigid over distances $r \lesssim \lambda$, and treating the contacts within each λ-area as one effective λ-contact with the parameters determined by a corresponding solution of the ME. Numerics also shows that most of the intercontact extra force arises in front and behind the broken contact, which means that the interface may be approximately considered as an effective 1D chain of λ-contacts.

Self-healing crack as a solitary wave. If the λ-contacts do not undergo elastic instability, then a local perturbation spreads smoothly over the interface. Otherwise, if it is subject to the elastic instability, i.e. if it breaks and slides at a certain threshold stress, then the nearest neighboring λ-contacts have a good chance to break too, and a sequence of breaks will propagate through the interface like in a domino effect. In the latter case the dynamics of the chain of λ-contacts can be addressed with the help of the FK model (see Sect. 10.3), where the sinusoidal substrate potential is replaced by a sawtooth-like potential of periodically repeated inclined pieces [209]. With this approach one can find analytically the maximum and minimum shear stress for crack propagation (the latter corresponds to the Griffith threshold) as well as the crack velocity as function of the applied stress. When the shear stress is uniform and a λ-contact breaks somewhere along the chain, two self-healing cracks propagate from the initial break point in opposite directions as solitary waves similarly to the kink-antikink pair of Fig. 10.3 until they reach the boundary or meet with another crack created somewhere else.

Onset of sliding. When an elastic slider is pushed from its trailing edge as in the experiments by Fineberg et al. [210–212], the nonuniform shear stress is maximal at the trailing edge and falls off with distance inside the block. As the pushing force is increased, the most likely starting event is the breaking of the leftmost λ-contact. Due to interaction between the contacts, this will result in the increase of the stress on the second λ-contact which will break too, and so on until the self-consistent stress will occur below the threshold. Thus, the self-healing crack created at the trailing edge, propagates through the interface over some distance Λ (which can be found analytically [213]), removing the stress at its tail but creating an extra stress in the region ahead. With a further increase of the pushing force, a second crack is created at the trailing edge. This second crack triggers the previously formed stressed state and propagates further to some distance, and so on until the cracks will reach the leading edge of the system. Besides, when the crack stops, the stress on the contact at crack's tip is close to the threshold value. When the trailing edge moves, the increasing stress is transferred through the interface and affects the contacts ahead of the arrested crack to break one by one. Therefore, between the propagation of fast cracks, the system exhibits a slow dynamics—a creep-like crack motion which may correspond to the slow crack mode observed experimentally [213]. The most important issue here is that, when the cracks propagate through the interface, the whole slider undergoes slight slips, the so-called precursors, which may be detected and used to predict the large earthquake [198, 213].

Real earthquakes. The EQ-like model described above has been invented initially by Burridge and Knopoff [214] (the famous BK spring-and-block model) to explain

real earthquakes, not friction. The physics of these two problems—friction and earthquakes—is essentially similar and differs mainly by the spatio-temporal scale: nanometers and seconds to hours in tribology in comparison to kilometers and years to centuries in geology. Real earthquakes are characterized by two laws—the Gutenberg-Richter (GR) law [215, 216] and the Omori law [217]. Both these laws are empirical, found through long-term statistical observations, and there are no more or less articulate explanations of these laws yet. EQ-like models discussed above may be one of the approaches which would allow to explain both laws. In particular, the GR law may be explained as emerging due to contact ageing [218], while the Omori law may be associated with a finite distance of crack propagation—after a large earthquake, not all the stress is released, but a part of it is stored at a distance Λ from the main shock. The eventual goal of these studies is to be able to predict earthquakes, but this has not been achieved yet.

10.6 Conclusions

Among provisional conclusions of this chapter we mention:

(i) All levels of modeling and simulation can be highly informative and predictive, provided that specific limitations are kept clear.

(ii) The simple PT and FK models are extremely useful in understanding several aspects of nanofriction, including superlubricity.

(iii) MD simulations are powerful and informative for qualitative and even quantitative descriptions of atomic stick-slip and high-speed smooth sliding. An advantage of MD is also that it can address extreme or otherwise unusual frictional situations, still unexplored experimentally [134, 219–221]. The main open problem in MD is the size and time limitations, in particular the complete omission of slow, logarithmic relaxations and ageing.

(iv) Earthquake-like (multi-contact) models are instrumental in describing mesoscopic friction and fracture, especially in bridging the gap between nano- and macroscale friction [222].

Among open problems we mention prospective mechanisms for the control of friction. One approach involves using natural or artificially-induced oscillations obtained by small normal or lateral mechanical vibrations which may, when applied at suitable frequency and amplitude ranges, help dislodge a contact, increasing surface mobility and diffusion and thus reducing friction and wear. Flexibility and accessibility are the main relevant features of this approach, since frictional properties can be tuned continuously by the frequency and the amplitude of the applied vibrations. This effect has been demonstrated experimentally with AFM [30, 50, 51, 223] and in sheared confined system [224–226] as well as numerically with MM [48, 49, 227, 228] approaches and with atomistic MD [229, 230]. Despite these promising numerical and experimental contributions, a realistic multi-contact analysis accounting for the friction dependence on vibrations is still to some extent lacking.

Another idea to control friction is to employ a substrate undergoing a phase transition. While it is obvious that friction will change in the presence of a phase transition,

it is more subtle to qualify and quantify precisely the effect. Surprisingly perhaps for such a basic concept, there are essentially no experimental data available—and no theory either. A PT-like MD nanofrictional simulation based on a point slider over a 2D model substrate with a built-in structural displacive transition recently predicted that stick-slip friction should actually peak near the substrate critical temperature [231].

Another interesting and practically important topic is AFM manipulation of surface-deposited clusters, which can serve as a useful method to measure the interfacial friction of structurally well-defined contacts of arbitrary size and material combinations. Indeed, one of the remarkable experimental observations of the last decade concerns the unexpected ability of relatively large metal clusters to execute friction-free motions and even long skids with size and shape conservation [232–236]. Gold clusters, comprising typically hundreds of atoms, have been repeatedly observed to diffuse on highly oriented pyrolytic graphite (HOPG) surfaces with surprisingly large thermally activated diffusion coefficients already at room temperature; a similar behavior was reported also for larger antimony clusters. Here, MD simulations are extremely useful in understanding depinning, diffusion, and frictional mechanisms of clusters on surfaces. MD simulations of the diffusive regime have shown the possible coexistence of sticking periods and of long jumps, reminiscent of so-called Levy flights [134, 237–239]. The sticking lasts so long as the cluster-substrate surfaces are orientationally aligned, and the long sliding jumps occur when a thermal fluctuation rotates the cluster destroying the alignment [134].

It is worth mentioning in closing that there remain fully open problems at the very basic theoretical level: we still do not have a proper theory of friction, namely a theory where the frictional work could be calculated quantitatively (not just simulated) in all cases—they are the majority—where linear-response theory is inapplicable. There are also many more outstanding challenges left in nanofriction, such as:

- The ageing of surface contacts at the nano and macroscales.
- Role of wear and adhesion at the nanoscale.
- Role of ball-shaped molecules (C_{60}) as additives to traditional lubricants [240–242].
- Rolling nanofriction: besides the known case of nanotubes: does it exist, and how to distinguish between rolling and sliding?
- Friction in dislocations and in granular systems.
- Water-based lubricants [243, 244].
- Friction in biological systems (motor proteins, cells membranes and pores, etc.).

Lively progress along these and newer lines is to be expected in the near future.

Acknowledgments We gratefully acknowledge helpful discussion and collaboration with A. Benassi, A.R. Bishop, R. Capozza, R. Guerra, B. Persson, M. Peyrard, G.E. Santoro, J. Scheibert, E. Tosatti, and M. Urbakh. This work is partly funded by the the Italian Ministry of University and Research, the Swiss National Science Foundation Sinergia CRSII2_136287, and the ERC Advanced Grant No. 320796-MODPHYSFRICT. O.B. acknowledges partial support from the EGIDE/Dnipro grant No. 28225UH and from the NASU "RESURS" program.

References

1. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, Nature (London) **430**, 525 (2004)
2. V. Bormuth, V. Varga, J. Howard, E. Schäffer, Science **325**, 870 (2009)
3. C.H. Scholz, Nature (London) **391**, 37 (1998)
4. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Oxford University Press, New York, 1950)
5. R.W. Carpick, M. Salmeron, Chem. Rev. **97**, 1163 (1997)
6. G. Binnig, C.F. Quate, Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986)
7. J.N. Israelachvili, Surf. Sci. Rep. **14**, 109 (1992)
8. J. Krim, A. Widom, Phys. Rev. B **38**, 12184 (1988)
9. J. Krim, Sci. Am. **275**, 74 (1996)
10. A. Vanossi, N. Manini, M, Urbakh, S. Zapperi, E. Tosatti. Rev. Mod. Phys. **85**, 529 (2013)
11. B.N.J. Persson, *Sliding Friction, Physical Properties and Applications* (Springer, Berlin, 2000)
12. C. Mathew Mate, Tribology on the Small Scale: A Bottom Up Approach to Friction, Lubrication, and Wear (Mesoscopic Physics and Nanotechnology), (Oxford University Press, Oxford, 2008)
13. E. Gerde, M. Marder, Nature (London) **413**, 285 (2001)
14. C. Marone, Annu. Rev. Earth Planet. Sci. **26**, 643 (1998)
15. S. Medyanik, W.K. Liu, I.-H. Sung, R. W. Carpick, Phys. Rev. Lett. **97**, 136106 (2006)
16. L. Prandtl, Z. Angew, Math. Mech. **8**, 85 (1928)
17. G.A. Tomlinson, Philos. Mag. **7**, 905 (1929)
18. Y. Sang, M. Dubé, M. Grant, Phys. Rev. Lett. **87**, 174301 (2001)
19. O.M. Dudko, A.E. Filippov, J. Klafter, M. Urbakh, Chem. Phys. Lett. **352**, 499 (2002)
20. C.E. Maloney, D.J. Lacks, Phys. Rev. E **73**, 061106 (2006)
21. J. Kurkijarvi, Phys. Rev. B **6**, 832 (1972)
22. A. Garg, Phys. Rev. B **51**, 15592 (1995)
23. W.L. Johnson, K.A. Samwer, Phys. Rev. Lett. **95**, 195501 (2005)
24. D.J. Lacks, J. Willis, M.P. Robinson, J. Phys. Chem. B **114**, 10821 (2010)
25. R. Berkovich, S. Garcia-Manyes, M. Urbakh, J. Klafter, J.M. Fernandez, Biophys. J. **98**, 2692 (2010)
26. J.S. Helman, W. Baltensperger, J.A. Holyst, Phys. Rev. B **49**, 3831 (1994)
27. R. Gilmore, *Catastrophe Theory for Scientists and Engineers* (Wiley, New York, 1981)
28. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)
29. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, H.-J. Güntherodt, Phys. Rev. Lett. **84**, 1172 (2000)
30. E. Riedo, E. Gnecco, R. Bennewitz, E. Meyer, H. Brune, Phys. Rev. Lett. **91**, 084502 (2003)
31. P. Reimann, M. Evstigneev, Phys. Rev. Lett. **93**, 230802 (2004)
32. S.Y. Krylov, K.B. Jinesh, H. Valk, M. Dienwiebel, J.W.M. Frenken, Phys. Rev. E **71**, 65101 (2005)
33. M. Müser, Phys. Rev. B **84**, 125419 (2011)
34. O.K. Dudko, A.E. Filippov, J. Klafter, M. Urbakh, Proc. Natl. Acad. Sci. U. S. A. **100**, 11378 (2003)
35. S. Stills, R. Overney, Phys. Rev. Lett. **91**, 095501 (2003)
36. A. Schirmeisen, L. Jansen, H. Fuchs, Phys. Rev. B **71**, 245403 (2005)
37. Q. Li, Y. Dong, D. Perez, A. Martini, R.W. Carpick, Phys. Rev. Lett. **106**, 126101 (2011)
38. P. Steiner, R. Roth, E. Gnecco, A. Baratoff, S. Maier, T. Glatzel, E. Meyer, Phys. Rev. B **79**, 045414 (2009)
39. I. Szlufarska, M. Chandross, R.W. Carpick, J. Phys. D **41**, 123001 (2008)
40. A. Schirmeisen, L. Jansen, H. Holscher, H. Fuchs, Appl. Phys. Lett. **88**, 123108 (2006)
41. X. Zhao, S.R. Phillpot, W.G. Sawyer, S.B. Sinnott, S.S. Perry, Phys. Rev. Lett. **102**, 186102 (2009)

42. I. Barel, M. Urbakh, L. Jansen, A. Schirmeisen, Phys. Rev. Lett. **104**, 066104 (2010)
43. I. Barel, M. Urbakh, L. Jansen, A. Schirmeisen, Trib. Lett. **39**, 311 (2010)
44. T. Gyalog, M. Bammerlin, R. Lüthi, E. Meyer, H. Thomas, Europhys. Lett. **31**, 269 (1995)
45. R. Prioli, A.F.M. Rivas, F.L. Freire Jr, A.O. Caride, Appl. Phys. A: Mater. Sci. Proc. **76**, 565 (2003)
46. C. Fusco, A. Fasolino, Appl. Phys. Lett. **84**, 699 (2004)
47. C. Fusco, A. Fasolino, Phys. Rev. B **71**, 045413 (2005)
48. M.G. Rozman, M. Urbakh, J. Klafter, Phys. Rev. E **57**, 7340 (1998)
49. V. Zaloj, M. Urbakh, J. Klafter, Phys. Rev. Lett. **82**, 4823 (1999)
50. A. Socoliuc, E. Gnecco, S. Maier, O. Pfeiffer, A. Baratoff, R. Bennewitz, E. Meyer, Science **313**, 207 (2006)
51. M.A. Lantz, D. Wiesmann, B. Gotsmann, Nature Nanotech. **4**, 586 (2009)
52. S.Y. Krylov, J.A. Dijksman, W.A. van Loo, J.W.M. Frenken, Phys. Rev. Lett. **97**, 166103 (2006)
53. Z. Tshiprut, A.E. Filippov, M. Urbakh, J. Phys.: Condens. Matter **20**, 354002 (2008)
54. O.M. Braun, YuS Kivshar, *The Frenkel-Kontorova Model: Concepts, Methods, and Applications* (Springer, Berlin, 2004)
55. U. Dehlinger, Ann. Phys. (Leipzig) **2**, 749 (1929)
56. Y. I. Frenkel, T. A. Kontorova, Phys. Z. Sowietunion **13**, 1 (1938)
57. T.A. Kontorova, Y. I. Frenkel, Zh. Eksp. Teor. Fiz. **8**, 89 (1938)
58. T.A. Kontorova, Y.I. Frenkel, Zh. Eksp. Teor. Fiz. **8**, 1340 (1938)
59. O.M. Braun, A.R. Bishop, J. Röder, Phys. Rev. Lett. **79**, 3692 (1997)
60. A. Vanossi, J. Röder, A.R. Bishop, V. Bortolani, Phys. Rev. E **67**, 016605 (2003)
61. O.M. Braun, Surf. Sci. **230**, 262 (1990)
62. J.D. Meiss, Rev. Mod. Phys. **64**, 795 (1992)
63. S.C. Ying, Phys. Rev. B **3**, 4160 (1971)
64. J.B. Sokoloff, Phys. Rev. B **16**, 3367 (1977)
65. V.L. Pokrovsky, A.L. Talapov, Zh. Exp. Teor. Fiz. **75**, 1151 (1978)
66. S. Aubry, in *Solitons and Condensed Matter Physics*, ed. by A.R. Bishop, T. Schneider. Springer Series in Solid State Sciences, vol 8 (Springer, Berlin, 1978), p. 264–277
67. S. Aubry, P.Y. Le Daeron, Phys. D **8**, 381 (1983)
68. M. Peyrard, S. Aubry, J. Phys. C: Solid State Phys. **16**, 1593 (1983)
69. L.M. Floría, J.J. Mazo, Adv. Phys. **45**, 505 (1996)
70. Y. Braiman, J. Baumgarten, J. Jortner, J. Klafter, Phys. Rev. Lett. **65**, 2398 (1990)
71. A. Benassi, A. Vanossi, E. Tosatti, Nature Commun. **2**, 236 (2011)
72. T. Pruttivarasin, M. Ramm, I. Talukdar, H. Haeffner, New J. Phys. **13**, 075012 (2011)
73. B. Lin, B. Hu, J. Stat. Phys. **69**, 1047 (1992)
74. R.S. MacKay, *Renormalization in Area-Preserving Maps* (World Scientific, Singapore, 1993)
75. S.N. Coppersmith, D.S. Fisher, Phys. Rev. B **28**, 2566 (1983)
76. S.R. Sharma, B. Bergersen, B. Joos, Phys. Rev. B **29**, 6335 (1984)
77. L. de Seze, S. Aubry, J. Phys. C **17**, 389 (1984)
78. S. Aubry, P. Quemerais, in *Low Dimensional Electronic Properties of Molybdenum Bronzes and Oxides*, ed. by C. Schlenker (Kluwer, Dordrecht, 1984)
79. R.S. MacKay, Phys. D **50**, 71 (1991)
80. M. Hirano, K. Shinjo, Phys. Rev. B **41**, 11837 (1990)
81. M. Hirano, K. Shinjo, Surf. Sci. **283**, 473 (1993)
82. K. Shinjo, M. Hirano, Surf. Sci. **283**, 473 (1993)
83. M. Dienwiebel, G.S. Verhoeven, N. Pradeep, J.W.M. Frenken, J.A. Heimberg, H.W. Zandbergen, Phys. Rev. Lett. **92**, 126101 (2004)
84. G.S. Verhoeven, M. Dienwiebel, J.W.M. Frenken, Phys. Rev. B **70**, 165418 (2004)
85. M. Dienwiebel, N. Pradeep, G.S. Verhoeven, H.W. Zandbergen, J.W.M. Frenken, Surf. Sci. **576**, 197 (2005)
86. T. Strunz, F.-J. Elmer, Phys. Rev. E **58**, 1601 (1998)
87. T. Strunz, F.-J. Elmer, Phys. Rev. E **58**, 1612 (1998)

88. J.E. Hammerberg, B.L. Holian, J. Röder, A.R. Bishop, S.J. Zhou, Phys. D **123**, 330 (1998)
89. M. Paliy, O. Braun, T. Dauxois, B. Hu, Phys. Rev. E **56**, 4025 (1997)
90. O.M. Braun, T. Dauxois, M.V. Paliy, M. Peyrard, B. Hu, Phys. D **123**, 357 (1998)
91. T. Bohlein, J. Mikhael, C. Bechinger, Nat. Mater. **11**, 126 (2012)
92. A. Vanossi, N. Manini, E. Tosatti, P. Natl, Acad. Sci. USA **109**, 16429 (2012)
93. M. Peyrard, M. Remoissenet, Phys. Rev. B **26**, 2886 (1982)
94. M. Remoissenet, M. Peyrard, Phys. Rev. B **29**, 3153 (1984)
95. A. Vanossi, J. Röder, A.R. Bishop, V. Bortolani, Phys. Rev. E **63**, 017203 (2000)
96. T.S. van Erp, A. Fasolino, O. Radulescu, T. Janssen, Phys. Rev. B **60**, 6522 (1999)
97. D. Cule, T. Hwa, Phys. Rev. Lett. **77**, 278 (1996)
98. D. Cule, T. Hwa, Phys. Rev. B **57**, 8235 (1998)
99. R. Guerra, A. Vanossi, M. Ferrario, Surf. Sci. **601**, 3676 (2007)
100. A. Vanossi, E. Tosatti, Nature Mater. **11**, 97 (2012)
101. T. Bohlein, C. Bechinger, Phys. Rev. Lett. **109**, 058301 (2012)
102. J. Röder, J.E. Hammerberg, B.L. Holian, A.R. Bishop, Phys. Rev. B **57**, 2759 (1998)
103. B.N.J. Persson, Phys. Rev. Lett. **71**, 1212 (1993)
104. B.N.J. Persson, Phys. Rev. B **48**, 18140 (1993)
105. B.N.J. Persson, J. Chem. Phys. **103**, 3849 (1993)
106. B.N.J. Persson, Phys. Rev. B **50**, 4771 (1994)
107. R.P. Mikulla, J.E. Hammerberg, P.S. Lomdahl, B.L. Holian, Mat. Res. Soc. Symp. Proc. **522**, 385 (1998)
108. O.M. Braun, M.V. Paliy, J. Röder, A.R. Bishop, Phys. Rev. E **63**, 036129 (2001)
109. S. Maier, E. Gnecco, A. Baratoff, R. Bennewitz, E. Meyer, Phys. Rev. B **78**, 045432 (2008)
110. M. Gauthier, M. Tsukada, Phys. Rev. Lett. **85**, 5348 (2000)
111. R. Bennewitz, A.S. Foster, L.N. Kantorovich, M. Bammerlin, C. Loppacher, S. Schär, M. Guggisberg, E. Meyer, A.L. Shluger, Phys. Rev. B **62**, 2074 (2000)
112. C. Loppacher, R. Bennewitz, O. Pfeiffer, M. Guggisberg, M. Bammerlin, S. Schär, V. Barwich, A. Baratoff, E. Meyer, Phys. Rev. B **62**, 13674 (2000)
113. P.M. Hoffmann, S. Jeffery, J.B. Pethica, H. Özgür, Özer, A. Oral, Phys. Rev. Lett. **87**, 265502 (2001)
114. C. Negri, N. Manini, A. Vanossi, G.E. Santoro, E. Tosatti, Phys. Rev. B **81**, 045417 (2010)
115. M.G. Rozman, M. Urbakh, J. Klafter, Phys. Rev. Lett. **77**, 683 (1996)
116. M.G. Rozman, M. Urbakh, J. Klafter, Phys. Rev. E **54**, 6485 (1996)
117. M.H. Müser, Phys. Rev. Lett. **89**, 224301 (2002)
118. M.G. Rozman, M. Urbakh, J. Klafter, Europhys. Lett. **39**, 183 (1997)
119. M.G. Rozman, M.M. Urbakh, J. Klafter, F.-J. Elmer, J. Phys. Chem. B **102**, 7924 (1998)
120. O.M. Braun, A. Vanossi, E. Tosatti, Phys. Rev. Lett. **95**, 026102 (2005)
121. A. Vanossi, N. Manini, G. Divitini, G.E. Santoro, E. Tosatti, Phys. Rev. Lett. **97**, 056101 (2006)
122. G.E. Santoro, A. Vanossi, N. Manini, G. Divitini, E. Tosatti, Surf. Sci. **600**, 2726 (2006)
123. A. Vanossi, N. Manini, F. Caruso, G.E. Santoro, E. Tosatti, Phys. Rev. Lett. **99**, 206101 (2007)
124. N. Manini, M. Cesaratto, G.E. Santoro, E. Tosatti, A. Vanossi, J. Phys.: Condens. Matter **19**, 305016 (2007)
125. R.L. Woulaché, A. Vanossi, N. Manini, Phys. Rev. E **88**, 012810 (2013)
126. I.E. Castelli, N. Manini, R. Capozza, A. Vanossi, G.E. Santoro, E. Tosatti, J. Phys.: Condens. Matter **20**, 354005 (2008)
127. I.E. Castelli, R. Capozza, A. Vanossi, G.E. Santoro, N. Manini, E. Tosatti, J. Chem. Phys. **131**, 174711 (2009)
128. C. Drummond, J. Israelachvili, Phys. Rev. E **63**, 041506 (2001)
129. C. Drummond, J. Israelachvili, P. Richetti, Phys. Rev. E **67**, 066110 (2003)
130. M. Weiss, F.-J. Elmer, Phys. Rev. B **53**, 7539 (1996)
131. M. Weiss, F.-J. Elmer, Z. Phys. B **104**, 55 (1997)
132. M. Igarashi, A. Nator, J. Nakamura, Phys. Rev. B **78**, 165427 (2008)
133. W.K. Kim, M.L. Falk, Phys. Rev. B **80**, 235428 (2009)

134. R. Guerra, U. Tartaglino, A. Vanossi, E. Tosatti, Nature Mater. **9**, 634 (2010)
135. M.O. Robbins, M.H. Müser, *Modern Tribology Handbook*, ed. by B. Bhushan (CRC Press, Boca Raton, 2001), p. 717
136. M.H. Müser, in *Computer Simulations in Condensed Matter Systems: From Materials to Chemical Biology*, vol. 2, ed. by M. Ferrario, G. Ciccotti, K. Binder. Lecture Notes in Physics, vol. 704 (Springer, Heidelberg, 2006), p. 65
137. R. Car, M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985)
138. C.H. Xu, C.Z. Wang, C.T. Chan, K.M. Ho, J. Phys.: Condens. Matter **4**, 6047 (1992)
139. B.J. Garrison, D. Srivastava, Ann. Rev. Phys. Chem. **46**, 373 (1995)
140. D.W. Brenner, O.A Sherendova, A. Areshkin, in *Quantum-Based Analytic Interatomic Forces and Materials Simulation*, ed. by K.B. Lipkowitz, D.B. Boyd. Reviews in Computational Chemistry (VCH Publishers, New York, 1998), p. 213
141. S.J. Weiner, P.A. Kollman, D.T. Nguyen, D.A. Case, J. Comp. Chem. **7**, 230 (1986)
142. J.H. Los, L.M. Ghiringhelli, E.J. Meijer, A. Fasolino, Phys. Rev. B **72**, 214102 (2005)
143. L.M. Ghiringhelli, J.H. Los, A. Fasolino, E.J. Meijer, Phys. Rev. B **72**, 214103 (2005)
144. S.J. Stuart, J. Chem. Phys. **112**, 6472 (2000)
145. D.W. Brenner, O.A. Shenderova, J.A. Harrison, S.J. Stuart, B. Ni, S.B. Sinnott, J. Phys.: Condens. Matter **14**, 783 (2002)
146. A.C.T. van Duin, S. Dasgupta, F. Lorant, W.A. Goddard III, J. Phys. Chem. A **105**, 9396 (2001)
147. M.S. Tomassone, J.B. Sokoloff, A. Widom, J. Krim, Phys. Rev. Lett. **79**, 4798 (1997)
148. O.M. Braun, M. Peyrard, Phys. Rev. E **63**, 046110 (2001)
149. O.M. Braun, R. Ferrando, Phys. Rev. E **65**, 061107 (2002)
150. V.B. Magalinskii, Sov. Phys. JETP **9**, 1381 (1959)
151. R.J. Rubin, J. Math. Phys. **1**, 309 (1960)
152. R. Zwanzig, J. Stat. Phys. **9**, 215 (1973)
153. X. Li, W. E, Phys. Rev. B 76, 104107 (2007)
154. L. Kantorovich, Phys. Rev. B **78**, 094304 (2008)
155. L. Kantorovich, N. Rompotis, Phys. Rev. B **78**, 094305 (2008)
156. A. Benassi, A. Vanossi, G.E. Santoro, E. Tosatti, Phys. Rev. B **82**, 081401 (2010)
157. A. Benassi, A. Vanossi, G.E. Santoro, E. Tosatti, Tribol. Lett. **48**, 41 (2012)
158. O.M. Braun, N. Manini, E. Tosatti, Phys. Rev. Lett. **110**, 085503 (2013)
159. O.M. Braun, J. Röder, Phys. Rev. Lett. **88**, 096102 (2002)
160. B. Luan, M.O. Robbins, Phys. Rev. Lett. **93**, 036105 (2004)
161. O.M. Braun, M. Peyrard, V. Bortolani, A. Franchini, A. Vanossi, Phys. Rev. E **72**, 056116 (2005)
162. O.M. Braun, A.G. Naumovets, Surf. Sci. Rep. **60**, 79 (2006)
163. A.F. Voter, F. Montalenti, T.C. Germann, Annu. Rev. Mater. Res. **32**, 321 (2002)
164. Y. Mishin, A. Suzuki, B.P. Uberuaga, A.F. Voter, Phys. Rev. B **75**, 224101 (2007)
165. W.K. Kim, M.L. Falk, Phys. Rev. B **84**, 165422 (2011)
166. J.H. Dieterich, B.D. Kilgore, Pure Appl. Geophys. **143**, 283 (1994)
167. A. Ruina, J. Geophys. Res. **88**, 10359 (1983)
168. F. Heslot, T. Baumberger, B. Perrin, B. Caroli, C. Caroli, Phys. Rev. E **49**, 4973 (1994)
169. Q. Li, T.E. Tullis, D. Goldsby, R.W. Carpick, Nature **480**, 233 (2011)
170. E. McGee, R. Smith, S.D. Kenny, Int. J. Mater. Res. **98**, 430 (2007)
171. B.Q. Luan, S. Hyun, J.F. Molinari, N. Bernstein, M.O. Robbins, Phys. Rev. E **74**, 046710 (2006)
172. W.E, W.Q. Ren, E. Vanden-Eijnden, J. Comput. Phys. 228, 5437 (2009)
173. A.Z. Szeri, *Modern Tribology Handbook*, ed. by B. Bhushan (CRC Press, Boca Raton, 2001), p. 384
174. H. Yosbizawa, J. Israelachvili, J. Phys. Chem. **97**, 11300 (1993)
175. J. Gao, W.D. Luedtke, U. Landman, Phys. Rev. Lett. **79**, 705 (1997)
176. J. Gao, W.D. Luedtke, U. Landman, J. Chem. Phys. **106**, 4309 (1997)
177. B.N.J. Persson, E. Tosatti, Phys. Rev. B **50**, 5590 (1994)

178. B.N.J. Persson, F. Mugele, J. Phys.: Condens. Matter **16**, R295 (2004)
179. U. Tartaglino, I.M. Sivebaek, B.N.J. Persson, E. Tosatti, J. Chem. Phys. **125**, 014704 (2006)
180. M.J. Stevens, M.O. Robbins, Phys. Rev. E **48**, 3778 (1993)
181. P.A. Thompson, M.O. Robbins, Science **250**, 792 (1990)
182. O.M. Braun, M. Peyrard, Phys. Rev. E **68**, 011506 (2003)
183. A.V. Zhukov, M.V. Paliy, O.M. Braun, T.F. George, Phys. Lett. A **361**, 437 (2007)
184. C.D. Lorenz, M. Chandross, G.S. Grest, J. Adhesion Sci. Technol. **24**, 2453 (2010)
185. C.D. Lorenz, M. Chandross, J.M.D. Lane, G.S. Grest, Model. Simul. Mat. Sci. Eng. **18**, 034005 (2010)
186. M. Chandross, E.B. Webb, M.J. Stevens, G.S. Grest, S.H. Garofalini, Phys. Rev. Lett. **93**, 166103 (2004)
187. M. Chandross, C.D. Lorenz, M.J. Stevens, G.S. Grest, Langmuir **24**, 1240 (2008)
188. M.H. Müser, M.O. Robbins, Phys. Rev. B **61**, 2335 (2000)
189. O.M. Braun, M. Paliy, S. Consta, Phys. Rev. Lett. **92**, 256103 (2004)
190. A.I. Volokitin, B.N.J. Persson, Rev. Mod. Phys. **79**, 1291 (2007)
191. A.I. Volokitin, B.N.J. Persson, Phys. Rev. Lett. **106**, 094502 (2011)
192. O.M. Braun, N. Manini, E. Tosatti, Phys. Rev. B **78**, 195402 (2008)
193. O.M. Braun, N. Manini, Phys. Rev. E **83**, 021601 (2011)
194. N. Manini, O.M. Braun, Phys. Lett. A **375**, 2946 (2011)
195. B.N.J. Persson, Phys. Rev. B **51**, 13568 (1995)
196. A.E. Filippov, J. Klafter, M. Urbakh, Phys. Rev. Lett. **92**, 135503 (2004)
197. O.M. Braun, E. Tosatti, Europhys. Lett. **88**, 48003 (2009)
198. O.M. Braun, I. Barel, M. Urbakh, Phys. Rev. Lett. **103**, 194301 (2009)
199. O.M. Braun, E. Tosatti, Philos. Mag. **91**, 3253 (2011)
200. R. Capozza, M. Urbakh, Phys. Rev. B **86**, 085430 (2012)
201. R.F. Smalley, D.L. Turcotte, S.A. Solla, J. Geophys. Res. **90**, 1894 (1985)
202. W.I. Newman, A.M. Gabrielov, Int. J. Fracture **50**, 1 (1991)
203. W.I. Newman, S.L. Phoenix, Phys. Rev. E **63**, 021507 (2001)
204. O.M. Braun, M. Peyrard, Phys. Rev. Lett. **100**, 125501 (2008)
205. O.M. Braun, M. Peyrard, Phys. Rev. E **82**, 036117 (2010)
206. O.M. Braun, M. Peyrard, Phys. Rev. E **83**, 046129 (2011)
207. O.M. Braun, M. Peyrard, D.V. Stryzheus, E. Tosatti, Collective effects at frictional interfaces. Tribol. Lett. **48**, 11 (2012)
208. C. Caroli, Ph Nozieres, Eur. Phys. J. B **4**, 233 (1998)
209. O.M. Braun, M. Peyrard, Phys. Rev. E **85**, 026111 (2012)
210. S.M. Rubinstein, G. Cohen, J. Fineberg, Nature (London) **430**, 1005 (2004)
211. S.M. Rubinstein, G. Cohen, J. Fineberg, Phys. Rev. Lett. **96**, 256103 (2006)
212. S.M. Rubinstein, I. Barel, Z. Reches, O.M. Braun, M. Urbakh, J. Fineberg, Pure Appl. Geophys. **168**, 2151 (2011)
213. O.M. Braun, M. Peyrard, J. Scheibert, Tribol. Lett. (2014)
214. R. Burridge, L. Knopoff, Bull. Seismol. Soc. Am. **57**, 341 (1967)
215. B. Gutenberg, C.F. Richter, Bull. Seismol. Soc. Am. **46**, 105 (1954)
216. B. Gutenberg, C.F. Richter, Ann. Geophys. **9**, 1 (1956)
217. F. Omori, J. Coll. Sci. Imp. Univ. Tokyo **7**, 111 (1894)
218. O.M. Braun, M. Peyrard, Phys. Rev. E **87**, 032808 (2013)
219. U. Tartaglino, T. Zykova-Timan, F. Ercolessi, E. Tosatti, Phys. Repts. **411**, 291 (2005)
220. T. Zykova-Timan, D. Ceresoli, U. Tartaglino, E. Tosatti, Phys. Rev. Lett. **94**, 176105 (2005)
221. T. Zykova-Timan, D. Ceresoli, E. Tosatti, Nature Mater. **6**, 231 (2007)
222. O.M. Braun, Tribol. Lett. **39**, 283 (2010)
223. S. Jeon, T. Thundat, Y. Braiman, Appl. Phys. Lett. **88**, 214102 (2006)
224. M. Heuberger, C. Drummond, J.N. Israelachvili, J. Phys. Chem. B **102**, 5038 (1998)
225. L. Bureau, T. Baumberger, C. Caroli, Phys. Rev. E **62**, 6810 (2000)
226. A. Cochard, L. Bureau, T. Baumberger, Trans. ASME **70**, 220 (2003)
227. Z. Tshiprut, A.E. Filippov, M. Urbakh, Phys. Rev. Lett. **95**, 016101 (2005)

228. R. Guerra, A. Vanossi, M. Urbakh, Phys. Rev. E **78**, 036110 (2008)
229. J. Gao, W. Luedtke, U. Landman, J. Phys. Chem. B **102**, 5033 (1998)
230. R. Capozza, A. Vanossi, A. Vezzani, S. Zapperi, Phys. Rev. Lett. **103**, 085502 (2009)
231. A. Benassi, A. Vanossi, G.E. Santoro, E. Tosatti, Phys. Rev. Lett. **106**, 256102 (2011)
232. L. Bardotti, P. Jensen, A. Hoareau, M. Treilleux, B. Cabaud, A. Perez, F. Cadete, Santos Aires. Surf. Sci. **367**, 276 (1996)
233. J. Brndiar, R. Turanský, D. Dietzel, A. Schirmeisen, I. Štich, Nanotechnology **22**, 085704 (2011)
234. D. Dietzel, C. Ritter, T. Mönninghoff, H. Fuchs, A. Schirmeisen, U.D. Schwarz, Phys. Rev. Lett. **101**, 125505 (2008)
235. G. Paolicelli, K. Mougin, A. Vanossi, S. Valeri, J. Phys.: Condens. Matter **20**, 354011 (2008)
236. G. Paolicelli, M. Rovatti, A. Vanossi, S. Valeri, Appl. Phys. Lett. **95**, 143121 (2009)
237. W. Luedtke, U. Landman, Phys. Rev. Lett. **82**, 3835 (1999)
238. L.J. Lewis, P. Jensen, N. Combe, J.-L. Barrat, Phys. Rev. B **61**, 16084 (2000)
239. Y. Maruyama, Phys. Rev. B **69**, 245408 (2004)
240. O.M. Braun, Phys. Rev. Lett. **95**, 126104 (2005)
241. O.M. Braun, E. Tosatti, J. Phys.: Condens. Matter **20**, 354007 (2008)
242. O.M. Braun, E. Tosatti, Phil. Mag. Lett. **88**, 509 (2008)
243. M. Paliy, O.M. Braun, S. Consta, Tribol. Lett. **23**, 7 (2006)
244. M. Paliy, O.M. Braun, S. Consta, J. Phys. Chem. C **116**, 8932 (2012)

Chapter 11
Theoretical Studies of Superlubricity

Martin H. Müser

Abstract In simulations, materials properties and boundary conditions can be varied at will, and the resulting changes in both macroscopic variables and the dynamics of individual atoms can be observed. This allows one to study systematically the effects of many different factors on friction and wear at the nano-scale. In this chapter, we describe how to set up meaningful simulations in the context of nanotribology and summarize some key results obtained so far.

11.1 Introduction

Every-day experience tells us that a finite threshold force, namely the static friction force, F_s, has to be overcome whenever we want to initiate lateral motion of one solid body relative to another. In order to maintain measurable motion, a force equal to or larger than the kinetic friction, F_k, has to be applied. Conversely, when attempting to drag a solid through a fluid medium, there is no threshold to initiate or another one to maintain the sliding motion. Instead, one only needs to counteract friction forces linear in the (final) sliding velocity v_0. It came as a surprise when Hirano and Shinjo suggested that kinetic friction between solids in ultra-high vacuum may essentially disappear as well [1, 2]. While their suggestion of super-low static friction, or superlubricity, contradicts our intuition based on every-day experience, it does not necessarily contradict classical mechanics. If the slider and substrate have homogeneous surfaces and wear and plastic deformation are negligible, then one may expect the same (free) energy at the beginning of the sliding process as at its end, because of translational invariance. Consequently, no work would have to be done on the system implying the possibility of very low friction. Not only the fundamental or theoretical considerations evolving around superlubricity are intriguing, but even more so the possibilities to exploit it technologically. For instance, in small mechanical devices,

M.H. Müser (✉)
Jülich Supercomputing Centre, Forschungszentrum Jülich, D52425 Jülich, Germany
e-mail: m.mueser@fz-juelich.de

M.H. Müser
Universität des Saarlandes,Campus C3, 66123 Saarbrücken, Germany

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_11

friction-induced wear and heat often cause the main limits to further miniaturization. These limits could be overcome if superlubric surfaces could be designed. Achieving technological exploitation of superlubricity will certainly benefit from a thorough theoretical understanding.

One of the key ingredient of (ideal) superlubricity is that all atoms within each solid must move in a correlated fashion, i.e., such that each atom in a solid only has one mechanically stable site in the vicinity of its current position while the center of mass of one solid is displaced with respect to that of another one. Ideally, the solids move like rigid plates and have no energy-dissipating boundary lubricant immersed between them. Significant heat could not be produced in this scenario, because there would be no random stick-slip motion of atoms, as it occurs, for instance when defects, such as dislocations, are created. Unless the surface corrugation of both solids matched by design or by chance, there could be no interlocking and hence no macroscopic static friction. Of course, solids do not behave like rigid plates but can be deformed. This makes it possible for two solids to interlock at the microscopic scale. It is interesting to note that Coulomb discussed elastic deformation as a potential scenario to lead to solid friction [3]: ". . .ou bien il faut supposer que les molécules des surfaces des deux plans en contact contracte, par leur proximeté, une cohérence qu'il faut vaincre pour produire le movement."[1] Today's pursuit for superlubricity still very much evolves around the question of how one can avoid the "coherence" of atoms or asperities in contacts between two surfaces. During the beginning of the last century, works by Prandtl [4] and Tomlinson [5] showed on a fundamental level how the "contraction" of surfaces may not only induce interlocking and hence static friction but also kinetic friction: Whenever individual atoms become mechanically unstable during sliding, they will "pop" into the next available potential energy minimum. In this process, lattice vibrations will be excited, whose energy will be irreversibly lost as heat. The argument applies to collective degrees of freedom as well [6]. The intriguing conclusion from Hirano and Shinjo's work is that the energy-dissipating "pops" do not have to be expected when two atomically smooth surfaces are in contact.

Since Hirano and Shinjo's pioneering works in the early 1990s, a lot of progress has been made on superlubricity by both theoreticians and experimentalists. The chapter by Dienwiebel and Frenken in this book focuses on the experimental aspects. This chapter is concerned with an overview of the theoretical concepts related to superlubricity. We describe these without any mathematical formalism first and then sketch the quantitative analysis. We also review some key computational studies of superlubric interfaces. One of the important messages resulting from many realistic computer simulations is that quite a few materials, or pairs of materials, may be good candidates to show superlow friction. The main limitation in real life appears to be to produce surfaces that are as flat and as flawless as those in the virtual computer experiments.

[1] . . . or one has to assume that the surface molecules of the two opposing planes contract due to their proximity into a coherence, which needs to be overcome to produce motion.

The remainder of this chapter is organized as follows: In Sect. 11.2, we first present some purely analytical concepts of superlubricity. This section opens with an attempt to define superlubricity, because there is no generally accepted definition as yet. The definition is based on measurable quantities and can thus serve theoreticians and experimentalists alike. The theory section, however, is mainly concerned with the analysis of two rigid solid bodies in contact as well as with the question of how elastic instabilities (as best described in the Prandtl-Tomlinson model) and long-range elastic deformations (as most easily introduced in the Frenkel-Kontorova model) alter the rigid-body picture. This includes a short treatment of drag forces in the superlubric regime. Section 11.3 gives an overview of computer simulations relevant to the topic. The section is subdivided according to the type of materials studied by computer experiments. The analyzed systems range from "generic materials" such as Lennard Jonesium and "finite-element materials" with excluded volume interaction to "realistic materials," including bare metals and hydrogen-terminated surfaces. One of the advantages of computer simulations is that geometries can be designed at will, allowing one to study effects of surface alignment, roughness at the nanometer scale, contamination, etc. with more rigor than with analytical or even experimental methods. Conclusions are drawn in Sect. 11.4.

11.2 Theory

The microscopic justification for the possibility of the virtual absence of lateral forces between solids can be supported by the following argument: There are as many bumps (or atoms) in the substrate pushing the slider to the right as there are surface irregularities in the substrate pushing the slider to the left. Hence, statistically speaking, there is the possibility of an almost perfect annihilation of lateral forces. In this section, we will investigate the nature of this process in more detail and explore how elasticity affects this picture. First, elasticity will lead to dissipative forces which behave similarly to the drag forces that a solid experiences when moving through a fluid. Thus, the interface may remain superlubric despite the abilities of the solids to deform. However, once the solids become sufficiently compliant, elasticity will destroy superlubricity and induce finite pinning forces. The concept of drag forces due to elasticity in the solids will be discussed within the framework of linear-response theory. As for the breaking of superlubricity, particular attention will be given to the Prandtl-Tomlinson model and the Frenkel-Kontorova model. We will also be concerned with the role of thermal fluctuations and mechanisms other than elastic instabilities that can lead to instabilities and thus to Coulomb-type friction.

11.2.1 Definition of Superlubricity

Defining superlubricity is not an easy task. There is no formation of a Bose Einstein condensate or of Cooper pairs as is the case in superfluids or superconductors. The mobility of solids does not diverge to infinity when they are superlubric, unlike the

mobility of helium atoms in a superfluid or that of electrons in a superconductor. Also, the friction forces between two solids in sliding motion remain finite at finite velocities v_0. As a side comment, one can note that the friction between two incommensurate *supercrystals* (i.e., a crystal in which a finite fraction of atoms form a Bose-Einstein condensate) would probably disappear completely. Because, in general, there is no true divergence of the mobility of the solids, it will be necessary to use an ad-hoc definition of superlubricity such as the kinetic friction coefficient $\mu_\mathrm{k} = F_\mathrm{k}/L$ must be less than, say, 10^{-3}.

Although μ_k may appear almost constant in many cases for a large range of loads L and sliding velocities v_0, μ_k will ultimately be a function of both L and v_0, e.g., μ_k may vanish at very small values of v_0 due to thermal activation. Therefore, it will be necessary to make two more ad-hoc assumptions on how to choose L and v_0. The (local) load should be high, but not as high as to plastically deform the system, e.g., in the order of 10 % of the indentation hardness of the softer material in contact. v_0 should also be sufficiently high so that the system cannot relax stress through creep or thermal relaxation, i.e., v_0 should be much larger than 0.1 μm/s. However, v_0 should certainly be small compared to the speed of sound c, which is in the order of 10^3–10^4 m/s in solids. A reasonable velocity where to evaluate superlubricity is half way between these two extreme, i.e., at $v_0 \approx 0.01$ m/s.

It is important to keep in mind that the kinetic friction coefficient μ_k and the static friction coefficient μ_s are not necessarily correlated. Kinetic friction is due to hysteresis while static friction is due to energy barriers. In this chapter we will encounter some model systems that have finite energy barriers and hence exhibit finite static friction, although hysteresis effects are negligible leading to an absence of kinetic friction. Given these dissipation-free systems, one may as well define superlubricity as the absence of (mechanical) hysteresis. For example, significant static friction was observed with an atomic force microscope in a recent experiment by Socoliuc et al. [7], yet, kinetic friction was very small. More details of that study and other experiments will be given in the chapter by Dienwiebel and Frenken.

11.2.2 Cancellation of Lateral Forces. Symmetry Considerations

When studying lateral forces exerted between two surfaces, symmetry is crucial. For instance, as will be shown in more formal detail further below, two flat, rigid, commensurate solids, i.e., solids that systematically share a common periodicity, have a friction coefficient that is independent of the area of contact, or, more generally speaking, independent of the number N of atoms in direct contact with the substrate. The reason is that the forces exerted on the individual atoms in the slider add up in a systematic fashion in commensurate interfaces. Conversely, for flat, rigid, incommensurate solids there is a systematic annihilation of lateral forces similar to the destructive interference in optics. This leads to friction coefficients that vanish linearly with the area of contact. For disordered surfaces, lateral forces have random direction or random sign. Consequently they add up stochastically so that the

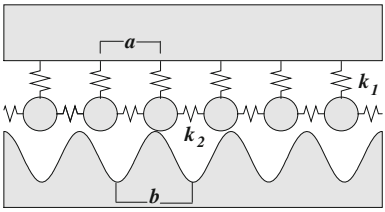


Fig. 11.1 Schematic representation of two solid bodies in contact. The substrate is considered rigid and to provide an external potential that is periodic in the lattice constant b. In the slider, each atom is coupled with a spring of stiffness k_1 to its ideal lattice site and with a spring of stiffness k_2 to its neighbor. Both springs act in lateral direction, i.e., they both lead to a restoring force linear in the displacement parallel to the interface

absolute lateral force only grows with $\sqrt{N}$ resulting in $\mu_s \propto 1/\sqrt{N}$, if we assume constant normal stress leading to $L \propto N$. The geometric arguments presented here are generally known among theoreticians, though it is not clear who presented them first and no original literature is known to the author. It may yet be beneficial to outline the concepts in more detail than in previous reviews [8, 9].

Let us start by making a minimalist model of two flat, crystalline solids in contact. It consists of a rigid substrate with lattice constant b and a slider with lattice constant a. A sketch is shown in Fig. 11.1. The surface atoms in the slider are connected to their lattice sites by springs of stiffness k_1 and to their neighbors by springs of stiffness k_2. To explore the effects of symmetry to a zero degree order, we will assume that both types of springs are infinitely stiff. In later sections, we will allow the springs to be flexible.

To model the effect of corrugation on the motion in lateral direction, it is often assumed that the potential energy in a system such as the one shown in Fig. 11.1 only depends on the lateral direction x and that the normal deflection of the atoms can be neglected. For purely periodic systems, in which the normal position of the atoms is assumed to be fixed, the potential energy of the slider V in the substrate potential can be expressed as a Fourier sum of the form [10].

$$V = \sum_{m=-\infty}^{\infty} \tilde{V}(G_m) \sum_{n=0}^{N-1} e^{iG_m(x_0+na)}, \tag{11.1a}$$

which can be simplified to yield

$$= \sum_{m=-\infty}^{\infty} \tilde{V}(G_m)\, e^{iG_m x_0} \times \begin{cases} N & \text{if } \frac{aG_m}{2\pi} \text{ is an integer} \\ \frac{1-e^{iNG_m a}}{1-e^{iG_m a}} & \text{otherwise.} \end{cases} \tag{11.1b}$$

Here, G_m are the reciprocal lattice vectors of the substrate with $G_m = 2\pi m/b$, $\tilde{V}(G_m)$ are the Fourier expansion coefficients of the substrate potential, and $x_0 + na$ is the position of atom n in the chain. One must ensure V to be a real function by requiring that $\tilde{V}(-G_m) = \tilde{V}^*(G_m)$. In the most simple case, where the ground harmonic dominates, (11.1a, 11.1b) reduces to

$$V = NV_0 + V_1 \sum_{n=1}^{N} \cos\left\{ \frac{2\pi}{b} (x_0 + na) + \varphi_1 \right\}, \tag{11.2}$$

where V_1 is real valued and $2\tilde{V}(G_1) = V_1 \exp(i\varphi_1)$. At a given position of the rigid slider, as denoted for example by the value of x_0, the (lateral) force F exerted from the substrate on the slider will be $F = -dV/dx_0$ and thus

$$F = \frac{2\pi}{b} V_1 \sum_{n=1}^{N} \sin\left\{ \frac{2\pi}{b} (x_0 + na) + \varphi_1 \right\} + \cdots, \tag{11.3}$$

where the periods indicate that the full Fourier series may have more terms than only the first harmonic. We will now explore the effects of symmetry on the way in which F grows with system size.

Surfaces with identical lattice constants. If $a = b$, the sum in (11.3) is easy to evaluate because all summands are identical, thus

$$F = N \frac{2\pi}{b} V_1 \sin\left(\frac{2\pi}{b} x_0 + \varphi_1 \right), \tag{11.4}$$

where we have neglected the effect of higher harmonics for reasons of simplicity. If we want to initiate sliding, the maximum lateral force has to be overcome, thus the static friction force reads

$$F_\mathrm{s} = N \frac{2\pi}{b} V_1. \tag{11.5}$$

Therefore, F_s is proportional to the number of atoms in contact. As the argument extends to higher dimensions as long as the identical surfaces are perfectly oriented, one may say that the friction force is proportional to the area of contact. This is not a contradiction to Amontons' laws, which state that static friction is proportional to the load irrespective of the (apparent) area of contact, because in our example, we would (implicitly) double the whole system and thus double the load L when we double N. In other words, the static friction coefficient would be independent of the area of contact for two ideally-oriented, rigid surfaces with identical lattice constants. Note that the *kinetic* friction force remains zero in this purely mechanical model without *internal* degrees of freedom, which would justify an ad-hoc damping term.

Commensurate surfaces with $a \neq b$. Two surfaces are called commensurate when the ratio a/b is a rational number, i.e., if there are two natural numbers p and q such that

$$\mathcal{L} = pa = qb, \tag{11.6}$$

where $\mathcal{L}$ is the smallest common period of the two surfaces. We have considered $p = q = 1$ above. Now we will be concerned with the case in which $\mathcal{L}$ is greater than both a and b. Consider again (11.1a, 11.1b). Terms related to Fourier components G_m that cannot be represented as an integer multiple of $2\pi/\mathcal{L}$ will not add up systematically as can be learned from (11.1b). Thus for sufficiently large interfaces, the terms that grow proportionally with N will dominate. These terms require that $a \cdot m$ be equal to an integer multiple of b, e.g., $a \cdot m$ is equal to $\mathcal{L}$, in which case $m = p$. Thus we can write the "corrugation potential" V_{corr} as

$$V_{\text{corr}} = N \sum_{m, \frac{aG_m}{2\pi} \in Z} \tilde{V}(G_m)\ e^{iG_m x_0}. \tag{11.7a}$$

$$= N v(x_0), \tag{11.7b}$$

where $v(x_0)$ can be interpreted as an average atomic potential. Typically, the expansion coefficients in surface potentials decay exponentially fast with increasing index [10]. Therefore, one has to expect static friction that is exponentially small in the smallest common period of the two surfaces in contact, although it does grow linearly with system size.

The exponential decay of the V_m's with increasing index m usually justifies to only keep the first term in the expansion, which is related to the smallest common period. Therefore, the first non-constant and non-vanishing term to contribute to V_{corr} is related to the lattice vector $G_p = 2\pi\mathcal{L}/ab$ and hence one often may approximate

$$V_{\text{corr}} \approx N V_p \cos\left(\frac{2\pi p}{b} x_0 + \varphi_p\right). \tag{11.8}$$

V_p and φ_p can be constructed from $\tilde{V}_p$ in a similar way as the equivalent terms related to the first harmonic introduced in (11.2).

As a side comment, I wish to note that two non-adhesive two Lennard Jones solids with $\mathcal{L} = 4a = 5b$ are sufficiently "out of tune" to produce friction coefficients of less than 10^{-3} when pressed against each other (unpublished data by the author). Mathematically speaking, one would not refer to such surfaces as superlubric, however, by means of practical applications it would be justified to classify such friction coefficients as extremely low.

Incommensurate surfaces. When the ratio a/b cannot be expressed by a rational number, the surfaces are said to be incommensurate. One may argue that the two surfaces have a common period $\mathcal{L}$ that is infinite and given the argument that the corrugation potential vanishes exponentially with $\mathcal{L}$, one would conclude that the corrugation potential per surface atom v/N tends to zero for large N. Alternatively, consider again (11.1a, 11.1b). There is no term in the corrugation potential that would increase with system size. Therefore, (11.7b) would have to be replaced with

$$V_{\rm corr} = v_{\rm inc}(N, x_0), \tag{11.9}$$

where $v_{\rm inc}$ can be calculated from (11.1a, 11.1b). While $v_{\rm inc}(N, x_0)$ does not vanish exactly, one can argue that it does not grow systematically with system size N, because the fractions on the r.h.s. of (11.1b) are oscillating functions of N and only the first few terms contribute in a significant way. Keeping only the first term related to $m = \pm 1$, one can easily see that the maximum lateral force per atom and hence $\mu_{\rm s}$ vanishes with $1/N$ for increasing system sizes or alternatively, the net friction force

$$F_{\rm s} \propto N^0. \tag{11.10}$$

Flat, disordered surfaces. For incommensurate surfaces, the annihilation of lateral forces is rather systematic, similar to the deconstructive interference in optics. In order to include the effect of disorder into our model, several avenues can be pursued. One is to introduce randomness into the substrate potential, another one is to allow for stochastic variations of the interatomic spacing within the slider. Both procedures yield similar results for rigid planes, but the latter approach is more easily discussed in the present context. Let us assume that two adjacent neighbors in the slider have a spacing of $a + \delta a_{n,n+1}$, where $\delta a_{n,n+1}$ is an (independent) random number of mean zero, $\langle \delta a_{n,n+1} \rangle = 0$, and well defined second moment of $\langle \delta a^2_{n,n+1} \rangle = \sigma_a^2$. The position of particle n can then be given by

$$x_n = x_0 + na + \tilde{a}_n \tag{11.11}$$

with $\tilde{a}_n = \sum_{n'=1}^{n-1} \delta a_{n',n'+1}$. Thus, the positions of two atoms m and n become more and more uncorrelated as the distance between m and n decreases. Specifically, the expectation values of $(x_m - x_n)$ and $(x_m - x_n)^2$ satisfy

$$\langle (x_m - x_n) \rangle \ = (m - n)\, a \tag{11.12a}$$

$$\langle (x_m - x_n)^2 \rangle = (m - n)^2 a^2 + |m - n|\, \sigma_a^2. \tag{11.12b}$$

We can rewrite (11.3) as

$$F = \frac{2\pi}{b} V_1 \Im \left(\sum_{n=0}^{N-1} \exp \left\{ i \left[\frac{2\pi}{b} (x_0 + na + \tilde{a}_n) + \varphi_1 \right] \right\} \right), \tag{11.13}$$

where $\Im(\bullet)$ denotes the imaginary part its argument. Higher harmonics were suppressed for simplicity. Due to the randomness in the configuration, we can only be concerned with statistical averages. For a given value of x_0, the expectation value of the force reads:

$$\langle F \rangle = \frac{2\pi}{b} V_1 \Im \left[\sum_{n=0}^{N-1} \exp \left\{ \frac{2\pi\, i}{b} (x_0 + na) + i\varphi_1 \right\} \left\langle \exp \left\{ \frac{2\pi\, i}{b} \tilde{a}_n \right\} \right\rangle \right]. \tag{11.14}$$

As we are interested in large systems, it will be sufficient to discuss the properties of $\tilde{a}_m$ for the many terms with $m \gg 1$. The standard deviation of $\tilde{a}_m$ is $\sqrt{m}\sigma_a$, which will exceed b for sufficiently large indices m. Therefore, the expectation value of $\exp(i2\pi\tilde{a}_m/b)$ on the right-hand-side of (11.14) will be zero, leading to

$$\lim_{N\to\infty} \frac{1}{N} \langle F \rangle = 0. \tag{11.15}$$

As for the second moment of the lateral force F, one can start from (11.13) and make use of the fact that the statistical properties of the expression would remain unaltered, if we took the real part instead of the imaginary part on the right hand side of the equation. Thus,

$$\langle F^2 \rangle = \left(\frac{2\pi V_1}{b}\right)^2 \frac{1}{2} \sum_{n,n'} \exp\left\{ i \frac{2\pi}{b}(n-n')a \right\} \left\langle e^{i2\pi(\tilde{a}_n - \tilde{a}_{n'})/b} \right\rangle. \tag{11.16}$$

Note that the expressions on the r.h.s. of (11.16) are translationally invariant, so that we can replace the sum over n' with a sum over Δn. Except for "surface" terms, one can therefore write

$$\langle F^2 \rangle = \frac{1}{2} N \left(\frac{2\pi V_1}{b}\right)^2 \sum_{\Delta \mathrm{n}} \exp\left\{ i \frac{2\pi a}{b} \Delta n \right\} \left\langle e^{i2\pi\tilde{a}_{\Delta \mathrm{n}}/b} \right\rangle \tag{11.17a}$$

$$= \frac{1}{2} N \left(\frac{2\pi V_1}{b}\right)^2 f\left(\frac{a}{b}, \frac{\sigma_a}{b}\right), \tag{11.17b}$$

where we can approximate the function $f\left(\frac{a}{b}, \frac{\sigma_a}{b}\right)$ with

$$f\left(\frac{a}{b}, \frac{\sigma_a}{b}\right) = \sum_{\Delta \mathrm{n}} \exp\left\{ i \frac{2\pi a}{b} \Delta n \right\} e^{-\frac{1}{2}(2\pi \Delta n \sigma_a/b)^2}, \tag{11.18}$$

which is exact for a Gaussian distribution of the random numbers $\delta a_{n,n+1}$ (as can be seen by seen by terms of a cumulant expansion).

An interesting aspect of (11.17b) is that the disorder introduced in (11.11) leads to an expectation value $\langle F^2 \rangle \propto N$ irrespective of the position x_0 and independent of the precise value of a/b. The function $f\left(\frac{a}{b}, \frac{\sigma_a}{b}\right)$ is merely a prefactor, which is maximal for $a = b$. However, the "average commensurability," $a = b$ with $\sigma_a > 0$, does not induce the $F \propto N$ behavior found for truly commensurate surfaces. Since (11.17b) is valid for any position x_0, it is also valid at the position where F is maximum, and hence we may conclude that

$$F_s \propto \sqrt{N}, \tag{11.19}$$

or in other words, $\mu_\mathrm{s} \propto 1/\sqrt{N}$.

Higher dimensions. The calculations done above generalize to higher dimension if we associate N with the number of atoms in the bottom most layer of the slider. Thus for two three-dimensional solids with a two-dimensional interface, the model of rigid surfaces would predict that

$$\mu_s \propto \begin{cases} A^0 & \text{commensurate} \\ A^{-1} & \text{incommensurate} \\ A^{-1/2} & \text{amorphous} \end{cases} . \tag{11.20}$$

It is important to reemphasize that the results strictly depend on the assumption that the planes are ideally rigid and atomically smooth. For commensurate surfaces in dimensions larger than one, it is not sufficient for the two surfaces to have identical periods, but the orientation of the two surfaces need to be aligned. In numerical simulations of atomically smooth but disordered surfaces, the $\mu_s \propto 1/\sqrt{A}$ hypothesis could be confirmed [11].

Extended defects and contact lines. Recent experiments [12–14] were concerned with the test of the predictions summarized in (11.20). Towards this end, small clusters of either antimony or gold were deposited onto graphite surfaces and their kinetic friction forces were measured in UHV. (For the moment, let us assume that static and kinetic friction correlate for these systems.) The results, which are reviewed in more detail in the chapter by Schwarz and Schirmeisen, are consistent with the observation that amorphous clusters increase their friction with $\sqrt{A}$. Moreover, it was found that trace amounts of dirt immediately induced finite friction, as predicted from computer simulations [15]. Any loosely-bonded contamination at the interface acts like a soft layer having the ability to deform such that it locks simultaneously into the two opposing surfaces.

While things seem to be established for amorphous clusters, the scaling of the friction force of finite, crystalline clusters on extended crystalline substrates calls for particular attention [16]. The reason is that we can no longer neglect defects (such as dislocations) or contact lines, whose effect can scale differently from the predictions in (11.20). When the contact line is rugged, annihilation of friction forces is no longer as systematic as for "perfect" crystals without surfaces. We now have to modify our summation such that we treat the contact line of length l_c as random leading to a $F_s \propto \sqrt{l_c}$. The scaling would now depend on the fractal dimension of the contact line. Assuming that it goes as $\sqrt{A}$, we find $F_s \propto A^{1/4}$. However, when the contact line is straight, the scaling can differ again. For example, if the contact line lies parallel to a crystallographic axis of the substrate, the energy barrier towards pushing this contact line over a row of atoms is linear in l_c. Thus, we would retrieve the $\sqrt{A}$ scaling valid for amorphous islands. However, when the straight contact line is not aligned with the crystallographic axis of the substrate, very slow scaling can be envisioned as well. Thus, the scaling of the static friction force can range from $F_s \propto A^0$ to $F_s \propto A^{1/2}$ for finite crystals on extended, crystalline substrates.

11.2.3 Role of Instabilities in Simple Models

The way in which elastic deformations can induce finite friction and thus break superlubricity has been casted in a semi-quantitative fashion within the Prandtl-Tomlinson (PT) [4, 5] model and within the Frenkel-Kontorova (FK) [6] model. The PT model describes the effects of elasticity within the Einstein model of solids, i.e., each (surface) atom in the slider is coupled to its ideal lattice site via a harmonic spring of strength k. Moreover surface atoms experience a sinusoidal interaction $V_s(x) = V_0 \cos(2\pi x/a)$ with the effectively rigid substrate plus a drag force linear in the atom's velocity, which will be motivated later. The equation of motion for the particle reads

$$m\ddot{x} + \gamma\dot{x} = \frac{2\pi V_0}{a} \sin(2\pi x/a) - k\,[x - x_d(t)]\,, \tag{11.21}$$

where $x_d(t)$ is the position of the atom's ideal lattice site. Above a critical value k^* for the spring constant k, the PT model predicts zero kinetic friction even in this athermal model. To be specific, if k exceeds the maximum curvature $\kappa = (2\pi/a)^2 V_0$ of the substrate potential, then no instabilities (hysteresis loops) are found and kinetic friction vanishes in the PT model. Thus, if the slider is sufficiently stiff, F_k may disappear and hence the interface can be superlubric. For $k < k^*$, zero-temperature F_k remains finite no matter how slowly we pull the spring, which is akin of Coulomb's law of friction, [3], The latter states that kinetic friction is (rather) independent of the sliding velocity v_0. The mechanism is described in more detail in Fig. 11.2.

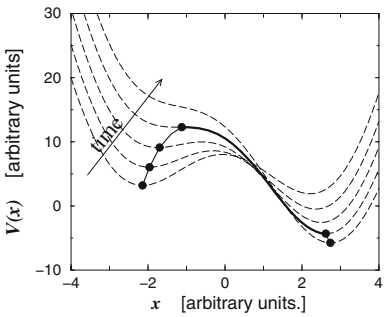


Fig. 11.2 Schematic representation of an instability in the PT model. The time-dependent potential $V(x)$ (*dashed lines*) is the superposition of a sinusoidal substrate potential, which is fixed in space, and a moving parabolic potential, which represents the elastic interaction between a degree of freedom (DOF), e.g., an atom or an AFM tip, and the moving driving device. The DOF, whose position at various instances of time is indicated by the *full circles*, becomes unstable when its mechanical equilibrium position is $x \approx -1.7$, from where it jumps to $x \approx 2.6$. In this process, an energy $\Delta E \approx 16.6$ is dissipated, provided the stage moves slowly and that thermal fluctuations can be neglected. The jumps occurs every time the driving device moves by one lattice constant, which is 2π in this example. Sliding over a long distance will therefore result in an average kinetic friction force of $F_k \approx 16.6/2\pi \approx 2.65$

The FK model is similar to the PT model, except that surface atoms are not coupled elastically to their lattice site but instead to their neighbors in a one-dimensional chain. The advantage of the FK model over the PT model that tribologists have seen is that the FK model incorporates the effect of long-range elastic deformation. However, one needs to keep in mind that the FK model, being a one-dimensional model, tremendously overestimates elastic deformations, artificially suppressing superlubricity.[2] The FK model and its generalizations to higher dimensions are yet useful to describe various tribological phenomena conceptually and sometimes even quantitatively. In fact, in one of their first studies, Hirano and Shinjo parameterized a generalization of the FK model with realistic interatomic potentials for copper and found that incommensurate copper surfaces should often remain superlubric [2].

Of course, there can be other ways than local, elastic instabilities that induce finite friction and break superlubricity. Adhesive jump-to-contact instabilities, plastic deformation, instabilities in confined boundary lubricants, chemical reactions, cold welding, and related processes are all characterized by sudden pops of microscopic degrees of freedom, which lead to energy dissipation [9].

11.2.4 Effect of Temperature

Temperature-induced relaxation can both increase or decrease friction [9]. It can increase friction because the real contact area can increase due to thermally-assisted plastic flow, which increases the effective load and consequently the friction. Moreover, thermally assisted aging can increase the strength of the junctions formed between two solids at rest. In superlubric systems, one should generally hope that plastic flow is negligible. Temperature then helps the instability (such as the one shown in Fig. 11.2) to occur prematurely and therefore temperature can reduce kinetic friction. When pulled sufficiently slowly, the system will be able to be close to thermal equilibrium, e.g., the coordinate x in the PT model will occur approximately with a probability $\exp(-V(x)/k_{\mathrm{B}}T)$ at a given position of the driving stage and not only close to the (previous) mechanically stable position as indicated in Fig. 11.2. In this case, linear-response theory is applicable and friction must be proportional to velocity.[3] In fact, any *finite* system at *finite* temperature will eventually approach the linear-response regime [17]. Even finite, *commensurate* surfaces, which according to any reasonable definition should not be classified as superlubric, will ultimately enter a regime in which $F_{\mathrm{k}} \propto v$, as demonstrated in numerical simulations of dry

[2] The PT model is the mean-field variant of the one-dimensional FK model and therefore more realistic in higher dimensions, in particular above the upper critical dimension for elastic manifolds sliding through external potentials.

[3] For a discussion of the velocity-dependence of F_{k} at more elevated velocities, see the chapter by Evstigneev and Reimann in this book.

contacts [18]. We refer the reader to [19] for a more in-depth description in particular with respect to the effect of thermal noise and the subsequent crossover from Stokes to Coulomb friction.

11.2.5 Damping in the Superlubric Regime

When friction is extremely small, it is tempting to assume that instabilities do not play a significant role and that the system is close to thermal equilibrium at all times. As argued in the previous section, friction would then have to be considered to be linear in velocity v_0 at small v_0. The response of the slider to a time-dependent external force can be calculated by using linear-response theory and the fluctuation-dissipation theorem, [17] i.e., the fluctuations of the lateral force at any fixed position of the driving device can be related to the the damping that the slider would exert on a slowly moving driving device. An excellent overview on the topic and tips how to make use of the concepts in computer simulations is given in the book *Molecular Simulation* by Frenkel and Smit [20]. Note that the damping coefficient of a degree of freedom (be it an atom or be it a collective degree of freedom such as an AFM tip) can be position dependent in principle.

The linear-response formalism has so far been employed in the so-called non-contact mode of atomic force microscope (AFM) tips [21, 22] (see also the chapter by Kantorovich and Trevethan in this book), however, the formalism is the same for intimate contacts [18]. It appears that the theoretically predicted values for non-contact damping are smaller than the ones measured experimentally. The reason for the discrepancy may be that most calculations assume sliding of ideal crystals. Surface defects, dislocations, and contamination are usually neglected. They will all lead to additional force fluctuations and hence to increased friction between substrate and slider. In essentially every real-life contact, there will be isolated points of contact where instabilities cannot be prevented. These will lead to dissipation that dominates the damping-induced interactions. Single-asperity contacts in sufficiently well-defined microcontacts may be an exception to this rule [7].

Lastly, it is worth mentioning that linear response does not always imply very low friction, for example, when internal deformation in solids is highly viscous—as is the case for rubber. Formalism to describe dissipation for sliding rubber contacts [23], or generalizations thereof, should form a promising starting point for formal theories of damping forces in superlubric contacts.

11.2.6 Long-Range Elastic Deformations

In real solids, atoms are not coupled elastically to their lattice sites but they interact with the other atoms in the solid. For many purposes, it is sufficient to treat this interaction within one solid as harmonic interactions between adjacent atoms. When one

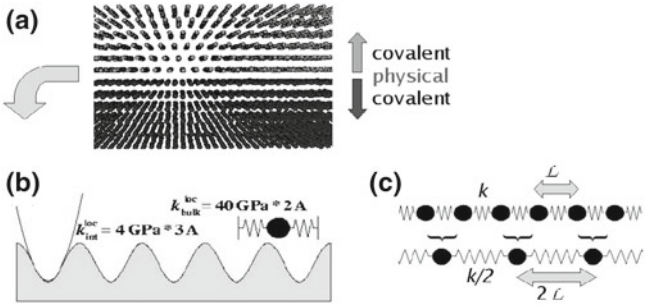


Fig. 11.3 Schematic view of the interaction between chemically passivated solids. Part **a** shows a snapshot of a simulation. Part **b** represents the coupling of a surface atom to its neighbors (reflected by the harmonic springs) and to the substrate (reflected by the sinusoidal *line*). The *parabola* indicates the maximum curvature of the atom–substrate potential, which corresponds to the curvature κ of the substrate potential in the PT model. Part **c** describes the scaling procedure for a one-dimensional elastic chain. From [26]

solid is placed on top of another one as shown schematically in Fig. 11.1, instabilities of surface atoms do not necessarily involve the sudden motion of single atoms but they may involve the collective motion of many atoms. In order to ascertain whether such collective instabilities occur, it is then necessary to analyze whether the intrabulk, elastic interactions dominate the interfacial interactions on length scales larger than atomic scales. If the answer is positive, the system can be superlubric.

The equations of motion for an elastic solid that is sliding with respect to a (rigid) substrate have the same structure as those that describe charge density waves (CDWs) [24]. These latter systems have been studied thoroughly [25]. One result is that the CDWs or, to be more precise, the vortices in type-II superconductors can be treated as essentially rigid domains within a characteristic length called the Larkin length. Adjacent Larkin domains are coupled only weakly so that (roughly speaking) its center of mass can have more than one (meta)stable position. This multistability would automatically break superlubricity in a way similar to that shown in Fig. 11.2.

In order to apply the concept of Larkin domains to tribological phenomena, one can proceed as follows: [26] Assume that a block of linear dimension $\mathcal{L}$ is rigid. Then assess how such a rigid block would couple elastically to its neighboring block (so that the solid's elastic constants remain independent of the coarse-grain length $\mathcal{L}$) and how the block manages to interlock with the substrate, see also Fig. 11.3. To get an order of magnitude estimate for the relevant variables at the atomic scale, $\mathcal{L} = \mathcal{L}_{\text{atomic}}$, it is reasonable to assume that the elastic coupling k between adjacent atoms is in the order of the bulk modulus B, say 40 GPa for a soft solid, times a lattice spacing, say 2 Å. Thus, $k(\mathcal{L}_{\text{atomic}}) \approx 8$ N/m would be a reasonable value for relatively soft solids. For hydrogen-terminated diamond, the estimate for $k(\mathcal{L}_{\text{atomic}})$ would be more than a factor 10 larger. A similar estimate can be done for the maximum local curvature of the substrate potential κ, which was introduced in the PT model following (11.21): The atoms of opposed surfaces interact via weak physical bonds, provided the surfaces

are chemically passivated. Solids that are tied together by physical bonds have bulk moduli in the order of 4 GPa and nearest neighbor separation are in the order of 3 Å, resulting in an estimate of $\kappa \approx 1.2$ N/m. Thus, at the atomic scale, $k > \kappa$, so that instabilities cannot be expected to occur. This is a necessary conditions for superlubricity.

Once the values for k and κ are estimated at the atomic scale, $\mathcal{L}_{\text{atomic}}$, it is possible to also estimate those values at any length scale $\mathcal{L}$. For instance, if we replace a linear chain of N beads with separation a by $N/2$ beads that are separated by $2a$, then we need to reduce the stiffness of the coarse-grained springs by a factor of 2, as illustrated in Fig. 11.3. Thus, the springs become softer upon coarse-graining in one-dimensional systems, just like the effective capacitance of capacitors connected in series is reduced. In two dimensions, we would not only connect springs (capacitors) in series but also in parallel. The net effect is that the coarse-grained spring (capacitor) has the same stiffness (capacitance) as the original one. Each dimension added in parallel makes the stiffness of the springs harder. If D is the dimensionality of the solid, one finds

$$k(\mathcal{L}) = \left(\frac{\mathcal{L}}{\mathcal{L}_{\text{atomic}}}\right)^{D-2} k(\mathcal{L}_{\text{atomic}}). \tag{11.22}$$

It is a bit more difficult to find a similar scaling law for the curvature of the slider-substrate potential. However, it is reasonable to assume that (apart from a prefactor) κ scales similarly as the static friction force $F_{\text{s}}(\mathcal{L})$ of a rigid domain of linear scale $\mathcal{L}$. In scaling studies, it is necessary to keep the intrinsic thermodynamic such as the (average) normal pressure $p_\perp$ constant so that we can say $\kappa(\mathcal{L}) \propto F_{\text{s}}(\mathcal{L}) = \mu_{\text{s}}(\mathcal{L}) \cdot (\mathcal{L}^{D_{\text{int}}} p_\perp)$, where D_{int} is the dimension of the interface and thus $(\mathcal{L}^{D_{\text{int}}} p_\perp)$ the load carried by an "area" A of linear dimension $\mathcal{L}$. The scaling of the friction coefficient, however, strongly depends on the order at the interface. Using (11.20) and the $\kappa \propto \mu L$ relationship for rigid or correlated domains, one finds that

$$\kappa(\mathcal{L}) \propto \begin{cases} \mathcal{L}^{D_{\text{int}}} & \text{commensurate} \\ \mathcal{L}^{0} & \text{incommensurate} \\ \mathcal{L}^{D_{\text{int}}/2} & \text{amorphous} \end{cases}. \tag{11.23}$$

Whenever the ratio of $k(\mathcal{L})/\kappa(\mathcal{L})$ increases systematically with $\mathcal{L}$, one should expect the intra-bulk elasticity to dominate and thus to have the possibility of superlubricity. In the technically relevant case of disordered, two-dimensional interfaces and three-dimensional solids, both $k(\mathcal{L})$ and $\kappa(\mathcal{L})$ increase linearly with $\mathcal{L}$. Allowing for some elastic deformation within the Larkin domains effectively yield logarithmic correction to the scaling laws, which would result in finite friction [27, 28]. However, as the Larkin domains would be relatively large and no local instabilities should be expected, these corrections can probably be seen as irrelevant, in particular for systems such as hydrogen-terminated diamond with large values of k/κ at the atomic scale.

11.2.7 Self-affine Rough Surfaces

So far we have neglected the highly non-uniform distribution of normal loads, which are the consequence of realistic surface topologies. Real contacts have roughness on many different length scales and the distribution of normal pressures sometimes allows for large normal values, which would increase the values for κ at the intimate point of contacts. It is plausible to expect elastic instabilities in such point of high pressure. However, elastic instabilities in bulk systems require pressures that are typically higher than the yield strength of the material, e.g., incommensurate interfaces between two atomically smooth gold surfaces only show friction in molecular dynamics simulations when $p_\perp$ is raised above 4 GPa [29]. (The periodic boundary conditions in lateral direction allow the gold solids to sustain these unusually large pressures.)

Campañá [30] investigated the friction between two surfaces, whose topography had been obtained from height measurements of steel surfaces characterized by a Hurst roughness exponent of $H = 0.84$, i.e., surface roughness lives predominantly on large wave lengths. As interaction between the walls, he used hard wall interactions as well as a model in which the energy (surface) density increased exponentially quickly as two surfaces approach each other locally. Campañá found that significant instabilities only occurred when driving the system into "ranges that are untypical for tribological experiments," i.e., mean (macroscopic) pressures of 0.1 times the Young's module and velocities being 0.2 times the velocity of sound, which would translate to a parameter range often assumed in FK-type modeling of solid friction. However, he also noted that friction exceeds 10^{-3} times the load when including roughness down to the atomic scale, i.e., all the way to the discretization of the elastic manifold. This confirms our analysis that it is the roughness and the interlocking at the atomic scale that needs to be considered most when designing superlubric systems.

Similar to Campañá, Luan and Robbins [31] considered the friction between rough surfaces. However, in addition to elastic manifolds, they also considered plastic deformation by using explicit Lennard Jones atoms plus a coupling to a finite-element foundation. Moreover, they varied the degree of commensurability between the surfaces. None of their results indicated as small friction coefficients as those found by Campañá, i.e., their friction coefficient roughly ranged between $2\cdot10^{-1} \lesssim \mu \lesssim 10^{-2}$. In comparison to values of every-day tribological systems, these friction coefficients can be labeled as small. They are nevertheless too large to qualify as superlubric. The relatively large values might be explained by the following observations. First, interactions were simple two-body interactions, which intrinsically cannot describe effects such as chemical passivation or directed bonds, which play a crucial role for real systems, for example, in graphite. Second, roughness was imposed by discrete steps favoring large roughness at small scales, which automatically leads to high friction, as Campañá already noted. Third, contact patches were very small, that is, in the order of five atoms, and moreover one-dimensional. Thus, lateral forces could not annihilate locally in a very systematic fashion. One therefore should not be discouraged by the results and instead learn that roughness at small scales should be avoided when designing superlubric systems. Interestingly, Luan and Robbins

found that plasticity can decrease friction between incommensurate surfaces. The plastic deformation flattens the top of the asperity during run-in thereby enlarging local contact areas, which enhances the systematic annihilation of lateral forces.

11.3 Simulations

Analytical calculations, such as the ones presented in the last section, can generally not be used to make quantitative predictions for specific systems. Even the most simple models without any chemical detail elude analytical tractability. For example, there are no closed-form solutions for the dynamical properties of the Frenkel-Kontorova (FK) model, which consists of a one-dimensional, elastic chain embedded into a simple, sinusoidal potential. Only some aspects of the continuum variant of the FK model, namely the sine-Gordon (SG) model, can be solved analytically. Due to its continuum character, there is no finite kinetic friction in incommensurate SG models. Discreteness corrections can be applied, however, the calculations are tedious [32]. Given the fact that most analytical models are one-dimensional and the paramount importance of dimensionality emphasized in the last section, it appears to be a sensitive choice to resort to computer simulations of sliding solids.

The main stream technique for these simulations is molecular dynamics (MD) [8]. Atomic configurations are set up in a computer experiment, interactions between atoms are assumed (or calculated with quantum chemical first-principle techniques) and Newton's equations of motion are solved for each individual atom. The boundary walls are typically coupled to a driving device and boundary conditions are chosen such that a tribological experiment is mimicked in the best possible way. Sliding produces heat, which then requires the use of (artificial) thermostats that remove the heat from a sliding contact. [8, 9] give an overview of the results of atomistic simulations, a more technical review on how to set up simulations will soon be available too [33].

11.3.1 Generic Models

A large part of the literature is concerned with so-called generic models, where the interaction between atoms are only simple two-body potentials. Results of these simulations can generally not be applied to specific systems. However, they allow one to study the effect of dimensionality, commensurability, contamination, surface geometry, etc. on tribological properties. Simulations of generic models can capture many tribological effects qualitatively or in some cases even semi-quantitatively. One popular approach of generic models is the study of "Lennard Jonesium," which is a (virtual) material in which atoms interact through the Lennard Jones (LJ) potential $V(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, where σ and ϵ are the LJ length and LJ energy, respectively. σ can be roughly associated with the sum of the radii of the two interacting atoms and ϵ with the cohesive energy of a dimer or nearest neighbors in a solid.

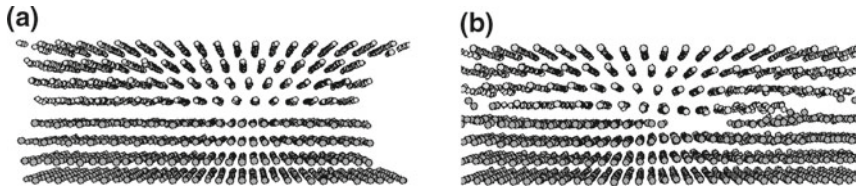


Fig. 11.4 **a** Low-load configuration of chemically-passivated solids with terraces. Both solids are fcc Lennard Jonesium with [111] surfaces. The surfaces are misoriented by 90°. Static and kinetic friction are too small to be defined, because thermal fluctuations at ambient conditions are strong enough to depin the contact. **b** Snapshot of a configuration that results from configuration (**a**) as a consequence of high normal pressure. Despite the onset of plastic deformation, the static friction coefficient for this contact remains smaller than 0.01. From [34]

In the context of superlubricity, one important question to address is whether solids generically have the tendency to show finite static and kinetic friction when brought into contact or whether one should expect superlubricity only for very specific materials. In order to evaluate at what interfacial strength perfect [111] surfaces of face-centered-cubic (fcc) solids would become unstable instantaneously if the models are more realistic than simple bead spring models, Müser placed two LJ solids on top of each other (see also Fig. 11.3a) [34]. All interactions between identical atoms, i.e., those atoms that originate from the same solid, were chosen such that $\epsilon = \sigma = 1$. The LJ parameters for pairs of atoms originating from opposed solids were chosen $\sigma_1 = \sigma$ and ϵ_1 was varied. In order to obtain *instantaneous* instabilities and thus pinning, the value ϵ_1 had to exceed that of ϵ by a factor of eight. These instabilities, however, were not elastic in nature, but they involved large rearrangement of the atoms which could be interpreted as cold welding. If the simulations could be run for much longer times, the mixing instabilities could of course be observed for any value of $\epsilon_1 > \epsilon$. Other simulations also support the idea that instabilities in solids without directed bonds are typically not elastic in nature, e.g., the elastic instabilities found by Lançon at normal pressures of 4 GPa in simulations of incommensurate solids of gold, [29] would imply plastic flow in laboratory experiments, see for instance Fig. 11.4b where the periodic boundary conditions in lateral direction prevent significant (further) plastic flow.

As long as the pressure conditions are not extreme and plastic deformation is absent or relatively small, all-atoms simulations of curved tips show that many of the analytical arguments advanced in Sect. 11.2.2 for flat contacts also apply to curved tips, [35], for instance that friction is largest for commensurate and smallest for incommensurate contacts and intermediate for disordered surfaces. One of the conclusions to be drawn from those simulations is that large radii of curvature lead to a smaller ratio of lateral and normal pressure. This means that layered materials may be one possibility to achieve ultra-small friction.

Recent approaches have incorporated roughness on more length scales and allowed to include the long-range elastic deformation in the solids by means of multi-scale methods [36]. While superlubricity is broken in these systems (the load

is carried at isolated locations only, which induces local pinning), it needs to be emphasized that (a) exceedingly small "terraces" were allowed in that study, sometimes as small as one atom, and (b) despite of the existence of these small terraces friction coefficients were typically still well below 0.1. When the solids were allowed to deform plastically as well, the small terraces were rubbed away resulting in smaller friction coefficients. This is a typical observation in MD simulations of unlubricated (non-metallic) surfaces that are driven at moderate or small loads. See also Fig. 11.4, which shows a contact with moderate plastic flow.

Many tribological aspects cannot be modeled by Lennard Jonesium, for instance whenever directed bonds are important. These play a crucial role in molecular solids, layered materials and other chemically complex solids such as the low-friction material Teflon. Also metals are poorly described by LJ potentials, for instance the well-known necking during contact formation between bare metals is a consequence of the particular many-body form of metals. Generally, metals have a larger propensity to annihilate free surface than Lennard Jonesium, which favors cold welding. Lennard Jonesium cannot reproduce these effects and thus more realistic studies are required to understand friction (or the absence thereof) in these materials.

11.3.2 Layered Materials

Given the analysis in the previous sections, one would expect that layered solids are natural candidates for superlubric materials. First, the elastic coupling within a sheet, i.e., k in the PT model, is very large, while the interaction with the opposed surface, i.e., κ in the PT model, is small. Moreover, long terraces are formed and as argued in the last section, large local radii of curvature lead to small friction in unlubricated and non-yielding contacts. Thus, almost all simulations of one layered material placed on another layered material show small friction.

An interesting aspect of many layered materials, in particular graphite, is that they have the ability to form nanotubes. These nanotubes allow one to study friction in a single-asperity contact for well-defined geometries as opposed to AFM tips, whose detailed atomic structure has remained elusive. Experimentally [37] and in simulations, [38, 39] registry or commensurability is found to play the role that one would expect from classical mechanics and the concept of geometric interlocking. The static friction is large if the nanotubes orient with the graphite substrate, while misaligned nanotubes show much reduced friction.

In many cases nanotubes are multiwalled. The friction between an inner tube and an outer tube could be measured experimentally [40]. Many times, inner tube and outer tube are incommensurate due to the different ways in which graphite sheets can fold to form a tube. When the tubes are incommensurate, non-extensive shear stresses are found suggestive of the dominant roles of surface effects in those systems.

Theoretical simulations of double-walled nanotubes often show extremely small friction indeed, which lead to the suggestion to use them as mechanical nano-oscillators [41]. The main force inducing lateral motion between inner and outer

tube is related to surface energy. The tubes try to minimize free surface, which can act in sliding direction or against sliding direction. In the absence of instabilities, only little energy is dissipated as heat within one oscillation cycle. It may be surprising that this last statement even holds for commensurate nanotubes, which have large instantaneous/static friction forces [42]. However, commensurate nanotubes also show large instantaneous "anti-friction." The net damping calculated for commensurate and incommensurate nanotubes turns out to be of similar magnitude—at least for the idealized geometries utilized in the simulations [42]. Thus, commensurability does not automatically lead to instabilities and significant kinetic friction. It may be worth pointing out that kinetic friction between slightly contaminated commensurate surfaces has been predicted to be even smaller than that between incommensurate forces [43].

Most simulations idealize the nanotubes initial geometry and underestimate the friction force. However, if chemical details of the fractured end are incorporated into the simulations, experimentally measured and calculated values for kinetic friction agree reasonably well [44]. It is probably safe to assume that the friction between nanotubes is dominated by defects, in particular those that involve chemical bonds between inner tube and outer tube. As the covalent bonds between carbon atoms are orders of magnitude stronger than the van der Waals interactions, these chemical bonds (and their rupture) between atoms in the inner and atoms in the outer tube will most significantly contribute to the dissipation in real nanotubes.

In practice, one disadvantage of layered materials certainly is that the sheets are easily rubbed off. This process leads to the generation of debris and consequently friction increases. This effect may become particularly severe when chemical point defects are present. Indeed, some of the first experimental evidence for superlubricity between layered molybdenum disulfide (MoS_2) was seen to be transient until wear particles probably increased the friction to coefficient to a value above 10^{-3} [45].

11.3.3 Metal on Metal Contacts

While the first realistic calculations suggestive of superlubricity [1, 2] were based on incommensurate contacts, it is probably safe to assume that bare metals are not good candidates for superlubric materials. One of the reason is the "vulnerability" of metal contacts in particular at their boundaries [46]. Upon sliding, dislocations and ultimately wear is easily generated from the contact boundaries, as discussed for ideally blunt copper tips sliding on ideal copper (111) surfaces. Thermal aging of the tips and hence producing more realistic contact geometries results in even more wear [46].

One can yet ask the question, how much friction one should expect in idealized situations, i.e., if it was possible to suppress the boundary effects. A particularly interesting study addressed the role of roughness in the very last layer. Qi et al. [47] studied atomically smooth Ni(100)/Ni(100) interfaces. Their idealized geometries display the same superlubric behavior as the idealized copper interfaces studied by

Hirano and Shinjo [1]. However, roughening the top layer with a mere 0.8 Å rms variation, changes the behavior completely, with friction coefficients increasing by several orders of magnitude. The calculated values for the surfaces with the 0.8 Å additional roughness matches the available experimental data extremely well [48].

In another study, the friction between chemically passivated aluminum surfaces (Al_2O_3 termination) was calculated [49]. An interesting result of that study is that incommensurate surfaces show small but yet non-negligible friction at moderate normal loads, although no wear occurred. As opposed to the commensurate case, the friction force evolved relatively smoothly as a function of time for incommensurate Al_2O_3 contacts, which would only be consistent with a picture in which the instabilities are localized. The "wear-less" instabilities must thus be a consequence of the directed bonds, which are usually not incorporated in a purely theoretical description of contacts. More studies would be needed to identify better candidates for superlubric terminated metals than aluminum, however, it may well be that this quest will not be successful.

11.3.4 Hydrogen-Terminated Surfaces

One of the most promising candidates for superlubric materials is hydrogen-terminated diamond-like carbon (DLC) [50]. Unfortunately, there have not yet been many simulations on hydrogen-terminated DLC, although results will soon be published [51]. Interesting new results are expected in the near future, partly due to the availability of force fields that allow one to model chemical reactions. An example for such a force field is the reactive empirical bond order potentials, [52] which was used to study the compression and sliding induced polymerization reactions within alkyne chains [53]. Chemical reactions may impede superlubricity, as the formation of bonds corresponds to instabilities, which can lead to Coulomb type friction if the reactions are reversible [54, 55].

A large fraction of computational studies of chemically-passivated, "realistic" surfaces has been concerned with self-assembled monolayers (SAM), which are much softer than chemically bonded solids, and thus inappropriate candidates for superlubricity. It has yet to be noted that SAMs have typically low friction and probably good resistance to wear, which make them interesting materials for tribological applications. Many of the results obtained in the literature can be interpreted with the concept of geometric interlocking. For instance, when n-alkane monolayers are grafted to a diamond surface with a period of two lattice constants ($2a$) to form a SAM, and this SAM is slid against a perfectly oriented diamond surface, then the friction force has a period of $2a$, with positive and negative contributions [56]. In these simulations, the maximum instantaneous lateral force, which can be associated with the static friction force F_s reaches 4 nN at normal loads of approximately 80 nN. The average, kinetic friction force F_k was not stated explicitly, but can be estimated from the graphs as being at least a factor of 10 smaller than F_s, thus the system is almost superlubric.

It would be interesting to know the friction between SAMs attached to diamond and hydrogen-terminated diamond for incommensurate interfaces. In simulations of generic models of two surfaces separated by one monolayer, it was found that static friction is much larger for commensurate than for incommensurate systems, while the opposite is true for kinetic friction [11].

The crucial role of orientation was also found in another study of friction involving monolayers. Commensurate fluorine-terminated alkanethiol SAMs exhibited a tremendous sensitivity of the (differential) kinetic friction coefficient $\tilde{\mu}_k = dF_\mathrm{k}/dL$ on line defects in one of the two SAMs [57]. When the layers were truly commensurate, $\tilde{\mu}_k$ turned out much larger than in those cases where line defects significantly reduced commensurability and thus $\tilde{\mu}_k$. Also the packing density in SAMs has the effect on friction that one would expect from the theoretical analysis, i.e., less densely packed and hence softer systems will become more easily unstable than dense, hard systems. For instance, friction between an amorphous carbon tip and a SAM decreased with increasing packing. Two systems had been compared [58]. When the polymers in the SAM were identical so that all of them contained $N = 14$ carbon atoms, the structures were dense, while they were less dense when N was either 12 or 16 (with equal probability). Kinetic friction in the latter system was twice as large than in the dense system.

11.4 Conclusions

Theoretical considerations and atomistic simulations clearly indicate that there should be the possibility of identifying many materials showing superlubricity. The ingredients favoring superlubricity are smoothness of the surfaces, in particular, smoothness of the last layer, and chemical passivation of the last layer. Studies investigating the relevance of roughness on larger length scales are still in their infancy, however, first results indicate that the local roughness is paramount. Thus, layered solids such as graphite and molybdenum disulfide appear as natural candidates for superlubric materials. However, the sheets in layered solids are easily rubbed off, which results in debris and ultimately increased friction.

The currently best candidate for superlubricity remain systems similar to diamond-like carbon, in which the dangling bonds are saturated. Currently, the surfaces seem to be predominantly passivated with hydrogen. From a theoretical point of view, it might be possible to make the surfaces smoother by terminating the dangling bonds of carbon with atoms that are larger than hydrogen. Thus, if feasible, theory would suggest a termination of the dangling bonds in diamond-like carbon with fluorine rather than with hydrogen. Potential alternatives are silicon saturated with either fluorine or perhaps chlorine.

Lastly, it is important to stress that superlubricity is not necessarily a well-defined term. For instance, commensurate systems typically have high static friction, but they do not necessarily show high kinetic friction, as was discussed in this chapter in the context of (idealized) commensurate carbon nanotubes. One may well classify

these structures as superlubric, because they dissipate little energy upon sliding. However, other systems have small kinetic friction *and* small static friction. The smallness in static friction is typically due to the incompatibility of geometries of the two opposed surfaces. For this scenario, the term structural lubricity was suggested. The terms structural lubricity and superlubricity can best be illustrated in the context of the experiments by Socoliuc et al. [7] and Dienwiebel et al. [59] In Socoliuc's experiment, the instantaneous lateral forces between the AFM tip and the substrate remained (relatively) large. This clearly indicates that the contact was not structurally lubric. However, due to the stiff cantilever, no instabilities occurred, and thus the contact can be identified as superlubric. In Dienwiebel's experiment, a graphite flake was brought out of registry. This resulted not only in a small kinetic friction but even in a small instantaneous force. Therefore, the contact was not only superlubric (absence of instabilities) but even structurally lubric (absence of geometric interlocking).

References

1. M. Hirano, K. Shinjo, Phys. Rev. B **41**, 11837 (1990)
2. K. Shinjo, M. Hirano, Surf. Sci. **283**, 473 (1993)
3. C.A. Coulomb, *Théorie des Machines Simples* (Memoirs de Mathematique et de Physique de l'Academie Royale, Paris, 1785), pp. 161–342
4. L. Prandtl, Z. Angew, Math. Mech. **8**, 85 (1928)
5. G.A. Tomlinson, Philos. Mag. Series **7**, 905 (1929)
6. Y.I. Frenkel, T. Kontorova, Zh Eksp, Teor. Fiz. **8**, 1340 (1938)
7. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)
8. M.O. Robbins, M.H. Müser, in *Modern Tribology Handbook*, ed. by B. Bhushan (CRC Press, Boca Raton, 2001)
9. M.H. Müser, M. Urbakh, M.O. Robbins, Adv. Chem. Phys. **126**, 187 (2003)
10. W. Steele, Surf. Sci. **36**, 317 (1973)
11. M.H. Müser, L. Wenning, M.O. Robbins, Phys. Rev. Lett. **86**, 1295 (2001)
12. C. Ritter, M. Heyde, B. Stegemann et al., Phys. Rev. B **71**, 085405 (2005)
13. D. Dietzel, C. Ritter, T. Monninghoff et al., Phys. Rev. Lett. **101**, 125505 (2008)
14. D. Dietzel, T. Monninghoff, C. Herding et al., Phys. Rev. B **82**, 035401 (2010)
15. G. He, M.H. Müser, M.O. Robbins, Science **284**, 1650 (1999)
16. A.S. Wijn, Phys. Rev. B **86**, 085429 (2012)
17. R. Kubo, Rep. Prog. Phys. **29**, 255 (1966)
18. M.H. Müser, M.O. Robbins, Phys. Rev. B **61**, 2335 (2000)
19. M.H. Müser, Phys. Rev. B **84**, 125419 (2011)
20. D. Frenkel, B. Smit, *Understanding Molecular Simulation* (Elsevier, San Diego, 2002)
21. M. Gauthier, M. Tsukada, Phys. Rev. B **60**, 11716 (1999)
22. L.N. Kantorovich, Phys. Rev. B **64**, 245409 (2001)
23. B.N.J. Persson, J. Chem. Phys. **115**, 3840 (2001)
24. B.N.J. Persson, E. Tosatti, Solid State Commun. **109**, 739 (1999)
25. G. Blatter, M.V. Feigelman, V.B. Geshkenbein et al., Rev. Mod. Phys. **66**, 1125 (1994)
26. M.H. Müser, Europhys. Lett. **66**, 97 (2004)
27. J.B. Sokoloff, Phys. Rev. Lett. **86**, 3312 (2001)
28. J.B. Sokoloff, Phys. Rev. B **65**, 115415 (2002)
29. F. Lançon, Europhys. Lett. **57**, 74 (2002)
30. C. Campaña, Phys. Rev. B **75**, 155419 (2007)

31. B. Luan, M.O. Robbins, Tribol. Lett. **36**, 1 (2009)
32. O.M. Braun, Y.S. Kivshar, Phys. Rep. **306**, 2 (1998)
33. M.H. Müser, N.J. Mosey, Rev. Comp. Che. **25**, 67 (2007)
34. M.H. Müser, Tribol. Lett. **10**, 15 (2001)
35. B.Q. Luan, M.O. Robbins, Nature **435**, 929 (2005)
36. B. Luan, S. Hyun, M.O. Robbins, N. Bernstein, Mater. Res. Soc. Symp. Proc. **841**, R7.4 (2005)
37. M.R. Falvo et al., Nature **397**, 236 (1999)
38. A. Buldum, J.P. Lu, Phys. Rev. Lett. **83**, 5050 (1999)
39. J.D. Schall, D.W. Brenner, Mol. Simulat. **25**, 73 (2000)
40. A.N. Kolmogorov, V.H. Crespi, Phys. Rev. Lett. **85**, 4727 (2000)
41. J.L. Rivera, C. McCabe, P.T. Cummings, Nano Lett. **3**, 1001 (2003)
42. P. Tangney, S.G. Louie, M.L. Cohen, Phys. Rev. Lett. **93**, 065503 (2004)
43. M.H. Müser, Phys. Rev. Lett. **89**, 224301 (2002)
44. Z. Xia, W.A. Curtin, Phys. Rev. B **69**, 233408 (2004)
45. J.M. Martin, C. Donnet, Th Le Mogne, Phys. Rev. B **48**, 10583 (1993)
46. M.R. Sørensen, K.W. Jacobsen, P. Stoltze, Phys. Rev. B **53**, 2101–2113 (1996)
47. Y. Qi, Y.-T. Cheng, T. Cagin, W.A. Goddard III, Phys. Rev. B **66**, 085420 (2002)
48. J.S. Ko, A.J. Gellman, Langmuir **16**, 8343 (2000)
49. Q. Zhang, Y. Qi, L.G. Hector Jr et al., Phys. Rev. B **72**, 045406 (2005)
50. A. Erdemir, Surf. Coat. Techn. **146**, 292 (2001)
51. J.A. Harrison (private ccommunication)
52. D.W. Brenner, O.A. Shenderova, J.A. Harrison et al., J. Phys. Condens. Matt. **14**, 783 (2002)
53. G.M. Chateauneuf, P.T. Mikulski, G.T. Gao, J.A. Harrison, J. Phys. Chem B **108**, 16626 (2004)
54. N.J. Mosey, T.K. Woo, M.H. Müser, Phys. Rev. B **72**, 054124 (2005)
55. N.J. Mosey, M.H. Müser, T. K. Woo, Science **307**, 1612 (2005)
56. P.T. Mikulski, J.A. Harrison, Tribol. Lett. **10**, 29 (2001)
57. B. Park et al., Langmuir **20**, 10007 (2004)
58. P.T. Mikulski, G.T. Gao, G.M. Chateauneuf, J.A. Harrison, J. Chem. Phys. **122**, 024701 (2005)
59. M. Dienwiebel et al., Phys. Rev. Lett. **922**, 126101 (2004)

Part III
Multiscale Friction

Chapter 12
On the Fractal Dimension of Rough Surfaces

Bo Persson

Abstract Most natural surfaces and surfaces of engineering interest, e.g., polished or sand blasted surfaces, are self affine fractal over a wide range of length scales, with the fractal dimension $D_f = 2.15 \pm 0.15$. We give several examples which illustrate this and a simple argument, based on surface fragility, for why the fractal dimension usually is <2.3. A kinetic model of sand blasting is presented, which gives surface topographies and surface roughness power spectra in good agreement with experiments.

12.1 Introduction

All natural surfaces and surfaces of engineering interest have surface roughness on many different length scales, sometimes extending from atomic dimensions to the linear size of the object under study. Surface roughness is of crucial importance in many engineering applications, e.g., in tribology [1–4]. For example, the surface roughness on a road surface influences the tire-road friction or grip [1]. It is therefore of great interest to understand the nature of roughness of surfaces of engineering interest. Several studies of the fractal properties of surface roughness have been presented, but mainly for surfaces produced by growth (atomic deposition) processes [5]. Many studies of surfaces produced by atomistic erosion processes, e.g., sputtering, have also been presented, see, e.g., [6–8]. In this article I will present several examples of power spectra of different surfaces with self-affine fractal-like surface roughness. All surfaces have fractal dimensions $D_f = 2.15 \pm 0.15$ and I will give a simple argument, based on surface fragility, for why the fractal dimension usually is <2.3. I also present a kinetic model of sand blasting which gives surface topographies and surface roughness power spectra in good agreement with experiments.

B. Persson (✉)
Peter Grünberg Institut, Forschungszentrum Jülich, 52425 Jülich, Germany
e-mail: b.persson@fz-juelich.de

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_12

12.2 Power Spectrum: Definition

We consider randomly rough surfaces where the statistical properties are transitionally invariant and isotropic. In this case the 2D power spectrum [3, 9]

$$C(\mathbf{q}) = \frac{1}{(2\pi)^2} \int d^2x \ \langle h(\mathrm{x}) h(\mathbf{0}) \rangle e^{-i\mathbf{q}\cdot\mathrm{x}}$$

will only depend on the magnitude q of the wavevector $\mathbf{q}$. Here $h(\mathrm{x})$ is the height coordinate at the point $\mathrm{x} = (x, y)$ and $\langle .. \rangle$ stands for ensemble averaging. From $C(q)$ many quantities of interest can be directly calculated. For example, the root-mean-square (rms) roughness amplitude h_{rms} can be written as

$$h_{\mathrm{rms}}^2 = 2\pi \int_{q_0}^{q_1} dq \ qC(q) \tag{12.1}$$

where q_0 and q_1 are the small and large wavevector cut-off. The rms-slope κ is determined by

$$\kappa^2 = 2\pi \int_{q_0}^{q_1} dq \ q^3 C(q). \tag{12.2}$$

For a self affine fractal surface

$$C(q) = C_0 \left(\frac{q}{q_0} \right)^{-2(1+H)} \tag{12.3}$$

Substituting this in (12.1) gives

$$h_{\mathrm{rms}}^2 = \frac{\pi C_0}{H} q_0^2 \left[1 - \left(\frac{q_1}{q_0} \right)^{-2H} \right] \tag{12.4}$$

and from (12.2) we get

$$\kappa^2 = \frac{\pi C_0}{1-H} q_0^4 \left[\left(\frac{q_1}{q_0} \right)^{2(1-H)} - 1 \right] \tag{12.5}$$

Usually $q_0/q_1 \ll 1$ and since $0 < H < 1$, unless H is very close to 0 or 1, we get

$$\kappa = q_0 h_{\mathrm{rms}} \left(\frac{H}{1-H} \right)^{1/2} \left(\frac{q_1}{q_0} \right)^{1-H} \tag{12.6}$$

Many surfaces of engineering interest, e.g., a polished steel surface, have rms-roughness of order $\sim 1\,\mu\mathrm{m}$ when probed over a surface region of linear size $L = \pi/q_0 \sim 100\,\mu\mathrm{m}$. This gives $q_0 h_{\mathrm{rms}} \approx 0.1$ and if the surface is self affine fractal the whole way down to the nanometer region (length scale a) then $q_1 = \pi/a \approx 10^{10}\,\mathrm{m}^{-1}$ and (12.6) gives $\kappa \approx 0.1 \times 10^{5(1-H)}$. I use this equation to argue that most surfaces of interest, if self affine fractal from the macroscopic length scale (say $L \sim 100\,\mu\mathrm{m}$) to the nanometer region, cannot have a fractal dimension larger than $D_{\mathrm{f}} \approx 2.3$ or so, as otherwise the average surface slope becomes huge which is unlikely to be the case as the surface would be very "fragile" and easily damaged (smoothed) by the mechanical interaction with external objects. That is, if we assume that the rms slope has to be below, say [3], we get that $H > 0.7$ or $D_{\mathrm{f}} = 3 - H < 2.3$. As we now show, this inequality is nearly always satisfied for real surfaces.

12.3 Power Spectra: Some Examples

I have calculated the 2D surface roughness power spectra of several hundred surfaces of engineering interest. Here I give just a few examples to illustrate the general picture which has emerged. Figure 12.1 shows the 2D power spectrum of a sand blasted PMMA surface obtained from 1D-stylus height profiles. The surface is self-affine fractal for large wavevectors and the slope of the dashed line corresponds to the Hurst exponent $H = 1$ or fractal dimension $D_{\mathrm{f}} = 2$. For $q < q_{\mathrm{r}} \approx 10^4\,\mathrm{m}^{-1}$ (corresponding to the roll-off wavelength $\lambda_{\mathrm{r}} = \pi/q_{\mathrm{r}} \approx 100\,\mu\mathrm{m}$) the power spectrum exhibits a roll-off which, however, moves to smaller wavevectors as the sand blasting time period increases (not shown).

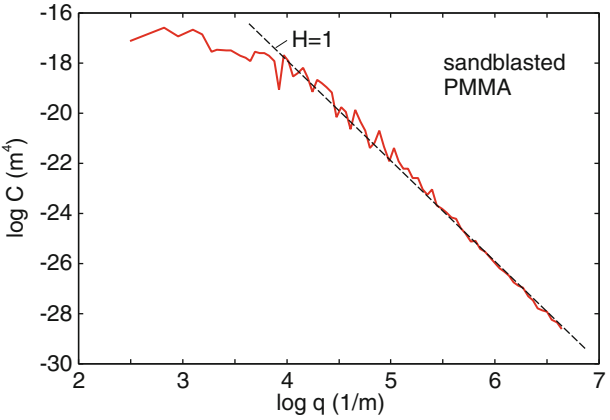


Fig. 12.1 The 2D power spectrum of a sand blasted PMMA surface based 1D-stylus height profiles [10] ($\log_{10} - \log_{10}$ scale). The slope of the *dashed line* corresponds to the Hurst exponent $H = 1$ or fractal dimension $D_{\mathrm{f}} = 2$

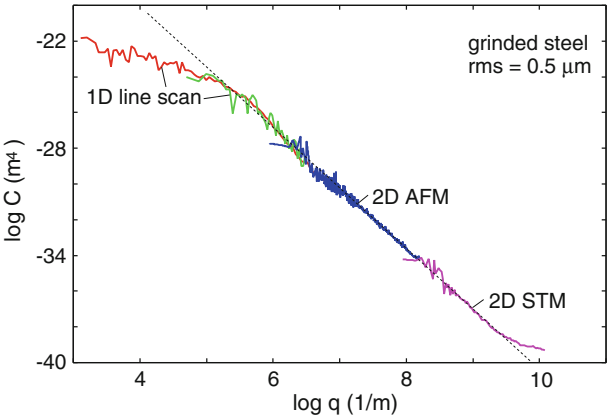


Fig. 12.2 The 2D power spectra of grinded steel surface [11]. The slope of the *dashed line* correspond to the Hurst exponent $H = 0.72$ or fractal dimension $D_f = 2.28$

Figure 12.2 shows the angular averaged power spectrum of a grinded steel surface. The surface topography was studied on different length scales using STM, AFM and 1D stylus. Note that the (calculated) power spectra using the different methods join smoothly in the wavevector regions where they overlap. The slope of the dashed line corresponds to the Hurst exponent $H = 0.72$ or fractal dimension $D_f = 2.28$

Figure 12.3 shows the power spectra of two asphalt road surfaces. Both surfaces are self-affine fractal for large wavevectors and exhibit a roll-off for small wavevectors which is related to the largest stone particles (diameter d) in the asphalt via $q_r \approx \pi/d$. The fractal dimension of both surfaces are $D_f \approx 2.20$.

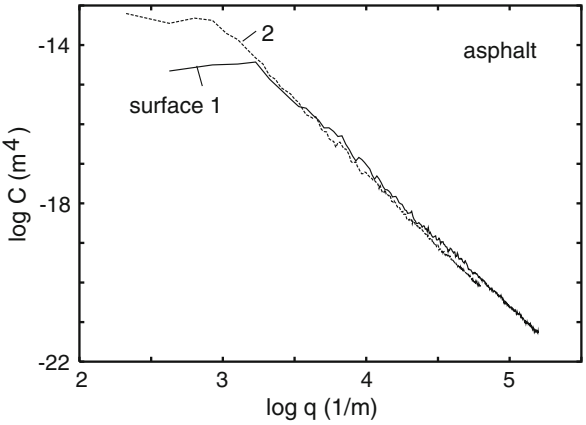


Fig. 12.3 The 2D power spectra of asphalt road surface [12]. The slope of the *dashed line* correspond to the Hurst exponent $H = 0.80$ or fractal dimension $D_f = 2.20$

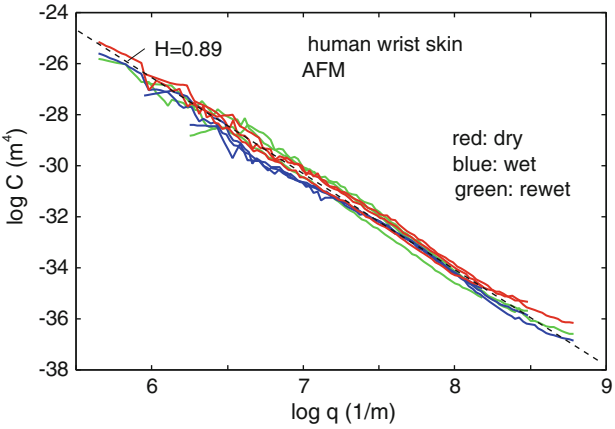


Fig. 12.4 The 2D power spectrum of human wrist skin obtained from AFM measurements [13]. The rms roughness is $h_{\rm rms} \approx 0.25$ μm within the studied wavevector region. The slope of the *dashed line* corresponds to the Hurst exponent $H = 0.89$ or fractal dimension $D_{\rm f} = 2.11$

Not only surfaces prepared by engineering methods (e.g., sand blasting or polishing) exhibit self-affine fractal properties with fractal dimensions $D_{\rm f} = 2.15 \pm 0.15$ but so do most natural surfaces. Thus, for example, surfaces prepared by crack propagation are usually self affine fractal with $D_{\rm f} \approx 2.2$. Here I give three more examples to illustrate this. Figure 12.4 shows the 2D power spectrum of human wrist skin obtained from AFM measurements. The rms roughness is $h_{\rm rms} \approx 0.25\,\mu$m in the studied wavevector region. The slope of the dashed line corresponds to the Hurst exponent $H = 0.89$ or fractal dimension $D_{\rm f} = 2.11$. Figure 12.5 shows 2D power

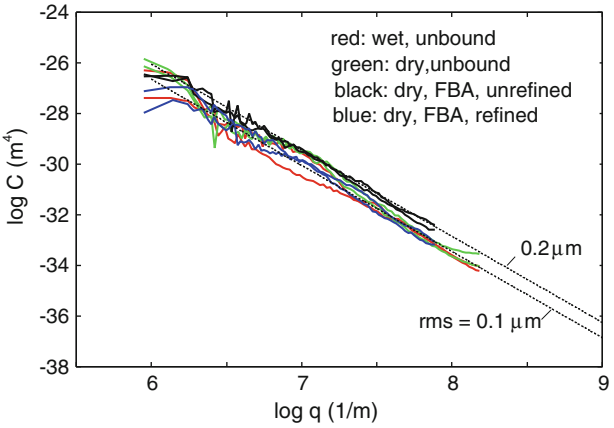


Fig. 12.5 The 2D power spectra of dry and wet cellulose fibers [15]. The surface topography was measured using AFM. The slope of the *dashed lines* correspond to the Hurst exponent $H = 0.7$ or fractal dimension $D_{\rm f} = 2.3$

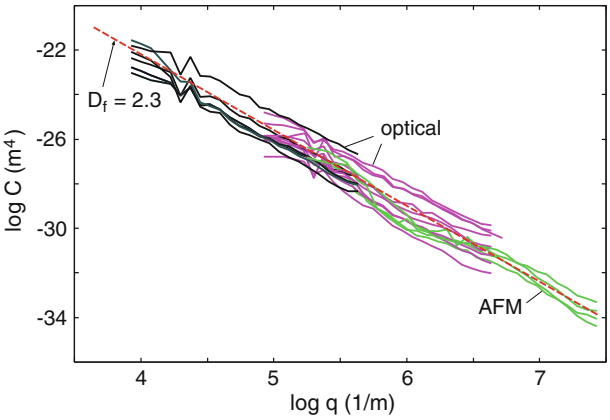


Fig. 12.6 The 2D power spectrum of pulled adhesive tape based on optical and AFM measurements [16]. The slope of the *dashed line* corresponds to the Hurst exponent $H = 0.7$ or fractal dimension $D_f = 2.3$

spectra of dry and wet cellulose fibers measured using AFM. The slope of the dashed lines correspond to the Hurst exponent $H = 0.7$ or fractal dimension $D_f = 2.3$. Finally, Fig. 12.6 shows the 2D power spectrum of pulled adhesive tape based on optical and AFM measurements. The slope of the dashed line corresponds to the Hurst exponent $H = 0.7$ or fractal dimension $D_f = 2.3$.

I have shown above that many engineering and natural surfaces exhibit self-affine fractal properties in a large wavevector range with fractal dimension $D_f = 2.15 \pm 0.15$. A fractal dimension larger than $D_f = 2.3$ is unlikely as it would typically result in surfaces with very large rms-slope, and such surfaces would be "fragile" and easily smoothed by the (mechanical) interaction with the external environment. However, this argument does not hold if the surface is self-affine fractal in a small enough wavevector region or if the prefactor C_0 in the expression $C(q) = C_0(q/q_0)^{-2(1+H)}$ is very small. In fact, self affine fractal surfaces with the fractal dimension $D_f = 3$ result when a liquid is cooled below its glass transition temperature where the capillary waves on the liquid surface gets frozen-in. For capillary waves (see, e.g., [2]):

$$C(q) = \frac{1}{(2\pi)^2} \frac{k_B T}{\rho g + \gamma q^2} \tag{12.7}$$

where ρ is the mass density, g the gravitation constant and γ the liquid surface tension. For $q \gg q_0 = (\rho g/\gamma)^{1/2}$ we have $C(q) \sim q^{-2}$ and comparing this with the expression for a self affine fractal surface $C(q) \sim q^{-2(1+H)}$ gives $H = 0$ and $D_f = 3$. In a typical case the cut-off $q_0 \approx 10^3$ m^{-1} is rather small, but the rms roughness and the rms slope are still rather small due to the smallness of $C_0 = k_B T/\rho g$, which results from the small magnitude of thermal energy $k_B T$. Using AFM, frozen capillary waves have recently been observed on polymer surfaces (polyaryletherketone, with the glass transition temperature $T_g \approx 423$ K and $\gamma \approx 0.03$ J/m) [18], see also [17].

The measured power spectrum was found to be in beautiful agreement with the theory prediction of (12.7). For this case, including all the roughness with $q > q_0$, one can calculate the rms roughness to be $h_{\rm rms} \approx (k_{\rm B}T/2\pi\gamma)^{1/2}[\ln(q_1/q_0)]^{1/2} \approx 1$ nm and the rms slope $\kappa \approx (k_{\rm B}T/4\pi\gamma)^{1/2}q_1 \approx 1$.

12.4 Simulation of Rough Surfaces: A Simple Erosion Process

I have argued above that if a surface is self affine fractal over a large wavevector region (as it is often the case) it usually has a fractal dimension <2.3, since otherwise the rms-slope would be so large ($\gg$1) as to make the surface fragile, and very sensitive to the impact of external objects which would tend to smooth the surface. Here I will consider a simple model of sand blasting, showing that if one assumes that material removal is more likely at the top of asperities rather than in the valleys (see Fig. 12.7), a surface with relatively low fractal dimension is naturally obtained. The model studied here has some similarities with growth models involving random deposition with surface relaxation. However, instead of adding atoms or particles I consider removal of material. In addition, while in growth models the surface relaxation is usually interpreted as a diffusive (thermal) motion of atoms, in the present case thermal effects are not directly involved (but may be indirectly involved in determining if the material removal involves plastic flow or brittle fracture).

We now present a model for sand blasting, where a beam of hard particles is sent on the surface orthogonal to the originally flat substrate surface, and with a laterally uniform probability distribution. The substrate is considered as a cubic lattice of

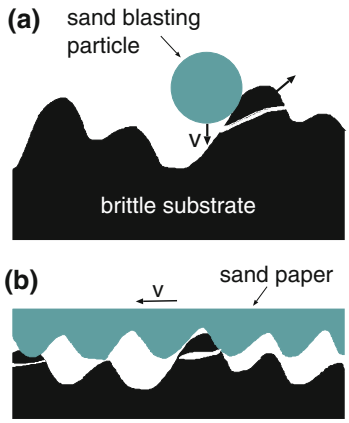


Fig. 12.7 Sand blasting (**a**) and lapping with sand paper (**b**) will roughen an initially flat surface but in such a way that high and sharp asperities never form i.e., the removal of material is easier at the top of asperities than at the valley. This will result in a rough surface with low fractal dimension

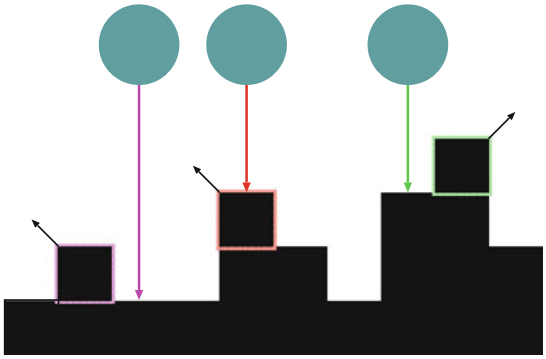


Fig. 12.8 Incoming particles (*arrows*) and the blocks removed by the impact (*black squares* surrounded by colored rims) for a 1D version of the simulation model used. For the 2D model I use, if a particle impact at site (i, j) (at position $(x, y) = (i, j)a$, where a is the lattice constant) then one of the blocks (i, j), $(i + 1, j)$, $(i - 1, j)$, $(i, j + 1)$ or $(i, j - 1)$ is removed. Of these blocks I assume that either the block which has the smallest number of nearest neighbors is removed (with probability 0.5), since this block is most weakly bound to the substrate, or the highest block is removed (with probability 0.5). In both cases, if several such blocks exist I choose randomly the one to be removed unless the block (i, j) is part of the set of blocks, in which case this block is removed

blocks (or particles) and every particle from the incoming beam removes a randomly chosen surface block on the solid substrate. As shown in Fig. 12.8, if an incoming particle impacts at site (i, j) (at position $(x, y) = (i, j)a$, where a is the lattice constant) then one of the blocks (i, j), $(i + 1, j)$, $(i - 1, j)$, $(i, j + 1)$ or $(i, j - 1)$ is removed. Of these blocks I assume that either (a) the block which has the smallest number of nearest neighbors is removed (with probability 0.5), since this block is most weakly bound to the substrate, or (b) the highest located block is removed (with probability 0.5). In both cases, if several such blocks exist I choose randomly the one to remove unless the block (i, j) is part of the set of blocks, in which case this block is removed. The substrate surface consists of 2048×2048 blocks and I assume periodic boundary conditions. We note that the processes (a) and (b) above are similar to the Wolf-Villain [19] and Family [20] grows models, respectively.

Figure 12.9 shows the topography of a surface produced by the eroding process described above (see also Fig. 12.8), after removing 76290 layers of blocks. The surface topography is practically undistinguished from that of sand blasted surfaces (not shown). Figure 12.10 shows the surface roughness power spectrum as a function of the wavevector (on a $\log_{10} - \log_{10}$ scale). The surface is self affine fractal with the Hurst exponent $H = 1$ (or fractal dimension $D_f = 2$), which has also been observed for sand blasted surfaces (see Fig. 12.1). Even the magnitude of $C(q)$ predicted by the theory is nearly the same as observed (see Fig. 12.1). For more results from simulations of surfaces roughened by erosion, see the Appendix.

Fig. 12.9 Topography picture of a surface produced by the eroding process described in Fig. 12.8 after removing 76,290 layers of blocks. The surface plane consists of 2,048 × 2,048 blocks. The surface is self affine fractal with the Hurst exponent $H = 1$ (or fractal dimension $D_f = 2$) (see Fig. 12.10). The width of the removed particles (or blocks) is $a = 0.1$ μm. The surface has the rms roughness $h_{rms} = 2.1$ μm and the rms slope $\kappa = 1.04$

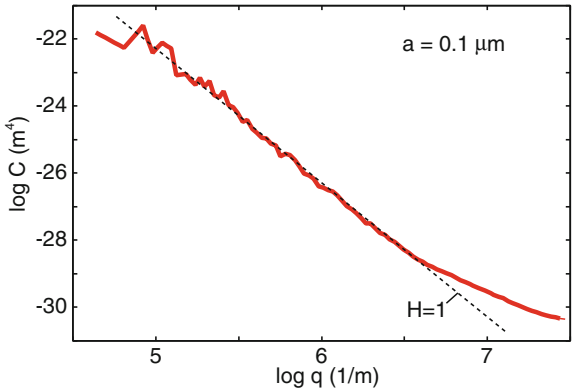


Fig. 12.10 The surface roughness power spectrum as a function of the wavevector ($\log_{10} - \log_{10}$ scale) after removing 76,290 layers of blocks (surface topography in Fig. 12.9). The surface plane consists of 2,048 × 2,048 blocks. The surface is self-affine fractal with the Hurst exponent $H = 1$ (or fractal dimension $D_f = 2$). We have assumed the linear size of the removed blocks to be $a = 0.1$ μm

12.5 Discussion and Summary

Surface roughness on engineering surfaces is important for a large number of properties such as the heat and electric contact resistance [21, 22], for mixed lubrication [23], wear and adhesion [24]. Thus, for example, one standard way to reduce adhesion is to roughen surfaces. In wafer bonding one instead wants the surfaces to be as smooth as possible and already surface roughness of order a few nanometer (when measured over a length scale of ∼100 μm) may eliminate adhesion.

Surfaces produced by brittle crack propagation tend to be self-affine fractal with the fractal dimension $D_f \approx 2.2$, but no generally accepted theory exists which can

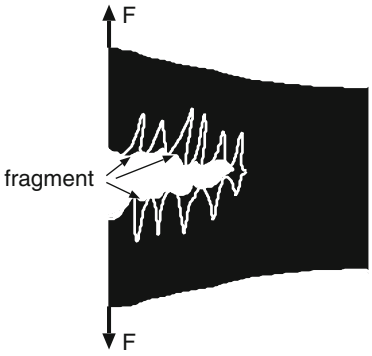


Fig. 12.11 Brittle fracture usually produces self-affine fractal surfaces with the fractal dimension $D_f \approx 2.2$. If (hypothetically) the fractal dimension would be much higher the surface slope would be very high too, which would result in sharp asperities broken-off forming fragments localized at the fracture interface

explain why [25, 26]. Fractured surfaces are usually very rough on macroscopic length scales. If such surfaces would have the fractal dimension $D_f > 2.3$ they would have huge rms-slope, i.e., very sharp asperities would appear at short enough length scales. It is intuitively clear that sharp asperities cannot form as they would not survive the cracking process, but would result in fragments of cracked material at the interface (see Fig. 12.11).

The argument presented in this paper for why the fractal dimension is close to 2 for most engineering surfaces assumes that the surfaces are produced by the mechanical interaction between solids and that the surfaces are fractal-like in a wide range of length scales. Many examples of surfaces with fractal dimension $D_f \approx 2.5$ or larger exist. For example, the surfaces resulting from electroreduction of Pd oxide layers have the fractal dimension $D_f \approx 2.57$ (see [27]). In this case no mechanical interaction with external objects (which could smooth the surface) has occurred. In addition, because of the relative thin oxide layer of the untreated surface, the self-affine fractal properties will only extend over a relative small range of length scales. Similarly, electrodeposition may result in surfaces with fractal dimension much larger than 2. Erosion by ion bombardment or exposure of a surface to plasma is another way of producing rough surfaces with self-affine fractal properties. In [8] it was shown that exposing a gold surface to oxygen or argon plasma produced self affine fractal surfaces with the fractal dimension $D_f = 2.1 \pm 0.1$. Ion bombardment (sputtering) of an iron surface produced a surface which was self-affine fractal over two decades in length scales (from 3 to 300 nm) with the fractal dimension $D_f = 2.47 \pm 0.02$ (see [7]). It is not obvious why the the gold and iron surfaces exhibit different fractal properties, but it may be related to the much higher mobility of Au atoms on gold as compared to Fe atoms on iron, which would tend to smooth the gold surface more than the iron surface [28].

To summarize, I have shown that most natural surfaces and surfaces of engineering interest, e.g., polished or sand blasted surfaces, are self affine fractal in a wide range

of length scales, with typical fractal dimension $D_f = 2.15 \pm 0.15$. I have argued that the fractal dimension of most surfaces <2.3, since surfaces with larger fractal dimension have huge rms-slopes and would be very fragile and easily smoothed by the interaction with external objects. I have also presented a simple model of sand blasting and showed that the erosion process I used results in self-affine fractal surfaces with the fractal dimension $D_f = 2$, in good agreement with experiments.

It is clear that a good understanding of the nature of the surface roughness of surfaces of engineering and biological interest, is of crucial importance for a large number of important applications.

Acknowledgments I thank J. Krim for useful comments on the text.

Appendix

Here I present some more results related to simulation of rough surfaces by erosion processes. Consider first the most simple picture of sand blasting where a beam of hard particles is sent on the surface orthogonal to the originally flat substrate surface, and with a laterally uniform probability distribution. The substrate is considered as a cubic lattice of blocks (or particles) and every particle from the incoming beam removes a randomly chosen surface block on the solid substrate. This process, which is similar to the random deposition model [5], will result in an extremely rough substrate surface with the Hurst exponent $H = -1$ and fractal dimension $D_f = 4$. This follows at once from the fact that the power spectrum of the generated surface is independent of the wavevector i.e., $C(q) = C_0$ (a constant) and using the definition $C(q) \sim q^{-2(1+H)}$ we get $H = -1$. The fact that $C(q)$ is constant in this case follows from the fact that the height $h(\mathbf{x})$ is uncorrelated with $h(\mathbf{0})$ for $\mathbf{x} \neq \mathbf{0}$. That is, $\langle h(\mathbf{x})h(\mathbf{0})\rangle = \langle h(\mathbf{x})\rangle\langle h(\mathbf{0})\rangle = 0$ for $\mathbf{x} \neq \mathbf{0}$. Thus we get

$$C(q) = \frac{1}{(2\pi)^2} \int d^2x \, \langle h(\mathbf{x})h(\mathbf{0})\rangle e^{-i\mathbf{q}\cdot\mathbf{x}} \sim \langle h^2(\mathbf{0})\rangle$$

Let us now consider the erosion processes (a), (b) and (a + b) discussed in Sect. 12.4. In Fig. 12.12 we show the power spectrum after removing 76,290, 19,070 and 2,384 layers of blocks assuming process (a + b). For short time of sand blasting a large roll-off region prevails which decreases towards zero as the sand blasting time increases. The same effect is observed in experiments (not shown) and reflects the fact that the correlation length ξ along the surface caused by the sand blasting extends only slowly as the sand blasting time t increases (as a power law $\xi \sim t^{1/z}$, see [5]).

In Fig. 12.13 I compare the surface roughness power spectrum as obtained using the random removal model with the random removal with relaxation models (a), (b) and ((a+b) (see Sect. 12.4) after removing 19,070 layers of blocks. The corresponding topography pictures for processes (a), (b) and ((a + b) are shown in Fig. 12.14. Note

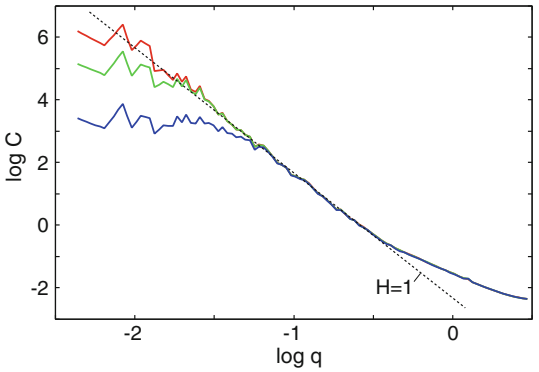


Fig. 12.12 The surface roughness power spectrum as a function of the wavevector ($\log_{10} - \log_{10}$ scale) for the erosion process (a + b), after removing 2,384 (*blue*), 19,070 (*green*) and 76,290 (*red*) layers of blocks. The wavevector is in units of $1/a$ and the power spectrum is in units of a^4

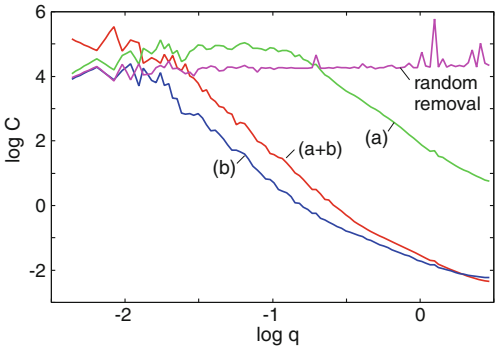


Fig. 12.13 The surface roughness power spectrum as a function of the wavevector ($\log_{10} - \log_{10}$ scale) for all the erosion processes considered, after removing 19,070 layers of blocks. The wavevector is in units of $1/a$ and the power spectrum is in units of a^4

that the random removal process gives a constant power spectrum which I have never observed for any real surface. The random removal with relaxation model (a) gives also unphysical surface topography with high sharp spikes. The ((a + b) model gives results in agreement with experiments, which shows, as expected, that both removal of high regions (asperity tops) and low coordinated surface volumes are important in sand blasting.

Note that random removal results in an interface which is uncorrelated (see above). The columns shrink independently, as there is no mechanism that can generate correlations along the interface. The other erosion processes [(a), (b) and ((a + b)] all involve correlated removal of material, allowing the spread of correlation along the surface.

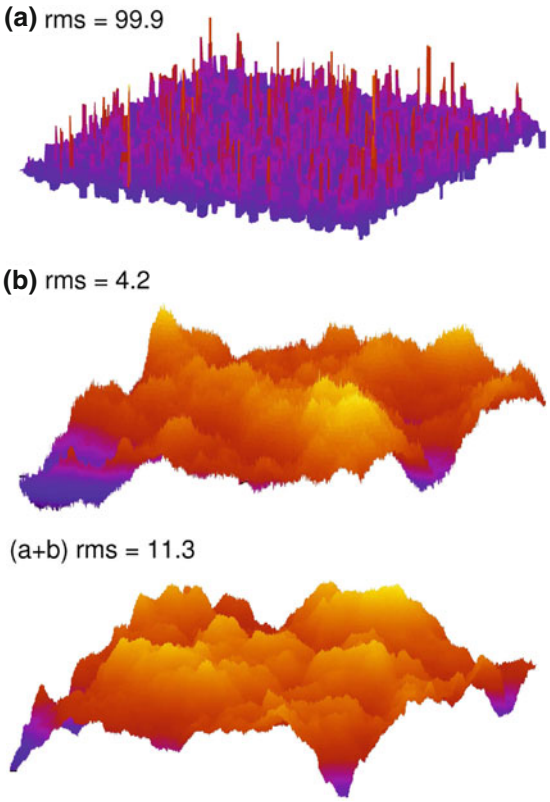


Fig. 12.14 Topography picture of surfaces produced by the eroding processes (a), (b) and (a + b) after removing 19,070 layers of blocks. The surface plane consist of $2{,}048 \times 2{,}048$ blocks. The rms roughness values are in units of a. Random removal without relaxation gives an extremely rough surface (not shown) with the rms roughness $h_{\mathrm{rms}} = 1{,}264a$

References

1. B.N.J. Persson, J. Chem. Phys. **115**, 3840 (2001)
2. B.N.J. Persson, Surf. Sci. Rep. **61**, 201 (2006)
3. B.N.J. Persson, O. Albohr, U. Tartaglino, A.I. Volokitin, E. Tosatti, J. Phys. Condens. Matter **17**, R1 (2005)
4. J. Krim, Adv. Phys. **61**, 155 (2012)
5. A.L. Barabasi, H.E. Stanley, *Fractal Concept in Surface Growth* (Cambridge University Press, Cambridge, 1995)
6. J. Krim, G. Palasantzas, Int. J. Modern Phys. B **9**, 599–632 (1995)
7. J. Krim, I. Heyvaert, C. Van Haesendonck, Y. Bruynseraede, Phys. Rev. Lett. **70**, 57–61 (1993)
8. D. Berman, J. Krim, Thin Solid Films **520**, 6201 (2012)
9. G. Carbone, B. Lorenz, B.N.J. Persson, A. Wohlers, Eur. Phys. J. **29**, 275 (2009)
10. B. Lorenz, PGI, FZ Jülich, Germany
11. A. Wohlers, IFAS, RWTH Aachen, Germany
12. O. Albohr, Pirelli Deutschland AG, 64733 Höchst/Odenwald, Germany

13. A. Kovalev, S.N. Gorb, Department of Functional Morphology and Biomechanics, Zoological Institute at the University of Kiel, Germany
14. B.N.J. Persson, A. Kovalev and S.N. Gorb, Tribology Letters 10.1007/s11249-012-0053-2.
15. C. Ganser, F. Schmied, C. Teichert, Institute of Physics, Montanuniversität Leoben, Leoben, Austria
16. A. Kovalev, S.N. Gorb, Department of Functional Morphology and Biomechanics, Zoological Institute at the University of Kiel, Germany
17. B.N.J. Persson, A. Kovalev, M. Wasem, E. Gnecco, S.N. Gorb, EPL **92**, 46001 (2010)
18. D. Pires, B. Gotsmann, F. Porro, D. Wiesmann, U. Duerig, A. Knoll, Langmuir **25**, 5141 (2009)
19. D.E. Wolf, J. Villain, Europhys. Lett. **13**, 389 (1990)
20. F. Family, J. Phys. A **19**, L441 (1986)
21. C. Campana, B.N.J. Persson, M.H. Muser, J. Phys. Condens. Matter **23**, 085001 (2011)
22. S. Akarapu, T. Sharp, M.O. Robbins, Phys. Rev. Letters **106**, 204301 (2011)
23. B.N.J. Persson, M. Scaraggi, European J. Phys. E **34**, 113 (2011)
24. N. Mulakaluri, B.N.J. Persson, EPL **96**, 66003 (2011)
25. E. Bouchaud, J. Phys. Condens. Matter **9**, 4319 (1997)
26. E. Bouchaud, G. Lapasset, J. Planes, Europhys. Lett. **13**, 73 (1990)
27. T. Kessler, A. Visintin, A.E. Bolzan, G. Andreasen, R.C. Salvarezza, W.E. Triaca, A.J. Arivia, Langmuir **12**, 6587 (1996)
28. J. Krim, private communication

Chapter 13
Contact Mechanics, Friction and Adhesion with Application to Quasicrystals

Bo Persson, Giuseppe Carbone, Vladimir N. Samoilov, Ion M. Sivebaek, Ugo Tartaglino, Aleksandr I. Volokitin and Chunyan Yang

Abstract We discuss the origin of friction and adhesion between hard solids such as quasicrystals. We emphasize the fundamental role of surface roughness in many contact mechanics problems, in particular for friction and adhesion between solid bodies. The most important property of rough surfaces is the surface roughness power spectrum $C(q)$. We present surface roughness power spectra of many surfaces of practical importance, obtained from the surface height profile measured using optical methods and the Atomic Force Microscope. We show how the power spectrum determines the contact area between two solids. We also present applications to contact mechanics and adhesion for rough surfaces, where the power spectrum enters as an important input.

13.1 Introduction

The first sample of a quasicrystal was produced in 1982 [1]. Intensive studies of this class of metallic materials have been conducted since that time. Quasicrystals display a unique combination of physical properties, namely low heat conductivity, relatively high hardness, and (under atmospheric condition) low friction coefficient and low

B. Persson (✉)
Peter Grünberg Institut, Forschungszentrum Jülich, 52425 Jülich, Germany
e-mail: b.persson@fz-juelich.de

G. Carbone
CEMeC Politecnico di Bari, Bari, Italy

V.N. Samoilov · C. Yang
IFF, FZ-Jülich, 52425 Jülich, Germany

I.M. Sivebaek
Department of Mechanical Engineering, Technical University of Denmark, Lyngby, Denmark

U. Tartaglino
Pirelli Tires, Milan, Italy

A.I. Volokitin
Samara State Technical University, Samara 443100, Russia

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_13

Fig. 13.1 A stainless steel pan coated by a quasicrystal material. The coating was made using electron beam vapor deposition in vacuum

surface energy. These properties make them promising candidates as coatings for, e.g., cookware (see Fig. 13.1), surgical tools and electrical shavers, automotive parts, and for air-space applications.

In this article we present results related to sliding friction, contact mechanics and adhesion. Most of the theory results are very general, and can be applied not only to quasicrystals but also to other materials. In Sect. 13.2 we study how sliding friction depends on the elastic modulus of the solids. In Sect. 13.3 we discuss sliding friction and adhesion for quasicrystals. Section 13.4 presents a general discussion about surface roughness, and in Sects. 13.5 and 13.6 we consider contact mechanics and adhesion. Section 13.7 contains the summary and an outlook.

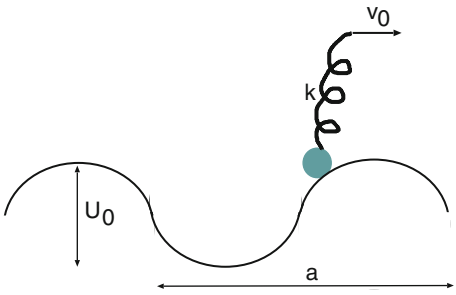


Fig. 13.2 A particle pulled by a spring (with the velocity v_0) in a periodical potential. If the spring k is weak enough or the barrier U_0 high enough ($ka^2 \ll U_0$), the particle will perform stick-slip motion. On the other hand, if $ka^2 \gg U_0$ no stick-slip occurs, and the friction force is very small (it will vanish as $v_0 \to 0$)

13.2 Sliding Friction—Role of Elasticity

Sliding friction for clean solid surfaces, or surfaces separated by a ~ 1 nm (or less) thick contamination film (boundary lubrication), usually originates from elastic instabilities occurring at the interface [2]. Elastic instabilities occur if the elastic modulus of the solids is low enough or if the lateral corrugation of the interaction potential at the interface is high enough. This is best illustrated by a one dimensional model, see Fig. 13.2. Here a particle is connected to a spring, and the free end of the spring is pulled with some (small) velocity v_0. If the spring constant is small enough or the potential well U_0 high enough, the particle will perform stick-slip (non-uniform) motion, where during slip the particle moves with a velocity $v(t)$ which is much higher than (and unrelated to) the driving velocity v_0. This will result in a large friction force (spring force averaged over time). On the other hand, if the spring is very stiff or the barrier very small, no stick-slip occurs and the velocity of the particle will be of the order of v_0, and proportional to v_0. In this case the friction force vanishes, at least when $v_0 \to 0$. In reality, the particle may represent some small group of atoms (block atoms and/or contamination atoms) at the interface, and the spring may represent some effective elastic properties which determine the force necessary to displace the group of atoms relative to the center of mass of the solid walls.

It is important to note that the elastic stiffness of solids depends on the length scale over which they are studied. Thus a solid elastic bar of length L will elongate by a distance proportional to L when exposed to some (fixed) forces F and $-F$ at its two ends. However, since hard solids also tend to have small contact areas (with small average diameter L) when squeezed together, this reduces the chances that elastic instabilities will occur at the interface during sliding. Thus, it is clear that hard materials, such as quasicrystals, may exhibit very low friction, in particular since the surfaces will always be incommensurate, thus lowering the barrier U_0.

As illustrations of the discussion above, let us present Molecular Dynamics simulations for an elastic block sliding on a rigid substrate when the wall atoms are (nearly) incommensurate with the substrate atoms. In Fig. 13.3 we show the center-of-mass coordinate of the bottom layer of block atoms as a function of time. Both the sliding layer and the substrate have square lattice structure, but with different lattice spacing to have (nearly) incommensurability (ratio 1.625 close to the golden mean). The upper surface of the block is moving with the constant speed $v = 0.1$ m/s. When the elastic stiffness of the block is small, stick-slip occurs (red curve), and the friction coefficient is nonzero. For a stiffer block (green curve), the stick and slip behaviour disappears and the friction coefficient gets negligibly small (below the noise level of the simulations).

Recently, a detailed study was performed of the friction between a Si tip and thin hard coatings [3]. As expected, it was observed that the friction coefficient decreases with increasing elastic modulus of the coating. An extreme case is the friction of diamond against diamond where the friction (when the diamond surfaces are passivated by hydrogen) is extremely small (of the order of 0.01).

13.3 Application to Quasicrystals

Quasicrystals differ radically from traditional crystalline materials because they have rotational symmetry which is incompatible with periodicity (translational symmetry). Due to the lack of translational symmetry, the plastic deformation properties of quasicrystals fundamentally differ from those of crystals. The plastic yield stress of most metal crystals is relatively low due to small barriers for motion of dislocations. This is not the case in quasicrystals because of the absence of long-range translational symmetry. Consequently, the plastic yield stress is much higher for quasicrystals than for most metallic crystals. Thus, in spite of the fact that quasicrystals only contain metal atoms, they form relatively hard and brittle-like materials. We believe that this is the main reason for the low sliding friction [4, 5] and wear usually observed for quasicrystal materials.

In one set of experiments [6], the adhesion and sliding friction were studied as a sharp tip coated with W_2C was in contact with a single grain tenfold decagonal $Al_{72.4}Ni_{10.4}Co_{17.2}$ quasicrystals. The coated tip had the radius of curvature $\sim 100\,\text{nm}$. For the clean surfaces in ultrahigh vacuum the work of adhesion was found to be $\approx 0.1\,\text{eV}/\text{Å}^2$, but this value is probably an overestimate of the change in the surface energy $\Delta\gamma = \gamma_1 + \gamma_2 - \gamma_{12}$ since some plastic deformation of the tip-sample contact takes place during rupture of the contact. If the quasicrystal surface is exposed to clean O_2 gas, a very thin oxide layer (one or at most two monolayers) is formed on the surface, and the work of adhesion drops to about $\approx 0.03\,\text{eV}/\text{Å}^2$. When the surface is air-oxidized the work of adhesion is only $\approx 3\,\text{meV}/\text{Å}^2$. Similarly, the

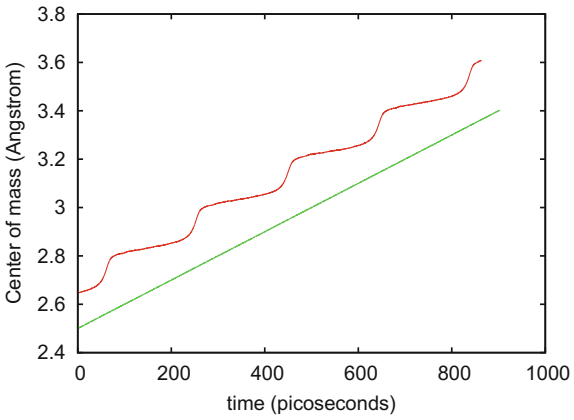


Fig. 13.3 Simulation results for an elastic block sliding on a rigid substrate. The atoms of the bottom surface of the block and of the top surface of the substrate form square lattices which are (nearly) incommensurate. The upper surface of the block is moving with the constant speed $v = 0.1\,\text{m/s}$. *Straight line* (*green*): Young modulus $E = 10\,\text{GPa}$, pressure 1 GPa. Stick and slip (*red*): Young modulus $E = 1\,\text{GPa}$, pressure 0.1 GPa. For the softer elastic solid stick-slip occurs at the interface while steady motion occurs for the stiffer block

friction coefficient drops from ≈ 0.4 for the clean surface to ≈ 0.2 for the surface exposed to O_2 and to ≈ 0.1 for the air-oxidized surface.

It has been reported that the oxide formed in air on the quasicrystal surface has a thickness of the order of 26 Å in dry air and 62 Å in humid air. This is much thicker than the in situ grown oxide (≈ 6 Å). Thus, the higher friction and work of adhesion on the very thin oxide formed in vacuum could be explained by the more fragile nature of the film that can be partly destroyed by the tip resulting in (weak) cold-welded regions [6]. In addition, the air exposed surface is likely to have a nanometer thick contamination layer consisting of organic molecules, water and other contamination molecules. This layer will also reduce the sliding friction although it may be at least partly removed after repeated sliding over the same surface area.

In another experiment two *macroscopic* $Al_{70}Pd_{21}Mn_9$ quasicrystals were brought into contact [7]. The crystal surfaces were polished to a mirror finish with 0.25 μm diamond pasta. The surface roughness amplitude was not measured but should be of the order of several 10 nm. In this case, even after lateral sliding, no adhesive force could be detected during pull-off. This may seem as a paradox taking into account the relatively large pull-off force measured in [6] when a tip was removed from a quasicrystal. However, the result is easy to understand based on the theoretical results presented in Sect. 13.6. Thus, when two macroscopic solid blocks of hard materials with randomly rough surfaces are brought into contact, the actual contact will only occur in very small, randomly distributed, asperity contact areas. For hard materials with low ductility, such as quasicrystals, a root-mean-square roughness of a few 10 nm (as in the present case) is enough to completely remove the (macroscopic) adhesion between the solids for the following reason. Since the asperities have different sizes they will have different amount of elastic deformation, and will act like elastic springs of different sizes. Thus during pull-off the different asperity contact areas will break at different times giving rise to a negligible adhesion even though breaking a single asperity contact region requires a non-negligible force as observed in the tip-substrate experiments reported on in [6]. We point out that a similar effect is observed in silicon wafer bonding (see Sect. 13.4).

For clean surfaces of more ductile metals such as Cu, Au or Al, strong adhesion is usually observed. This is the case even for oxide coated surfaces if sliding occurs before pull-off, as the sliding will break up the oxide coating and result in the formation of cold welded contact areas. During pull-off, because of the high ductility of Cu, Au or Al (and most other metals), "long" metallic bridges may be formed between the solids so that instead of having junctions popping one after another during pull-off, a large number of adhesive junctions may simultaneously impede the surface separation during pull-off (see Fig. 13.4), leading to a large pull-off force.

In [8] sliding friction measurement was performed both for clean surfaces (in ultra high vacuum) and for O_2 exposed surfaces and for surfaces oxidized in the air. For clean surfaces the friction coefficient was of order ≈ 0.6 which dropped to ≈ 0.4 when exposed to O_2. The friction coefficient of air exposed surfaces was only ≈ 0.1.

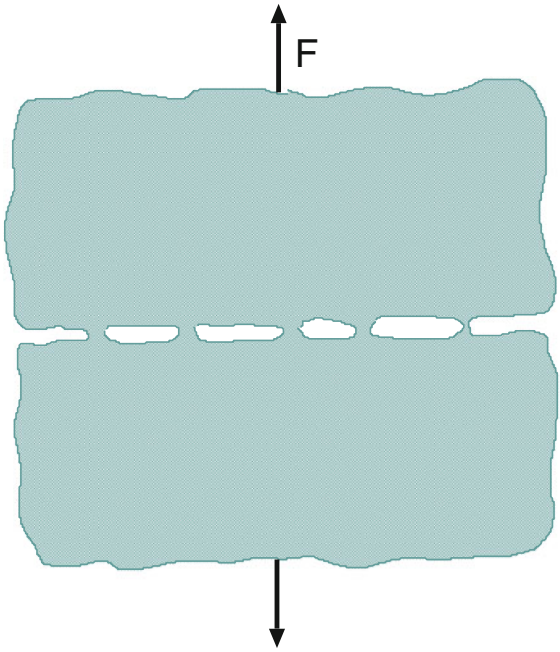


Fig. 13.4 When two ductile metals, e.g., Au or Al, are separated after being in contact, metallic bridges will occur in many asperity contact areas giving rise to a nonzero pull-off force. For (plastically) harder and more brittle metal because of elastic deformation of the asperities, the asperity contact regions will break one after another during pull-off and no adhesion (or pull-off force) will be observed

13.4 Surface Roughness

Surface roughness has a huge influence on many important physical phenomena such as contact mechanics, sealing, adhesion and friction. Thus, for example, experiments have shown that already a substrate roughness with a root-mean-square (*rms*) roughness of order $\sim 1\,\mu m$ can completely remove the adhesion between a rubber ball and a substrate, while nanoscale roughness will remove the adhesion between most hard solids, e.g., metals and minerals; this is the reason why adhesion is usually not observed in most macroscopic phenomena. Similarly, rubber friction on most surfaces of practical interest, e.g., road surfaces, is mainly due to the pulsating forces which act on the rubber surface as it slides over the substrate asperities.

Let us illustrate the importance of surface roughness with three modern applications. At present there is a strong effort to produce small mechanical devises, e.g., micromotors. The largest problem in the development of such devices is the adhesion and, during sliding, the friction and wear between the contacting surfaces [9]. As an example, in Fig. 13.5 we show the simplest possible micro device, namely a micrometer cantilever beam. (Suspended micromachined structures such as plates

Fig. 13.5 **a** Micrometer sized cantilever beam. **b** If the beam is too long or too thin the minimum free energy state corresponds to the beam partly bound to the substrate. Surface roughness lowers the binding energy (per unit area) and hence stabilizes the non-bonded state in (**a**)

and beams are commonly used in manufacturing of pressure and accelerator sensors.) If the beam is too long or too thin the free beam state in (a) will be unstable, and the bound state in (b) will correspond to the minimum free energy state [10]. Roughly speaking, the state (b) is stable if the binding energy to the substrate is higher than the elastic energy stored in the bent beam. The binding energy to the substrate can be strongly reduced by introducing (or increasing) the surface roughness on the substrate (see Sect. 13.6.1). In addition, if the surfaces are covered by appropriate monolayer films the surfaces can be made hydrophobic thus eliminating the possibility of formation of (water) capillary bridges.

A second application is the formation of hydrophobic coatings on surfaces by creating the appropriate type of surface roughness [11]. This involves copying Nature where many plant surfaces are found to be highly hydrophobic (Fig. 13.6) as a result of the formation of special types of surface roughness (Fig. 13.7). The surface roughness allows air to be trapped between the liquid and the substrate, while the liquid is suspended on the tips of the asperities. Since the area of real liquid-substrate contact is highly reduced, the contact angle of the drop is determined almost solely by the surface tension of the liquid, leading to a very large contact angle. New commercial products based on this "Lotus effect", such as self-cleaning paints and glass windows, have been produced.

Finally, we discuss the effect of surface roughness on direct wafer bonding [12]. Wafer bonding at room temperature is due to relatively weak interatomic attraction forces, e.g., the van der Waals interaction or hydrogen bonding, giving (for perfectly

Fig. 13.6 A water droplet on a superhydrophobic surface: The droplet touches the leaf only in a few points and forms a ball. It completely rolls off at the slightest declination. From [11]

flat surfaces) an interfacial binding energy of order $6\,\text{meV}/\text{Å}^2$. The wafer surface roughness is the most critical parameter determining the strength of the wafer bonding. In particular, when the surface roughness exceeds a critical value, the wafers will not bind at all, in agreement with the theory presented in Sect. 13.6.1. Primary grade polished silicon wafer surfaces have *rms* roughness of order $\sim 0.1\,\text{nm}$ when measured over a $10 \times 10\,\mu\text{m}$ surface area, and such surfaces bind spontaneously. However, when the surface roughness amplitude is of order 1 nm the surfaces either bind (slowly) when squeezed together at high enough pressure, or they do not bind at all depending on the detailed nature of the surface roughness power spectra.

Surfaces with "ideal" roughness, e.g., prepared by fracture or by some growth process, have been studied intensively for many years [13–17]. However, much less information has been presented for more common surfaces of engineering interest. In what follows we discuss the nature of the power spectra of some surfaces of practical importance. As illustrations we discuss contact mechanics and adhesion.

13.4.1 Surface Roughness Power Spectra: Definition and General Properties

The influence of roughness on the adhesion and frictional properties described above is mainly determined by the surface roughness power spectra $C(q)$ defined by [18]

$$C(q) = \frac{1}{(2\pi)^2} \int d^2x \, \langle h(\mathbf{x}) h(\mathbf{0}) \rangle e^{-i\mathbf{q}\cdot\mathbf{x}}. \tag{13.1}$$

Here $h(\mathbf{x})$ is the substrate height measured from the average plane defined so that $\langle h \rangle = 0$. The $\langle \ldots \rangle$ stands for ensemble averaging, or averaging over the surface area (see below). We have assumed that the statistical properties of the substrate are translational invariant and isotropic so that $C(q)$ only depend on the magnitude $q = |\mathbf{q}|$ of the wave vector $\mathbf{q}$. Note that from (13.1) follows

Fig. 13.7 A leaf surface with roughness on several length scales optimized (via natural selection) for hydrophobicity and self-cleaning. Through the combination of micro- (cells) and nanostructure (wax crystals) the water contact angle θ_0 is maximized. From [11]

$$\langle h(\mathbf{x})h(\mathbf{0})\rangle = \int d^2q\ C(q)e^{i\mathbf{q}\cdot\mathbf{x}},$$

so that the root-mean-square roughness amplitude $\sigma = \langle h^2\rangle^{1/2}$ is determined by

$$\langle h^2\rangle = \int d^2q\ C(q) = 2\pi \int_0^\infty dq\ qC(q). \tag{13.2}$$

In reality, there will always be an upper, q_1, and a lower, q_0, limit to the q-integral in (13.2). Thus, the largest possible wave vector will be of order $2\pi/a$, where a is some lattice constant, and the smallest possible wave vector is of order $2\pi/L$ where L is the linear size of the surface. In general, one may define a root-mean-square roughness amplitude which depends on the range of roughness included in the integration in (13.2):

$$\langle h^2\rangle(q_0, q_1) = 2\pi \int_{q_0}^{q_1} dq\ qC(q). \tag{13.3}$$

For a randomly rough surface, when $h(\mathbf{x})$ are Gaussian random variables, the statistical properties of the surface are completely defined by the power spectra $C(q)$. In this case the height probability distribution

$$P_h = \langle \delta[h - h(\mathbf{x})]\rangle$$

will be a Gaussian

$$P_h = \frac{1}{(2\pi)^{1/2}\sigma} e^{-h^2/2\sigma^2}.$$

The height distribution of many natural surfaces, e.g., surfaces prepared by fracture, or surfaces prepared by blasting with small particles (e.g., sand blasting or ion sputtering) are usually nearly Gaussian. On the other hand, rough surfaces, e.g., a surface prepared by fracture, which have been (slightly) polished have a non-symmetric height distribution (i.e., no symmetry as $h \to -h$) since the asperity tops have been more polished than the bottom of the valleys, and such surfaces (which are of great practical importance—see below) have non-Gaussian height distribution. For such

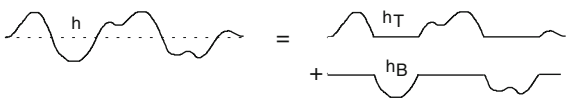


Fig. 13.8 The surface profile $h(x)$ is decomposed into a *top* $h_T(x)$ and a *bottom* $h_B(x)$ profile

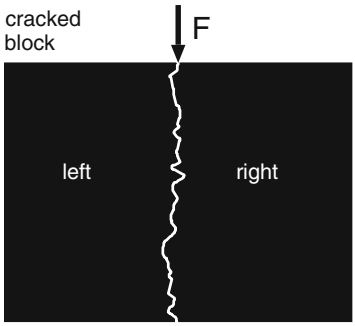


Fig. 13.9 Rough surfaces prepared by crack propagation have surface roughness with statistical properties which must be invariant under the replacement of $h \to -h$. This follows from the fact that what is a valley on one of the crack surfaces (say the *left*) is an asperity with respect to the other crack surface (*right*). Thus the *top* and *bottom* power spectra must obey $C_T(q) = C_B(q)$

surfaces it is interesting to study the *top*, C_T, and the *bottom*, C_B, power spectra's defined by

$$C_T(q) = \frac{1}{(2\pi)^2} \int d^2x \, \langle h_T(\mathbf{x}) h_T(\mathbf{0}) \rangle e^{-i\mathbf{q}\cdot\mathbf{x}}, \tag{13.4a}$$

$$C_B(q) = \frac{1}{(2\pi)^2} \int d^2x \, \langle h_B(\mathbf{x}) h_B(\mathbf{0}) \rangle e^{-i\mathbf{q}\cdot\mathbf{x}}, \tag{13.4b}$$

where $h_T(\mathbf{x}) = h(\mathbf{x})$ for $h > 0$ and zero otherwise, while $h_B(\mathbf{x}) = h(\mathbf{x})$ for $h < 0$ and zero otherwise, see Fig. 13.8. It is easy to show that $C \approx C_T + C_B$. It is also clear by symmetry that for a surface prepared by fracture, $C_T(q) = C_B(q)$, since what is top on one of the cracked block surfaces is the bottom on the other (opposite) crack surface, and vice versa, see Fig. 13.9. However, if the cracked surface is (slightly) polished then, since the polishing will be stronger at the top of the asperities than at the bottom of the valleys [the contact pressure with the polishing object (e.g., sand paper) is highest at the asperity top], $C_B > C_T$. If n_T and n_B are the fraction of the nominal surface area (i.e., the surface area projected on the xy-plane) where $h > 0$ and $h < 0$, respectively, with $n_T + n_B = 1$, then we also define $C_T^*(q) = C_T/n_T$ and $C_B^* = C_B/n_B$. In general, $n_T \approx n_B \approx 0.5$ and for surfaces prepared by fracture $n_T = n_B = 0.5$. Roughly speaking, C_T^* would be the power spectra which would result if the surface profile in the large valleys (for $h < 0$) is replaced by a surface profile with similar short-wavelength roughness as occurs on the large asperities (for $h > 0$). A similar statement holds for C_B^*.

Many surfaces tend to be nearly self-affine fractal. A self-affine fractal surface has the property that if part of the surface is magnified, with a magnification which in general is appropriately different in the perpendicular direction to the surface as compared to the lateral directions, then the surface "looks the same", i.e., the

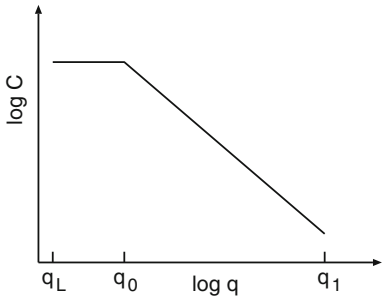


Fig. 13.10 Surface roughness power spectra of a surface which is self affine fractal for $q_1 > q > q_0$. The long-distance roll-off wave vector q_0 and the short distance cut-off wave vector q_1 depend on the system under consideration. The slope of the $\log C - \log q$ relation for $q > q_0$ determines the fractal exponent of the surface. The lateral size L of the surface (or of the studied surface region) determines the smallest possible wave vector $q_L = 2\pi/L$

statistical properties of the surface are invariant under the scale transformation. For a self-affine surface the power spectrum has the power-law behavior

$$C(q) \sim q^{-2(H+1)},$$

where the Hurst exponent H is related to the fractal dimension D_f of the surface via $H = 3 - D_\mathrm{f}$. Of course, for real surfaces this relation only holds in some finite wave vector region $q_0 < q < q_1$, and in a typical case $C(q)$ has the form shown in Fig. 13.10. Note that in many cases there is a roll-off wavelength q_0 below which $C(q)$ is approximately constant. We will discuss this point further below.

Finally, note that while the root-mean-square roughness usually is dominated by the longest wavelength surface roughness components, higher order moments of the power spectra such as the average slope or the average surface curvature are dominated by the shortest wavelength components. For example, assuming a self affine fractal surface, (13.3) gives

$$\langle h^2 \rangle (q_0, q_1) \sim \int_{q_0}^{q_1} dq\, q^{-2H-1} \sim q_0^{-2H} - q_1^{-2H} \approx q_0^{-2H}$$

if $q_1/q_0 \gg 1$. However, the average slope and the average curvature have additional factors of q^2 and q^4, respectively, in the integrand of the q-integral, and these quantities are therefore dominated by the large q (i.e., short wavelength) surface roughness components.

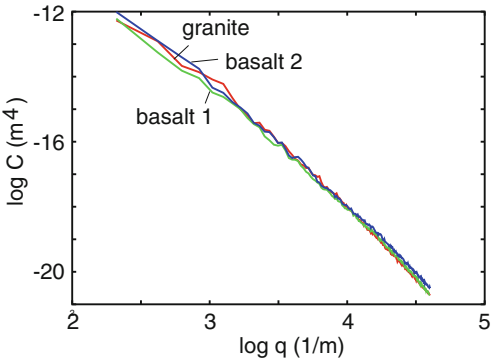


Fig. 13.11 The surface roughness power spectra for two freshly cleaved basalt surfaces and a fresh granite surface

13.4.2 Surface Roughness Power Spectra: Experimental Results

In this section we present power spectra for different surfaces of practical importance. The power spectra have been calculated using (13.1), (13.4a) and (13.4b), where the height profile $h(\mathbf{x})$ has been measured using either an optical method or the Atomic Force Microscope.

13.4.2.1 Surfaces Produced by Crack Propagation

Figure 13.11 shows the power spectra $C(q)$ for three freshly cleaved stone surfaces, namely a granite and two basalt stone surfaces. Here, and in what follows, we show the power spectra on a log-log scale. Note that the granite and basalt surfaces, in spite of the rather different mineral microstructure (see below), give identical power spectra within the accuracy of the measurement. It has been stated (see, e.g., [19]) that surfaces produced by crack propagation have self affine fractal structure with the universal fractal dimension $D_{\rm f} \approx 2.2$. However, our measured $\log C - \log q$ relations are not perfectly straight lines, i.e., the surfaces in the studied length-scale range cannot be accurately described as self affine fractal, and the average slope of the curves in Fig. 13.11 correspond to the fractal dimension $D_{\rm f} \approx 2$ rather than 2.2.

Note the similarity of the power spectra for the basalt and granite surfaces in Fig. 13.11. Granite and basalt both result from magma and have a similar composition, consisting mainly of minerals from the silicate group. However, granite results from magma which is trapped deep in the crust, and it takes very long time to cool down enough to crystallize into solid rock. As a result granite is coarse-textured rock in which individual mineral grains are easily visible. Basalt, on the other hand, results from fast cooling of magma from, e.g., volcanic eruptions, and is therefore fine grained, and it is nearly impossible to see the individual minerals without magnification. In spite of these differences, the surface roughness power spectra of freshly cleaved surfaces are nearly identical. This may indicate some kind of universal power spectrum for surfaces resulting from cleaving of mineral stones of different types.

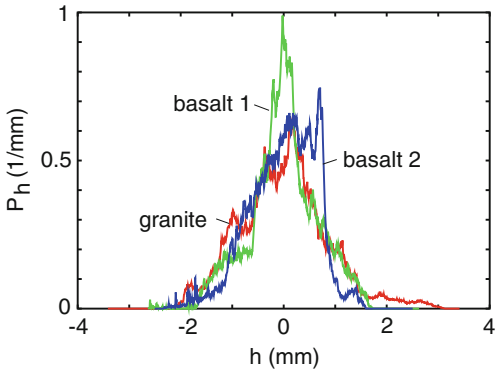


Fig. 13.12 The height distribution P_h for two freshly cleaved (cobble stone) basalt surfaces and a fresh granite surface. Note the random non-Gaussian nature of the height profiles

Note that there is no roll-off region for surfaces produced by fracture (crack propagation), and the surfaces remains fractal-like up to the longest length scale studied, determined by the lateral size L of the surfaces (or of the regions experimentally studied), i.e., with reference to Fig. 13.10, $q_0 = q_L$. One consequence of this is that the *rms*-roughness amplitude is determined mainly by the $\lambda \sim L$ wavelength fluctuations of the surface height, and will therefore depend on the size L of the surface, and the height distribution P_h obtained for any given realization of the rough surface will not be Gaussian, but will exhibit random fluctuations as compared to other realizations (see Fig. 13.12, which illustrates this point for the three stone surfaces discussed above). However, the ensemble averaged height distribution (not shown) should be Gaussian or nearly Gaussian. Thus, when there is no roll-off region in the measured power spectra, averaging over the surface area is not identical to ensemble averaging. However, when there is a roll-off wave vector $q_0 = 2\pi/\lambda_0$, and if the surface is studied over a region with the lateral size $L \gg \lambda_0$, ensemble averaging and averaging over the surface area $L \times L$ will give identical results for P_h, and the *rms*-roughness amplitude will be independent of L for $L \gg \lambda_0$.

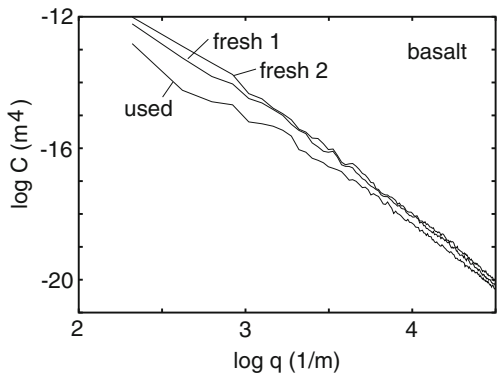


Fig. 13.13 The surface roughness power spectra $C(q)$ for two freshly cleaved cobble stone (basalt) surfaces, and for a used surface

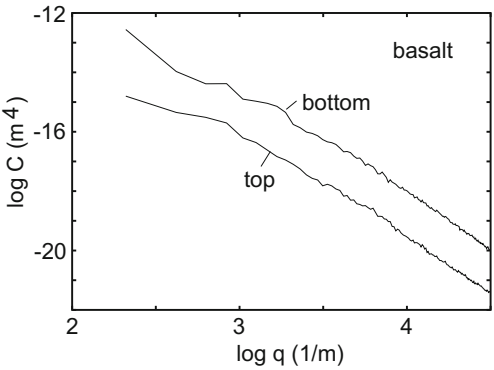


Fig. 13.14 The *top* C_T^* and the *bottom* C_B^* surface roughness power spectra $C(q)$ for a used cobble stone (basalt) surface

13.4.2.2 Polished Crack Surfaces

In the past, cobble stones, made from granite or basalt, were frequently used for road surface pavements. However, these surfaces do not exhibit good frictional properties against rubber. In particular, with increasing time the cobble stone surfaces become polished by the road–tire interaction, which results in a reduced rubber-road friction, even during dry driving conditions. Figure 13.13 illustrates this polishing effect. It shows the power spectrum of a strongly used (basalt) cobble stone, and of two freshly cleaved surfaces (from Fig. 13.11), from the same cobble stone. At long wavelength the power spectrum of the strongly used surface is nearly one decade smaller than that of the freshly prepared surfaces. The effect of the polishing is even better illustrated by calculating the top and bottom power spectra, C_T^* and C_B^*, as shown in Fig. 13.14. The top power spectrum is a factor ~ 30 times smaller than the bottom power spectrum for all wave vectors studied. This arises from the higher polishing of the road asperities than of the valleys (the tire–road contact pressure is highest at the road asperities, resulting in the strongest polishing of the asperity tops during breaking on the road). It is important to take this polishing effect into consideration when designing road pavements.

13.4.2.3 Surfaces with Long-Distance Roll-off

As pointed out above, surfaces prepared by fracture have no natural long-distance cut-off and the *rms* roughness amplitude increases continuously (without limit) as the probed surface area increases. This is similar to Brownian motion where the mean square displacement increases without limit (as $\sim t^{1/2}$) as the time t increases. However, most surfaces of engineering interest have a long distance cut-off or roll-off wavelength λ_0 corresponding to a wave vector $q_0 = 2\pi/\lambda_0$, as shown in Fig. 13.10. For example, if a flat surface is sand blasted for some time the resulting rough surface will have a long distance roll-off length, which increases with the time of sand blasting. Similarly, if atoms or particles are deposited on an initially flat surface

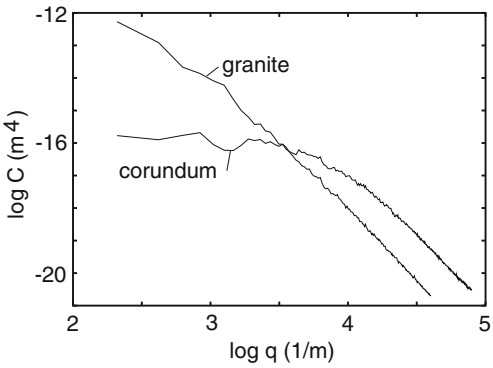


Fig. 13.15 The surface roughness power spectra for a fresh granite surface and a fresh particle-made corundum surface

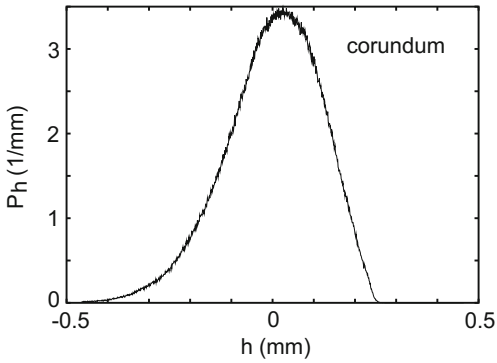


Fig. 13.16 The height distribution P_h as a function of the height h for a particle-made corundum surface

the resulting rough surface will have a roll-off wavelength which increases with the deposition time, as has been studied in detail in recent growth models. Another way to produce a surface with a long-distance roll-off wavelength is to prepare the solid from small particles. A nominally flat surface of such a solid has still roughness on length scales shorter than the diameter of the particles, which therefore may act as a long distance roll-off wavelength. We illustrate this here with a solid produced by squeezing together corundum particles at high temperature and pressure (Fig. 13.15), and for a sand paper surface (Fig. 13.17). For both surfaces the height distribution P_h is smooth and nearly Gaussian (see Figs. 13.16 and 13.18), since averaging over a surface area with lateral size $L \gg \lambda_0$ is equivalent to ensemble averaging.

The sand paper surface in Fig. 13.17 was studied using the AFM at two different resolutions over square areas 20×20 and 100×100 μm as indicated by the two different lines in Fig. 13.17. The height distribution P_h (and hence also the *rms*-roughness

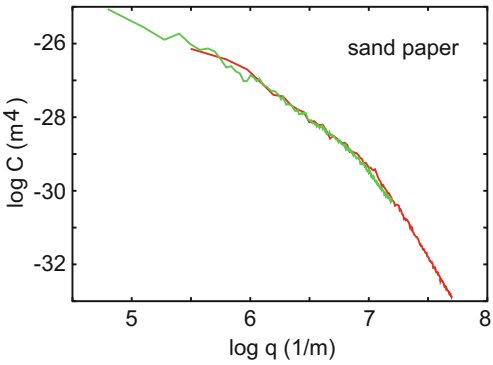


Fig. 13.17 The surface roughness power spectra $C(q)$ for a sand paper surface. The two curves are based on the height profiles measured with an AFM at two different spatial resolutions over 20×20 and 100×100 μm square areas

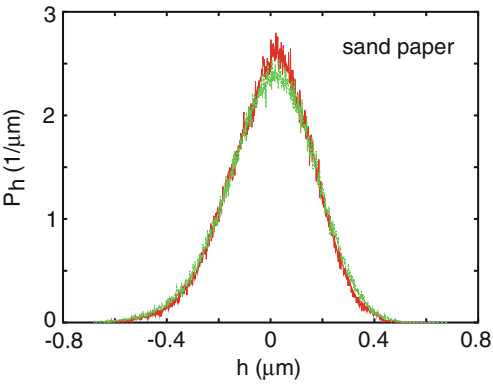


Fig. 13.18 The surface roughness height probability distribution P_h for a sand paper surface. The two curves are based on the height profiles measured with an AFM at two different spatial resolution over 20×20 and 100×100 μm square areas

amplitude) calculated from these two different measurements over different surface areas, see Fig. 13.18, are nearly identical, as indeed expected when L is larger than the roll-off length λ_0.

13.5 Contact Mechanics

Practically all macroscopic bodies have surfaces with roughness on many different length scales. When two bodies with nominally flat surfaces are brought in contact, real (atomic) contact will only occur in small randomly distributed areas, and the area of real contact is usually an extremely small fraction of the nominal contact area. We can visualize the contact regions as small areas where asperities from one solid are squeezed against asperities of the other solid; depending on the conditions the asperities may deform elastically or plastically.

How large is the area of *real* contact between a solid block and the substrate? This fundamental question has extremely important practical implications. For example,

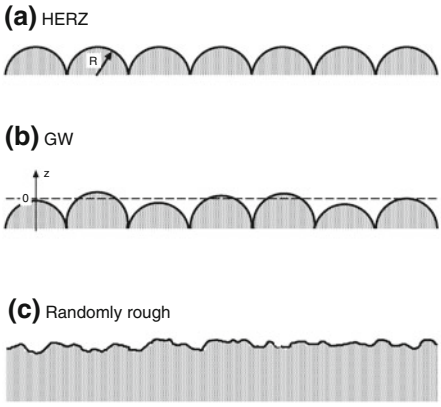


Fig. 13.19 Three models of "rough" surfaces. In case **a** all the "asperities" are equally high and have identical radius of curvature. In this case, according to the Hertz contact theory, the area of real contact ΔA between a solid with a flat surface and the shown surface depends non-linearly on the squeezing force (or load) F_{N} according to $\Delta A \sim F_{\mathrm{N}}^{2/3}$. If the asperities have a random distribution of heights as in (**b**) then, for small F_{N}, ΔA is *nearly* proportional to the squeezing force. If the surface roughness is random with "asperities" of different heights and radius of curvature as in (**c**), the area of real contact for small F_{N} is *exactly* proportional to the squeezing force

it determines the contact resistivity and the heat transfer between the solids. It is also of direct importance for wear and sliding friction [20], e.g., the rubber friction between a tire and a road surface, and it has a major influence on the adhesive force between two solid blocks in direct contact.

Contact mechanics has a long history. The first study was presented by Hertz [21]. He gave the solution for the frictionless normal contact of two elastic bodies of quadratic profile. He found that the area of real contact ΔA varies nonlinearly with the load or squeezing force: $\Delta A \propto F_N^{2/3}$. In 1957 Archard [22] applied the Hertz solution to the contact between rough surfaces and showed that for a simple fractal-like model, where small spherical bumps (or asperities) where distributed on top of larger spherical bumps and so on, the area of real contact varies *nearly linearly* with F_N. A similar conclusion was reached by Greenwood [23], Greenwood and Williamson [24], Johnson [25] who again assumed asperities with spherical summit (of identical radius) with a Gaussian distribution of heights, as sketched in Fig. 13.19b. A more general contact mechanics theory has been developed by Bush et al. [26, 27]. They approximated the summit by paraboloids and applied the classical Hertzian solution for their deformation. The height distribution was described by a random process, and they found that at low squeezing force F_{N} the area of real contact increases linearly with F_{N}.

Figure 13.20 shows the contact between two solids at increasing magnification ζ. At low magnification ($\zeta = 1$) it looks as if complete contact occurs between the solids at many *macro asperity* contact regions, but when the magnification is increased smaller length scale roughness is detected, and it is observed that only

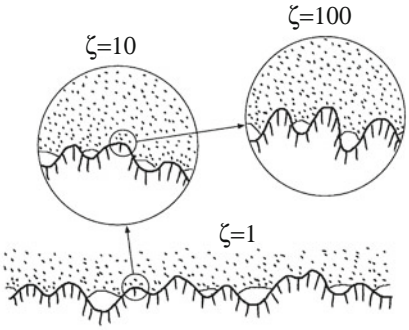


Fig. 13.20 A rubber block (*dotted area*) in adhesive contact with a hard rough substrate (*dashed area*). The substrate has roughness on many different length scales and the rubber makes partial contact with the substrate on all length scales. When a contact area is studied at low magnification ($\zeta = 1$) it appears as if complete contact occurs in the macro asperity contact regions, but when the magnification is increased it is observed that in reality only partial contact occurs

partial contact occurs at the asperities. In fact, if there would be no short distance cut-off the true contact area would vanish. In reality, however, a short distance cut-off will always exist since the shortest possible length is an atomic distance. In many cases the local pressure at asperity contact regions at high magnification will become so high that the material yields plastically before reaching the atomic dimension. In these cases the size of the real contact area will be determined mainly by the yield stress of the solid.

From contact mechanics (see, e.g., [25]) it is known that in the frictionless contact of elastic solids with rough surfaces, the contact stresses depend only upon the shape of the gap between them before loading. Thus, without loss of generality, the actual system may then be replaced by a flat elastic surface [elastic modulus E and Poisson ratio ν, related to the original quantities via $(1-\nu^2)/E = (1-\nu_1^2)/E_1+(1-\nu_2^2)/E_2$] in contact with a rigid body having a surface roughness profile which results in the same undeformed gap between the surfaces.

One of us (Persson) has recently developed a theory of contact mechanics [28, 29], valid for randomly rough (e.g., self affine fractal) surfaces. In the context of rubber friction, which motivated this theory, mainly elastic deformation occurs. However, the theory can also be applied when both elastic and plastic deformations occur in the contact areas. This case is, of course, relevant to almost all materials other than rubber.

The basic idea behind the new contact theory is that it is very important not to a priori exclude any roughness length scale from the analysis. Thus, if $A(\lambda)$ is the (apparent) area of contact on the length scale λ [30] (see Fig. 13.21), then we study the function $P(\zeta) = A(\lambda)/A(L)$ which is the relative fraction of the surface area where contact occurs on the length scale $\lambda = L/\zeta$ (where $\zeta \geq 1$), with $P(1) = 1$. Here $A(L) = A_0$ denotes the macroscopic contact area [L is the diameter of the macroscopic contact area so that $A_0 \approx L^2$].

Consider the system at the length scale $\lambda = L/\zeta$, where L is the diameter of the nominal contact area. We define $q_L = 2\pi/L$ and write $q = q_L\zeta$. Let $P(\sigma, \zeta)$ denote the stress distribution in the contact areas under the magnification ζ. The function $P(\sigma, \zeta)$ satisfies the differential equation (see [28, 29]):

$$\frac{\partial P}{\partial \zeta} = f(\zeta)\frac{\partial^2 P}{\partial \sigma^2}, \tag{13.5}$$

where $f(\zeta) = G'(\zeta)\sigma_0^2$ with

$$G(\zeta) = \frac{\pi}{4}\left(\frac{E^*}{\sigma_0}\right)^2 \int_{q_L}^{\zeta q_L} dq\, q^3 C(q), \tag{13.6}$$

where $E^* = E/(1-\nu^2)$.

Equation (13.5) is a diffusion type of equation, where time is replaced by the magnification ζ, and the spatial coordinate with the stress σ (and where the "diffusion constant" depends on ζ). Hence, when we study $P(\sigma, \zeta)$ on shorter and shorter length scales (corresponding to increasing ζ), the $P(\sigma, \zeta)$ function will become broader and broader in σ-space. We can take into account that detachment actually will occur when the local stress reaches $\sigma = 0$ (we assume no adhesion) via the boundary condition [31]:

$$P(0, \zeta) = 0. \tag{13.7}$$

In order to solve the (13.5) we also need an "initial" condition. This is determined by the pressure distribution at the lowest magnification $\zeta = 1$. If we assume a constant pressure σ_0 in the nominal contact area, then $P(\sigma, 1) = \delta(\sigma - \sigma_0)$.

We assume that only elastic deformation occurs (i.e., the yield stress $\sigma_Y \to \infty$). In this case

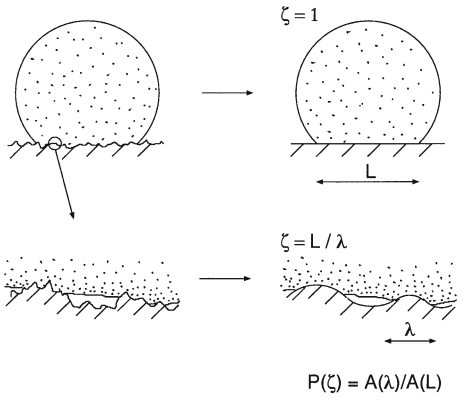


Fig. 13.21 An elastic ball squeezed against a hard, rough, substrate. *Left*: the system at two different magnifications. *Right*: the area of contact $A(\lambda)$ on the length scale λ is defined as the area of real contact when the surface roughness on shorter length scales than λ has been removed

$$P(\zeta) = \int_0^\infty d\sigma P(\sigma, \zeta).$$

When adhesion is taken into account, tensile stresses can occur at the interface between the two solids, and the boundary condition (13.7) is no longer valid [32, 33], see Sect. 13.6.1. It is straightforward to solve (13.5) with the boundary conditions $P(0, \zeta) = 0$ and $P(\infty, \zeta) = 0$ to get

$$P(\zeta) = \frac{2}{\pi} \int_0^\infty dx \frac{\sin x}{x} e^{-x^2 G(\zeta)} = \text{erf}\left(\frac{1}{2\sqrt{G}}\right). \tag{13.8}$$

Note that for small load σ_0, $G \gg 1$ and in this case (13.8) reduces to $P(\zeta) \approx P_1(\zeta)$ where

$$P_1(\zeta) = [\pi G(\zeta)]^{-1/2}. \tag{13.9}$$

Since $G \sim 1/\sigma_0^2$ it follows that the area of real contact is *proportional* to the load for small load. Using (13.8) and (13.9) we can write in a general case

$$P(\zeta) = \text{erf}\left(\frac{\sqrt{\pi}}{2} P_1(\zeta)\right). \tag{13.10}$$

The physical meaning of (13.5) is as follows: When the system is studied at the lowest magnification $\zeta = 1$ no surface roughness can be observed and the block makes (apparent) contact with the substrate everywhere in the nominal contact area. In this case, if we neglect friction at the interface, the stress at the interface will everywhere equal the applied stress σ_0, see Fig. 13.22a, so that $P(\sigma, 1) = \delta(\sigma - \sigma_0)$. When we increase the magnification we observe surface roughness with wavelength down to $\lambda = L/\zeta$. In this case one may observe some non-contact regions as shown in Fig. 13.22b. Since the stress must go continuously to zero at the edges of the boundary between the contact and non-contact regions, it follows that the stress distribution $P(\sigma, \zeta)$ will have a tail extending the whole way down to the zero stress as indicated in Fig. 13.22b (right). There will also be a tail toward larger stresses $\sigma > \sigma_0$ because the average stress must be equal to σ_0. Thus with increasing magnification, the stress distribution will broaden without limit as indicated in Fig. 13.22 (right).

The theory presented above predicts that the area of contact increases linearly with the load for small load. In the standard theory of Greenwood and Williamson [24] this result holds only approximately and a comparison of the prediction of their theory with the present theory is therefore difficult. Bush et al. [26, 27] have developed a more general and accurate contact theory. They assumed that the rough surface consists of a mean plane with hills and valleys randomly distributed on it. The summits of these hills are approximated by paraboloids, the distribution of heights and principal curvatures of which is obtained from the random process theory.

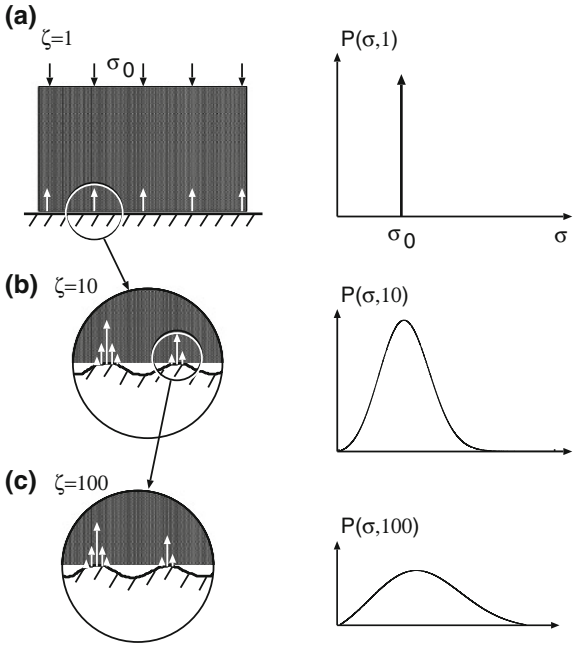


Fig. 13.22 The stress distribution $P(\sigma, \zeta)$ in the contact region between a (rigid) block and an elastic substrate at increasing magnification ζ. **a** At the lowest (engineering) magnification $\zeta = 1$ the substrate surface looks smooth and the block makes (apparent) contact with the substrate in the whole nominal contact area. **b, c** As the magnification increases, we observe that the area of (apparent) contact decreases, while the stress distribution becomes wider and wider

This is to be compared with the GW assumption that the caps of the asperities are spherical each having the same mean radius of curvature. As a result of the more random nature of the surface, Bush et al. found that at small load the area of contact depends linearly on the load accordingly to

$$\frac{A}{A_0} = \kappa \frac{F_N}{E^*} \left(\int d^2q \; q^2 C(q) \right)^{-1/2}, \tag{13.11}$$

where $\kappa = (2\pi)^{1/2}$. This result is very similar to the prediction of the present theory where, for small load, from (13.6) and (13.9), A/A_0 is again given by (13.11) but now with $\kappa = (8/\pi)^{1/2}$. Thus our contact area is a factor of $2/\pi$ smaller than predicted by the theory of Bush et al. Both the theory of Greenwood and Williamson and of Bush et al., assume that the asperity contact regions are independent. However, as discussed in [31], for real surfaces (which always have surface roughness on many different length scales) this will never be the case even at a very low nominal contact pressure. We have argued [31] that this may be the origin of the $2/\pi$-difference

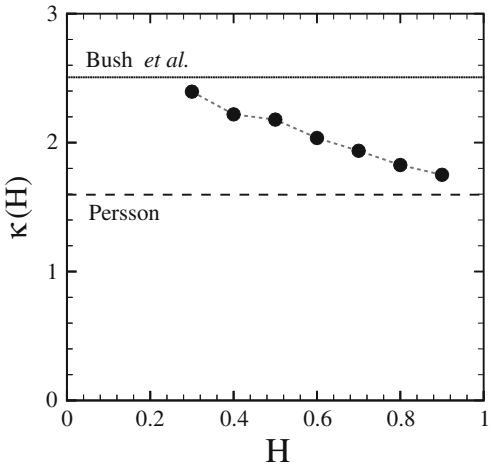


Fig. 13.23 The factor κ as a function of Hurst exponent H for self affine fractal surfaces. The two horizontal lines gives the predictions of the theories of Bush et al. (*solid line*) and Persson (*dashed line*). From [34]

between our theory (which assumes roughness on many different length scales) and the result of Bush et al.

The predictions of the theories of Bush et al. [26, 27] and Persson [28, 29] have been compared to numerical calculations (see [31, 34, 35]). Borri-Brunetto et al. [36] have studied the contact between self affine fractal surfaces using an essentially exact numerical method. They found that the contact area is proportional to the squeezing force for small squeezing forces. Furthermore, it was found that the slope $\alpha(\zeta)$ of the line $A = \alpha(\zeta)F$ decreased with increasing magnification ζ. This is also predicted by the analytical theory (13.11). In fact, it was found a good agreement between the theory and the computer simulations for the change in the slope with magnification and its dependence on the fractal dimension D_f.

Hyun et al. have performed a finite-element analysis of contact between elastic self-affine surfaces. The simulations are done for a rough elastic surface contacting a perfectly rigid flat surface. The elastic solid is discretized into blocks and the surface nodes form a square grid. The contact algorithm identifies all nodes on the top surface that attempt to penetrate the flat bottom surface. The total contact area A was obtained by multiplying the number of penetrating nodes by the area of each square associated with each node. As long as the squeezing force is so small that the contact area is below 10 % of the nominal contact area, i.e., $A/A_0 < 0.1$, the area of real contact is found to be proportional to the squeezing force in accordance with (13.11). In Fig. 13.23 we present the results for the factor κ in (13.11) as a function of Hurst exponent H for self affine fractal surfaces. The two horizontal lines gives the predictions of the theories of Bush et al. (solid line) and Persson (dashed line). The agreement with the analytical predictions is quite good considering the ambiguities in discretization of the surface. The algorithm only considers nodal heights and assumes that contact of a node implies contact over the entire corresponding square. This procedure would be accurate if the spacing between nodes where much smaller than the typical size of asperity contacts. However, the majority of the contact area consists of clusters

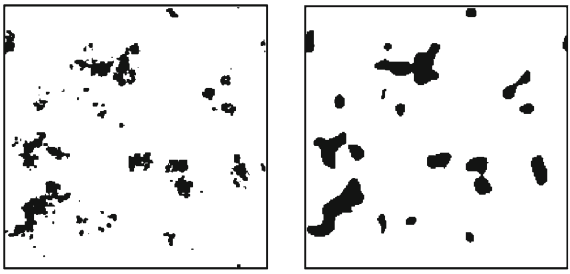


Fig. 13.24 The contact area between an elastic solid block and a randomly rough hard substrate at high (atomic) magnification (*left*), and at a lower magnification (*right*)

containing only one or a few nodes. Since the number of large clusters grows as $H \to 1$, this may explain why the numerical results approach Persson's prediction in this limit.

Hyun et al. also studied the distribution of connected contact regions and the contact morphology. In addition, the interfacial stress distribution was studied and it was found that the stress distribution remained non-zero as the stress $\sigma \to 0$. This violates the boundary condition (13.7) that $P(\sigma, \zeta) = 0$ for $\sigma = 0$. However, it has been shown analytically [31] that for "smooth" surface roughness this latter condition must be satisfied, and we believe that the violation of this boundary condition in the numerical simulations reflects the way the solid was discretized and the way the contact area is defined in the numerical procedure.

Yang et al. [35] have studied contact mechanics using Molecular Dynamics. They also found that the contact area varies linearly with the load for small load, and that the contact area at low magnification is larger than at high magnification (see Fig. 13.24), as predicted by the theory (13.11). The detailed comparison of the simulation results with the theory will be presented elsewhere [35].

Elastic contact theory and numerical simulations show that in the region where the contact area is proportional to the squeezing force, the stress distribution at the interface is independent of the squeezing force. In addition, for an infinite system the distribution of sizes of the contact regions does not depend on the squeezing force (for small squeezing forces). Thus, when the squeezing force increases, new contact regions are formed in such a way that the distribution of contact regions and the pressure distribution remains unchanged. This is the physical origin of Coulombs friction law which states that the friction force is proportional to the normal (or squeezing) force [20], and which usually holds accurately as long as the block-substrate adhesional interaction can be neglected [2].

13.6 Adhesion

In this section we discuss adhesion between rough surfaces. We point out that even when the force to separate two solids vanishes, there may still be a finite contact area (at zero load) between two solids as a result of the adhesional interaction between

the solids. We also study the adhesion between a thin elastic film and a randomly rough, rigid substrate.

13.6.1 Adhesion Between Rough Surfaces

A theory of adhesion between an elastic solid and a hard randomly rough substrate must take into account that partial contact may occur between the solids on all length scales. For the case where the substrate surface is self affine fractal theory shows that when the fractal dimension is close to 2, complete contact typically occurs in the macro asperity contact areas (the contact regions observed when the system is studied at a magnification corresponding to the roll-off wavelength $\lambda_0 = 2\pi/q_0$ of the surface power spectra, see Fig. 13.10), while when the fractal dimension is larger than 2.5, the area of (apparent) contact decreases continuously when the magnification is increased. An important result is that even when the surface roughness is so high that no adhesion can be detected in a pull-off experiment, the area of real contact (when adhesion is included) may still be several times larger than when the adhesion is neglected. Since it is the area of real contact which determines the sliding friction force, *the adhesion interaction may strongly affect the friction force even when no adhesion can be detected in a pull-off experiment*.

The influence of surface roughness on the adhesion between rubber (or any other elastic solid) and a hard substrates has been studied in a classic paper by Fuller and Tabor [37] (see also [38–44]). They found that already a relative small surface roughness can completely remove the adhesion. In order to understand the experimental data they developed a very simple model based on the assumption of surface roughness on a single length scale. In this model the rough surface is modeled by asperities all of the same radius of curvature and with heights following a Gaussian distribution. The overall contact force was obtained by applying the contact theory of Johnson et al. [45] to each individual asperity. The theory predicts that the pull-off force, expressed as a fraction of the maximum value, depends upon a single parameter, which may be regarded as representing the statistically averaged competition between the compressive forces exerted by the higher asperities trying to prize the surfaces apart and the adhesive forces between the lower asperities trying to hold the surfaces together. This picture of adhesion developed by Tabor and Fuller would be correct *if* the surfaces had roughness on a single length scale as assumed in their study. However, when roughness occurs on many different length scales, a qualitatively new picture emerges [32, 33], where, e.g., the adhesion force may even vanish (or at least be strongly reduced), if the rough surface can be described as a self affine fractal with fractal dimension $D_f > 2.5$. Even for surfaces with roughness on a single length scale, the formalism used by Fuller and Tabor is only valid at "high" surface roughness, where the area of real contact (and the adhesion force) is very small. The theory presented below is particularly accurate for "small" surface roughness, where the area of real contact equals the nominal contact area.

13.6.1.1 Qualitative Discussion

Let us estimate the energy necessary in order to deform a rubber block so that the rubber fills up a substrate cavity of height h and width λ. The elastic energy stored in the deformation field in the rubber is given by

$$U_{\rm el} \approx \frac{1}{2} \int d^3x \; \sigma\epsilon,$$

where the stress $\sigma \approx E\epsilon$, where E is the elastic modulus. The deformation field is mainly localized to a volume $\sim \lambda^3$ (see Fig. 13.25) where the strain $\epsilon \approx h/\lambda$. Thus we get $U_{\rm el} \approx \lambda^3 E(h/\lambda)^2 = E\lambda h^2$.

Let us now consider the role of the rubber–substrate adhesion interaction. As shown above, when the rubber deforms and fills out a surface cavity of the substrate, an elastic energy $U_{\rm el} \approx E\lambda h^2$ will be stored in the rubber. Now, if this elastic energy is smaller than the gain in adhesion energy $U_{\rm ad} \approx \Delta\gamma\lambda^2$, where $\Delta\gamma = \gamma_1 + \gamma_2 - \gamma_{12}$ is the change of surface free energy (per unit area) upon contact due to the rubber-substrate interaction (which usually is mainly of the van der Waals type), then (even in the absence of an external load $F_{\rm N}$) the rubber will deform *spontaneously* to fill out the substrate cavities. The condition $U_{\rm el} = U_{\rm ad}$ gives $h/\lambda \approx (\Delta\gamma/E\lambda)^{1/2}$. For example, for very rough surfaces with $h/\lambda \approx 1$, and with parameters typical for rubber $E = 1\,{\rm MPa}$ and $\Delta\gamma = 3\,{\rm meV/\AA^2}$, the adhesion interaction will be able to deform the rubber and completely fill out the cavities if $\lambda < 0.1\,\mu{\rm m}$. For very smooth surfaces $h/\lambda \sim 0.01$ or smaller, so that the rubber will be able to follow the surface roughness profile up to the length scale $\lambda \sim 1\,{\rm mm}$ or longer.

The argument given above shows that for elastic solids with surface roughness on a *single length scale* λ, the competition between adhesion and elastic deformation is characterized by the parameter $\theta = Eh^2/\lambda\delta \approx U_{\rm el}/U_{\rm ad}$, where h is the amplitude of the surface roughness and $\delta = 4(1-\nu^2)\Delta\gamma/E$ the so called *adhesion length*, ν being the Poisson ratio of the rubber. The parameter θ is the ratio between the elastic energy and the surface energy stored at the interface, assuming that complete contact occurs. When $\theta \gg 1$ only partial contact occurs, where the elastic solids make contact only close to the top of the highest asperities, while complete contact occurs when $\theta \ll 1$.

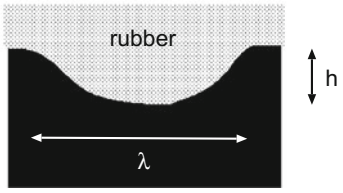


Fig. 13.25 A rubber surface is "pulled" into a cavity in a hard solid by the rubber-substrate adhesional interaction. The elastic energy stored in the deformation field is of order $E\lambda h^2$

13.6.1.2 Pull-off Force

Consider a rubber ball (radius R_0) in adhesive contact with a perfectly smooth and hard substrate. The elastic deformation of the rubber can be determined by minimizing the total energy which is the sum of the (positive) elastic energy stored in the deformation field in the rubber ball, and the (negative) binding energy between the ball and the substrate at the contact interface. The energy minimization gives the pull-off force [45, 46]

$$F_c = (3\pi/2)R_0\Delta\gamma. \tag{13.12}$$

Consider now the same problems as above, but assume that the substrate surface has roughness described by the function $z = h(\mathbf{x})$. We assume that the surface roughness power spectra has a roll-off wavelength $\lambda_0 = 2\pi/q_0$ (see Fig. 13.10) which is smaller than the diameter of the nominal contact area between the two solids. In this case we can still use the result (13.12), but with $\Delta\gamma$ replaced by $\gamma_{\rm eff}$. The effective interfacial energy $\gamma_{\rm eff}$ is the change in the interfacial free energy when the elastic solid is brought in contact with the rough substrate. $\gamma_{\rm eff}(\zeta)$ depends on the magnification ζ, and the interfacial energy which enters in the rubber ball pull-off experiment is the macroscopic interfacial energy, i.e., $\gamma_{\rm eff}(\zeta)$ for $\zeta = 1$. If A_0 is the nominal contact area and A_1 the true atomic contact area, then

$$A_0\gamma_{\rm eff}(1) = A_1\Delta\gamma - U_{\rm el}, \tag{13.13}$$

where $U_{\rm el}$ is the elastic energy stored at the interface as a result of the elastic deformations necessary in order to bring the solids in atomic contact in the area A_1.

13.6.1.3 Stress Probability Distribution

The theory in [32, 33] is based on the contact mechanics formalism described in Sect. 13.4.1. Thus, we focus on the stress probability distribution function $P(\sigma, \zeta)$ which satisfies (13.5):

$$\frac{\partial P}{\partial \zeta} = f(\zeta)\frac{\partial^2 P}{\partial \sigma^2}.$$

We assume that detachment occurs when the local stress on the length scale L/ζ reaches $-\sigma_a(\zeta)$. Thus, the following boundary condition is valid in the present case

$$P(-\sigma_a(\zeta), \zeta) = 0.$$

This boundary condition replaces the condition $P(0, \zeta) = 0$ valid in the absence of adhesion (see Sect. 13.4.1).

Let us consider the system on the characteristic length scale $\lambda = L/\zeta$. The quantity $\sigma_a(\zeta)$ is the stress necessary to induce a detached area of width λ. This stress can be

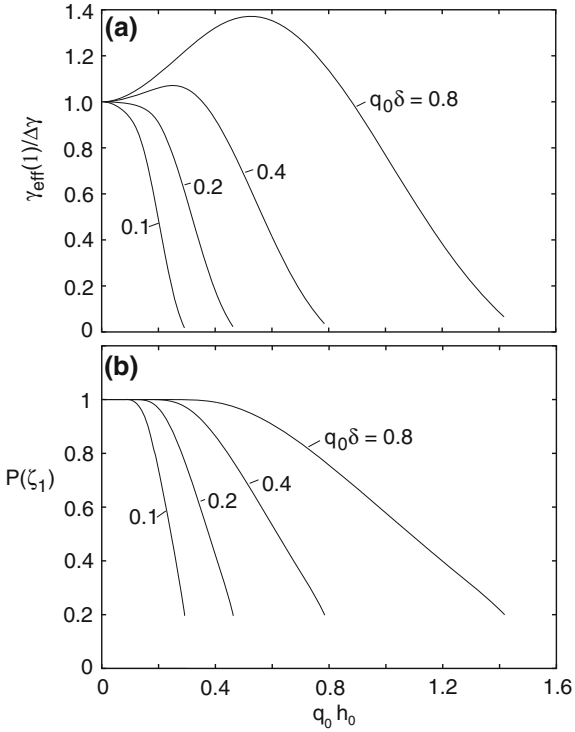


Fig. 13.26 **a** The macroscopic interfacial energy as a function of the dimensionless surface roughness amplitude q_0h_0. **b** The normalized area of real contact, $P(\zeta_1) = A(\zeta_1)/A_0$, as a function of q_0h_0. The curves correspond to different adhesion energies: $q_0\delta = 0.1$, 0.2, 0.4 and 0.8 as indicated. For $H = 0.8$ and $q_1/q_0 = \zeta_1 = 100$

obtained from the theory of cracks, where for a penny-shaped crack of diameter λ

$$\sigma_a = \left[\frac{\pi\gamma_{\rm eff}(\zeta)E}{(1-\nu^2)\lambda}\right]^{1/2} = \left[\frac{\gamma_{\rm eff}(\zeta)Eq}{2(1-\nu^2)}\right]^{1/2}, \tag{13.14}$$

where $q = 2\pi/\lambda = \zeta q_L$. In [32, 33] we have derived two equations for $\gamma_{\rm eff}(\zeta)$ and $P(\zeta)$ which determine how these quantities depend on the magnification ζ; those equations are the basis for the numerical results presented below.

13.6.1.4 Numerical Results

Figure 13.26 shows (a) the effective interfacial energy $\gamma_{\rm eff}(\zeta)$ ($\zeta = 1$) and (b) the normalized area of real contact, $P(\zeta_1) = A(\zeta_1)/A_0$, as a function of q_0h_0, h_0 being the surface r.m.s. roughness and q_0 the roll-off wave vector. Results are shown for different adhesion lengths $\delta = 4(1-\nu^2)\Delta\gamma/E$: $q_0\delta = 0.1, 0.2, 0.4$ and 0.8. We will refer to $\gamma_{\rm eff}(1)$ at the magnification $\zeta = 1$ as the *macroscopic* interfacial free energy which can be deduced from, e.g., the pull off force for a ball according to (13.12). Note that for $q_0\delta = 0.4$ and 0.8 the macroscopic interfacial energy first increases

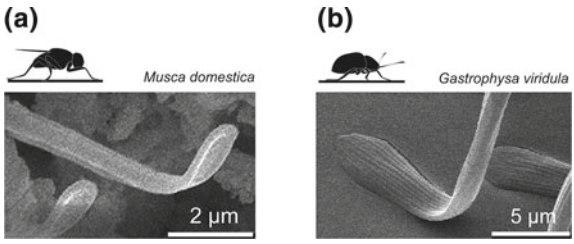


Fig. 13.27 Insect attachment systems consist of fibers or hair which terminates with leaf-like plates which can easily deform (without storing a lot of elastic energy) to bind strongly even to very rough substrates

with increasing amplitude h_0 of the surface roughness, and then decreases. The increase in $\gamma_{\rm eff}$ arises from the increase in the surface area. As shown in Fig. 13.26b, for small h_0 the two solids are in complete contact, and, as expected, the complete contact remains to higher h_0 as $\delta \sim \Delta\gamma/E$ increases. Note also that the contact area is nonzero even when $\gamma_{\rm eff}(1)$ is virtually zero: the fact that $\gamma_{\rm eff}(1)$ (nearly) vanish does not imply that the contact area vanish (even in the absence of an external load), but imply that the (positive) elastic energy stored at the interface just balance the (negative) adhesion energy from the area of real contact. *The stored elastic energy at the interface is given back when removing the block, and when* $\gamma_{\rm eff}(1) \approx 0$ *it is just large enough to break the block-substrate bonding.*

13.6.1.5 Plate Adhesion

In this section we discuss the adhesion of a thin elastic plate to a randomly rough hard substrate. This topic is important for many applications such as thin films used as protective coatings [47], for the manufacturing of multilayered wafer structures [48], or in bio-films for orthopedic implants [49]. The problem under consideration is also of great importance for understanding the adhesion of flies, bugs, and lizards to a rough substrate (see Fig. 13.27), [50, 51] or the adhesive behavior of recently biologically-inspired adhesive films [52].

Here we consider in detail the case of a thin plate in partial contact with a hard substrate with a self-affine fractal rough surface. Figure 13.28 (thick lines) shows (a) the macroscopic interfacial energy γ_{eff} (1), i.e. the effective interfacial energy calculated at the magnification $\zeta = 1$, and (b) the normalized area of real contact $P(\zeta_1)$ at the maximum magnification $\zeta = \zeta_1$, as a function of the dimensionless roughness amplitude $q_0 h_0$. We show results for three different values of $q_0\delta$. The results are for $H = 0.8$, i.e. $D_f = 2.2$, and for a dimensionless thickness of the plate equal to $q_0 d = 0.63$. Note that the macroscopic interfacial energy initially increases with the amplitude h_0 of the rough profile up to a maximum value, and after decreases with h_0. This is caused by the increase of the real contact area produced by the fine structure of the rough profile. Figure 13.26b shows, indeed, that at small h_0 the plate adheres in full contact to the substrate, so that an increase of the surface roughness produces a corresponding increases of the area of contact and, hence, of the surface energy. However this

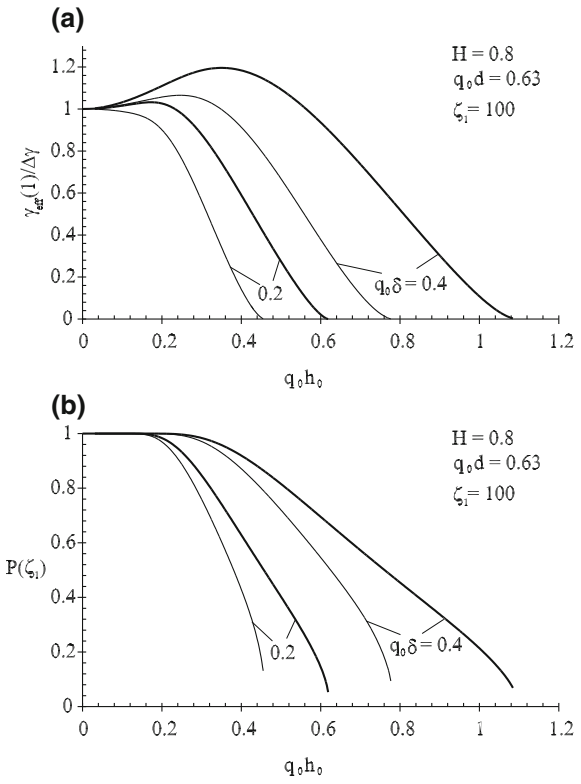


Fig. 13.28 **a** The normalized macroscopic interfacial energy and **b** the normalized area of real contact, as a function of the dimensionless surface roughness amplitude $q_0 h_0$. Thick lines are for the plate case and thin lines are for the semi-infinite solid case. Results are for $H = 0.8$ and $q_0 d = 0.63$, and for two different values of $q_0 \delta$

is no more true at large h_0, because of the reduction of the area of real contact. Figure 13.26 also shows that, as expected, the roughness-induced increment of the macroscopic interfacial energy grows by increasing the adhesion length $\delta \sim \Delta\gamma / E$, and that the full contact condition remains to higher amplitude h_0 as δ increases.

In Fig. 13.28 we compare the results obtained for the plate case (thick lines) with those of the semi-infinite solid (thin lines). As expected, because of the higher compliance of the plate, both the macroscopic interfacial energy $\gamma_{\text{eff}}(1)$ and the normalized area of real contact $P(\zeta_1)$ are larger than for the semi-infinite solid case.

To summarize, at small magnification (long length scales) the plate, because of its higher compliance, is able to adhere in apparent full contact to the long wavelength corrugation of the underlying surface. That is, at length scales longer than the plate thickness, the gain in the adhesion energy upon the contact with the substrate overcomes the repulsive elastic energy produced by the elastic deformations, and the plate is able to fill out the large cavities of the rigid substrate. This produces a larger area of contact and an enhanced capability to adhere to a rough surface in comparison to the semi-infinite elastic solid case. However, at large enough magnification (small length scales) the plate behaves as a semi-infinite solid, and, depending on the

roughness statistical properties, the area of true atomic contact may be much smaller than the nominal contact area.

13.6.1.6 Experimental Manifestations

Unfortunately, the surface roughness power spectrum has not been measured for any surface for which adhesion has been studied in detail. Instead only the roughness amplitude (center line average) and the radius of curvature of the largest surface asperities was determined. Nevertheless, the experimental data of Fuller and Tabor [37], Briggs and Briscoe [40] and Fuller and Roberts [41] are in good qualitative agreement with our theoretical results. In Fig. 13.29 we show the macroscopic interfacial energy for "hard" and "soft" rubber in contact with Perspex, as a function the substrate (Perspex) roughness amplitude as obtained by Briggs and Briscoe [40]. It is not possible to compare these results quantitatively with the theory developed above since the power spectrum $C(q)$ was not measured for the Perspex substrate. Even if the surfaces would be self affine fractal as assumed above, not only the surface roughness amplitude will change from one surface to another, but so will the long distance cut off length λ_0 and hence also the ratio $\zeta_1 = q_1/q_0$. In the experiments reported on in [40] the Perspex surfaces where roughened by blasting with fine particles. The roughness could be varied through the choice of the particles and the air pressure.

One practical problem is that most rubber materials have a wide distribution of relaxation times, extending to extremely long times. This effect is well known in the context of rubber friction (see Sect. 13.6.1), where measurements of the complex elastic modulus show an extremely wide distribution of relaxation times, resulting in large sliding friction even at very low sliding velocities, $v < 10^{-8}$ m/s.

The effect of the stored elastic energy on adhesion has recently been studied using a polyvinylsiloxane rubber block squeezed against a smooth glass surface for

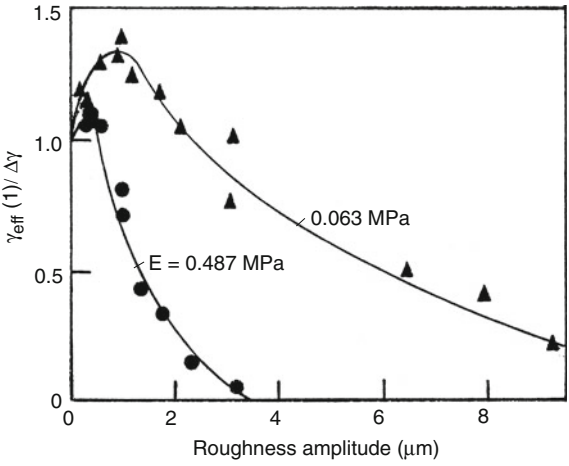


Fig. 13.29 The macroscopic interfacial energy (obtained from the pull-off force) for a smooth rubber surface (ball) in contact with Perspex surface as a function of the roughness (center line average) of the Persplex. Results are shown for a "soft" rubber ($E = 0.063$ MPa) and a "hard" rubber ($E = 0.487$ MPa). From [40]

a fixed time period before measuring the pull-off force [53]. The square-symbols in Fig. 13.30 show the pull-off force as a function of the squeezing force. For squeezing forces $F_N > 850$ mN the pull off force decreases. This may be explained by a drastic increase of the elastic energy stored in the rubber because of the strong deformation of the rubber (which remains even when the load is removed as a result of the rubber–glass friction at the interface), see Fig. 13.31 (top). This energy, freed during the process of unloading, will help to break the adhesive bonds at the interface. This effect is even stronger when the surface is structured. Thus, the triangles in the figure shows the pull-off force when the rubber surface is covered by a regular array of rubber cylindrical asperities. In this case the pull-off force drops to nearly zero for $F_N > 700$ mN. Visual inspection shows that in this case the cylindrical asperities at high load bend and make contact with the glass on one side of the cylinder surface, see Fig. 13.31 (bottom). This again stores a lot of elastic energy at the interface which is given back during pull-off, reducing the pull-off force to nearly zero.

13.6.1.7 The Role of Plastic Yielding on Adhesion

When the local stress in the asperity contact regions between two solids becomes high enough, at least one of the solids yields plastically. This will tend to increase the effective adhesion (or pull-off force) for the following three reasons. First, the area of real contact between the solids will increase as compared to the case where the deformations are purely elastic. Secondly, the amount of stored elastic energy in the contact regions (which is given back during pull-off) will be reduced because of the lowered elastic deformations. Finally, for many materials plastic yielding will strengthen the junctions [54]. For example, most metals are protected by thin oxide layers, and as long as these are intact the main interaction between the surfaces in the contact areas may be of the van der Waals and electrostatic origin. However, when plastic yielding occurs it may break up the oxide films resulting in direct metal-metal contact and the

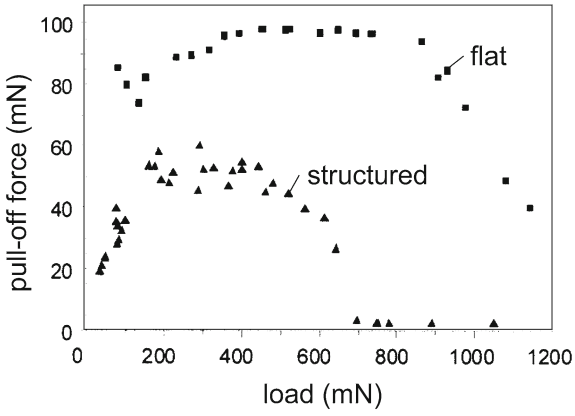


Fig. 13.30 The pull-off force as a function of the squeeze force or load. For silicon rubber in contact with a smooth glass surface. From [53]

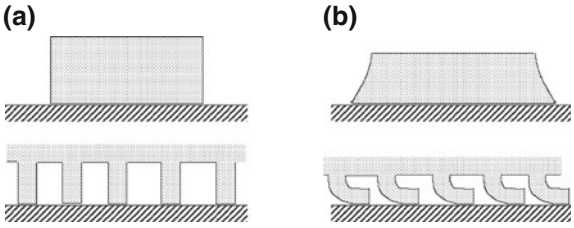


Fig. 13.31 Elastic deformation of a rubber block with a smooth surface (*top*) and a structured surface (*bottom*). **a** shows the initial state before applying a squeezing force, and **b** the new state (without load) after applying (and then removing) a very large squeezing force. In state (**b**) a lot of elastic energy is stored in the rubber which is "given back" during pull-off resulting in a nearly vanishing pull-off force

formation of "cold-welded" junctions. When this occurs, because of the high ductility of many metals, during pull-off "long" metallic bridges may be formed between the solids so that instead of having junctions popping one after another during pull-off, a large number of adhesive junctions may simultaneously impede the surface separation during pull-off, leading to a large pull-off force. However, experiment have shown [8] that just squeezing before pull-off will in general only result in very few cold welded junctions, while squeezing *and* sliding will break up the oxide film, resulting in the formation of many more cold welded contact regions, and will hence result in a much larger pull-off force.

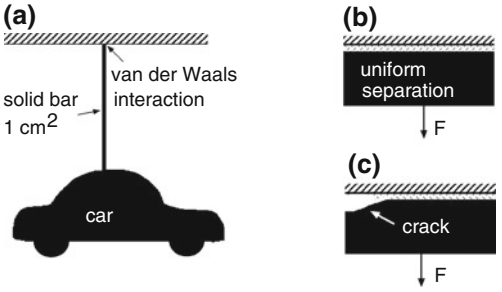


Fig. 13.32 Even the weakest force in Nature of relevance in condensed matters physics, namely the van der Waals force, is relative strong on a macroscopic scale. Thus, for example, if the bond breaking occur uniformly over the contact area as in (**b**), already a contact area of order 1 cm^2 can sustain the weight of a car (i.e., a force of order 10^4 N) [see (**a**)]. However, on a macroscopic scale the bond-breaking does not usually occur uniformly over the contact area, but by crack propagation, see (**c**), which drastically reduce the pull-off force. In addition, interfacial surface roughness drastically reduces the pull-off force

13.6.2 The Adhesion Paradox

The biggest "mystery" related to adhesion is not why it is sometimes observed but rather why it is usually not observed. Thus, even the weakest force in Nature of relevance in condensed matter physics, namely the van der Waals force, is relatively strong on a macroscopic scale. For example, even a contact area of order $1\,\mathrm{cm}^2$ can sustain the weight of a car (i.e., a force of order 10^4 N) [see Fig. 13.32a] also when only the van der Waals interaction operates at the interface. [Here we have assumed that the bond breaking occurs uniformly over the contact area as illustrated in Fig. 13.32b.] However, this is never observed in practice and this fact is referred to as the *adhesion paradox*.

There are several reasons why adhesion is usually not observed between macroscopic bodies. For example, on a macroscopic scale the bond-breaking usually does not occur uniformly as in Fig. 13.32b, but occurs by crack propagation, see Fig. 13.32c. The local stress at the crack tip is much higher than the average stress acting in the contact area, and this drastically reduces the pull-off force. Another reason, already addressed in Sect. 13.6.1, is the influence of surface roughness. Thus, for elastically hard surfaces the true (atomic) contact between the solids at the interface is usually much smaller than the nominal contact area. In addition, the elastic energy stored in the solids in the vicinity of the contact regions is given back during pull-off and helps to break the interfacial bonds between the solids (see Sect. 13.6.1).

It is interesting to note that for very small solid objects, typically of order $100\,\mu\mathrm{m}$ or smaller, the bond breaking may occur uniformly over the contact area (no crack propagation) so that adhesion between smooth surfaces of small objects, e.g., in micromechanical applications (MEMS), may be much stronger than for macroscopic bodies, and this fact must be taken into account when designing MEMS [55, 56].

13.6.3 The Role of Liquids on Adhesion Between Rough Solid Surfaces

As explained in Sect. 13.6.1, surface roughness reduces the adhesion between clean surfaces. First, it lowers the area of real contact. Since the adhesion interaction comes almost entirely from the area where the solids make atomic contact, it is clear that the surface roughness may drastically reduce the adhesion. Secondly, elastic deformation energy is stored in the vicinity of the asperity contact regions. During pull-off the elastic energy is "given back" to the system, usually resulting in a drastic reduction in the effective adhesion and the pull-off force.

Most surfaces have at least nano-scale roughness, and hard solids in the normal atmosphere have at least a monolayer of liquid-like "contamination" molecules, e.g., water and hydrocarbons. Small amount of (wetting) lubricant or contamination liquids between rough solid walls may drastically enhance the adhesion. Thus, for surfaces with nanoscale roughness, a monolayer of a wetting liquid may result in the

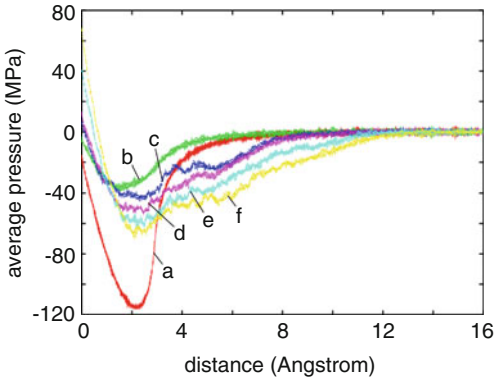


Fig. 13.33 The variation of the average pressure during retraction developed as the block moves a distance of 16 Å away from the substrate. Octane C_8H_{18} was used as lubricant. Pull-off (retraction) velocity was $v_z = 1$ m/s. **a** For the flat substrate without lubricant. **b** For the corrugated substrate without lubricant. Curves **c**–**f** show results for the corrugated substrate with about 1/8, 1/4, 1/2 and 1 monolayer of octane in the contact region, respectively. For clarity, the curve for the flat substrate (**a**) is displaced to the right, by 2 Å

formation of a large number of nano-bridges between the solids, which increases the pull-off force. This effect is well known experimentally. For example, the adhesion force which can be detected between gauge blocks (steel blocks with very smooth surfaces) is due to the formation of many very small capillary bridges made of water or organic contamination. For thicker lubrication or contamination films the effective adhesion will be more long-ranged but the pull-off force may be smaller. The thickness of the lubricant or contamination layer for which the pull-off force is maximal will in general depend on the nature of the surface roughness, but is likely to be of order the root-mean-square roughness amplitude. In fact, it is an interesting and important problem to find out at exactly what liquid thickness the pull-off force is maximal.

Some insects such as flies or crickets inject a thin layer of a wetting liquid in the contact region between the insect attachment surfaces and the (rough) substrate. The optimum amount of injected liquid will depend on the nature of the substrate roughness, and it is likely that the insect can regulate the amount of injected liquid by a feedback system involving the insect nerve system.

Here we consider the adhesion between two solid elastic walls with nanoscale roughness, lubricated by octane [43, 44, 57]. We consider two types of substrates (bottom surface)—flat and nano-corrugated (corrugation amplitude 1 nm and wavelength of the corrugation in x and y direction, 4 nm)—and varied the lubricant coverage from $\sim 1/8$ to ~ 4 monolayers of octane. The upper surface (the block) is assumed to be atomically smooth but with a uniform cylinder curvature with a radius of curvature $R \approx 100$ nm (see Fig. 13.35). The results presented here have been obtained using standard molecular dynamics calculations [43].

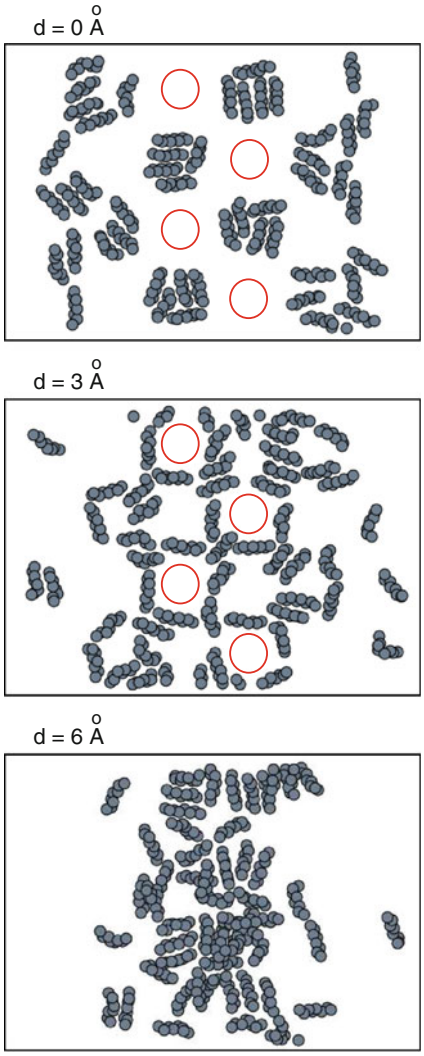


Fig. 13.34 Snapshot pictures (for three different block positions $d = 0$, 3 and 6 Å) of the lubricant layer during retraction. We only show the lubricant molecules in the central part of the contact area between the block and the substrate surfaces (*top view*, surfaces parallel to the plane of the image). Octane C_8H_{18} was used as lubricant. Pull-off (retraction) velocity was $v_z = 1$ m/s. For the corrugated substrate with about 1/4 monolayer of octane in the contact region. The *circles* indicate the position of several asperity tops of the corrugated substrate surface

Figure 13.33 shows the variation of the average pressure during retraction as the block moves a distance of 16 Å away from the substrate. The pull-off (retraction) velocity was $v_z = 1$ m/s. We have varied the lubricant coverage from 0 to 1 monolayer in the contact region. The pull-off force is maximal when the adsorbate coverage

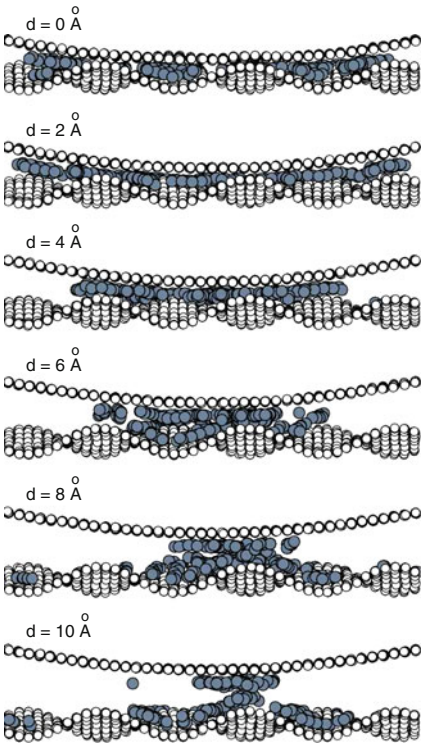


Fig. 13.35 Snapshot pictures (for six different block positions) during retraction. The snapshot pictures show the side view of the central 108 Å × 50 Å section (in the xy-plane) of the contact area. Octane C_8H_{18} was used as lubricant. Pull-off (retraction) velocity was $v_z = 1$ m/s. For the corrugated substrate with about 1/4 monolayer of octane in the contact region

is of the order of one monolayer [curve (f)]. However, the pull-off force is still smaller than for a *flat* substrate without lubricant [curve (a)]. As a function of the octane coverage (for the corrugated substrate) the pull-off force first increases as the coverage increases from zero to ~ 1 monolayer, and then decreases as the coverage is increased beyond monolayer coverage (not shown).

At low octane coverage, the octane molecules located in the substrate corrugation wells during squeezing, are pulled out of the wells during pull-off, forming a network of nano capillary bridges around the substrate nanoasperities, thus increasing the adhesion between two surfaces, see Figs. 13.34 and 13.35. For greater lubricant coverages a single capillary bridge is formed.

Let us discuss the nature of the adhesion for the corrugated substrate, with about 1/4 monolayer of octane in the contact region. Figure 13.34 shows snapshot pictures of the lubricant layer during retraction, as the block moves away from the substrate for three different block positions $d = 0, 3$ and 6 Å. Only the central part of the contact between the block and the substrate is shown, top view, after removing the block and substrate atoms. In the beginning ($d = 0$ Å) octane molecules are located in the substrate corrugation wells, or cavities with direct metal–metal contact between the block and the top of the substrate nano asperities (see Fig. 13.35). During retraction ($d = 3$ Å) the octane molecules are pulled out of the wells forming an

almost symmetric network of nano-bridges around the asperity tops, increasing the adhesion between the two surfaces. This configuration corresponds to the maximal adhesion force, see curve (d) in Fig. 13.33. Thus maximal adhesion is achieved via the formation of many small capillary nano-bridges, involving just a few molecules for each bridge (see Fig. 13.35). Further retraction ($d = 6$ Å) results in the collapse of the nano-bridges and the formation of a single "large" capillary bridge in the center of the contact region.

13.7 Summary and Outlook

We have discussed the origin of friction and adhesion between hard solids such as quasicrystals. However, most of the results presented above are valid for all types of solids. We have emphasized the fundamental role of surface roughness on many contact mechanics problems, in particular for friction and adhesion between solids bodies.

Surface roughness has a huge influence on many common phenomena. It is the main reason for why macroscopic bodies usually do not adhere to each other with any measurable strength. For example, if the floor and the sole of the shoes would be atomically smooth and clean, it would be impossible to walk on the floor! The (nearly) absence of adhesion in most situations is crucial for the function of many man-made constructions.

The surface to volume ratio of solid objects increases as the lateral size of the object decreases. The role of surface roughness becomes therefore more important as the size of objects decreases. The present drive toward the miniaturization of mechanical devices, e.g., MEMS, requires a better understanding of the role of surface roughness on, e.g., contact mechanics and adhesion.

Surface roughness is also of great importance for the function of many biological systems. Thus, flies, bugs, crickets and lizards have developed very soft layers on their attachment organs which allow them to attach and move on both very smooth and rough vertical solid walls, e.g. stone walls or leafs. Another example is non-wetting coatings on plant surfaces based on surface roughness on many different length scales (the so called Lotus effect) [11].

The roughness of surfaces can today be studied strait forwardly using standard equipments based on optical methods and stylus methods, e.g., the atomic force microscope (AFM). These methods cover the whole length scale from atomic dimension to macroscopic distances. Thus, the AFM can probe the surface profile from ~ 1 nm to 100 μm and optical methods from ~ 1 μm to kilometers. For randomly rough surfaces, the most important quantity which can be deduced from the measured height profile is the surface roughness power spectra. We have shown in this paper how the power spectra determines the contact mechanics and adhesion for solid objects in direct contact. It also govern rubber friction on rough substrates, e.g., tires on a road surfaces, and influence other phenomena of technological importance, e.g., the roughness induced leaking of sealings.

Thus, studies of surface roughness is important not only for understanding many natural and biological phenomena, but also for many technological processes. The present drive toward miniaturization and the design of optimal systems by transfer of ideas from studies on biological systems to materials science (bionics) [11], is likely to accelerate the interest and efforts to study and predict the influence of surface roughness on many phenomena.

References

1. D. Shechtman, I. Blech, D. Gratias, J.W. Cahn, Phys. Rev. Lett. **53**, 1951 (1984)
2. B.N.J. Persson, O. Albohr, F. Mancosu, V. Peveri, V.N. Samoilov, I.M. Sivebaek, Wear **254**, 835 (2003)
3. E. Riedo, H. Brune, Applied Physics Letters **83**, 1986 (2003)
4. J.M. Dubois, S.S. Kang, J. von Stebut, J. Mater. Sci. Lett. **10**, 537 (1991)
5. J.M. Dubois, P. Brunet, W. Costin, A. Merstallinger, J. Non-Cryst. Solids **334–335**, 475 (2004)
6. J.Y. Park, D.F. Ogletree, M. Salmeron, C.J. Jenks, P.A. Thiel, Trib. Lett. **17**, 629 (2004)
7. J.S. Ko, A.J. Gellman, T.A. Lograsso, C.J. Jenks, P.A. Thiel, Surface Science **423**, 243 (1999)
8. C.F. McFadden, A.J. Gellman, Trib. Lett. **1**, 201 (1995)
9. Y.-P. Zhao, L.S. Wang, T.X. Yu, J. Adhesion Sci. Technol. **17**, 519 (2003)
10. C.H. Mastrangelo, Trib. Lett. **3**, 223 (1997)
11. F. Bottiglione, G. Carbone, Langmuir **29**, 599 (2013)
12. C. Gui, M. Elwenspoek, N. Tas, J.G.E. Gardeniers, J. Appl. Phys. **85**, 7448 (1999)
13. See, e.g., A.-L. Barabasi, H.E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, 1995)
14. See, e.g., J. Krug, Advances in Physics **46**, 139 (1997)
15. J. Feder, *Fractals* (Plenum Press, New York, 1988)
16. M.V. Berry, Z.V. Lewis, Proc. R. Soc. London A **370**, 459 (1980)
17. J. Krim, G. Palasantzas, Int. J. of Modern Phys. B **9**, 599 (1995)
18. P.R. Nayak, ASME J. Lubrication Technology **93**, 398 (1971)
19. E. Bouchaud, J. Phys.: Condens. Matter **9**, 4319 (1997)
20. B.N.J. Persson, *Sliding Friction: Physical Principles and Applications*, Sec. edn. (Springer, Heidelberg, 2000)
21. H. Hertz, J. Reine Angew. Math. **92**, 156 (1882)
22. J.F. Archard, Proc. R. Soc. A **243**, 190 (1957)
23. J.A. Greenwood, in Fundamentals of Friction, Macroscopic and Microscopic Processes, Ed. by I.L. Singer and H.M. Pollack (Kluver, Dordrecht, 1992)
24. J.A. Greenwood, J.B.P. Williamson, Proc. R. Soc. A **295**, 300 (1966)
25. K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985)
26. A.W. Bush, R.D. Gibson, T.R. Thomas, Wear **35**, 87 (1975)
27. A.W. Bush, R.D. Gibson, G.P. Keogh, Mech. Res. Commun. **3**, 169 (1976)
28. B.N.J. Persson, Phys. Rev. Lett. **87**, 1161 (2001)
29. B.N.J. Persson, J. Chem. Phys. **115**, 3840 (2001)
30. We define the apparent contact area $A(\lambda)$ on the length scale λ to be the area of real contact if the surface would be smooth on all length scales shorter than λ. That is, considering the Fourier transform of the surface profile, all the components whose wave vector is larger that $2\pi/\lambda$ have to be set to 0, and the contact area with this new surface is by definition $A(\lambda)$
31. B.N.J. Persson, F. Bucher, B. Chiaia, Phys. Rev. B **65**, 184106 (2002)
32. B.N.J. Persson, Eur. Phys. J. E **8**, 385 (2002)
33. B.N.J. Persson, Phys. Rev. Lett. **89**, 245502 (2002)
34. S. Hyun, L. Pei, J.-F. Molinari, M.O. Robbins, Phys. Rev. E **70**, 026117 (2004)

35. C. Yang, U. Tartaglino, B.N.J. Persson, Eur. Phys. J. E **19**, 47 (2006)
36. M. Borri-Brunetto, B. Chiaia, M. Ciavarella, Comput. Methods Appl. Mech. Engrg. **190**, 6053 (2001)
37. K.N.G. Fuller, D. Tabor, Proc. R. Soc. London A **345**, 327 (1975)
38. K. Kendall, Molecular Adhesion and it Applications (Kluwer, New York, 2001); D. Maugis, Contact, Adhesion and Rupture of Elastic Solids (Springer, Berlin, 1999)
39. K. Kendall, J. Phys. D: Appl. Phys. **4**, 1186 (1971); **6**, 1782 (1973); **8**, 115 (1975). See also the beautiful review article of K. Kendall, Contemp. Phys. **21**, 277 (1980)
40. G.A.D. Briggs, B.J. Briscoe, J. Phys. D: Appl. Phys. **10**, 2453 (1977)
41. K.N.G. Fuller, A.D. Roberts, J. Phys. D: Appl. Phys. **14**, 221 (1981)
42. S. Zilberman, B.N.J. Persson, J. Chem. Phys. **118**, 6473 (2003)
43. V.N. Samoilov, I.M. Sivebaek, B.N.J. Persson, J. Chem. Phys. **121**, 9639 (2004)
44. J.N. Israelachvili, *Intermolecular and Surface Forces* (Academic Press, London, 1995)
45. K.L. Johnson, K. Kendall, A.D. Roberts, Proc. R. Soc. Lond. A **324**, 301 (1971)
46. G. Sperling, PhD thesis, Karlsruhe Technical University (1964)
47. N.P. Padture, M. Gell, E.H. Jordan, Science **296**, 280 (2002)
48. K.T. Turner, S.M. Spearing, J. Appl. Phys. **92**, 7658 (2002)
49. L.D. Piveteau, B. Gasser, L. Schlapbach, Biomaterials **21**, 2193 (2000)
50. S. Gorb, B.N.J. Persson, J. Chem. Phys. **119**, 11437 (2003)
51. B.N.J. Persson, J. Chem. Phys. **118**, 7614 (2003)
52. A.K. Geim, S.V. Dubonos, I.V. Grigorieva, K.S. Novoselov, A.A. Zhukov, S.Yu. Shapoval, Nature Materials **2**, 461 (2003)
53. A. Peressadko, S.N. Gorb, Journal of Adhesion **80**, 247 (2004)
54. K.L. Johnson, Tribology International **31**, 413 (1998); L. Zhang, Y.-P. Zhao, J. Adhesion Sci. Technol. **18**, 715 (2004)
55. B.N.J. Persson, Wear **254**, 832 (2003)
56. H.J. Gao, H.M. Yao, Proceedings of the National Academy of Sciences of the United States of America **101**, 7851 (2004)
57. B.N.J. Persson, F. Mugele, J. Phys.: Condens. Matter **16**, R295 (2004)

Chapter 14
MD/FE Multiscale Modeling of Contact

Srinivasa Babu Ramisetti, Guillaume Anciaux and Jean-Francois Molinari

Abstract Limitations of single scale approaches to study the complex physics involved in friction have motivated the development of multiscale models. We review the state-of-the-art multiscale models that have been developed up to date. These have been successfully applied to a variety of physical problems, but that were limited, in most cases, to zero Kelvin studies. We illustrate some of the technical challenges involved with simulating a frictional sliding problem, which by nature generates a large amount of heat. These challenges can be overcome by a proper usage of spatial filters, which we combine to a direct finite-temperature multiscale approach coupling molecular dynamics with finite elements. The basic building block relies on the proper definition of a scale transfer operator using the least square minimization and spatial filtering. Then, the restitution force from the generalized Langevin equation is modified to perform a two-way thermal coupling between the two models. Numerical examples are shown to illustrate the proposed coupling formulation.

14.1 Introduction

Traditional friction experiments are particularly difficult to comprehend since they involve a wide variety of physical mechanisms that interact at several length and time scales. Amongst those mechanisms, one can list for instance long range elastic deformations, plasticity, third body interactions, lattice dynamics and heat transfer [1]. An additional difficulty comes from their interactions with surface topology. Experimental, theoretical and numerical studies have shown that surface roughness is a key determining factor for friction. Roughness being present at all length scales

S.B. Ramisetti
University of Edinburgh, Sanderson Building, Edinburgh, UK
e-mail: s.ramisetti@ed.ac.uk

G. Anciaux
Ecole Polytechnique Fédérale de Lausanne, EPFL ENAC IIC LSMS, GC A2 484
(Bâtiment GC), Station 18, 1015 Lausanne, Switzerland
e-mail: guillaume.anciaux@epfl.ch

J.F. Molinari (✉)
e-mail: jean-francois.molinari@epfl.ch

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_14

[2], developing a fundamental understanding of how microscopic contact clusters develop under load remains an important question [3, 4].

In the last few decades, with the development of nanotechnologies and nano science, interfaces and surfaces have started to dominate over the more classical and thus better understood bulk mechanisms. This challenges our traditional design tools and in particular continuum mechanics predictions which can be shown to break down at atomistic size asperities [5]. On the other hand, novel experimental techniques such as surface force apparatus, atomic force microscopy, friction force microscopy and quartz-crystal micro-balance are now extensively used to understand the atomic origins of friction [6, 7]. These techniques provide new insights and give renewed hope that we will one day have fully predictive tools for friction.

An essential component of those tools will be numerical modeling. Simulations are not only a useful complement to experiments as they can answer several experimental unknowns, but they can be used to explore a range of conditions out of reach of experiments. While numerical contact mechanics models have traditionally relied on a macroscopic description with empirical or semi-empirical phenomenological laws (Coulomb friction law, Archard wear law), recent modeling efforts increasingly attempt to represent the accurate atomistic mechanisms and capture the statistics of contact forces at small-scale contact asperities. However, a true separation of scales does not exist in most applications, and thus it is important to couple the small scale atomic mechanisms with long-range elastic forces and a proper handling of far field boundary conditions. This can be achieved with the rapid and recent developments in multiscale methods paralleled by a continuing expansion of computational power.

This chapter will introduce the reader to current methods in multiscale modeling of contact. An emphasis will be put on sliding contact and thus we narrow the focus to methods that couple an atomistic domain (*Molecular Dynamics*, MD, to capture atomic mechanisms at contacting asperities) to a continuum domain (*Finite Elements*, FE, for an accurate representation of long range elastic forces). Incidentally, MD/FE multi-scaling is also by far the most researched and versatile approach. The chapter begins by a review of some of the main variants of FE/MD class of direct (i.e. fully coupled) multiscale model. Furthermore, the important generation of heat during sliding friction, due to plastic activity at contacting asperities, will be demonstrated to challenge most current multiscale approaches. This will show the need for novel coupling strategies capable of handling heat fluxes through interfaces between distinct scale models. The last section of this chapter will turn to the description and validation of a novel thermo-mechanical coupling method, that shows great prospect for contact simulations.

14.2 Modeling Techniques of Contact at Nanoscale

A large amount of numerical studies of contact problems are based on single scale approaches. Numerical techniques such as Ab-Initio [8], *Discrete Element Method* [9, 10], *Discrete Dislocation Dynamics* [11], *Finite Element Method* [12, 13] and

Molecular Dynamics [14–16] have been used to study contact/friction problems. Two of the most classical techniques are the *Finite Element Method* [17, 18] and the *Molecular Dynamics* [19]. A large literature has had recourse to the *Finite Element Method*, which is a computationally efficient strategy, to model contact at the asperity level [12, 20–22].

Nevertheless, recent *Finite Element Method* simulations [23] as well as atomistic studies [5] show that contact mechanics is dominated by nanoscale asperities. Continuum mechanics is unable to capture the details of force profiles at this scale. In order to represent efficiently the atomic organization and forces at contact clusters, one can resort to *Molecular Dynamics* (MD).

Classical MD is a well-established numerical approach that is used to simulate materials at nanoscales. According to the Born-Oppenheimer approximation [24, 25], atomic nuclei are treated as point particles, because they are much heavier than surrounding electrons. Therefore, classical MD consists in driving N particles with the following Newtonian equation of motion:

$$m_i \frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{f}_i = \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{F}_{ij} \tag{14.1}$$

where m_i is the mass of the ith atom, $\mathbf{r}_i$ is its position and $\mathbf{F}_{ij}$ is the force acting on atom i exerted by atom j. The forces perceived by particles are described by an inter-particle potential, since the force $\mathbf{f}_i$ acting on atom i is equal to the gradient of the total potential energy of the system with respect to the ith atom position:

$$\mathbf{f}_i = -\nabla_i \Phi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{14.2}$$

MD simulations are used to investigate nanoscale mechanisms at the origin of adhesive and friction forces [7, 14–16, 26–34]. Besides the refined mechanical description achieved by MD models, severe limitations should be noted. First, the stable timestep is usually of the order of a femtosecond which restricts long (>100 ns) simulation runs. Secondly, the number of atoms to materialize a small chunk of matter is restricted by the computational time [35].

The limitations of purely atomistic or purely continuum simulations, which make extremely difficult the link of simulations and experiments, have motivated research in multi-scale simulations that bridge atomistic and continuum modeling [36–44]. In these multiscale approaches, atoms can be used at contacting asperities to capture in great details contact forces, whereas *Finite Elements* (FE) are used away from the interface to accurately represent elastic forces. The main purpose of the coarse domain is to reduce significantly the number of unknowns to handle.

14.3 Multiscale Coupling Applied to Contact

Multiscale modeling has captured tremendous attention from different fields such as materials science, mechanics and high performance computing, which is due to its potential to perform numerical simulations that were impossible or difficult with the full atomistic simulations. Thus, during the past decades several multiscale methods have been developed to investigate material problems.

A broad classification of multiscale approaches is done in [45], which separates the field in two categories. The first contains hierarchical methods which model the different scales separately but with information flow between fine and coarse scales. While this approach avoids the technical difficulty of direct coupling between scales, which explains the reason for its wide usage, the necessary scale separation assumption can be a too strong approximation for real applications. The second category considers direct/concurrent multiscale approaches which simulate simultaneously the different length scale models. Atomistic models are used in critical regions to capture atomistic processes, while coarser models are used in regions away from complex behavior. The coherency between the atomistic and the coarser models is enforced in an interface or overlap region as illustrated on Fig. 14.1.

Concurrent approaches are relevant for sliding friction simulations where complex deformations occur at the contacting interface, while coarser scales handle long range elastic interactions and provide boundary conditions. There exist different concurrent coupling methods such as finite element atomistics method (FEAt) [50], quasicontinuum method (QC) [46, 47], coupling of length scales (CLS) [36], coupled

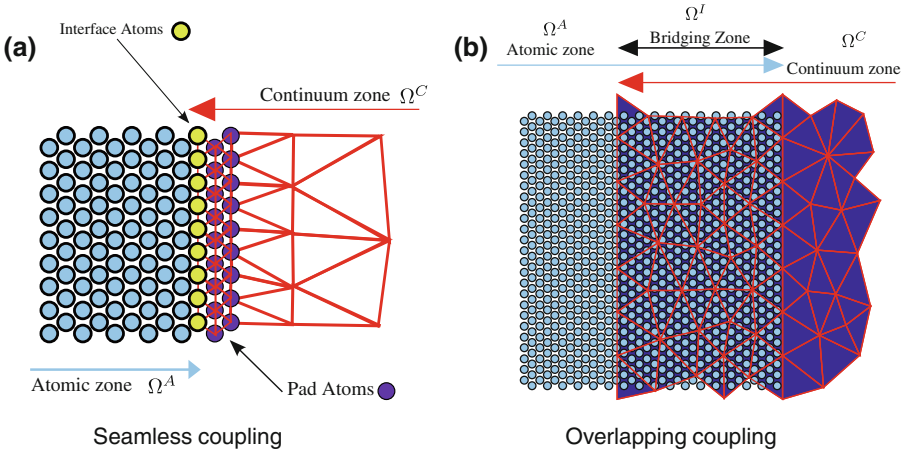


Fig. 14.1 Illustration of coupling interfaces between molecular dynamics and finite elements. **a** Typical interface zone in a *seamless coupling* such as the *Quasi-Continuum* [46, 47], or the *Coupled Atomistic and Discrete Dislocation* [48] methods. **b** Typical bridging/overlapping zone employed in methods such as the *Bridging Domain* [49]. It should be noted that in both cases, the introduction of coarser elements leads to distinct dispersion relations and spectral decompositions

atomistic and discrete dislocation (CADD) [48], bridging scale method (BSM) [51] and bridging domain method (BDM) [49]. This list is not complete and should not be taken as the only reference. There also exists a good number of review articles, that one can refer to, which address the different multiscale methods [42, 44] and their comparison with each other [43]. Nevertheless, in the following section we will review four methods, which are now classical in the literature.

14.3.1 State of the Art of Multiscale Methods

The **Quasicontinuum** (QC) method, developed by Tadmor et al. [46], was first used to investigate two-dimensional quasi-static single crystal deformation problems. In the atomic/refined region, the energy is computed using interatomic potentials. At the interface between the atoms and the FE's, the energy of the interface atoms is calculated by introducing neighboring atoms (known as pad atoms) which are deformed accordingly to interpolated FE displacement fields. In the FE region, the strain energy density W is computed from the atomistic potential using the Cauchy-Born rule [46]. The total energy of the coupled system is written as:

$$E = \sum_{i \in \Omega^A} E_i + \sum_{e \in \Omega^C} w_e E_e \tag{14.3}$$

where E_i is the energy of atom i, E_e is the energy of element e, and w_e is a weighting function to correct an energy unbalance. Indeed, the last free atoms at the interface and the first FE (pad atoms/nodes) bear an overlapped contribution to the total energy which is corrected with the weight w_e.

During the recent years, various improved versions of the QC method, including the treatment of multigrains and three-dimensional deformation problems, have been developed [37, 47, 52–54]. Several finite temperature extensions of the QC method exists [55–59]. For instance, the hot-QC method [56, 60] uses a temperature dependent Hamiltonian, based on the idea of the potential of mean force, to approximate the contributions of missing atoms in the continuum region.

The **Coupled Atomistic and Discrete Dislocation** (CADD) method, developed by Shilkrot et al. [48, 61], allows the direct coupling of an atomistic region with a continuum region containing dislocations. The key feature of this method is that it can pass dislocations from the atomistic region to a continuum region without confining the plastic deformation to the atomistic region in contrast to the QC method. The dislocations passed into the continuum region are represented using the discrete dislocation method [62].

As described in [48, 61], the solution to the boundary value problem is obtained by dividing it into three problems: an infinite elastic continuum with dislocations, a linear elastic continuum without any defects and a full atomistic region. The solution to the first problem is obtained by superposing the analytical elastic fields due to the network of dislocations. The stress, strain and displacement contributions from the

discrete dislocations are denoted as $\tilde{\boldsymbol{\sigma}}$, $\tilde{\boldsymbol{\varepsilon}}$, $\tilde{\mathbf{u}}$ respectively. The solution to the second problem is found by using corrective tractions $\hat{\boldsymbol{\sigma}}$ and displacements $\hat{\mathbf{u}}$. The corrective strain field is denoted by $\hat{\boldsymbol{\varepsilon}}$. And the third problem consisting of the atomistic region is solved by using interatomic potentials. The atoms near the continuum-atomistic interface are treated in the same way as in the QC method.

The total energy of the boundary value problem is expressed as:

$$E = \frac{1}{2}\int_{\Omega^C} (\tilde{\boldsymbol{\sigma}} + \hat{\boldsymbol{\sigma}}) : (\tilde{\boldsymbol{\varepsilon}} + \hat{\boldsymbol{\varepsilon}}) dV - \int_{\partial\Omega^C} \mathbf{t_0} \cdot \mathbf{u} dA + E_A - \mathbf{f}_A \cdot \mathbf{u}_A \tag{14.4}$$

where $\mathbf{t}_0$ is the prescribed traction, $\mathbf{u} = \tilde{\mathbf{u}} + \hat{\mathbf{u}}$ is the total displacement, E_A is the atomistic energy, $\mathbf{f}_A$ refers to the atomic forces along the traction boundary and $\mathbf{u}_A$ refers to the atomic displacements.

The detection of the dislocations nucleated in the atomistic region and their passing to the continuum region is accomplished in two steps: (i) In 2D, a detection band of triangular elements inside region Ω^A and close to the interface is defined to monitor the Lagrangian finite strain and to allow the identification of dislocations based on their recognizable slip strains within the crystal. (ii) After the detection step, the dislocation is passed to the continuum region by adding the displacement fields associated with a dislocation dipole. This shifts the dislocation core along its slip plane from its location in the detection band to a location across the interface in the continuum region.

Currently, the approach is only validated in two-dimensional case. Extension of CADD to finite temperature simulations can be found in [63, 64], where the stadium damping method with a Langevin based thermostat is used to maintain a constant temperature of the system. However, it is not yet suitable to treat non-isothermal processes. Extension to three-dimensional systems is the subject of active research.

The **Bridging Domain method** (BDM) uses the concepts of the Arlequin approach [65–68] which can intermix energies of several continuum mechanical models and constrain consistent displacements within an overlaping zone (also termed as the bridging domain). Xiao et al. applied this strategy for coupling continuum models with molecular dynamics (MD) [49, 69].

The total Hamiltonian of the system is considered to be equal to the sum of the weighted Hamiltonians of both the atomistic and continuum models:

$$H = (1 - \alpha) E^A + \alpha E^C \tag{14.5}$$

where E^A and E^C are the atomic and continuum Hamiltonian contributions and where α is an arbitrary weighting function. The displacement continuity is enforced between the two models in the overlap region (Ω^I) by constraining the degrees of freedom using the Lagrangian multiplier method. The constraints on the velocities are expressed as:

$$\mathbf{g} = \mathbf{N}\dot{\mathbf{u}} - \dot{\mathbf{d}} = 0 \tag{14.6}$$

where $\mathbf{g}$ is the vector containing per atom constraints, $\mathbf{u}$ is the FE nodal displacement, $\mathbf{d}$ is the atomic displacement and $\mathbf{N}$ is the matrix containing FE shape functions evaluated at all initial atomic positions. The governing equations for degrees of freedom inside the overlap region are formulated using the Lagrangian multiplier method. The Lagrange multipliers λ (L multipliers, with L the number of coupled atoms) are obtained by solving the linear system of equations

$$\mathbf{H}\lambda = \mathbf{g}^{\star} \tag{14.7}$$

where $\mathbf{g}^{\star}$ is the constraint vector before correction and $\mathbf{H}$ is the $L \times L$ constraint matrix defined as

$$\mathbf{H} = \mathbf{N}^T \hat{\mathbf{M}}^{-1} \mathbf{N} - \hat{\mathbf{m}}^{-1} \tag{14.8}$$

where $\hat{\mathbf{M}} = \alpha \mathbf{M}$ with $\mathbf{M}$ is a coarse scale lumped mass matrix and where $\hat{\mathbf{m}} = \alpha \mathbf{m}$ with $\mathbf{m}$ is a diagonal atomic mass matrix. Finally, the discrete governing equations of the two models are expressed as follows:

$$\begin{cases} \hat{\mathbf{M}}\ddot{\mathbf{u}} = \alpha \mathbf{f}(\mathbf{u}) - \lambda \mathbf{N}^T \\ \hat{\mathbf{m}}\ddot{\mathbf{d}} = (1-\alpha)\mathbf{f}(\mathbf{d}) - \lambda \end{cases} \tag{14.9}$$

where $\mathbf{f}(\mathbf{d})$ and $\mathbf{f}(\mathbf{u})$ are the atomic and nodal forces. Details concerning the derivation of the above equations are presented in [49, 69, 70]. The arbitrary weighting is remarkably suited to dissipate spurious wave reflections [49, 70] at small temperatures ($\sim 0K$) and material problems such as fracture were successfully simulated using this method [49]. However, the application of this method to simulate finite temperature problems is difficult [71].

The **Bridging Scale method** (BSM) was developed by Wagner and Liu to concurrently couple atomistic and continuum models [51]. The idea of this method is to decompose the total displacement field $\mathbf{u}(\mathbf{x})$ into coarse and fine scales as:

$$\mathbf{u}(\mathbf{x}) = \bar{\mathbf{u}}(\mathbf{x}) + \mathbf{u}'(\mathbf{x}) \tag{14.10}$$

The coarse scale displacement field in matrix form is defined as:

$$\bar{\mathbf{u}} = \mathbf{N}\mathbf{d} \tag{14.11}$$

where $\mathbf{N}$ is the matrix containing FE shape functions evaluated at all initial atomic positions and $\mathbf{d}$ is the FE nodal displacements. The fine scale displacement field is defined as the projection of MD displacements $\mathbf{q}$ on the FE basis functions subtracted from the total solution $\mathbf{q}$ and is expressed as:

$$\mathbf{u}' = \mathbf{q} - \mathbf{P}\mathbf{q} \tag{14.12}$$

where $\mathbf{P}$ is a projection matrix defined as:

$$\mathbf{P} = \mathbf{N}\mathbf{M}^{-1}\mathbf{N}^T\mathbf{M}_A \tag{14.13}$$

Here $\mathbf{M}_A$ is a diagonal atomic mass matrix and $\mathbf{M} = \mathbf{N}^T\mathbf{M}_A\mathbf{N}$ is a coarse scale mass matrix.

The final equations of motion for the MD and FE models are derived using the Lagrangian form. More details about the derivation of these equations can be found in [72]. The key equations are:

$$\mathbf{M}_A\ddot{\mathbf{q}} = \mathbf{f}(\mathbf{q}) + \mathbf{f}^{imp} + \mathbf{R}^f \tag{14.14a}$$

$$\mathbf{M}\ddot{\mathbf{d}} = \mathbf{N}^T\mathbf{f}(\mathbf{u}) \tag{14.14b}$$

where $\mathbf{f}^{imp}$ is an impedance force and $\mathbf{R}^f$ is a random force. The impedance force is defined as:

$$\mathbf{f}_i^{imp} = \sum_{j \in neighbors(i)} \int_0^t \Theta_j(t-\tau) \times \left(\mathbf{q}_j(\tau) - \bar{\mathbf{u}}_j(\tau) - \mathbf{R}_j^d(\tau)\right) d\tau \tag{14.15}$$

where $\Theta_j(t-\tau)$ and $\mathbf{R}_j^d(t)$ are a time history kernel and a random displacement term respectively. The purpose of the random force $\mathbf{R}^f$ is to restitute the energy dissipated by the impedance force and thus ensuring energy conservation.

The important point to note in this method is that the impedance force has the role of dissipating the short wavelengths that cannot be represented by the FE mesh. This energy dissipation is based on the generalized Langevin equation (GLE) [73–75]. One limitation is that the time history kernel is usually derived for a given lattice structure which restricts their usage to crystalline materials.

Several other concurrent atomistic-continuum coupled approaches have been developed using the idea of GLE to dissipate short wavelengths that are reflected at the MD-FE interface. However, these approaches differ in the way the time history kernel function Θ is derived. For instance, Cai et al. [76] computed Θ from several MD simulations. E and Huang [77, 78] have computed analytically the kernel coefficients by minimizing the reflection coefficients at each wavenumber. Wagner et al. [79] have computed Θ using the Laplace and the Fourier transforms. Most of these approaches assume the temperature of the coarse scale to be zero to ignore the random force term $\mathbf{R}^f$ and thus are not suitable for thermal transfer applications. A few methods based on the idea of GLE also exists that are suitable for studying problems with non-equilibrium processes. For instance, Karpov et al. [80] have developed a concurrent atomistic continuum model by using analytical expressions for Θ and including a random force term to allow the passage of thermal energy between the atomistic and continuum regions. Mathew et al. [81] have used a time

dependent friction force and a weighted random force to treat thermal fluxes across the atomistic-continuum interface. The common feature in all these methods is the time history kernel function, which is built using different techniques. Recently, a parametric study focused on the influence of time and spatial kernels on the dynamics of one-dimensional MD systems was conducted [82] and revealed that spatial filters present interesting features, when compared to time filters, which can be exploited as will be demonstrated in a later section.

14.3.2 Sliding Friction and Heat Generation

Sliding friction between rough surfaces generates intense heat fluxes because of the large plastic deformations. This can put to the test any direct multiscale method [83]. For instance, when rough surfaces carved from two cubic-like copper crystals at zero Kelvin with self-affine fractal [84] generated with a Voss [85, 86] algorithm (as presented in Fig. 14.2) are pressed against each other and sled, a temperature rise is to be expected. In order to demonstrate the artificial impact of the Bridging Domain algorithm on sliding contact dynamics, three different models are compared:

- Full MD model (the continuum zone is replaced by atoms and serves as a reference)
- Coupled model (as described in Fig. 14.2)
- Reduced MD model (the continuum zone is eliminated).

To quantify the effect of the coupling with regards to phonon emission, the kinetic energy of the top zone of the deformable body is measured. This zone contains the energy close to the asperities, without any contribution of the overlap region atoms.

These measures are presented in Fig. 14.3. The coupled approach always leads to a minimal residual kinetic energy, while the reduced case stores a lot more vibrational energy in the contacting zone. It is noteworthy that the coupling scheme fails in recovering the full MD behavior: the kinetic energy profile remains almost flat and at a low value.

Interestingly, most work in the literature has sought to prevent wave reflections without necessarily considering that the damping of the problematic (high frequency) waves could impact the uncoupled zones. Indeed, the Bridging Domain method, when handling properly the undesired high frequency waves incoming from the molecular domain, is damping a part of the kinetic energy in an ad hoc way [70, 87]. Here, with an initial state of zero Kelvin, and with asperities of various sizes and shapes, colliding and scratching at contact points, thermal vibrations are being generated at an important rate. The resulting heat increase is an integral part of the contact problem and for some problems should not be damped entirely by the coupling zone. Thus, the sliding friction problem calls for a thermo-mechanical multiscale model with the potential to address heat fluxes.

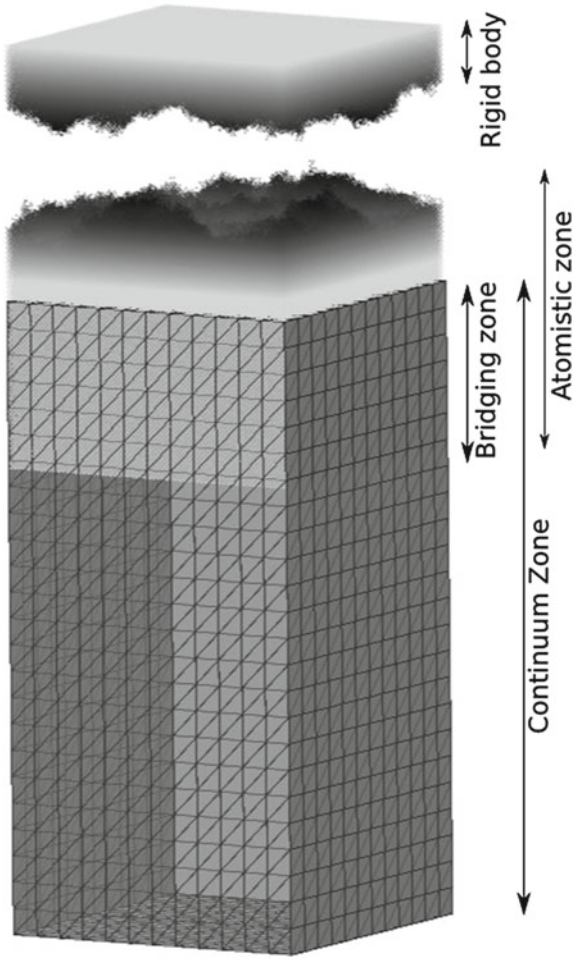


Fig. 14.2 3D View of the mesh and atomic zone coupled together

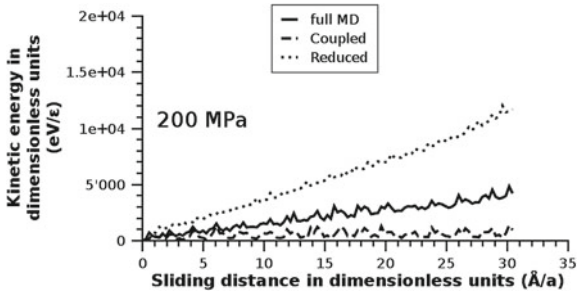


Fig. 14.3 Residual kinetic energy in the zone near contacting asperities (thickness 24 a, a = 3.615Å is the lattice constant) for the full MD (*solid line*), coupled (*dashed line*) and reduced cases (*dotted line*)

14.4 Finite Temperature Coupling

This section begins with the description of a novel multiscale model coupling MD and FE. A scale transfer operator using a spatial filter is described. Then the coupling formulation, which uses GLE to damp selective frequency modes in the coupling region, is presented. Later on, the thermal coupling formulation to treat thermal fluxes across the MD-FE interface is introduced.

In order to illustrate the geometry of the coupling zone needed for this strategy, a schematic is presented in Fig. 14.4. Ω^A and Ω^C are used to refer to the pure atomistic and the pure continuum regions, while $\Omega^{C(A)}$ and $\Omega^{A(C)}$ represent the coupling and boundary regions respectively [49, 88].

While the dynamics of an atom in region Ω^A follow the classical Newtonian equation, the mechanical and thermal fields within the continuum model are represented with two different partial differential equations: the evolution of the displacement field is described using the equations of motion combined with a linear elastic law, while Fourier's thermal conductivity equations are called upon to represent the temperature field. It should be noted that the heat propagation within the system is assumed to be only due to conduction without taking into account convection and radiation. The dynamics inside the coupling $\Omega^{A(C)}$ and the boundary $\Omega^{C(A)}$ regions, where both atoms and finite elements coexist, need three components such as a *scale transfer operator*, a *selective thermostat* and a *heat balance equation*, which are presented in the following sections.

14.4.1 Scale Transfer Operator

In the boundary region $\Omega^{A(C)}$, the atomic displacements and velocities are simply computed from the interpolated FE fields, whereas the coupling is more complex when information has to pass from the fine to the coarse scale. For this operation, a scale transfer operator is used to define the transmission of information, such as

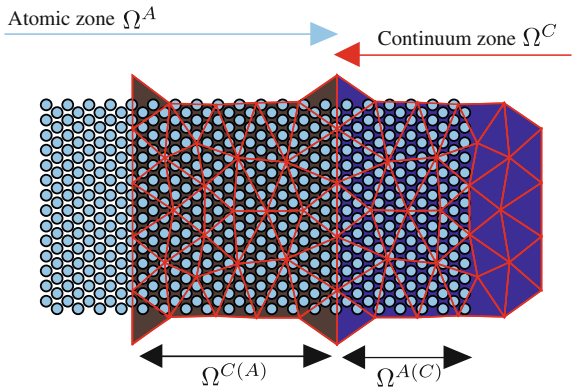


Fig. 14.4 Illustration of the MD-FE multiscale domain

displacement and velocity, from one length scale to another without corrupting the system dynamics in either of the scales.

For example, Fackeldey et al. [89, 90] have developed an atomistic-continuum coupled approach using a weighted least square projection as a scale transfer operator to decompose the atomic displacements into low and high frequency components. They provided numerical examples demonstrating the seamless transmission of displacements from MD to FE at zero Kelvin. However, it can be shown that the least square projection has poor filtering properties [91].

In order to improve this scale transfer operator, a least-square projection can be combined [91] with a spatial filter to define an improved scale transfer operator. The continuum displacement U_J of any node J in the coupling region $\Omega^{C(A)}$ is then formally obtained with:

$$U_J = \sum_I A^{-1}_{I,J} \sum_{i \in \Omega^{C(A)}} \hat{u}_i \, N_I(X_i) \tag{14.16}$$

$$\text{with } A_{I,J} = \sum_{i \in \Omega^{C(A)}} N_I(X_i) N_J(X_i)$$

where $A_{I,J}$ is the least square projection matrix [40], N_I is the shape function described by a linear polynomial for node I and $\hat{u}_i$ is the spatially filtered displacement of any atom i inside the $\Omega^{C(A)}$ region defined by:

$$\hat{u}_i = \sum_{j \in neighbors(i)} \gamma(|X_i - X_j|) u_j \tag{14.17}$$

where γ is a spatial filter (memory kernel) function, which can be chosen so that the *finite elements* receive only waves for which they have enough degrees of freedom to represent. The continuum velocity field is defined in a similar manner. The scale transfer operator hence defined allows to transmit precisely the band of frequencies that the mesh can handle with its coarser representation. The waves not mechanically transmitted have to be transformed into thermal energy which would allow the coupling of heat fluxes. This is the role of the selective thermostat presented below.

14.4.2 Selective Thermostat

The presented method stands on the generalized Langevin equation (GLE) [73–75]. The dynamics of atoms inside the coupling region is described using the GLE which incorporates spatial filters as expressed by the following equations:

$$m\dot{v} = -\nabla\Phi - \frac{m(v - \hat{v})}{\alpha} + R \tag{14.18}$$

where m is the mass, v is the velocity, $\dot{v}$ is the acceleration, $\hat{v}$ is the spatially filtered velocity and Φ is the potential function. The second term on the right side of equation (14.18) is the frictional force, α is a damping parameter to decide the strength of the frictional force, and R is a random or fluctuating force which is correlated in both space and time. The purpose of the random force is to balance both the energy dissipation of the friction force and the heat exchange with the continuum model. For instance the random force $R(x, t)$ can be derived analytically for a one-dimensional mono-atomic lattice:

$$R(x,t) = \frac{1}{\alpha}\sqrt{\frac{2mk_BT}{N}}\sum_k cos(\omega(k)t + k\,x + \phi(k)) \tag{14.19}$$

where k_B is the Boltzmann constant, T is the desired temperature, N is the number of restitution modes, $\phi(k)$ is a random phase sampled in the interval $[0, 2\pi]$ and $\omega(k)$ is the angular frequency associated with the wave vector k, taken from the dispersion relation. In the restitution, a temperature T has to be defined.

In the case of thermal equilibrium, the temperature T can be taken as a constant. However, in the non-equilibrium case, this temperature is given by the continuum model and ensures that the continuum can exchange heat with the atomic region. As an additional component, the energy balance presented in next section allows heat fluxes to be introduced in the continuum.

14.4.3 Heat Balance Equation

The governing equation used to describe the thermal transfer assuming Fourier's law ($\mathbf{q} = -\kappa\,\nabla T$) within the finite-element model is given by:

$$\rho\,C_v\,\dot{T} = \nabla\cdot(\kappa\,\nabla T) + Q \tag{14.20}$$

where ρ is the mass density, C_v is the specific heat capacity, $\dot{T}$ is the temperature rate, κ is the thermal conductivity, T is the temperature and Q is the volumetric heat source per unit time. Classically [17, 18], the finite element resolution of (14.20) leads to:

$$\mathbf{C}_{IJ}\dot{T}_J + \mathbf{K}_{IJ}T_J = Q_I \tag{14.21}$$

with $\mathbf{C}$ the capacity matrix, $\mathbf{K}$ the conductivity matrix and Q_I the heat rate associated with node I which is described as:

$$Q_I = \int_{\Omega^{C(A)}} N_I\,Q\,d\Omega^C \tag{14.22}$$

where N_I is the shape function associated with node I. The balance of the thermal energy inside the $\Omega^{C(A)}$ region is achieved from the difference between the heat rate of the atomistic and the continuum models as described by:

$$Q(x) = \sum_{i \in \Omega^{C(A)}} (q_i^F - q_i^R)\, \delta(x - x_i) \tag{14.23}$$

where $\delta(x - x_i)$ is the Dirac delta function equal to infinity at the position x_i of an atom i and zero elsewhere, q_i^F and q_i^R are the per atom heat rate due to the friction and random forces respectively, which are expressed as:

$$q_i^F = \frac{1}{\alpha}\, m\, (v - \hat{v})\, v_i \qquad q_i^R = R_i\, v_i \tag{14.24}$$

Thus, the heat rate Q_I associated with node I is expressed as:

$$Q_I = \int_{\Omega^{C(A)}} N_I\, Q d\Omega^C = \sum_{i \in C(A)} N_I(x_i)\, (q_i^F - q_i^R) \tag{14.25}$$

Because of the shape functions scope, only interface nodes will receive a flux coming from the MD model, which turns out to be a boundary condition for the FE region.

14.5 Validation and Application

In this section, three different numerical examples illustrate the method. In the first example, the method is validated by passing a mechanical wave pulse while maintaining the system at a constant finite temperature. The second example includes a mechanical wave propagation from the FE region into the MD region in addition to transient heat propagation. Finally, the case of a dynamic contact is shown.

In what follows, the material is a FCC aluminum crystal thin sheet which has a hexagonal lattice corresponding to the (111) plane of bulk aluminum. We resort to a simple harmonic potential with first neighbor interactions to prevent any plastic deformation. For the continuum model, the equations of motion described by an elastic orthotropic material law and the Fourier's heat conduction law are used to describe mechanical and thermal fields respectively. The parameters for both MD and FE models are found in Tables 14.1 and 14.2.

As previously stated, the dispersion relation between the angular frequency $\omega(k_x, k_y, b)$ and the wave vector $\mathbf{k}$ is called upon to construct the random force $R(\mathbf{x}, t)$. In the case of the considered two-dimensional hexagonal lattice it follows:

Table 14.1 Parameters of the MD model

Parameters	Value
m	26.98 g.mol^{-1}
ε	1.36 Kcal.mol^{-1}
σ	2.54 Å
r_0	$2.016\sqrt{2}$ Å
r_1	$\sqrt{3}r_0$ Å
r_{cut}	3.89 Å

Table 14.2 Parameters of the FE model

Parameters	Value
ρ	3.83 g.mol^{-1}.Å^{-2}
$E1$	9.78 Kcal.mol^{-1}.Å^{-3}
$E2$	9.78 Kcal.mol^{-1}.Å^{-3}
$\nu 12$	0.33
$G12$	3.67 Kcal.mol^{-1}.Å^{-3}
C_v	1.47e^{-4} Kcal.g^{-1}.K^{-1}
κ	1.23e^{-3} Kcal.mol^{-1}.Å^{-1}.fs^{-1}.K^{-1}

$$\omega^2(k_x, k_y, b) = \frac{C}{m}\Bigg[3 - \cos(k_x) - 2\zeta + (-1)^b \sqrt{(\cos(k_x) - \zeta)^2 + 3\sin^2\left(\frac{k_x}{2}\right)\sin^2\left(\frac{\sqrt{3}k_y}{2}\right)}\Bigg] \tag{14.26}$$

where $\zeta = \cos\left(\frac{k_x}{2}\right)\cos\left(\frac{\sqrt{3}k_y}{2}\right)$, b is the acoustic branch number, $\mathbf{k} = (k_x, k_y)$ is the wave vector and m is the mass of each atom.

14.5.1 Mechanical Wave Propagation at Finite Temperature

The coupled model is shown in Figure 14.5. The dimensions of the MD region (Ω^A) is $400r_0 \times 40r_1$, where r_0 is the inter-atomic spacing and $r_1 = \frac{\sqrt{3}}{2}r_0$. Two FE meshes each with 3520 linear triangular elements with a characteristic size $h = 5.0r_0$, as shown in Fig. 14.5, are used on either side of the MD region. Periodic boundary conditions are imposed along the y-direction for both models. Along the x-direction the size of the coupling region $\Omega^{C(A)}$ is $20r_0$. Each coupling region contains 4 and 16 finite elements along x and y directions respectively. A boundary region with 8 finite elements along x direction on both ends of Ω^A is used. For the initial condition, a

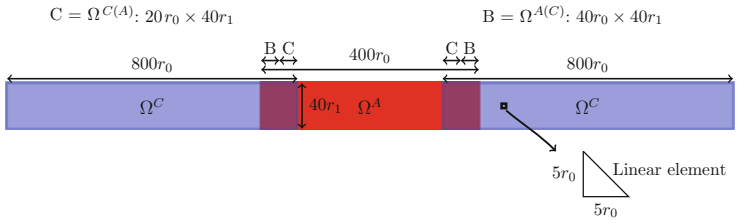


Fig. 14.5 Illustration of the coupled model. A uniform FE mesh with linear triangle elements is used on either side of the MD region

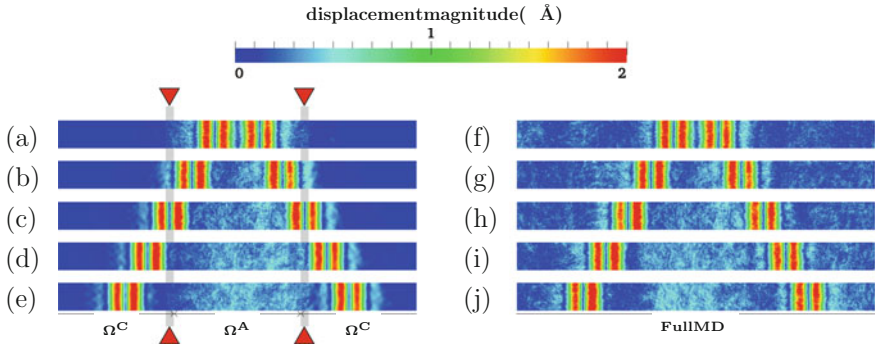


Fig. 14.6 Comparison of the displacement field in the coupled model (*Left*) with the full MD simulation (*Right*) at time $t = 4$, 8, 12, 16 and 20 ps. The temperature of the system is maintained at 50 K. The overlap region in the coupled model is indicated using the *triangle marks* in *red color*. For the sake of visualization, only 50 % of the complete mesh on either side of the MD region is shown

low frequency wave packet is introduced in the region Ω^A, while the high frequency modes in the MD region are only due to the injected temperature which is set to 50 K. The energy dissipated in the coupling region due to the damping is balanced by the random force.

The displacement profiles extracted from the coupled simulation with those obtained from the full MD results are compared in Fig. 14.6 and show a good agreement.

As expected the small wavelengths are damped and restituted in the coupling region with the help of a spatial filter. This ensures that the energy of the entire model, i.e. the sum of kinetic, potential and thermal energies, remains constant during the entire simulation [91].

14.5.2 Thermo-Mechanical Wave Propagation

In this example, a transient heat problem superposed with an impulse wave is considered to validate the coupling approach in the case of non-equilibrium processes. Figure 14.7, shows the MD-FE coupled model used in this example.

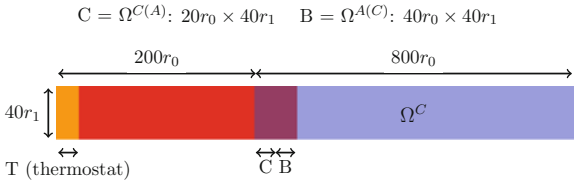


Fig. 14.7 Setup of the coupled model with a thermostat at 200 K imposed on a group of atoms on the left side of the MD region

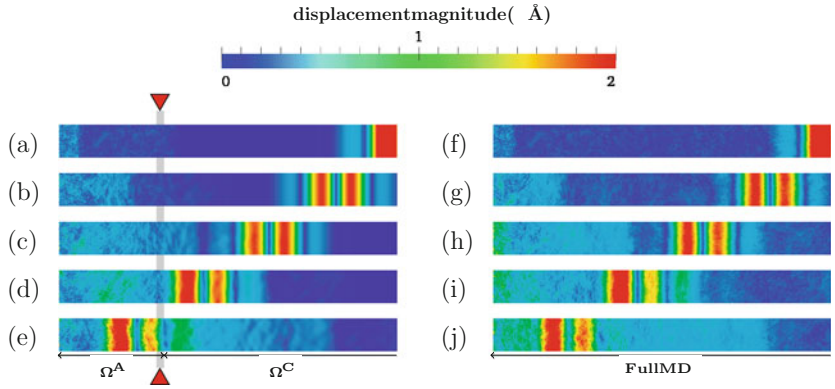


Fig. 14.8 Comparison of the displacement field in the coupled model (*Left*) with the full MD simulation (*Right*) at time $t = 0$, 10, 20, 30 and 40 ps. The overlap region in the coupled model is indicated using the *triangle marks* in *red* color

The size of the MD region Ω^A is $200r_0 \times 40r_1$, composed of 16,000 atoms. Similarly to the previous example, triangular finite elements ($h = 5r_0$) are used to represent the coarse scale model. Furthermore, the coupling and boundary regions share similar sizes with the previous example.

An initial temperature of 10 K is imposed everywhere in the model. Then, a Langevin thermostat of temperature 200 K is applied on a group of 3200 atoms on the left side of the MD region. The thermostat creates a thermal flux within the entire system which initiates transient heat propagation from the MD to FE region. At the same time, an impulse wave is imposed on the right side of the FE region. The dynamics of the entire system is allowed to evolve for a total time of 50 ps.

The snapshots of the displacement profile of the coupled MD–FE model and of the full MD simulations are shown in Fig. 14.8. A smooth transition of the large wavelength from the FE to the MD region can be observed. In addition to the displacements, the time averaged temperature profiles for both the coupled and the full MD models at time $t = 50$ ps are shown in Fig. 14.9a. Also, the total energy of the coupled model is compared with the full MD simulation and found to be in good agreement (see Fig. 14.9b).

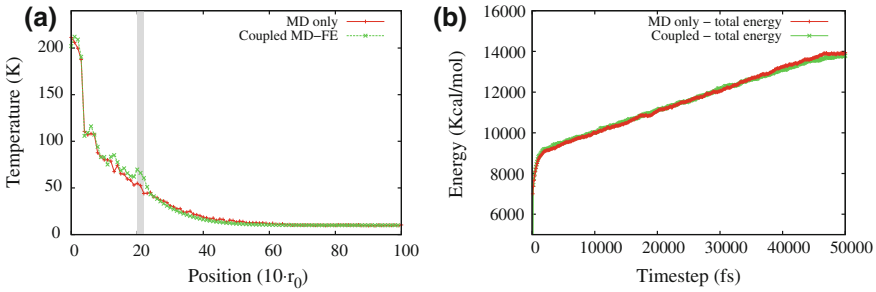


Fig. 14.9 **a** Comparison of the temperature profile of the coupled model with the full MD simulation at time $t = 50$ ps. The overlap region is indicated with a *light gray* background. **b** Comparison of the total energy of the coupled model with the full MD simulation. The total energy increases with time as it is a non-equilibrium process

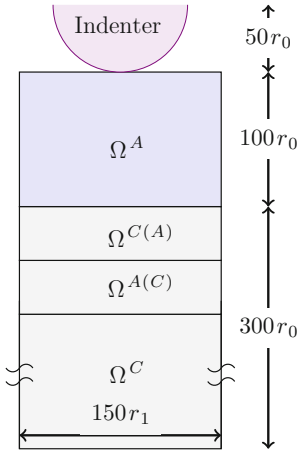


Fig. 14.10 Illustration of the multiscale model used for dynamic contact

14.5.3 Application to Dynamic Contact

Once again, the aluminum material from the previous two examples is considered to model the deformable substrate, which is subjected to an impact by a rigid circular indenter of radius $50r_0$. The schematic of the coupled MD–FE model is shown in Fig. 14.10. An initial temperature of $T = 20$ K is set within the MD–FE model. After reaching thermal equilibrium, the indenter impacts the substrate at a velocity of 5 Å.ps^{-1}.

The displacement profile of the MD–FE model is extracted at various timesteps and compared with the displacement profile of a full MD model as shown in Fig. 14.11. Again, a good agreement with the reference full MD is achieved. During the impact, the indenter tip creates waves propagating into the substrate. Waves with large wavelengths propagate through the FE mesh, while high frequency waves that cannot be represented by the FE mesh are transmitted as a thermal flux. Thus, the

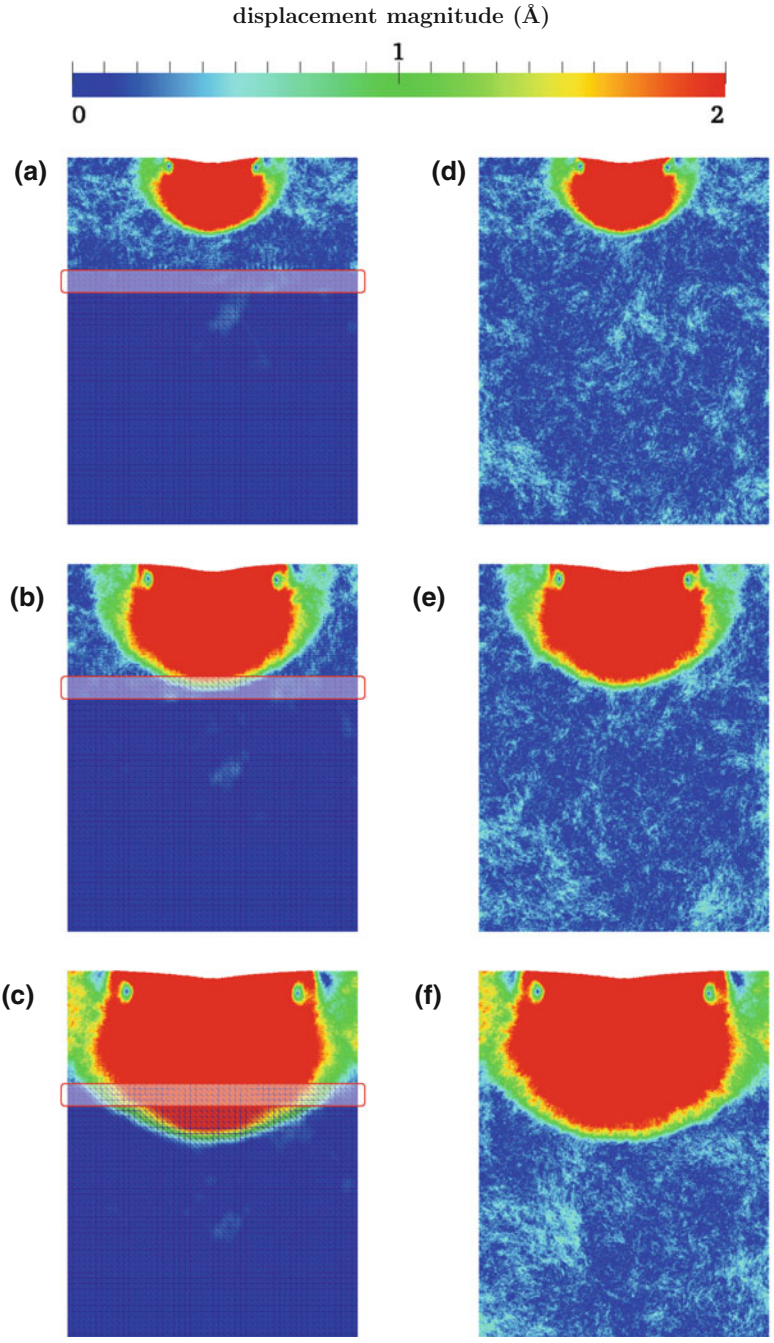


Fig. 14.11 Snapshots of the displacement field in the coupled model (*Left*) and the full MD model (*Right*) at time $t = 6$, 9 and 12 ps. The overlap region in the coupled model is shown using a transparent rectangle

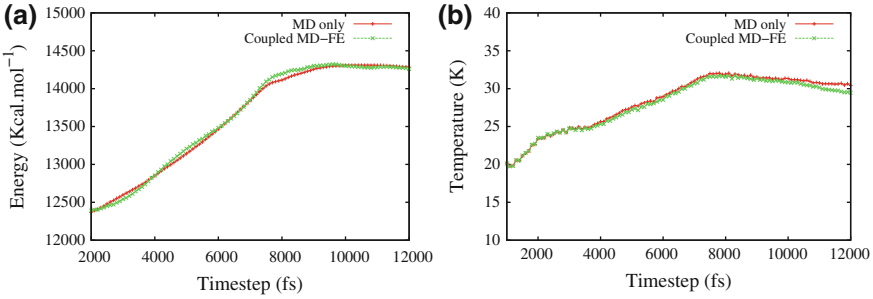


Fig. 14.12 **a** Total energy comparison between the coupled model and the full MD simulation and **b** Temperature in the MD region (Ω^A) in the coupled and the full MD models

total energy of the coupled model is found to be in good match with the full atomistic solution as shown in Fig. 14.12a. The atomistic temperature in region (Ω^A) is also measured for the coupled and the full MD models, which is found to have a reasonable agreement as shown in Fig. 14.12b. Thus, the proposed coupling approach can be applied to a dynamic contact problem and produce satisfactory results.

14.6 Conclusion

This chapter discussed modeling techniques of contact at the nanoscale with a special emphasis on molecular dynamics. The limitations of single scale approaches motivated the development of multiscale methods. A review on the state-of-the-art multiscale methods was presented, which was followed by a discussion brought by a rough-on-rough sliding problem simulated using a now classical coupling method (Bridging Domain method). The influence of the coupling scheme was quantified by measuring the kinetic energy of atoms close to the asperities and was compared with a full MD and a reduced MD models. The results show clearly that an ad-hoc damping of high frequency waves changes the dynamics of sliding friction. This is an important limitation of most current multiscale approaches, and prevents their wide usage in sliding contact simulations, in which one expects large thermal fluxes to be generated.

Consequently, an alternative multiscale approach was proposed to concurrently couple molecular dynamics and a finite element model at finite temperatures. The proposed approach is based on the generalized Langevin equation and resorts to spatial filters. The thermal coupling that handles the heat flux between the atomistic and continuum models was presented. The fundamental idea is that the high frequency waves that are not represented by the finite elements are damped by the friction force through spatial filtering. The balance with the damped energy is performed through the random force. Finally, we presented two-dimensional numerical examples: i) wave propagation at constant finite temperature, ii) thermo-mechanical wave

propagation, and iii) a dynamic contact problem. In all cases the coupled simulations were compared with full MD simulations and found to be in good agreement. While finite-temperature multiscale approaches show great prospect for friction simulations, it is important to emphasize that more research is needed to improve the computing performance (especially in three dimensions), and to explore the thermomechanical mechanisms contributing to friction within this new framework.

Acknowledgments This material is based on the work supported by the Swiss National Foundation under Grant no 200021_122046/1 and the European Research Council Starting Grant no 240332.

References

1. H. Czichos, *Tribology* (Elsevier, Amsterdam 1978)
2. A. Majumdar, B. Bhushan, Role of fractal geometry in roughness characterization and contact mechanics of surfaces. J. Tribol. **112**(2), 205–216 (1990)
3. B.N.J. Persson, Contact mechanics for randomly rough surfaces. Surf. Sci. Rep. **61**(4), 201–227 (2006)
4. S.B. Ramisetti, C. Campa, Anciaux, J.F. Molinari, M.H. Mser, M.O. Robbins, The autocorrelation function for island areas on self-affine surfaces. J. Phys. Condens. Matter **23**(21), 215004 (2011)
5. B. Luan, M.O. Robbins, The breakdown of continuum models for mechanical contacts. Nature **435**(7044), 929–932 (2005)
6. G.V. Dedkov, Experimental and theoretical aspects of the modern nanotribology. Phys. Status Solidi A **179**(1), 375 (2000)
7. J. Gao, W.D. Luedtke, D. Gourdon, M. Ruths, J.N. Israelachvili, U. Landman, Frictional forces and Amontons' law: From the molecular to the macroscopic scale. J. Phys. Chem. B **108**(11), 3410–3425 (2004)
8. J.O. Koskilinna, M. Linnolahti, T.A. Pakkanen, Friction coefficient for hexagonal boron nitride surfaces from ab initio calculations. Tribol. Lett. **24**(1), 37–41 (2006)
9. M. Renouf, F. Massi, N. Fillot, A. Saulot, Numerical tribology of a dry contact. Tribol. Int. **44**(78), 834–844 (2011)
10. J.F. Jerier, J.F. Molinari, Normal contact between rough surfaces by the discrete element method. Tribol. Int. **47**, 1–8 (2012)
11. V.S. Deshpande, A. Needleman, E. Van der Giessen, Discrete dislocation plasticity analysis of static friction. Acta Mater. **52**(10), 3135–3149 (2004)
12. S. Hyun, L. Pei, J.F. Molinari, M.O. Robbins, Finite-element analysis of contact between elastic self-affine surfaces. Phys. Rev. E **70**(2), 026117 (2004)
13. P. Wriggers, T.A. Laursen, *Computational Contact Mechanics* (Springer, Dordrecht, 2008)
14. B. Luan, M.O. Robbins, Contact of single asperities with varying adhesion: Comparing continuum mechanics to atomistic simulations. Phys. Rev. E **74**(2), 026111 (2006)
15. Y. Mo, I. Szlufarska, Roughness picture of friction in dry nanoscale contacts. Phys. Rev. B **81**(3), 035405 (2010)
16. P. Spijker, G. Anciaux, J.F. Molinari, The effect of loading on surface roughness at the atomistic level. Comput. Mech. **50**(3), 273–283 (2011)
17. T.J.R. Hughes, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis* (Dover Publications, New York, 2000)
18. O.C. Zienkiewicz, R.L. Taylor, J.Z. Zhu, *The Finite Element Method: Its Basis & Fundamentals* (Elsevier Butterworth-Heinemann, Amsterdam, 2005)
19. D.C. Rapaport, *The Art of Molecular Dynamics Simulation* (Cambridge University Press, 2004)

20. K. Komvopoulos, J. Yang, Dynamic analysis of single and cyclic indentation of an elasticplastic multi-layered medium by a rigid fractal surface. J. Mech. Phys. Solids **54**(5), 927–950 (2006)
21. K. Komvopoulos, Z.Q. Gong, Stress analysis of a layered elastic solid in contact with a rough surface exhibiting fractal behavior. Int. J. Solids Struct. **44**(78), 2109–2129 (2007)
22. K. Komvopoulos, Effects of multi-scale roughness and frictional heating on solid body contact deformation. C. R. Mnique **336**(12), 149–162 (2008)
23. S. Hyun, M.O. Robbins, Elastic contact between rough surfaces: Effect of roughness at large and small wavelengths. Tribol. Int. **40**(10–12), 1413–1422 (2007)
24. H.J.C. Berendsen, *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics* (Cambridge University Press, 2007)
25. M. Griebel, S. Knapek, G. Zumbusch, *Numerical Simulation in Molecular Dynamics: Numerics, Algorithms, Parallelization, Applications* (Springer, November 2010)
26. J. Rottler, M.O. Robbins, Macroscopic friction laws and shear yielding of glassy solids. Comput. Phys. Commun. **169**(13), 177–182 (2005)
27. O.M. Braun, A.G. Naumovets, Nanotribology: Microscopic mechanisms of friction. Surf. Sci. Rep. **60**(67), 79–158 (2006)
28. H.H. Yu, P. Shrotriya, Y.F. Gao, K.S. Kim, Micro-plasticity of surface steps under adhesive contact: Part I surface yielding controlled by single-dislocation nucleation. J. Mech. Phys. Solids **55**(3), 489–516 (2007)
29. C. Campañá, M.H. Müser, Contact mechanics of real vs. randomly rough surfaces: a Green's function molecular dynamics study. EPL (Europhysics Letters) **77**(3), 38005 (2007)
30. H.J. Kim, W.K. Kim, M.L. Falk, D.A. Rigney, MD simulations of microstructure evolution during high-velocity sliding between crystalline materials. Tribol. Lett. **31**(1), 67–67 (2008)
31. C. Yang, B.N.J. Persson, Contact mechanics: contact area and interfacial separation from small contact to full contact. J. Phys.: Condens. Matter **20**(21), 215214 (2008)
32. T. Liu, G. Liu, P. Wriggers, S. Zhu, Study on contact characteristic of nanoscale asperities by using molecular dynamics simulations. J. Tribol. **131**(2), 022001–022001 (2009)
33. P. Spijker, G. Anciaux, J.F. Molinari, Dry sliding contact between rough surfaces at the atomistic scale. Tribol. Lett. **44**(2), 279–285 (2011)
34. P. Spijker, G. Anciaux, J.F. Molinari, Relations between roughness, temperature and dry sliding friction at the atomic scale. Tribol. Int. **59**, 222–229 (2013)
35. F.F. Abraham, R. Walkup, H. Gao, M. Duchaineau, T.D.D.L. Rubia, M. Seager, Simulating materials failure by using up to one billion atoms and the world's fastest computer: work-hardening. Proc. Nat. Acad. Sci. **99**(9), 5783–5787 (2002)
36. J. Broughton, F. Abraham, N. Bernstein, E. Kaxiras, Concurrent coupling of length scales: Methodology and application. Phys. Rev. B **60**(4), 2391–2403 (1999)
37. R. Miller, E.B. Tadmor, R. Phillips, M. Ortiz, Quasicontinuum simulation of fracture at the atomic scale. Modell. Simul. Mater. Sci. Eng. **6**(5), 607–638 (1998)
38. V.B. Shenoy, R. Miller, E.B. Tadmor, R. Phillips, M. Ortiz, Quasicontinuum models of interfacial structure and deformation. Phys. Rev. Lett. **80**(4), 742–745 (1998)
39. W.A. Curtin, R.E. Miller, Atomistic/continuum coupling in computational materials science. Modell. Simul. Mater. Sci. Eng. **11**(3), R33–R68 (2003)
40. W.K. Liu, E.G. Karpov, S. Zhang, H.S. Park, An introduction to computational nanomechanics and materials. Comput. Methods Appl. Mech. Eng. **193**(17–20), 1529–1578 (2004)
41. H.S. Park, W.K. Liu, An introduction and tutorial on multiple-scale analysis in solids. Comput. Methods Appl. Mech. Eng. **193**(17–20), 1733–1772 (2004)
42. G. Lu, E. Kaxiras, *An overview of multiscale simulations of materials. Handbook of Theoretical and Computational Nanotechnology* (American Scientific Publishers, Stevenson Ranch, 2005), p. 10
43. R.E. Miller, E.B. Tadmor, A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. Modell. Simul. Mater. Sci. Eng. **17**(5), 053001 (2009)
44. J.M. Wernik, S.A. Meguid, Coupling atomistics and continuum in solids: status, prospects, and challenges. Int. J. Mech. Mater. Des. **5**(1), 79–110 (2009)

45. E. Weinan, B. Engquist, X. Li, W. Ren, E. Vanden-Eijnden, Heterogeneous multiscale methods: a review. Commun. Comput. Phys. **2**(3), 367–450 (2007)
46. E.B. Tadmor, M. Ortiz, R. Phillips, Quasicontinuum analysis of defects in solids. Philos. Mag. A **73**(6), 1529–1563 (1996)
47. R. Miller, E.B. Tadmor, The quasicontinuum method: overview, applications and current directions. J. Comput. Aided Mater. Des. **9**(3), 203–239 (2002)
48. L.E. Shilkrot, R.E. Miller, W.A. Curtin, Coupled atomistic and discrete dislocation plasticity. Phys. Rev. Lett. **89**(2), 025501 (2002)
49. S.P. Xiao, T. Belytschko, A bridging domain method for coupling continua with molecular dynamics. Comput. Methods Appl. Mech. Eng. **193**(17–20), 1645–1669 (2004)
50. S. Kohlhoff, P. Gumbsch, H.F. Fischmeister, Crack propagation in b.c.c. crystals studied with a combined finite-element and atomistic model. Philos. Mag. A **64**(4), 851–878 (1991)
51. G.J. Wagner, W.K. Liu, Coupling of atomistic and continuum simulations using a bridging scale decomposition. J. Comput. Phys. **190**(1), 249–274 (2003)
52. R. Miller, M. Ortiz, R. Phillips, V. Shenoy, E.B. Tadmor, Quasicontinuum models of fracture and plasticity. Eng. Fract. Mech. **61**(3–4), 427–444 (1998)
53. V.B. Shenoy, R. Miller, E.B. Tadmor, D. Rodney, R. Phillips, M. Ortiz, An adaptive finite element approach to atomic-scale mechanicsthe quasicontinuum method. J. Mech. Phys. Solids **47**(3), 611–642 (1999)
54. J. Knap, M. Ortiz, An analysis of the quasicontinuum method. J. Mech. Phys. Solids **49**(9), 1899–1923 (2001)
55. V. Shenoy, V. Shenoy, R. Phillips, Finite temperature quasicontinuum methods. MRS Online Proc. Libr. **538** (1998)
56. L.M. Dupuy, E.B. Tadmor, R.E. Miller, R. Phillips, Finite-temperature quasicontinuum: molecular dynamics without all the atoms. Phys. Rev. Lett. **95**(6), 060202 (2005)
57. Z. Tang, H. Zhao, G. Li, N.R. Aluru, Finite-temperature quasicontinuum method for multiscale analysis of silicon nanostructures. Phys. Rev. B **74**(6), 064110 (2006)
58. Y. Kulkarni, J. Knap, M. Ortiz, A variational approach to coarse graining of equilibrium and non-equilibrium atomistic description at finite temperature. J. Mech. Phys. Solids **56**(4), 1417–1449 (2008)
59. J. Marian, G. Venturini, B.L. Hansen, J. Knap, M. Ortiz, G.H. Campbell, Finite-temperature extension of the quasicontinuum method using langevin dynamics: entropy losses and analysis of errors. Modell. Simul. Mater. Sci. Eng. **18**(1), 015003 (2010)
60. E.B. Tadmor, F. Legoll, W.K. Kim, L.M. Dupuy, R.E. Miller, Finite-temperature quasicontinuum. Appl. Mech. Rev. **65**(1), 010803–010803 (2013)
61. L.E. Shilkrot, W.A. Curtin, R.E. Miller, A coupled atomistic/continuum model of defects in solids. J. Mech. Phys. Solids **50**(10), 2085–2106 (2002)
62. E. Van der Giessen, A. Needleman, Discrete dislocation plasticity: a simple planar model. Modell. Simul. Mater. Sci. Eng. **3**(5), 689 (1995)
63. B. Shiari, R.E. Miller, W.A. Curtin, Coupled atomistic/discrete dislocation simulations of nanoindentation at finite temperature. J. Eng. Mater. Technol. **127**(4), 358–368 (2005)
64. S. Qu, V. Shastry, W.A. Curtin, R.E. Miller, A finite-temperature dynamic coupled atomistic/discrete dislocation method. Modell. Simul. Mater. Sci. Eng. **13**(7), 1101 (2005)
65. H.B. Dhia, Problémes mécaniques multi-échelles: la méthode arlequin. Comptes Rendus de l'Académie des Sciences - Series IIB - Mechanics-Physics-Astronomy, 326(12):899–904 (1998)
66. H.B. Dhia, G. Rateau, The arlequin method as a flexible engineering design tool. Int. J. Numer. Meth. Eng. **62**(11), 14421462 (2005)
67. P.T. Bauman, H.B. Dhia, N. Elkhodja, J.T. Oden, S. Prudhomme, On the application of the arlequin method to the coupling of particle and continuum models. Comput. Mech. **42**(4), 511–530 (2008)
68. S. Prudhomme, H.B. Dhia, P.T. Bauman, N. Elkhodja, J.T. Oden, Computational analysis of modeling error for the coupling of particle and continuum models by the arlequin method. Comput. Methods Appl. Mech. Eng. **197**(4142), 3399–3409 (2008)

69. T. Belytschko, S.P. Xiao, Coupling methods for continuum model with molecular model. Int. J. Multiscale Comput. Eng. **1**(1), 115–126 (2003)
70. G. Anciaux, O. Coulaud, J. Roman, G. Zerah, *Ghost force reduction and spectral analysis of the 1D bridging method*, Technical report (INRIA, HAL, 2008)
71. G. Anciaux, S.B. Ramisetti, J.F. Molinari, A finite temperature bridging domain method for MD-FE coupling and application to a contact problem. Comput. Methods Appl. Mech. Eng. **205208**, 204212 (2011)
72. H.S. Park, E.G. Karpov, P.A. Klein, W.K. Liu, Three-dimensional bridging scale analysis of dynamic fracture. J. Comput. Phys. **207**(2), 588–609 (2005)
73. S.A. Adelman, Generalized langevin equation approach for atom/solid-surface scattering: collinear atom/harmonic chain model. J. Chem. Phys. **61**(10), 4242–4246 (1974)
74. S.A. Adelman, Generalized langevin theory for gas/solid processes: dynamical solid models. J. Chem. Phys. **65**(9), 3751–3762 (1976)
75. S.A. Adelman, Generalized langevin equation approach for atom/solid-surface scattering: general formulation for classical scattering off harmonic solids. J. Chem. Phys. **64**(6), 2375–2389 (1976)
76. W. Cai, M. de Koning, V.V. Bulatov, S. Yip, Minimizing boundary reflections in coupled-domain simulations. Phys. Rev. Lett. **85**(15), 3213–3216 (2000)
77. E. Weinan, Z. Huang, Matching conditions in atomistic-continuum modeling of materials. Phys. Rev. Lett. **87**(13), 135501 (2001)
78. E. Weinan, Z. Huang, A dynamic atomisticcontinuum method for the simulation of crystalline materials. J. Comput. Phys. **182**(1), 234–261 (2002)
79. G.J. Wagner, E.G. Karpov, W.K. Liu, Molecular dynamics boundary conditions for regular crystal lattices. Comput. Methods Appl. Mech. Eng. **193**(1720), 1579–1601 (2004)
80. E.G. Karpov, H.S. Park, W.K. Liu, A phonon heat bath approach for the atomistic and multiscale simulation of solids. Int. J. Numer. Meth. Eng. **70**(3), 351–378 (2007)
81. N. Mathew, R.C. Picu, M. Bloomfield, Concurrent coupling of atomistic and continuum models at finite temperature. Comput. Methods Appl. Mech. Eng. **200**(5–8), 765–773 (2011)
82. S.B. Ramisetti, G. Anciaux, J.F. Molinari, Spatial filters for bridging molecular dynamics with finite elements at finite temperatures. Comput. Methods Appl. Mech. Eng. **253**, 28–38 (2013)
83. G. Anciaux, J.F. Molinari, Sliding of rough surfaces and energy dissipation with a 3D multiscale approach. Int. J. Numer. Meth. Eng. **83**(8–9), 1255–1271 (2010)
84. B.N.J. Persson, O. Albohr, U. Tartaglino, A.I. Volokitin, E. Tosatti, On the nature of surface roughness with application to contact mechanics, sealing, rubber friction and adhesion. J. Phys.: Condens. Matter **17**(1), R1–R62 (2005)
85. H.O. Peitgen, D. Saupe, Y. Fisher, M. McGuire, R.F. Voss, M.F. Barnsley, R.L. Devaney, B.B. Mandelbrot, *The Science of Fractal Images*, 1st edn. (Springer, New York, 1988)
86. R.F. Voss, Random fractal forgeries, in *Fundamental Algorithms for Computer Graphics*, ed. by R.A. Earnshaw (Springer, Heidelberg, 1985), pp. 805–835
87. G. Anciaux, Simulation multi-échelles des solides par une approche couplée dynamique moléculaire/éléments finis. De la modélisation á la simulation haute performance. Ph.D. thesis, University of Bordeaux (INRIA, CEA), France, July 2007
88. J. Fish, M.A. Nuggehally, M.S. Shephard, C.R. Picu, S. Badia, M.L. Parks, M. Gunzburger, Concurrent AtC coupling based on a blend of the continuum stress and the atomistic force. Comput. Methods Appl. Mech. Eng. **196**(4548), 4548–4560 (2007)
89. K. Fackeldey, R. Krause, Multiscale coupling in function spaceweak coupling between molecular dynamics and continuum mechanics. Int. J. Numer. Meth. Eng. **79**(12), 15171535 (2009)
90. K. Fackeldey, The Weak Coupling Method for Coupling Continuum Mechanics with Molecular Dynamics. Ph.D. thesis, Bonn, February 2009
91. S.B. Ramisetti, G. Anciaux, J.F. Molinari, A concurrent atomistic and continuum coupling method with applications to thermo-mechanical problems. Submitted, 2013

Chapter 15
Effect of Capillary Condensation on Nanoscale Friction

Rosario Capozza, Itay Barel and Michael Urbakh

Abstract While formation of capillary bridges significantly contributes to the adhesion and friction at micro- and nanoscales, many key aspects of dynamics of capillary condensation and its effect on friction forces are still not well understood. Here, by analytical model and numerical simulations, we address the origin of reduction of friction force with velocity and increase of friction with temperature, which have been experimentally observed under humid ambient conditions. We demonstrate that adding a low amplitude oscillatory component to the pulling force, when applied at the right frequency, can significantly suppress condensation of capillary bridges and thereby reduce friction. The results obtained show that frictional measurements performed in this mode can provide significant information on the mechanism of frictional aging.

15.1 Introduction

The ability to control and manipulate friction during sliding is extremely important for a large variety of applications. Development of novel efficient methods to control friction requires understanding microscopic mechanisms of frictional phenomena. One of the main difficulties in understanding and predicting frictional response is the intrinsic complexity of highly non-equilibrium processes going on in any tribological contact, which include detachment and reattachment of multiple microscopic junctions (bonds) between the surfaces in relative motion [1–8]. Even for an apparently sharp AFM tip sliding on a crystalline surface, the actual interface consists of

R. Capozza (✉)
International School for Advanced Studies (SISSA), Via Bonomea 265, I-34136 Trieste, Italy
e-mail: rosario.capozza@gmail.com

I. Barel
Department of Chemistry and Biochemistry, University of California, Santa Barbara, USA
e-mail: ibarel@chem.ucsb.edu

M. Urbakh
School of Chemistry, Tel Aviv University, Ornstein 208, Tel Aviv, Israel
e-mail: urbakh@post.tau.ac.il

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_15

an ensemble of individual junctions [9, 10]. On larger scales the multicontact picture becomes even more obvious. Friction is not simply the sum of single-junction responses, but is influenced by temporal and spatial dynamics across the entire ensemble of junctions that form the frictional interface. The way how individual junctions can be averaged to yield friction response has been the focus of intense research in the past decades [7, 8, 11–16]. These junctions may represent molecular bonds, capillary bridges, asperities between rough surfaces, and for lubricated friction they can mimic patches of solidified lubricant or its domains.

In this chapter we focus on the contribution of capillary bridges to frictional response. Important examples include the flow of granular materials [17], adhesion of insects or Geckos to surfaces [18], friction in micro- and nanoelectromechanical systems (MEMS/NEMS) [19]. Condensation of capillary bridges may be responsible for frictional ageing that is manifested as logarithmic increase of static friction with time during which two surfaces are held in stationary contact [20]. Capillary bridges play a crucial role in the operation of atomic force microscopy (AFM) under humid ambient conditions and their formation often dominates the measured forces [21–25]. Recent investigations of dependencies of frictional force on hold time, Δt, velocity, V, and temperature, T, suggested that capillary condensation is thermally activated [20–24]. It has been found that under ambient conditions the nanoscopic friction force decreases linearly with ln V at low velocities, and increases with T [21–23]. These observations disagree with predictions of thermal Prandtl-Tomlinson model of friction [26–28] and with results of experiments carried out under ultrahigh vacuum, which show a logarithmic increase of friction force with velocity and reduction of friction with temperature [29, 30]. In order to explain the observed logarithmic decrease of friction force with velocity a phenomenological model has been proposed [21] that is based on the thermally activated mechanism of nucleation of capillary bridges in a gap between asperities on the contacting surfaces. The model suggested that higher velocities correspond to shorter contact times and hence to smaller number of water bridges formed between the tip and substrate, and as a result the friction force decreases with V.

Despite importance of water bridges in frictional phenomena and the growing efforts in the field, many key aspects of dynamics of capillary condensation and its effect on friction forces are still not well understood. The present study addresses this problem through an analytical model and numerical simulations.

One unique path to controlling and ultimately manipulating the friction forces between material interfaces is through externally imposed oscillations of small amplitude and energy. Validity of this approach has been demonstrated experimentally at nano [31–33] and macroscales [34–36] and numerically with minimal models [37, 38] and molecular dynamics simulations [39, 40].

In this Chapter we demonstrate that adding a low amplitude oscillatory component to the pulling force, when applied at the right frequency, can significantly suppress the formation of capillary bridges and thereby reduce friction. The results obtained show that frictional measurements performed in this configuration can provide significant information on the kinetics of bridge formation and their stiffness.

15.2 Model

In our model and simulations we consider a rigid tip with mass M and center-of-mass coordinate X that interacts with the underlying surface through an array of contacts representing capillary bridges. The tip is pulled along the surface with a velocity V_d, through a linear spring of spring constant, K_d (see Fig. 15.1). In accordance with recent observations [21–23], we assume that nucleation of capillary bridges occurs preferentially between asperities at the tip and substrate surfaces, whose contact radius is much smaller than the AFM tip radius.

Nucleation of capillary bridges is a thermally activated process [20–24], and its rate, k_{on}, can be described by the equation

$$k_{on}^{(i)} = \omega_{on} \exp\left[-\Delta E_{on}^{i}/k_B T\right], \tag{15.1}$$

where ω_{on}, $\Delta E_{on}^{(i)}$ are the attempting frequency and the barrier height for the ith bridge. The energy barrier is proportional to the liquid volume needed to nucleate the liquid bridge, $v = a_0^2 h\,(r_i)$, where a_0^2 is a typical nucleation area and h is the gap between the surfaces at the nucleating site. The capillary bridges are characterized by a broad distribution of the barrier heights [15, 20–23, 41, 42] that leads to time-dependent increase of frictional strength, or frictional aging. Here, for simplicity we assume a uniform distribution of barrier heights above a bottom threshold, $\Delta E_{on}^{\min}$. Then a fraction of junctions with $\Delta E_{on}^{(i)} < \Delta E_{on}^{*}$ is equal to $s_E\left(\Delta E_{on}^{*} - \Delta E_{on}^{\min}\right)$, with s_E being the density of the distribution. Qualitative conclusions of this work are not sensitive to a particular choice of the distribution. It is worth pointing out that this model also describes a single capillary bridge growing with time. In this case the number of contacts represents the bridge size, while the distribution of activation energies constitutes the barrier to overcome in order to increase the size.

As long as a bridge is intact, it responds elastically to the applied force, and it can be modeled as an elastic spring with stiffness, κ [43]. Under the action of the pulling force the bridges are stretched in the lateral direction with a velocity equal to the velocity of the tip, $\dot{X}$, and therefore the tip experiences the surface force,

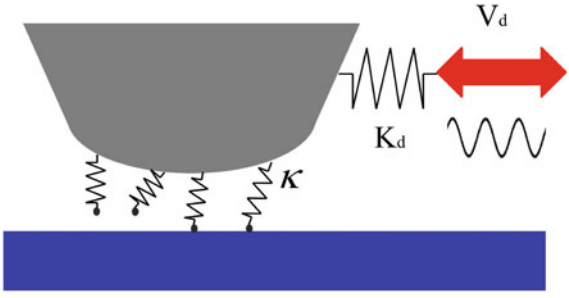


Fig. 15.1 Schematic sketch of a model geometry

$F_{cap} = -\sum_{N_s}^{i} f_i$, where $f_i = \kappa \left(l_i\,(t) - l_i^0\right), l_i(t)$ is the time-dependent bridge length, l_i^0 is the bridge length at moment of nucleation and a sum is taken over all bridges.

When the force, f_i, acting on the ith bridge exceeds the threshold value, f_c, corresponding to a maximum force that a capillary bridge can hold the bridge detaches from the tip. It should be noted that the height of the potential barrier for rupture of condensed bridges is much higher than $k_B T$ [44, 45], and because of this the effect of temperature on rupture processes is negligible.

Here for the estimation of the threshold force f_c we use the value of the maximum adhesion force (the capillary force) between the asperity and the tip that is given by the equation

$$f_c = 2\pi\gamma a_0 \left(\cos\theta_T + \cos\theta_S\right) \tag{15.2}$$

where γ is the liquid surface tension, a_0 is the average radius of asperity, and θ_T and θ_S are the contact angles at the tip and the substrate, respectively. (15.2) should be applied with caution at microscopic scales [44–46], because it relies on macroscopic concepts. It should be also noted that the maximum force that a capillary bridge can hold in the lateral direction before it breaks can be different from the maximum adhesion force in vertical direction, f_c.

In the framework of the model described above the motion of the driven tip is described by the following equation:

$$M\ddot{X} + \eta\dot{X} - F_{cap} + K_d(X - V_d t) = 0 \tag{15.3}$$

where η is a damping coefficient responsible for the dissipation of the tip kinetic energy to phonons and other degrees of freedom which are not considered explicitly here. In a wide range of parameters the results of calculations are independent of the value of η.

The instantaneous lateral spring force, which is the main observable in friction experiments, reads as $F = -K_d(X - V_d t)$, and its time average is equal to the friction force $\langle F \rangle$.

15.3 Temperature and Velocity Dependencies of Friction

The essential difference of the model described above from the previous works [14, 47, 48], where friction was described in terms of rupture and reattachment of nanoscale contacts, is a broad distribution of nucleation barriers. This leads to a long time scale dependence of number of condensed bridges on the time of contact between the tip and the substrate, and thereby to the effect of frictional aging. The average number of bridges, $N_{mf}(t)$, which are formed during the time t, can be estimated using the mean-field approach that gives [42]:

$$N(t) = N_0 \times \begin{cases} t/\tau, & t < \tau \\ ln(t/\tau) + \gamma, & t > \tau \end{cases} \tag{15.4}$$

where $N_0 = s_E k_B T$, $1/\tau = \omega_{on} \exp[-\Delta E_{on}^{\min}/k_B T]$ and $\gamma \approx 0.5572$ is the Euler-Mascheroni constant. The number of condensed bridges as a function of time shows a linear behavior at short times and logarithmic one at longer times. A logarithmic increase of number of condensed bridges with time has been already suggested in previous studies, which considered the effect of capillary condensation on friction at nano and macro scales [20, 21].

The mean field description presented above allows to describe the effect of capillary condensation on velocity and temperature dependencies of friction. When the tip is pulled with velocity V_d, the force, f_i, acting on a water bridge grows with a time-dependent rate $K_{eff} V_d / N(t)$, where $K_{eff}(t) = \frac{N(t)\kappa K_d}{N(t)\kappa + K_d}$ is the effective stiffness of the system that includes the pulling spring, K_d, and capillary bridges with total stiffness $N(t)k$. The bridge will be ruptured at the time, t_0, for which $f_i = f_c$. This yields the following equation for t_0:

$$\int_{\tau_0}^{t_0} \frac{K_{eff}}{N(t)} V_d \, dt = f_c \,, \tag{15.5}$$

where τ_0 is a time needed for nucleation of one bridge, $N(\tau_0) = 1$, and it can be estimated as $\tau_0 \approx \tau (N_0)^{-1}$. Assuming that $N(t)k \gg Kd$, (15.5) can be rewritten as $\int_{\tau_0}^{t_0} \frac{dt}{N(t)} = \frac{f_c}{K_d V_d}$. Then considering a stick-slip regime of motion, the time averaged friction force, $\langle F \rangle$, can be calculated as

$$\langle F \rangle \approx \frac{1}{2} K_d V_d (t_0 - \tau_0) \tag{15.6}$$

It should be noted that (15.5)–(15.6) have been derived under the assumption that all condensed bridges are ruptured simultaneously at the time t_0, when the force acting on one of them approaches f_c. The validity of this assumption will be tested below using stochastic simulations.

At low pulling velocities or relatively high temperatures, for which a large number of bridges are condensed during the stick phase of motion $N(t_0) \gg N_0$, the friction force can be written as

$$\langle F \rangle = f_c N_0 \Im \left(\frac{f_c N_0}{K V \tau} \right), \tag{15.7}$$

where $\Im$ is a scaling function, which depends only on one parameter, $\zeta = \frac{f_c N_0}{K V \tau}$, and increases with increasing ζ. It should be noted that both N_0 and τ are functions of temperature, and thus (15.7) provides a direct link between the temperature and velocity dependencies of friction. In particular, in the range of parameters, where $\langle F \rangle$ grows linearly with T, the average friction force decreases approximately logarithmically with velocity, $\langle F \rangle \approx const - ln V_d$. This result can be obtained substituting

(15.4) for $N(t)$ in (15.5) and neglecting the terms of the form of $ln\, ln\left[\frac{f_c N_0}{K_d V_d \tau}\right]$, which are weakly dependent on temperature and velocity.

Equation (15.7) exhibits the same characteristic friction-velocity and friction-temperature dependences as observed in the experiments [21–23]: (i) decrease of the friction force with velocity and increase with temperature, (ii) two different regimes of velocity dependence of friction for high and low pulling velocities, (iii) increase of critical velocity, V_c, for a transition between the two velocity regimes with T. In accordance with experimental observations [21–23], at low velocities $\langle F \rangle$ shows a logarithmic-like decrease with V_d, and it levels off at high velocities.

The above results have been obtained in the framework of mean field description, which is based on consideration of average number of condensed bridges, $N(t)$, and assumes that all condensed bridges are ruptured simultaneously. This is an approximation, and a more accurate approach requires a consideration of stochastic dynamics of nucleation and rupture of capillary bridges that is governed by the nucleation rates $k_{on}^{(i)}$ and coupled to the motion of tip in (15.3) [42]. Results of stochastic simulations are presented in Figs. 15.2 and 15.3, and they show velocity and temperature dependencies of friction force, and of time-averaged number of condensed bridges. One can see that the observed variations of friction force with V_d and T are entirely determined by the corresponding variations of number of condensed bridges. The proposed stochastic model enables us to describe all features of friction-velocity and friction-temperature curves, which have been observed experimentally under humid ambient conditions. The simulation parameters, f_c, ΔE_{on}^0, n_0, have been chosen to fit qualitatively the experimental data [21–23], and their values are close to those suggested by experimental studies.

The results of simulations shown in Figs. 15.2 and 15.3 are qualitatively similar to the predictions of the mean field approach [42]. However, for the same values of system parameters the friction forces obtained in the stochastic simulations are significantly higher (almost two-times) than those derived in the mean field approximation. This discrepancy results from complex dynamics of rupture and reattachment of capillary bridges, which has been observed in simulations and neglected in the approximate description. We found that part of condensed bridges is ruptured at the values of the applied force, which are considerably lower than the value needed to initiate a slip of the tip, and then they are nucleated again during the same stick interval. This effect leads to an enhancement of friction force peaks in the stick-slip series compared to those calculated in (15.5)–(15.6).

The kinetics of capillary condensation is characterized by a long-scale, logarithmic increase of number of condensed bridges with time that results in frictional ageing [20]. However, measurements of velocity and temperature dependencies of friction do not allow to distinguish this behavior from the condensation of molecular contacts characterized by a narrow distribution of nucleation barriers [42]. The only essential condition for observation of friction-velocity and friction-temperature curves similar to those in Figs. 15.2 and 15.3 is that formation of contacts (bridges) is a thermally activated process, while kinetics of rupture is temperature-independent, or only slightly influenced by T. This is different from the major assumption of

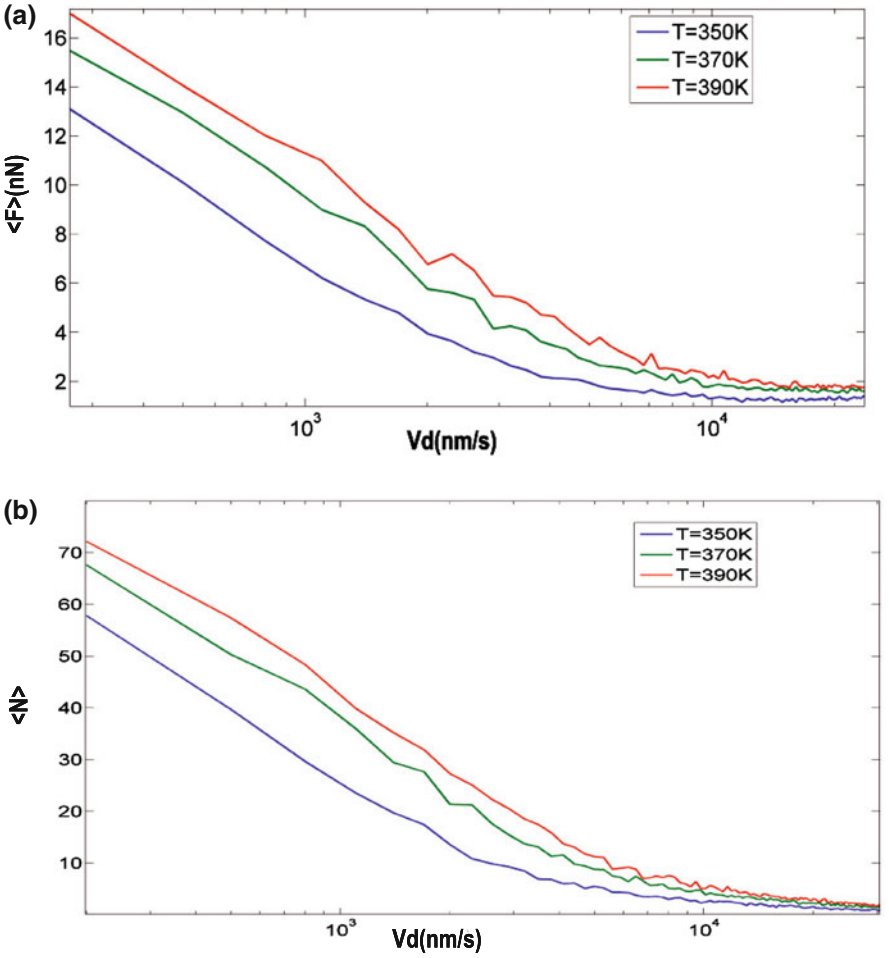


Fig. 15.2 Results of stochastic simulations for velocity dependence of the average friction force (**a**) and the average number of condensed bridges (**b**) calculated for three different temperatures. Parameter values: $M = 5 \times 10^{-11}$ kg, $\eta = 2 \times 10^{-5}$ kg/s, $f_c = 0.8$ nN, $n_0 = 2.5 \times 10^{18}$ m^{-2}, $\Delta E_{on}^0 = 4.2 \times 10^{-20}$ J, $\omega_{on}^0 = 10^7$ s^{-1}, $K = 6$ N/m, $k = 1$ N/m, $R = 30$ nm, $h_0 = 1$ nm

the standard thermally activated model of friction [26–28], according to which the rupture of nanoscale contacts (barrier crossing) is a thermally activated process. The above consideration shows that measurements of velocity and temperature dependencies of friction do not allow to draw definite conclusions on contribution of ageing processes to friction.

Nevertheless, friction measurements can provide an efficient way to study kinetics of bridge formation. This can be done using slide-hold-slide experiments [41], which are similar to slide-hold-slide tests on rocks [49, 50], where the external drive

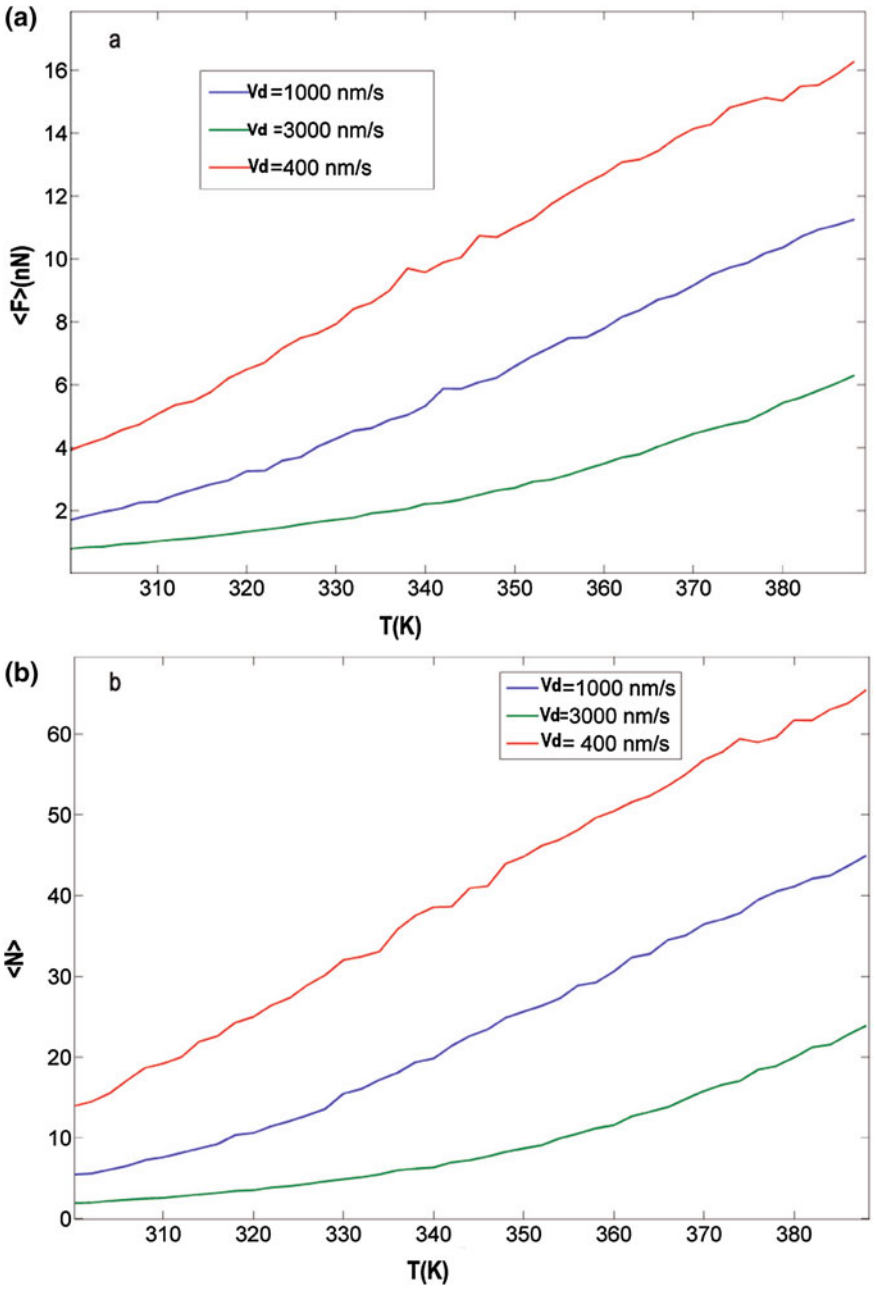


Fig. 15.3 Results of stochastic simulations for temperature dependence of the average friction force (**a**) and average number of condensed bridges (**b**) calculated for three different velocities. Parameter values as in Fig. 15.2

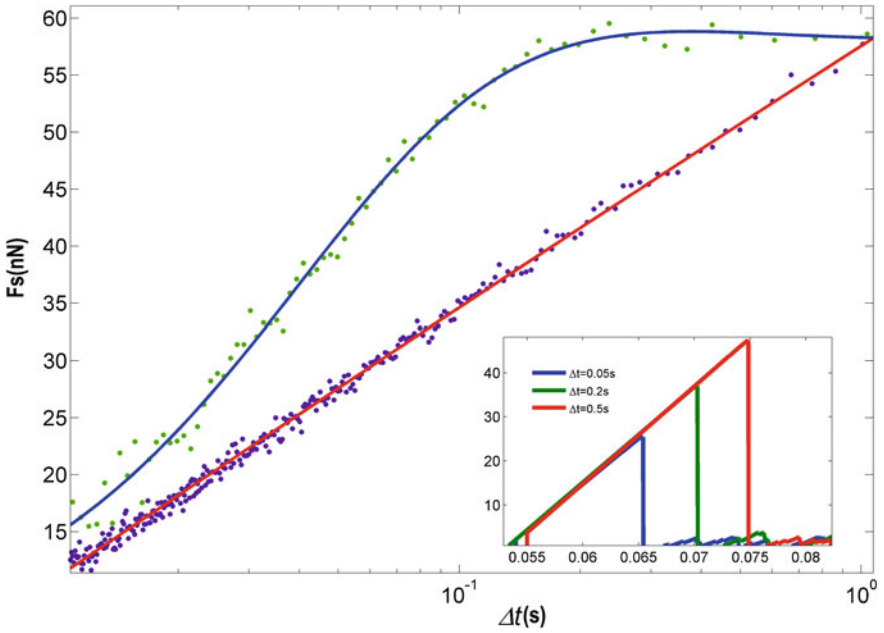


Fig. 15.4 Maximum friction force as a function of the hold time calculated for the case of capillary condensation (*green*) and for identical contacts (*purple*). The *red* and *blue* lines represent polynomial fitting curves to the calculated data. Inset shows the lateral force versus time calculated for three values of the hold time Δt. Pulling velocity, $V = 400\,\text{nm/s}$, and other parameters as in Figs. 15.4 and 15.6

is stopped for a certain hold time, Δt, and then reinitiated with the same pulling velocity, V_d. Figure 15.4 shows results of calculations of the maximum force, F_S, following specified hold periods. The presented forces have been found by averaging over hundred realizations. The maximum force is larger than the time-averaged friction force corresponding to the same velocity by an amount $\Delta F = F_s - \langle F \rangle$, called friction drop. For capillary condensation of bridges F_S, grows linearly with the logarithm of the hold time showing a strong effect of ageing on nanoscopic friction that results from a broad distribution of activation energy barriers for capillary condensation. In the case of identical contacts significant deviations from the logarithmic dependence are observed already for $\Delta t \geq 3\tau$. Thus, our simulations demonstrate that slide-hold-slide experiments can provide important information on kinetics of bridge formation and mechanism of frictional ageing.

Figure 15.4 shows that in the case of capillary condensation the calculated relative friction drop, $\Delta F / \langle F_s \rangle$ increases about ten times as the hold time increased by two orders of magnitude. This result is consistent with recent nanoscopic friction measurements [41], which found ten times increase of relative friction drop, when the hold time increased by two orders of magnitude. The magnitude of the ageing effect for the nanoscale single asperity contacts that has been found in experiments [41] and

in our simulations is much larger than for macroscopic multi-asperity friction experiments [49–51]. However, it should be noted that in the cited experiments the large ageing effect has been observed also for low humidities showing that the capillary condensation is not the only mechanism for frictional ageing. Further experimental and theoretical studies are needed to address the mechanisms that account for this.

15.4 Effect of Inplane Oscillations

In order to study the effect of inplane oscillatory modulation on friction, we add a low amplitude oscillatory component to the ramped forces. Then the displacement of the stage reads as $X_d = V_d t + A \cos(2\pi \upsilon t)$ where V_d is a constant velocity, υ and A are the frequency and amplitude of oscillations. When the tip is pulled with constant velocity, time series of the spring force exhibit stick-slip behavior corresponding to collective rupture and reattachment of capillary bridges (see Fig. 15.5a). Once small harmonic perturbations are introduced to the ramped loading, this picture changes significantly. Figure 15.5b shows that for frequencies of oscillations above a threshold one, υ_{th}, the force traces represent a set of alternating segments of stick-slip oscillation and low friction sliding, which are marked as t_{s-s} and t_{lag}, respectively [52]. For high frequencies (green curve in Fig. 15.6a) the stick-slip oscillations are completely suppressed and the spring force remains low over the entire time of

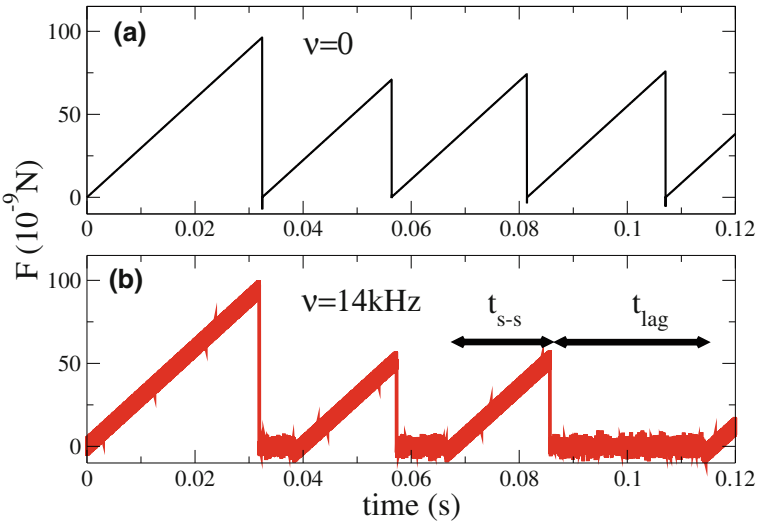


Fig. 15.5 **a** Typical stick-slip profile obtained when the tip is pulled with constant velocity. **b** Force trace corresponding to lateral modulations with the amplitude A = 1 nm and υ = 14 kHz. The trace shows alternating segments of stick-slip and low friction motion with durations t_{s-s} and t_{lag}, respectively

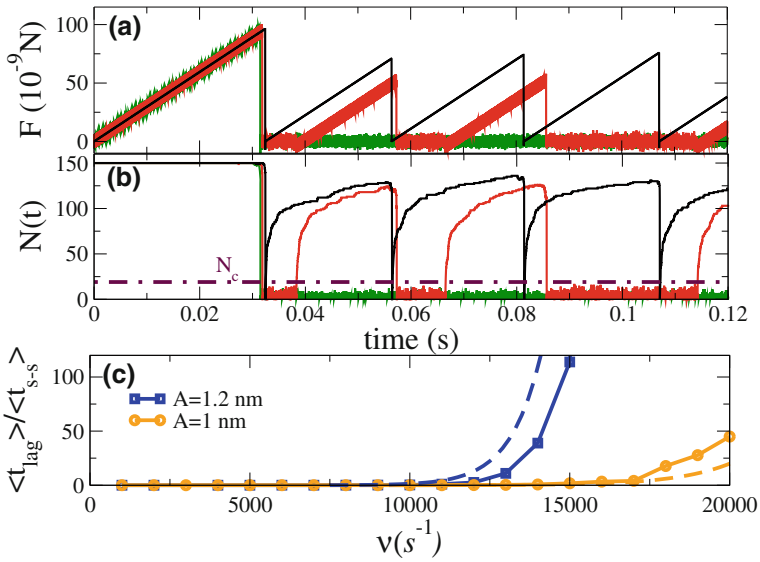


Fig. 15.6 **a** Force traces and **b** time-dependent number of condensed bridges, N(t), calculated in the absence of oscillations (*black curves*) and including the oscillatory component of the force with A = 1nm and υ = 14 kHz (*red curves*) and υ = 19.6 kHz (*green curves*). *Dashed-dotted line* in the panel (**b**) indicates the maximal number of bridges, N_c, which can be ruptured by the oscillatory modulations with the amplitude A = 1 nm estimated using (15.8). **c** Ratio $\langle t_{lag} \rangle / \langle t_{s-s} \rangle$ as a function of υ for two amplitudes of oscillations which are indicated in the figure. *Solid* and *dashed curves* present results of simulations and analytical theory, respectively

simulations shown in Fig. 15.6a. In order to elucidate the mechanism of reduction of friction we show in Fig. 15.6b the effect of oscillations on the time-dependent number of condensed bridges, *N(t)*. While in the absence of oscillations *N(t)* starts to grow monotonically directly after the slip event, application of small-amplitude oscillations suppresses the formation of capillary bridges during finite time-intervals, t_{lag}. This effect results in a low friction regime of motion corresponding to uncorrelated rupture of small clusters of bridges. Because of stochastic nature of bridge formation, the regime of low friction motion persists only for a finite time, t_{lag} until the number of bridges formed during a half-period of oscillations becomes larger than a critical value N_c that cannot be ruptured by the oscillatory component of force, as shown by dashed-dotted line in Fig. 15.6b. Then, oscillatory modulations become inefficient, and the number of bridges grows until the ramped component of the loading force causes the collective rupture of bridges similar to what happens in the absence of modulations. The value of loading force corresponding to the collective rupture in the presence of modulation is only slightly below the maximal force for constant velocity pulling (see Fig. 15.6a) and it depends weakly on υ. However, the length of the time-intervals of low friction, t_{lag}, increases rapidly with υ and as a result the average friction force decreases.

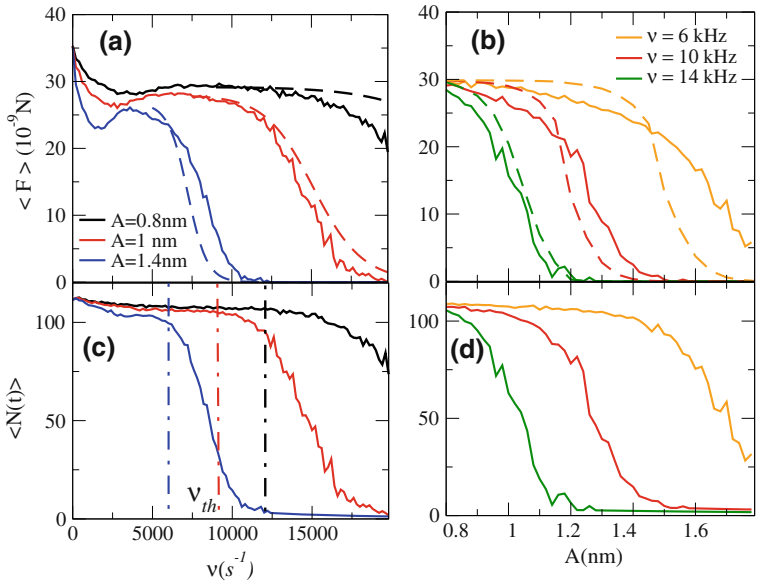


Fig. 15.7 Average friction force, $\langle F \rangle$, as a function of frequency (**a**) and amplitude (**b**) of oscillations, and the corresponding variations of average number of condensed bridges, $\langle N(t) \rangle$, reported in panels (**c**) and (**d**). *Solid and dashed curves* in (**a**) and (**b**) show results of simulations and analytical calculations according to (15.12), respectively. *Vertical dashed-dotted lines* in (**c**) present analytical estimations of the threshold frequency, υ_{th}, corresponding to different values of the amplitude A

Figure 15.6c shows the ratio $\langle t_{lag} \rangle / \langle t_{s-s} \rangle$ as a function of υ for two amplitudes of oscillation, where $\langle t_{lag} \rangle$ and $\langle t_{s-s} \rangle$ are the mean values of t_{lag} and t_{s-s} which have been calculated by averaging over a large number of realizations. Inplane oscillations induce nonzero time-intervals of low friction, t_{lag}, only for frequencies exceeding a threshold value, υ_{th}, which decreases with the amplitude of oscillations. Above the threshold frequency, $\langle t_{lag} \rangle / \langle t_{s-s} \rangle$ increases sharply with υ and the average friction force is reduced. It should be noted that over the entire range of data presented, the amplitude of the applied force oscillations was lower than the tenth of the force needed to initiate a slip of the tip in the case of constant velocity pulling. Thus, despite strong effects on stick-slip dynamics the perturbations are decidedly small.

Figure 15.7 presents the average friction force, $\langle F \rangle$, and the average number of condensed capillary bridges, $\langle N(t) \rangle$, as functions of frequency and amplitude of oscillations. One can see that for a given amplitude, $N(t)$ decreases steeply above a threshold frequency, υ_{th} (A), and the friction force follows this behavior. Figure 15.7b demonstrate a similar reduction of $\langle N(t) \rangle$ and $\langle F \rangle$ with increase of the amplitude of oscillations for a given frequency. Thus, application of small amplitude oscillations with frequencies of few kHz allows to reduce the friction force by more than one order of magnitude.

The mechanism of reduction of friction discussed here differs significantly from those suggested in previous works [32, 40] on the effects of oscillatory modulations on friction. It can operate only in tribological contacts exhibiting aging where times for bridge formation are widely distributed or there is a long time-scale strengthening of bridges. In these systems a frequency of force modulations can be chosen in a way that only a small number of bridges is formed during the half-period of oscillations, and these "fresh" bridges can be ruptured by the oscillatory component of the loading force. Thus, small inplane oscillations are able to prevent the formation of multiple bridges and reduce friction. In order to achieve similar reduction of friction in contacts characterized by a narrow distribution of times for bridge formation (in the absence of aging) much higher amplitudes or/and frequencies of modulation are required.

The main features of numerical results presented above can be reproduced by an analytical model based on a mean field description of ensemble of bridges that has been discussed above. When the tip is pulled with the velocity $\dot{X}_d$, the force, f_i, acting on a bridge grows with a time-dependent rate $K_{eff}\dot{X}_d/N(t)$, where $K_{eff}(t) = \frac{N(t)\kappa K_d}{N(t)\kappa+K_d}$ is the effective stiffness of the system. Assuming that all bridges are ruptured simultaneously at time t, when the force acting on one of them approaches f_c, the condition for a collective rupture is given by (15.5). This equation allows to estimate the key parameters which define the effect of oscillations on friction, such as: the maximal number of capillary bridges, N_c, which can be ruptured by the oscillatory modulations, and the threshold frequency, υ_{th}, above which the force oscillations produce low friction segments of motion. Considering that rupture occurs at a time corresponding to the maximum of the oscillatory force, $t = T_\upsilon = 1/(2\,\upsilon)$, and using (15.4) for *N(t)* we get the following approximate equation for $N_c = N(T_\upsilon)$:

$$N_c \frac{\kappa}{K_d}\left(1 - \frac{f_c}{4AK_d}N_c\right) = ln\left(1 + \frac{\kappa}{K_d}N_c\right) \tag{15.8}$$

Then, for a linear regime of growth of *N(t)* in (15.4) the threshold frequency can be estimated as $\upsilon_{th} \simeq \frac{N_0}{2\tau N_c}$. Estimations of N_c and $\upsilon_{th}(A)$ reported in Figs. 15.6b and 15.7c, respectively, show a good agreement with the results of simulations.

In order to describe the effect of oscillations on friction we have to consider a stochastic nature of bridge formation. The number of bridges formed during the half-period of oscillations fluctuates around an average value $N(T_\upsilon)$ given by (15.4). While *N(t)* is below the critical value, N_c, the spring force, *F(t)*, remains low but if during one of the oscillations $N(T_\upsilon)$ exceeds N_c the force oscillations become inefficient and *F(t)* grows. The probability, *P(m,t)*, of formation of m bridges in time t is given by a sum of probabilities of formation of all possible clusters of bridges of size m. In the case of identical bridges, *P(m,t)* can be easily calculated taking into account that there are $\frac{N_s!}{(N_s-m)!m!}$ clusters of size m, where N_s is a total number of available bridges. However, in systems exhibiting aging, different bridges have different barrier heights ΔE^i_{on}, and correspondingly different probabilities of formation which can be calculated as, $P_i(t) = 1 - exp\left(-k^i_{on}\,t\right)$, where k^i_{on} is the rate of formation of i-th bridge given by (15.1). Then, the probability $P(m,t)$ can

be found using the recursive equation [53]

$$P(m,t) = \begin{cases} \prod_{i=1}^{N_s} (1 - P_i(t)), & m = 0 \\ \frac{1}{m} \sum_{i=1}^{N_s} (-1)^{i-1} P(m-i,t) Q_i(t), & m > 0 \end{cases} \tag{15.9}$$

where $Q_i(t) = \sum_{j=1}^{N_s} \left(\frac{P_j(t)}{1-P_j(t)} \right)^i$. Then the average length of the time-interval, $\langle t_{lag} \rangle$, during which the number of condensed bridges is below N_c, and the spring force remains low, can be calculated as

$$\langle t_{lag} \rangle = \frac{T_\upsilon}{\tilde{P}(N_c, T_\upsilon)} \tag{15.10}$$

where $\tilde{P}(N_c, T_\upsilon) = \sum_{i=1}^{+\infty} P(i, T_\upsilon)$ is a probability that no less than N_c bridges are formed during the time T_υ.

As the frequency increases, the probability $\tilde{P}(N_c, T_\upsilon)$ decreases rapidly and $\langle t_{lag} \rangle$ increases. Dashed curves in Fig. 15.6c present results of analytical calculations of $\langle t_{lag} \rangle$ in (15.10), which agree qualitatively with the results of numerical simulations discussed above.

Considering that in the presence of oscillatory modulations the force series represent the set of alternating segments of stick-slip oscillations and low friction sliding (see Figs. 15.5 and 15.6), the average friction force can be estimated as

$$\langle F \rangle = \frac{\langle F_{s-s} \rangle \langle t_{s-s} \rangle}{\langle t_{s-s} \rangle + \langle t_{lag} \rangle} \tag{15.11}$$

where $\langle t_{s-s} \rangle / \left(\langle t_{s-s} \rangle + \langle t_{lag} \rangle \right)$ is a fraction of time corresponding to the stick-slip state of motion, $\langle F_{s-s} \rangle$ is the average force experienced by the tip in that state, and the contribution of the low friction sliding was neglected. Approximating by $\langle F_{s-s} \rangle \cong \frac{K_d V_d \langle t_{s-s} \rangle}{2}$, we get the following equation for the average friction force

$$\langle F \rangle = \frac{2 \langle F_{s-s} \rangle^2}{2 \langle F_{s-s} \rangle + K_d V_d \langle t_{lag} \rangle} \tag{15.12}$$

With increase of A or/and υ the length of low friction segments $\langle t_{lag} \rangle$ grows and the friction force decreases rapidly. Dashed curves in Fig. 15.7a, b show $\langle F \rangle$ as functions of frequency and amplitude of modulations which have been calculated according to the (15.12). The analytical results are in qualitative agreement with numerical simulations. It should be noted that the mean field description given by

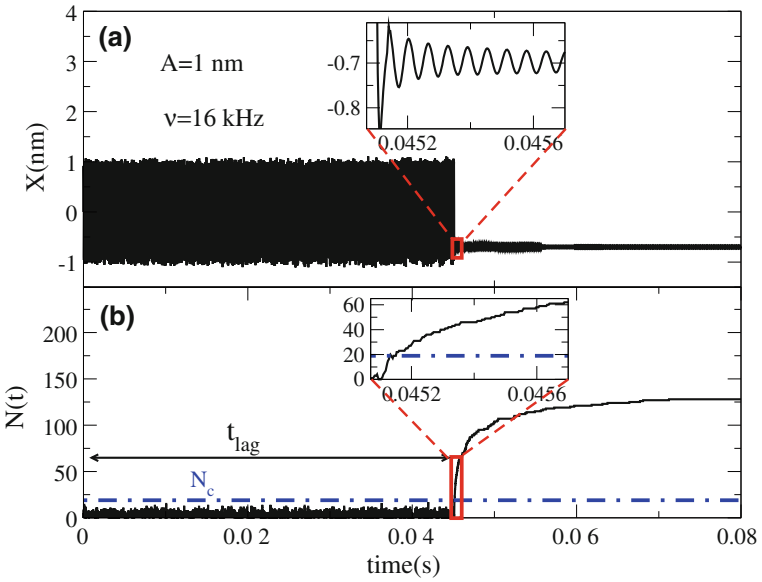


Fig. 15.8 **a** Displacement of the tip, X(t), as a function of time in response to harmonic driving. The tip is brought in contact with a substrate at t = 0. **b** Dynamics of bridge formation, N(t). *Dashed dotted line* in the panel (**b**) indicates the maximal number of junctions, N_c, which can be ruptured by the oscillatory modulations with the amplitude A = 1.0 nm. The insets zoom in the region corresponding to the transition from the low to high friction state

(15.5), (15.8)–(15.12) assumes that bridges are ruptured simultaneously under the action of the pulling force and $\langle F_{s-s} \rangle$ is independent on A. These assumptions are inaccurate for high amplitudes and frequencies of oscillation, and in this range of parameters the analytical results deviate essentially from the numerical ones (see Fig. 15.7a, b).

Our simulations suggest that applying small-amplitude inplane oscillations to the stage of AFM one can give important information on the kinetics of frictional aging and the stiffness of bridges. In order to do this we propose to bring the oscillating tip in contact with the surface and to follow the time variation of the amplitude of the tip oscillations, A_{tip}. Figure 15.8 shows the tip motion and the kinetics of capillary bridges formation for the tip that is driven at the velocity $V_d = -2\pi\nu A \sin(2\pi\nu t)$ and brought in the contact at t = 0. The tip exhibits high amplitude oscillations for the time interval $t < t_{lag}$ during which the formation of bridges is suppressed by oscillations. In this regime the amplitude of tip oscillations, A_{tip}, is only slightly below the driving amplitude A = 1 nm, as shown in Fig. 15.8a. When the number of bridges formed during a half-period of oscillations exceeds the critical one, N_c, (inset in Fig. 15.8b) the cluster of condensed bridges starts to grow and the amplitude, A_{tip}, is greatly reduced as shown in Fig. 15.8a. Considering the balance of forces acting on the tip, $A_{tip}(t)\, N(t)\kappa = \left(A - A_{tip}(t)\right) K_d$, the time variation of the amplitude of tip

oscillations can be related to the time-dependent stiffness of the cluster of condensed bridges,

$$N(t)\kappa = K_d\left(\frac{A}{A_{tip}(t)} - 1\right) \tag{15.13}$$

Thus the proposed measurements can provide direct information on the kinetics of frictional aging and the stiffness of bridges. Additional information on the distribution of heights of barriers for bridge formation can be obtained comparing the measured values of t_{tag} with the results of calculations according (15.10). This comparison allows to estimate the main parameters of the distribution of the barrier heights, such as the minimal barrier height, $\Delta E_{on}^{\min}$, and the density, S_E.

15.4.1 Summary

Capillary bridges play a crucial role in the operation of atomic force microscopy under humid ambient conditions and their formation often dominates the measured forces. It has been found that under ambient conditions the nanoscopic friction force decreases with pulling velocities and increases with temperature [21–23]. These observations differ significantly from the results of friction experiments carried out under ultrahigh vacuum, and disagree with predictions of thermal Prandtl-Tomlinson model of friction. Here, by analytical model and numerical simulations, we demonstrate that the observed friction phenomena are caused by the fact that formation of capillary bridges is a thermally activated process, while kinetics of their rupture is temperature-independent, or only slightly influenced by T. Our calculations show that measurements of velocity and temperature dependencies of friction force do not allow to draw definite conclusions on contribution of ageing processes to friction. However, we found that this information can be provided by slide-hold-slide measurements.

We demonstrated that adding a low amplitude oscillatory component to the pulling force, when applied at the right frequency, can significantly suppress formation of capillary bridges and thereby reduce friction. Our simulations suggest that applying small-amplitude inplane oscillations to the stage of AFM one can get direct information on the kinetics of frictional aging and the stiffness of capillary bridges.

Acknowledgments We grateful to R. W. Carpick, A.E. Filippov, C. Greiner, P.-E. Mazeran and O. Noel for helpful discussions. R.C. acknowledges support from the Swiss National Science Foundation SINERGIA Project CRSII2 136287\1. The work was supported by DIP (German-Israeli Project Cooperation Program) and the Israel Science Foundation (1109/09).

References

1. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, The nonlinear nature of friction. Nature **430**, 525–528 (2004)
2. V. Bormuth, V. Varga, J. Howard, E. Schaffer, Protein Friction Limits Diffusive and Directed Movements of Kinesin Motors on Microtubules. Science **325**, 870–873 (2009)
3. C.H. Scholz, Earthquakes and friction laws. Nature **391**, 37–42 (1998)
4. R. Budakian, S.J. Putterman, Correlation between charge transfer and stick-slip friction at a metalinsulator interface. Phys. Rev. Lett. **85**, 1000 (2000)
5. E. Gerde, M. Marder, Friction and fracture. Nature (London) **413**, 285 (2001)
6. A.E. Filippov, J. Klafter, M. Urbakh, Friction through dynamical formation and rupture of molecular bonds. Phys. Rev. Lett. **92**, 135503 (2004)
7. S.M. Rubinstein, G. Cohen, J. Fineberg, Detachment fronts and the onset of dynamic friction. Nature **430**, 1005 (2004)
8. A. Vanossi, N. Manini, M. Urbakh, S. Zapperi, E. Tosatti, Colloquium: modeling friction: from nanoscale to mesoscale. Rev. Mod. Phys. **85**, 529 (2013)
9. Y. Mo, K.T. Turner, I. Szlufarska, Friction laws at the nanoscale. Nature **457**, 1116–1119 (2009)
10. B. Gotsmann, M.A. Lantz, Quantized thermal transport across contacts of rough surfaces. Nature Mater. **12**, 59–65 (2012)
11. B.N.J. Persson, *Sliding Friction: Physical Principles and Applications* (Springer, Berlin, 1998)
12. O.M. Braun, M. Peyrard, Modeling friction on a mesoscale: master equation for the earthquakelike model. Phys. Rev. Lett. **100**, 125501 (2008)
13. O.M. Braun, I. Barel, M. Urbakh, Dynamics of transition from static to kinetic friction. Phys. Rev. Lett. **103**, 194301 (2009)
14. I. Barel, M. Urbakh, L. Jansen, A. Schirmeisen, Multibond dynamics of nanoscale friction: the role of temperature. Phys. Rev. Lett. **104**, 066104 (2010)
15. Y. Liu, I. Szlufarska, Chemical origins of frictional aging. Phys. Rev. Lett. **109**, 186102 (2012)
16. R. Capozza, M. Urbakh, Static friction and the dynamics of interfacial rupture. Phys. Rev. B **86**, 085430 (2012)
17. T.C. Halsey, A.J. Levine, How sandcastles fall. Phys. Rev. Lett. **80**, 3141 (1998)
18. S.N. Gorb, *Attachment Devices of Insect Cuticle* (Kluwer Academic Publishers, Dordrecht, 2001)
19. B. Bhushan, *Handbook of Nanotribology* (Springer, New York, 2007)
20. L. Bocquet, E. Charlaix, S. Ciliberto, J. Crassous, Moisture-induced ageing in Granular media and the kinetics of capillary condensation. Nature (London) **396**, 735 (1998)
21. E. Riedo, F. Le'vy, H. Brune, Kinetics of capillary condensation in nanoscopic sliding friction. Phys. Rev. Lett. **88**, 185505 (2002)
22. R. Szoszkiewicz, E. Riedo, Nucleation time of nanoscalewater bridges. Phys. Rev. Lett. **95**, 135502 (2005)
23. C. Greiner, J.R. Felts, Z. Dai, W.P. King, R.W. Carpick, Local nanoscale heating modulates single-asperity friction. Nano Lett. **10**, 4640 (2010)
24. O. Noel, P.-E. Mazeran, H. Nasrallah, Sliding velocity dependence of adhesion in a nanometer-sized contact. Phys. Rev. Lett. **108**, 015503 (2012)
25. L. Zitzler, S. Herminghaus, F. Mugele, Capillary forces in tapping mode atomic force microscopy. Phys. Rev. B **66**, 155436 (2002)
26. Y. Sang, M. Dube, M. Grant, Thermal effects on atomic friction. Phys. Rev. Lett. **87**, 17430 (2001)
27. O.K. Dudko, A.E. Filippov, J. Klafter, M. Urbakh, Dynamic force spectroscopy: a Fokker−Planck approach. Chem. Phys. Lett. **352**, 499 (2002)
28. I. Szlufarska, M. Chandross, R.W. Carpick, Recent advances in single-asperity nanotribology. J. Phys. D **41**, 123001 (2008)
29. E. Gnecco, R. Bennewitz, T. Gyalog, C. Loppacher, M. Bammerlin, E. Meyer, H.J. Guntherodt, Velocity dependence of atomic friction. Phys. Rev. Lett. **84**, 1172–1175 (2000)

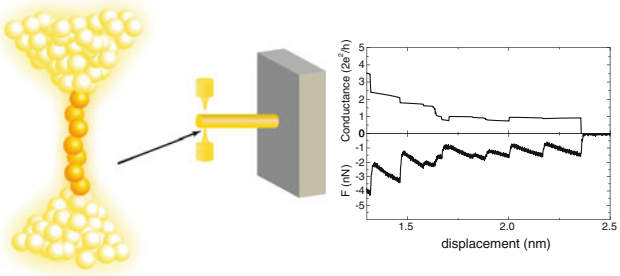
30. L. Jansen, H. Holscher, H. Fuchs, A. Schirmeisen, Temperature dependence of atomic-scale stick-slip friction. Phys. Rev. Lett. **104**, 256101 (2010)
31. M. Heuberger, C. Drummond, J.N. Israelachvili, Coupling of normal and transverse motion during frictional sliding. J. Phys. Chem. B **102**, 5038 (1998)
32. A. Socoliuc et al., Atomic-scale control of friction by actuation of nanometer- sized contacts. Science **313**, 207 (2006)
33. S. Jeon, T. Thundat, Y. Braiman, Effect of normal vibration on friction in the atomic force microscopy experiment. Appl. Phys. Lett. **88**, 214102 (2006)
34. A. Cochard, L. Bureau, T. Baumberger, Stabilization of frictional sliding by normal load modulation: a bifurcation analysis. Trans. ASME **70**, 220 (2003)
35. V.L. Popov, J. Starcevic, A.E. Filippov, Influence of ultrasonic in-plane oscillations on static and sliding friction and intrinsic length scale of dry fiction. Tribol. Lett. **39**, 25 (2010)
36. R. Capozza, S.M. Rubinstein, I. Barel, M. Urbakh, J. Fineberg, Stabilizing stick-slip friction. Phys. Rev. Lett. **107**, 024301 (2011)
37. M.G. Rozman, M. Urbakh, J. Klafter, Controlling chaotic frictional forces. Phys. Rev. E **57**, 7340 (1998)
38. Z. Tshiprut, A.E. Filippov, M. Urbakh, Tuning diffusion and friction in microscopic contacts by mechanical excitations. Phys. Rev. Lett. **95**, 016101 (2005)
39. J.P. Gao, W.D. Luedtke, U. Landman, Friction control in thin-film lubrication. J. Phys. Chem. B **102**, 5033–5037 (1998)
40. R. Capozza, A. Vanossi, A. Vezzani, S. Zapperi, Suppression of friction by mechanical vibrations. Phys. Rev. Lett. **103**, 085502 (2009)
41. Q. Li, T.E. Tullis, D. Golldsby, R.W. Carpick, On the origins of rate and state friction: frictional ageing from interfacial bonding. Nature (London) **480**, 233 (2011)
42. I. Barel, A.E. Filippov, M. Urbakh, Formation and rupture of capillary bridges in atomic scale friction. J. Chem. Phys. **137**, 164706 (2012)
43. H. Choe, M.-H. Hong, Y. Seo, K. Lee, G. Kim, Y. Cho, J. Ihm, W. Jhe, Formation, manipulation, and elasticity measurement of a nanometric column of water molecules. Phys. Rev. Lett. **95**, 187801 (2005)
44. M. He, A.S. Blum, D.E. Aston, C. Buenviaje, R.M. Overney, R. Luginbuhl, Critical phenomena of water bridges in nanoasperity contacts. J. Chem. Phys. **114**, 1355 (2001)
45. J. Crassous, M. Ciccotti, E. Charlaix, Capillary, force between wetted nanometric contacts and its application to atomic force microscopy. Langmuir **27**, 3468 (2011)
46. H.-J. Butt, Capillary forces: influence of roughness and heterogeneity. Langmuir **24**, 4715 (2008)
47. I. Barel, M. Urbakh, L. Jansen, A. Schirmeisen, Temperature dependence of friction at the nanoscale: when the unexpected turns normal. Trib. Lett. **39**, 311 (2010)
48. I. Barel, M. Urbakh, L. Jansen, A. Schirmeisen, Unexpected temperature and velocity dependencies of atomic-scale stick-slip friction. Phys. Rev. B **84**, 115417 (2011)
49. J.H. Dieterich, Modeling of rock friction: 1. experimental results and constitutive equations. J. Geophys. Res. **84**, 2161 (1979)
50. V.L. Popov, *Contact Mechanics and Friction: Physical Principles and Applications* (Springer, Berlin, 2010)
51. K.M. Frye, C. Marone, Effect of humidity on granular friction at room temperature. J. Geophys. **107**, 2309 (2002)
52. R. Capozza, I. Barel, M. Urbakh, Probing and tuning frictional aging at the Nanoscale. Sci. Rep. **3**, 1896 (2013)
53. X.H. Chen, A.P. Dempster, J.S. Liu, Weighted finite population sampling to maximize entropy. Biometrika **81**(3), 457 (1994)

Part IV
Nanomanipulation

Chapter 16
Mechanical Properties of Metallic Nanocontacts

Gabino Rubio-Bollinger, Juan J. Riquelme, Sebastian Vieira and Nicolas Agraït

Abstract The mechanical properties of the reduced number of atoms forming the apex of a tip are interesting both from a fundamental point of view and for the interpretation of experiments related to scanning local probe methods. These mechanical properties can be studied by establishing a very small contact, a nanocontact, between a tip and a surface. The elasticity and fracture events during the controlled breaking of a nanocontact as the tip is separated from the surface provide information about the mechanical properties of the tip apex. In the case of metallic tips, electron transport through the nanocontact also provides information on its mechanical properties, because at the scale of a few atoms forming the nanocontact the mechanical and electron transport properties are strongly related.



Abbreviations

AFM	Atomic force microscope
DFT	Density functional theory

G. Rubio-Bollinger (✉) · J.J. Riquelme · S.Vieira · N. Agraït
Departamento de Física de la Materia Condensada C-III, Universidad Autónoma de Madrid, 28049 Madrid, Spain
e-mail: gabino.rubio@uam.es

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_16

DOS	Density of states
MCBJ	Mechanically controlled break-junction
MD	Molecular dynamics
MFM	Magnetic force microscopy
PC	Point contact
PCS	Point contact spectroscopy
SNOM	Scanning near-field optical microscopy
SPM	Scanning probe microscopy
STM	Scanning tunneling microscope
TF	Tuning fork
UHV	Ultra high vacuum

16.1 Introduction

Local probe techniques are often based on the interaction of a tip in close proximity with a surface. Highest imaging resolution is achieved when the tip apex is brought as close as possible to the surface but avoiding direct contact, so that neither the tip apex nor the surface are modified or damaged during scanning. The transition from the non-contact to the contact regime plays therefore a crucial role in the achievement of high resolution images and their interpretation. In most of the cases this transition in not smooth but there is a mechanical instability which results in a jump to contact process [1, 2]. The instability occurs at a probe to sample distance at which the elasticity of the probe becomes lower than the force gradient of the interaction between the tip and the surface. There are two main contributions to the elasticity of the probe. The first, which could be named extrinsic, has to be considered if the tip is mounted on an elastic device, such as the cantilever beam of an atomic force microscope. Second, a non-negligible intrinsic contribution has its origin in the finite elasticity of the apex of tip and the spot on the surface close to the tip. While the extrinsic elasticity is macroscopic in nature, the intrinsic elasticity is at the nanoscale because it involves very few atoms located at the tip apex and its surroundings. These nanoscale mechanical properties are the subject of this chapter. Some scanning probe methods are used to study friction and wear between a tip and a surface. In such a situation, an atomic scale contact is established between the tip and the surface. The mechanical properties of this kind of structure, a nanocontact, are different from that of macroscopic contacts. The detailed arrangement of the atoms in the contact can result in a variety of mechanical behaviors and quantum effects become relevant [3, 4]. At the very smallest contact, a one-atom contact, electron transport and mechanical properties have been shown to be intimately related. Theoretically, molecular dynamics [5] and density functional theory calculations have provided deeper insight into the behavior of matter at the nanoscale. Despite the apparent simplicity of structures composed of a small number of atoms, theoretical modeling still does not provide a full quantitative way of finding the mechanical properties of atomic-size structures, and further research is still necessary.

In addition, there is an increasing interest in techniques that allow establishing a reliable interface between macroscopic or mesoscopic electric circuits, and atomic-sized structures tailored to have functional and practical mechanical or electron transport properties. That is the case of functional molecules and carbon based electronic devices, such as carbon nanotubes or graphene nanoribbons [6]. These tiny, but subtle, structures will probably be interfaced with (scanning or static) atomic-sized tips whose mechanical and electron transport properties will play a significant role in the functionality and performance of these nanoscale devices.

16.2 Experimental Tools

We will focus in this section on devices which are especially well suited for the study of metallic nanocontacts at low temperatures. Despite many successful low temperature scanning tunneling microscopy experiments the number of results obtained by low temperature scanning force microscopy is still small. Moreover, only a few commercial manufacturers have presented a low temperature scanning force microscope so far [7–9] because of the difficulties to adapt the usual force detection techniques to cryogenic environments.

Binnig and Rohrer [10] developed the Scanning Tunneling Microscope (STM) more than 20 years ago. The STM has allowed studying the topography and electronic properties of metallic surfaces with atomic resolution. Moreover, the STM is the first of a family of local probes, being the Atomic Force Microscope (AFM) the most important [11]. Soon after its invention, Gimzewski and Möller employed the STM to fabricate a nano-sized contact [2]. They measured the electrical resistance while approached from the tunneling regime and touching gently the surface with the STM tip. Then a nano-sized protuberance was observed. Subsequently, Dürig et al. [1] investigated the adhesion forces using an AFM.

Local elastic properties are studied by both static and dynamic AFM methods. The static methods include force-distance curves, while the dynamic techniques track the frequency shift of a resonator versus the distance. In order to measure both mechanical and electrical properties the microscope has to be supplemented with a force sensor. Conventional AFMs with conductive tips are usually not well suited for these experiments because the cantilever elastic constant is too low ($\leq 1\ \mathrm{N\ m^{-1}}$) to be able to fabricate and manipulate metallic nanocontacts, whose effective elastic constant is above $10\ \mathrm{N\ m^{-1}}$ [12–14]. Therefore, modified versions of combined AFM/STM techniques are required [4], using bending beams with elastic constants above $100\ \mathrm{N\ m^{-1}}$.

There are different devices to detect the deflection of the force sensor. The most commonly used method in commercial AFMs measures the deflection of a laser beam on the backside of the cantilever using a photodiode [15]. Interferometric detection is a very precise technique that detects the motion of the cantilever using the interference of light reflected from the cantilever with that reflected from the end of a fiber, which is mounted very close to the backside of the cantilever [16, 17].

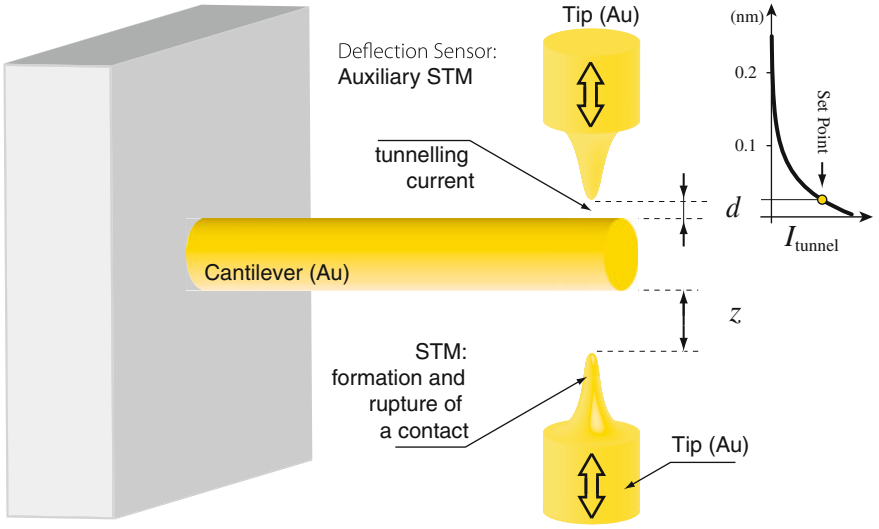


Fig. 16.1 Sketch of a STM supplemented with a force sensor. The cantilever deflection is followed by an auxiliary STM working in constant current mode, providing an extremely high resolution measurement of the deflection of the cantilever

Piezoresistive methods measure the change of the resistance path on the backside of the cantilever and extract its deflection [18]. Furthermore, a STM can also be used not only as a tunneling current probe but also for displacement sensing applications, see Fig. 16.1. It was used by Binnig, Quate and Gerber in the first AFM [11]. A tunneling tip on the conductive backside of a cantilever is used to measure deflection. Such a sensor features extreme sensitivity for the detection of displacement, given that the current decays exponentially over a decay length of 44 pm for an apparent tunneling barrier height of 5 eV. Among the possible force detection methods cited above, we present an implementation that uses the STM as a force sensor, because it enables us to study the mechanical properties of metallic nanocontacts at low temperatures with extremely high resolution in the cantilever deflection measurement.

In addition to the static cantilever deflection measurement, it is possible to measure the force gradient if the cantilever is forced to oscillate. The dynamic methods use some signal derived from the force between tip and sample as a feedback parameter to track the topography of a surface (for a review, see for example [19, 20]). The detection devices listed above are also used in these dynamic methods. Besides, piezoelectric detectors based on quartz tuning forks (TF) are also used as sensors [21, 22]. The TFs feature high stiffness and a high quality factor. The first feature maintains the stability of the microscope, while the latter allows precise tracking of the resonance frequency. This high quality factor is a consequence of low coupling between the preferred vibrational mode of the tuning fork where both legs move in the same plane and in opposite phase, and other modes, leading to a high quality factor. In contrast to micromachined silicon cantilevers, the quartz forks are large enough to

permit attachment to one prong of a wide variety of conductors, thus enabling the use of a bulk metallic tip. An additional advantage of piezoelectric sensors when used for low-temperature applications is that the electric dissipation is negligible when compared with laser beam bounce, interferometric or piezoresistive methods. All this goes to show that such a sensor is a very suitable tool to explore the mechanical properties of metallic nanocontacts at low temperature.

16.2.1 The Scanning Tunneling Microscope Supplemented with a Force Sensor

Although very accurate, there are some important issues to be considered when using the STM as a deflection sensor, since the position of the tunneling tip has to be maintained within fractions of an Angstrom during operation. First, in order to measure the deflection of the cantilever with subatomic accuracy, one needs to keep the tip positioned over the same atom. Thus, it is necessary consider the diffusion of atoms over surfaces since it may mislead the measurement of the cantilever deflection. Second, the STM must be carefully designed to avoid external disturbances in the tunnel current signal. This is best accomplished in those designs that results in a rigid, compact STM. A rigid STM with a high resonance frequency helps to attenuate the external low frequency excitations while a compact design enables to suspend the STM from springs. Third, another issue in SPM comes from the rather complicated relation between the mechanical deformation of the piezoelectric ceramic material used for moving the tip over the sample and the applied voltage. One may consider that the piezoelectric actuator deformation is proportional to the applied voltage. However, a hysteretic behavior may considerably modify this proportional dependence. This behavior arises from piezoelectric creep. Fourth, differential thermal contractions of the mechanical parts of the microscope can result in small but unavoidable displacement drift. Modern SPMs overcome these problems by using different correction strategies implemented as closed loop feedback systems.

These issues are naturally overcome when using a setup in a cryogenic environment, because at liquid helium temperature (4.2 K) all thermally activated processes are hampered. Additionally, as the sensor uses the tunneling current as the control signal for the feedback loop, it is crucial to prevent the presence of adsorbates, contamination, and oxides both in the tip and in the backside of the cantilever. When operated at low temperature the STM is in a cryogenic vacuum environment that minimizes the exposition of fresh surfaces to contaminants. Furthermore, the piezoelectric actuator's nonlinear behavior is strongly temperature dependent. As a result, these nonlinearities are strongly reduced when the microscope is operated at low temperatures.

However, there are some remarkable technical difficulties at low-temperature. Usually, materials with similar thermal expansion coefficients are chosen in order to minimize undesired stress that may lead to fracture of a STM part. The coarse

positioning system of the STM needs some care when it is designed to work at low temperatures. There are several designs of piezoelectric inertial motors [23–25] or friction motors [26–28] in STMs at low temperatures. A design that follows the one by Pan [28] is used here because is has been found to be highly reliable at low temperature while allowing a compact STM design.

In order to reduce the generation of mechanical noise it is advisable to use a superinsulated cryostat that does not require liquid N_2 for thermal shielding. A compact, low mass STM design and a low heat loss setup minimize the liquid He consumption and thus the bubbling noise. Moreover, the high mechanical stiffness of the dewar vessel and the insert does not amplify mechanic and acoustic vibrations. In addition to that, acoustic disturbances are amenable to being attenuated using foam adhered to its enclosure. These cautions are important because at low temperatures there are few possibilities to use a viscoelastic damping system.

16.2.2 The Mechanically Controllable Break-Junction Technique

The mechanically controllable break-junction technique (MCBJ) has its origin in the work done by Moreland and Ekin [29], but the configuration outlined below resembles more closely the one developed by Muller et al. [30]. It has been widely used during the last decade to study electron transport through metallic nanocontacts, and has been recently supplemented with a tuning fork force sensor in order to simultaneously study their mechanical properties [12, 31].

The MCBJ is depicted in Fig. 16.2. It consists of a conducting thin wire (diameter 0.01–0.25 mm) of the material that one wants to investigate, which is fixed at two closely spaced spots on top of a flexible insulated substrate. In order to take care that the wire will break it has a notch at the position between the two spots. The notch will decrease the force needed to break the wire and can be fabricated straightforwardly

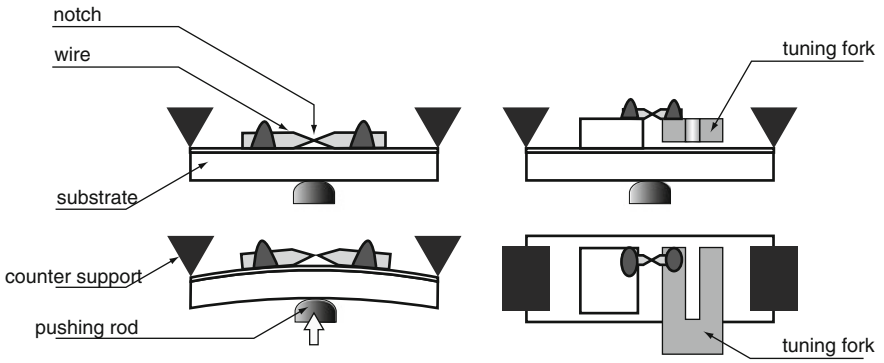


Fig. 16.2 Schematic *top* and *side view* of a MCBJ (*left*) and a MCBJ supplemented with a tuning fork (TF) resonator. The TF resonance frequency is sensitive to the force gradient between the tips

in the case of ductile materials using a knife. The substrate is mounted in a three-point bending configuration, where either the central or the counter supports can be displaced resulting in controlled bending of the substrate. This bending causes the top surface of the substrate to expand, resulting in the elongation of the weakest part of the wire at the notch until rupture. In this way, two clean fracture surfaces are exposed. A contact between the fractured surfaces can be reestablished by relaxing the bending force on the substrate, hence the name 'break junction'. A combination of micrometric screws and piezoelectric actuators are commonly used for fine control of the opening at the notch.

One main advantage of the MCBJ technique is that the freshly exposed surfaces are free of contamination, and can be kept clean if the device is kept in UHV or cryogenic vacuum. A second advantage of the MCBJ is the mechanical stability of the two electrodes with respect to each other that results from the short mechanical loop between electrodes. That is, the distance between the two fixed points of the wire at both sides of the notch can be as small as 0.1 mm. Lithographically fabricated MCBJs present an outstanding stability since the interelectrode distance is reduced to $\sim 1\,\mu$m [32]. On the contrary, STM has scanning capabilities and permits the fabrication of contacts between two different metals, two features that are not easily incorporated in MCBJ implementations.

The MCBJ supplemented with a force sensor, based on a microfabricated quartz tuning fork (TF) resonator, enables simultaneous current and force gradient measurements [12, 31]. The frequency shift of the TF can be accurately tracked using a frequency modulation technique [33]. Tuning forks have recently been introduced in different scanning probe microscopy setups, such as AFM [21, 34–36], Magnetic Force Microscopy (MFM) [37] and Scanning Near-field Optical Microscopy (SNOM) [38, 39]. It has very good properties as a resonator because its high quality factor leads to high frequency resolution, while its high stiffness avoids jumps to contact [40] and anharmonic vibration.

A description of a MCBJ supplemented with a tuning fork follows, as shown in Fig. 16.2. Commercially available tuning forks (TF) are inside of a metal cap. This cap is partially removed leaving only a little ring around the base. The ring is then soldered onto a substrate, so that the prongs of the fork are freely standing with the preferred oscillation mode parallel to the substrate (see Fig. 16.2). Next to the tuning fork, a little metal block is mounted on the substrate, such that the top of the block and the top of the tuning fork are at the same height. A thin metal wire with is soldered on both one prong and the block with a notch in between the prong and the block.

In the experiments by Rubio-Bollinger et al. [12] the TF electrodes where used for both the excitation and the resonance detection. Valkering et al. [31] have used an external excitation source based on a magnetic actuator. This resonant frequency is tracked by implementing a phase locked loop oscillator [12, 20, 39, 41, 42]. The shift of the resonance frequency Δf of the TF excited with vanishing amplitude is proportional to the force gradient or stiffness k of the interaction between the electrodes: the measured resonance frequency $f_0 + \Delta f$ is given by

$$f_0 + \Delta f = \frac{1}{2\pi}\sqrt{\frac{k_{\text{lever}} + k}{m_{\text{eff}}}} \tag{16.1}$$

with m_{eff} the effective mass in resonance. Since the factor $\alpha = \Delta f/k$ depends on the mechanical properties of the TF, a high spring constant k_{lever} reduces the sensitivity of the force derivative measurement.

Compared with previous measurements of atomic contact forces using atomic force microscope cantilevers, the use of a TF has a series of advantages. First, the stiffness (10^3 N m^{-1} [21, 42]) of the lever supporting the contact is high enough to prevent early jump to contact. Second, TF has a very high quality factor [21] (larger than 10^4 in the experimental configuration in vacuum at low temperature [36, 42]), permitting a precise measurement of the frequency shift. Moreover, it allows for a low power electrical measurement of the resonance, which is particularly convenient for low temperature experiments. In addition, it is possible to use a TF to form and break the contact. This avoids the usual uncertainties in the calibration of the interelectrode distance in MCBJ experiments. One can control the interelectrode distance using the piezoelectric effect of the TF itself by applying a dc voltage in addition to the ac drive. This piezoelectric displacement can be calibrated at room temperature. This calibration does not change at low temperature because the TF is fabricated from a single crystal of quartz.

To sum up, this setup allows us to make use of the high mechanical stability of the MCBJ while the tuning fork has a high spring constant 10,000 N m^{-1}, which in combination with a high quality factor, permits detection of very small frequency shifts (smaller than 100 mHz) and at small vibration amplitudes (smaller than 10 pm).

16.3 Electron Transport Through Metallic Nanocontacts

It is possible to extract information on the contact mechanical properties from magnitudes related with the electronic transport, because at the scale of a few atoms forming the nanocontact these mechanical and electronic transport properties are strongly related. Mostly, these electronic transport properties are determined by only a few atoms and thus provide rich information related to the size, shape, state of strain and elastic deformation of metallic nanocontacts [3].

In metallic contacts the electronic transport description changes as their dimensions are reduced down to the nanoscale. Macroscopic conductors are characterized by Ohm's law, which establishes that the conductance G of a given sample is directly proportional to its transverse area S and inversely proportional to its length D, i.e.

$$G = \sigma S/D, \tag{16.2}$$

where σ is the electrical conductivity of the sample.

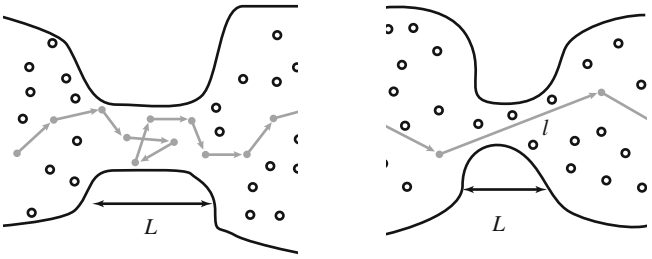


Fig. 16.3 Scheme of the electronic transport in metallic conductors: diffusive (*left*) and ballistic regimes

As the typical length scale of our sample L is reduced, one can identify different transport regimes. These regimes are defined by the relative size of various length scales. The differences between these regimes will become clear as we move through them one by one. In metallic point contacts an important length scale is related to the phase-coherence length L_ϕ, that measures the distance over which quantum coherence is preserved. As long as $L \gg L_\phi$ the contact is in the macroscopic regime, whereas if $L \ll L_\phi$ the contact is in the so-called mesoscopic regime. Another relevant scale is associated with the elastic mean free path l which defines the change from the diffusive regime if $L \gg l$ to the ballistic regime if $L \ll l$ (see Fig. 16.3).

In the classical limit, the conductance of a constriction was calculated by Maxwell for a constriction of hyperbolic geometry [43], and is proportional to the constriction radius a and the conductivity $G \propto a\sigma$. When the dimensions of a contact are much smaller than their mean free path, the electrons will pass through ballistically. In such contacts, the large potential gradient near the contact accelerates the electrons within a short distance. The conduction through this type of contacts resembles the effusion of a dilute gas [44]. Sharvin first solved this problem using a semiclassical approximation [45], providing an expression of the conductance for a pure ballistic contact, the so-called Sharvin's conductance

$$G = \frac{2e^2}{h}\left(\frac{k_\mathrm{F}a}{2}\right)^2, \tag{16.3}$$

where h is the Planck constant, e is the electron charge, k_F is the Fermi wave vector, and a is the contact radius. The Sharvin's conductance depends on the material only through k_F, and it is totally independent of conductivity and the mean free path, contrary to the conductance in the diffusive regime. Sharvin's formula may be used in ballistic contacts to estimate the area of minimal cross-section from the conductance. The quantity $2e^2/h$ is called the conductance quantum G_0. Corrections to this formula where given by Torres, Pascual, and Saenz, using an exact quantum calculation for a circular cross-section [46].

The ultimate atomic contact has one atom between two electrodes. In this case, a fully quantum description is needed because the quantum effects are increasingly important as the size of the constriction is comparable to the Fermi wavelength of the conduction band electrons. It has been shown that the conductance of a one atom contact depends dramatically on its chemical nature [47].

16.4 Mechanical Properties of Metallic Nanocontacts

Nanocontacts are structures where the number of atoms in the minimum cross-section of the constriction ranges from tens of atoms to a single atom. Although there are several ways or preparing such structures [48, 49], the experiments at low temperature usually involve the use of probes related to STM or MCBJ. This section is dedicated to the fabrication procedure of nanocontacts between gold electrodes, whose electronic and mechanical properties have been extensively studied in the past decade.

16.4.1 Fabrication of Metallic Nanocontacts

Single atomic-sized contacts between metals can be produced with a scanning tunneling microscope. In its constant current operation mode the tip is scanned over the sample surface without making contact. The tip-sample separation is kept constant by controlling the current that flows between tip and sample, due to the tunneling effect, when a fixed bias voltage is applied between them. Typical currents are $\sim$1 nA, corresponding to a tunneling junction resistance in the G Ω range, for tip-sample voltages $\sim$1 V and the tip-sample distances $\sim$0.1 nm.

However, the STM was soon used to modify the sample surface on a nanometer scale. In the experiment by Gimzewski and Möller [2] the surface was gently touched with the tip and the transition from the tunneling regime to metallic contact was observed as an abrupt jump in the current. This jump in current was due to a change in resistance of about 13 kΩ. According to Sharvin's formula, a conductance of 1 G_0 corresponds to a contact diameter of 0.25 nm, suggesting a mechanical jump to contact forming single atom bridge between the tip and the surface.

After the jump to contact, the indentation continues and the conductance increases showing a characteristic staircase pattern. In the case of gold subsequent retraction of the tip results in breaking the contact: the neck between the two electrodes gets thinner as the sense of the tip motion is reversed. Again, the conductance has intervals in which it is relatively constant (plateaus) separated by jumps. The last plateau, in the case of gold, shows a quite well defined conductance with a value of approximately of $1G_0$. For other metals the conductance during the indentation cycle will look somewhat different, depending on the electronic structure of the metal [47] but the stepwise behavior of the conductance is still observed. However, it has been reported that in some cases the jump does not occur. Instead, there is a continuous increase

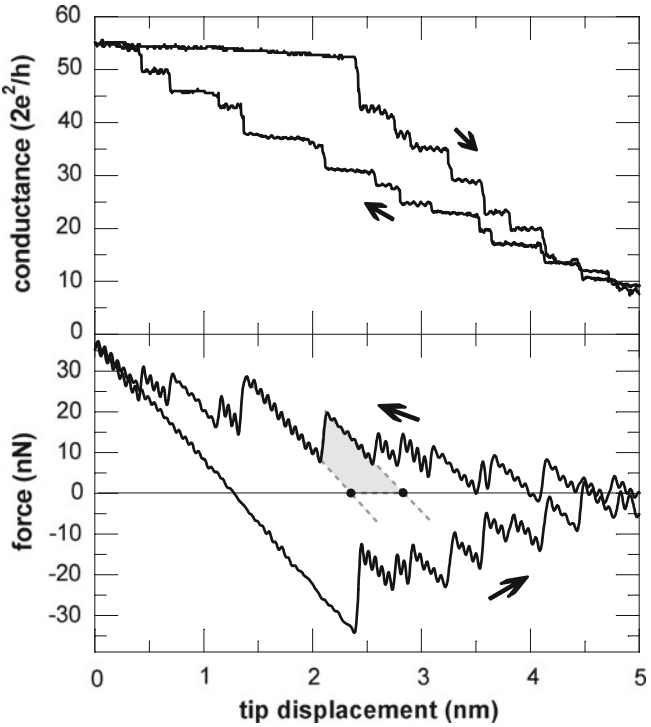


Fig. 16.4 Simultaneous measurement of conductance and force during an indentation cycle, for Au at 300 K. The *arrows* indicate compression, (positive forces) and elongation of the nanocontact. Reprinted figure with permission from [52]. Copyright (1997) by the American Physical Society

of current from tunneling to contact: it occurs for some metals, in particular for Ni, W, and Ir [3].

Figure 16.4 shows an indentation cycle in gold without breaking the contact. The bias voltage is 10 mV using a STM supplemented with a force sensor. As the contact size increases, the conductance increases showing a step like behavior. Nevertheless, the measured curve shows a force dependence on the distance that looks like saw-tooth signal. There is a sequence of linear stages separated by sudden relaxations. The simultaneously measured conductance remains almost constant in the elastic stages, and there is an abrupt conductance change when the force relaxes. Landman et al. [5] predicted this behavior in early MD simulations. This yielding results in a sudden change of the minimal cross-section due to the atomic rearrangement. These processes results in a sharp jump in the contact conductance. The relaxation in the tensile force is the result of a change in the length of the constriction. This change of length is the distance between successive relaxed configurations in the nanocontact separated by a yielding event, marked as two black dots in Fig. 16.4. In gold nanocontacts, this plastic relaxation length ranges from 0.2 to 1 nm, which implies that only a few atomic layers participate in the plastic deformation process.

This dependence of both conductance and force versus the electrode separation is unique since the arrangements are different for every indentation cycle.

16.4.2 Elasticity and Fracture of Metallic Nanocontacts

The simultaneous measurement of conductance and tensile force during the indentation cycle described above provides information about the different mechanical behavior of macroscopic specimens and metallic nanocontacts. Macroscopic objects change their shape when they are deformed by a stress. As long as they regain their original shape when the load that causes the stress is removed, they are subjected to an elastic deformation. In the elastic regime, the stress applied to an isotropic material is linearly related to the strain by the Young modulus and the Poisson's ratio by the generalized Hooke's law (see for example [50]). Consider the simple case of a rod of length L_z and diameter b subjected to a stress through its longitudinal axis z. The stress s is proportional to the stress uniaxial strain $\delta L_z/L_z$, being the proportionality constant the Young modulus. At the nanoscale, elastic strain is associated to a change of the interatomic distance. Thus, Young modulus measures the resistance of the bonds between atoms to deform. The effective stiffness of the nanocontacts can be extracted from the slope of the elastic stages of the force versus separation curve.

When a macroscopic object subjected to a load does not recover its original shape after the load is removed, but is deformed in a permanent, no recoverable fashion, the deformation lies into the plastic regime. Elastic regime holds for most metallic materials when the strain is smaller than 0.005. Higher strain values cause a transition from elastic to plastic regime or yielding. In macroscopic samples is difficult to fix the lower limiting stress below which no plastic deformation is appreciable because the transition from plastic to elastic regime takes place gradually. The yield strength is conventionally defined as the stress necessary to produce a plastic strain of 0.002 under uniaxial stress. This yield strength depends on heat treatment, sample purity and prior deformation of the sample. On the contrary, Young modulus is not sensitive to these factors.

Plastic deformation is equivalent at the nanoscale to permanent change in the positions of the atoms: they do not recover their original positions after the load is removed. Hence, when subjected to a plastic deformation, the atoms break the bonds with their neighbors and form new bonds with new neighbors. The simplest model of plastic deformation of a perfect crystal, that is, one with no defects, considers the sliding of two compact planes with respect to each other. Frenkel [51] calculated the maximum shear stress required for this process to occur and found a value of $\tau_{\max} \approx G_{\text{shear}}/30$, where G_{shear} is the shear modulus of the material. This value of the shear modulus is much larger than those observed in macroscopic metal specimens. This discrepancy is explained by the presence of dislocations, which can glide at low stress values. Experiments on whiskers find high values of the shear stress, and even

higher values are expected in nanocontacts, where dislocations are expelled from nm-volume regions [5].

The maximum pressure on nanocontacts can be extracted from the combined force and conductance measurements before the relaxation, see Fig. 16.4. In ballistic contacts, the area of minimum cross-section and the conductance are related through Sharvin formula. The apparent pressure found in gold nanocontacts ranges from 3 to 6 GPa for the compression and elongation parts of the indentation cycle. This value is much larger that the maximum pressure that a macroscopic sample can sustain. Moreover, it is of the same order of magnitude as the theoretical value in the absence of dislocations and is consistent with the theoretical maximum shear stress value for gold. A comparison of the elastic constant of nanocontacts with continuum mechanics models shows good agreement [4].

The energy dissipated in each force relaxation, that is, the energy necessary to produce a configurational change, can be directly obtained from the force cycle. In Fig. 16.4, the energy to pass from one configuration to the next is given by the grayed area. The value of this energy is of the order of 0.1 eV per atom in the minimal cross-section of the contact. If we compare with the heat of fusion (0.13 eV/atom) we find that configurational changes take place only at the zone around the narrowest part of the nanocontact.

16.4.3 The Shape of Metallic Nanocontacts

The shape of nanocontacts has been studied using a STM. Local modifications on a surface have been produced by approaching a Pt-Ir tip to a Ag substrate in UHV, touching the surface and imaging after tip indentation [2]. The surface after a gentle indentation cycle of the tip into substrate exhibits a nanometer sized protrusion. This is attributed to the formation of a small neck while pushing the tip that subsequently is stretched and broken.

In order to estimate the shape of the constriction Untiedt et al. [52] used a slab model for the constriction. Experimentally, the conductance versus displacement curves are recorded during an indentation cycle (Fig. 16.5) [52]. When the constriction is subjected to a force, the stresses are highest in the narrowest part, with cross-sectional area A_i. Therefore, one may assume that plastic deformation takes place in that narrowest slab, in a zone of depth λ, leaving the rest of the neck unmodified. In this model, a new slab is formed from a piece of length λ_i and area A_i. The new slab cross-sectional area A_{i+1} is given by volume conservation, being $\lambda_i + \Delta l$ the length of the new slab, where Δl is the deformation. Given that only the narrowest slab is modified, the shape of the constriction after a number of plastic deformation processes results in a sequence of values of A_i and λ_i. The plastic deformation length λ_i can be obtained from the experimental $G(z)$ curve noting that for the limit $\Delta l \to 0$, $\lambda = -(d \ln A/dl)^{-1}$, where A is the cross-section of the narrowest portion of the constriction.

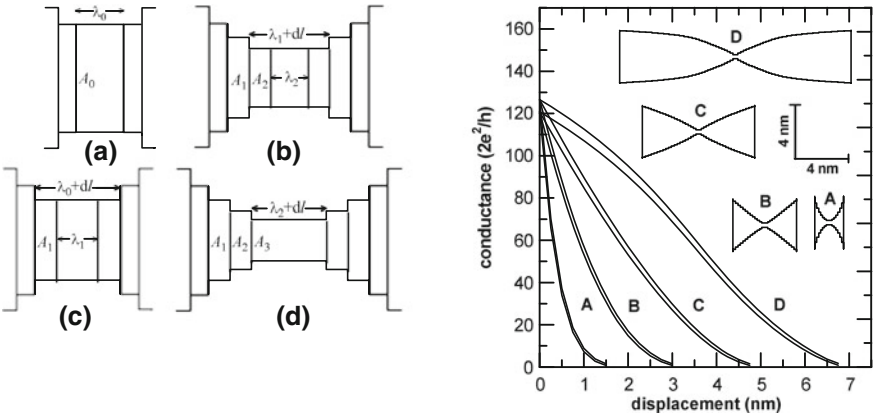


Fig. 16.5 The *left panel* shows the plastic deformation of a nanocontact using the slab model. From one configuration to the next only the central slab elongates, while the rest of the neck does not change. The *right panel* shows calculated shapes and conductance for the contacts of Fig. 16.6 using the slab model. Reprinted figure with permission from [52]. Copyright (1997) by the American Physical Society

As is shown in Fig. 16.6, gold nanocontacts may exhibit different behaviors in the conductance curve corresponding to different indentation cycles. The have markedly different slopes in the conductance curves. From the slopes of the conductance curves Untiedt et al. obtained the constriction shape (Fig. 16.5, right panel) using the described slab model. The analysis indicates that the steeper $G(z)$ curves correspond to constrictions with larger opening angle and involve shorter plastic deformation lengths. Hence, only one atomic layer is involved in plastic deformation at yielding events in constrictions with this shape.

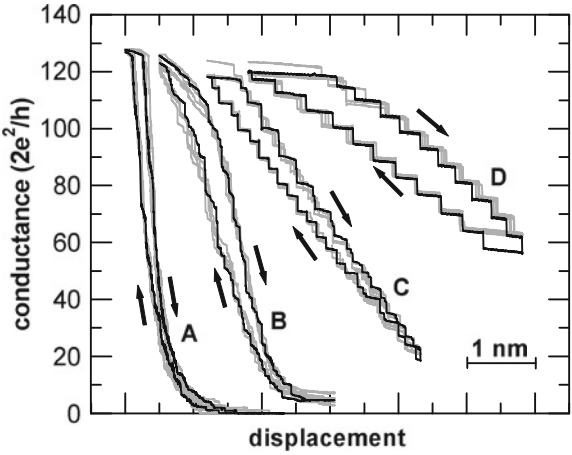


Fig. 16.6 Conductance curves $G(z)$ for four different sets of indentation cycles for Au at 4.2 K. Each set consists of five indentations. Reprinted figure with permission from [52]. Copyright (1997) by the American Physical Society

16.4.4 Inelastic Scattering by Phonons in Nanocontacts

Up to this point, only elastic scattering has been considered, but when applying a bias voltage in order to measure conductance, there are also effects derived from inelastic scattering that can be related to mechanical properties. In ballistic point contacts such as nanocontacts, the dominant inelastic scattering mechanism is electron-phonon scattering. This interaction has been shown to be a useful tool when studying nanocontacts. The derivative of the differential conductance of a point contact (PC) contains information about the inelastic electron backscattering. The Eliashberg function for the electron-phonon interaction in the point-contact situation, or point-contact spectroscopic (PCS) curve, is proportional to the density of states of phonons at a given energy F(eV) times a factor related to the electron-phonon interaction α^2, as shown in [53], by

$$\alpha^2 F = -\frac{3}{32\sqrt{2}} \frac{h^{3/2} k_{\mathrm{F}}^2}{4\pi^2 m_{\mathrm{e}}} G^{-3/2} \frac{\mathrm{d}^2 I}{\mathrm{d}V^2}, \tag{16.4}$$

where $G = (\mathrm{d}I/\mathrm{dV})$ is the differential resistance and m_{e} is the electron mass.

The amplitude of the phonon-induced peaks is reduced if there is elastic scattering, for example, due to impurities or defects. Consequently, large PCS amplitude indicates that the constriction is indeed ballistic. Hence, it is possible to obtain experimental information on the degree of disorder in the constriction using point-contact spectroscopy [54].

Untiedt et al. [52] found that in a point contact at low temperature the PCS curve does not change with the conductance as the derivative of the conductance scales with $G^{-3/2}$ as given by (16.4). Figure 16.7 shows the PCS curves corresponding to contacts obtained from indentation cycles similar to D, C, and B in Fig. 16.6. The PCS curve does not change with the conductance since $\mathrm{d}^2 I/\mathrm{dV}^2$ scales with $G^{-3/2}$,

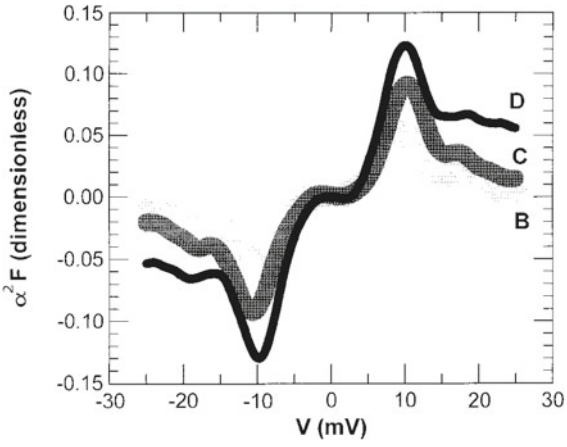


Fig. 16.7 Point contact spectroscopy curves at 4.2 K for Au nanocontacts. The *black curve* corresponds to a contact such as *D* in Fig. 16.6. The *gray curve* corresponds to type *C* and the *light gray curve* to type *B*. Reprinted figure with permission from [52]. Copyright (1997) by the American Physical Society

but the amplitudes of the phonon peaks are different for different realizations of the contact.

For a contact such as *D*, the amplitude of the phonon peaks is maximum (the black curve in Fig. 16.7) and similar to that of previously reported spectra [55] for ballistic point contacts. This is an indication of the degree of order in these contacts that can be attributed to a crystalline contact [52]. In contacts such as *B* and *C*, the amplitude of the phonon peaks in the PCS curve is reduced (the gray curves in Fig. 16.7) due to elastic scattering in the neck region, indicating the presence of defects, but still the necks are far from being disordered. From this evidence, they concluded that nanocontacts are ballistic, and are not disordered. This idea is supported by the results of the MD simulations [5, 56].

A comparison with the plastic deformation length λ for these necks at the largest radius (approximately 1.9 nm) corresponding to contacts with conductance 120 G_0 shows large plastic deformation lengths in contacts with high crystalline order. For curve *A*, $\lambda = 0.2$ nm; for curve *B*, $\lambda = 1$ nm; for curve *C*, $\lambda = 2.8$ nm. In addition, for curve *D* λ about 6 nm. Given that λ is related to the amount of material involved in the plastic deformation that in the contact *D* there are many layers involved in the process of plastic deformation, many more that those involved in the case of the contact *A*.

16.5 Suspended Chains of Single Gold Atoms

Experiments by Ohnishi et al. using a Transmission Electron Microscope [48] and Yanson et al. [57] using a MCBJ and a STM, showed that when a gold nanocontact is stretched, it may form a chain of single atoms. These structures are not only the ultimate nanowire, but they are also an ideal test bed for mesoscopic physics. Atomic chains of gold atoms have distinctive physical properties: they can sustain very large current densities of up to 8×10^{14} A m^{-2}. This fact supports a main assumption of the ballistic transport, since it is possible only if the transmission through the chain is ballistic and that most of the power is dissipated in the electrodes, far away from the contact. Furthermore, given that these atomic gold structures are stable for as long as 1 hour or even longer times at low temperature, it is possible to test peculiar physical properties since atomic chains are close to ideal one-dimensional metallic systems. Finally, not only gold can form chains: Smit et al. [58] reported that platinum and iridium spontaneously can form chains of atoms when pulling a one atom-contact.

16.5.1 Fabrication of Chains of Atoms Using Local Probes

The evolution of conductance during the breaking of a gold nanocontact is shown in Fig. 16.8. The conductance decreases while pulling from the contact, down to

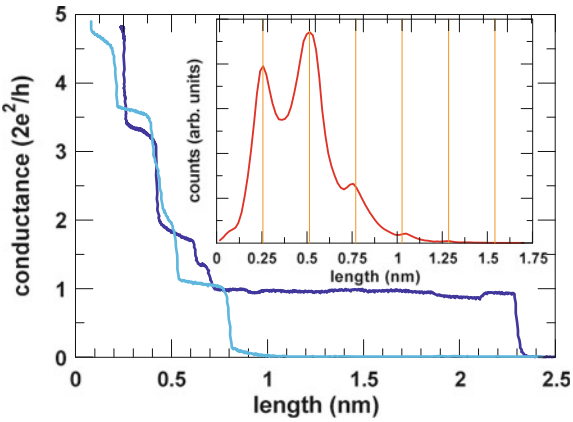


Fig. 16.8 Evolution of the conductance while pulling from the gold electrodes, extracting an atomic chain (*black curve*) at 4.2 K. The last plateau of the conductance corresponds to an atomic chain about six atoms long. *Inset* histogram of the last plateau length made from 10,000 indentation cycles that exhibit peaks at 0.25 nm. Reprinted figure with permission from [79]. Copyright (2002) by the American Physical Society

a conductance value close $1G_0$. As was discussed already in Sect. 16.3, a conductance of $1G_0$ in gold corresponds to a contact with a cross-section of one atom. The experiments show that a one-atom contact accommodates a maximum elastic deformation below 0.25 nm. However, it was discovered that gold nanocontacts exhibit sometimes a different behavior: this one atom-contact can be further stretched by a distance larger than 1 nm, without the conductance deviating appreciably from $1G_0$, showing up in Fig. 16.8 as a very long conductance plateau.

As was discussed in Sect. 16.4, the breaking process of metallic nanocontacts by controlled separation of the electrodes takes place in a sequence of elastic deformation and abrupt yield stages resulting in a non-continuous reduction of the minimal cross-section of the contact. It has been shown that for some metals this process takes place down to the smallest contact, a single atom contact between the electrodes [4]. Further separation of the electrodes usually results in breakage of the metallic contact and an abrupt jump into the tunneling regime. The explanation is that this stretching takes place at the position of this one atom contact. Yanson et al. [57] concluded that a single atom does not have to break, but that two atoms forming a short chain between the electrodes can replace it. As the contact is stretched even further, this chain will often break, but it does have a finite probability to be replaced by a chain containing more atoms. Once the chain starts being pulled the conductance never exceeds $1G_0$, confirming that the chain acts as one-dimensional nanowire. When the chain finally breaks, the electrodes have to travel back a return distance to re-establish metallic contact. This return distance is almost equal to the length of the last plateau itself, suggesting that after the chain breaks, and its constituent atoms collapse onto the electrodes on either side. In addition to that, Yanson et al. [57] showed by the use of a STM that they can swing one of the electrodes sideways at any given position

at the last plateau. They found that the lateral displacement that the structure can support increases as the length of the plateau increases, just the behavior that one would expect for a chain.

The conductance traces for successive nanocontact ruptures do not reproduce in detail, as they depend on the exact atomic positions in the contact and not every contact rupture results in the formation of a chain of atoms [57]. The definition of the length of the plateau, and thus the length of the chain, was therefore given as the distance between the points at which the conductance drops below $1.1G_0$ and at which conductance drops below $0.5G_0$. The probability of formation of such a structure can be quantified by constructing a histogram of last plateau length. As is shown in Fig. 16.8, the histogram exhibits a series of equidistant peaks rather than a smooth distribution. The peaked structure of the histogram shows that atomic chains tend to be elongated by integer multiples of 0.25 ± 0.2 nm, which is close to the nearest neighbor spacing of gold atoms in the crystal. In addition, it is shown that the probability of pulling a chain of length L decreases rapidly for large L.

Despite the low probability of formation of chains, once an atomic chain is pulled, the retraction of the electrode can be stopped and the chain remains very stable at liquid helium temperature (4.2 K): some of the longest chains obtained in the experiments have been held stable for at least 1 hour. This makes atomic chains suitable for investigation of one-dimensional electron transport and for studies of wear and fracture on low-coordinated metallic nanostructures.

16.5.2 Mechanical Processes During Formation of Atomic Chains

The mechanical processes involved during the formation and rupture of atomic chains of gold have been studied both experimentally and theoretically, see [3] and references therein. A STM supplemented with a force sensor (see Sect. 16.2) has been used to study the force in gold chains at low temperature [13]. Using this probe it is feasible to study the force evolution simultaneously with the conductance while drawing out the chain.

Figure 16.9 shows a simultaneous measurement of force and conductance. The force shows a sequence of linear stages separated by sudden relaxations. The conductance on the last plateau remains just below to $1G_0$. There are small conductance jumps related to force relaxations, but their magnitude is much smaller than $1G_0$. In stages with a linearly growing tensile force the chain is stretched, while at the force jumps abrupt atomic rearrangements occur.

If an atomic chain is formed or not while pulling a nanocontact relates to the relative strength of different bonds in specific atomic configurations. The reason is that breaking a nanocontact involves breaking many individual atomic bonds. One simple model consists of a chain connected to the electrodes on each side. In an elastic stage, as the chain is drawing out the force increases. When the system is stretched so that the tensile force value is higher than a critical force the weakest of the bonds will break. If the slip of an atom from an electrode into the chain requires

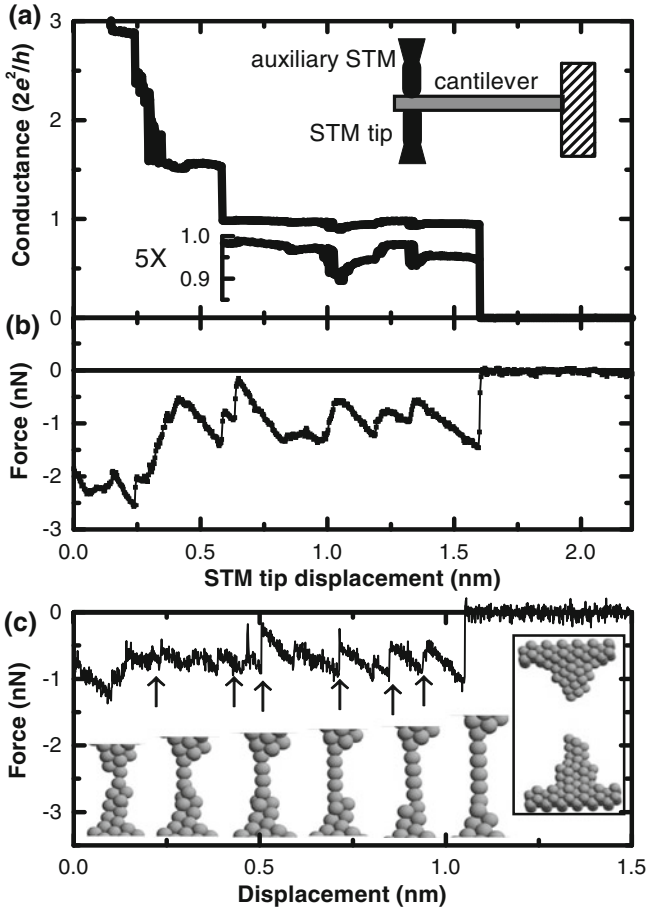


Fig. 16.9 *Top panel* conductance (**a**) and tensile force (**b**) measured simultaneously while making and breaking a chain of gold atoms at 4.2 K. *Bottom panel* force calculated from a MD simulation. *Arrows* indicate where a new atom pops into the chain and snapshots of the structure at these positions are shown. *Inset* scheme of the experimental setup, a combined STM-AFM. Reprinted figure with permission from [13]. Copyright (2001) by the American Physical Society

a smaller force than the force needed to break the chain, then atom is added and the chain grows. This is not usually the case because pulling an atom out of a surface implies breaking more bonds than to break the chain. Nevertheless, MD simulations sustain the hypothesis that an atom could slip into the chain at lower forces [13].

In order to examine a more realistic model it is important to take in account the fact that the bond strength increases as the coordination number is reduced. Suppose a contact where the electrodes have a pyramidal shape. The atom sitting at the apex of one of these pyramids has three bonds with its neighbors placed in the pyramid and one with the chain. Given that the bond strength increases as the coordination

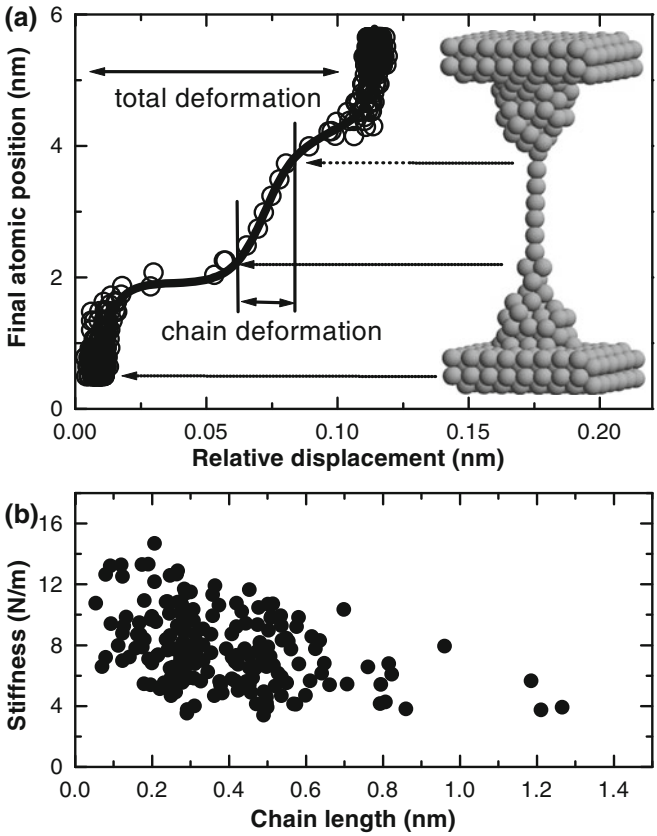


Fig. 16.10 MD simulation of the relative displacement during an elastic deformation stage. The *bottom panel* shows the measured chain stiffness just before rupture as a function of chain length. Reprinted figure with permission from [13]. Copyright (2001) by the American Physical Society

number is reduced it may be favorable for the atom in the apex break a bond with the underlying atoms. Moreover, in this model this atom can be incorporated into the chain breaking only one bond. MD calculations (see [13, 59] and references in [3]) show that generally larger force jumps correspond to the incorporation of an atom into the bridging atomic chain. An experimental fact that supports this hypothesis is that mechanical relaxations take place at force values smaller than the final breaking force. A further analysis needs to take in consideration the detailed configuration of atoms at the apex of the electrode as well as the relative strength of their bonds.

MD simulations [13] show that the elastic deformation is accumulated in the electrodes, see Fig. 16.10. This could be explained because the electrodes can be deformed not only by stretching the interatomic distance. It is also possible that atoms in the electrodes are sitting in arrangements where the breaking of bonds is the result of a more concerted motion of atoms that requires smaller forces given its

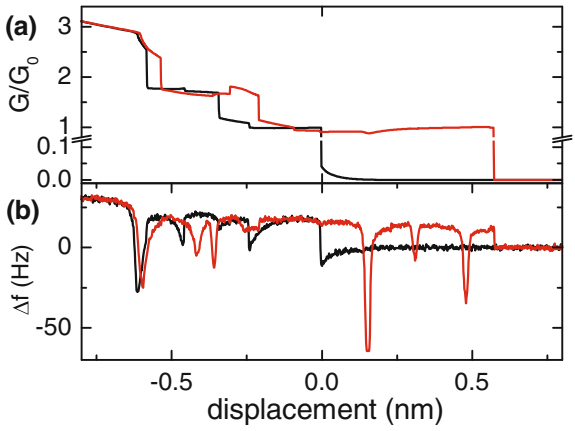


Fig. 16.11 Simultaneous measurement of conductance (**a**) and the stiffness (**b**) during elongation of a chain of atoms. Reprinted figure with permission from [12]. Copyright (2004) by the American Physical Society

longer paths. In addition, the atoms in the chain have stronger bonds due to its low coordinated situation. The sum of these two effects results in a peculiar feature of this nanostructure: thinner is actually stronger.

Some further mechanisms behind the formation of an atomic chain have been recently studied in Pt chains at low temperature [60]. Using a MCBJ supplemented with a force sensor, they have found a correlation between the stiffness and the number of atoms in the chain. A comparison between the mean total stiffness (i.e. the electrodes plus the chain) is found to be higher if the chain breaks when pulling than if one atom could be added to the chain. They can conclude that longer chains have lower stiffness. This could be supported from the fact that a chain can sustain a maximum force that is independent of its length, as was reported by force measurements on Au chains. Roughly speaking, for the same elongation a larger force acts on the chain that has a larger starting stiffness, breaking it at a shorter length.

Rubio Bollinger et al. [12] have used the MCBJ technique and obtained high-resolution measurements of the contact stiffness. Figure 16.11 shows simultaneous measurement of the conductance and the force gradient while pulling a chain of gold atoms. The atomic rearrangements result in sudden drops in the stiffness that reflect bond weakening due to extreme strains close to yield. These drops are also observed during the atomic chain formation, whose conductance, in contrast remains close to $1G_0$. This mechanical behavior agrees with direct force measurements in atomic chains and is consistent with the observed longitudinal phonon frequency decrease in one-atom contacts and chains, as is showed in the next section.

16.5.3 Phonons in Atomic Chains

Dynamical mechanical properties of nanocontacts and chains can be probed measuring electron transport and using point contact spectroscopy, because the finite

electron-phonon scattering. Using PCS techniques, Agraït et al. have studied the evolution of these phonon modes when a chain of gold atoms is stretched [61].

In atomic chains of gold, the electron flow should be ballistic. As the mean free path of the electrons is larger than the device length, electrons are supposed to move freely: there is neither scattering nor defects to inhibit resistance-free current. In a simplified picture, the zero-bias conductance $G(0)$ of an atomic chain is the conductance quantum G_0 [62–64]. Neglecting further effects [65], this zero-bias conductance is a consequence of the coupling between the chain and the reservoirs of electrons, the electrodes. This is clear from the Landauer approach to electronic transport [66, 67], which can be applied to atomic chain. When a small bias voltage is applied between the two electrodes, there is unbalance of population at the electrodes and a net current flows. Subsequently there is heat dissipated from the resistance. The chain is considered as a perfect conductor, and this dissipation is attributed to the relaxation of the electrons to the Fermi level in the electrodes and far away from the chain.

Experimentally, it has been found that the zero-bias conductance of an atomic gold chain at low temperatures is close to $1G_0$ [57]. This chain has one open channel in agreement with theoretical calculations [62–64]. In experiments at liquid He temperatures by Agraït et al. [61, 68], the conductance is measured using the lock-in technique, and the derivative of the differential conductance $\mathrm{d}G/\mathrm{d}V$ was calculated numerically. As soon as the voltage is swept a range of 20 mV from zero, the conductance does depend on the voltage. Typical differential conductance curves for short and long atomic wires are shown in Fig. 16.12b–d. The differential conductance G features a hump at zero bias, dropping about 1 % in the range (20 mV). Often the differential conductance curves look asymmetrical, showing oscillations due to elastic scattering [69, 70].

A voltage-dependent conductance is observed in ballistic point-contacts of much larger size, where it is associated to inelastic scattering of electrons with phonons and other elementary excitations taking place in the bulk [53, 71]. These processes are voltage dependent because the electrons must be injected with enough energy to emit an excitation. In the case of phonons, the derivative of the conductance shows peaks which correspond to peaks in the phonon density of states. The amplitude of the signal for contacts of different sizes is proportional to $G(0)^{3/2}$. This reflects the fact that only those electrons scattered in the immediate vicinity of the contact have a significant probability of coming back through the contact. The standard spectra for Au have peaks at 10 and 18 mV, corresponding to the maxima in the transverse and longitudinal phonon Density of States (DOS). In those spectra, the transverse peak is found to be stronger than the longitudinal peak.

Atomic chains exhibit markedly different features. First, the peaks in the PC spectrum vary for each chain as is shown in Fig. 16.12. In the case of a one atom long atomic wire, labeled *S*, the hump in conductance signal is narrower than in a larger wire, labeled *M*, supposed to be 0.3 nm longer. This distance is enough to accommodate one more atom in the chain. The peak of the spectrum is about three times larger than that given by the semiclassical theory of PC spectroscopy [53, 71].

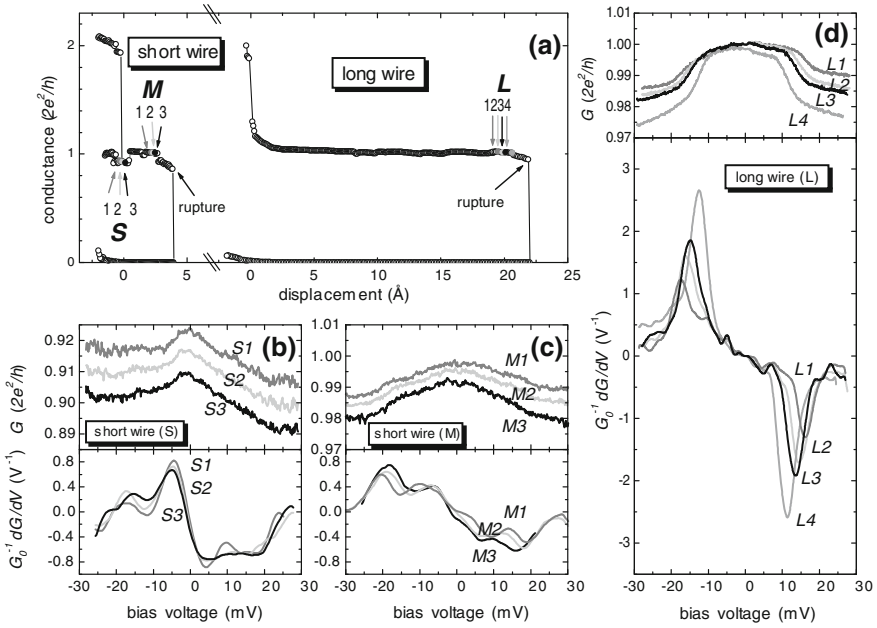


Fig. 16.12 *Panel* **(a)**: Conductance at finite bias of two gold chains of different lengths. The short chain is about 0.4 nm and the long chain is 2.2 nm long. *Panels* **(b)**, **(c)**, and **(d)**: differential conductance and its derivative at points *S*, *M*, and *L*, respectively, marked by the *arrows*. The various *curves* in **(b)**, **(c)**, and **(d)** were acquired at intervals of 0.03, 0.03, and 0.05 nm respectively. Note that the *vertical scale thee panels* are identical. Reprinted figure with permission from [61]. Copyright (2002) by the American Physical Society

The wire labeled L corresponds to a long wire about seven atoms long, see Fig. 16.12a. In this long wire, Agraït et al. [61, 68] recorded the curve $G(V)$ at different elongations of the chain. The conductance curves, and subsequently the spectra, show a behavior that is considered a fingerprint of a one-dimensional system. The conductance curves show drops that take place quite sharply, and result in sharp peaks in the spectra.

Figure 16.12d shows that the position of the peak shifts as a function of the strain in the wire. Like the pitch of a guitar string, there a shift as a function of the tension, but for atomic wires the frequency decreases because of the decreasing bond strength between the atoms. As the frequency decreases, the amplitude increases in an elastic stage until an atomic rearrangement takes place. The increase in the amplitude is attributed partially to the softening of the phonon modes with tension.

In a one-dimensional system at zero temperature, momentum is conserved, and, consequently, electrons can only excite longitudinal vibrations. Since momentum is conserved, electrons can only excite longitudinal vibrations of the atomic chain whose wave number is twice the Fermi wave number k_F, hence only one phonon mode will be shown in the spectra. The Fermi wavevector k_F in a linear atomic chain

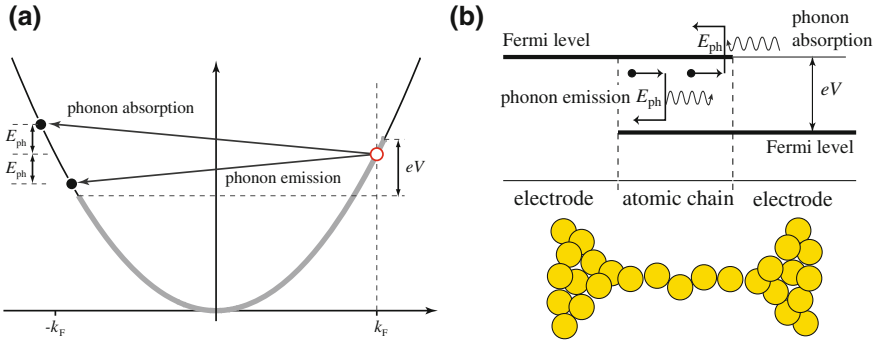


Fig. 16.13 *Left panel* allowed inelastic transitions in a one-dimensional band. The only allowed transitions are from a state with wavenumber k_F to a state with wavenumber $-k_F$, that is, electrons can only interact with one phonon whose wavenumber is $2k_F$. For an atomic wire with a single conduction electron per atom, $k_F = \pi/2a$, where a is the interatomic distance. Note that the figure is not drawn to scale: phonon energies are about two orders of magnitude smaller than electron energies. *Right panel* schematic representation of phonon emission and absorption processes in a one-dimensional ballistic wire. In the wire, there are two Fermi levels, which define the occupation of electronic states: for the *right*-going electron states up to the Fermi level of the *left* electrode are occupied, while for the *left*-going electrons the occupation of states is up to the Fermi level of the *right* electrode. The separation of these Fermi levels is eV, where V is the voltage difference applied to the electrodes. Electronic transitions are possible when the final electronic state is unoccupied; consequently, phonon emission is possible only for energies higher than a threshold $\hbar\omega_{2kF}$

with interatomic distance a and one conduction electron per atom is equal to $\pi/2a$. Then, if the energy of electrons is below a threshold voltage $\hbar\omega_{2k_F}/e$, they will not interact with phonons (see Fig. 16.13). A sudden decrease in conductance marks the onset of the phonon emission process. In the experimental conductance curves the drop in conductance is somewhat rounded due to the nonzero temperature, the thermal smearing is 2 meV for 4.2 K, see [72].

Given that the chain length is finite, there is an uncertainty in the momentum of the electron that also contributes to round the drop in the conductance signal. A linear monoatomic chain of N atoms will have N longitudinal vibrational modes. For an atomic chain of length L coupled to rigid electrodes, the wavelengths of the different modes n are simply given by $\lambda_n = 2L/n$. However, there is a non-rigid coupling to the electrodes that results in broader resonances. The width of the resonances is linked to the mechanical coupling between the electrodes and the chain.

Furthermore, momentum conservation in the electron-phonon interaction will not hold strictly in a finite system. Therefore, electrons may interact with phonons at different energies and then, other vibrational modes appear in the spectra. This effect is more important in shorter chains.

While the position of the peak in the spectra V_{ph} gives the frequency of the ω_{2k_F} phonon, and its height A_{ph}, which is related to the conductance drop, is proportional to the probability of the phonon emission process. The magnitude of the conductance drop (about 1 % for a chain of 2 nm in length) is consistent with an inelastic

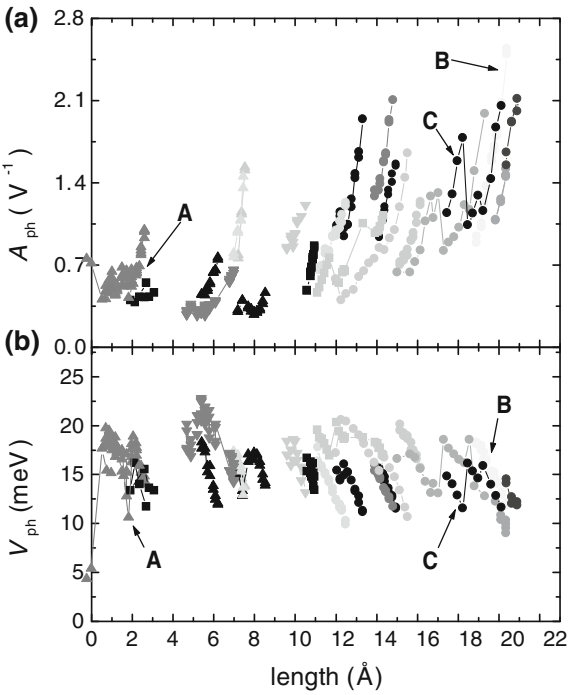


Fig. 16.14 **a** Magnitude A_{ph} and **b** position V_{ph} of the phonon peak in the PC spectrum as a function of chain length. For clarity, only 22 representative chains out of more than 100 studied are shown. Each chain is represented by a different symbol. The length of the chain is estimated from the length of the last plateau. Reprinted figure with permission from [61]. Copyright (2002) by the American Physical Society

mean-free path of about 200 nm in an infinite wire, which is reasonable for a metal at low temperatures.

Experimentally, it is observed that stretching the wire results in an increase in the emission probability. The increase in the emission probability, see Fig. 16.14, indicates an enhancement of the electron-phonon interaction. A plot of the amplitude for many different atomic wires is shown in Fig. 16.14. The emission probability increases with the length of the wire, and the variations due to stretching are much larger than for short wires [61, 68]. The background conductance as well as the zero bias conductance remains mostly unchanged. This behavior strongly suggests that the background features correspond to backscattering processes in the bulk while the peak is related to backscattering processes in the chain itself.

16.6 Metallic Adhesion in Atomic-Sized Tunneling Junctions

The study of mechanical properties of small tunneling junctions between single asperities or tips is of fundamental importance. These forces are involved in STM imaging mechanisms and tribological imaging of surfaces at the nanoscale [73, 74]. Moreover, the same situation occurs both after the rupture and before the formation of a one-atom contact. When the separation between two tips is small enough, there

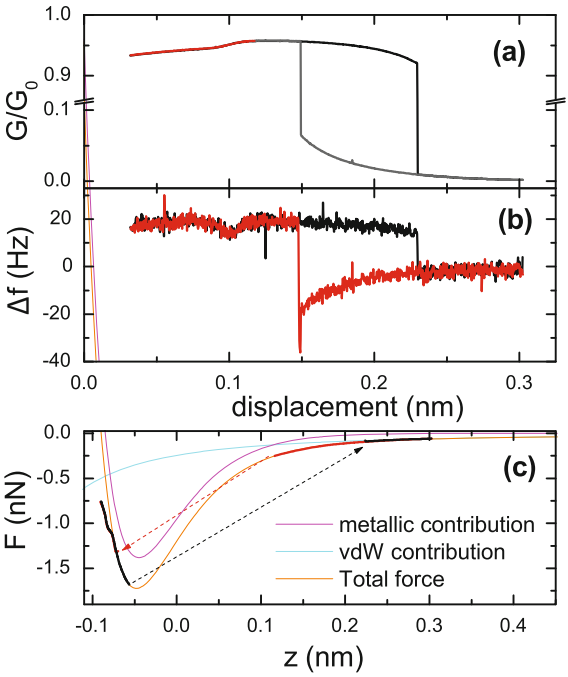


Fig. 16.15 Simultaneous **a** conductance and **b** frequency shift measurements for a one-atom junction. **c** Force, with van der Waals and metallic force components. Reprinted figure with permission from [12]. Copyright (2004) by the American Physical Society

is a spontaneous jump to contact. Beyond a critical distance, the electrode separation becomes unstable and consequently a one-atom contact is established. This is reflected as a sudden increase of the conductance. The process described above is common in the formation of a metallic contact despite there are some exceptions [75]. For example in W contacts there are sometimes a continuous increase of current from tunneling to contact [76].

This section is focused on experiments between two gold tips, where a spontaneous jump to contact occurs. Both the force gradient and the tunneling current has been measured simultaneously using a MCBJ supplemented with a force sensor described in Sect. 16.2 [12, 31]. The MCBJ technique provides freshly fractured surfaces that are only exposed to cryogenic vacuum, reducing their contamination to an absolute minimum. The force gradient of the interaction between the electrodes is proportional to the shift of the resonance frequency of the TF if excited with vanishing amplitude. This resonant frequency is tracked by implementing a phase locked loop oscillator, and results in a high sensitivity measurement. With such a tool different contacts realizations modifies the atomic configuration of the tip apex. Hence, different tips are fabricated by making a large contact and breaking it again.

In Fig. 16.15, an indentation cycle is shown. The stiffness of the contact can be obtained straightforwardly from the curve. It decreases while pulling the contact. The stiffness of the one-atom contact is 5.8 N m^{-1} in this case. When the one-atom contact is broken, there is an abrupt jump in both the current and the force gradient (jump out

contact). After that, the two tips are approached from the tunneling regime. The forces acting in these regimes are attractive since the frequency shift, and consequently the force gradient, is always negative during the approach. Note that the decay length of this force is below 0.05 nm, indicating a short-range metallic interaction. This metallic adhesion is due to the overlap of the electronic wave functions. Various theoretical models [77, 78] for metallic adhesion predict a decay length similar to the tunneling current decay length. For smaller distance, there is a spontaneous jump to contact and again we recover a one-atom contact.

Acknowledgments We would like to acknowledge fruitful discussions with C. Untiedt, R.H.M. Smit and P. Joyez. This work was partially supported by MINECO (MAT2011-25046) and by Comunidad de Madrid (Spain) through program Citecnomik (S-0505/ESP/0337).

References

1. U. Dürig, O. Züger, D.W. Pohl, Phys. Rev. Lett. **65**, 349 (1990)
2. J.K. Gimzewski, R. Möller, Phys. Rev. B **36**, 1284 (1987)
3. N. Agraït, A.L. Yeyati, J.M. van Ruitenbeek, Phys. Rep. Rev. Sect. Phys. Lett. **377**, 81 (2003)
4. G. Rubio, N. Agraït, S. Vieira, Phys. Rev. Lett. **76**, 2302 (1996)
5. U. Landman, W.D. Luedtke, N.A. Burnham, R.J. Colton, Science **248**, 454 (1990)
6. P. Avouris, Z. Chen, V. Perebeinos, Nat. Nano **2**, 605 (2007)
7. attoAFM. Attocube systems AG
8. CryogenicSFM. Omicron Nanotechnolgy GmbH
9. LT-SPM. NanoMagnetics Instruments.
10. G. Binnig, H. Rohrer, Helvetica Physica Acta **55**, 726 (1982)
11. G. Binnig, C.F. Quate, C. Gerber, Phys. Rev. Lett. **56**, 930 (1986)
12. G. Rubio-Bollinger, P. Joyez, N. Agraït, Phys. Rev. Lett. **93**, 116803 (2004)
13. G. Rubio-Bollinger, S.R. Bahn, N. Agraït, K.W. Jacobsen, S. Vieira, Phys. Rev. Lett. **87**, 26101 (2001)
14. S.P. Jarvis, M.A. Lantz, H. Ogiso, H. Tokumoto, U. Durig, Appl. Phys. Lett. **75**, 3132 (1999)
15. G. Meyer, N.M. Amer, Appl. Phys. Lett. **53**, 1045 (1988)
16. D. Rugar, H.J. Mamin, R. Erlandsson, J.E. Stern, B.D. Terris, Rev. Sci. Instrum. **59**, 2337 (1988)
17. D. Rugar, H.J. Mamin, P. Guethner, Appl. Phys. Lett. **55**, 2588 (1989)
18. U. Stahl, C.W. Yuan, A.L. Delozanne, M. Tortonese, Appl. Phys. Lett. **65**, 2878 (1994)
19. R. Garcia, R. Perez, Surf. Sci. Rep. **47**, 197 (2002)
20. F.J. Giessibl, Rev. Mod. Phys. **75**, 949 (2003)
21. F.J. Giessibl, Appl. Phys. Lett. **73**, 3956 (1998)
22. J. Rychen, T. Ihn, P. Studerus, A. Herrmann, K. Ensslin, Rev. Sci. Instrum. **70**, 2765 (1999)
23. K. Besocke, Sur. Sci. **181**, 145 (1987)
24. D.W. Pohl, Rev. Sci. Instrum. **58**, 54 (1987)
25. C. Renner, P. Niedermann, A.D. Kent, O. Fischer, Rev. Sci. Instrum. **61**, 965 (1990)
26. I.B. Altfeder, A.P. Volodin, Rev. Sci. Instrum. **64**, 3157 (1993)
27. S.H. Pan, Patent WO 9319494 (1993)
28. S.H. Pan, E.W. Hudson, J.C. Davis, Rev. Sci. Instrum. **70**, 1459 (1999)
29. J. Moreland, J.W. Ekin, J. Appl. Phys. **58**, 3888 (1985)
30. C.J. Muller, J.M. Vanruitenbeek, L.J. Dejongh, Physica C **191**, 485 (1992)
31. A.M.C. Valkering, A.I. Mares, C. Untiedt, K.B. Gavan, T.H. Oosterkamp, J.M. van Ruitenbeek, Rev. Sci. Instrum. **76**, 103903 (2005)

32. J.M. van Ruitenbeek, A. Alvarez, I. Pineyro, C. Grahmann, P. Joyez, M.H. Devoret, D. Esteve, C. Urbina, Rev. Sci. Instrum. **67**, 108 (1996)
33. T.R. Albrecht, P. Grutter, D. Horne, D. Rugar, J. Appl. Phys. **69**, 668 (1991)
34. F.J. Giessibl, Appl. Phys. Lett. **76**, 1470 (2000)
35. W.H.J. Rensen, N.F. van Hulst, A.G.T. Ruiter, P.E. West, Appl. Phys. Lett. **75**, 1640 (1999)
36. J. Rychen, T. Ihn, P. Studerus, A. Herrmann, K. Ensslin, H.J. Hug, P.J.A. van Schendel, H.J. Guntherodt, Rev. Sci. Instrum. **71**, 1695 (2000)
37. H. Edwards, L. Taylor, W. Duncan, A.J. Melmed, J. Appl. Phys. **82**, 980 (1997)
38. W.A. Atia, C.C. Davis, Appl. Phys. Lett. **70**, 405 (1997)
39. K. Karrai, R.D. Grober, in *Near-Field Optics (Proceedings of the Society of Photo-Optical Instrumentation Engineers (Spie))*, vol. 2535, ed. by M.A. Paesler, P.J. Moyer (1995), p. 69
40. F.J. Giessibl, S. Hembacher, M. Herz, C. Schiller, J. Mannhart, Nanotechnology **15**, S79 (2004)
41. J. Rychen, T. Ihn, P. Studerus, A. Herrmann, K. Ensslin, H.J. Hug, P.J.A. van Schendel, H.J. Guntherodt, Appl. Surf. Sci. **157**, 290 (2000)
42. R.H.M. Smit, R. Grande, B. Lasanta, J.J. Riquelme, G. Rubio-Bollinger, N. Agrait, Rev. Sci. Instrum. **78**, 113705 (2007)
43. J.C. Maxwell, *A Treatise on Electricity and Magnetism*, vol. 1 (Courier Dover Publications, 1954)
44. M. Knudsen, *The Kinetic Theory of Gases*. (Methuen, London, 1934)
45. Y.V. Sharvin, Zh. Eksp. Teor. Fiz **48**, 984 (1965)
46. J.A. Torres, J.I. Pascual, J.J. Saenz, Phys. Rev. B **49**, 16581 (1994)
47. E. Scheer, N. Agrait, J.C. Cuevas, A.L. Yeyati, B. Ludoph, A. Martin-Rodero, G.R. Bollinger, J.M. van Ruitenbeek, C. Urbina, Nature **394**, 154 (1998)
48. H. Ohnishi, Y. Kondo, K. Takayanagi, Nature **395**, 780 (1998)
49. V. Rodrigues, D. Ugarte, Phys. Rev. B **6307**, 3405 (2001)
50. M.H. Sadd, *Elasticity Theory, Applications, and Numerics* (Elsevier Butterworth Heinemann 2005)
51. J. Frenkel, Z. Physik **37**, 572 (1926)
52. C. Untiedt, G. Rubio, S. Vieira, N. Agraït, Phys. Rev. B **56**, 2154 (1997)
53. A.M. Duif, A.G.M. Jansen, P. Wyder, J. Phys. Condens. Matter **1**, 3157 (1989)
54. A.G.M. Jansen, A.P. Vangelder, P. Wyder, J. Phys. C-Solid State Phys. **13**, 6073 (1980)
55. A.G.M. Jansen, F.M. Mueller, P. Wyder, Phys. Rev. B **16**, 1325 (1977)
56. R.M. Lyndenbell, Science **263**, 1704 (1994)
57. A.I. Yanson, R. Bollinger, Nature **395**, 783 (1998)
58. R.H.M. Smit, C. Untiedt, A.I. Yanson, J.M. van Ruitenbeek, Phys. Rev. Lett. **87**, 266102 (2001)
59. E.Z. da Silva, F.D. Novaes, A.J.R. da Silva, A. Fazzio, Phys. Rev. B **69**, 115411 (2004)
60. T. Shiota, A.I. Mares, A.M.C. Valkering, T.H. Oosterkamp, J.M. van Ruitenbeek, Phys. Rev. B **77**, 125411 (2008)
61. N. Agraït, C. Untiedt, G. Rubio-Bollinger, S. Vieira, Phys. Rev. Lett. **88**, 216803 (2002)
62. M. Brandbyge, N. Kobayashi, M. Tsukada, Phys. Rev. B **60**, 17064 (1999)
63. E.G. Emberly, G. Kirczenow, Phys. Rev. B **60**, 6028 (1999)
64. M. Okamoto, K. Takayanagi, Phys. Rev. B **60**, 7808 (1999)
65. R.H.M. Smit, C. Untiedt, G. Rubio-Bollinger, R.C. Segers, J.M. van Ruitenbeek, Phys. Rev. Lett. **91**, 76805 (2003)
66. R. Landauer, IBM J. Res. Dev. **1**, 223 (1957)
67. R. Landauer, Philos. Mag. **21**, 863 (1970)
68. N. Agraït, C. Untiedt, G. Rubio-Bollinger, S. Vieira, Chem. Phys. **281**, 231 (2002)
69. B. Ludoph, M.H. Devoret, D. Esteve, C. Urbina, J.M. van Ruitenbeek, Phys. Rev. Lett. **82**, 1530 (1999)
70. C. Untiedt, G.R. Bollinger, S. Vieira, N. Agraït, Phys. Rev. B **62**, 9962 (2000)
71. I.K. Yanson, Zhur. Eksper. Teoret. Fiziki **66**, 1035 (1974)
72. A.V. Khotkevich, I.K. Yanson, *Atlas of Point Contact Spectra of Electron-Phonon Interactions in Metals* (Kluwer Academic Publishers, Boston, 1995)
73. O. Pfeiffer, L. Nony, R. Bennewitz, A. Baratoff, E. Meyer, Nanotechnology **15**, S101 (2004)

74. A. Stalder, U. Durig, Appl. Phys. Lett. **68**, 637 (1996)
75. C. Untiedt, M.J. Caturla, M.R. Calvo, J.J. Palacios, R.C. Segers, J.M. van Ruitenbeek, Phys. Rev. Lett. **98**, 206801 (2007)
76. A. Halbritter, S. Csonka, G. Mihaly, E. Jurdik, O.Y. Kolesnychenko, O.I. Shklyarevskii, S. Speller, H. van Kempen, Phys. Rev. B **68**, 035417 (2003)
77. C.J. Chen, J. Phys. Condens. Matter **3**, 1227 (1991)
78. J.H. Rose, J.R. Smith, J. Ferrante, Phys. Rev. B **28**, 1835 (1983)
79. C. Untiedt, A.I. Yanson, R. Grande, G. Rubio-Bollinger, N. Agraït, S. Vieira, J.M. van Ruitenbeek, Phys. Rev. B **66**, 85418 (2002)

Chapter 17
Nanotribological Studies by Nanoparticle Manipulation

Dirk Dietzel, Udo D. Schwarz and André Schirmeisen

Abstract Friction of extended nanocontacts has lately become a subject of growing interest in nanotribology. The related length scales, which are not accessible to conventional friction force microscopy, can best be analyzed by measuring the friction of nanoparticles sliding over flat surfaces. By pushing nanoparticles with an AFM tip a large range of materials combinations and contact areas can be studied under well-defined interface conditions, therefore offering new insight into atomistic concepts of friction.

17.1 Nanoparticle Manipulation: An Alternative Route to Nanotribology

Since its invention in 1987, friction force microscopy (FFM) [1] has become a widespread technique for the investigation of frictional processes. By detecting the lateral forces acting between the tip of an atomic force microscope (AFM) and the sample surface, this technique has proven to be a versatile tool for the analysis of a wide variety of nanoscale frictional phenomena [2, 3].

In conventional FFM operation, the lateral force signal, which is proportional to the friction-induced cantilever torsion, is recorded during the sliding of the cantilever across the surface. The lateral force therefore originates from the contact between tip and surface. Measurements have been performed as a function of a wide variety of parameters such as the externally applied cantilever load [4–9], radius and shape of the AFM tip [4, 6, 7], sliding velocity [10–13], the temperature [14–16], the relative orientation between scan direction and substrate lattice [17–20], or the chemical nature of the sample [21–23]. Often, FFM studies led to the successful analysis of frictional processes at the atomic scale [1, 10, 13, 16].

D. Dietzel · A. Schirmeisen (✉)
Institute of Applied Physics (IAP), Justus-Liebig-Universität, Giessen, Germany
e-mail: schirmeisen@uni-giessen.de

U.D. Schwarz
Department of Mechanical Engineering and Materials Science and Center for Research on Structures and Phenomena (CRISP), Yale University, New Haven, CT, USA

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_17

But despite the indisputable successes of many FFM studies, this method still has severe inherent limitations. Mainly four aspects are of concern.

1. While the nature of the samples can be freely selected, commercially available AFM tips are usually limited to a very narrow set of materials, mostly silicon, silicon oxide, silicon nitride, and diamond. This limits the number of material combinations that can be investigated. To overcome that limitation, other materials can be evaporated onto AFM cantilevers. This strategy, however, generally leads to AFM tips of inferior quality and/or unknown geometry unless performed in a very controlled way. In addition, many materials that are applied as thin films to cover the AFM tip will wear out fast due to the significant shear stresses during scanning.
2. AFM tips used for FFM generally feature amorphous or disordered tip ends. Therefore, it is very difficult to investigate the effect of ordered structures on friction, which is expected to have a dramatic influence under certain circumstances. Most prominently, an effect denoted as *superlubricity* [24] or, more precisely, *structural lubricity* [25] is expected to occur at specific relative orientations of extended atomically flat contacts that show crystalline long-range order [26–29].
3. Related to the above point is the question how friction depends on the 'true' contact area at the atomic scale. It can be argued that the contact area dependence of friction is one of the most fundamental yet unsolved issues in nanotribology, as its understanding is crucial for successfully bridging the conceptual gap between nanoscale and microscale friction. Unfortunately, the fixed tip radius of commercially available cantilevers makes it difficult to analyze effects as a function of the contact area. This leaves a 'gap' in the experimental accessibility of contact areas between the tens of nm^2 realized in FFMs and the typically ten thousands of μm^2 found in surface force apparatus measurements [30, 31]. In addition, the determination of the contact area has to rely on the realization of a specific contact geometry (usually the Hertzian contact geometry, representing a spherical tip apex on a flat surface) and on the validity of certain assumptions of the contact mechanical models [32].
4. Recent theoretical studies also indicate that the contact area is not necessarily a suffficient parameter to describe the geometry of the interface [33]. The shape of the interface can influence friction as well, meaning that two nanoscopic contacts of same size but different shape can show decidedly different friction. And while some attempts have been made to analyze the contact area of friction experimentally [6, 7, 34, 35], so far there are no experimental studies related to the shape of the contact area.

In order to overcome these limitations, it would be desirable to have a method available that measures the interfacial friction of structurally well-defined contacts of arbitrary sizes, shapes, and material combinations. One possible solution is to use the AFM tip as a manipulation tool for controlled lateral manipulation of nanoscale particles supported by flat substrates [36] (see Fig. 17.1).

The concept of nanoparticle manipulation was first demonstrated with the example of C_{60} islands grown on a NaCl surface [37] and later used to investigate frictional

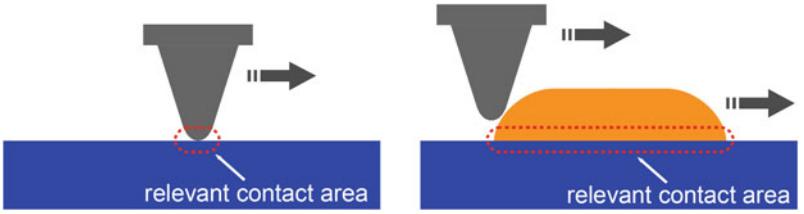


Fig. 17.1 The difference between conventional friction force microscopy studies and particle manipulation schemes lies in the relevant interface. While FFM is limited to friction occurring at the interface between tip and surface (*left*), the particle manipulation method allows to study the much larger, but well-defined particle/surface interface (*right*)

anisotropies for MoO nanoparticles [38]. Nanoscale objects, like nanotubes, have been pushed to distinguish sliding and rolling motion [39]. Recently, there has been an increase of systematic friction studies using nanoparticle manipulation, highlighting the influence of surface structure on particle trajectories [40] as well as the influence of parameters like surface chemistry and temperature [41]. The influence of relative orientation between particles, substrate, and direction of manipulation was analyzed for ligand-capped CdSe nanorods [42]. The fundamental question of how friction is related to contact area has been addressed by a systematic variation of the size of of metallic nanoparticles [34, 35]. Furthermore, nanoparticle manipulation experiments have been used to analyze the difference between static and sliding fricton. While some experimental approaches are mainly sensitive to either static friction [34, 43] or sliding friction [35], recent experiments have demonstrated, how nanoparticle manipulation can be used to monitor the transition from static to sliding friction [44].

The range of contact areas that is accessible to nanoparticle manipulation experiments is indicated in Fig. 17.2 in comparison to other experimental tools commonly used in nano- and microtribology, namely the friction force ficroscope (either with

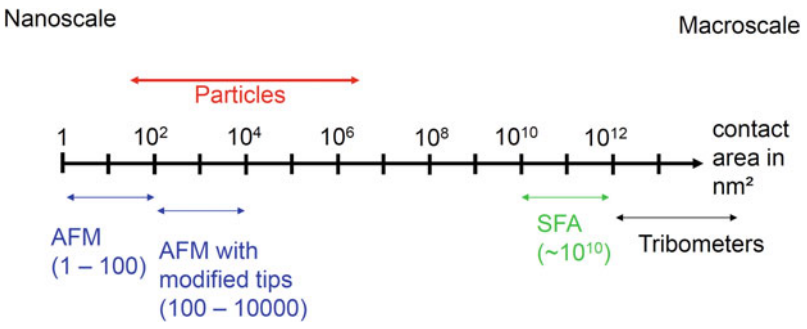


Fig. 17.2 Overview over the different ranges of contact areas covered by tribological techniques on the nano- and meso-scale

standard or modified tips) [1, 6, 7, 10], the surface force apparatus [45, 46], and the the quartz crystal microbalance (QCM) [47–49]. As Fig. 17.2 illustrates, nanoparticle manipulation offers unique access to lengthscales at the transition between the nano- and meso-scale.

17.2 Friction Measurements by Nanoparticle Manipulation: Experimental Approach

Since the sharp tip of an atomic force microscope is an ideal tool to push nanoobjects on a surface, the AFM has become the common basis for all nanoparticle manipulation schemes. For conventional topography measurements, an AFM is typically operated in two main modes: In the *contact* (or static) *mode* tip and sample are in direct mechanical contact. This technique can be used to obtain nanometer resolution images on a wide variety of surfaces. Higher resolution, however, is often achieved using dynamic modes like tapping mode [50], or noncontact mode [51], where the cantilever oscillates near the sample surface. For both the contact and the dynamic mode, it has been shown, that they can be been succesfully applied to facilitate nanoparticle sliding with simultaneous assessment of energy dissipation. The resulting different experimental approaches to nanoparticle manipulation are described in the following subsections.

17.2.1 Dynamic AFM Techniques for Nanoparticle Manipulation

In the dynamic mode, the cantilever is typically oscillated close to its resonance frequency while the oscillation amplitude serves as feedback parameter. The energy that is dissipated during one oscillation cycle ΔE is then a function of cantilever spring constant c_z, the quality factor Q, and the drive and oscillation amplitudes a_d and A, respectively [52]:

$$\Delta E = \pi c_z \left(A\, a_\mathrm{d} \sin(\phi) - \frac{A^2}{Q} \frac{f_\mathrm{d}}{f_0} \right). \tag{17.1}$$

Here, f_d and f_0 are the driving and oscillation frequencies and ϕ is the phase shift between them. Controlled manipulation of latex spheres on highly oriented pyrolythic graphite (HOPG) substrates in the dynamic mode was demonstrated by Ritter et al. [53]. Basically, when the oscillating tip hits the rim of the nanoparticle, energy is transferred causing a lateral movement of the particle. The efficiency of this process is determined, among other factors, by the a priori unknown impact angle (see Fig. 17.3). For controlled manipulation, the tip is placed at the side of the particle and the oscillation amplitude is increased until particle motion is observed.

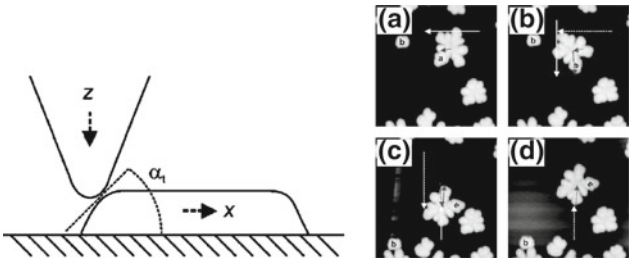


Fig. 17.3 *Left* Sketch of the tip-particle coupling. The impact angle at between tip and antimony particle determines the normal (z) and lateral (x) components of the acting force. *Right* Illustration of the dynamic mode manipulation procedure for Sb on HOPG (image size $1 \times 1\ \mu m^2$). **a** Overview of the particle of interest (labeled with a) and the surrounding area. A *white* and a *gray arrow* indicate the path of the subsequent tip motion and the resulting dislocation of the particle, respectively. **b** Topography after the manipulation, showing a lateral translation of 83 nm and an in-plane rotation of 58°. **c** Result of the second manipulation step, and **d** final result after the third manipulation step (adapted from [34])

The particles can be translated when the power input exceeds a threshold value necessary to overcome the friction force of the adsorbed particle. By changing the amplitude of the dither piezo that drives the cantilever oscillations while the feedback loop is continuously working, it is possible to switch between an imaging mode and a manipulation mode with variable power input into the sample. Thus, an individual adaptation to the sample properties is feasible.

In the case of a free cantilever, increasing the amplitude of the dither piezo leads to an increase of the effective oscillation amplitude, which scales linearly with the excitation. During the manipulation experiments, however, the excitation is increased when the cantilever is still in feedback. The feedback system tries to maintain the preselected setpoint amplitude of the cantilever by decreasing the distance between cantilever and sample. Recording the dither amplitude, the setpoint, and the phase angle allows one to calculate the power dissipation during manipulation by using (17.1). Theoretical analysis shows that this value is in fact a measure of the lateral forces occurring during manipulation [54].

The dynamic mode is characterized by a very high degree of flexibility, since particle motion in arbitrary directions can be performed. Furthermore, the range of lateral forces that can be applied to the particle for translation is related to the square of the oscillation amplitude, which yields a very large dynamic range. The excitation amplitude can indeed be adjusted over orders of magnitude, if necessary, to switch between gentle imaging and manipulation of even the biggest particles.

The above discussion shows that very controlled manipulation of individual particles is possible. For its successful realization, however, it is necessary to have an electronic AFM control system available, such as the one used to perform the particle manipulation shown in Fig. 17.3, where arbitrary tip motions along user-defined trajectories have been carried out under full feedback control. Unfortunately, this is often impossible with commercial AFM systems. A different approach relies on the

statistical movement of a large ensemble of particles of similar size. Mougin et al. [41] and Paolicelli et al. [43, 55] have systematically analyzed the amplitude threshold necessary to induce particle motion in dynamic mode during surface scanning. For gold nanoclusters deposited on.

Similarly, Gnecco et al. report a detailed analysis of particle trajectories due to the impact between the oscillating tip and the particle within one scan frame [40].

Unfortunately, the tapping mode, which has been commonly used for nanoparticle manipulations using tapping mode AFM operation, is not suitable for UHV conditions [51]. Instead, the frequency modulation mode (FM-AFM) must be used under UHV conditions. However, due to its self excitation principle, this mode is very sensitive against perturbations. Therefore, trying to manipulate typical nanoparticles with contact areas larger than a few $100\,nm^2$ will often result in the breakdown of cantilever oscillations, making this mode inapt for nanoparticle manipulations. However, it has been shown how the FM-AFM mode can be applied to move extremely small structures like single atoms [56] or PTCDA-molecules [57] with simultaneous assessment of the forces required to move the atoms or molecules.

17.2.2 Contact Mode AFM Techniques for Nanoparticle Manipulation

17.2.2.1 Pushing Nanoparticles from the Side

As an alternative to nanoparticle manipulation performed in tapping mode, manipulation can be carried out during contact mode operation [38, 39, 58–60]. The two most crucial parameters influencing the manipulation of the particles in contact mode are the stiffness of the cantilever and the component of the tip force exerted during the scan along the surface normal. As the cantilever stiffness is set once a particular cantilever has been chosen, the normal force represents the most important factor in the manipulation process. By increasing or decreasing the normal force, one can switch between imaging and manipulation in a controlled manner: If the normal force is below a certain threshold, the cantilever, which is always scanned with the feedback loop on, follows the topography of the sample. Only if the normal force is above the threshold, the feedback loop does no longer compensate for the height difference of the island, but rather pushes the island along with the cantilever. These two possible ways of interaction between cantilever, island, and substrate are depicted schematically in Fig. 17.4.

This manipulation technique was optimized by scanning with a constant normal force that is very close to the threshold of manipulation. An example for such manipulation is presented in Fig. 17.5, where the island is moved sidewards during one particular scan line. Quantitative information can then be gained from the topography and friction signals acquired during the translation process (Fig. 17.7e). Essentially,

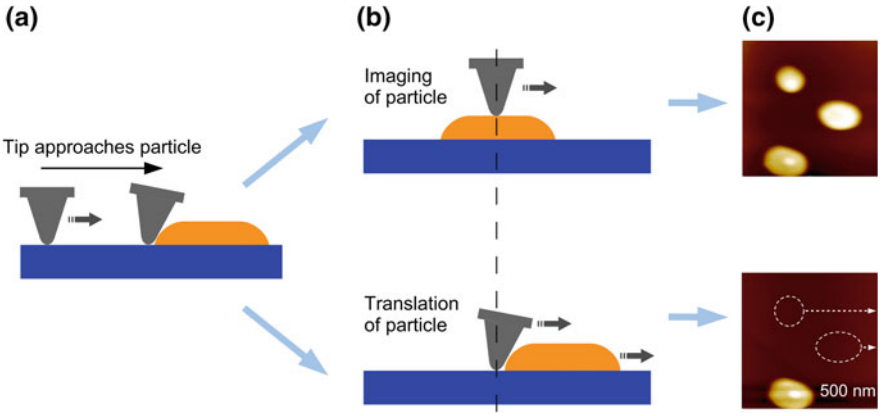


Fig. 17.4 Schematic representation of the two different interaction scenarios that might occur when the tip reaches a particle **a** Tip approaches the particle. In both subsequent scenarios, the tip will experience additional torsion once it reaches the rim of the particle. **b** *Top* if the normal force is below the dislocation threshold, the tip will trace the topography without moving the particle. The cantilevers additional torsion then represents the tip-particle friction on the island. *Bottom* if the dislocation threshold is exceeded, the feedback loop does not follow the topography but instead the tip will start pushing the particle. In this case, an additional lateral force corresponding to the particle-surface friction can be observed by monitoring the cantilever torsion. **c** Topography scans taken before and after the two upper particles were moved to the right

the sudden increase of the lateral force signal during the pushing process represents the friction of the manipulated island.

The main advantages of this manipulation approach is, that it allows to measure friction for a number of particles in a relatively short time, because often several displacements can be observed during on image. But since there is no way to precisely control the tip-particle interaction, it is often difficult to avoid unwanted nanoparticle motion. Thus, translation of a specific nanoparticle in a well defined way is challenging. In order to improve the manipulation procedure, the AFM can be operated in non-contact mode (either constant amplitude mode [51] or constant excitation mode [61, 62]) for recording topography images. From these images, a particle suitable for manipulation can be chosen and the tip can be positioned beside the nanoparticle, before switching back to contact mode. By moving the AFM tip along an straight pathway with a sufficiently high normal force, the nanoparticle can now be displaced, before the AFM is switched back to non-contact mode [60, 63]. The lateral force signal obtained during such a manipulation is found to be very similar to the shape shown in Fig. 17.5e).

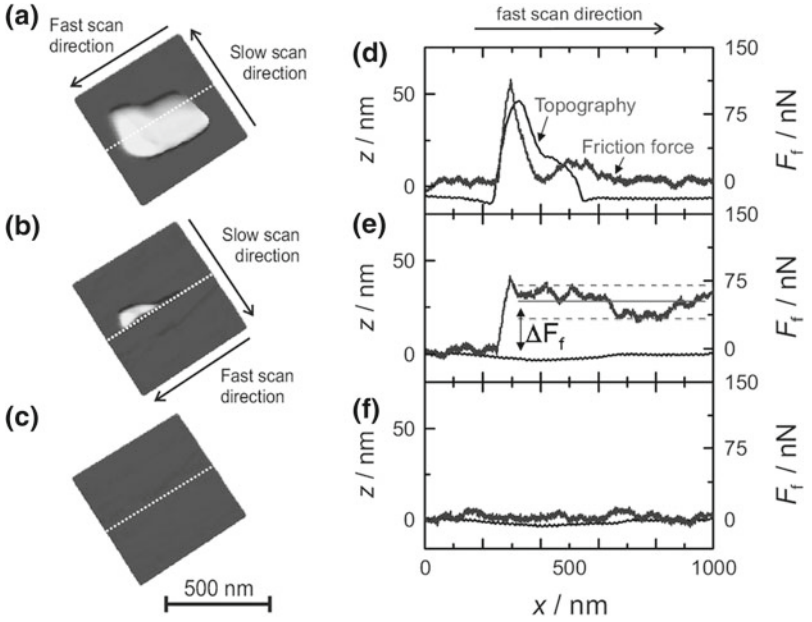


Fig. 17.5 Illustration of the particle manipulation procedure based on scanning with a constant normal force close to manipulation threshold. **a** Imaging of a nanoparticle at low external loading force. **b** A slight increase of the load initiates particle motion. The particle is imaged for several line scans before it is pushed out of the field of view (along the white dotted scan line), thus showing a 'cut' particle. **c** Surface image after translation, confirming that the particle has been moved out of the field of view. **d** Topography (*left axis*) and lateral force (*right axis*) of the last scan line before translation. The lateral force signal is mainly topography-induced, as the cantilever twists at the particle's edges. **e** Scan line during displacement. The topography now reflects the flat graphite surface, while the average frictional resistance of the particle can be determined from the lateral force signal. **f** First scan line after manipulation (adapted from [35])

17.2.2.2 Nanoparticle Trajectories During Manipulation

Ideally, when pushing a nanoparticle from the side, both the tip and the nanoparticle should move the same distance along a straight line. However, the shapes of tip and nanoparticle can lead to force components perpendicular to the tip path [40, 43], which can cause the contact between tip and nanoparticle to break. Once this happens, the AFM tip will continue its path but leave the nanoparticle behind. During nanomanipulation experiments, this effect can be minimized by trying to direct the tip trajectory through the center of mass of the nanoparticle. It can nonetheless be problematic with respect to the accuracy of particle positioning and, more importantly, it can also affect the friction force measured during the particle manipulation. For the simplest case of a round particle and a round tip, the geometrical configuration used to calculate the particle trajectory is depicted in Fig. 17.6. For the calculation of the particle trajectories it is assumed that any dependence of

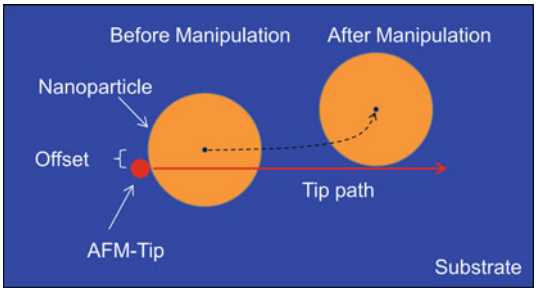


Fig. 17.6 Schematic of geometry during off-center manipulation. The position of the nanoparticle is shown in relation to the AFM-tip and the tip path directly before and after the manipulation takes place. The *dotted line* indicates the particle trajectory during manipulation. The most crucial parameter do describe the manipulation process is the offset between the AFM tip and the center of mass of the nanoparticle measured perpendicular to the tip path

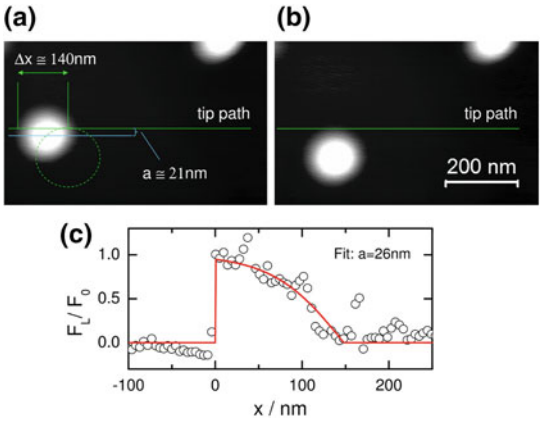


Fig. 17.7 Example of a manipulation event where the tip was moved from left to right and pushed an nanoparticle downwards during manipulation. **a** Topography image of the Sb nanoparticle on HOPG substrate before manipulation, **b** topography image after manipulation. The tip path and the position of the nanoparticle after manipulation are indicated in **(a)** and allow to estimate the offset $a \cong 21$ nm. **c** Lateral force signal measured during the manipulation. A fit to the experimental data yields a $= 26$ nm (image adapted from [63])

friction on sliding direction and sliding velocity can be neglected. Furthermore, the tip radius was assumed to be pointlike. The theoretical equations derived from this configuration [63] have subsequently been used to analyze the friction signal measured for an off center manipulation of an antimony nanoparticle on HOPG (see Fig. 17.7). In this experiment an antimony nanoparticle of about 150 nm diameter was pushed from the side by the AFM tip and after manipulation, a considerable displacement perpendicular to the tip path was found (Fig. 17.7a, b). The corresponding lateral force signal shows the typical steep increase, when the tip hits the nanoparticle, but starts to decrease immediately until the fricton is back to the initial level after

approximately 150 nm, meaning that the tip has lost contact with the nanoparticle. This behaviour can be well fitted by a theoretical friction profile calculated for an offset a of 26 nm between nanoparticle and tip. From the AFM images measured before and after nanoparticle manipulation, the offset a can be estimated to 21 nm. Thus fit parameter and directly measured offset are in good agreement. If the lateral force signal can be described by a theoretical model, as shown in Fig. 17.7, the interfacial friction can still be precisely determined. However, such a calculations become increasingly complicated for more irregularly shaped particles and from a practical point of view, it is therefore preferably to limit any quantitative analysis to particle manipulations with straight trajectories.

In other experiments by Gnecco et al., the continuous off-center manipulation of nanoparticles during imaging has been used to force groups of nanoparticles onto common resulting trajectories, an approach which can be used for arranging nanoparticles on surfaces or for sorting of nanoparticles. In order to do so, Rao et al. have scanned a Si surface covered with a number of round colloidal gold nanoparticles in tapping mode, while the tip sample interaction was chosen to be well above the threshold of manipulation [40]. This way, whenever the AFM tip hits a nanoparticle, the nanoparticle is displaced according to a theory similar to the one used to describe contact mode measurements [40, 64]. Again, the determining factor is the offset between the tip path and the nanoparticles center of mass. Rao et al. have shown, that this parameter can be tuned by the line spacing during imaging and the nanoparticles can thereby be forced onto straight passes, where the effective angle of the trajectories is directly related to the line spacing (see Fig. 17.8).

Recently Nita et al. have applied the manipulation concept developed by Rao et al. to push antimony nanoparticles of complex shape by contact mode AFM techniques on MoS_2 [65], where the shape of the particles led to trajectories far more irregular,

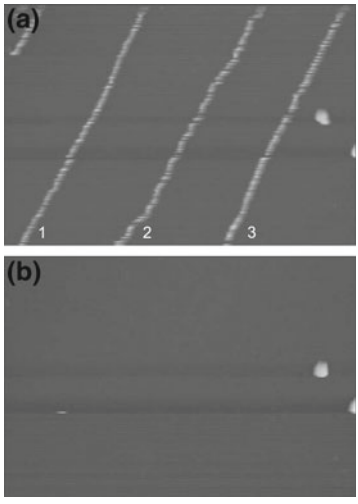


Fig. 17.8 Topography images recorded in tapping mode during forward scan **(a)** and backward scan **(b)**. The forward scan **(a)** shows three parallel trajectories of the gold nanospheres on a Si substrate. No trajectories are found in the backward scan **(b)**, which means that the particle are efficiently pushed from the tip path during the forward scan and are not interacting with the tip during the backward scan (image taken from [64])

than the ones shown in Fig. 17.8a. However, by using the precise particle shape as input for numerical simulations, the particle trajectory could accurately be described and quantitative data for the interfacial friction was extracted [65]. These quantitative results obtained in contact mode hint toward a promising route of extracting friction from nanoparticles manipulations, since simple a imaging procedure, as it is possible with even very basic AFMs, allows to record data, from which friction values can be obtained.

17.2.2.3 'Tip-on-Top'-Approach

A slightly different approach for particle manipulation is realized by placing the tip on top of the particle during manipulation instead of placing it at the side. In this approach, which is illustrated in Fig. 17.9, the tip is first positioned approximately in the center of the nanoparticle's top surface. If then tip motion is initiated, two scenarios are possible: (1) The tip slides over the surface of the nanoparticle with the lateral force signal reflecting the friction between tip and nanoparticle, or (2) the nanoparticle is following the tip motion by gliding over the substrate. In this case, which we will refer to as the 'tip-on-top' manipulation mode, the measured torsional signal is directly proportional to the interfacial friction between particle and substrate.

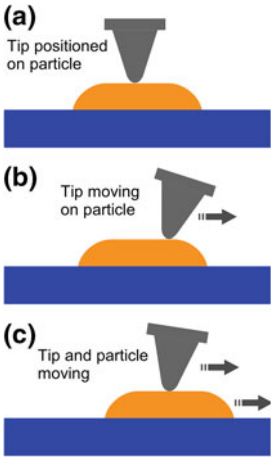


Fig. 17.9 Figure illustrating nanoparticle manipulation by employing the 'tip-on-top' approach described in the text. **a** *Top* The tip is positioned on top of a particle (starting position). **b** If motion of the cantilever is initiated and the cantilever normal force is below a certain threshold value, the cantilever is sliding on the particle, profiling the nanoparticle's top surface. **c** If the cantilever load is above the threshold, the tip remains on a fixed position on top of the particle and tip and particle will move together

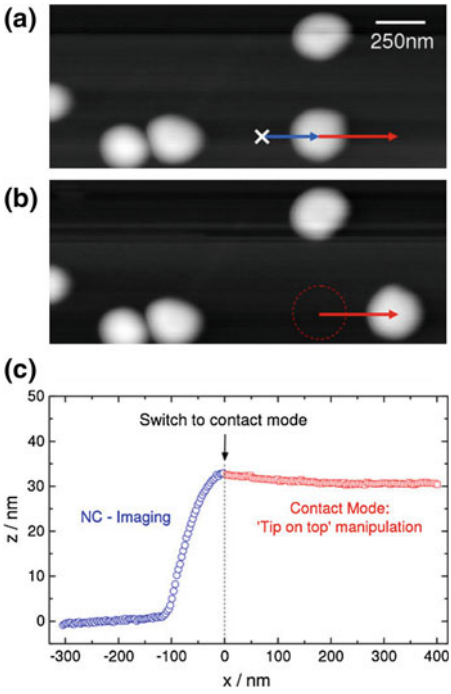


Fig. 17.10 Example of a controlled manipulation performed in the 'tip-on-top' mode. **a** Non-contact topography image before manipulation. The *cross* indicates the initial cantilever position, whereas the two *arrows* mark the paths used to position the cantilever on top of the particle and to perform the manipulation. **b** Non-contact topography image after the nanoparticle manipulation along the vector path. **c** Topography signal measured during the two vector pathways. First the tip is positioned on top of the nanoparticle in noncontact mode ($x \leq 0$) and subsequently the manipulation is done in contact mode ($x \geq 0$) (figure adapted from [63])

The crucial parameter that distinguishes between the two scenarios is the ratio of the shear forces in the tip-particle-substrate system. Only if the lateral force needed to shear the tip-particle interface is larger than the force required to shear the particle-substrate interface, the particle moves together with the tip. If this is the case nanoparticles can be moved over large distances while the tip is placed on top of them. An example is given in Fig. 17.10e, where an antimony particle has been displaced under UHV conditions in the 'tip-on-top' mode.

One strategy to perform nanoparticle manipulations using the tip on top mode is to first operate the AFM in non-contact mode and placing the tip either on the left or right hand side of the chosen nanoparticle (Fig. 17.7a, position marked by the cross). Then the tip is first scanned across half of the particle (as indicated in Fig. 17.7a), placing it directly on top of the nanoparticle (Fig. 17.10a, c for x = 0). At this position the AFM is switched in situ from non-contact to contact mode [60, 63] and the cantilever normal force is slowly increased, allowing to exert a sufficient lateral

force for moving the particle. Once the normal load has been set, the tip is moved along a second vector (indicated in Fig. 17.10a) to perform the controlled nanoparticle movement. During the manipulation the topography signal and the lateral force signal are recorded. Given a flat substrate surface the topography signal remains flat over the whole pathway of the manipulation (Fig. 17.10c, $x \leq 0$), as long as the tip remains firmly on top of the nanoparticle. After the particle movement is completed, the AFM is switched back from contact to non-contact mode and a control image is recorded (Fig. 17.10b) verifying the nanoparticle's manipulation path. If the normal load is sufficient, we find that nanoparticle firmly follows the tip movements allowing controlled long-distance manipulations of over 1 μm [44].

If quantitative values for interfacial friction are to be extracted from 'tip-on-top' manipulations, one has to keep in mind that in this case a single one-directional nanoparticle manipulation lacks an absolute reference level necessary to quantify the interfacial friction. Therefore, quantitative friction data must be extracted from the forward *and* backward motion of the nanoparticle, a procedure similar to recording friction loops in conventional friction force microscopy (see Sect. 17.2.3).

17.2.3 Identifying Static Friction in Nanoparticle Manipulation Experiments

One important aspect in attaining a coherent picture of friction processes at the nanoscale is the correlation between static and sliding friction. While this difference is a well known fact for friction experiments on the macro-scale, it is less clear if or how this concept is applicable to nano- or mesoscale contacts, where stick-slip motion is considered to be the dominant process. Currently, interest is especially spurred by new concepts that take ageing of nanocontacts into account [16, 66, 67]. In many cases, nanotribological ageing effects can principally be analyzed by velocity dependent measurements [13, 16], since the stick phase during stick-slip motion can be considered as a hold time, during which contact ageing can occur. However, velocity-dependent measurements only allow to vary the hold times in a certain range. To achieve longer hold times with eventually saturated contact ageing, it might be required to suspend the sliding motion alltogether and reinitiate it after the desired amount of time.

Again, the well-defined interfaces between nanoparticles and substrates can form ideal model systems to undertand nanoscale processes related to static friction. To measure the static friction of nanoparticles, different strategies have successfully been employed. One of the first examples was presented by Luethi et al., where the C_{60} nanoparticle on MoS_2 was pushed from the side and the torsion of the cantilever interacting with the nanoparticle was used as a measure of interfacial friction (see Fig. 17.4b). In this case, the static friction resulted in a high cantilever torsion that was built up right before the nanoparticle started moving [58]. Once the particle was sliding steadily, a reduced lateral force signal was measured. However, the

steep increase of the lateral force signal when the cantilever hits the particle can make it difficult to identify the exact maximum, which can be interpreted as static friction, especially if typical point densities for data acquisition are used. Moreover, the process of contact formation between the tip and the nanoparticle can influence the measurement, resulting in unreliable information about static friction.

Tripathi et al. have used tapping mode manipulation techniques to assess the temperature dependence of static friction for gold nanoclusters on HOPG [55]. By measuring the threshold of amplitude reduction, at which detachment occurs, they could quantify static friction and found that the detachment of small gold nanoclusters with diameters of about 27 nm can be described as a thermally activated process, resulting in lower detachment energies measured at higher temperatures [55].

Another possible approach to distinguish between static and sliding friction is based on the 'tip on top' manipulation scheme [44]. With the tip resting on top of the nanoparticles, there are two possible modes of tip and cantilever movement (see Fig. 17.9): The tip can either move on top of the nanoparticles, in which case the lateral force signal represents the friction between tip and particle, or the contact between tip and particle is firm, in which case the tip drags the particle along and the lateral force signal represents the friction between particle and substrate. The key parameter to control the sliding behavior is the cantilever normal force, which can be used to facilitate the transition between static and sliding friction. The principle scheme is depicted in figure Fig. 17.11 for an Sb nanoparticles on HOPG: First the tip is positioned on top of the nanoparticle and a contact mode scan of a small area (typically $A = 20\,nm^2$) in the center of the particle is initiated. This scan starts at a low cantilever load, in which case the lateral force between tip and particle is not sufficient to overcome the static friction between particle and substrate. By gradually increasing the cantilever load, the friction between tip and nanoparticle will increase (Fig. 17.11c), until the lateral force is finally sufficient to overcome the static friction of the nanoparticles ($t = 1.85$ in Fig. 17.11c). Now, the nanoparticles is moving together with the tip and the friction level in Fig. 17.11c (red part) represents the sliding friction of the nanoparticles, whereas the maximum of friction (at $t = 1.85$, blue curve) can be interpreted as static friction of the nanoparticles. From Fig. 17.11c it can be seen, that in case of the moving nanoparticles no further load dependence can be observed, which can be understood by assuming that the nanoparticles adhesion is much larger than any applied cantilever load. The transition of dynamic states becomes also obvious from Fig. 17.11b. Right before the transition, the friction loop is fairly wide, while the related topography signal has a considerable slope related to the shape of the particle, which is not flat. After the transition, however, the friction loop is not only significantly narrower, but also the slope in topography has vanished, since now the topography signal is related to the nanoparticle sliding on the flat HOPG substrate.

It was shown, that the transition from static to sliding friction is reproducible, meaning, that the transition from static to sliding friciton can be repeated several times for the same nanoparticles. Interestingly, also the ratio between static and sliding friction seemed to be constant for several particle of the same size and a typical ratio of $F_{\text{sliding}}/F_{\text{static}} \cong 0.5$ is measured.

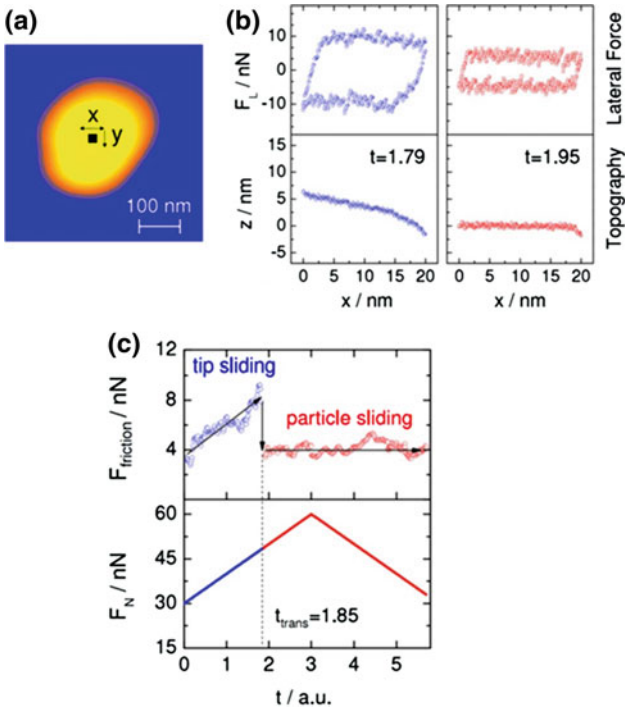


Fig. 17.11 Distigushing static and sliding friction using the 'tip on top' approach. **a** Sb nanoparticle ob HOPG substrate, where the typical scan area is indicated. **b** Friction loops (*top*) and topography (*bottom*) just before (*left hand panels*) and after (*right hand panels*) particle sliding was initiated. **c** Plots of the effective friction force F_{friction} and cantilever normal force F_{N} as a function of the time, while continuously scanning the AFM tip on top of the nanoparticle ($A_{\text{contact}} = 68,000\,\text{nm}^2$) with a scan range of 20 nm × 20 nm. The sudden drop at t_{trans} indicates the transition to particle sliding. (Figure taken from [44])

So far, the exact reason for the observed difference between static and sliding friciton remains unclear. Due to the reporducibility of effects for the same nanoparticle, any wear related interface changes can be ruled out. However, theory predicts that the behavior of a layer of mobile molecules trapped between the moving surfaces can dominate the phenomenon of static and kinetic friction due to a shear force induced transition from a solidlike to a liquidlike structure of the interface layer [68]. Persson found that the ratio between kinetic and static friction is consistently one half for a wide variety of simulation parameters [68] in good agreement with the experimental results. The model also predicts that if the lateral force is reduced again, the liquidlike state prevails until much below the initial threshold, giving rise to a hysteretic behavior as observed experimentally [44]. Nonetheless, in our UHV experiments it is difficult to imagine the presence of a layer of additional interface molecules. Still, it is astonishing that the model from Persson is very consistent with the experimental observations, including a kinetic/static ratio of one half and hysteretic behavior.

This suggests that if no contamination particles are present, the last layer of Sb atoms in contact with the HOPG substrate might act as a de facto boundary lubrication layer.

17.2.4 Comparison of Manipulation Strategies

The various manipulation schemes introduced above have different advantages and drawbacks. In manipulations based on dynamic AFM modes, a high range of forces can be applied to the particles by simply adjusting the oscillation amplitude, and arbitrary translation paths for the particles can be chosen. On the downside, a direct measurement of frictional force is not possible; instead, the momentum transfer to the particle is quantified through monitoring the system's energy dissipation during manipulation. Here, it often remains unclear how much of this energy is actually transferred into the nanoparticle motion.

In contact mode manipulation, on the other hand, the fixed lateral spring constant of the specific cantilever used significantly limits the range of frictional forces that can be accurately detected. However, the restriction might be outweighed by the ability to measure the frictional force between particle and substrate directly, allowing a straightforward quantitative analysis of interfacial friction. The manipulation pathways are more limited, since friction can only be measured perpendicular to the cantilever, but in principle static and dynamic friction can be distinguished. When placing the tip on top of the particles the manipulation control is even better, since a fixed contact between nanoparticle and tip exists. In this configuration, it is possible to perform multiple nanoparticle manipulations without braking the contact between tip and sample. First experiments showed that up to 100 consecutive manipulations are possible. This opens the door for measurements regarding two fundamental key parameters in nanotribology, namely load and velocity. For load dependent measurements, the cantilever can be used to exert a varying normal load on the particle during sliding. This would allow to measure the true load dependence of friction, an issue of considerable fundamental interest. Furthermore, it is possible to vary the sliding velocity of the nanoparticle within an uninterrupted series of nanoparticle manipulations. Such experiments might clarify, how the basic model of thermal activation can be transferred from small contact areas of AFM tips to the extended contacts of nanoparticles.

17.3 Nanoparticles for Manipulation Experiments

Of course, for any nanoparticle manipulation experiments, the nanoparticles itself are of paramount importance. In principle, any mechanically stable particle that has been transferred onto a flat substrate can be investigated by AFM-based manipulation techniques. In order to investigate the frictional properties of very clean interfaces, however, it is mandatory to prepare nanoparticles under ultrahigh vacuum conditions

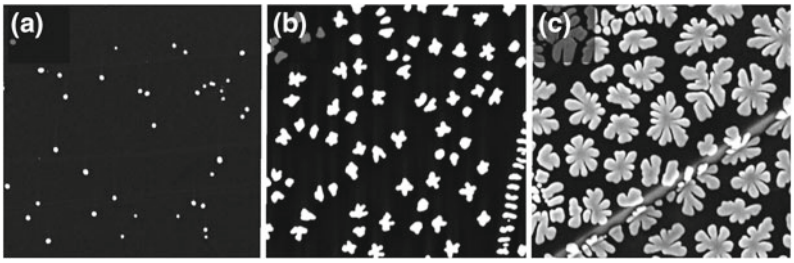


Fig. 17.12 Scanning electron microscopy images of the sample surface morphology after deposition of the equivalent of **a** 2 monolayers (ML), **b** 10 ML, and **c** 40 ML of antimony on a HOPG(0001) substrate surface at a deposition rate of 0.1 nm/s. Image size is $3.6 \times 3.6\ \mu m^2$ in all cases (*images* courtesy of B. Stegemann, HTW-Berlin)

(UHV) and transfer those samples to an UHV-AFM without breaking the vacuum. One approach that achieves this goal is the in-situ thermal evaporation of metals onto a flat substrate.

An example is shown in Fig. 17.12, representing metallic antimony particles grown by thermal evaporation on highly oriented pyrolytic graphite (HOPG). The precise growth parameters and their atomic structure have been studied in detail before [69]. It was found that small, round shaped particles are amorphous, whereas particles with diameters larger than ≈60 nm can already be crystalline (note that the exact characteristic particle size where the transition between the two states occurs depends on the exact evaporation parameters, e.g. surface temperature or evaporation rate [70]). Since it is expected that nanoscale friction is governed by the atomistic structure of the particle-substrate interface, the occurrence of this structural transition allows for a unique study case. And indeed, Ritter et al. have recently reported on two distinct shear stresses observed during nanomanipulation experiments on antimony nanoparticles on HOPG under ambient conditions. For larger particles a shear stress three times as high as for smaller particles was found, with the contact area of transition ($A_{transition} \cong 20{,}000\,nm^2$) coinciding with the structural transition from compact to more branched particles [71].

Recent theoretical studies also suggest that in the case of crystalline nanoparticles, the particle orientation can be of considerable importance for the frictional behaviour of nanoparticles [33], especially due to the occurrence of commensurate and incommensurate otientation between nanoparticle and substrate. But although antimony nanoparticles become crystalline, their large size and branched structure make them inappropriate for systematic analysis of orientation dependence. A far more suitable model system is formed by gold nanoparticles on HOPG. Similar to antimony nanoparticles, gold particles can also be grown by thermal evaporation under UHV conditions, resulting in particles with clean and well defined interfaces. The crystalline structure of such gold nanoparticles is directly obvious from SEM measurements (Fig. 17.13). Topography images of the nanoparticles thus allow to directly determine the orientation of the particles, while the orientation of the

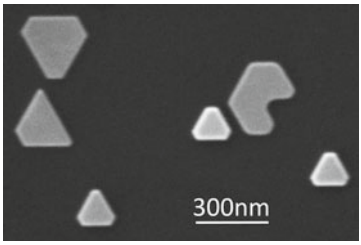


Fig. 17.13 Scanning electron microscopy image of a selection of gold nanoparticles evaporated onto freshly cleaved HOPG under UHV conditions. The evaporation time was 2 min with the gold filled crucible heated to 1380 °C while the HOPG substrate was kept at room temperature

substrate can be determined from atomically resolved stick-slip measurements in direct vicinity of the nanoparticles. The additional option to rotate the nanoparticles by off-center manipulations makes this sample system an ideal candidate to systematically analyze the directional dependence of interfacial friction in case of crystalline interfaces.

Despite the potential for the analysis of fundamental effects in interfacial friciton, the approach of particle preparation by thermal evaporation is rather limited. Chemical methods can yield a much higher variety of different shapes, sizes, and surface functionalization. For example, gold particles can be prepared in very wide range of geometries, including spheres, rods, and even star-like shapes (see Fig. 17.14). Furthermore, they can be coated with self-assembled monolayers terminated with hydrophobic (e.g., methyl, -CH_3) or hydrophilic groups (e.g., hydroxyl, -OH). This allows to study the influence of the hydrophobicity of the coatings on the mobility of the nanoparticles [41]. Tranvouez et al. have studied ligand-capped cadmium selenide nanorods deposited on HOPG by AFM manipulation techniques and found a distinct anisotropy in nanoparticle movement, depending whether the rod was moved parallel or perpendicular to its main axis. These observations could then be

Fig. 17.14 Gold nanoparticles with different shapes suitable for manipulation experiments: *Spheres* (image size 280 × 300 nm^2), *rods* (image size 370 × 390 nm^2); and even complex geometries like stars can be prepared by chemical methods (image size 630 × 650 nm^2, image courtesy of Karine Mougin, Institute de Science des Matériaux de Mulhouse, CNRS-LRC)

linked to the alignment between the organic ligands surrounding the nanorod and the substrate [42].

In addition to these ways of preparing nanoparticles for manipulation experiments, recent publications have shown that nanostructures suitable for manipulation experiments can also be gained from a variety of other inventive experimental strategies. For example, Bombis et al. have shown that after evaporation of NaCl on Cu(111), NaCl nanoparticles suitable for manipulation can be cut from larger NaCl structures [72]. This approach is especially interesting, since size and shape of such nanoparticles can in principle directly be controlled. Additionally, Feng et al. have reported on the mobility of graphite flakes on graphene, where the flakes have been prepared by H_2O assisted cracking of graphene sheets [73]. The results confirmed the strong dependence of interfacial friction on the relative orientation between flake and substrate, as previously reported by Dienwiebel et al. for the case of a graphite flake trapped between AFM-tip and HOPG substrate [29].

17.4 Friction of Extended Nanocontacts: Theoretical Concepts

Since the advent of friction force microscopy in 1987, most research in the field of nanotribology has concentrated on the analysis of very small contact areas, so-called nano-asperities. The interest to focus on nano-asperities was spurred because any real surface can be described as a complex system of multiple asperities, which was first recognized in the 1950's by Bowden and Tabor [74]. Experimentally, the contact between an AFM tip and a surface can be considered an ideal model system for such single asperities. In this configuration, friction force microscopy allows to analyze the basic friction processes of point contacts on the atomic scale and in most cases, the experimental results could be explained by theoretical models, which reduced the contact to only a few atoms or even a single atom [10, 13, 16]. Considerably less research has, however, been done on extendend nanocontacts, even though they represent the most frequent building block of realistic surfaces in most current theories describing the contact mechanics of realistic surfaces. As a consequence, a lot of aspects of frictional behavior of extended nanocontacts are still not yet fully explored. One of the most fundamental - yet unresolved - question in current nanotribology concerns the question of how the frictional force F_{friction} experienced at a finite, atomically flat interface of nanoscopic dimensions scales with the actual contact area A_{contact}. The answer might even affect our understanding of the widely accepted classical friction laws of Amontons, who stated that friction is proportional to the normal force, but independent of the apparant contact area:

$$F_{\text{friction}} = \mu F_{\text{load}}. \tag{17.2}$$

Here F_{load} represents the external loading force and μ the friction coefficient, which depends only on the actual combination of materials in contact. Only under the assumption of a linear dependence between true contact area and friction, this

law can be understood by the commonly acknowledged model first introduced by Greenwood and Williamson [75], where a linear dependence between the real contact area between two surfaces and the applied load on the interface is assumed.

However, not only the unresolved questions regarding contact area dependence of friction make extended nanocontacts an intriguing problem. Another aspect of fundamental interest is the fact that in contrast to atomic point contacts where mostly the interaction of single atoms is considered, the collective behaviour of a multitude of atoms *within* the two surfaces sliding relative to each other can be of significant importance. The relevance of such collective behaviour becomes most obvious in the case of an effect called 'structural lubricity', which, as described below in detail, depends on the degree of interlocking between the atoms of two extended surfaces.

One key parameter determining interfacial friction between two sliders is the ratio of lattice constants. Let us first consider the case of incommensurate lattices, i.e., the lattice constants do not match. In this case, if two flat surfaces move relative to each other, on average, for every asperity or atom going up a ramp, another one is going down. As a consequence, the mean friction between rigid surfaces must vanish unless they happen to have the same periodicity and alignment. In this case, all atoms have to go up or down at the same time, resulting in a high effective energy barrier that is scaling proportional to the contact area [76]. Basically, elastic deformations of the surfaces might alter this behaviour, but detailed calculations show that elastic deformations are generally too small to modify these general conclusions [24, 77–79]. The effect of low friction due to non matching interfaces has originally been termed '*superlubricity*' [24, 80]; however, as it is a purely structural effect and to distinguish it from other effects that may lower the interfacial friction, it has been suggested by Müser to denote it more adequately as '*structural lubricity*' [25].

Figure 17.15 illustrates the situation from slightly different viewpoint. It shows the surface atoms of a substrate and the surface potential they cause. If only one atom was placed in this surface potential (Fig. 17.15***I***), it drops into a deep minimum and a huge energy barrier has to be overcome to displace it by one lattice constant a.

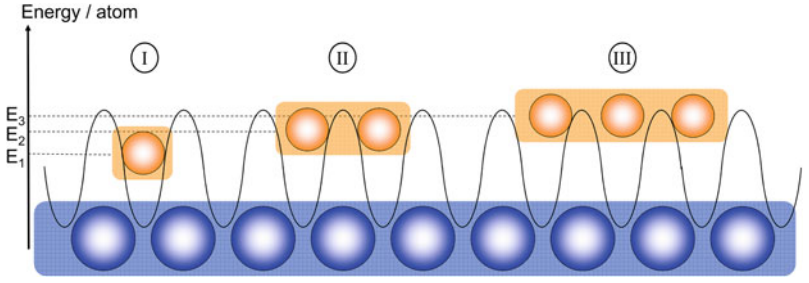


Fig. 17.15 Figure illustrating the effect of incommensurability onthe average barrier between potential minima: While the number of atoms that have to overcome a barrier increases from one to three for (***I***) to (***III***), the height of the individual barriers shrinks significantly. For increasingly larger contacts, the effective overall barrier height will approach zero even though a large number of atoms contributes to the frictional resistance

However, if we placed *two* atoms in a fixed structural relation with each other on the surface, featuring a lattice constant $b \neq a$, the energy barrier that has to be overcome by each individual atom to move the entire two-atom cluster by a has shrunk considerably (Fig. 17.15*II*). This principle continues for increasing contact sizes (see Fig. 17.15*III*) and ultimately results in a vanishing energy barrier and therefore ultra-low friction as long as substrate and slider have incommensurate lattices or feature disordered (amorphous) surfaces. While it is difficult to give an analytic expression describing the case of incommensurate lattices, one can show that the static friction F_{static} between dry, amorphous, and flat surfaces depends on the contact area A with

$$F_{\text{static}} \propto F_{\text{load}}/\sqrt{A_{\text{contact}}} \tag{17.3}$$

due to the averaging effect of probalbility theory's central limit theorem [76].

In case of an amorphous interface structure, the contact area A_{contact} is the *only* parameter determining the scaling of interfacial friction, meaning that the interfacial friction is not affected by altering the particle's shape or orientation as long as the contact area remains constant. For crystalline interfaces, however, the situation is more complex. In recent theoretical studies, deWijn [33] has analytically calculated the friction for the case of triangular nanocrystals on a hexagonal substrate, a configuration suitable to describe, e.g., gold nanoparticles on HOPG. It is found that the scaling of friction with contact area sensitively depends on the particle's shape and orientation. For triangular particles with non-matching lattice constants, incommensurate and pseudo-commensurate orientations have been identified, which result in different power laws describing the friction versus contact area. While the friction scales $F_{\text{friction}} \propto A^0$ for incommensurate orientations, an increase of friction described by $F_{\text{friction}} \propto A^{0.5}$ is found for the pseudo-commensurate case. If irregularly shaped nanoparticles are considered, the powers can increase significantly, e.g. from 0 to 0.25 for incommensurate orientations, indicating that an irregular particle shape effectively weakens the efficiency of structural lubricity.

Curently, many aspects of structural lubricity are still under debate, such as the exact circumstances under which a superlubric state can actually be established. However, a growing number of experimental studies already seems to corroborate its existence [26–29, 35]. The results from [29] show a significant increase of the friction for relative angles between a tip and a graphite surface of about 0° and 60°, but ultra-low friction is found for all other angles. In order to explain their results, the authors assumed that the tip has picked up a small graphite flake during scanning. This flake would then be commensurate with the underlying lattice for sliding angles of 0° and 60°, but incommensurate for all angles in between.

Nonetheless, a structure-induced superlubric state still seems to be more exotic than widespread. Not only that FFM experiments have been reported where the friction-load dependence appears to be linear, but also macroscopically measured friction coefficients are always substantial and do not vanish with increasing contact area. What could be the reason that structural lubricity is so difficult to observe? According to a suggestion by He et al. [81] the problem is that realistic contacts

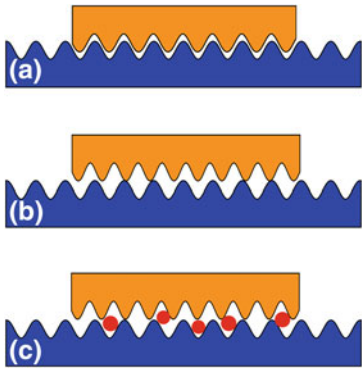


Fig. 17.16 Figure illustrating the effect of contamination on the resulting friction coefficient. **a** The atomic structure of two commensurate surfaces in contact can interlock, resulting in an area-independent friction coefficient. **b** Structural lubricity: Two incommensurate, atomically flat surfaces. The barrier between stable potential minima, and thus the friction coefficient, decreases with increasing contact size. **c** If the contact in b) is contaminated with mobile interface molecules, the friction coefficient is again independent of the contact size. In this case, the mobile molecules can always lock at suitable potential minima, acting as molecular mediators between the incommensurate structures

almost always feature adsorbed, but mobile molecules that are trapped between the sliding surfaces. Examples for such molecules are small hydrocarbon or water molecules that adhere on any surface (this effect can be only prevented by ultrahigh vacuum conditions). Often referred to as 'dirt particles', these molecules prevent a direct interaction of the surface potentials of the sliding interfaces by acting as spacers. Since their mobility allows them to simultaneously lock at surface potential minima for both sliders, an area-independent friction coefficient is obtained for any surface geometry (see Fig. 17.16). Due to the ubiquity of contamination, this effect even represents the assumed reason for the linear scaling between contact area and friction and thus for the apparent universality of Amontons' law [76].

Another reason, that may prevent systems displaying structural lubricity, even under very clean interface conditions, may be the nature of interaction between the interfaces. If, e.g., strong chemical bonds between the interfaces are possible, the interfacial friction should depend mainly on the number of bonds. Szlufarska et al. have used MD simulations to analyze this szenario for amorphous carbon tips on diamond substrates (both terminated with hydrogen) [82] and could determine for all Si atoms at the interface whether a bond to the substrate was formed or not, which was dependent on the proximity between Si atom and substrate. In this case, the contact area is defined by the region in which bonds are build and the number of bonds determines the friction. In case of an atomically flat contact, friction should then scale proportional to the contact area with no reduction of friction related to structural lubrication effects.

When analyzing the friction of nanoparticles, it has also to be considered whether the contact area of the nanoparticle can really be considered to be completely rigid.

In a rigid configuration, a lot of fundamental characteristics regarding commensurate or incommensurate interfaces can be predicted. However, elastic deformations might change the particle's behaviour. For example, Reguzzoni et al. have used MD simulations, to describe the onset of a slip process in a Xe film on a Cu substrate [83], a system where a commensurate contact can be assumed. If an external load is applied to this system, this can result in a frictional slip that originates from the nucleation of a small commensurate domain. By this simulations, Reguzzoni et al. could understand QCM experiments results obtained by Coffey et al., where particle movement was observed in spite of an energy barrier that was definetely too high to be overcome in the absence of the nucleation of structually distinct domains [49].

17.5 Frictional Duality of Sliding Nanoparticles

17.5.1 Contact Area Dependence of Friction Analyzed by Nanoparticle Manipulation

The above-raised fundamental questions related to the area dependence of frictional forces are ideally addressed by particle manipulation. Since crystalline or amorphous particles of different sizes can be prepared on atomically flat surfaces, a systematic assessment of friction as a function contact area is possible. This strategy was followed in two recent studies where the friction of metallic antimony particles on a flat graphite surface was studied [34, 35].

In the work of Ritter et al. [34] particles were investigated under ambient conditions at room temperature using dynamic mode AFM manipulation. In these experiments the energy dissipation at the manipulation threshold was analyzed as a function of the contact area between particle and substrate. The results for antimony particles on two different substrates, HOPG (triangles) and MoS_2 (circles), are shown in Fig. 17.17. The contact areas of the various particles range from 10,000 to 110,000 nm^2. The data displayed in Fig. 17.17 implies a linear relationship for the dependence of the dissipated power P_{tip} (which is equivalent to ΔE in (17.1)) on the particle's contact area with the substrate.

As outlined above, theoretical considerations suggest that for crystalline, non-commensurate surfaces in contact, friction should scale with the square root of the contact area, which is contradictory to the above results. Furthermore, an offset in the friction-size relation is observed in Fig. 17.17, which remains puzzling. However, one has to consider that there might be an influence of particle geometry on the momentum transfer of the oscillating tip to the particles. Further, these measurements were done under ambient conditions, and surface contamination may play an important role, as discussed in the previous section (see Fig. 17.16c). Therefore, we turn in the following to another set of experiments, which were performed under ultrahigh vacuum conditions to ensure clean interfaces [35]. Additionally, the experiments were carried out using the contact mode manipulation method. This yields directly

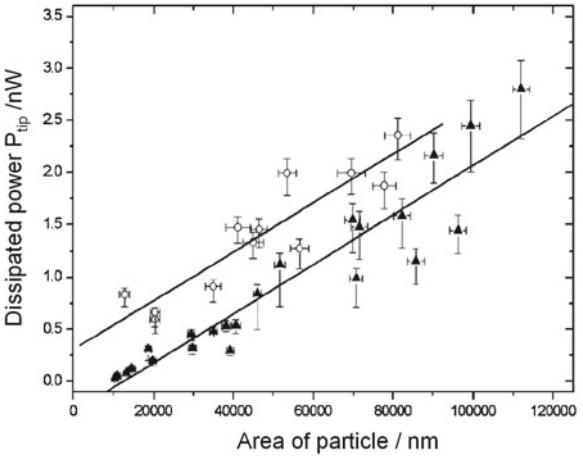


Fig. 17.17 Plot of the minimum values of power dissipation needed for translation of differently-sized Sb nanoparticles on HOPG (*filled triangles*) and MoS_2 (*empty circles*), respectively. The threshold values for both substrates are in the same range and scale linearly with the contact area of the translated particles. The *straight lines* represent linear fits of the measured data. The heights of the translated particles had an average value of 26.2 nm for the 23 particles moved on HOPG and 21.5 nm for the 12 particles moved on MoS_2 (adapted from [34])

quantitative friction values, thus avoiding possible ambiguities from the unknown impact angle in dynamic AFM manipulation.

The results of several individual dislocation events using particles featuring contact areas of up to 200,000 nm^2 are presented in Fig. 17.18. These events can be categorized in two distinct regimes: While the majority featured substantial frictional resistance (regime 1; solid symbols), about 1/4 of the events showed almost no detectable friction (regime 2; open symbols), causing an apparent 'frictional duality'. The results in regime 1 suggest a linear dependence and a constant shear stress $\tau = F_l/A = (1.04 \pm 0.06)$ MPa. Since the normal force experienced by the particles is due to adhesion, which scales linearly with area, an area-independent friction coefficient follows, reinforcing Amontons' law at the nanoscale.

But what is the reason for the unexpected vanishing friction behaviour of one quarter of the particles? First, we consider two possible artifacts to account for the occurrence of vanishing friction: Those particles may be picked up by the tip during translation or they may be stuck on a graphite flake that slides superlubric [29]. The pick-up hypothesis can be discarded since frictionless displaced particles could still be imaged directly after translation (see Fig. 17.19). The case of a graphite flake stuck underneath the particle is also unlikely since the images recorded after particle manipulation show no sign of missing graphite flakes. Thus, the vanishing friction behaviour must be correlated to the physical properties of the particles and/or the interface.

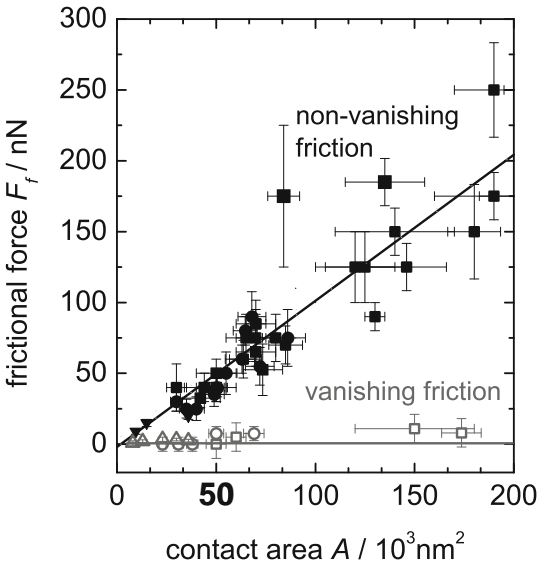


Fig. 17.18 Contact area dependence of the lateral forces experienced during nanoparticle manipulation in ultrahigh vacuum. The lateral force component of the total force is plotted versus the previously determined size of the manipulated islands, which corresponds directly to the contact area between island and substrate. Two regimes are visible: While the majority featured substantial frictional resistance (regime 1; *solid symbols*), about 1/4 of the events showed almost no detectable friction (regime 2; *open symbols*), causing an apparent 'frictional duality' (*Figure* taken from [35])

Close observation of the island structure by AFM imaging shows no systematic correlation of particle features (e.g., degree of ramification, structure, or height) with the occurrence of vanishing friction. Indeed, islands of comparable size and shape can show completely different frictional behavior within one scan frame. Also, the friction measurements show a high degree of reproducibility during multiple translation of the same particle (cf. Fig. 8 in [60]). This suggests that the properties of the particle-surface interface are decisive for the observed duality.

So what about the atomic structure of the interface? If the particles are crystalline and exhibit well-ordered, crystalline interfaces, the observation of finite friction (commensurate interfaces) or vanishing friction (incommensurate interfaces) is expected [27, 29, 80] (note that the occurrence of commensurability and incommensurability also depends on the orientation of the particle lattice relative to the substrate). The compact shape of most of the particles, however, suggests them being amorphous [69]. And even if some of the particles were crystalline, the atomic lattices of Sb and HOPG do not match, i.e., the interfaces are incommensurate under all circumstances. As a result, superlubric behavior should prevail, which is in contradiction to the observed duality.

Thus, let us recall that theoretical investigations predict superlubric behaviour to break down if small amounts of mobile molecules are trapped between the sliding

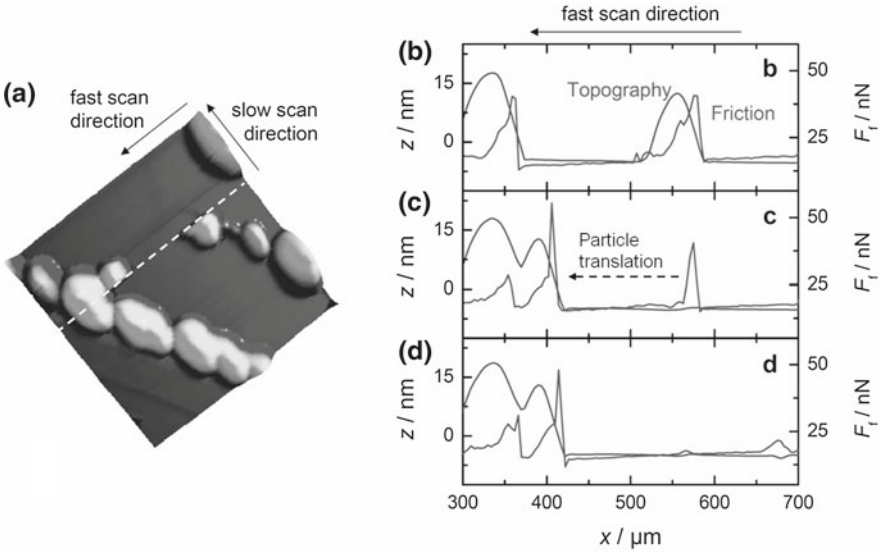


Fig. 17.19 **a** Translation event of an Sb nanoparticle with a contact area of 8 000 nm^2 (topography signal). **b–d** The corresponding scan lines just before **(b)**, during **(c)**, and right after **(d)** the translation of the particle. In contrast to Fig. 17.5e, the friction signal here only shows a peak where the tip hits the island at its initial position ($x = 580$ nm) and remains flat afterwards (below 1 nN) until the island reaches its new resting position (*Image* taken from [35])

surfaces [76, 81]. Those molecules act as mediators between the top and bottom atomic lattices and an area independent friction coefficient was found in simulations[76]. This effect seems not to depend strongly on the level of contamination [81]. In the experiments of Dietzel et al. [35] even under 'clean' UHV conditions, a fair number of such mobile adsorbates can accumulate on HOPG surfaces over extended measurement times. One could imagine that the interfaces of some particles are atomically clean, while others experience a break-down of superlubricity due to contamination, which would explain the observed duality of nanoparticle friction.

17.5.2 The Role of Interface Contaminations: Theoretical Calculations

Although the universal presence of interface contamination seems to be a very plausible explanation for the observed duality in friction of nanoparticles, the question remains if conceivable concentrations of likely contaminants are really sufficient to quantitatively explain the friction levels observed in nanomanipulation experiments. In order to clarify this question, Brndiar et al. used density functional (DFT)

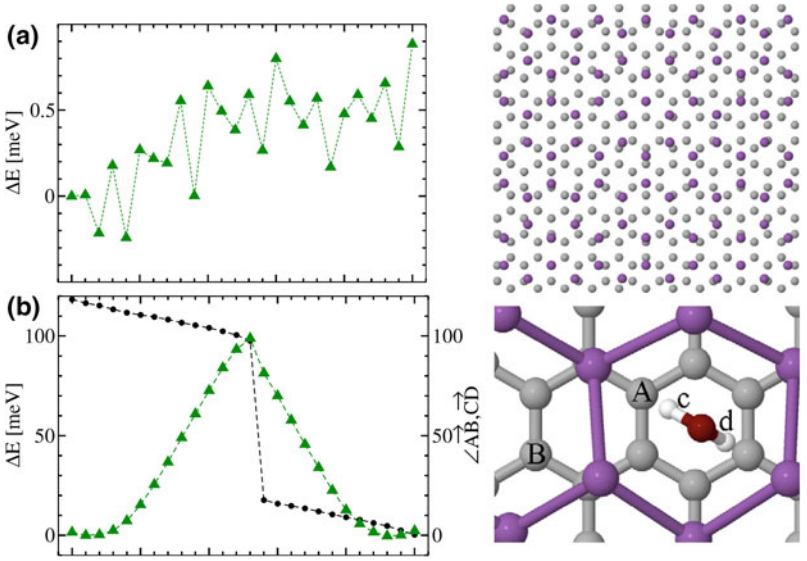


Fig. 17.20 Potential energy landscape (*left panels*, *green triangles*) calculated for a clean, crystalline Sb/HOPG interface **(a)** and an Sb/HOPG interface with a H_2O molecule trapped in the interface (*right panels*). The *black spheres* in **(b)** indicate the relative orientation between the H_2O molecule and the substrate, showing, that the trapped molecule is indeed mobile (figure adapted from [84])

modeling and analyzed the behaviour of different contamination atoms or molecules at the interface between HOPG and crystalline antimony [84]. By calculating the potential energy surface (PES) for the sliding process with and without contaminants, information could be gained about the additional energy barriers that need to be overcome in the presence of contaminants. Figure 17.20 illustrates the results for the case of a clean interface and an interface with an H_2O molecule trapped between the surfaces of 2.6 nm^2 size. While the clean interface results in a very low energy barrier ($E_{barrier} < 1$ meV), which is consistent with the expectations for structural lubricity, adding an H_2O molecule into the interface results in a substantial increase of the energy barrier, which is found to be approximately 100 meV. Similar effects and energy barriers were found when adding an Sb_4-clusters or one oxygen atom or Sb_4O_6 into the interface.

The simulations by Brndiar et al. revealed that the trapped H_2O molecule is moving along with the antimony surface, but still remains essentially mobile. This mobility is indicated in Fig. 17.20b, where the rotation of H_2O molecule during translation is plotted. Particle mobility is an important condition for 'dirt' molecules to function as mediators between incommensurate interfaces [76]. If dirt molecules would be fixed and rigidly bound to, e.g., the particle, the resulting new interface would again show superlubricity.

In an attempt to estimate the number of contamination molecules at the interfaces, it was found, that approximately 1 H_2O molecule/75 nm^2 is required to explain the experimentally observed friction levels. This value seems to be conceivable, especially, when considering, that the HOPG used for experiments was cleaved under ambient conditions [35]. On the other hand, a concentration of 1 Sb_4/7 nm^2 was found to be required. While this concentration in itself seems to be very high, it might be understood by assuming that Sb_4 clusters loosely attached to the sliding surface of the particle might act as self contaminants. In any case, these simulations affirm the plausibility of interface contamination to be responsible for the observed frictional duality.

17.6 Conclusion and Outlook

In recent years, analyzing friction by nanoparticle manipulation has become a field of growing interest in nanotribology. The quantitative extraction of interfacial friction from particle manipulation experiments opens the door for the analysis of many current issues in the field of nanoscale friction. Compared to conventional fricton force microscopy, the accessible range of materials is greatly enhanced and only limited by the ability to create nanoscale particles on flat surfaces. Due to the well defined and clean interface conditions that can be achieved during nanoparticle manipulation experiments under UHV conditions, it is possible to verify fundamental tribological theories describing the friction between extended nanocontacts. Most prominently, the contact area dependence of friction can be analyzed in detail, which was previously hindered by the geometric limitations of tip-sample contacts in conventional friction force microscopy. Experiments show that such measurements are also suitable to approach other current issues in nanotribology like the occurrence of structural lubricity and the influence of interface contamination on friction.

Over the years, different experimental strategies have evolved, the choice of which mostly depends on the experimental conditions or the nanoparticles to be analyzed. Nowadays, especially the ‘tip on top’-approach seems to be very promising for the analysis of fundamental friction processes. It was initially introduced to measure the difference between static and sliding friction, but the permanent contact between tip and nanoparticle also makes it an interesting approach to perform systematic measurements with, e.g., variation of the normal force or sliding velocity. Through this approach, it can be analyzed how fundamental concepts describing atomic friction, such as the thermally activated Prandtl-Thomlinson-Model, can be transferred to extended nanocontacts.

Furthermore, the analysis of friction by nanoparticle manipulations is not only interesting for analyzing fundamental friction laws, but can also be relevant for technological applications. Compared to conventional friction force microscopy with contact sizes of about 10–100 nm^2, nanoparticle manipulation allows the analysis of far more realistic contact sizes that can also be found in micro- or nanoelctromechanical systems (MEMS, NEMS), where friction and wear currently still

limit the perspective for widespread application. Understanding friction for such mesoscale contacts might therefore have a considerable impact on technological applications.

Acknowledgments U.S. acknowledges primary financial support by the National Science Foundation through the Yale Materials Research Science and Engineering Center (Grant No. MRSEC DMR-1119826).

References

1. M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**, 1942–1945 (1987)
2. E. Gnecco, E. Meyer (eds.), *Fundamentals of Friciton and Wear on the Nanoscale* (Springer, Berlin, 2007)
3. R.W. Carpick, M. Salmeron, Chem. Rev. **97**, 1163 (1997)
4. R.W. Carpick, N. Agraït, D.F. Ogletree, M. Salmeron, J. Vac. Sci. Technol. B **14**, 1289 (1996)
5. M.A. Lantz, S.J. O'Shea, M.E. Welland, K.L. Johnson, Phys. Rev. B **1997**(55), 10776 (1997)
6. U.D. Schwarz, O. Zwörner, P. Köster, R. Wiesendanger, Phys. Rev. B **56**, 6987 (1997)
7. U.D. Schwarz, O. Zwörner, P. Köster, R. Wiesendanger, Phys. Rev. B **56**, 6997 (1997)
8. E. Meyer, R. Lüthi, L. Howald, M. Bammerlin, M. Guggisberg, H.-J. Güntherodt, J. Vac. Sci. Technol. B **14**, 1285 (1996)
9. M. Enachescu, R.J.A. van den Oetelaar, R.W. Carpick, D.F. Ogletree, C.F.J. Flipse, M. Salmeron, Phys. Rev. Lett. **1998**, 81 (1877)
10. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, H.-J. Güntherodt, Phys. Rev. Lett. **84**, 1172–1175 (2000)
11. M. Evstigneev, A. Schirmeisen, L. Jansen, H. Fuchs, P. Reimann. Phys. Rev. Lett. **97**, 240601 (2006)
12. O. Zwörner, H. Hölscher, U.D. Schwarz, R. Wiesendanger, Appl. Phys. A **66**, S263 (1998)
13. L. Jansen, H. Hölscher, H. Fuchs, A. Schirmeisen, Phys. Rev. Lett. **104**, 256101 (2010)
14. A. Schirmeisen, L. Jansen, H. Hölscher, H. Fuchs, Appl. Phys. Lett. **88**, 123108 (2006)
15. X. Zhao, M. Hamilton, W.G. Sawyer, S.S. Perry, Trib. Lett. **27**, 113 (2007)
16. I. Barel, M. Urbakh, L. Jansen, A. Schirmeisen, Phys. Rev. Lett. **104**, 066104 (2010)
17. R.M. Overney, H. Takano, M. Fujihira, W. Paulus, H. Ringsdorf, Phys. Rev. Lett. **72**, 3546 (1994)
18. H. Bluhm, U.D. Schwarz, K.-P. Meyer, R. Wiesendanger, Appl. Phys. A **61**, 525 (1995)
19. H. Shindo, K. Shitagami, T. Sugai, S.-I. Kondo, Phys. Chem. Chem. Phys. **1**, 1597 (1999)
20. J.Y. Park, D.F. Ogletree, M. Salmeron, R.A. Ribeiro, P.C. Canfield, C.J. Jenks, P.A. Thiel, Science **309**, 1354 (2005)
21. E. Meyer, R. Overney, D. Brodbeck, L. Howald, R. Lüthi, J. Frommer, H.-J. Güntherodt, Phys. Rev. Lett. **69**, 1777 (1992)
22. R.M. Overney, E. Meyer, J. Frommer, D. Brodbeck, R. Lüthi, L. Howald, H.-J. Güntherodt, M. Fujihira, H. Takano, Y. Gotoh, Nature **359**, 133 (1992)
23. U.D. Schwarz, W. Allers, G. Gensterblum, R. Wiesendanger, Phys. Rev. B **52**, 14976 (1995)
24. M. Hirano, K. Shinjo, Phys. Rev. B **41**, 11837 (1990)
25. M.H. Müser, Europhys. Lett. **66**, 97 (2004)
26. J.M. Martin, C. Donnet, T.L. Mogne, T. Epicier, Phys. Rev. B **48**, 10583 (1993)
27. M. Hirano, K. Shinjo, R. Kaneko, Y. Murata, Phys. Rev. Lett. **78**, 1448 (1997)
28. A. Crossley, E.H. Kisi, J.W.B. Summers, S. Myhra, J. Phys. D **32**, 632 (1999)
29. M. Dienwiebel, G.S. Verhoeven, N. Pradeep, J.W.M. Frenken, J.A. Heimberg, H.W. Zandbergen, Phys. Rev. Lett. **92**, 126101 (2004)
30. J.N. Israelachvili, J. Colloid Interface Sci. **44**, 259 (1973)

31. A.M. Homola, J.N. Israelachvili, P.M. McGuiggan, M.L. Gee, Wear **136**, 65 (1990)
32. U.D. Schwarz, J. Colloid Interface Sci. **261**, 99 (2003)
33. A.S. de Wijn, Phys. Rev. B **86**, 085429 (2012)
34. C. Ritter, M. Heyde, B. Stegemann, K. Rademann, U.D. Schwarz, Phys. Rev. B **71**, 085405 (2005)
35. D. Dietzel, C. Ritter, T. Mönninghoff, H. Fuchs, A. Schirmeisen, U.D. Schwarz, Phys. Rev. Lett. **101**, 125505 (2008)
36. A. Schirmeisen, U.D. Schwarz, Chem. Phys. Chem **10,** 2373 (2009)
37. R. Lüthi, E. Meyer, H. Haefke, L. Howald, W. Gutmannsbauer, H.-J. Güntherodt, Science **1994**, 266 (1979)
38. P.E. Sheehan, C.M. Lieber, Science **272**, 1158 (1996)
39. M.R. Falvo, J. Steele, R.M. Taylor II, R. Superfine, Phys. Rev. B **62**, R10665 (2000)
40. A. Rao, E. Gnecco, D. Marchetto, K. Mougin, M. Schönenberger, S. Valeri, E. Meyer, Nanotechnology **20**, 115706 (2009)
41. K. Mougin, E. Gnecco, A. Rao, M.T. Cuberes, S. Jayaraman, E.W. McFarland, H. Haidara, E. Meyer, Langmuir **24**, 1577 (2008)
42. E. Tranvouez, A. Orieux, E. Boer-Duchemin, C.H. Devillers, V. Huc, G. Comtet, G, Dujardin. Nanotechnology **20**, 165304 (2009)
43. G. Paolicelli, M. Rovatti, A. Vanossi, S. Valeri, Appl. Phys. Lett. **95**, 143121 (2009)
44. D. Dietzel, M. Feldmann, H. Fuchs, U.D. Schwarz, A. Schirmeisen, Appl. Phys. Lett. **95**, 053104 (2009)
45. J.N. Israelachvili, D. Tabor, Wear **24**, 386–390 (1972)
46. B.J. Briscoe, D.C.B. Evans, Proc. Roy. Soc. Lond. A **1982**(380), 389 (1982)
47. J. Krim, A. Widom, Phys. Rev. B **38**, 12184–9 (1988)
48. J. Krim, D.H. Solina, R. Chiarello, Phys. Rev. Lett. **66**, 181–184 (1991)
49. T. Coffey, J. Krim, Phys. Rev. Lett. **95**, 076101 (2005)
50. Q. Zhong, D. Inniss, K. Kjoller, V. Elings, Surf. Sci. Lett. **290**, L688 (1993)
51. T.R. Albrecht, P. Grtter, D. Horne, D. Rugar. J. Appl. Phys. **69**, 668 (1991)
52. B. Anczykowski, B. Gotsmann, H. Fuchs, J.P. Cleveland, V.B. Elings, Appl. Surf. Sci. **140**, 376–382 (1999)
53. C. Ritter, M. Heyde, U.D. Schwarz, K. Rademann, Langmuir **18**, 7798–7803 (2002)
54. D.A. Aruliah, M. Müser, U.D. Schwarz, Phys. Rev. B **71**, 085406 (2005)
55. M. Tripathi, G. Paolicelli, S. D'Addato, S. Valerie, Nanotechnology **23**, 245706 (2012)
56. M. Ternes, C.P. Lutz, C.F. Hirjibehedin, F.J. Giessibl, A.J. Heinrich, Science **319**, 1066 (2008)
57. G. Langewisch, J. Falter, H. Fuchs, A. Schirmeisen, Phys. Rev. Lett. **110**, 036101 (2013)
58. R. Lüthi, E. Meyer, H. Haefke, L. Howald, W. Gutmannsbauer, H.-J. Güntherodt, Science **266**, 1979 (1994)
59. M. Palacio, B. Bhushan, Nanotechnology **19**, 315710 (2008)
60. D. Dietzel, T. Mönninghoff, L. Jansen, H. Fuchs, C. Ritter, U.D. Schwarz, A. Schirmeisen, J. Appl. Phys. **102**, 084306 (2007)
61. H. Ueyama, Y.Sugawara, S.Morita, Appl. Phys. A Mater. Sci. Process. **66**, 295 (1998)
62. A. Schirmeisen, H. Hölscher, B. Anczykowski, D. Weiner, M.M. Schäfer, H. Fuchs, Nanotechnology **16**, S13 (2005)
63. D. Dietzel et al., Trib. Lett. **39**, 273 (2010)
64. A. Rao, M.L. Wille, E. Gnecco, K. Mougin, E. Meyer, Phys. Rev. B **80**, 193405 (2009)
65. P. Nita, S. Casado, D. Dietzel, A. Schirmeisen, E. Gnecco, Nanotechnology **24**, 325302 (2013)
66. Y. Liu, I. Szlufarska, Phys. Rev. Lett. **109**, 186102 (2012)
67. Q. Li, T.E. Tullis, D.L. Goldsby, R.W. Carpick, Nature **480**, 233 (2011)
68. B.N.J. Persson, Phys. Rev. Lett. **71**, 1212 (1993)
69. B. Stegemann, C. Ritter, B. Kaiser, K. Rademann, J. Phys. Chem. B **108**, 14292 (2004)
70. B. Kaiser, B. Stegemann, H. Kaukel, K. Rademann, Surf. Sci. **496**, L18–L22 (2002)
71. C. Ritter, M.Z. Baykara, B. Stegemann, M. Heyde, K. Rademann, J. Schroers, U.D. Schwarz, Phys. Rev. B. **88**, 045422 (2013)

72. C. Bombis, F. Ample, J. Mielke, M. Mannsberger, C.J. Villagmez, C. Roth, C. Joachim, L. Grill, Phys. Rev. Lett. **104**, 185502 (2010)
73. X. Feng, S. Kwon, J.Y. Park, M. Salmeron, ACS Nano **7**, 1718 (2013)
74. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Oxford University Press, New York, 1950)
75. J.A. Greenwood, J.B.P. Williamson, Proc. R. Soc. Lond. A **295**, 300 (1966)
76. M.H. Müser, L. Wenning, M.O. Robbins, Phys. Rev. Lett. **86**, 1295 (2001)
77. M. Hirano, K. Shinjo, Wear **168**, 121–125 (1993)
78. M.R.Sørensen, K.W. Jacobsen, P. Stoltze, Phys. Rev. B **53**, 2101–2113 (1996)
79. M.H. Müser, M.O. Robbins, Phys. Rev. B **61**, 2335–2342 (2000)
80. K. Shinjo, M. Hirano, Surf. Sci. **283**, 473 (1993)
81. G. He, M.H. Müser, M.O. Robbins, Science **284**, 1650 (1999)
82. Y. Mo, K.T. Turner, I. Szlufarska, Nature **457**, 1116 (2009)
83. M. Reguzzonia, M. Ferrarioa, S. Zapperia, M.C. Righia, PNAS **107**, 1313 (2010)
84. J. Brndiar, R. Turansky, D. Dietzel, A. Schirmeisen, I. Stich, Nanotechnology **22**, 085704 (2011)

Chapter 18
Tribological Aspects of In Situ Manipulation of Nanostructures Inside Scanning Electron Microscope

Boris Polyakov, Leonid Dorogin, Sergei Vlassov, Ilmar Kink and Rünno Lõhmus

Abstract This chapter is dedicated to manipulation of nanostructures inside a scanning electron (SEM) microscope employed for real-time tribological measurements. Different approaches to force registration and calculation of static and kinetic friction are described. Application of the considered methodology to Au and Ag nanoparticles, as well as ZnO and CuO nanowires, is demonstrated. Advantages and limitations of the methodology in comparison to traditional AFM-based manipulation techniques are discussed.

18.1 Introduction

Macroscopic laws of friction had been thoroughly studied during last centuries. Recently significant progress was made also in research of friction phenomena at the atomic scale. However, it is important to improve our understanding of tribological behaviour of nanoscale and mesoscale structures and systems, in order to provide both fundamental and practical knowledge for emerging numerous applications in nanotechnology. These structures are typically too complex to be modelled from

B. Polyakov (✉)
Institute of Solid State Physics, University of Latvia, Kengaraga Street 8, Riga 1063, Latvia
e-mail: boris.polyakov@cfi.lu.lv

L. Dorogin · S. Vlassov · I. Kink · R. Lõhmus
Institute of Physics, University of Tartu, Ravila 14c, 50412 Tartu, Estonia
e-mail: dorogin@ut.ee

S. Vlassov
e-mail: vlassovs@ut.ee

I. Kink
e-mail: ilmar.kink@ut.ee

R. Lohmus
e-mail: rynno@fi.tartu.ee

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_18

atomic/molecular point of view, and at the same time too small that the macroscopic models often break down. Therefore strong demands arise for advanced experimental and modelling techniques aimed to provide more information on characteristics of investigated nanoobjects (nanoparticles, nanorods, nanowires, nanotubes, etc.).

Tribological characterization of nanostructures often implies controllable displacement of investigated objects. In contrast to macroscopic objects, where large fraction of applied energy is spent to overcome inertia of a body to put it to motion, at the nanoscale most of the energy is dissipated on adhesion and static friction forces. Taking into account the contact area dependence of the nanoscale friction [1], it is extremely important that the investigated nanostructures have well-defined geometry and atomically smooth surfaces. Typical objects in the nanotribological studies comprise various types of nanoparticles (NPs) or nanowires (NWs).

NP material can be chosen in a wide range from noble metals to metal oxides, sulphides, halides, etc. By choosing a proper synthesis method, the NP shape and size can be well controlled. For example, gold NPs can be produced in tetrahedral, decahedral, icosahedral shapes and size from several tens of nanometres to a few hundred nanometres. Using thermal or laser assisted annealing the faceted NPs can be transformed into spherical ones [2]. Moreover, complex core-shell structures [3, 4] with core consisting of inorganic nanoparticles surrounded by inorganic or organic materials as a shell can be investigated.

NWs of metal (e.g. gold, silver, etc.) or covalent materials (e.g. silicon, indium arsenide, zinc oxide, etc.) can be fabricated monocrystalline having well-defined shape [5–8], which makes them good candidates for nanotribological experiments. Being nanoscale in one dimension, nanowires are ideal objects bridging nano and microscale. Moreover, the profile of NW during the manipulation experiments can be used as a tool for measuring friction force, which significantly expands assortment of experimental techniques [9].

The most common instrument used for the manipulation of nanostructures is an atomic force microscope (AFM). AFM enables a high level of control over the applied force and precise manipulation. However, it has several limitations. AFM can be used either for visualisation or manipulation. Therefore only initial and final positions of the object can be determined, while intermediate data are missing.

It is possible to integrate AFM with SEM and perform insitu manipulations and frictional measurements for nanostructures. There are several commercial SEM-compatible AFM instruments [10, 11]. In this approach prominent AFM features are complemented with advantages of SEM. Independent visualization and manipulation functions provide real-time observation of the behaviour of the objects during the manipulations and simultaneous recording of the interaction force. Visual guidance is especially important for manipulation of 1D nanostructures like nanowires or nanotubes (NTs). It is possible to observe the profile evolution of a NW during its initial displacement and find degree of elastic deformation of the NW translated over substrate surface. In addition, SEM enables to identify rolling, sliding, or rotational motions of nano objects during manipulations. Moreover, SEM provides vacuum environment that excludes the humidity effect on tribological measurements. In combined AFM-SEM systems imaging can be realized both with scanning electron beam

or AFM probe. However, for inside-SEM tribological studies AFM is mainly used for manipulations and force measurements. In this case the feedback is not essential and therefore the whole set-up can be significantly simplified. The last concept together with application examples will be introduced in more details in the following sections.

18.2 Section I: Instrumentation

18.2.1 Nanomanipulators

A typical nanomanipulation system consists of a sharp needle (tip) and a positioning device that enables a displacement of the tip. Very precisely controlled movement of the tip is an essence of nanoscale experiments. Nanomanipulation system should provide both coarse positioning (mm range) and precise (nm range) motion of the probe relative to substrate. Some commercially available nanomanipulators, such as e.g. 3D Manipulators by SmarAct GmbH [12], combine both functions in one device using inertial piezo slider principle. However, separation of coarse and fine motions is preferable because of stability and precision reasons. The coarse movements can be realized either manually using precise screws, or can be automated by means of piezo-, stepper- or servo motors. For fine movements piezo actuators (scanners), providing down to subangstrom motion resolution, are commonly utilized. Among them, a tubular scanner design combining *X-Y-Z* motion in a single piezotube has emerged as a simple and widely used technique. The tube scanners can be relatively easily installed inside any electron microscopes and can serve as a basis of a simple home-made manipulation system [13]. A serious drawback of the piezotube scanners is a crosstalk and spherical coordinates of *X-Y* movement. A special design of piezoscanner mechanics should be used to exclude all degrees of freedom but one. As an example of advantageous solutions a flexure guidance system actuated by a piezoelement can be mentioned. Modular principle of stacking and combining piezoscanners enables building systems according to specific user requirements using commercially available components like e.g. Attocube scanners [14]. If independent *X*, *Y* and *Z* scanners are used, then the linear motion in Cartesian coordinates can be realized.

There are several general drawbacks inherent to all piezoelectric scanners, including nonlinear response to applied electric field, hysteresis, creep and depolarization with time [15]. Various hardware and software solutions can be used to minimize these effects such as closed loop scanner equipped with position sensors (electrical, optical, etc.) or alternatively special scanner calibration, which compensate its hysteresis [10]. In order to minimize a thermal drift most parts of the manipulation system should have similar thermal expansion coefficients.

General requirement for nanomanipulation system is a rigid, compact and stable design. In order to minimize a thermal drift most parts of the manipulation system should have similar thermal expansion coefficients. Additional requirements arise

for manipulation inside an electron microscope. No magnetic components should be used in close proximity to the electron beam to avoid distortions of the image. Both samples and probes should have sufficient electrical conductivity to exclude charging effect. Manipulator details should be vacuum compatible, and detail surface should be of high processing quality, and details design should exclude closed volumes to minimize degassing and accelerate vacuum pumping.

18.2.2 Force Measurements

Utilization of nanomanipulation methods for tribological studies requires force detection capability. One way is to use conventional AFM cantilevers with an external deflection detection systems [16–19] like e.g. optical readout, where laser beam is reflected by a cantilever and angular deflection of the cantilever can be read out by a position sensing photodetector. Another concept is based on self-actuating and self-sensing force probes where electronic excitation and readout are used [20, 21], like piezoresistive cantilevers or electromechanical resonators. The self-sensing probes are more convenient for applications inside SEM due to compact design, fully digital-electronic control and simple integration into vacuum chamber.

18.2.2.1 Force Sensors Based on Quartz Tuning Fork

One of the best electromechanical resonators is a quartz tuning fork (QTF), originally designed to serve as a frequency standard in various devices. The QTF can be used as force sensors since oscillation parameters (amplitude, frequency and phase) are correlated to the applied force. Due to the piezoelectric properties of quartz crystal, oscillations in QTF sensors can be excited and registered electrically, making them small, robust and simple to operate compared to optical force detection methods. Moreover, QTF force sensors have several advantages: high mechanical quality Q factor, large spring constant and adjustable oscillation amplitude and sensitivity [22]. For obtaining numerical values of force, QTF based sensor should be calibrated [23, 24].

To make a force sensor suitable for measurements and manipulations inside SEM, sharp tip should be attached to one prong of QTF and electrically connected with one of its electrodes [28]. General recommendation for the tips used for the in situ nanomanipulations is sufficient conductivity to dissipate an electric charge, which will unavoidably arise due to the electron beam and will disturb an image in case of nonconductive objects.

The tip can be fabricated on QTF directly using Focused Ion Beam (FIB) technique [25], or the tip can be glued to QTF sensor [26]. Tips can be made from different materials. For example, tungsten tips can be easily prepared on-site by electrochemical etching [27]. However, sharp metallic tips can be plastically deformed and therefore are suitable only for very fine manipulations where the forces in order of a few nN or smaller are involved. Alternatively, conventional AFM cantilevers can be glued

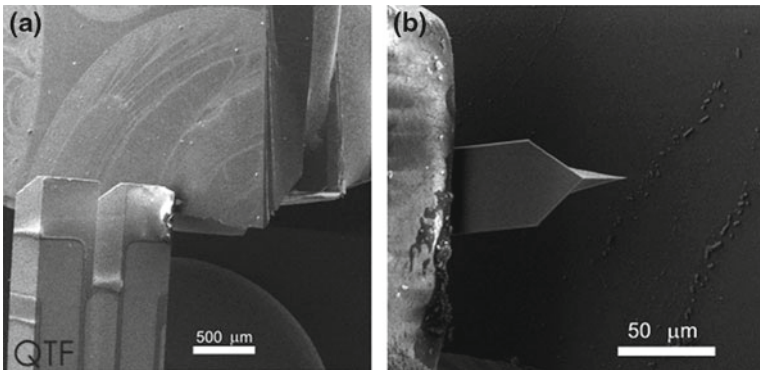


Fig. 18.1 SEM image of QTF force sensor with cantilever having tip visible from the *top* (**a**). Zoomed image (**b**)

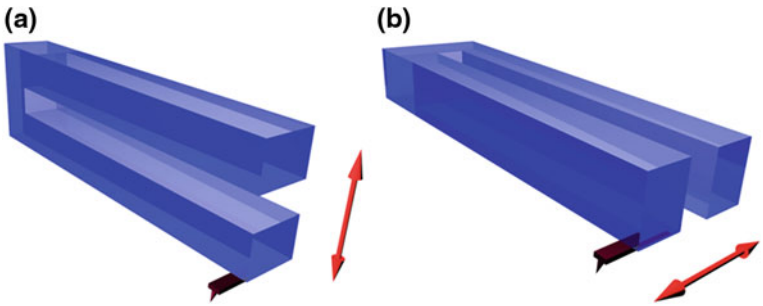


Fig. 18.2 Schematics of QTF force sensor for normal (**a**) and shear regime (**b**)

to QTF using optical microscope and micromanipulators [28]. The tip of AFM cantilever is usually hidden below the cantilever and is not seen on the SEM image, therefore both cantilever and the substrate should be tilted to make the tip visible from the top during manipulations [28]. More convenient is to use cantilevers with the tip protruding from its end [29]. Such configuration makes the tip visible from the top (Fig. 18.1).

Depending on how the tip is glued to QTF it is possible to work in normal and shear modes. In normal mode the tip oscillates perpendicular to the substrate surface (Fig. 18.2a), whereas in shear mode the tip oscillates parallel to the surface (Fig. 18.2b). Conventional AFM microscopes operate in normal mode only. In normal mode the interpretation of force interaction of the tip with manipulated object is complicated, since the directions of the tip movements and tip oscillations do not match. In shear mode the directions of the tip movement and oscillation coincide, which is favourable for the displacement of the object and interpretation of the interaction force. Shear mode is also approx. 5–10 times more sensitive at the same driving voltage.

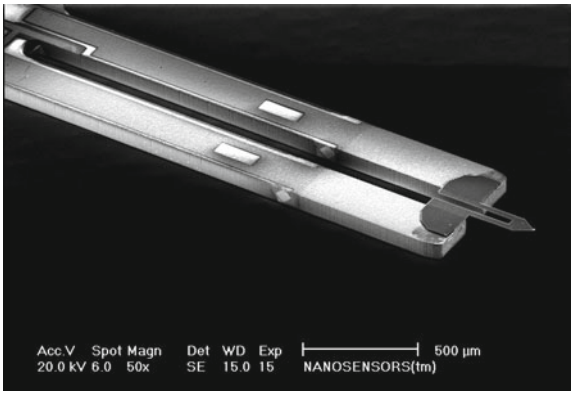


Fig. 18.3 SEM image of Akyama probe. Copyright by NanoWorld AG

18.2.2.2 Akyama-Probes

Recently self-sensing and self-actuating probes utilizing QTF in a combination with micromachined cantilevers became commercially available. The original probe was invented by Terunobu Akiyama and therefore called "Akiyama-probe" [30]. Akiyama-probe features a symmetrical arrangement of a U-shaped silicon cantilever attached to both prongs of a QTF (Fig. 18.3). The tuning fork serves as a high Q-factor mechanical oscillatory force sensor. The ends of the QTF prongs are moving in-plane and have opposite phases. This motion causes a small vibration at the glued ends of the cantilever, which always contains a *Z*-component (i.e., the axis of the tip) due to the twisting motion of the prongs (i.e., the cantilever disturbs the symmetry of the TF). The cantilever amplifies the vibration and out-of-plane motion of the tip is obtained. The direction of the Akiyama-probe tip oscillation is similar to conventional tapping mode AFM cantilevers. However, the tip oscillates at QTF resonance frequency and not determined by the force constant of the cantilever.

18.2.2.3 Piezoresistive Cantilevers

Working principle of piezoresistive cantilevers is based on a property of piezoresistive materials to change their bulk resistivity when mechanical stress is applied. Silicon is a well known piezoresistive material and suitable for fabrication of cantilevers. There are different designs of piezoresistive cantilever.

In general, a silicon resistor made from doped silicon is placed on a U-shaped or single slab cantilever and electrically isolated from it. The resistor is connected to an external DC-biased Wheatstone bridge, which directly measures the deflection by measuring the cantilever's resistance. Elastic deformation in piezoresistor is proportional to voltages difference $V_1 - V_2 = -V\Delta R/4R$. The resistor can be placed near to the most highly stressed region close to the base of the cantilever. The resistor

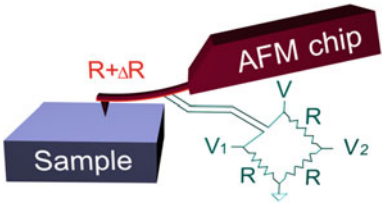


Fig. 18.4 Detection of piezoresistive cantilever deflection using Wheatstone bridge

should be made as thin as possible to ensure maximal stress and consequently high piezoresistive response.

Alternatively, whole cantilever can be fabricated as a piezoresistor (Fig. 18.4). In this case cantilever will produce zero net piezoresistive response due to mutual compensation of stress on the top and bottom side of the cantilever and zero total contribution to the change in resistance [31, 32].

18.3 Section II: Manipulation of Nanoparticles

Frictional properties of NPs were extensively studied in AFM-based manipulation experiments either in contact or dynamic (oscillating) modes. Following a common AFM manipulation protocol, NPs are first visualized in dynamic mode using a very gentle force. Then several manipulation strategies can be applied. According to one method a tip is positioned behind the NP to be manipulated and the particle is pushed in contact mode with disabled feedback whilst lateral force signal is recorded. Then image is taken again to examine the result of manipulation [33]. Another contact method, so-called "tip-on-top", was introduced by Dietzel et al. [34]. In this method, the AFM tip is first positioned on top of the NP approximately at its centre and certain normal force is applied. The tip is then moved aside and then the NP follows the tip motion. The measured torsional signal (lateral force) is directly proportional to the interfacial friction between the particle and the substrate. NP should be flattened to be manipulated by "tip-on-top" method. In oscillating mode NPs can be moved during the scanning process with enabled feedback when the amplitude of the tip oscillations is increased above a certain threshold value [35]. Force applied on particle is proportional to the tip oscillation amplitude. Tip oscillates perpendicular to the substrate plane, and lateral component of the applied force pushing a NP parallel to the substrate plane depends on NP diameter and tip radius. Increasing the scan rate above a certain value rather than increasing the oscillation amplitude also lead to displacement of NPs [36]. If experiments are performed inside SEM then simultaneous force registration and visualisation of particle movement is possible providing additional information and higher degree of control over the manipulation procedure.

18.3.1 Contact Area

It is known that the friction force at the nanoscale is proportional to the contact area (in the absence of wear) [37]. When measuring friction in nanomanipulation experiments, roughness of a substrate is commonly neglected and the substrate is assumed to be atomically flat. Then, depending of the particle shape, two cases can be distinguished. If particle have well-defined facets, the contact area is considered to be simply equal to the surface area of the facet in contact with the substrate. The same is valid for surface area of the base of flattened particles (also called "nanoislands"[34]). Static friction force can then be calculated from simple relation $F_{friction} = \tau S_{\mathrm{r}}$, where τ is the interfacial shear stress/strength and S_{r} is the contact area [38]. The shear strength is defined as an ultimate shear stress τ before the object starts to move or slips and can be estimated using the relation $\tau_{theo} = G*/Z$ between the theoretical shear strength and the combined shear modulus, $G^* = [(2-\nu_1)/G_1 + (2-\nu_2)/G_2]^{-1}$, where $G_{1,2} = E_{1,2}/2(1+\nu_{1,2})$ [39]. Z is an empirical material-dependent coefficient ranging from 5 to 30 [40].

The situation is different for a spherical particle. Spherical particle (nanoball) is a special case of a particle having geometrical point contact with the substrate. However for real systems the contact area is always finite and typically can be estimated on the basis of continuum elasticity models for deformable spheres such as JKR [41] or DMT-M model [42]. According to Tabor [43], the choice of the most suitable model is determined by the parameter:

$$\eta = \left(\frac{16R\gamma^2}{9K^2 z_0^3}\right)^{1/3}, \tag{18.1}$$

where R is the radius of the sphere, γ is the work of adhesion, and z_0 is the equilibrium spacing for the Lennard-Jones potential of the surfaces. K is the combined elastic modulus of the sphere and substrate, defined as $K = 4/3\left[\left(1-\nu_1^2\right)/E_1 + \left(1-\nu_2^2\right)/E_2\right]^{-1}$ in which $\nu_{1,2}$ and $E_{1,2}$ are the Poisson ratios and Young moduli of the substrate and sphere, respectively. For small η, the DMT-M theory is more appropriate. According to the DMT-M model, the contact area A_{DMT-M} is expressed:

$$A_{\mathrm{DMT-M}} = \pi\left(\frac{2\pi\gamma}{K}\right)^{2/3} R^{4/3}. \tag{18.2}$$

In the next sections two particular works on real-time manipulation of faceted and rounded NPs are described.

18.3.2 Manipulation of Polyhedron-Like Nanoparticles

Vlassov et al. [28] applied in situ nanomanipulation technique to study static friction of chemically synthesized Au NPs of 150 nm diameter on oxidized silicon wafers

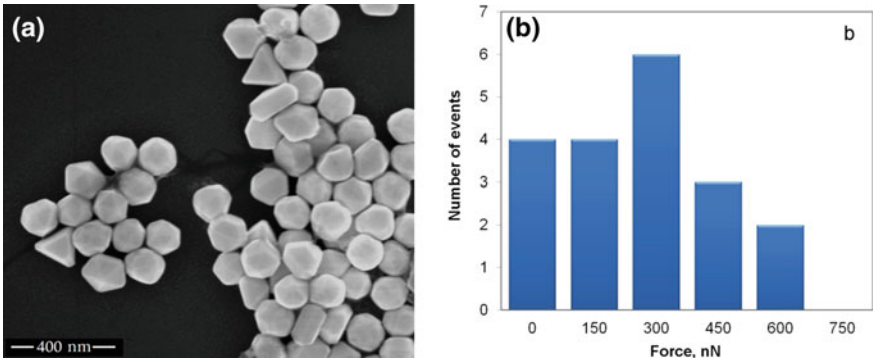


Fig. 18.5 SEM micrograph of Au NPs (**a**); Distribution of the static friction force for Au NPs (**b**)

Table 18.1 Estimated static friction forces for 150 nm Au particles of different geometries

Shape	Contact area, nm^2	Static friction, nN
Tetrahedral	9743	2768
Decahedral	3652	1038
Icosahedral	2693	765

inside SEM. Manipulation set-up consisted of a nanopositioner equipped with a self-made QTF-based force sensor.

Gold particles used in the experiment were, in general, not spherical and had facets of different geometries (Fig. 18.5a). Therefore the contact area between the particle and the surface also varied from particle to particle. Contact areas and static friction values were estimated analytically considering common geometries of the particles used in experiment [6] (Table 18.1). It should be noted that the real particles often exhibit truncated edges and apexes. Moreover NPs were annealed at 500°C to remove surfactant. Thermal treatment may lead to rounding of particles [28]. Therefore, the contact areas and static frictions should be generally lower than the maximal values listed in Table 18.1.

In the manipulation experiment the particles were firstly displaced ("kicked") by an abrupt tip motion to a fresh site in order to reduce the initial adhesion. Initial displacement was followed by controlled manipulation of the particle by pushing it with oscillating tip in the scan regime with simultaneous force recording. During manipulation, the tip moved parallel to the surface along a straight line without feedback loop.

Significant force should be applied to overcome the static friction and displace an Au NP. After displacement NPs move easily at very small applied force. This finding is in agreement with other studies that have demonstrated that the kinetic friction vanishes for clean surfaces in vacuum [44]. Nonstop motion of NPs during the manipulation was an essential condition for staying in the vanishing friction regime; rested NPs had a tendency to adhere strongly to the substrate. Static friction was measured for 19 NPs and ranged from 50 to 750 nN (Fig. 18.5b).

It should be noted that for some particles the oscillation amplitude dropped to zero during manipulation, which corresponded to a force higher than 1500–2500 nN (depending on the particular sensor). Variations in the experimental values of static friction were explained by the different contact areas between a particle and a substrate due to different shapes of the NPs.

18.3.3 Manipulation of Silver Nanoballs and Nanodumbbells

Similar manipulation technique was used for manipulation of silver nanoballs and dumbbells. Balls and dumbbells were produced by pulsed laser induced partial melting of 100 nm diameter silver NWs deposited on an oxidized silicon wafer [45]. Due to concentration of electric field on a NW ends, silver melts and affected by surface tension rounded structures are formed on both ends of a NW. If a NW was short enough, a single "nanoball" was created (Fig. 18.6).

Friction was measured with a QTF-based force sensor operated in the shear mode. Nanoballs were first displaced by abrupt motion of the tip and then static friction was measured in one-line slow scan. A typical manipulation force curve is shown in Fig. 18.7.

As it can be seen that the force needed to overcome the static friction (F_{st} = 162 nN) and displace the nanoball (region b-c in Fig. 18.7) is significantly higher than the force needed to maintain the motion (region c-d-e in Fig. 18.7). Kinetic friction force in this particular case spontaneously decreased from $F_{kin} = 26$ nN to 7 nN (Fig. 18.7), which can be attributed to switching between sliding and rolling types of motion. Such behaviour was not typical and was observed only in few cases.

Assuming the following parameters for silicon and silver: $E_1 = 71.7$ GPa, $\nu_1 = 0.17$, $E_2 = 82.5$ GPa, $\nu_2 = 0.36$, $\gamma = 50\text{mJ/m}^2$ [36, 46], $R = 200$ nm and $z_0 = 0.3$ nm, we obtained $\eta \approx 0.22$. For such η, the DMT-M theory is more appropriate. According to the DMT-M model, the contact area for 400 nm spherical Ag NPs is $A_{DMT-M} \approx 117\,\text{nm}^2$. Then for Z in the range from 5 to 30 [40] calculated ultimate static friction $F_{friction} = \tau\, S_\text{r}$ ranging from 34 to 205 nN. Experimentally measured

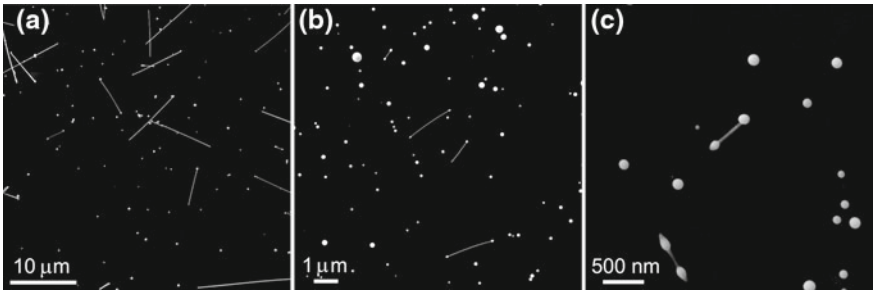


Fig. 18.6 Nanoballs and dumbbells nanostructures produced by laser processing of Ag NWs (**a**) and Au NWs (**b**, **c**)

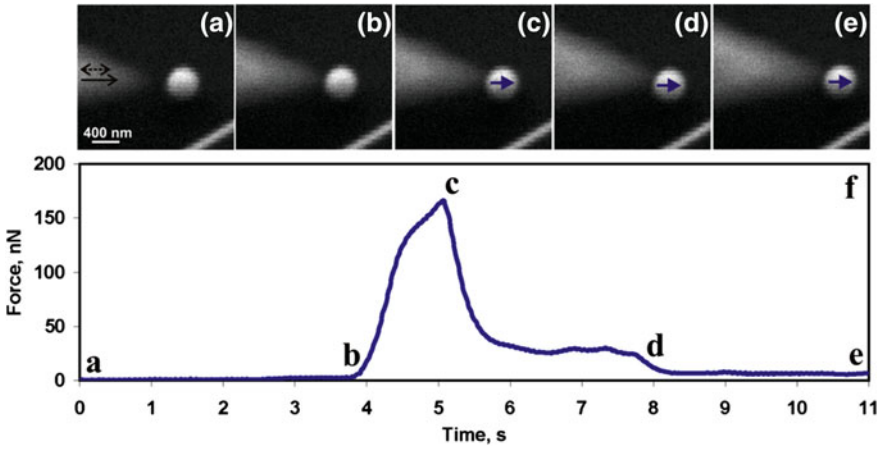


Fig. 18.7 Manipulation of a silver nanoball. The tip approaches the nanoball (**a**, **b**). Nanoball is displaced (**c**) and continuously moved by the tip (**c**–**d**–**e**). Corresponding tip-nanoball force interaction data recorded by QTF sensor (**f**). *Solid black arrow* indicates tip motion, while *dashed line* indicates tip oscillation direction. Direction of NP movement indicated by *short arrows*

static friction force having average value 65 nN agrees well with the calculated range of friction force.

The same manipulation technique was then applied to Ag nanodumbbells. Nanodumbbells of 1–2 μm in length can be roughly considered as two spheroidal NPs connected by a NW. It should be noted, that rounded ends of dumbbells usually had significant deviations from ideal spherical shape in most cases (due to contact angle of liquid metal droplet with substrate) and DMT-M model for sphere was not valid anymore. However, dumbbell ends ensure relatively small contact area and therefore reduced adhesion and static friction. For example, displacement of untreated uniform Ag NWs on a flat silicon substrate is almost impossible without severe damage and plastic deformation of NW. Moreover, for dumbbell-type structures both rolling and sliding motions are possible, which can be distinguished visually when manipulated inside SEM. It makes them attractive objects for nanotribological studies by manipulations.

Typical DB manipulation experiment is shown at Fig. 18.8. Significant static friction force ($F_{st} = 1,050$ nN) should be overcame to displace the DB. In particular case the DB rolled over first (Fig. 18.8a, b), and then rotated around one end (Fig. 18.8c, d). In some cases one end completely lost contact with the substrate and dumbbell rotated around adhered end out of substrate plane. If both ends of dumbbell lost contact with the substrate, dumbbell flew away. In many cases static friction was high enough to keep one end fixed which led to plastic deformation of the dumbbell during manipulation. Long range rolling of Ag dumbbells was never observed, while rolling up to approx. 90° however was registered frequently.

As it was demonstrated above, SEM assisted visual guiding is equally useful for manipulation of both NPs and elongated objects by probe-equipped force sensor.

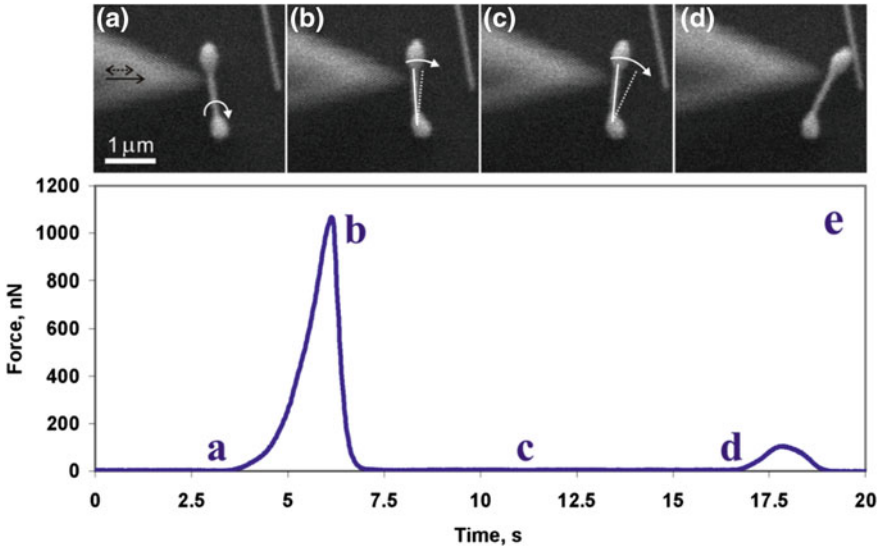


Fig. 18.8 Manipulation of a silver nanodumbbell. The dumbbell rolls over approx. 90° (**a**, **b**), then rotates around one of its "balls" (bottom ball, **b**, **c**, **d**), and finally runs into NW (**d**). *Arrows* indicate the motion type. Corresponding tip-dumbbell force interaction data recorded by a QTF sensor (**e**)

18.4 Section III: Manipulation of Nanowires

There is a number of experimental methods and theoretical models for investigation and description of mechanical properties of NWs and nanotubes (NTs). Nevertheless, only a very few works have focused on measuring the static or kinetic friction of NWs (or NTs) on a flat substrate. Manoharan et al. performed a study that examined NW-substrate adhesion and friction of long ZnO NWs on a silicon substrate [47]. They examined the friction force during dragging of a NW (parallel to its axis) at various loadings. However, most of the friction measurements on NTs and NWs were performed using a conventional AFM at ambient conditions. Hsu et al. manipulated BN NW on a silicon substrate and simultaneously measured the lateral force [48]. Falvo et al. studied sliding and the rotation of multiwall carbon NTs on HOPG using lateral force AFM microscopy [49]. It is necessary to emphasise that visual guidance for NW manipulation experiments is even more critical in comparison to NPs manipulation. Only a short or very stiff NW behaves as a rigid rod during manipulation (slides or rotates as a rigid object without bending). High aspect ratio NWs can be bent significantly during manipulation and even broken if static or kinetic friction force exceeds NW bending strength. In situ manipulation can reveal evolution of the NW profile bending in contrast to AFM manipulation in which only the initial and final positions of the NW are obtained. In the following section we will demonstrate, that combination of in situ SEM visual observation and mathematical

modelling of NWs elastic bending during manipulation can be a powerful tool for tribological investigations.

18.4.1 Elastic Beam Theory Employed for Tribomechanical Studies of Nanowires

Theoretical methods for investigation of NW's mechanical and tribological properties commonly employ the framework of classical physics and range from simplest Newton's mass point models [50] to continuum models and finite element method (FEM) calculations [51]. Due to the relatively large sizes of the systems (which may account in millions of atoms), application of molecular dynamics is limited to NW with dimensions of several nm. Therefore, the continuum approach is attractive for NW modelling because of extensively developed applicable analytical and numerical methods with different levels of accuracy.

One of the most typical deformations of NW is elastic bending. Using the assumption of pure elastic bending, it is possible to investigate many important properties of NW exposed to the friction and other forces occurring during manipulations. The advantage of this assumption is its relative simplicity, which we will demonstrate throughout the current section.

The most known model based on the assumption of pure elastic bending is Timoshenko elastic beam theory (EBT) [52]. It combines the equilibrium equations of a continuum body with the laws of linear elasticity approximated for one-dimensional structure. In the particular case considered, the bending profile of a NW comes as a result of the interplay between elastic, friction, and other external forces distributed along the NW with certain densities. The EBT can assist to measure NW-substrate friction.

The present section will include description of several methods of theoretical approach to NW-on-substrate friction during and after manipulations together with some experimental examples [53–55]. These methods enable to gather information on NW-substrate distributed static and kinetic friction as a functions along NW's length, ultimate static friction, as well as NW flexural strength at the breaking point. Additionally, the case of a NW half-suspended on a trench will be considered that enables measuring Young's modulus for individual NWs and can be used as an alternative way to measure the ultimate static friction.

18.4.1.1 Elastic Forces

Let us consider a prismatic-shaped NW of length L and diameter D, bent under external in-plane forces (Fig. 18.9). For the implementation of the EBT we will need to introduce variables of the force and momentum of elastic stresses in a cross section of the NW. They represent the force and momentum between the elastically deformed

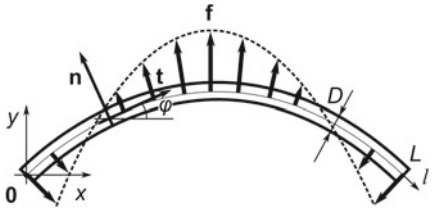


Fig. 18.9 Schematics of a NW of length L and diameter D held in a bent state by in-plane distributed lateral force **f**. Fixed coordinates system Oxyz and local coordinate basis (**t**,**n**) along the NW axis l are used. Angle between the tangent vector **t** and axis Ox is denoted as φ

sides of the NW at an imaginary cut, and will be designated as **F** and **M**, respectively. Their components can be written as integrals over the cross section area S at any given point l of NW axis [52, 56]:

$$F_i = \int_s \sigma_{i\gamma} n_\gamma dS \tag{18.3}$$

$$M_i = \int_s e_{i\alpha\beta} r_\alpha \sigma_{\beta\gamma} n_\gamma dS \tag{18.4}$$

where $\sigma_{\alpha\beta}$ are the components of stress tensor, n_γ are the components of the normal vector to the element of cross section area dS, r_α are the components of the radius vector from the axial point l and $e_{\alpha\beta\gamma}$ represents the unit anti-symmetric tensor. Both the elastic force **F** and the momentum **M** are considered as functions of the coordinate l along the NW axis (Fig. 18.9).

Here and below we assume that the NW is undergoing a pure bending deformation. This can be a good approximation for the given problem provided the external forces are close to the normal direction of the NW axis and lie on the same plane as the NW itself. It assures the negligibly small tensile, compressive and twisting deformations. The case of the pure bending of prismatic shaped NW yields the following equation for momentum in the framework of linear isotropic elasticity [56]:

$$\mathbf{M} = EI\mathbf{t} \times \frac{d\mathbf{t}}{dl}, \tag{18.5}$$

where E is the Young modulus of the NW, I is the area moment of inertia of the NW, and **t** is the tangent vector of the NW axis. Equation (18.5) describes the momentum of elastic forces inside a NW bent purely with the given curvature. It can be projected to the Oz axis and rewritten as

$$M = EI\frac{d\varphi}{dl} = EI\kappa, \tag{18.6}$$

where $\varphi(l)$ is the tangent angle function over the length of the NW and κ (l) = 1/R(l) = dφ/dl is the curvature function related to the local radius of curvature *R (l)*. Physical meaning of (18.5) and (18.6) for pure bending deformation is of similar kind as Hooke's law for an elastic spring. In fact, this is the "Hooke's law" written for the specific case, where κ plays a role of "displacement", *M* means "elastic force" of restoration, and *EI* is a "spring constant" (which is sometimes referred to as *flexural rigidity*).

By following this analogy we can obtain the expressions for elastic energy per unit length u^{st} and total elastic energy U^{st} of the bent NW can be calculated, respectively, as

$$u^{\mathrm{st}} = \frac{EI}{2}\kappa^2 \tag{18.7}$$

$$U^{\mathrm{st}} = \int_0^L \frac{EI}{2}\kappa^2 dl \tag{18.8}$$

It must be noted, that single crystalline solids possess anisotropic elastic properties and therefore should be described with multiple elastic moduli. However, due to simplicity, the isotropic model is often used. Although this approximation of isotropic body is acceptable in the most cases, it is important to keep in mind that isotropic elastic moduli describe only "effective" elastic properties in the specific axes of deformation.

18.4.1.2 Cross-Section and Contact Area

Elastic response of a NW is directly related to its geometry (cross section), particularly the diameter *D*. In its turn the NW cross section determines the NW-substrate contact area that essentially determines the NW-substrate friction.

Typical cross section geometries of NWs can be roughly classified as polygonal and round cross sections. For example, hexagonal cross section yields $I = 5\sqrt{3}D^4/256$. Contact regions for polygonal shapes are easily defined. Determination of contact region for circular cross section is a non-trivial problem that requires employment of contact mechanics.

It should be noted that from the expression for *I,* which scales as power 4 of the NW lateral dimensions, it is evident that an accurate measurement of NW cross section geometry is crucial for correct application of EBT. In the literature for simplicity reasons a round cross section is often assumed instead of the real polygonal one. Moreover, possible distortions of the NW cross section are commonly not taken into account. However, even an accurate measurement of NW diameter is not trivial task due to technical issues not mentioning the measurements of the cross section distortions.

18.4.1.3 Friction and Other External Forces

External forces acting on a NW may have different origins: friction, adhesion, actuator load etc. In the case of pure bending of a NW on a flat substrate the lateral forces will play the major role, whereas normal forces can be neglected.

The forces can be distributed or concentrated. A distributed force can be defined as a continuous function **f** (*l*) along the NW. Then **f** (l) *dl* corresponds to the force exerted on an infinitesimal element *dl* in the point *l* of the NW. Friction forces are a typical example of the distributed external forces, which may vary from point to point. The concentrated force $\mathbf{f}_0$ can be expressed in the same manner via Dirac's delta-function as $\mathbf{f}(l) = \mathbf{f}_0\delta\ (l - l_0)$. Concentrated forces in NWs commonly occur as a result of a load from an external actuator (e.g. manipulator's tip).

18.4.1.4 Boundary Conditions

Boundary conditions for the elastic momentum and force in the NW depend on the specific situation. For a free end of NW the conditions **F** = 0 and **M** = 0 must be satisfied. In a typical case if the NW end is pushed by an actuator, then the boundary value of **F** is defined and non-zero, while **M** = 0 as there is no tip-induced momentum.

18.4.1.5 Equilibrium of the Nanowire

Let us now consider a NW in mechanical equilibrium due to the interplay of intrinsic elastic force in the NW and the external forces of various origins. The full system of equilibrium equations for the NW involve conditions for the elastic force **F** and momentum **M** [56]:

$$\frac{d\mathbf{F}}{dl} = -\mathbf{f}, \tag{18.9}$$

$$\frac{d\mathbf{M}}{dl} = -\mathbf{t} \times \mathbf{F} \tag{18.10}$$

where **f** is the distributed external force per unit length. Equation (18.9)can be interpreted as the condition of zero total force $d\mathbf{F} + \mathbf{f}dl$=0 acting on a prismatic element *dl* of the NW. The other equation reflects the condition of zero total momentum, respectively. It can be clearly seen since **M** has always the same direction *Oz* (perpendicular to the bending plane) and the right-hand side of the equation (10) has magnitude of the normal component of the force $F_n = \mathbf{F} \cdot \mathbf{n}$. Therefore, it can be rewritten as $dM + F_n dl = 0$, which had to be demonstrated.

General view on the NW flexural mechanics can be summarized in a form of diagram (Fig. 18.10), disclosing the internal factors of elastic response and external forces of various nature being in mutual equilibrium. In other words, the bending

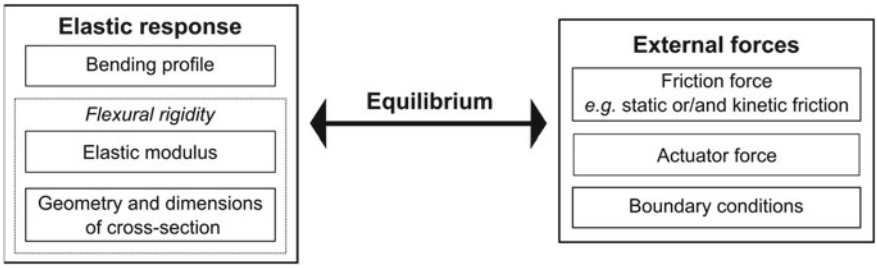


Fig. 18.10 Logical scheme describing the intrinsic and external factors for elastically bent NW to be in mechanical equilibrium

profile of resting, being displaced or uniformly moving NW can be calculated from condition of balance of NW elastic response (if Young modulus and cross section are known) and external forces (friction forces and applied actuator force). Respectively, an inverse problem can be solved, determination of friction force if bending profile and elastic response of NW are known. It should be noted that even the kinetic friction can be analysed with the equilibrium equations providing the NW is moving uniformly.

18.4.2 Nanowire Loaded at One End

In this subsection it will be demonstrated that situation of NW loaded at one end can be used for calculation of both static friction and Young modulus.

18.4.2.1 Model

Let us consider a specific situation, when initially straight NW was loaded at the end and a part of the NW was fixed to the substrate by static friction (see Fig. 18.11). The mobile part of the NW undergoes kinetic friction and the concentrated load force. The distributed force $\mathbf{f}$ consists of the kinetic friction force $\mathbf{q}^{\text{kin}}$ per unit length distributed uniformly along the moving part of the NW:

$$\mathbf{f} = \mathbf{q}^{\text{kin}}. \tag{18.11}$$

Zero elastic force and momentum at the NW's free end dictate the boundary conditions:

$$\mathbf{F}|_{l=L_{\text{kin}}} = \mathbf{F}_{\text{apl-lat}} \tag{18.12}$$

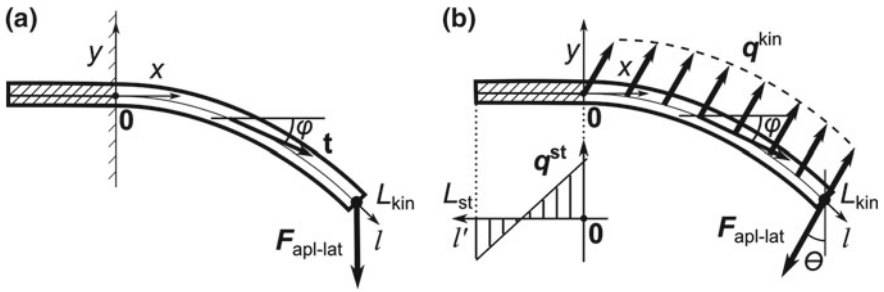


Fig. 18.11 Schematics of a NW loaded by concentrated force $F_{apl-lat}$ at the end in point L_{kin}. NW is half-suspended and friction-free at the protruding part (**a**). NW is affected by distributed kinetic friction force q^{kin} on the interval $(0, L_{kin})$ along l axis (**b**). The left end is fixed by strong static friction force q^{st} on the interval $(0, L_{st})$ along l' axis. Angle between the tangent vector and axis Ox is denoted as φ and $F_{apl-lat}$ is directed by angle θ to Oy axis

$$\mathbf{M}|_{l=L_{\text{kin}}} = 0. \tag{18.13}$$

And elastic force **F** according to (18.9) yields together with (18.11):

$$\mathbf{F} = -\mathbf{q}^{\text{kin}} \cdot (l - L_{\text{kin}}) + \mathbf{F}_{\text{apl-lat}} \cdot \tag{18.14}$$

Finally the equation for $\varphi(l)$ describing the profile of NW with boundary conditions:

$$EI\frac{d^2\varphi}{dl^2} = \{q^{\text{kin}}(l - L_{\text{kin}}) - F_{\text{apl-lat}}\} \cdot \cos(\varphi + \theta) \tag{18.15}$$

$$\varphi|_{l=0} = 0 \tag{18.16}$$

$$\frac{d\varphi}{dl}\Big|_{l=L_{kin}} = 0 \tag{18.17}$$

where (18.16) defines orientation of the NW with respect to the coordinate system, (18.17) follows directly from (18.13) and (18.12) is automatically satisfied.

Motionless ("static") part of the NW is assumed to be a rigid rod of length L_{st} with 2 degrees of freedom: axis Oy and rotation angle in the plane Oxy. Values of elastic force $F_y|_{l=0}$ and momentum $M|_{l=0}$ at the start point of the motionless region follow from the solution of (18.15), (18.16) and (18.17) for continuity reasons. For simplicity we assume that the interfacial stress between the NW and substrate and consequently the static friction force is distributed linearly along the linearly rotated static part of the NW: $q_y^{\text{st}}(l') = a \cdot l' + b$, where a and b are unknown constants. For the equilibrium of the NW, conditions for total force and momentum yield a system

of two equations:

$$F_{y|_{l=0}} = \int_0^{L_{\mathrm{st}}} q_y^{\mathrm{st}}(l')dl' \tag{18.18}$$

$$M_{|_{l=0}} = \int_0^{L_{\mathrm{st}}} q_y^{\mathrm{st}}(l')l'dl' \tag{18.19}$$

which easily determine a and b. This allows to find such parameters as maximal static friction $q_{\mathrm{max}}^{\mathrm{st}} = |b|$ and averaged static friction $q_{\mathrm{avg}}^{\mathrm{st}} = \frac{1}{L_{\mathrm{st}}} \int_0^{L_{\mathrm{st}}} |q_y^{\mathrm{st}}(l')|dl'$.

In the particular case, when the kinetic friction is absent or negligible $\mathbf{q}^{kin} = 0$, we have situation of half-suspended NW loaded at the protruding end. This case is sometimes referred as "clamped" or "cantilevered beam" in textbooks.

18.4.2.2 Young Modulus Measurements

As it was previously demonstrated, Young modulus is an essential parameter in the framework of EBT. Typically bulk values or values found in literature are used in tribological studies of NWs. However, considering high scattering of mechanical properties [57] and the fact that they depend on a quality of material and synthesis conditions, it is more accurate to measure Young modulus for a particular set of NWs or even for each single NW separately.

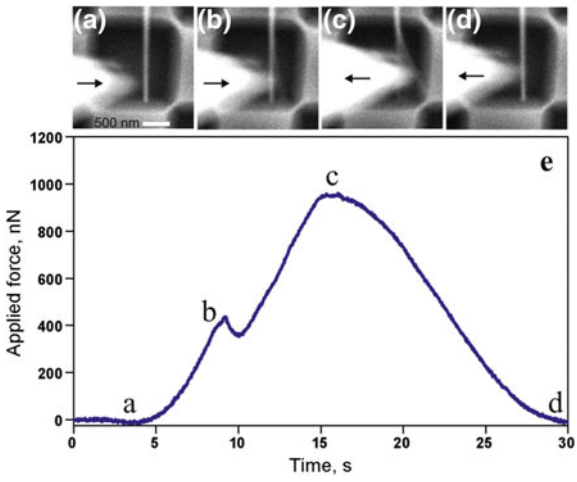


Fig. 18.12 Young's modulus measurement of a CuO NW. Arrows indicate direction of the tip movement. Sequence of SEM images of CuO NW bending (**a**)–(**d**) and the corresponding force curve (**e**)

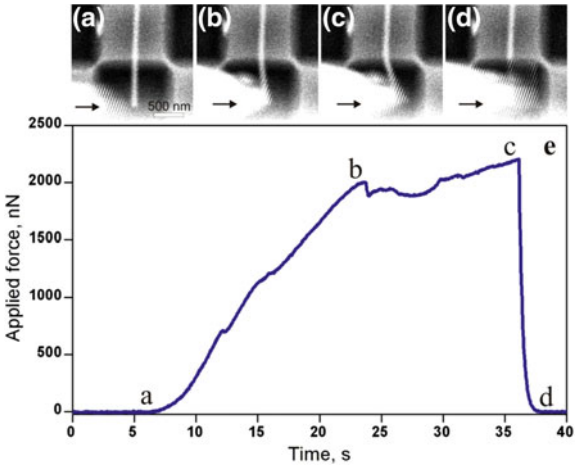


Fig. 18.13 Bending strength measurement of a CuO NW. *Arrows* indicate the direction of tip movement. Sequence of SEM images of bent and broken NW (**a**)–(**d**) and the corresponding force curve (**e**)

Polyakov et al. measured elastic properties and ultimate strength of ZnO and CuO NWs [53, 57] half-suspended on the edge of a trench in a patterned silicon wafer as shown in Fig. 18.12. The cantilevered beam bending technique [58] was then applied to determine Young's modulus and bending strength. The fixed part of the NW stayed motionless due to high adhesion. The suspended part of the NW was pushed by the tip of the QTF-based force sensor. Sensor signal and the grabbed SEM images were recorded simultaneously during the experiment. To process the data in order to find the value of Young's modulus, the SEM image of the NW profile was numerically fitted to the curve given by (18.15). A typical force curve and the corresponding SEM images are presented in Fig. 18.12. Integral static friction force in adhered part was usually higher than applied force and no displacement of NW was observed besides few cases for very short adhered part. Often, internal stress inside NW generated by applied force exceeded NW ultimate strength and caused rupture of NW (Fig. 18.13). Measurements carried out on 9 CuO NWs and 14 ZnO NWs resulted in an average Young's moduli of 204 GPa and 58 GPa respecively. The average bending strengths of CuO and ZnO NWs were found to be 8.2 GPa and 4.4 GPa.

18.4.2.3 Static Friction

High friction involved in NW manipulation on a flat substrate restrains the use of QTF based force sensors. However if experiment performed inside SEM and elastic properties of NW are known the NW itself can serve as a force sensor for measurements of friction forces.

In the method introduced by Polyakov et al. [54] the NW-substrate static friction can be calculated by fitting of an experimentally found bending profile to the theoretical profile of a NW lying on a flat substrate and pushed from one end, while the second end is fixed by static friction. The adhered end of the NW stays motionless as

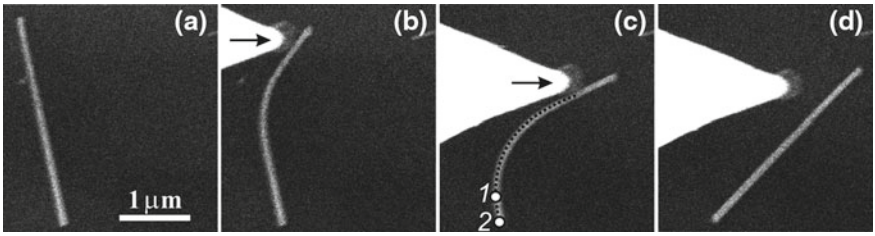


Fig. 18.14 Series of SEM images of the ZnO NW bent by AFM tip from one end while other end is adhered to the substrate by static force. Intact NW (**a**); partially displaced NW (**b**); "most bent state" of the NW with fitting profile lied over the NW image (**c**);completely displaced NW (**d**)

long as the applied external force $\mathbf{F}_{\text{apl-lat}}$ does not exceed the total static friction force. When the elastic force overcomes the total static friction force, the whole NW is displaced. The NW bending profile (or the profile in the "most bent state") just before complete displacement of the NW is then used for fitting to find the maximum static friction and average static friction, which determine the total static friction force.

The described method was applied to study the static friction of ZnO NWs on oxidized silicon wafer. Figure 18.14 represents a typical manipulation experiment. The NW "most bent state" profile of the whole NW was calculated according to (18.15) and laid over the SEM image (Fig. 18.14c). From the elastic deformation of the NW the generated elastic force F and momentum M were calculated, enabling to find counteracting friction forces in the adhered part.

The median value of the averaged static friction was $q^{\text{st}}_{\text{avg}} \sim 5$ nN/nm for 39 measured ZnO NWs. The corresponding value of averaged static shear stress was $\sigma^{\text{st}}_{\text{avg}} \sim 67$ MPa. Many NWs were bent up to 90° deg before displacement, but some NWs broke during bending. Fractured NWs did not necessarily break in the maximal tensile strain region but in any region with a structural defect. The median value of tensile stress in successfully displaced NWs was $\sigma_{\text{tens-displ}} \sim 2.6$ GPa (for a set of 32 NWs), while in broken NWs, $\sigma_{\text{tens-fract}} \sim 3.3$ GPa (for a set of 21 NWs).

18.4.2.4 Effect of a Substrate Roughness on Static Friction

It is known, that for typical surfaces, macroscopic dry friction is usually nearly independent of the surface roughness, unless the surfaces are either very rough or very smooth [59]. However, in nanoscopic systems, roughness of the contacting surfaces is directly associated to the actual contact area and the average distance between them, determining magnitude of van der Waals (vdW) force, which plays the same role as external loading force or gravity at the macroscale. The effect of nanoscale roughness on adhesion for MEMS has been addressed in many works [60, 61].

Polyakov et al. applied previously described "NW loaded at one end" approach for investigation of the effect of surface roughness on the CuO NW-substrate static

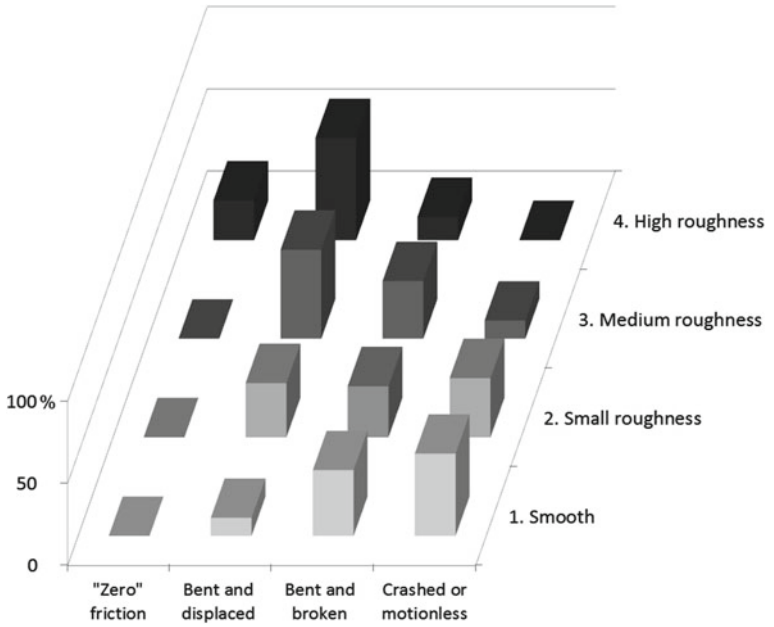


Fig. 18.15 Statistics of NW behaviour during manipulation on substrates of different roughness

friction [54]. Four substrates of roughness R_{RMS} ranging from 0.7 to 16.1 nm were prepared by chemical etching of amorphous silicon. CuO NWs were manipulated according to the scheme shown at Fig. 18.14.

Four distinct scenarios of the NW behaviour during manipulation were observed. At high static friction forces, NWs either could not be displaced at all or broke into shorter fragments without bending ("crashed or motionless"). At lower static friction, NWs were gradually displaced from one end of the NW with the other end fixed by the static friction ("bent and displaced"). If the static friction force exceeded the ultimate strength of the NW, the NW broke ("bent and broken"). If static friction force was smaller than the NW interfacial shear strength, the NW was displaced as a rigid object ("zero friction"). Only in the second scenario ("bent and displaced") the "most bent state" of the NW could be registered and static friction can be analysed quantitatively. Nevertheless, visual observation of NW behaviour on substrate with different roughness provides valuable qualitative information of static friction. Frequency of occurrence of one or another scenario is summarized on the graph in Fig. 18.15.

From the given graph it is evident, that nanoscale roughness has a great impact on static friction for nanostructures. It can be explained considering that the smaller the roughness of the surfaces, the "closer" the two surfaces can be. It leads to higher vdW force, and consequently, higher adhesion and static friction. Numerical

Table 18.2 Effect of substrate roughness on NW-substrate static friction

Substrate Nr.	R_{RMS}, nm	Average static friction, nN/nm
1. Smooth	0.7	5.8
2. Small roughness	2.1	3.9
3. Medium roughness	5.4	1.4
4. High roughness	16.1	0.6

values of the static friction forces calculated for the cases when the NW was bent and displaced using its "most bent state" profile gives an approximately order of magnitude difference between smooth and most rough substrates. Data on averaged static friction $q^{\mathrm{st}}_{\mathrm{avg}}$ for all substrates are summarized in Table 18.2. It is necessary to note that used statistics does not account the case of "zero static friction" (friction below the sensitivity of the method) and motionless NWs (friction beyond the applicability of the method), which means that actual difference should be even higher.

18.4.3 Nanowire Pushed in the Midpoint: Kinetic Friction

At the moment when the nanomanipulator's tip starts to push the NW in the middle point, the static friction force q_{st} is distributed over a portion of the NW of unknown length, which may be smaller than the entire NW length L (Fig. 18.16a). Only a fraction l'of the NW where the applied force exceeds the ultimate value of the distributed static friction force (corresponding to the maximum shear strength) is displaced by the tip (Fig. 18.16b). Static friction acts on the intact part of the NW of length L - l', while kinetic friction affects the displaced portion l'of the NW. For a partially displaced NW, the effective friction force is the combination of the static q_{st} and kinetic q_{kin} friction forces. It is difficult to measure length of the kinetic part of the NW and to analyse the contribution of the static and kinetic friction forces during the initial stage of NW displacement. Upon further displacement of the NW (uniform translation), we consider the friction force as purely kinetic and uniformly distributed along the length of the NW (Fig. 18.16c).

18.4.3.1 Model

If a NW is being uniformly dragged at its midpoint without rolling and all parts of the NW have the same constant velocity, the equilibrium equations (18.13) and (18.14) are still applicable due to the principle of relativity.

In this case, the profile of the deformed NW is determined by the balance of the external driving tip force, the kinetic friction between the NW and the substrate and

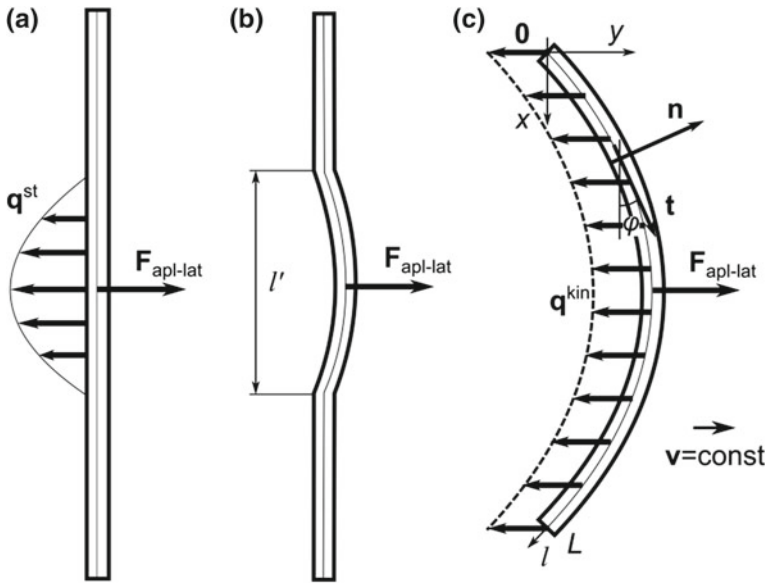


Fig. 18.16 NW pushed at the midpoint by concentrated force $F_{apl-lat}$. NW does not move until the applied force exceeds maximal static friction or shear strength (**a**). Partially displaced NW (**b**). Entirely displaced NW. Kinetic friction force q_{kin} equally distributed over NW length (**c**). Fixed coordinate system Oxyz and local coordinate basis (**t**,**n**) along the NW axis l. Angle between the tangent vector **t** and axis Ox is denoted as φ

the intrinsic elastic forces of the NW. The distributed driving force $\mathbf{F}_{\text{apl-lat}}$ can be modelled via the delta function, and the kinetic friction $\mathbf{q}^{\text{kin}}$ maintains a constant vector opposite to the direction of motion and $\mathbf{F}_{\text{apl-lat}}$ (Fig. 18.11):

$$\mathbf{f} = \mathbf{q}^{\text{kin}} + \mathbf{F}_{\text{apl-lat}} \cdot \delta\left(l - \frac{L}{2}\right), \tag{18.20}$$

where $\delta(x)$ is Dirac's delta-function.

The condition of zero total force yields $\mathbf{F}_{\text{apl-lat}} = -\mathbf{q}^{\text{kin}} L$.

Zero elastic force and momentum at the free ends of the NW dictate the boundary conditions:

$$\mathbf{F}|_{l=0} = \mathbf{F}|_{l=L} = 0 \tag{18.21}$$

$$\mathbf{M}|_{l=0} = \mathbf{M}|_{l=L} = 0. \tag{18.22}$$

The differential equation of "kinetic" equilibrium of the NW on the interval (0, L) directly follows from (18.5), (18.9) and (18.10):

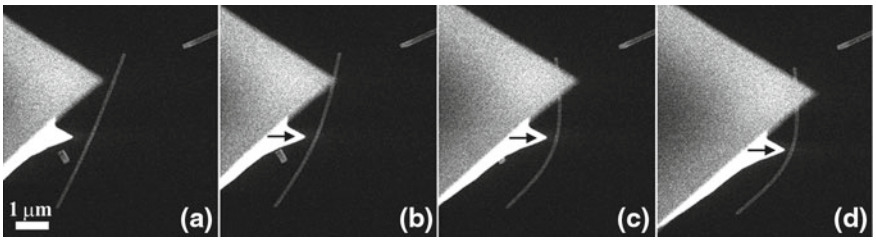


Fig. 18.17 Sequence of SEM images of the ZnO NW dragged by AFM tip on HOPG substrate. The *arrow* indicates the direction of tip movement (**a**); partially displaced NW (**b**); completely displaced NW (**c**); constant characteristic profile during NW motion (**d**)

$$IE\frac{d^2\varphi}{dl^2} = -q^{\text{kin}}\left[l - LH\left(l - \frac{L}{2}\right)\right]\cos\varphi, \tag{18.23}$$

where $H(x)$ is the Heaviside step function. The (18.23) can be solved numerically in order to obtain the NW profile. It is easy to check that the solution of (18.23) together with the initial condition $\varphi\prime(0) = 0$ fully complies with the free boundary conditions (18.21) and (18.22).

18.4.3.2 Kinetic friction

Polyakov et al. utilized the described model [62] to calculate distributed kinetic friction of ZnO NWs on a flat surfaces (HOPG or silicon wafer) inside SEM. The NW was translated along the substrate surface by pushing it at its midpoint with AFM tip (Fig. 18.17). NW profile during translation was used to determine the kinetic friction force distributed along the NW by fitting it to theoretical profile calculated using (18.23). The measurements were performed for a set of 8 NWs on silicon wafer and for 6 NWs on HOPG. Interfacial shear stress σ_{kin} were found to be 3.2 MPa and 2.75 MPa respectively. These values are in average about twenty times lower than static interfacial stress ($\sigma^{\text{st}}_{\text{avg}} \sim 67$ MPa) for ZnO NWs on silicon wafer [54]. It agrees well with known fact that for nanoscale objects static friction is commonly significantly higher than kinetic friction [63].

18.4.4 Redistributed Static Friction of a Bent Nanowire Relaxed After Manipulation

Static friction can hold an elastically bent NW on a flat surface in a deformed state after the removal of external load. The first attempt to use the profile of the elastically bent NW for calculation of static friction was made by Bordag et al. [64]. They assumed the most bent part of an arc-shaped InAs NW to be circular and used the

bending curvature as a basis for the static friction calculation. Then Strus et al. [65] introduced a method, where an AFM image of bent CNT was used for determining the static friction and elastic stresses in the framework of the EBT. Stan et al. [66] used parabolas to fit through a defined region of the NW centreline for the analysis with Strus' equilibrium expressions.

Below we will consider elaborated method proposed by Dorogin et al. [9, 67] for modelling of the distributed static friction force in the frameworks of the EBT. Method benefits from fully satisfied boundary conditions and complete equilibrium equations, which enables to apply it for analysis of complex shaped NWs.

18.4.4.1 Model

The equations of equilibrium for an elastically bent NW held in a bent state by distributed static friction force $\mathbf{f}(l) = \mathbf{q}^{\mathrm{st}}(l)$ give (see Fig. 18.9):

$$\frac{dF_t}{dl} - \kappa F_n - q_t^{\mathrm{st}} \tag{18.24}$$

$$\frac{dF_n}{dl} + \kappa F_t - q_n^{\mathrm{st}} \tag{18.25}$$

$$EI\frac{d\kappa}{dl} = -F_n \tag{18.26}$$

where F_t and F_n are the projections of elastic force $\mathbf{F}$, q_t^{st} and q_n^{st} are the projections of $\mathbf{q}^{\mathrm{st}}$ to the local coordinates ($\mathbf{t}$, $\mathbf{n}$).

The complete set of boundary conditions (18.21) and (18.22) applied to the system of (18.24), (18.25) and (18.26) yields:

$$\kappa|_{l=0} = \kappa|_{l=L} = 0 \tag{18.27}$$

$$\frac{d\kappa}{dl}|_{l=0}\,\frac{d\kappa}{dl}|_{l=L} = 0. \tag{18.28}$$

We will extend the system of 3 equilibrium equation (18.24), (18.25) and (18.26) for 4 unknown functions (F_t, F_n, q_t^{st} and q_n^{st}) with the condition of absence of the tangential component of the friction $q_t^{\mathrm{st}} = 0$. Thus the system of equations has become complete and can be easily solved for F_t and q_t^{st}:

$$F_t = -EI\int_0^l \kappa\frac{d\kappa}{dl}dl = -EI\frac{\kappa^2}{2} \tag{18.29}$$

$$q_n^{\mathrm{st}} = EI\left(\frac{d^2\kappa}{dl^2} + \frac{\kappa^3}{2}\right) \tag{18.30}$$

with the initial condition $F_t|_{l=0} = 0$. The absence of a tangential friction component does not lead to the vanishing of F_t which is directly linked with F_n according to the (18.24) and is necessary for the exact NW equilibrium.

It is important to note that the assumption was dictated by the consideration that the direction of $\mathbf{q}^{\mathrm{st}}$ should be close to the direction at which the NW tends to unbend. This "unbending" direction correspondingly lies close to normal to the NW's line. Formally it means that the integral contribution of along the length of the NW is much smaller than that of q_n^{st}. Moreover, the tangential component of the force does not contribute to the bending in the framework of the current model (as can be seen from (18.10)) and is considered to be a small effect of higher order.

18.4.4.2 Redistributed Static Friction

By using the above described model the equilibrium of NW elastically bent on a flat surface can be employed for calculation of static force distribution $\mathbf{q}^{\mathrm{st}}(l)$ as shown in (18.30), which in turn is determined by the curvature $\kappa(l)$. Therefore, extraction of a differentiable curvature function is needed, which can be realized in the following procedure: (1) NW centreline vectorization; and (2) interpolation of tangent angle $\varphi(l)$ and curvature $\kappa\ (l)$.

The first stage of centreline vectorization belongs merely to digital image processing and will not be considered. The output of the vectorization stage is a set of n discrete points (x_i, y_i), representing the NW centreline.

In the second stage, differential analysis of the centreline is needed to convert the Cartesian coordinates (x_i, y_i) to natural coordinates (l_i, φ_i), which are more suitable for the subsequent calculations. The discrete mesh for subsequent interpolation of $\varphi\ (l)$ is produced with the following expressions:

$$\varphi_i = \tan^{-1}\left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i}\right), \quad 1 \le i < n-1 \tag{18.31}$$

$$l_i = l_{i-1} + \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}, \quad 2 \le i < n \tag{18.32}$$

where $l_1 = 0$.

Correct choice of the interpolation function is important due to the errors, which may arise during the identification and differentiation of the centreline curve. Regular polynomial interpolation would likely break the boundary conditions (18.27) and (18.28). The interpolation of $\varphi\ (l)$ can be performed through the curvature $\kappa\ (l)$ with the linear combination of the selected polynomial functions as follows:

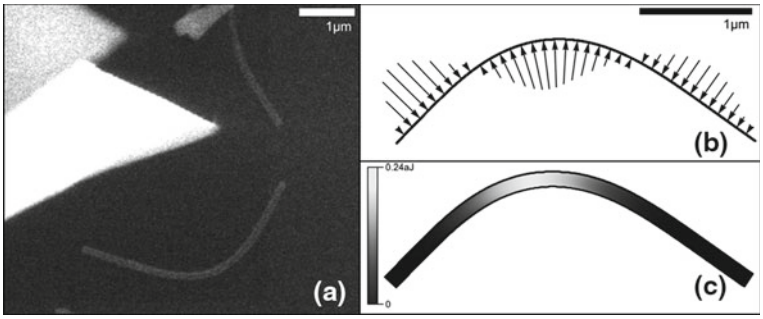


Fig. 18.18 SEM image of ZnO NW in a bent state self-balanced by static friction (**a**). Calculated distributed static friction along the NW (**b**). Elastic energy distribution along the NW (**c**)

$$\frac{d\varphi}{dl} = \kappa(l) = \sum_{i=0}^{N} A_i l^{i+2} (L-l)^2, \tag{18.33}$$

where A_i are coefficients to be determined by the numerical interpolation. It is apparent that φ (l) defined according to (18.33) complies with the boundary conditions (18.27) and (18.28).

NW-substrate interaction can be characterized directly by the friction distribution and indirectly by the elastic energy stored in the bent NW. The stronger the NW-substrate friction interaction, the higher the strain energy can be conserved in the elastically bent NW. It is notable that energy distribution can be calculated with higher accuracy compared to the friction force distribution due to its dependence on the NW curvature κ (l) and not on the curvature's derivatives.

Dorogin et al. [9] applied the described method to ZnO NW manipulated to arc shape on a silicon wafer and calculated the static friction force q_n^{st} together with distributed elastic energy u^{st} along the wire. NW was pushed near its midpoint perpendicular to the NW axis by an AFM tip inside SEM. Over several microns of travel, the initially straight profile of the NW transformed into the characteristic arc shape. After the removal of the tip, the NW relaxed and came to equilibrium non-straight shape determined by the interplay of elastic and friction forces (Fig. 18.18a).

From the Fig. 18.18b it is clearly seen that static friction distribution has non-trivial wavy profile. While u^{st} has only one extremum near the NW centre, the distributed friction q_n^{st} has three extrema of different signs. This evidences a non-local character of the NW-substrate friction. Moreover, the model enables certain cases when $\left|q_n^{\mathrm{st}}\right|$ has absolute maxima near the NW ends, rather than in the middle where the absolute maximum strain is located. Consequently, it is impossible to judge where the maximum friction is applied without detailed analysis in the framework of the presented distributed NW friction model.

18.4.5 Specific Problems of Manipulations Inside SEM

There are several issues to keep in mind when planning experiments inside SEM. First, typical scanning rate of the electron beam is limited to a few Hz, therefore only relatively slow processes can be visualized. Another limitation is associated with SEM resolution. Resolution and signal intensity are reversely proportional to scanning speed. Visible dimensions of the nanostructures at magnification close to the resolution limit are sensitive to brightness/contrast settings of a microscope. It can be problematic to identify the exact shape of the NW cross section, for example, to distinguish between square and rectangular cross sections or estimate degree of hexagon or pentagon distortions. Precise determination of NWs diameter and cross section, however, is essential for correct measurement of the contact area and frictional forces in the framework of EBT.

One more thing that should be taken into account is effect of electron beam on a substrate surface and nanostructures. As it is well known, melting temperature of nanostructures decreases with decreasing their diameter [68]. Focused e-beam is capable to introduce significant energy and cause partial melting of nanostructures. Additional effects can be substrate surface activation or electrostatic charging, which also can influence nanotribological experiments [69]. Degradation of some materials under electron beam should be considered as an another issue. For example, electron beam produce severe defects in carbon based materials like graphene, graphite and carbon nanotubes, or self-assembled monolayers of organic molecules, which are frequently used for tribological experiments [70].

Another important problem is e-beam assisted carbon deposition [71]. High energy electrons can decompose hydrocarbons molecules, which are present in the vacuum chamber and amorphous carbon can be deposited on the areas exposed to electron beam for prolonged periods. Amount of deposited carbon depends on the vacuum level, e-beam energy, the presence of hydrocarbon inside the chamber and exposition time.

18.5 Outlook

It was demonstrated that visual guidance during the nanomanipulations inside SEM is extremely helpful and provides essential information on behaviour of manipulated nanostructures. Method is well suitable for manipulation of nanoparticles and especially fruitful for elongated objects. Inside SEM technique is not limited to in-plane manipulations, but also capable of performing experiments in three dimensions. Moreover, experiments are less time-consuming in comparison to AFM manipulations due to separation of visualization and manipulation functions. Besides to the apparent advantages, there are certain peculiarities and limitations that should be taken into account when planning experiment inside SEM.

References

1. Y. Mo, K.T. Turner, I. Szlufarska, Nature **457**, 1116–1119 (2009). doi:10.1038/nature07748
2. L. Liu, P. Peng, A. Hu, G. Zou, W.W. Duley et al., Appl. Phys. Lett. **102**, 073107 (2013). doi:10.1063/1.4790189
3. L. Wang, H.-Y. Park, S. Lim, M. Schadt, D. Mott, J. Luo, X. Wang, C.-J. Zhong, J. Mater. Chem. **18**, 2629–2635 (2008)
4. L. Dorogin, S. Vlassov, A. Kolesnikova, I. Kink, R. Lõhmus, A. Romanov, Crystal mismatched layers in pentagonal nanorods and nanoparticles. Physica Status Solidi B-Basic. Solid State Physics **247**(2), 288–298 (2010)
5. D. Hidayat, A. Purwanto, W.-N. Wangc, K. Okuyama, Mater. Res. Bull. **45**, 165–173 (2010)
6. D. Seo, C.I. Yoo, I.S. Chung, S.M. Park, S. Ryu, H. Song, J. Phys. Chem. **C112**, 2469–2475 (2008)
7. C.-L. Chiang, M.-B. Hsu, L.-B. Lai, J. Solid State Chem. **177**, 3891–3895 (2004)
8. C. Cao, S. Park, S. Sim, J. Colloid Interface Sci. **322**, 152–157 (2008)
9. L.M. Dorogin, B. Polyakov, A. Petruhins, S. Vlassov, R. Lõhmus, I. Kink, A.E. Romanov, J. Mater. Res. **27**, 580–585 (2012)
10. BRR—A SEM-AFM integration for Zeiss scanning electron microscope. DME Danish Micro Engineering A/S. http://www.dme-spm.dk/
11. Combined AFM FIB and AFM SEM. Nanonics Imaging Ltd. http://www.nanonics.co.il/products/nsom-spm-systems/combined-afm-fib-and-afm-sem.html
12. http://www.smaract.de/
13. D. Erts, A. Lõhmus, R. Lõhmus, H. Olin, A.V. Pokropivny, L. Ryen, K. Svensson, Appl. Surf. Sci. **188**(3–4), 460–466 (2002)
14. Attocube systems AG. http://www.attocube.com/
15. Fundamentals of Piezoelectric Actuators, PI Ceramic GmbH. http://www.piceramic.com/
16. R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (Cambridge University Press, Cambridge, 1994)
17. M. Troyon, H.N. Lei, Z. Wang, G. Shang, A scanning force microscope combined with a scanning electron microscope for multidimensional data analysis. Scan. Microsc. **12**(1), 139–148 (1998)
18. K. Fukushima, D. Saya, H. Kawakatsu, Development of a Versatile atomic force microscope within a scanning electron microscope. Jp. J. Appl. Phys. **39**, 3747–3749 (2000)
19. I. Joachimsthaler, R. Heiderhoff, L.J. Balk, A universal scanningprobe-microscope-based hybrid system. Meas. Sci. Technol. **14**(1), 87–96 (2003)
20. U. Mick, V. Eichhorn, T. Wortmann, C. Diederichs, S. Fatikow, Combined Nanorobotic AFM/SEM System as Novel Toolbox for Automated Hybrid Analysis and Manipulation of Nanoscale Objects. *2010 IEEE International Conference on Robotics and Automation*
21. U. Stahl, C.W. Yuan, A.L. Lozanne, M. Tortonese, Atomic force microscope using piezoresistive cantilevers and combined with a scanning electron microscope. Appl. Phys. Lett. **65**(28), 2878–2880 (1994)
22. M. Barbic, Sens. Actuators, A **136**, 564–566 (2007)
23. S. Fain Jr, K. Barry, M. Bush, B. Pittenger, R. Louie, Appl. Phys. Lett. **76**, 930 (2000)
24. C. Su, L. Huang, K. Kjoller, Ultramicroscopy **100**, 233–239 (2004)
25. H. Hidaa, M. Shikida, K. Fukuzawa, S. Murakami, Ke. Sato, K. Asaumi, Y. Iriye, Ka. Sato, Sens. Actuators, A **148**, 311–318 (2008)
26. V. ThanhTung, S.A. Chizhik, T. XuanHoai, N. TrongTinh, V.V. Chikunov, in *Tuning Fork Scanning Probe Microscopes*, ed. by V. Nalladega, Applications for the Nano-Analysis of the Material Surface and Local Physico-Mechanical Properties, Scanning Probe Microscopy-Physical Property Characterization at Nanoscale. ISBN: 978-953-51-0576-3, InTech (2012)
27. S. Kerfriden, A. Nahlé, S. Campbell, F. Walsh, J. Smith, Electrochimica Acta **43**(12–13), 1939–1944 (1998)
28. S. Vlassov, B. Polyakov, L.M. Dorogin, A. Lohmus, A.E. Romanov, I. Kink, E. Gnecco, R. Lohmus, Solid State Commun. **151**, 688 (2011)

29. AdvancedTEC ^{TM} Cont. http://www.nanosensors.com/
30. Akiyama-Probe (A-Probe) guide. NANOSENSORS, NanoWorld AG. http://www.akiyamaprobe.com
31. J.L. Arlett, J.R. Maloney, B. Gudlewski, M. Muluneh, M.L. Roukes, Nano Lett. **6**, 1000 (2006)
32. M. Li, H.X. Tang, M.L. Roukes, Nat. Nanotechnol. **2**, 114 (2007)
33. C. Baur, A. Bugacov, B.E. Koel, A. Madhukar, N. Montoya, T.R. Ramachandran, A.A.G. Requicha, R. Resch, P. Will, Nanotechnology **9**, 360 (1998)
34. D. Dietzel, M. Feldmann, H. Fuchs, U.D. Schwarz, A. Schirmeisen, Appl. Phys. Lett. **95**(5) (2009)
35. K. Mougin, E. Gnecco, A. Rao, M. Cuberes, S. Jayaraman, E. McFarland, H. Haidara, E. Meyer, Langmuir **24**, 1577 (2008)
36. S. Kim, D.C. Ratchford, X. Li, ACS Nano **3**, 2989–2994 (2009)
37. A.M. Homola, J.N. Israelachvili, M.L. Gee, P.M. McGuiggan, J. Tribol. **111**, 675 (1989)
38. R.W. Carpick, M. Salmeron, Chem. Rev. **97**, 1163–1194 (1997)
39. A.H. Cottrell, *Dislocations and Plastic Flow in Crystals* (Oxford University Press, Oxford, UK, 1953)
40. J.P. Hirth, J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1968). 780 p
41. K.L. Johnson, K. Kendall, A.D. Roberts, Proc. B. Soc. Lond. A. **324**, 301–313 (1971)
42. B.V. Derjaguin, V.M. Müller, Y.P. Toporov, J. Colloid Interface Sci. **53**, 314 (1975)
43. D. Tabor, J. Colloid Interface Sci. **58**, 2–13 (1977)
44. D. Dietzel, C. Ritter, T. Monninghoff, H. Fuchs, A. Schirmeisen, U.D. Schwarz, Phys. Rev. Lett. **101**, 125505 (2008)
45. L. Kondic, J.A. Diez, Phys. Rev. **E79**, 026302 (2009)
46. D.R. Smith, F.R. Fickett, Low-temperature properties of silver. J. Res. Natl. Inst. Stand. Technol. **100**, 119 (1995)
47. M. Manoharan, A. Desai, G. Neely, M. Haque, J. Nanomater. Article ID 849745 (2008)
48. J. Hsu, S. Chang, Surface adhesion between hexagonal boron nitride nanotubes and silicon based on lateral force microscopy. Appl. Surf. Sci. **256**, 1769–1773 (2010)
49. M. Falvo, J. Steele, I.I. Taylor, R.R. Superfine, Evidence of commensurate contact and rolling motion: AFM manipulation studies of carbon nanotubes on HOPG. Tribol. Lett. **9**, 73–76 (2000)
50. R. Mohan, Y. Liang, Cutting Edge Nanotechnology (InTech, 2010), Chap. 10
51. M.A. Schubert, S. Senz, M. Alexe, D. Hesse, U. Gösele, Appl. Phys. Lett. **92**, 122904 (2008)
52. S. Timoshenko, J. N. Goodier, *Theory of Elasticity*, 2nd. ed. (McGraw-Hill Book Company, 1951), pp. 316–342
53. B. Polyakov, L. Dorogin, S. Vlassov, I. Kink, A. Lohmus, A. Romanov, R. Lohmus, Solid State Commun. **151**, 1244–1247 (2011)
54. B. Polyakov, L. Dorogin, S. Vlassov, A.E. Romanov, R. Lohmus, I. Kink, Micron **43**, 1140–1146 (2012)
55. B. Polyakov, L.M. Dorogin, A. Lõhmus, A.E. Romanov, R. Lõhmus, Appl. Surf. Sci. **258**, 3227 (2012)
56. L. Landau, E. Lifshitz, *Theory of Elasticity*, 3rd edn. (Butterworth-Heinemann, Oxford, 1986), pp. 70–76
57. B. Polyakov, L.M. Dorogin, S. Vlassov, M. Antsov, P. Kulis, I. Kink, R. Lohmus, In situ measurements of ultimate bending strength of CuO and ZnO nanowires. Eur. Phys. J. B **85**, 366 (2012)
58. E.C.C.M. Silva, L. Tong, S. Yip, K.J. Van Vliet, Small **2**, 239–243 (2006)
59. B. Persson, *Sliding Friction*, 2nd edn. (Springer, Berlin Heidelberg New York, 2000)
60. N. Tambe, B. Bhushan, Nanotechnology **15**, 1561 (2004)
61. F. Delrio, M. de Boer, J. Knapp, E. Davidreed, P. Clews, M. Dunn, Nat. Mater. **4**, 629 (2005)
62. B. Polyakov et al., Appl. Surf. Sci. **258**, 3227 (2012)
63. D. Dietzel, M. Feldmann, H. Fuchs, U. Schwarz, A. Schirmeisen, Transition from static to kinetic friction of metallic nanoparticles. Appl. Phys. Lett. **95**, 053104 (2009)

64. M. Bordag, A. Ribayrol, G. Conache, L.E. Frcberg, S. Gray, L. Samuelson, L. Montelius, H. Pettersson, Small **3**, 1398–1401 (2007)
65. M. Strus, R. Lahiji, P. Ares, V. Lopez, A. Raman, R. Reifenberger, Nanotechnology **20**, 385709 (2009)
66. G. Stan, S. Krylyuk, A.V. Davydov, R.F. Cook, Bending manipulation and measurements of fracture strength of silicon and oxidized silicon nanowires by atomic force microscopy. J. Mater. Res. **27** (2012)
67. L.M. Dorogin, S. Vlassov, B. Polyakov, M. Antsov, R. Lõhmus, I. Kink, A.E. Romanov, Real-time manipulation of ZnO nanowires on a flat surface employed for tribological measurements: experimental methods and modeling. Phys. Status Solidi B, 1–13 (2012)
68. W.H. Qi, Phys. B: Condens. Mater **368**(1–4), 46–50 (2005)
69. G. Conache, A. Ribayrol, L.E. Froberg, M.T. Borgstrom, L. Samuelson, L. Montelius, H. Pettersson, S.M. Gray, Phys. Rev. B **82**, 035403 (2010). doi:10.1103/PhysRevB.82.035403
70. D. Teweldebrhan, A.A. Balandin, Appl. Phys. Lett. **94**, 013101 (2009)
71. S. Zaitsev, O. Shtempluck, E. Buks, Sens. Actuators, A: Phys. **179**, 237–241 (2012)

Chapter 19
Driven Colloidal Monolayers: Static and Dynamic Friction

Andrea Vanossi, Nicola Manini and Erio Tosatti

Abstract Trapping and dragging colloidal monolayers in two-dimensional optical lattices is offering the possibility to mimic friction between crystals (or even quasicrystals) visualizing directly the intimate mechanisms of sliding friction, with the additional possibility to change parameters freely, and to compare directly experiment with theory. Realistic simulations, which we review here, make a number of predictions about static features and dynamic sliding and reproduce well recent observations. Together, they provide a first demonstration of the potential impact of colloid dynamics in nanotribology.

19.1 Introduction

The intimate understanding of sliding friction, a central player in the physics and technology of an enormous variety of systems, ranging from nanotribology to mesoscale and macroscale sliding [1, 2], is historically hampered by a number of well konwn difficulties. One of them is the practical inaccessibility of the buried interface between the moving bodies—with few exceptions, we can only hypothesize about its nature and behavior during sliding. Another is the general impossibility to fully control the

A. Vanossi (✉) · E. Tosatti
CNR-IOM Democritos National Simulation Center,
Via Bonomea 265, 34136 Trieste, Italy
e-mail: vanossi@sissa.it

A. Vanossi · E. Tosatti
International School for Advanced Studies (SISSA),
Via Bonomea 265, 34136 Trieste, Italy

N. Manini
Dipartimento di Fisica,Università degli Studi di Milano,
Via Celoria 16, 20133 Milano, Italy
e-mail: nicola.manini@mi.infm.it

E. Tosatti
International Center for Theoretical Physics (ICTP),
Strada Costiera 11, 34151 Trieste, Italy
e-mail: tosatti@sissa.it

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_19

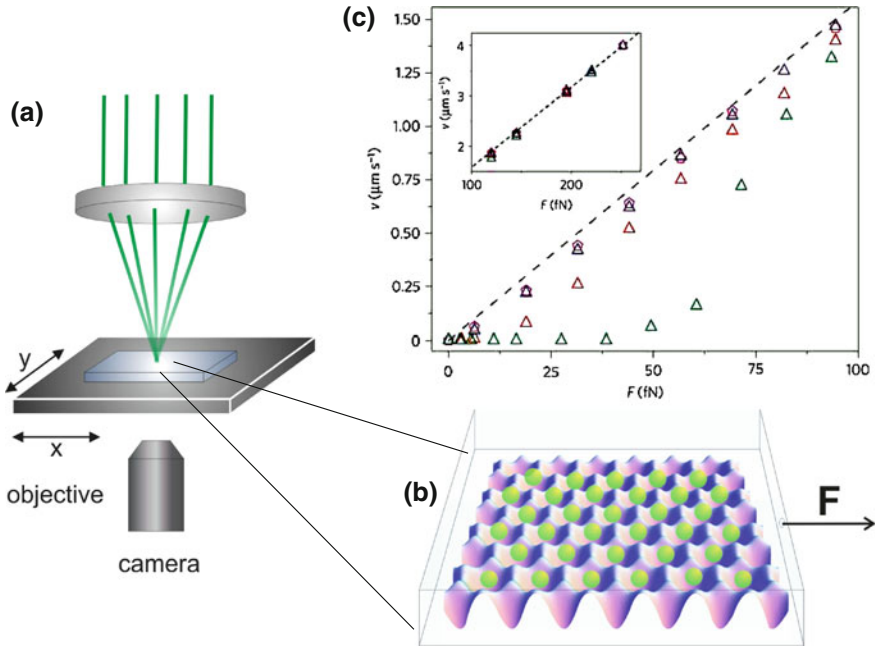


Fig. 19.1 **a** A scheme of the experimental setup of [3] with the sample cell mounted on the piezo table, the interfering laser beams, and the camera interfaced to a microscope. **b** The colloids (*spheres*) in the triangular-symmetry corrugated potential-energy profile. **c** The observed mean colloid velocity as a function of the driving force F for different (commensurate or antisolitonic) configurations

detailed nature, morphology, and geometric parameters of the sliders; thus for example, even perfectly periodic, defect-free sliding surfaces have essentially only been accessible theoretically. If we knew and, even better, if we could control the properties and the relative asperity parameters of the sliders, our physical understanding would be greatly enhanced, also disclosing possibilities to tune friction in nano and mesoscopic systems and devices.

As Bohlein et al. [3] showed (Fig. 19.1), two dimensional (2D) colloid crystalline monolayers can be forced by the flow of their embedding fluid to slide against a laser-generated optical lattice potential mimicking the interface "corrugation" potential in ordinary sliding friction. The external driving force, the interparticle interactions, and especially the corrugation potential are all under control, the latter ranging from weak to strong, and from periodic, to quasi-periodic [4, 5], and in principle to more complex types too. Contrary to established techniques in meso and nanosize sliding friction (Atomic Force Microscope, Surface Force Apparatus, Quartz Crystal Microbalance) [6], which address the tribological response in terms of averaged physical quantities (overall static and kinetic friction, mean velocities, slip lengths and slip times, etc.),

in colloid sliding every individual particle can be visualized and followed in real time, sharing a privilege hitherto restricted to molecular-dynamics (MD) simulations [7–9].

Materializing concepts long-anticipated theoretically [10, 11], the data of Bohlein et al. show that the sliding of a flat crystalline lattice on a perfectly periodic substrate takes place through the motion of soliton or antisoliton superstructures (also known in one dimension (1D) as kinks or antikinks), density modulation domain walls which represent the misfit dislocations of the two lattices that are incommensurate in their mutual registry. While forming regular static Moiré superstructure patterns when at rest, the solitons constitute the actual mobile entities during depinning and sliding, and are essential for "superlubricity" [12]—i.e., zero static friction—of hard incommensurate sliders. When the two periodic sliders are perfectly lattice matched, and thus solitons are absent, or when they are incommensurate and solitons exist but are pinned due to the system softness, the colloids and the periodic potential are initially stuck together. Only after a static friction force F_s is overcome, solitons appear (or depin if they exist but are pinned) unlocking the colloids away from the corrugation potential, after which the sliding can initiate.

In the present chapter we review recent work where a full understanding of the colloidal patterns and dissipation mechanism was achieved [13, 14]. Section 19.2 describes, in short, the driven colloid sliding setup as a model system for nanotribology. Sections 19.3 and 19.4 detail the simulation model and the protocol adopted. Section 19.5 outlines the main simulation results. A detailed analysis in terms of friction and dissipation is carried out in Sect. 19.6. The chapter closes with a final discussion, Sect. 19.7.

19.2 Sliding of a Colloid Monolayer on Laser-Created Periodic Potentials

Recent experiments show that the field of friction can now benefit from the opportunities offered by trapping and handling nano (or micro) particles with potentials artificially created by interfering lasers; a technique originally applied to cold atoms [15]. Soft-matter systems, and in particular colloidal suspensions of particles with tunable interactions, provide a way to study condensed-phase phenomena with a direct optical single-particle resolution usually unavailable to atomistic systems.

Bohlein and colleagues [3] have shown how colloidal particle crystals trapped in laser interference fields can cast new light on elementary frictional processes in ideally controlled sliding geometries and operative conditions. Specifically, they used carboxylated polystyrene particles with a radius $R = 1.95\,\mu\text{m}$ and a polydispersity of 2.7 %. Owing to functionalization with –COOH groups, the particles are negatively charged in water and interact through a screened Coulomb repulsion. The Debye screening length was determined to be $\lambda_D = 160$ nm [16]. Optical interference patterns are created using a laser beam divided into three or five beams, depending

on the desired substrate geometry, and overlapped in the sample plane. The resulting length scales of the interference pattern are adjusted by the angle of incidence. With a further optical forces pushing the colloids toward the bottom of the sample cell, the out-of-plane particle fluctuations are reduced to less than 5 % of the particle diameter. Thus, the colloids form an almost perfect 2D triangular crystal at the bottom of the sample cell. During motion, the trajectory of each colloid particle can be determined in real space and time with great precision thanks to digital video microscopy.

To apply lateral forces to the colloidal monolayer to enact the sliding, the sample cell is mounted on a piezo table able to be laterally translated with nanometer accuracy. When the sample cell moves with velocity $\mathbf{u}$, a viscous Stokes force $\eta(\mathbf{u}-\mathbf{v})$ acts on a particle of velocity $\mathbf{v}$. The effective drag coefficient η is determined by measuring the short-time particle diffusion coefficient D within the monolayer and exploiting the Einstein relation $\eta = k_B T/D$. The tribological response of the colloidal monolayer can be characterized by its mobility $\mu = |\langle \mathbf{v} \rangle|/|\mathbf{F}|$ where $\langle \mathbf{v} \rangle$ is the mean drift velocity of the colloids at the applied driving force $\mathbf{F} = \eta \mathbf{u}$.

19.3 Molecular Dynamics Simulation Model

To simulate these experiments, Vanossi et al. describe the colloidal particles as classical point-like objects moving under the action of three separate forces: the external force, the mutual repulsion, and the interaction with the viscous fluid in which they are immersed.

The equation of motion for the j-th particle is:

$$m\ddot{\mathbf{r}}_j + \eta(\dot{\mathbf{r}}_j - \mathrm{v}_d\hat{\mathbf{x}}) = -\nabla_{\mathbf{r}_j}(U_2 + U_{\mathrm{ext}}) + \tilde{\mathbf{f}}_j \,. \tag{19.1}$$

Here $\mathbf{r}_j$ is a 2-dimensional displacement vector relative to the center of the cell; $\mathbf{u} = \mathrm{v}_d\hat{\mathbf{x}}$ is the fluid drift velocity, giving rise to the Stokes' driving force $F = \eta v_d$ experienced by all the colloidal particles. The viscous term $\eta\dot{\mathbf{r}}_j$ accounts for the dissipation of the colloids kinetic energy into the thermal bath provided by the surrounding water.

Given the slow motion ($v_d \simeq 1\ \mu\mathrm{m\,s}^{-1}$) of a colloidal particle in the liquid, the inertial term $m\ddot{\mathbf{r}}_j$ can be neglected, and a diffusive motion can be assumed, with an appropriate choice of η (given in Table 19.1).

Table 19.1 Numerical parameters adopted in the simulation, (19.3), (19.5), (19.6), and (19.7), expressed in model units, to be scaled according to Table 19.3

	η	N	Q	λ_D	A_c	σ	n	U_0	L_x	L_y
Cluster	2.8	28861	10^{13}	0.03	1200	1200	3	0.1	500	500
Bulk	2.8	28080	10^{13}	0.03	0.0	1200	3	0.1	156	$90\sqrt{3}$

L_x and L_y are the sides of the rectangular simulation supercell respectively

Thermal (Brownian) motion of individual colloids due to random forces from the fluid can be simulated by a Langevin random force $\tilde{\mathbf{f}}_j$, which as usual is assumed to be Gaussian distributed and uncorrelated: $\langle \tilde{f}_{j,\alpha}(t)\, \tilde{f}_{j',\alpha'}(t') \rangle = \delta_{jj'}\delta_{\alpha\alpha'} 2k_{\mathrm{B}} T \eta \delta(t-t')$. As shown by Hasnain et al. [14] in the fully commensurate case, room-temperature thermal effect are of great importance, since they control the nucleation process which is at the origin of sliding. Conversely, in the hard incommensurate regime, which is the case of Bohlein's experiments, thermal fluctuations play a lesser role, and indeed most phenomena are accounted for by $T = 0$ deterministic equations, and the resulting simulations offer a somewhat neater picture.

The 2-body interaction energy is

$$U_2 = \sum_{j<j'}^{N} V(|\mathbf{r}_j - \mathbf{r}'_j|)\,. \tag{19.2}$$

The repulsion between charged particles is a screened Coulomb repulsion, which varies with interparticle distance as a Yukawa-type potential

$$V(r) = \frac{Q}{r} \exp(-r/\lambda_D)\,. \tag{19.3}$$

Typical nearest-neighbor experimental colloid separations are $r \simeq 5.7\ \mu\mathrm{m} \simeq 30\lambda_D$ [3]. Colloids rarely approach much closer than that, because of the violently increasing repulsion $V(r)$ that keeps them apart well before the hard-core repulsion sets in at the colloid diameter $\simeq 3.9\ \mu\mathrm{m}$. In simulation an explicit inclusion of the hard-core term is thus unnecessary, because it would cause no change in the colloid trajectories. Given the moderate volume fraction occupied by the colloidal particles in the solution and the adiabatically slow motions under study, it is also appropriate to neglect hydrodynamic forces [17], that would become relevant only at much denser/faster regimes.

The 1-body external potential energy

$$U_{\mathrm{ext}} = \sum_{j}^{N} V_{\mathrm{ext}}(\mathbf{r}_j) \tag{19.4}$$

is introduced by a laser field, and can be experimentally shaped with substantial freedom. The following spatial variation is assumed:

$$V_{\mathrm{ext}}(\mathbf{r}) = G(\mathbf{r})[-A_c + U_0\, W_n(\mathbf{r})]\,, \tag{19.5}$$

where

$$G(\mathbf{r}) = \exp\left(-\frac{|\mathbf{r}|^2}{2\sigma^2}\right)\,, \tag{19.6}$$

Table 19.2 Values of parameters that generate $\mathbf{k}_l$ of (19.8)

n	2	3	4	5	6
c_n	1	4/3	$\sqrt{2}$	2	$4/\sqrt{3}$
α_n	0	0	$\pi/4$	0	$-\pi/6$

is an unnormalized Gaussian of (large) width σ, accounting for the overall intensity envelope of the laser beam, and

$$W_n(\mathbf{r}) = -\frac{1}{n^2}\left|\sum_{l=0}^{n-1}\exp(i\,\mathbf{k}_l\cdot\mathbf{r})\right|^2, \tag{19.7}$$

is a periodic (for $n = 2, 3, 4$, or 6) or quasi-periodic (e.g., $n = 5$) potential of n-fold symmetry [5, 18] produced by the interference of n laser beams, representing the substrate corrugation. The appropriate 2D interference pattern is realized by taking

$$\mathbf{k}_l = \frac{c_n\pi}{a_{\text{las}}}\left[\cos\left(\frac{2\pi l}{n}+\alpha_n\right),\ \sin\left(\frac{2\pi l}{n}+\alpha_n\right)\right]. \tag{19.8}$$

The numerical constants c_n are chosen in order to match the potential lattice spacing to the laser interference periodicity a_{las}, and α_n are chosen so that one of the primitive vectors of the periodic potential $W_n(\mathbf{r})$ is directed along the x axis. These numerical coefficients are reported in Table 19.2.

The two positive amplitudes A_c and U_0 set respectively the intensity of the overall potential confining the colloids near the simulation-cell center, and the intensity of the corrugated, spatially oscillating term. Note that the same overall Gaussian intensity modulation, (19.5) accounting for the lateral intensity profile of the laser beam, shapes both the confining well and the amplitude modulation, consistently with experiment [3]. The simulation is carried out in dimensionless units, defined in terms of the physical quantities of Table 19.3. The simulated colloid island has a small extension compared with the Gaussian width σ, reported in Table 19.1.

Table 19.3 Basic units for various quantities in the model, with typical values appropriate for the setup of [3]

Physical quantity	Model expression	Typical value
Length	a_{coll}	5.7 μm
Force	$F_0 = 9F_{s1}/(8\pi)$	18 fN
Viscosity coefficient	η	6.3×10^{-8} kg/s
Energy	$F_0\,a_{\text{coll}}$	1.0×10^{-19} J
Time	$\eta a_{\text{coll}}^2/U_0$	20 s
Mass	$\eta^2 a_{\text{coll}}/F_0$	1.3×10^{-6} kg
Velocity	F_0/η	0.284 μm/s
Power	F_0^2/η	5.1×10^{-21} W

Vanossi et al. simulate the colloidal system in a periodically repeated simulation box where periodic boundary conditions are implemented for particle-particle interactions, but not for the external potential $V_{\text{ext}}(\mathbf{r})$ which originates in the central cell only. This choice of boundary conditions is appropriate for both these situations:

1. With $A_c > 0$, an island of particles is located near the center of a much wider supercell, with few or no particle ever crossing the cell boundary. This describes a finite trapped "droplet" of colloidal particles, surrounded by a rarefied cloud of isolated particles. The 2D density of the triangular 2D lattice is fixed by N ($\sim$30,000) and by the balance of the confining energy G and the 2-body repulsion energy. This balance is tuned to produce unit average colloid lattice spacing a_{coll} (before submittal to the corrugation potential W).[1]
2. With $A_c = 0$ and very large σ, the supercell is filled with a compact monolayer, and a periodic potential is felt by the colloids. In this scheme one simulates an essentially infinite system, characterized by an average colloid density equaling N divided by the supercell area. Here the supercell size must be chosen carefully to produce an overall periodic geometry, both relative to the colloid lattice and to the periodic modulation potential W_n.

In either version, by choosing appropriate ratios of the average particle spacing to the periodic laser potential spacing a_{las} one can realize a variety of static superstructures, each with its soliton array and pattern. Both approaches are adopted, with numerical parameters listed in Table 19.1. Unlike experiments, where the colloid density is finite everywhere, with a soft edge characterized by a minor decrease outside the Gaussian attractive radius σ, in the simulations of [13] the density vanishes rapidly at the island edge. While the edge details may differ in experiment and in simulation, the central part of the simulated island (which is the only region being monitored in experiment) is directly comparable.

As in experiment, Vanossi et al. consider a triangular lattice potential, i.e. $n = 3$. The shape of the periodic potential $W(\mathbf{r}) = W_3(\mathbf{r})$

$$W(\mathbf{r}) = -\frac{2}{9}\,U_0\left[\frac{3}{2} + 2\cos\frac{2\pi x}{a_{\text{las}}}\,\cos\frac{2\pi y}{\sqrt{3}a_{\text{las}}} + \cos\frac{4\pi y}{\sqrt{3}a_{\text{las}}}\right] \tag{19.9}$$

is depicted in Fig. 19.2a. The corrugation profile along $\hat{\mathbf{x}}$, where the energy barrier is lowest is

$$W(x,0) = -U_0\left(\frac{5}{9} + \frac{4}{9}\cos\frac{2\pi x}{a_{\text{las}}}\right), \tag{19.10}$$

[1] The spacing of the fully relaxed colloid configuration varies smoothly from $a \simeq 0.984$ at the sample center to $a \simeq 1.05$ at the periphery, with an average density equal to that of a triangular crystal of spacing $a_{\text{coll}} = 1$.

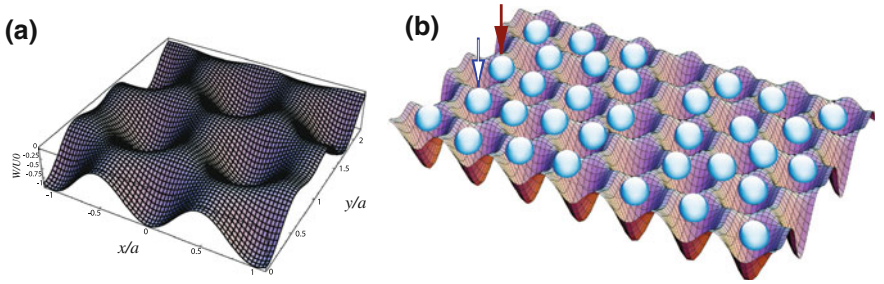


Fig. 19.2 **a** The triangular-symmetry corrugated potential-energy profile $W(x, y) = W_3(x, y)$ as a function of position. Observe the minima at energy $-U_0$, the saddle points at energy $-U_0/9$, and the maxima at energy 0. The resulting lowest energy barrier in the x direction thus equals $\frac{8}{9}U_0$. **b** A sketch of the model for the colloid particles interacting with the periodic potential W. The *red filled arrow* and the *blue empty arrow* point at examples of colloids at an unstable and stable position respectively

with a barrier amplitude $\frac{8}{9}U_0$. Accordingly, the static friction force for an isolated colloid, i.e. the minimum force that a single colloid requires in order to slide in this one dimensional potential, is

$$F_{s1} = \frac{8\pi U_0}{9a_{\text{las}}} = \frac{8\pi}{9} F_0. \tag{19.11}$$

The depinning force per colloid is evaluated in realistic many-colloid calculations and compared with this elementary barrier.

With the adopted parameters, the balance of the Gaussian confinement and the repulsive colloid-colloid repulsion leads in the absence of corrugation ($U_0 = 0$) to an equilibrium spacing $a_{\text{eq}} = 0.984$ at the center of the sample, and an overall average density compatible with nearest-neighbor separation $a_{\text{coll}} = 1$, taken as reference unit distance. It is convenient to define an appropriate length ratio $\rho = a_{\text{las}}/a_{\text{coll}}$, such that when $\rho = 1$ a fully matched configuration is realized.

19.4 The Simulation Protocol

The initial configuration is obtained by cutting a circular island out of a perfect 2D triangular lattice with spacing $a_{\text{coll}} = 1$. A first simulation is run with zero external force $F = 0$, in the presence of the confining Gaussian and of the periodic potential of lattice spacing a_{las}, simultaneously turned on. Three representative cases are considered, namely: underdense, $\rho = 0.95$ (antisoliton-incommensurate—AI; the starting state at rest is shown in Fig. 19.3a); ideally dense, $\rho = 1.0$ (nearly

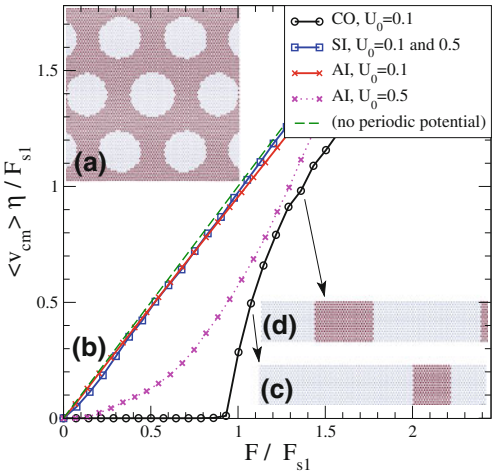


Fig. 19.3 **a** The starting static configuration at $\rho = 0.95$. Three families of antisoliton lines (*darker areas*) cross at 120°. **b** Velocity-force characteristics for various colloid densities, with a lattice-potential corrugation commensurability ratio $\rho = a_{\rm las}/a_{\rm coll} = 1.0$ (CO), 1.05 (SI), and 0.95 (AI). The CO case always displays static friction. For weak corrugation ($U_0 = 0.1$), $F_{\rm s} = 0$ in both AI and SI incommensurate cases. At larger corrugation ($U_0 = 0.5$) a major asymmetry appears between the AI and SI configurations: only the AI case exhibits a finite depinning threshold with static friction. **c**, **d** Snapshots of the central region of the initially commensurate colloid during motion, illustrating sliding-generated solitons, whose density increases as F is increased. In all snapshots, colloids located at repulsive spots of the corrugation potential (defined by $W(\mathbf{r}) > -U_0/2$, e.g. the colloid pointed at by the *red filled arrow* in Fig. 19.2b) are drawn as *dark red spots*, while colloids nearer to potential minima ($W(\mathbf{r}) \leq -U_0/2$, e.g. the colloid pointed at by the *blue empty arrow* in Fig. 19.2b) are *light blue*

commensurate—CO, which becomes exactly commensurate after turning on W); overdense, $\rho = 1.05$ (soliton incommensurate—SI).[2]

In the mismatched cases, still at $F = 0$, the colloid structure relaxes forming lattice-matched regions separated by a hexagonal network of soliton/antisoliton regions where the misfit accumulates, in the form of a longer/shorter average colloid-colloid separation. Figure 19.2b sketches a typical mutual pattern of the colloids and corrugation potential. For soliton/antisoliton global patterns see Appendix, Figs. 19.11 and 19.9 respectively. The high mobility of these soliton defects allows the colloidal system to heal most of initial tensile strain of the circular island, thus compressing to an average nearest-neighbor spacing $a_{\rm coll} \simeq a_{\rm eq}$. As the figures in Appendix show, this leads to a visibly smaller spacing of solitons than antisolitons. In between these overdense and underdense cases, by taking $a_{\rm las} = 1$, Vanossi et al. [13] simulate a nominally commensurate case, as in experiment [3]. As illustrated by the $a_{\rm las} = 1$ point of Fig. 19.7, the matched configuration is energetically favorable, with all colloids sitting near to a W potential minimum, and undergoing little or

[2] Similar models were studied in the past with a view to understand 2D Frenkel-Kontorova models and adsorbate monolayers physics [19–22].

no initial rearrangement during the initial $F = 0$ simulation. Actually, because the island center equilibrium spacing $a_{\rm eq} = 0.984$ is smaller than $a_{\rm las} = 1$ there is still, in this nominally commensurate case, an overall tensile strain. On account of a strong "epitaxial" effect of the periodic potential, this strain is however too weak to give rise to the formation of antisolitons, and is left unhealed in the equilibrated island, that remains fully commensurate. Only under sliding, when soliton structures move in from the edge boundary and sweep across the central region (which is the only region being monitored in experiment) under the action of the dragging force, does the tensile strain have a chance to heal out, giving rise in the very same system that was commensurate at rest to a sparse but regular and nonzero soliton density in the running state. This sliding induced soliton proliferation in statically commensurate colloids agrees very well with what has been reported experimentally [3].

Mimicking experiment, an $\hat{\mathbf{x}}$-directed force F acting on each particle is turned on as in (19.1). This force is kept fixed for a finite time t_F, after which its sign is reversed for the same time duration. Afterward, this reciprocation process is repeated with an increased force magnitude $F + \Delta F$. To mimic experiment, each forward and backward run is carried out for an amount of time t_F inversely proportional to the force value F itself, and thus the cell speed. The product of $F\,t_F$ is selected in such a way that, under the action of F for a time t_F, an isolated unconfined particle would move by a few lattice spacings $F\,t_F/\eta \approx (2 \div 3)a_{\rm las}$ typically, corresponding to at most 3 solitons/antisolitons crossing a given $\hat{\mathbf{y}}$-directed line during a simulation.

Focusing on that region and thus excluding undesired edge effects, a square central region of size 80×80, containing ~7400 colloids, is selected. An initial transient of approximately 30 % of the simulation time is dropped, and the $\hat{\mathbf{x}}$-component of the velocity of this central block is time-averaged over the rest of the simulation. This procedure defines the quantity $\langle v_{\rm cm} \rangle$ of Fig. 19.3b.

19.5 Simulation Results

19.5.1 Force-Velocity Characteristics

Figure 19.3b displays the mean speed $\langle v_{\rm cm} \rangle$ of the central portion of the colloid system as a function of the driving force F. Fully reproducing experiment [3], the simulated force-velocity characteristics of Fig. 19.3b show a large static friction force threshold in the $\rho \simeq 1$ CO case, where the colloid and corrugation lattices are pinned together. Static friction is lost in case of incommensurability and moderate corrugation, where preformed mobile solitons or antisolitons are present. The snapshots of Fig. 19.4 illustrate the patterns of solitons/antisolitons sliding in opposite directions under the same driving force $F > 0$. For a weak external force and a ~5 % lattice mismatch, the static friction drops essentially to zero, and a nearly free viscous sliding is realized, reflecting a situation of "superlubricity" [23–25]. However, under the same mismatched conditions, not all incommensurate geometries are superlubric. Whereas for weak corrugation the overall colloid mobility $\langle v_{\rm cm} \rangle / F$ is remarkably constant for both incommensurate densities, Vanossi et al. find in fact that by

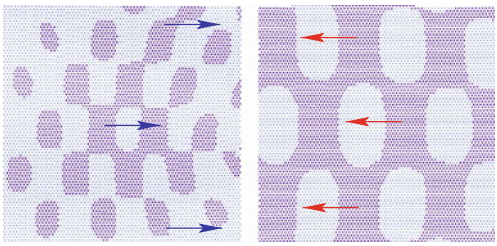


Fig. 19.4 Depinned, moving particles (*darker*) of a 2D colloid in a periodic potential under the action of a *rightward* force F. **a** *Rightward* propagating solitons of overdense colloids ($\rho = 1.05$, SI); **b** *Leftward* propagating antisolitons of underdense colloid ($\rho = 0.95$, AI)

increasing the corrugation amplitude U_0 the mobility of the AI configuration drops to zero at small force, and pinning with static friction reemerges despite incommensurability. By contrast, SI configurations remain superlubric up to much larger U_0.

19.5.2 Aubry-Like Pinning-Unpinning Transition

Borrowing results of the 1D Frenkel-Kontorova (FK) model [10], the single-soliton width $d \simeq g^{1/2} a_{\rm las}$, where $g = \xi a_{\rm las}^2 k/U_0$ [here $k = V''(a_{\rm coll})$ and ξ is a constant of order unity], is large for a hard layer on a weak corrugation, and small for a soft layer on a strong corrugation. Between these two extremes, the 1D incommensurate FK model crosses the so-called Aubry transition [23] where superlubricity is lost, and pinning sets in with finite static friction despite incommensurability. Even in the present 2D case it is qualitatively expected that all incommensurate colloids, both underdense ($\rho \lesssim 1$) and overdense ($\rho \gtrsim 1$) will undergo an Aubry-like superlubric-to-pinned transition for increasing corrugation.

This expectation is indeed confirmed in this 2D model colloid system. Figure 19.5 (obtained by independent simulations of the infinite-size system with periodic boundary conditions) shows the Aubry-like pinning transition crossed by an AI ($\rho = 0.95$) underdense colloid at a critical corrugation, here $U_0^{\rm crit} \simeq 0.2 - 0.3$. The threshold Aubry corrugation depends upon ρ, and is much larger for overdense SI than for underdense AI colloids.

19.5.3 Soliton-Antisoliton Asymmetry

This strong asymmetry of static friction—and of all other properties—between overdense ($\rho \gtrsim 1$) and underdense ($\rho \lesssim 1$) colloids can be rationalized, in the limit of strong corrugation $g \ll 1$, in terms of the large physical difference between solitons,

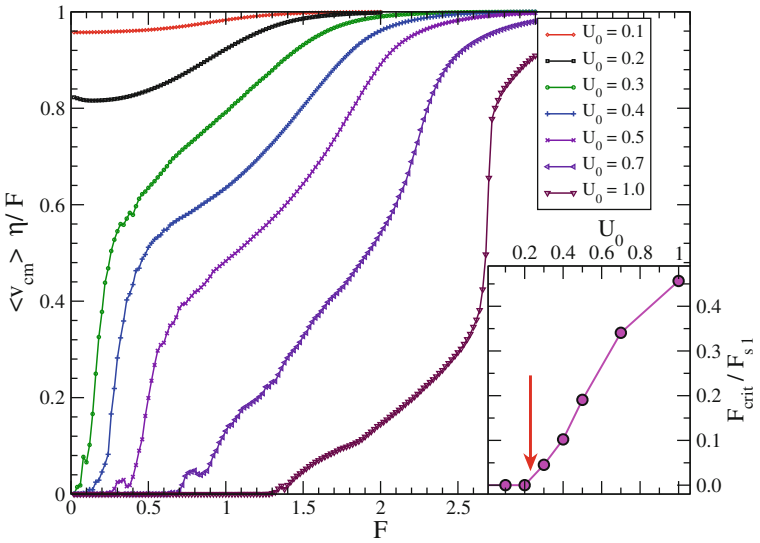


Fig. 19.5 2D Aubry transition for antisolitons at $\rho = 0.95$, in an infinite-size colloid system. *Main panel* colloid mobility as a function of the applied force, for increasing corrugation amplitude U_0. Note the appearance of pinning with static friction just above U_0= 0.2. *Inset* static friction (depinning) force F_s, normalized to the single-colloid force barrier $F_{s\,1}$, as a function of U_0, with an *arrow* indicating the critical Aubry corrugation

defects formed by lines of lattice interstitials, and antisolitons, lines of vacancies. This asymmetry is retained even for weak corrugation ($g \gg 1$), when solitons/antisolitons involve relative displacements far smaller than those of proper interstitials or vacancies. A small variation δ in the inter-colloid separation a is sufficient to produce a large relative variation of the effective spring constant, i.e. the interaction curvature

$$\frac{V''(a \pm \delta)}{V''(a)} \simeq \frac{V'(a \pm \delta)}{V'(a)} \simeq \frac{V(a \pm \delta)}{V(a)} \simeq \exp(-\delta/\lambda_D). \tag{19.12}$$

For a realistic $\lambda_D \simeq 0.03\, a_{\rm coll}$, this highly nonlinear and asymmetric relation, implies a huge 460 % increase whenever two colloids are approached by 5 % of their average separation, but only a 82 % reduction for a 5 % increased separation. This asymmetry is held responsible for the much weaker propensity of solitons to become pinned and to localize compared to antisolitons.

19.5.4 The Sliding State

Under sliding, the shapes and geometries of solitons/antisolitons and their motion are of most immediate interest, as they are directly comparable with experiment.

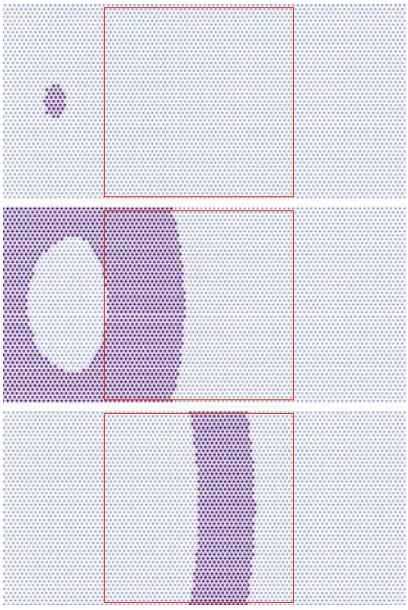


Fig. 19.6 Three successive snapshots of the initial depinning instants of the commensurate ($\rho = 1$) configuration for the $F \simeq F_{s\,1}$ simulation, see Fig. 19.3. The horizontally extended window visualizes the the nucleation and separation of a soliton-antisoliton pair, *left* of the central observation region (*square*). Pair nucleation constitutes the depinning mechanism of all commensurate sliders

Figure 19.4 shows the large-scale checkerboard structure of solitons/antisolitons of the sliding colloid lattice. They move with a speed v much larger than the average lattice speed $\langle v_{\rm cm}\rangle$, because $v/\langle v_{\rm cm}\rangle \sim \rho/(\rho - 1)$ by particle conservation. The moving structure is a distortion of the original triangular soliton/antisoliton pattern (Fig. 19.3a) induced by the circular shape of the confining potential, and by the directional sliding. With increasing F, the soliton arrangements elongate into a stripe-like pattern perpendicular to the driving direction. Comparison with experimental pictures is quite realistic, especially when focusing (as done in experiment) on the central sample region, far from boundaries.

In the AI superlubric colloid $\rho \lesssim 1$, preformed antisolitons fly (leftward) across the colloid lattice antiparallel to the (rightward) force. They are eventually absorbed at the left edge boundary, while new ones spawn at the right edge boundary to replace them, sustaining a steady-state mobility. In the SI superlubric colloid $\rho \gtrsim 1$ conversely, preformed solitons fly rightward, parallel to the force. Solitons, unlike antisolitons, are not automatically spawned at the boundary, owing to the decreasing density. Instead, an antisoliton/soliton pairs must nucleate first, near the boundary, and this is possible only if the force overcomes the nucleation barrier. Below this threshold, a steady DC external force eventually sweeps out all the preformed solitons transforming the colloid to an artificially pinned, immobile CO state.

Finally, the pinned CO colloid $\rho \gtrsim 1$ only moves after static friction is overcome. As illustrated in Fig. 19.6, motion starts off here by nucleation of soliton-antisoliton pairs inside the bulk—here, driving the monolayer rightward, close to the left edge because the central region tends to be slightly overdense. The antisolitons flow

leftwards and are absorbed by the left edge, becoming undetectable to the optically monitored central part of the colloid, where only solitons transit, as seen in experiment. This type of commensurate nucleation has been described in considerable detail in the literature, including finite-temperature effects [11, 14, 26]. Note that in the pinned CO colloid the soliton or antisoliton density, initially zero, actually increases with increasing sliding velocity (see, e.g., Fig. 19.3c, d), as opposed to frankly incommensurate cases, where it is nearly constant.

19.5.5 Phase-Diagram Evolution with Sliding

Much can be learned about the habit of sliding colloids from their behavior and their structural phase diagram, first at rest and then under sliding. With $\rho \simeq 1$, close to commensurate but not exactly commensurate, the colloid monolayer can realize in the periodic potential two alternative static arrangements which are local minima of the overall free energy: a fully lattice-matched CO state, or a weakly incommensurate state characterized by a sparse soliton (AI or SI) superstructure, with a density fluctuating around the local value prescribed by the $G - V$ balance. Comparing the potential energy of these two states as a function of ρ, the static phase diagram contains, as sketched in Fig. 19.7, a fully commensurate extended CO region separated from the AI and SI regions by commensurate-incommensurate transitions,

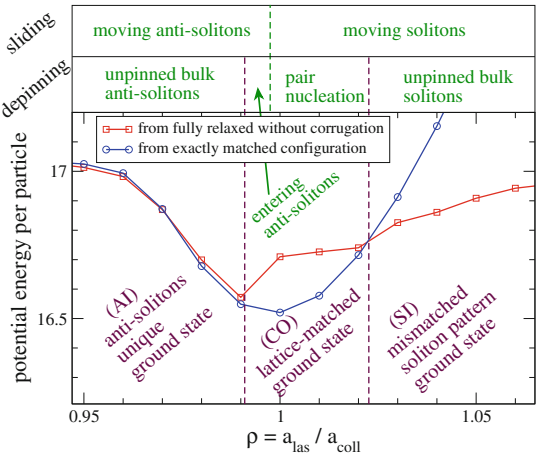


Fig. 19.7 Effective "phase diagram" of the finite colloid crystal as a function of the lattice spacing mismatch ρ, for $U_0 = 0.1$. The two potential energy curves (colloid-colloid repulsion plus W interaction) characterize the static phases: (*red squares*) relaxation started from the slightly inhomogeneous configuration produced by a previous relaxation for $U_0 = 0$; (*blue circles*) relaxation started with a fully matched lattice of spacing a_{las}. Two AI (underdense) and SI (overdense) phases surround a commensurate phase (CO). The depinning mechanisms and the soliton structures sustaining sliding are illustrated at *top* of figure for the different regions

well known in adsorbed surface layers [27–30]. The CO region is wider on the SI side ($\rho > 1$) than the AI side ($\rho < 1$), another manifestation of the SI-AI asymmetry discussed above. The CO range naturally widens or shrinks when the corrugation amplitude U_0 is increased or decreased, respectively.

Under an external force F, sliding effectively tilts the balance between the two static phases (loosely speaking, for of course under sliding the physical significance of a "phase" is not the same as at rest) sliding populates the former CO phase with solitons/antisolitons, turning it effectively into SI or AI. In the running state, the colloid average density increases or decreases from 1 to a value closer to the nominal ρ of the colloid at $U_0 = 0$. This explains why in a quasi-commensurate configuration with $\rho \gtrsim 1$ such as that shown by Bohlein et al. [3], solitons (and not, e.g., soliton-antisoliton pairs) sweep the colloid upon depinning, as also seen in Fig. 19.3c, d.

It is curious to note here the different fate of solitons in the slightly overdense CO and in the SI phases. In the CO phase they do not exist at rest, but they appear after depinning and under sliding. In the SI phase they exist at rest, but they could be swept out under DC sliding, when a weak external force can turn the SI colloid into effectively CO. This sweepout phenomenon was never observed on the AI side.

19.6 Friction of Colloid Sliding on the Optical Lattice

Coming to frictional work, this is a quantity of crucial importance for the tribological significance of colloid sliding. The overall power balance is set by the scalar product of the instantaneous velocity $\mathbf{v}_i$ of each colloid i times the net force acting on it, $\eta\,(\mathbf{v}_d - \mathbf{v}_i)$, including both the bare external force $\mathbf{F} = \eta\mathbf{v}_d$ and the viscous drag $-\eta\mathbf{v}_i$. This product vanishes instantaneously at any time when either colloids are stuck ($\mathbf{v}_i = \mathbf{0}$) or else when the corrugation potential is absent, so that $\mathbf{v}_i \equiv \mathbf{v}_d$. After averaging over a very long trajectory, the balance reads

$$\begin{aligned} P_{\rm tot} &= \sum_i \eta \langle (\mathbf{v}_d - \mathbf{v}_i)\cdot \mathbf{v}_i \rangle \\ &= N\left(\mathbf{F}\cdot\langle \mathbf{v}_{\rm cm}\rangle - \eta\langle |\mathbf{v}_{\rm cm}|^2\rangle\right) - \eta\sum_i \langle |\mathbf{u}_i|^2\rangle \\ &= P_{\rm frict} - P_{\rm kin}\,, \end{aligned} \tag{19.13}$$

where $\mathbf{u}_i = \mathbf{v}_i - \mathbf{v}_{\rm cm}$. Under steady-state sliding conditions where $P_{\rm tot} = 0$, the effective friction power $P_{\rm frict}$ is exactly balanced by an internal kinetic energy excess rate. Per colloid particle, $P_{\rm frict}$ is

$$p_{\rm frict} = \frac{P_{\rm frict}}{N} \simeq \mathbf{F}\cdot\langle \mathbf{v}_{\rm cm}\rangle - \eta|\langle \mathbf{v}_{\rm cm}\rangle|^2, \tag{19.14}$$

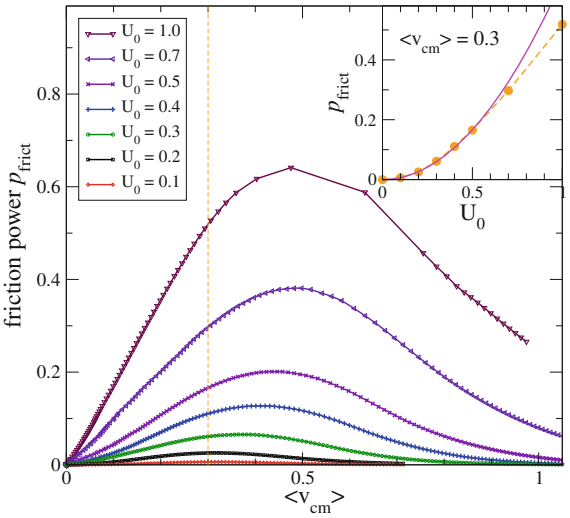


Fig. 19.8 Sliding friction p_{frict} per particle as a function of speed $\langle v_{\text{cm}} \rangle$ for an underdense AI colloid $\rho = 0.95$ for increasing corrugation amplitude U_0. *Inset* U_0 dependence of friction for speed $\langle v_{\text{cm}} \rangle = 0.3$, showing the quadratic rise for weak corrugation behavior, followed by a roughly linear growth

where small center of mass fluctuations are neglected, by assuming $\langle \mathbf{v}_{\text{cm}}{}^2 \rangle \simeq |\langle \mathbf{v}_{\text{cm}} \rangle|^2$.

Figure 19.8 shows p_{frict} (briefly referred to as "friction" in the following) for the special case of the AI underdense phase as extracted as a function of $\langle \mathbf{v}_{\text{cm}} \rangle$ through a "bulk" simulation (with periodic boundary conditions as in Fig. 19.5). The main features found are (i) a linear rise at low CM speed; (ii) a decline at large speed; (iii) a maximum at some intermediate corrugation-dependent speed. Observe also that (iv) the dissipated power increases (not unexpectedly) with corrugation; and (v) the corresponding frictional maximum simultaneously shifts to larger speed.

The qualitative interpretation of these results is relatively straightforward, and yet revealing. (i, iv) At low sliding velocities the motion of solitons/antisolitons involves the viscous motion of individual particles with a velocity distribution whose spread toward higher values rises proportionally to the sliding speed and inversely proportional to their spatial width. As shown, e.g. within the 1D FK model [10], but also in the present simulations, the width d of solitons/antisolitons increases roughly as $d \sim a_{\text{las}}\sqrt{g}$ with the dimensionless interparticle interaction strength $g \propto a_{\text{las}}^2 V''(a_{\text{coll}})/U_0$ measured relative to the periodic corrugation amplitude. The decrease in width with increasing corrugation U_0 requires an increasing instantaneous speed of individual particles in the soliton/antisoliton, yielding an increasing viscous friction, and a decreasing overall mobility, as observed. (ii, iii, v) At high sliding velocities, the colloid relaxation time exceeds the soliton/antisoliton transit time across the Peierls-Nabarro barrier [10] so that their spatial structure is gradually washed out by the sliding motion. The critical speed where the smoothening behavior takes over, roughly corresponding to maximal friction, increases as corrugation increases, corresponding to narrower solitons/antisolitons that are harder to wash out. The increase of friction with corrugation strength U_0, plotted in the inset for

a chosen speed, is found to be quadratic at weak corrugation, gradually turning to linear for larger values. Linear response theory naturally accounts for the quadratic increase, a behavior first discussed by Cieplak et al. [31] and observed in quartz crystal microbalance experiments [32].

Demonstrated for a specific AI case with antisolitons, the above results appear of general validity for infinitely extended sliders of controlled colloid density, and apply equally well although with great quantitative asymmetry to SI with solitons once their larger widths, greater mobilities, and weaker Peierls-Nabarro barriers are taken into account.

19.7 Summary and Discussion

We have reviewed the sliding of 2D colloids on optical lattices, and in particular initial simulation results and theory that strongly vouch in favor of this technique as a promising tool for future tribological advances. The motion of solitons and antisolitons known from experiment is reproduced and understood, unraveling the subtle depinning mechanisms at play. The presence of Aubry transitions is pointed out for future verification, along with a strong asymmetry between underdense and overdense incommensurate layers. Of direct tribological interest, the simulations anticipate a behavior of friction with corrugation (mimicking "load") and with sliding velocity, with results which, while of course generally very different from the classic laws of macroscopic friction, are highly relevant to friction at nano and mesoscopic scales. A strong complementarity between theory-simulation and experiment is underlined, and should be pursued further.

There are many lines of future research that these studies implicitly suggests. One line will be to pursue the analogy of the sliding over a periodic potential with other systems such as driven Josephson junctions [33], and sliding charge-density waves [34]. Time-dependent nonlinear phenomena such as the Shapiro steps [33, 34] should become accessible to colloid sliding too. A second line is to include non-periodic complications to the corrugation potential, including the quasicrystal geometry such as that recently realized [35] and beyond that, random, or pseudo-random corrugations to be realized in the future. A third line involves the investigation of the lubricant speed quantization phenomena, characterized so far only theoretically [36–39]. A fourth line involves the investigation of rotated orientational epitaxy of the colloid lattice relative to the laser potential [40, 41] and its effects on friction.

A further very important development will be to address colloidal friction in larger, mesoscopic or macroscopic size systems, whose phenomenology is accessible so far only by a few, very ingenious, but very limited, methods [6, 42–46]. A major scope in that case will be to realize and study stick-slip friction and aging phenomena, at the heart of realistic physical and technological tribology.

Acknowledgments We gratefully acknowledge helpful discussions with C. Bechinger, T. Bohlein, O.M. Braun, C. Dellago, M. Invernizzi, D. Mandelli, C. Reichhardt, G.E. Roat, and G.E. Santoro. This work is supported in part by COST Action MP1303, the Italian Ministry of University and

Research, the Swiss National Science Foundation Sinergia CRSII2_136287, and the ERC Advanced Grant No. 320796-MODPHYSFRICT.

Appendix: Static Configurations

Figures 19.9, 19.10, 19.11 display three overall views of the static fully relaxed $F = 0$ configurations for different values of a_{lub} (or, equivalently, of ρ). These pictures represent the lowest-energy configurations of the three regions in the phase diagram—Fig. 19.7. Soliton/antisoliton patterns are highlighted by coloring colloids differently for different positions relative to the potential profile of Fig. 19.2: dark,

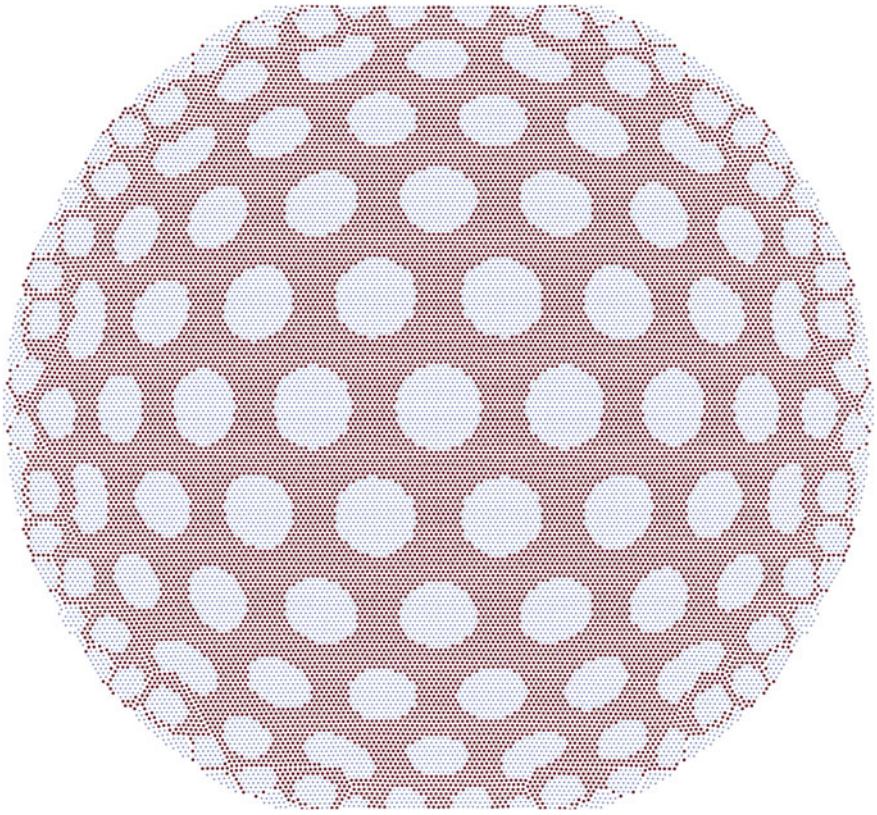


Fig. 19.9 The static initial configuration for $U_0 = 0.1$, $F = 0$, $a_{\mathrm{las}} = 0.95$, i.e. $\rho = 0.95$ (antisoliton incommensurate pattern: AI). *Darker dots* indicate colloidal particles sitting at a repulsive point of the corrugation landscape, namely with $W(\mathbf{r}) > -U_0/2$. This configuration is essentially unique, since an extremely similar configuration is retrieved at the end of a long relaxation, regardless of the initial condition. The central portion of this figure is represented in Fig. 19.3a

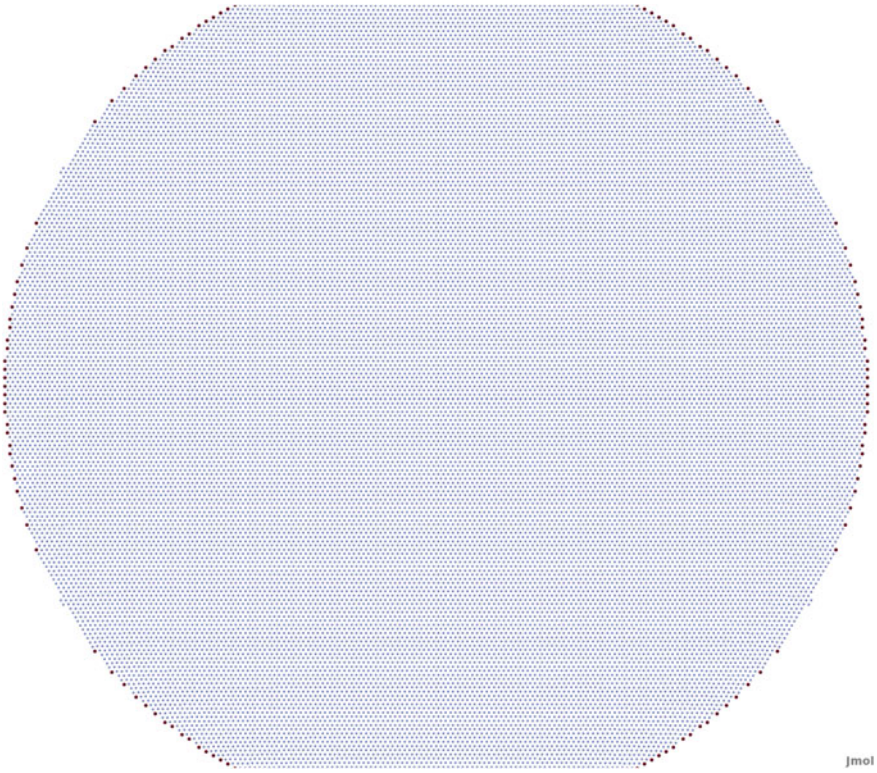


Fig. 19.10 The static initial configuration for $U_0 = 0.1$, $F = 0$, $a_{\text{las}} = 1.00$, i.e. $\rho = 1.02$ (commensurate: CO). *Darker dots* have the same significance as in Fig. 19.9. At this nearly matched value $\rho = 1$, this configuration is the lowest-energy state, on the *blue-circle curve* in the phase diagram of Fig. 19.7. The metastable high-energy state of the *red-square curve* in Fig. 19.7 is qualitatively similar to the one depicted in the subsequent Fig. 19.11

bolder colloids occupy locally unfavorable repulsive regions for the corrugation profile W. Figure 19.12 illustrates the antisoliton pattern for a larger amplitude of the corrugation U_0, to be compared with Fig. 19.9 obtained with smaller corrugation. Note that the antisoliton lines are narrower in Fig. 19.12, but they form the same pattern as in Fig. 19.9.

Finally, Fig. 19.13 is to be compared with Fig. 19.9 to appreciate the effect of the random thermal motions characteristic of 300 K: (i) Brownian fluctuations smear the boundaries between in-registry and antisolitonic regions and (ii) a small thermal expansion is marked by a reduction in ρ, and therefore in the separation between antisolitons.

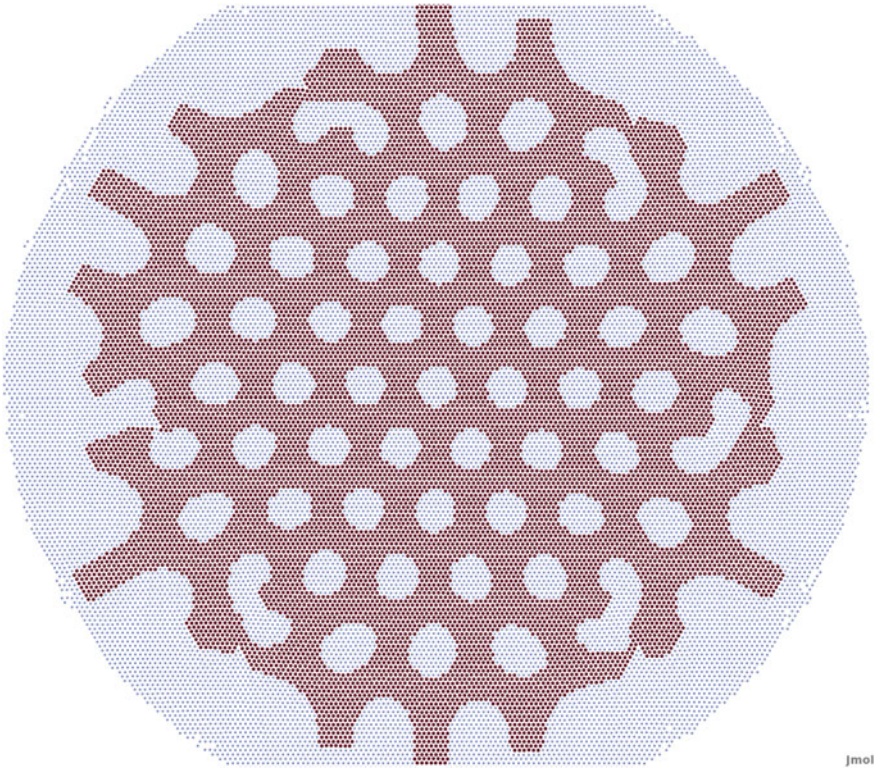


Fig. 19.11 The static initial configuration for $U_0 = 0.1$, $F = 0$, $a_{\mathrm{las}} = 1.05$, i.e. $\rho = 1.05$ (soliton incommensurate pattern: SI). *Darker dots* have the same significance as in Fig. 19.9. At this comparably large mismatch value $\rho = 1.05$, this configuration is the lowest-energy state, on the *red-square curve* in the phase diagram reported in Fig. 19.7. The metastable high-energy state of the *blue-circle curve* looks very similar to the one represented in the previous Fig. 19.10

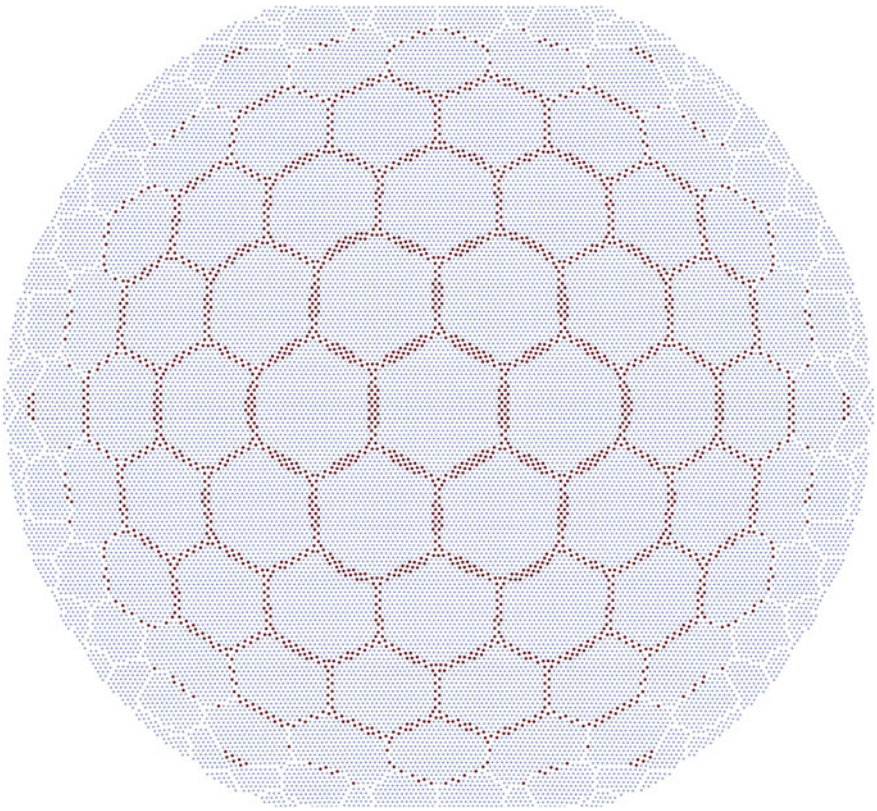


Fig. 19.12 A typical initial configuration for a stronger ($U_0 = 0.5$) corrugation potential. The other parameters ($F = 0$, $a_{\mathrm{las}} = 0.95$, AI) and the color notation are the same as in Fig. 19.9. By comparison with the weaker corrugation, here antisolitons are much narrower, intersecting and isolating well-faceted in-registry regions. Note that the pattern formed by the *center* of the *antisoliton lines* is the same in both figures

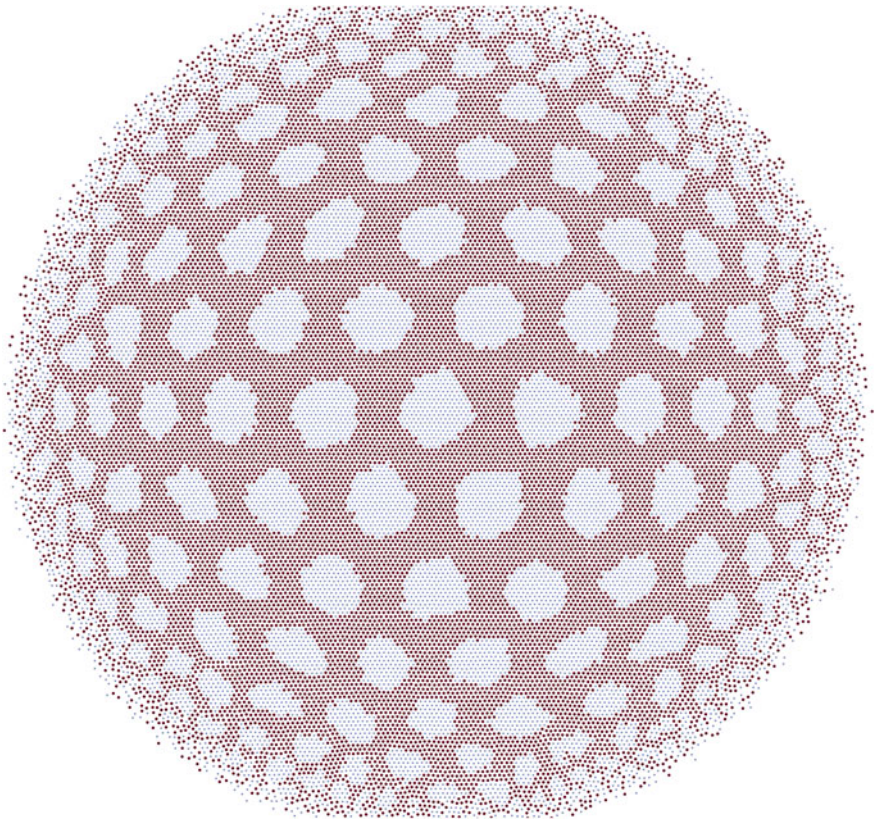


Fig. 19.13 The effects of thermal fluctuations. A typical $F = 0$ snapshot of a Langevin simulation at $k_B T = 0.04$, corresponding to room temperature in model units. The parameters ($U_0 = 0.1$, $F = 0$, $a_{las} = 0.95$, AI) and the color notation are the same as in Fig. 19.9. By comparison with the $T = 0$ configuration, the antisoliton pattern is only marginally affected by thermal noise. A small thermal expansion is responsible for a slight reduction of ρ, producing a visibly denser antisolitonic pattern than at $T = 0$

References

1. A. Vanossi, N. Manini, M. Urbakh, S. Zapperi, Tosatti E, Rev. Mod. Phys. **85**, 529 (2013)
2. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, Nature (London) **430**, 525 (2004)
3. T. Bohlein, J. Mikhael, C. Bechinger, Nat. Mater. **11**, 126 (2012)
4. J. Mikhael, J. Roth, L. Helden, C. Bechinger, Nature (London) **454**, 501 (2008)
5. J. Mikhael, M. Schmiedeberg, S. Rausch, J. Roth, H. Stark, C. Bechinger, Proc. Natl. Acad. Sci. USA **107**, 7214 (2010)
6. R.W. Carpick, M. Salmeron, Chem. Rev. **97**, 1163 (1997)
7. A. Vanossi, E. Tosatti, Nat. Mater. **11**, 97 (2012)
8. C. Reichhardt, C.J. Olson, Phys. Rev. Lett. **88**, 248301 (2002)
9. C. Reichhardt, C.J. Olson Reichhardt, Phys. Rev. Lett. **106**, 060603 (2011)
10. O.M. Braun, YuS Kivshar, *The Frenkel-Kontorova Model: Concepts, Methods, and Applications* (Springer, Berlin, 2004)
11. O.M. Braun, A.R. Bishop, J. Röder, Phys. Rev. Lett. **79**, 3692 (1997)
12. A. Erdemir, J.-M. Martin (eds.), *Superlubricity* (Elsevier, Amsterdam, 2007)
13. A. Vanossi, N. Manini, E. Tosatti, P. Natl, Acad. Sci. USA **109**, 16429 (2012)
14. J. Hasnain, S. Jungblut, C. Dellago, Soft Matter **9**, 5867 (2013)
15. I. Bloch, J. Dalibard, W. Zwerger, Rev. Mod. Phys. **80**, 885964 (2008)
16. M. Brunner, C. Bechinger, W. Strepp, V. Lobaskin, H.H. von Grunberg, Europhys. Lett. **58**, 926 (2002)
17. P.T. Korda, G.C. Spalding, D.G. Grier, Phys. Rev. B **66**, 024504 (2002)
18. M.J. Ablowitz, B. Ilan, E. Schonbrun, R. Piestun, Phys. Rev. E **74**, 035601 (2006)
19. P.S. Lomdahl, D.J. Srolovitz, Phys. Rev. Lett. **57**, 2702 (1986)
20. D.J. Srolovitz, P.S. Lomdahl, Physica D **23**, 402 (1986)
21. Y.N. Gornostyrev, M.I. Katsnelson, A.V. Kravtsov, A.V. Trefilov, Phys. Rev. B **60**, 1013 (1999)
22. O.M. Braun, M.V. Paliy, J. Röder, A.R. Bishop, Phys. Rev. E **63**, 036129 (2001)
23. M. Peyrard, S. Aubry, J. Phys. C: Solid State Phys. **16**, 1593 (1983)
24. M. Dienwiebel, G.S. Verhoeven, N. Pradeep, J.W.M. Frenken, J.A. Heimberg, H.W. Zandbergen, Phys. Rev. Lett. **92**, 126101 (2004)
25. A.E. Filippov, M. Dienwiebel, J.W.M. Frenken, J. Klafter, M. Urbakh, Phys. Rev. Lett. **100**, 046102 (2008)
26. M. Reguzzoni, M. Ferrario, S. Zapperi, M.C. Righi, Proc. Natl. Acad. Sci. USA **107**, 1311 (2010)
27. S.N. Coppersmith, D.S. Fisher, B.I. Halperin, P.A. Lee, W.F. Brinkman, Phys. Rev. Lett. **46**, 549 (1981)
28. P. Bak, Rep. Prog. Phys. **45**, 587 (1982)
29. A. Patrykiejew, S. Sokołowski, T. Zientarski, K. Binder, Surf. Sci. **421**, 308 (1999)
30. K. Mangold, P. Leiderer, C. Bechinger, Phys. Rev. Lett. **90**, 158302 (2003)
31. M. Cieplak, E.D. Smith, M.O. Robbins, Science **265**, 1209 (1994)
32. T. Coffey, J. Krim, Phys. Rev. Lett. **95**, 076101 (2005)
33. M. Tinkham, *Introduction to Superconductivity* (McGraw Hill, New York, 1996)
34. G. Grüner, Rev. Mod. Phys. **60**, 1129 (1988)
35. T. Bohlein, C. Bechinger, Phys. Rev. Lett. **109**, 058301 (2012)
36. A. Vanossi, N. Manini, G. Divitini, G.E. Santoro, E. Tosatti, Phys. Rev. Lett. **97**, 056101 (2006)
37. M. Cesaratto, N. Manini, A. Vanossi, E. Tosatti, G.E. Santoro, Surf. Sci. **601**, 3682 (2007)
38. A. Vanossi, N. Manini, F. Caruso, G.E. Santoro, E. Tosatti, Phys. Rev. Lett. **99**, 206101 (2007)
39. I.E. Castelli, R. Capozza, A. Vanossi, G.E. Santoro, N. Manini, E. Tosatti, J. Chem. Phys. **131**, 174711 (2009)
40. A.D. Novaco, J.P. McTague, Phys. Rev. Lett. **38**, 1286 (1977)
41. G.V. Uĩmin, L.N. Shur, JETP Lett. **28**, 18 (1979)
42. J. Krim, D.H. Solina, R. Chiarello, Phys. Rev. Lett. **66**, 181 (1991)
43. M.S. Tomassone, J.B. Sokoloff, A. Widom, J. Krim, Phys. Rev. Lett. **79**, 4798 (1997)
44. C. Drummond, J. Israelachvili, Phys. Rev. E **63**, 041506 (2001)
45. S.M. Rubinstein, G. Cohen, J. Fineberg, Nature (London) **430**, 1005 (2004)
46. S.M. Rubinstein, G. Cohen, J. Fineberg, Phys. Rev. Lett. **96**, 256103 (2006)

Part V
Layered Materials, Polymers

Chapter 20
Micro- and Nanotribology of Graphene

Martin Dienwiebel and Roland Bennewitz

20.1 Introduction

Graphene is a two-dimensional layer of carbon atoms bound by covalent sp^2-bonds. It is the building block of graphite, in which the graphene layers are stacked and held together by van der Waals forces. Graphene has become subject of intense research since the discovery of facile preparation procedures and the demonstration of outstanding electronic and mechanical properties [1]. The mechanical properties include remarkable values for the stiffness and the intrinsic strength, which were measured by nanoindentation into a free-standing graphene membrane [2]. Since graphene is the building block of the widely used solid lubricant graphite, its tribological properties are of fundamental scientific and possibly technological interest. Some limitations of graphite as solid lubricant are related to the chemical reactivity of the edges of graphitic flakes in the compound, for example the failure under vacuum conditions. In recent years, a number of procedures for the preparation of large-flake graphene have been developed [3–5]. The coverage of substrates with coherent, defect-free layers of graphene is a promising new direction in solid lubrication. In this chapter we review

M. Dienwiebel (✉)
Karlsruhe Institute of Technology, Institute for Applied Materials—Reliability of Systems and Components, Microtribology Center μTC, Kaiserstr. 12, 76131 Karlsruhe, Germany
e-mail: martin.dienwiebel@iwm.fraunhofer.de

R. Bennewitz
INM—Leibniz-Institute for New Materials, Campus D2 2, 66123 Saarbrücken, Germany
e-mail: roland.bennewitz@inm-gmbh.de

R. Bennewitz
Physics Department, Saarland University, Campus D2 2, Saarbrücken, Germany

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_20

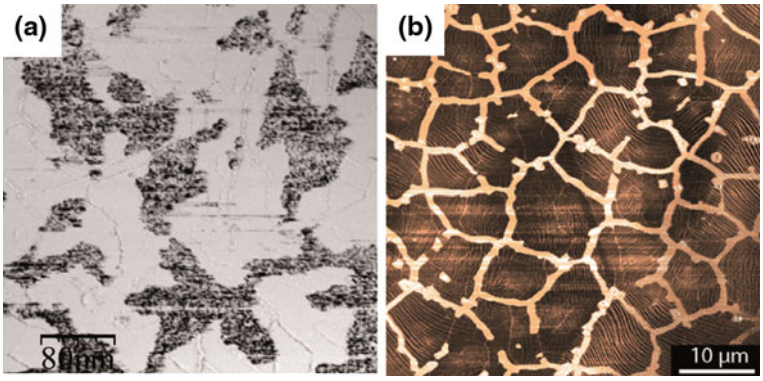


Fig. 20.1 Identification of graphene patches by reduced friction as measured by means of friction force microscopy (*darker color* indicates lower friction). **a** Epitaxial graphene on SiC(0001), **b** Graphene grown by chemical vapor deposition on copper, adapted from [10]

investigations into the mechanisms of graphene tribology by focusing on nanometer- and micrometer-scale studies. Note that graphene has also become of great interest as a constituent of composite materials with tribological applications [6].

20.2 Friction Force Microscopy of Graphene

Friction force microscopy provides one with spatially resolved maps of friction on inhomogeneous surfaces. This technique has been used to demonstrate the lubrication by graphene on a number of substrates, where the friction was found to be lower on graphene patches as compared to the surrounding substrate surface. Examples include graphene patches on SiC(0001) [7], silicon oxide [8, 9], and copper foils [10]. In the latter work of Marsden et al. on graphene growth on copper foils, the simultaneous recording of topography and lateral force allowed for a correlation of substrate surface structure and graphene growth.

The small contact area and high lateral resolution of friction force microscopy gives the opportunity to perform friction experiments on free-standing graphene membranes, which are suspended over micro-fabricated holes silicon wafers. First results have indicated that friction on free-standing graphene is similar to friction on graphene attached to rough substrates with little interaction between surface and graphene [11]. The comparison of friction on supported graphene with friction on free-standing graphene also allows to evaluate the role of van der Waals-interactions with subsurface material for friction and adhesion [12].

Friction force microscopy has also been employed to investigate changes in friction for chemically modified graphene layers. Somewhat unexpectedly, Kwon et al. found that fluorination of graphene causes a six fold increase of friction, despite a decrease in adhesion [13]. The authors attributed the additional dissipation to changes

in the flexural dynamics of the graphene upon fluorination. Experiments on hydrogenated and oxidized graphene have resulted in doubled and seven fold friction [14].

20.3 Graphene Versus Graphite

Although we restrict our description to friction studies on graphene, we would like to mention a few earlier results for graphite, which have important implications for the tribology of graphene. Graphite has been used as a model material for nanometer-scale friction studies since the first observation of atomic scale phenomena [15]. Highly oriented pyrolytic graphite (HOPG) was used to explore friction forces with atomic periodicity and to understand the stick-slip motion of the scanning tip of a friction force microscope [16]. The Prandtl-Tomlinson model was employed to explain stick-slip motion and the friction at atomic steps [17–19]. It also predicts the effects of commensurability between a flake of graphite and an HOPG surface [20] (for detailed information see also chapter *Superlubricity of Graphite*).Using a dedicated friction force microscope it was found that when rotating a graphite surface under the tungsten scanning tip, distinctive peaks of high friction appear at 0 and 60 ° of rotation, while for other orientations the friction dropped below the detection threshold of the instrument. The behavior was attributed to structural superlubricity between a graphite flake at the tip and the HOPG surface. Using in-situ transmission electron microscopy, Merkle et al. showed that transfer of graphitic layers to a sliding counterpart is a common phenomenon. The typical thickness of a transferred flake was found to be in the order of 10 basal planes [21]. It has been early on suggested by John Pethica that even scanning tunneling microscopy of graphite may result in the transfer of a graphitic flake to the probing tip, giving enhanced atomic contrast [22]. The tribology of graphene may thus have played an important role in the study of graphite before the systematic isolation of graphene sheets became available.

20.4 Atomic-Scale Friction of Graphene

High-resolution friction force microscopy on surfaces with long-range atomic order often results in regular pattern in the lateral force with atomic-scale periodicity. The pattern originates in a stick-slip mechanism, where the tip apex is stuck in a certain atomic position on the surface until the increasing lateral force is strong enough to initiate a slip to the corresponding atomic position of the next crystallographic unit cell. The earliest demonstration of an atomic friction experiment on graphene was reported by Enachescu et al. for an experiment on a single atomic layer of carbon on a Pt(111) surface [23]. Atomic friction experiments have attracted attention because they allow for the study of elementary processes in sliding friction. Until now, the analytical descriptions of the stick-slip dynamics (Prandtl-Tomlinson model) have not been complemented with a predictive model for the dissipation term which takes into account parameters of the respective surface materials. Atomic friction experiments

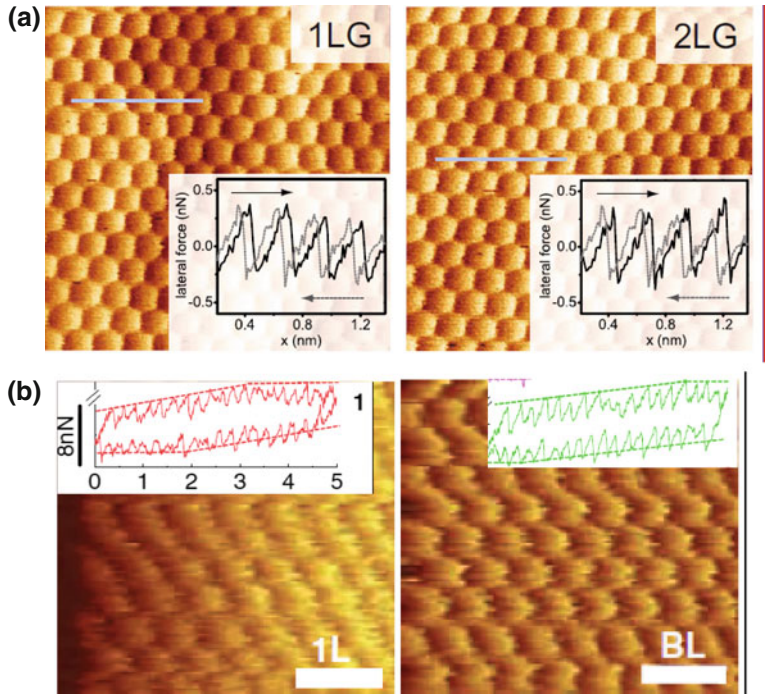


Fig. 20.2 Atomic friction experiments on graphene single and bilayers. **a** Graphene epitaxially grown on SiC(0001), adapted from [24], **b** Exfoliated graphene on silicon oxide, adapted from [11]

are nevertheless very useful in revealing the atomic structure of the relevant glide plane.

Filleter et al. have reported atomic friction experiments on graphene grown on SiC(0001) by thermal decomposition [24]. The results show that graphene grows smoothly across substrate steps, as they found a coherent hexagonal stick-slip patternin the lateral force signal. For the average friction signal, Filleter observed a significant contrast between single and bilayers of graphene on SiC(0001) [25]. Atomic friction results helped to exclude some possible sources of this contrast. Single and bilayer graphene exhibited a regular hexagonal stick-slip pattern with the same orientation and periodicity, see also Fig. 20.2a. The lateral contact stiffness, as determined from the slope of the lateral force vs. displacement curves, was the same for single and bilayer. Furthermore, all results could be reproduced with oxidized silicon and diamond-coated tips [24]. The authors suggested that the friction contrast is related to a difference in coupling between electrons and phonons for single and bilayer graphene on SiC(0001), both possible systems for dissipating energy of the sliding tip [25].

A very helpful contribution of atomic friction experiments to our understanding of friction on graphene has been described by Lee et al. [11]. They observed a distortion

of the regular stick-slip pattern which revealed an out-of-plane deformation of the graphene sheet around the sliding tip, see Fig. 20.2b. In particular for systems with a weak interaction between graphene and substrate, such as in graphene transferred to SiO_2 surfaces, the graphene tends to develop a fold in front of the sliding tip, which increases friction by increasing the contact area between tip and graphene. This effect becomes less pronounced for few-layer graphene, in which the stronger interaction between the graphene layers results in a stiffer layer with less tendency to wrap around the tip apex. Similarly, the effect is weaker for systems with a strong interaction between single-layer graphene and the substrate, for example on a mica substrate [26, 27]. Additional dissipation due to out-of-plane folding was found to be a mechanism effective on several weakly bound layered materials, namely MoS_2, $NbSe_2$, and h-BN [11].

The folding of loosely bound graphene around the sliding tip is also referred to a puckering effect. Choi et al. reported that the puckering effect leads to significant friction anisotropy when the graphene film is rippled as a result of the sample preparation [8]. The folding of graphene around the tip can also lead to quite exotic effects such as an increase of friction while the tip is retracted from the substrate [28, 29]. Stick-slip friction with a periodicity of a few nanometer, much larger than the unit cell of graphite, has also been attributed to graphene puckering [30]. However, the latter results are probably explained by lateral jumps of the tip between atomic rows when the tip slides in some small angle with respect to the atomic rows [31].

Atomic friction results allow to identify graphene patches on partially covered substrates with resolution that is better than scanning electron microscopy [10]. Marsden et al. have shown that symmetry and periodicity of the stick-slip signal reveal not only the material but also the azimuthal orientation of the graphene patches. Finally, atomic stick-slip pattern reveal superstructures resulting from the interaction of the graphene sheet with the underlying substrate structure [32].

20.5 Atomistic Simulations of Graphene Tribology

The dramatic effects of commensurability on the friction between a graphene flake and a graphite surface have stimulated several atomistic simulation studies of the situation. Bonelli et al. have implemented a Prandtl-Tomlinson model based on interactions described by a tight-binding model to explore friction of a deformable, rotating graphene flake on graphite [33]. They found good agreement with experiments by Dienwiebel et al. with respect to the stick-slip mechanism and the effects of rotational orientation, but some differences in force values. De Wijn et al. have studied the lubrication of two sliding surface by mobile rotating graphene flakes and found that a superlubric state can be established for incommensurate surfaces [34]. Guo et al. have performed molecular force field simulations and added one important aspect to the body of simulation studies, namely the increase in friction with decreasing distance between graphene flake and substrate, mimicking the effects of external normal pressure [35].

For the case of friction of a single-asperity contact sliding on graphene attached to a substrate surface, Ye et al. tested the hypothesis that the dependence of friction on the number of graphene layers is the result of graphene wrinkling around the sliding tip by molecular dynamics simulations. Their results support the hypothesis and explain the decreasing friction by the ability of multilayers to act as single material and resist wrinkling [36]. Dong et al. have analyzed the results of atomistic simulations of friction on hydrogenated graphene [37]. They concluded that the effects of hydrogenation on friction are mostly due to the enhanced atomic-scale roughness, excluding other suggested mechanisms such as changes in flexural rigidity or in adhesion [38].

20.6 Friction and Wear of Graphene at the Microscale

First experiments at larger normal forces were conducted by Kim et al. graphene films grown by CVD on copper (1 layer) and nickel films (1–10 layers) and then transferred to a SiO_2/Si substrate [39]. Using a silica lens with a diameter of 25.8 mm as counter surface they observed a reduction of the friction coefficient (COF) from 0.68 to 0.22 for Cu-grown graphene on SiO_2 and 0.12 for Ni-grown graphene. An investigation of the wear track after the sliding experiments by optical microscopy, XPS, and Raman spectroscopy showed that the graphene layer on Cu was removed and possibly transferred to the silica lens. On the nickel substrate, an amorphous carbon film stayed in the wear track.

Marchetto and coworkers studied the friction and wear behavior of epitaxial graphene grown on silicon carbide SiC-6H(0001) using a commercial microtribometer [40]. As counter face they used ruby spheres with a diameter of 500 μm. The range of normal forces was varied from 0.1 to 1 mN and the sliding speed was 50 μm/s and the length of the sliding track was 400 μm. The ruby sphere was scanned over the graphene sample in a reciprocating motion. In the friction experiments the graphene layer was found to reduce the friction strongly compared to the bare SiC substrate and was also found to be lower that the friction of a highly oriented pyrolytic graphite (HOPG) sample.

However, unlike the results presented in [39] the authors found a profound running-in effect. During the first few cycles the COF starts with 0.02 but then it rises quickly to approx. 0.04, see Fig. 20.3. By looking at the friction force during one cycle it was found that the raise of COF occurred at specific locations within the wear track. Therefore the graphene sample was investigated with AFM and FFM. The topography and friction maps showed that the graphene layer was torn away locally. At some places within the wear track however, graphene was still present. Consecutive experiments with a similar setup revealed that the friction vs. cycle data as shown in Fig. 20.3 could not always be reproduced. For a better understanding Wählisch et al. [7] used a similar microtribometer as in [40] and combined it with an Agilent 5400 AFM in order to investigate the wear track without transferring the substrate from the microtribometer to the AFM. Using this setup it was possible

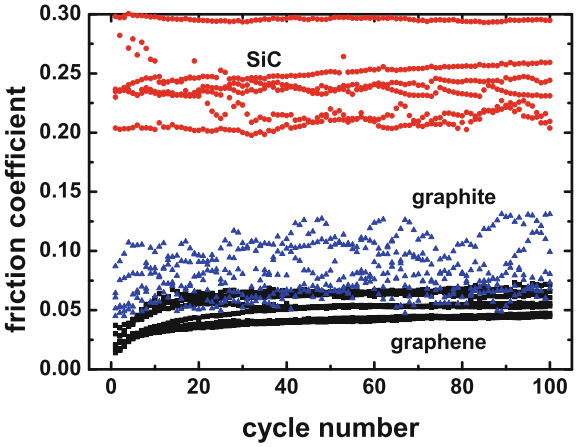


Fig. 20.3 Evolution of the friction coefficient on SiC (*red*), graphene (*blue*), and graphite (*black*) as function of cycle number for applied loads ranging from 0.1 to 1 mN. Reproduced from [40]

to correlate the friction force with the contact area between ruby sphere and the graphene coated sample (see Fig. 20.4). This was done by approximating the contact area by the width of the wear track as measured with the AFM.

It was found that the friction force scales linearly with the contact area, as predicted in the now classical description by Bowden and Tabor [41]. The scatter in the

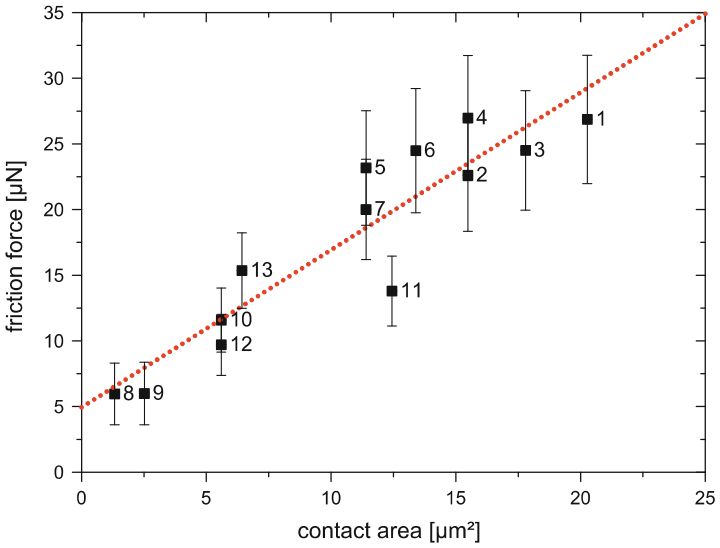


Fig. 20.4 Friction force of a hard ruby sphere sliding against epitaxial graphene plotted against the approximated contact area obtained by AFM measurements of the wear track. The numbers indicate the order of experiments under equal conditions. From [7]

friction force was caused by different contact geometries between the ruby sphere and the graphene sample during experiments although the experimental conditions were kept identical. Moreover the authors found that the graphene layer is quickly worn in the wear track due to high contact pressures at step edges. The graphitic interlayer which is present between the first layer of graphene and the SiC crystal is more stable and remains in the wear track. This leads to the conclusion that epitaxial graphene could be used for friction reduction for special applications where low contact pressures can be realized.

Shin et al. used a Hysitron triboindenter to perform scratch tests on exfoliated graphene on silicon and on epitaxial graphene on SiC [42]. The counterbody was a diamond tip with 1 μm radius. During the scratch test the normal force was ramped from 0 to 500 μN within 60 s. Unlike previous FFM studies, no differences in friction were found for one, two, or three layers of graphene. Failure of the graphene films was encountered between 450 and 2,250 μN. Nevertheless the authors concluded that graphene films can be used for wear protection.

This opinion is shared by Berman et al. They recently studied the lubrication and corrosion protection by solution processed few layer graphene that was spread from a graphene ethanol solution on a steel surface [43]. This type of graphene is produced by reduction of graphite-oxide. In macroscopic friction tests at a normal load of 2 N and a sliding speed of 90 mm/s it was found that an intermittent supply of graphene solution every 400 cycles lead to a friction reduction from 0.91 without any graphene to 0.14. Also the wear was found to be strongly reduced. Using Raman spectroscopy the authors concluded that this reduction of friction and wear might have been caused by a graphitic tribofilm formation within the wear track.

20.7 Summary

Nano- and microtribologcial studies of graphene demonstrate a great potential as solid lubricant, with an obvious limitation in wear resistance. Tribological methods with high resolution reveal the microscopic mechanisms of lubrication and failure. The excellent lubrication and strength of the graphene-like layer on the Si-face of SiC(0001) after thermal treatment indicates an interesting route for future research: lubrication by single layers of graphitic structure which are covalently bound to the underlying substrate.

Acknowledgments We thank our colleagues Diego Marchetto, Felix Wählisch, Tobin Filleter, Christian Held, Florian Hausen, Judith Hoth, and Eduard Arzt for stimulating discussions.

References

1. A.K. Geim, Science **324**, 1530 (2009)
2. C. Lee, X.D. Wei, J.W. Kysar, J. Hone, Science **321**, 385 (2008)

3. X.S. Li, W.W. Cai, J.H. An, S. Kim, J. Nah, D.X. Yang, R. Piner, A. Velamakanni, I. Jung, E. Tutuc, S.K. Banerjee, L. Colombo, R.S. Ruoff, Science **324**, 1312 (2009)
4. A. Reina, X.T. Jia, J. Ho, D. Nezich, H.B. Son, V. Bulovic, M.S. Dresselhaus, J. Kong, Nano Lett. **9**, 30 (2009)
5. K.V. Emtsev, A. Bostwick, K. Horn, J. Jobst, G.L. Kellogg, L. Ley, J.L. McChesney, T. Ohta, S.A. Reshanov, J. Rohrl, E. Rotenberg, A.K. Schmid, D. Waldmann, H.B. Weber, T. Seyller, Nat Mater **8**, 203 (2009)
6. S. Stankovich, D.A. Dikin, G.H.B. Dommett, K.M. Kohlhaas, E.J. Zimney, E.A. Stach, R.D. Piner, S.T. Nguyen, R.S. Ruoff, Nature **442**, 282 (2006)
7. F. Wählisch, J. Hoth, C. Held, T. Seyller, R. Bennewitz, Wear **300**, 78 (2013)
8. J. Choi, J.-S. Kim, I.-S. Byun, D. Lee, M. Lee, B. Park, C. Lee, D. Yoon, H. Cheong, K. Lee, Y.-W. Son, J. Park, M. Salmeron, Science **333**, 607 (2011)
9. H. Lee, N. Lee, Y. Seo, J. Eom, S. Lee, Nanotechnology **20**, 325701 (2009)
10. A.J. Marsden, M. Phillips, N.R. Wilson, Nanotechnology **24**, 255704, 1–10 (2013)
11. C. Lee, Q. Li, W. Kalb, X.-Z. Liu, H. Berger, R. Carpick, J. Hone, Science **328**, 76 (2010)
12. Z. Deng, N.N. Klimov, S.D. Solares, T. Li, H. Xu, R.J. Cannara, Langmuir **29**, 235 (2013)
13. S. Kwon, J.-H. Ko, K.-J. Jeon, Y.-H. Kim, J.Y. Park, Nano Lett. **12**, 6043 (2012)
14. J.-H. Ko, S. Kwon, I.-S. Byun, J. Choi, B. Park, Y.-H. Kim, J. Park, Tribol. Lett. **50**, 137 (2013)
15. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987)
16. H. Hölscher, U.D. Schwarz, O. Zwörner, R. Wiesendanger, Phys. Rev. B **57**, 2477 (1998)
17. H. Hölscher, D. Ebeling, U.D. Schwarz, Phys. Rev. Lett. **101**, 246105 (2008)
18. Y.L. Dong, X.Z. Liu, P. Egberts, Z.J. Ye, R.W. Carpick, A. Martini, Tribol. Lett. **50**, 49 (2013)
19. P. Egberts, Z.J. Ye, X.Z. Liu, Y.L. Dong, A. Martini, R.W. Carpick, Phys. Rev. B **88**, 035409 (2013)
20. M. Dienwiebel, G.S. Verhoeven, N. Pradeep, J.W.M. Frenken, J.A. Heimberg, H.W. Zandbergen, Phys. Rev. Lett. **92**, 126101 (2004)
21. A.P. Merkle, L.D. Marks, Appl. Phys. Lett. **90**, 064101 (2007)
22. J.B. Pethica, Phys. Rev. Lett. **57**, 3235 (1986)
23. M. Enachescu, D. Schleef, D.F. Ogletree, M. Salmeron, Phys. Rev. B **60**, 16913 (1999)
24. T. Filleter, R. Bennewitz, Phys. Rev. B **81**, 155412 (2010)
25. T. Filleter, J.L. McChesney, A. Bostwick, E. Rotenberg, K.V. Emtsev, T. Seyller, K. Horn, R. Bennewitz, Phys. Rev. Lett. **102**, 086102 (2009)
26. Q.Y. Li, C. Lee, R.W. Carpick, J. Hone, Phys. Status Solidi B-Basic Solid State Phys. **247**, 2909 (2010)
27. D.-H. Cho, L. Wang, J.-S. Kim, G.-H. Lee, E.S. Kim, S. Lee, S.Y. Lee, J. Hone, C. Lee, Nanoscale **5**, 3063 (2013)
28. A. Smolyanitsky, J.P. Killgore, V.K. Tewary, Phys. Rev. B **86**, 035412 (2012)
29. Z. Deng, A. Smolyanitsky, Q. Li, X.-Q. Feng, R.J. Cannara, Nature Mater **11**, 1032 (2012)
30. M.V. Rastei, B. Heinrich, J.L. Gallani, Phys. Rev. Lett. **111**, 084301 (2013)
31. S.G. Balakrishna, A.S. De Wijn, R. Bennewitz, in preparation (2014)
32. P. Steiner, E. Gnecco, T. Filleter, N.N. Gosvami, S. Maier, E. Meyer, R. Bennewitz, Tribol. Lett. **39**, 321 (2010)
33. F. Bonelli, N. Manini, E. Cadelano, L. Colombo, Eur. Phys. J. B Condens. Matter Complex Syst. **70**, 449 (2009)
34. A.S. de Wijn, A. Fasolino, A.E. Filippov, M. Urbakh, EPL **95**, 66002 (2011)
35. Y. Guo, W. Guo, C. Chen, Phys. Rev. B (Condens. Matter Mater. Phys.) **76**, 155429 (2007)
36. Z.J. Ye, C. Tang, Y.L. Dong, A. Martini, J. Appl. Phys. **112**, 116102 (2012)
37. Y. Dong, X. Wu, A. Martini, Nanotechnology **24**, 375701 (2013)
38. E.J. Sandoz-Rosado, O.A. Tertuliano, E.J. Terrell, Carbon **50**, 4078 (2012)
39. K.-S. Kim, H.-J. Lee, C. Lee, S.-K. Lee, H. Jang, J.-H. Ahn, J.-H. Kim, H.-J. Lee, ACS Nano **5**, 5107 (2011)
40. D. Marchetto, C. Held, F. Hausen, F. Wahlisch, M. Dienwiebel, R. Bennewitz, Tribol. Lett. **48**, 77 (2012)
41. F.P. Bowden, D. Tabor, *The Friction and Lubrication of Solids* (Clarendon Press, Oxford, 2008)
42. Y.J. Shin, R. Stromberg, R. Nay, H. Huang, A.T.S. Wee, H. Yang, C.S. Bhatia, Carbon **49**, 4070 (2011)
43. D. Berman, A. Erdemir, A.V. Sumant, Carbon **54**, 454 (2013)

Chapter 21
Superlubricity in Layered Nanostructures

Seymur Cahangirov and Salim Ciraci

Abstract Interaction between two surfaces in relative motion can give rise to energy dissipation and hence sliding friction. A significant portion of the energy is dissipated through the creation of non-equilibrium phonons. Recent advances in material synthesis have made the production of specific single layer honeycomb structures and their multilayer phases, such as graphene, graphane, fluorographene, MoS_2 and WO_2. When coated to the moving surfaces, the attractive interaction between these layers is normally very weak and becomes repulsive at large separation under loading force. Providing a rigorous quantum mechanical treatment for the 3D sliding motion under a constant loading force within Prandtl-Tomlinson model, we derive the critical stiffness required to avoid stick-slip motion. Also these nanostructures acquire low critical stiffness even under high loading force due to their charged surfaces repelling each other. The intrinsic stiffness of these materials exceeds critical stiffness and thereby the materials avoid stick-slip regime and attain nearly dissipationless continuous sliding. Remarkably, layered WO_2 a much better performance as compared to others and promises a potential superlubricant nanocoating. The absence of mechanical instabilities leading to conservative lateral forces is also confirmed directly by the simulations of sliding layers. Graphene coated metal surfaces also attain superlubricity and hence nearly frictionless sliding through a charge exchange mechanism with metal surface.

21.1 Introduction

Organisms in both micro and macro scales use friction to move. The way nature uses friction has frequently been an inspiration for scientific community. While the presence of friction is so important in our life, its absence is desperately demanded in

S. Cahangirov
Nano-Bio Spectroscopy Group, Departamento Fisica de Materiales, Centro de Fisica de Materiales CSIC-UPV/EHU-MPC and DIPC, Universidad Del Pais Vasco,
Av. Tolosa 72, 20018 San Sebastian, Spain
e-mail: seycah@gmail.com

S. Ciraci (✉)
Department of Physics, Bilkent University, 06800 Ankara, Turkey
e-mail: ciraci@fen.bilkent.edu.tr

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_21

most of our technological applications. We loose a substantial portion of the energy and material that we produce due to friction. Achieving superlubricity would help us to stop wasting enormous resources for nothing but producing heat of no use. Superlubricity can increase stability and sustainability of our technology alongside increasing its efficiency.

The key to achieve superlubricity is to understand friction in the most fundamental level. To this end, one expect that theoretical studies can help us to design superlubricant materials. Following this spirit, we organized this chapter starting with discussions of fundamental aspects of friction, building theoretical and computational methods to design materials and finally predicting novel materials that could be used as superlubricants. To achieve our objective we considered the exceptional properties of single layer honeycomb structures to exploit whether the surfaces in relative motion can attain superlubricity when coated with these single layer materials.

21.1.1 Dissipation Phenomena

Friction is not a fundamental physical force like gravity or electromagnetic interaction. It is a manifestation of a deeper phenomena called dissipation. Dissipation arises in systems having large number of interacting degrees of freedom. In sufficiently large systems, the macroscopic behavior of the system is generally irreversible even though the microscopic interactions are governed by reversible dynamics. In such systems, there is a probability distribution of states corresponding to equilibrium which occupies the largest volume in the configurational phase space. Any non-equilibrium distributions of states are less probable and eventually they are equilibrated through microscopic interactions.

In the case of friction, the dissipation of nonequilibrium phononic and electronic states have important role and it is governed by interactions among themselves and with each other. For example, phonons arise from the harmonic interaction between ions. However, the interaction between phonons themselves is governed by anharmonic terms of ionic interactions. In a perfectly harmonic crystal, non-equilibrium phononic states would remain in their initial distribution while in a highly anharmonic crystal they are quickly equilibrated through phononic dissipation.

21.1.2 Adiabatic Versus Sudden Processes

In sliding friction the non-equilibrium phononic and electronic states can be generated in two ways. The first way is generation through adiabatic processes. This kind of processes are generally not localized in space or time, which means that one can't show a microscopical region where such events start or end. For example, when two identical surfaces with weak interactions are sliding over each other

with equal but opposite center of mass velocities (which are well below the atomic vibration velocities), the system possesses an adiabatic non-equilibrium. The equilibrium configuration for such system is reached when atoms in both surfaces have same velocity distribution. The system will dissipate until this equilibrium is reached and all mechanical energy is transformed into random atomic vibrations. The surfaces will feel a force opposite and proportional to their center of mass velocity. This is, in general, the case for dissipative systems with adiabatic non-equilibrium generation mechanisms. The phononic and electronic dissipation mechanisms discussed above can be considered as examples in which the non-equilibrium state is generated adiabatically.

The second way is to generate non-equilibrium states through sudden processes. In this case, the process of generation can be traced by looking in a certain microscopical region in a certain time. As an example, consider a surface with asperities which have adhesive interaction with each other. When such surfaces are sliding over each other asperities that come close will cling to each other. Then they will be stretched due to the relative motion of surfaces that they are attached to. At some critical strain asperities will suddenly detach from each other and release all energy they have stored during the stretching as a burst of non-equilibrium phonons. Sliding systems which have this kind of microscopic events are said to be in the stick-slip regime.

Here it is important to point out that, if a non-equilibrium state is generated then it will be certainly dissipated no matter how weak the dissipation mechanisms are. The strength of the dissipation mechanisms only determine the time interval in which the equilibrium is reached. There is no way to regain the energy that comes off during the generation of non-equilibrium states. For this reason, the sudden processes determine the frictional properties of the systems which are in the stick-slip regime. In such systems the microscopic dissipation mechanisms have less importance and are usually lumped in a single coefficient as in (21.1).

21.1.3 Prandtl-Tomlinson Model

In the late 1920s when Einstein and Bohr were having a debate on philosophical issues brought by the newly developing quantum theory, Prandtl [1] and Tomlinson [2] separately developed concepts which are still at the heart of much older problem of friction. Their model explained the weak dependence of friction on sliding velocity. Here we present the important aspects of this model.

We start by pointing out that, Prandtl-Tomlinson model is one of the simplest yet being one of the most effective models describing the generation of non-equilibrium phonons. In this model, an elastic spring is dragged over a surface, as shown in Fig. 21.1a. The tip of this spring is interacting with the surface according to some potential energy landscape depending on the position of the tip on the surface. The other end of the spring (or the head of the spring) is pulled with constant velocity in certain direction. The restoring force on the tip is assumed to be directly proportional with the deviation of the tip from its equilibrium position. As mentioned in the

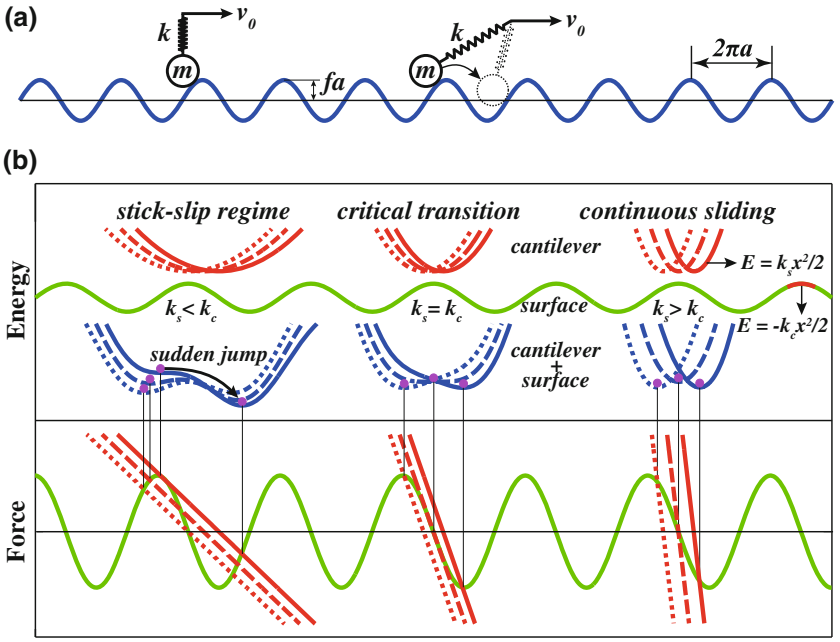


Fig. 21.1 **a** Prandt-Tomlinson model. **b** Schematic representation of stick-slip regime (*left*), critical transition (*middle*) and continuous sliding regime (*right*) in Prandtl-Tomlinson model. *Upper part* the potential energy curves of the surface (*green lines*) and of the tip (+cantilever) (*red lines*); *lower part* force variation of the surface (*green lines*) and of the tip (*red lines*). *Blue lines* represent the potential energy of the tip and surface. The *magenta dot* shows the position of the tip on the surface, while its other end is positioned at the minimum of the parabola shown with *red lines* in the *upper part*. The *dotted*, *dashed* and *solid lines* correspond to three different tip positions moving to the *right*

previous section, the dissipation due to interactions of the tip with surface atoms is lumped in a microscopic friction term proportional to the velocity of the tip. Assuming that the interaction between the tip and the surface has a sinusoidal profile one can write an equation of motion for the tip as following;

$$m\ddot{x} = -m\gamma\dot{x} + k(v_0 t - x) + f\sin(x/a) \tag{21.1}$$

which can be written in dimensionless form as;

$$\ddot{\tilde{x}} = \tilde{\gamma}\dot{\tilde{x}} + \tilde{k}(\tilde{v_0}\tilde{t} - \tilde{x}) + \sin(\tilde{x}) \tag{21.2}$$

where dimensionless variables are defined as $\tilde{x} = x/a$, $\tilde{\gamma} = \gamma\sqrt{ma/b}$, $\tilde{k} = ka/f$, $\tilde{v_0} = v_0\sqrt{m/ba}$ and $\tilde{t} = t\sqrt{b/ma}$. Here the most important parameter is $\tilde{k}$ which represents the ratio of the stiffness of the tip to the curvature of surface energy at its maximum points. If $\tilde{k} > 1$ the total energy of the tip-surface system always have one

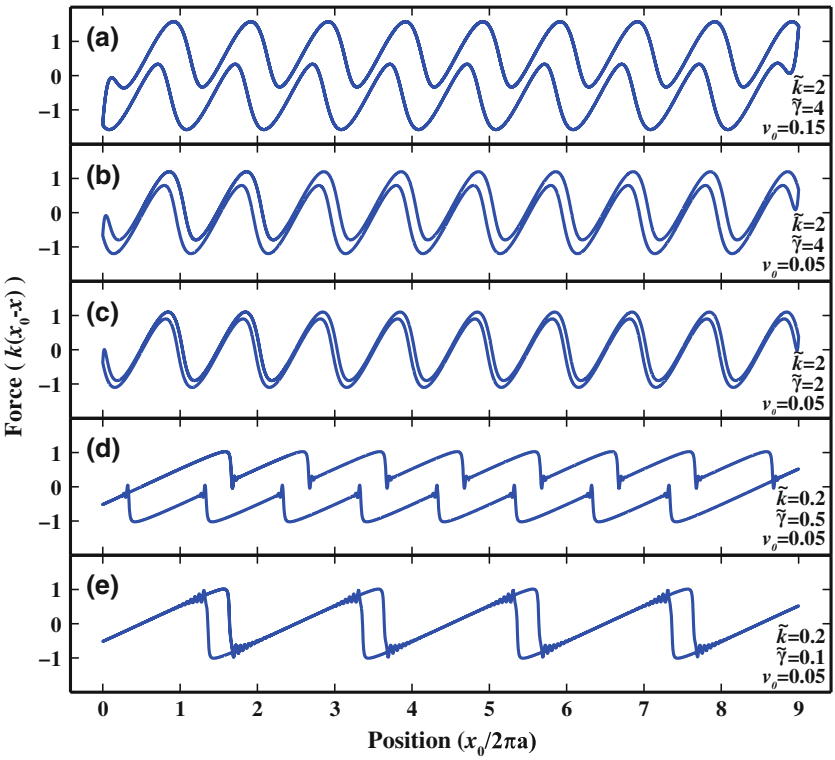


Fig. 21.2 Friction force felt by the tip during forward and backward sliding. The model parameters in each case are; **a** $\tilde{\gamma} = 4, \tilde{k} = 2, \tilde{v}_0 = 0.15$. **b** $\tilde{\gamma} = 4, \tilde{k} = 2, \tilde{v}_0 = 0.05$. **c** $\tilde{\gamma} = 2, \tilde{k} = 2, \tilde{v}_0 = 0.05$. **d** $\tilde{\gamma} = 0.5, \tilde{k} = 0.2, \tilde{v}_0 = 0.05$. **e** $\tilde{\gamma} = 0.1, \tilde{k} = 0.2, \tilde{v}_0 = 0.05$

minimum. In this case, for small velocities the tip gradually follows this minimum without making any sudden jumps. However when $\tilde{k} < 1$ the system possess multiple local minima and the tip can jump from one to another during sliding. This issue is explored further in forthcoming sections. For this simple system transition from stick-slip to continuous sliding regime occurs around $\tilde{k} = 1$. Here another important parameter is $\tilde{\gamma}$. The system shows overdamped and underdamped behavior when $\tilde{\gamma} \ll 1$ and $\tilde{\gamma} \gg 1$, respectively.

One can trace the force needed to slide the tip by calculating $\tilde{k}(\tilde{x_0} - \tilde{x})$ where $\tilde{x_0} = \tilde{v_0}\tilde{t}$ is position of the other end of the tip. The result of such calculation is presented in Fig. 21.2. Here the tip is slid forward and backward to complete a friction loop. The area covered by this loop correspond to the dissipated energy. One can see in Fig. 21.2a–c that when $\tilde{k} > 1$ the force loop is smooth which corresponds to continuous sliding. Comparing the area covered by loops presented here one can see that the friction force in continuous sliding regime is proportional to $\tilde{\gamma}\tilde{v_0}$.

Conversely, force loops having sudden changes can be seen in Fig. 21.2d, e which corresponds to stick-slip regime with $\tilde{k} < 1$. Comparing Fig. 21.2c, d one can see that

the area covered by the friction loop is much larger in (d), despite the decrease in $\tilde{\gamma}$. This shows that in stick-slip regime the friction force has much larger contribution from sudden processes compared to $\tilde{\gamma}\tilde{v_0}$ term that appear in continuous sliding regime. Another interesting phenomena is observed when Fig. 21.2d, e are compared. One can see that when $\tilde{\gamma} = 0.1$ the system is in underdamped stick-slip regime where double slips occur.

In the light of the above model the following trends are found when one looks at the variation of the average friction force $\tilde{F}$ with sliding velocity $\tilde{v_0}$. When $\tilde{k} > 1$ the average friction force linearly varies with sliding velocity and approaches zero as the velocity goes to zero. However, when $\tilde{k} < 1$ the average friction force converges to some finite value as sliding velocity approaches zero. Interestingly, when $\tilde{k}$ and $\tilde{\gamma}$ are small while $\tilde{v_0}$ is above some critical value, the system starts to show multiple slips which in turn lowers the average friction force dramatically.

21.1.4 Motivation

Advances in atomic scale friction [3–5] have provided insight on dissipation mechanisms. The stick-slip phenomena is the major process, which contributes to the dissipation of the mechanical energy through sudden or non-adiabatic transitions between bi-stable states of the sliding surfaces [1, 2, 6, 7]. During a sudden transition from one state to another, the velocities of the surface atoms exceed the center of mass velocity sometimes by orders of magnitudes [8]. Local vibrations are created thereof evolve into the non-equilibrium system phonons via anharmonic couplings [9] within picoseconds [10, 11]. In specific cases, even a second state in stick-slip can coexist [7].

In Fig. 21.1, two regimes of sliding friction are summarized within the framework of Prandtl-Tomlinson model [1, 2, 8], where an elastic tip (+cantilever) moves over a sinusoidal surface potential. The curvature of this potential at its maximum gives the value of the critical stiffness k_c. If the intrinsic stiffness of the tip k_s is higher than this critical stiffness i.e. $k_s/k_c > 1$, the total energy of the tip-surface system always has one minimum. The sliding tip gradually follows this minimum, which results in the continuous sliding regime. Conversely, if the tip is softer than the critical value, then it is suddenly slipped from one of the bi-stable states to the other. This slip event can be activated by thermal fluctuations even before the local minimum point becomes unstable [12]. Experimentally, using friction force microscope, Socoliuc et al. [13] showed that the transition from stick-slip regime to continuous sliding attaining ultralow friction coefficient can be achieved by tuning the loading force on the contact.

Superlubricant materials composed of weakly interacting two-dimensional (2D) layers have been a central figure of intense studies in tribology. In this respect, recently synthesized, two-dimensional, single layer honeycomb structures, which have been synthesized, hold the promise of being potential lubricant material between two sliding flat surfaces.

21.2 Superlubricity Between Two Layers of Graphene Derivatives and Transition Metal Dichalcogenides

Here we investigate the sliding friction between two same pristine layers of nanostructures, such as graphane, [14, 15] fluorographene, [16, 17] molybdenum disulfide, [18] and tungsten dioxide, [19] (abbreviated according to their stoichiometry as CH, CF, MoS_2 and WO_2 respectively) using the Density Functional Theory [20, 21]. We find that these nanostructures avoid stick-slip even under high loadings and execute continuous sliding. Consequently, the sliding occurs without friction that would originate from the generation of non-equilibrium phonons. Our approach mimics the realistic situation, where the total energy and forces are calculated from first-principles as two-dimensional (2D) two layers undergo a 3D sliding motion under a constant (normal) loading force. This is the most critical and difficult aspect of our study. In this respect, our results provide a 3D rigorous *quantum mechanical* treatment for the 1D and empirical Prandtl-Tomlinson model [1, 2].

The nanostructures considered in the present study are recently discovered insulators having honeycomb structure, which can form suspended single layers as well as multilayers. The unusual electronic, magnetic and elastic properties of these layers have been the subject of recent numerous studies. In particular, they have large band gaps to hinder the dissipation of energy through electronic excitation and have high in-plane stiffness ($C = (1/A)\partial^2 E_s/\partial\epsilon^2$, i.e. the second derivative of the strain energy relative to strain per unit area, A being the area of the unit cell) [17–19, 22]. Analysis based on the optimized structure, phonon and finite temperature molecular dynamics calculations demonstrate that each suspended layer of these nanostructures are planarly stable [15, 17–19]. In graphane, positively charged three hydrogen atoms from the top side and another three from the bottom are bound to the alternating and buckled carbon atoms at the corners of hexagons in graphene to form a uniform hydrogen coverage at both sides (see Fig. 21.3a). Recently synthesized CF [16] is similar to CH, but F atoms are negatively charged. Tribological properties of carbon based fluorinated structures have been the focus of interest [23, 24]. In the layers of MoS_2 or WO_2, the plane of positively charged transition metal atoms is sandwiched between two negatively charged outer S or O atomic planes. It was shown that MoS_2 structure can have ultralow friction [25]. Theoretically, the static energy surfaces are calculated during sliding at $MoS_2(001)$ surfaces [26]. Apparently, the interaction energy between two single layers of these nanostructures is mainly repulsive due to charged outermost planes except very weak Van der Waals attractive interaction around the equilibrium distance. In Fig. 21.3, each layer being a large 2D sheet consisting of three atomic planes mimics one of two sliding surfaces. In practice, sliding surfaces can be coated by these single layer nanostructures as one achieved experimentally [27].

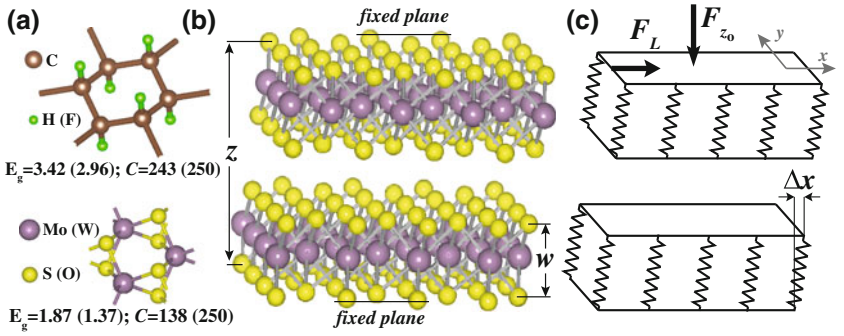


Fig. 21.3 **a** Ball and stick model showing the honeycomb structure of graphane CH (fluorographene CF) (*top*) and MoS_2 (WO_2) (*bottom*). Calculated values of energy gaps E_g and in-plane stiffness C are also given in units of eV and J/m^2 respectively. **b** Two MoS_2 layers sliding over each other have the distance z between their outermost atomic planes. **c** Each layer is treated as a separate elastic block. Lateral F_L and normal (loading) F_{z_o} forces, the shear of bottom atomic plane relative to top atomic plane in each layer $\Delta x(y)$, and the width of the layer w, are indicated

21.2.1 Methods

To investigate the sliding friction between surfaces coated with single layer nanostructures we used an approach, which is based on quantum mechanics. Our results are obtained by state-of-the-art Density Functional Theory (DFT) plane wave calculations within the Generalized Gradient Approximation (GGA) [28] including van der Waals corrections [29] and using PAW potentials [30]. All structures have been treated within supercell geometry using the periodic boundary conditions. A plane-wave basis set with kinetic energy cutoff of 400 and 500 eV is used for transition metal and carbon based structures respectively. In the self-consistent potential and total energy calculations the Brillouin zone is sampled by fine meshes. All atomic positions and lattice constants are optimized by using the conjugate gradient method where total energy and atomic forces are minimized. The convergence for energy is chosen as 10^{-5} eV between two steps, and the maximum force allowed on each atom is less than 10^{-4} eV/Å. Numerical plane wave calculations have been performed by using VASP package [31, 32]. Further details of the calculations can be obtained from the [33, 34].

21.2.2 Critical Curvature

We consider two layers of the same nanostructures in relative motion, where the spacing z between the bottom atomic plane of the bottom layer and the top atomic plane of the top layer is fixed. These layers mimics two surfaces coated by these nanostructures. These layers mimics two surfaces, which are coated by Here the

frictional behavior of the system is dictated mainly by C–H(F), Mo–S and W–O bonds and their mutual interactions. These layers are represented by periodically repeating rectangular unit cells. We calculate the value of the equilibrium lattice constants, which increase as z decreases. For each value of z the fixed atomic layer at the top is displaced by x and y on a mesh within the quarter of the rectangular unitcell. Then all possible relative positions (displacements) between fixed atomic layers are deduced using symmetry. At each mesh point all atoms of the system except those of fixed top and bottom planes are relaxed and the total energy of the system $E_T(x, y, z)$ (comprising both layers) is calculated. We have also derived $\Delta x(x, y, z)$ and $\Delta y(x, y, z)$ data which correspond to the shear (deflection) from the equilibrium position of the relaxed atomic planes relative to the fixed atomic plane of the same layer as illustrated in Fig. 21.3c. The matrices of these data are arranged for each nanostructure using the mesh spacing of ∼0.2 Å in x and y directions. The forces exerting on the displacing top layer in the course of relative motion of layers are calculated from the gradient of the total energy of the interacting system, namely $F(x, y, z) = -\nabla E_T(x, y, z)$ at each mesh point (x, y). These forces are in agreement with the resultant of the atomic forces calculated for the top layer using Hellman-Feynman theorem. Eventually, the matrices of all data, namely $E_T(x, y, z)$, $\Delta x(x, y, z)$, $\Delta y(x, y, z)$ and $\mathbf{F}(x, y, z)$ are made finer down to mesh spacing of ∼0.05 Å using spline interpolation.

The properties affecting the friction between layers should be derived under a given constant loading force. First of all we preset the value of applied loading, F_{z_o}, which corresponds to the operation pressure when divided by the cell area A, namely $\sigma_N = F_{z_o}/A$. We obtain the normal force from $F_z(x, y, z) = -\partial E_T(x, y, z)/\partial z$ and for each x and y we calculate the value of z where $F_z(x, y, z) = F_{z_o}$ and abbreviate it as $z_o(x, y)$. Then by using spline interpolation in z direction we calculate the x and y dependence of $F_{x_o}[x, y, z_o(x, y)]$ and $F_{y_o}[x, y, z_o(x, y)]$, as well as $\Delta x_o[x, y, z_o(x, y)]$ and $\Delta y_o[x, y, z_o(x, y)]$ for a given F_{z_o}. The lateral force is then $\mathbf{F}_L[x, y, z_o(x, y)] = F_{x_o}\hat{i} + F_{y_o}\hat{j}$. Integrating the lateral force over the rectangular unitcell we obtain,

$$E_I[x, y, z_o(x, y)] = \int_0^x \int_0^y \mathbf{F}_L(x, y, z_o(x, y)) \cdot \mathbf{dr} \tag{21.3}$$

where $E_I[x, y, z_o(x, y)]$ is the interaction energy for displacement (x, y) in the cell under applied constant loading force F_{z_o}. It should be noted that E_I is different from $E_T(x, y, z)$ (but $E_I \to E_T$ for $z \gg 1$) and is essential to reveal the friction coefficient. Contour plots of E_I of two sliding MoS_2 layers calculated for $\sigma_N = 15$ GPa are shown in Fig. 21.4a and those of CH, CF, WO_2 in Fig. 21.5a. The profile of E_I is composed of hills arranged in a triangular lattice. These hills correspond to the relative positions when the charged atoms of adjacent layers have the minimum distance. The hills are surrounded by two kind of wells. The difference between these two wells is enhanced with increasing pressure. The wells form a honeycomb

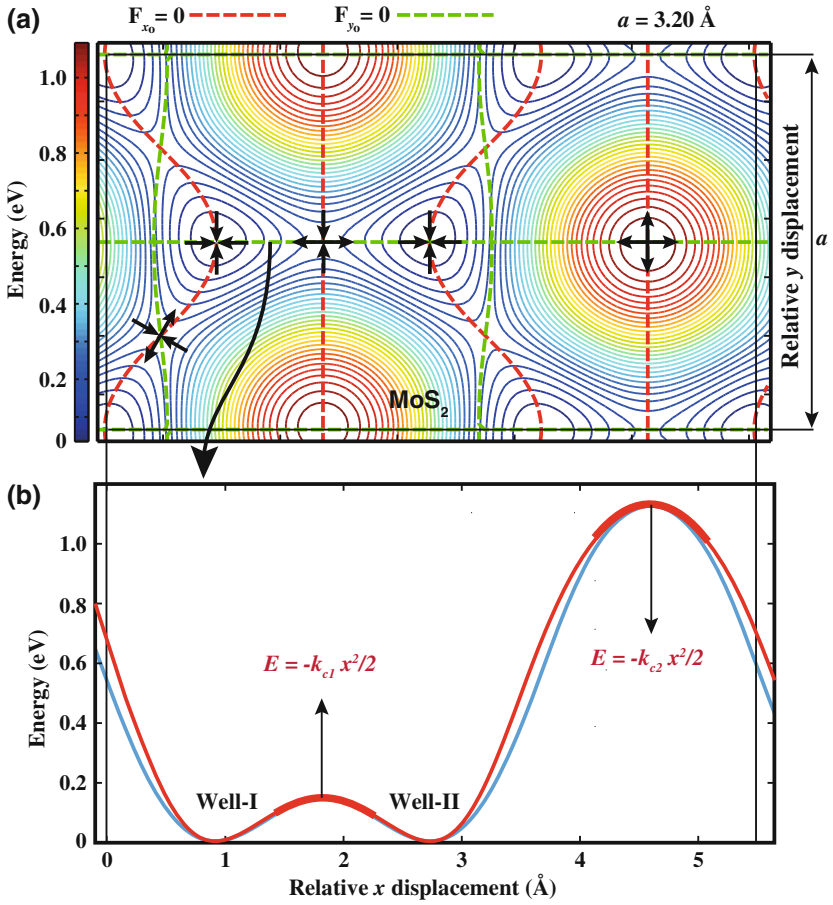


Fig. 21.4 **a** The contour plot of interaction energy E_I of two sliding layers of MoS_2. The zero of energy is set to $E_I[0, 0, z_o(0, 0)]$. The energy profile is periodic and here we present the rectangular unitcell of it. The width of this unitcell in y-direction is equal to the lattice constant a of the hexagonal lattice. Forces in x- (y-) direction is zero along the *red* (*green*) dashed lines, respectively. There are several points at which the lateral force $\mathbf{F_L}$, is zero. The arrows at these critical points indicate the directions where the energy decreases. **b** The energy profiles of E_I (*blue line*) and E_I^o (*red line*) along the *horizontal line* with $F_y = 0$ for MoS_2. Loading pressure in all cases is $\sigma_N = 15$ GPa

structure and are connected to each other through the saddle points (SP). When the layers are moved over each other they will avoid the relative positions corresponding to the hills. For example, if the layers are pulled in the y-direction they will follow the curved $F_x = 0$ path passing through the wells and SP but not the straight one passing through the hills as shown in the Fig. 21.4b. This makes SP very important because moving from one well to the adjacent one requires to overcome the barriers at these points. We note that the critical stiffness can be calculated from the curvature of E_I^o, which is obtained by subtracting the strain energies of two sliding MoS_2 layers,

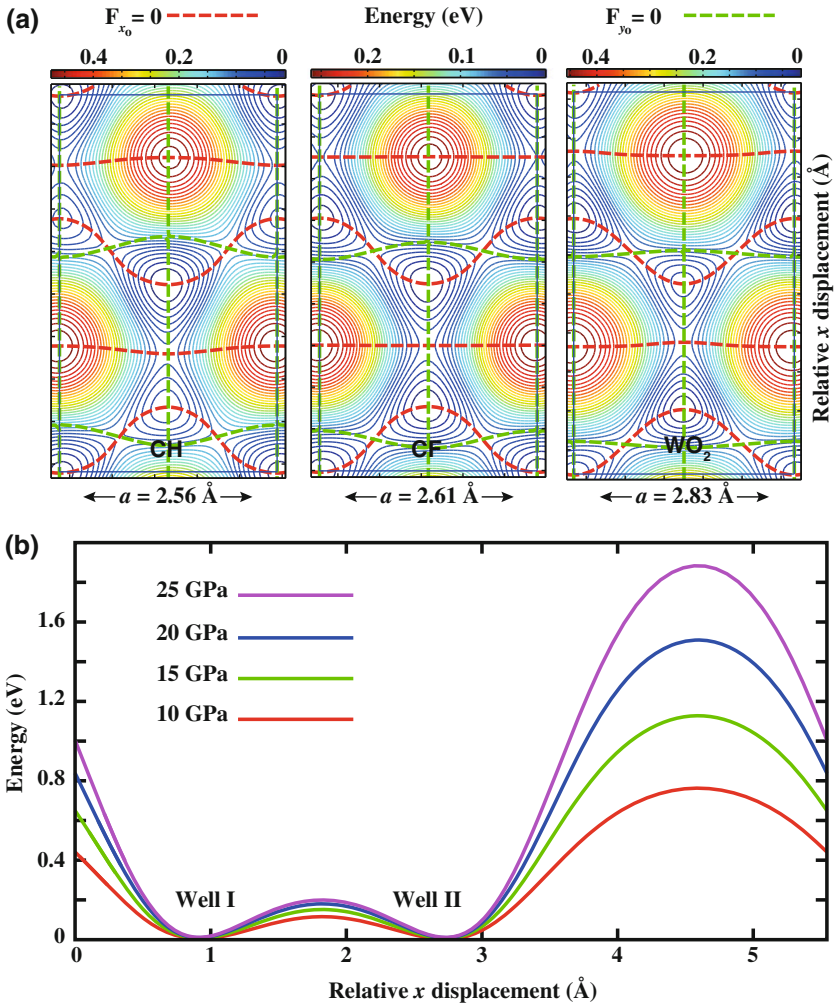


Fig. 21.5 **a** Contour plots of interaction energy E_I of two layers of CH, CF, and WO_2 executing sliding motion under constant loading pressure are presented in their rectangular unit cells. The zero of energy is set to $E_I[0, 0, z_o(0, 0)]$. Loading pressure in all cases is $\sigma_N = 15$ GPa. **b** Variation of interaction energy E_I^0 with applied loading for MoS_2 structure along the straight $F_y = 0$ line passing through two wells, saddle point and one hill

namely $E_I^o = E_I - k_s(\Delta x_o^2 + \Delta y_o^2)$ and by replacing x by $x - 2\Delta x_o$. While the SP serves as a barrier in the direction joining the nearby wells it acts as a well in the perpendicular direction joining the hills. Since we are interested in the curvature of the SP in the former direction we have made a plot along the $F_y = 0$ line which passes through the hill, the wells and the SP in between as shown in the Fig. 21.4b. We derive two critical stiffness values from E_I^o curve for a given normal loading

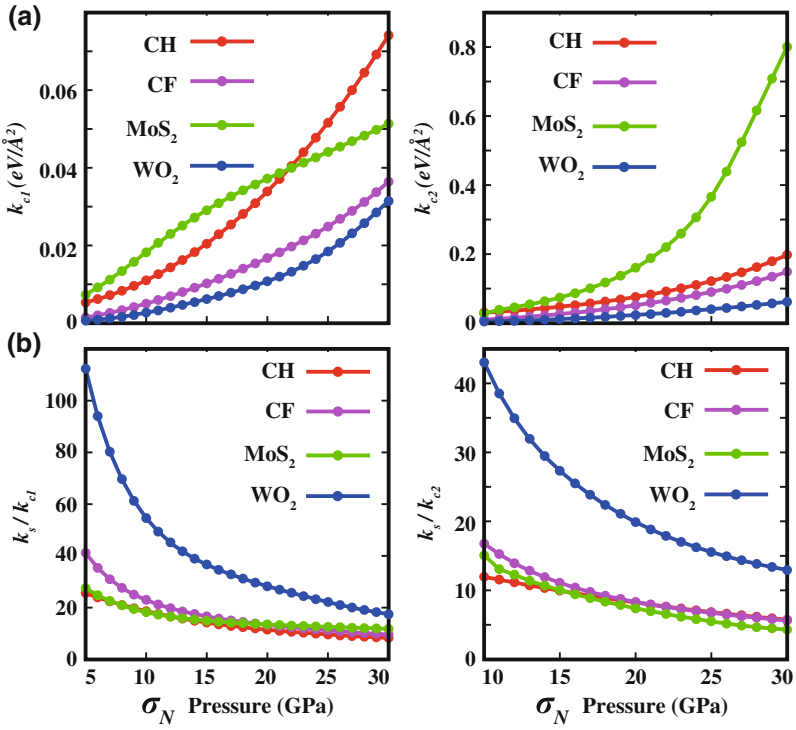


Fig. 21.6 **a** The variation of k_{c1} and k_{c2} with loading pressure. **b** The variation of the frictional figures of merits k_s/k_{c1} and k_s/k_{c2}, with loading pressure calculated for CH, CF, MoS_2 and WO_2

force; namely k_{c1} at the SP and k_{c2} at the hill by fitting the curve at the maxima of the barriers to a parabola. Although the hills will be avoided during sliding motion the curvature at these points are calculated for completeness. We also present the variation of E_I^0 with applied loading for MoS_2 structure in Fig. 21.5b. Note that, the variation of the amplitude at the saddle point is minute. In Fig. 21.6a the variation of k_{c1} and k_{c2} of CH, CF, MoS_2 and WO_2 with loading pressure σ_N is presented. Generally, the critical stiffness, in particular k_{c1} is low due to repulsive interaction between sliding layers. This facilitates the transition to continuous sliding.

21.2.3 Intrinsic Stiffness

Next we calculate the intrinsic stiffness k_s of individual MoS_2 layers using the force and the displacement data. For each x and y the lateral forces $F_{x_o}[x, y, z_o(x, y)]$ and $F_{y_o}[x, y, z_o(x, y)]$ versus the displacements $\Delta x_o[x, y, z_o(x, y)]$ and $\Delta y_o[x, y, z_o(x, y)]$, respectively are plotted. As shown in Fig. 21.7, this data falls on a straight

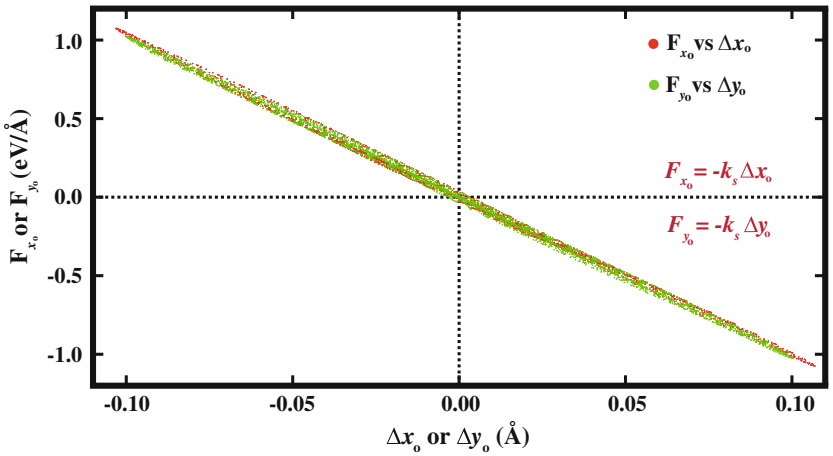


Fig. 21.7 The force versus shear values along x- and y-directions for each mesh point by *red* and *green dots*, respectively

line having a negative slope as expected from Hook's law of elasticity. We note that the elastic properties of layers having honeycomb structure is uniform and is independent of the direction of displacement and force [22]. The magnitude of the slope, $k_s = -F_{x(y)_o}/\Delta x(y)_o$ gives us the stiffness of the layers. Note that, normally the stiffness is defined as stress over strain and has units of energy per volume. Here we only need the ratio of material stiffness to the critical stiffness and should have the same units. The critical stiffness was calculated as second order spatial derivative of energy in the unitcell and it has units of energy per unitcell per unit area. As defined above, the stiffness of layers, k_s, also has units of energy per unitcell per unit area. Calculated intrinsic stiffness values of CH, CF, MoS_2 and WO_2 in the range of σ_N from 5 to 30 GPa are found to be 6.15 ± 0.15, 4.5, 10.0 ± 0.3 and 15.2 ± 0.3 eV/Å^2, respectively. Clearly, these values of k_s, in particular those of MoS_2 and WO_2 are rather high.

21.2.4 Frictional Figure of Merit

Based on the discussion at the beginning, the ratios k_s/k_{c1} and k_s/k_{c2} give us a dimensionless measure of performance of our layered structures in sliding friction. When these ratios are above two (since both layers in relative motion contribute), the stick-slip process is replaced by continuous sliding, whereby the dissipation of mechanical energy through phonons is ended. Under these circumstances the friction coefficient diminish, if other mechanisms of energy dissipation were neglected. For this reason one may call these ratios as a frictional figures of merit of the layered materials. In Fig. 21.6b we present the variations of the ratios k_s/k_{c1} and k_s/k_{c2} with

normal loading forces. Even for very large σ_N, $k_s/k_{c1} > 2$ and $k_s/k_{c2} > 2$. For usual loading pressures, the stiffness of MoS_2, CF and CH is an order of magnitude higher than corresponding critical values. Interestingly, for WO_2 this ratio can reach to two orders of magnitudes at low pressures. The absence of mechanical instabilities has been also tested by performing extensive simulations of the sliding motion of layers in very small displacements. C–H, C–F, Mo–S and W–O bonds in each case of two layers in relative motion under significant loading force did not display the stick-slip motion.

21.2.5 Stick-slip in Silicane: A Counter Example

Conversely, we now examine the sliding of two silicane [35, 36] layers (abbreviated as SiH and composed of silicene [37] saturated by hydrogen atoms from both sides, like graphane) with $k_s = 2.1 \pm 0.1$ eV/Å^2 for 2 GPa $\leq \sigma_N \leq$ 8 GPa. This is an interesting material because the onset of stick-slip occurs already at low loading pressures and exhibits a pronounced asymmetry in the direction of sliding between two wells. In Fig. 21.8 we present the lateral force variation calculated for two different loading pressures. For small loading pressure, σ_N=2 GPa the stick-slip is absent since approaching the SP from Well-I, the curvature is $k_{c,I} = 0.28$ eV/Å^2 and from Well-II it is $k_{c,II} = 0.16$ eV/Å^2, thus $k_s/k_{c,I or II} > 2$ for both directions. Whereas, once the pressure is raised to $\sigma_N = 8$ GPa stick-slip already governs the sliding friction, since $k_{c,I}$ reaches 1.38 eV/Å^2. Interestingly, since $k_{c,II}$ is only 0.28 eV/Å^2 for $\sigma_N = 8$ GPa, going from Well-II to Well-I a slip event occurs at SP. Eventually, one sees in Fig. 21.8 a hysteresis in the variation of F_L leading to energy dissipation.

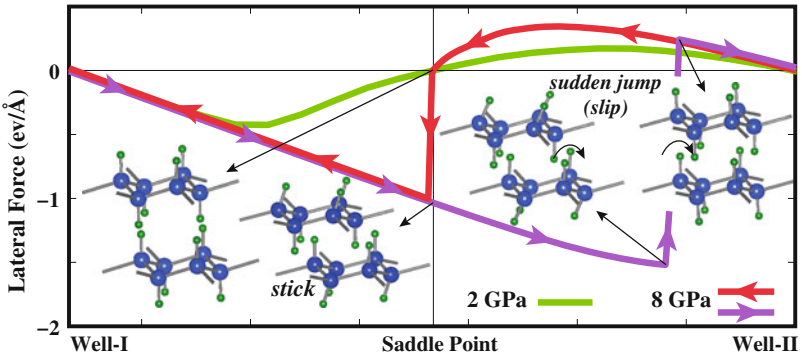


Fig. 21.8 Calculated lateral force variation of two single layer SiH under two different σ_N. The *top layer* is moving to the *right* or to the *left* between two wells. Atomic positions of two SiH layers in stick and slip stages are shown by inset. The movement of SiH layers under loading pressure of $\sigma_N = 8$ GPa is presented as a supplemental material of our work [33]

21.3 Superlubricity Between Graphene Coated Metal Substrates

Bulk counterparts of molybdenum disulfide and graphite flakes were used as a solid lubricant in industrial applications long before the 2D layers constituting them were isolated. The key features which make these materials so important in friction science are strong covalent intralayer bonds in contrast to weak van der Walls interlayer interactions. As discussed in the previous sections, the contrast between these intralayer and interlayer interactions can be quantified in terms of frictional figure of merit [33].

Experimental and theoretical studies have shown that friction force between graphite layers can be very small when the layers are slid with a certain angle to each other [38, 39]. In this case the rotated layers are incommensurate with respect to each other and the corrugation potential between such layers is flat. It was shown that, the torque felt by rotated layers can twists them until the layers become commensurate which increases the friction force [40]. Transition between commensurate and incommensurate states of graphene flakes on graphite was investigated in detail by several theoretical studies [41–43].

Recently, several experimental works investigated the variation in friction force when the number of 2D layers are varied from single layer to many layers representing the bulk structure [44–46]. Lee et al. used friction force microscope with a SiN tip to investigate atomic friction on graphite and graphene flake prepared on silicon oxide layer [44]. They have found that friction on graphene was lower than on silicon oxide but higher than that on graphite. Their results show that friction force monotonically decreases as the number of graphene layers are increased and approaches the bulk value found for graphite. They have attributed this trend to long ranged van der Walls interactions between layers.

Filleter et al. used atomic force microscope with cantilevers coated in a polycrystalline diamond film and found that friction force on single layer graphene epitaxially grown on SiC is lower than that on SiC surface but higher than bilayer graphene on SiC [45]. In contrast to other works, they have found the friction force to be higher on graphite compared to bilayer graphene. They have found the similar trends when oxidized single crystal silicon cantilevers with sharper tips were used. They assert that, the lower friction measured in bilayer graphene is related to suppressed electron-phonon coupling which plays important role in dissipation.

Much recently, the variation of friction force with number of layers was investigated for structures composed of graphene, molybdenum disulfide, niobium diselenide and hexagonal boron nitride layers [46]. Similar to results of Lee et al., it was found that the friction force decreases as the number of layers increase. It was also found that, this trend continues to be observed in suspended layers while it is suppressed when graphene layers are deposited on highly adherent mica surface. Here the observed trend was attributed to higher compliance of thin layers to perpendicular stress. Together with finite element simulations they show that as the tip is slid over, the layers are puckered and piled up in direction of sliding, which increases the contact area and resistance to sliding. The puckering is more pronounced when

the number of layers are decreased. Due to similar reasons, the trend is suppressed in mica because it prevents graphene layers from puckering.

The potential of graphene as a lubricant material can be revealed realistically, when it is placed between two flat sliding surfaces. Here we investigate the energy dissipation and the strength of the potential corrugation between two Ni(111) surfaces having $n = 0 - 5$ layers of graphene in between. We treat infinite surfaces using periodic boundary conditions, which also minimizes effects such as puckering or rippling [46]. Our approach mimics a realistic situation where the metallic surfaces are coated by graphene layers and the radii of asperities are much larger compared with atomic scales. Similar to what detailed in the previous section, the interaction energy, as well as lateral forces are calculated using quantum mechanical treatments as 2D layers execute a 3D sliding motion under a given constant normal force. We found that strong adhesive forces between Ni(111) surfaces, which lead to strong energy dissipation and wear are substantially suppressed when a single layer of graphene is inserted between the surfaces. However, the system enters into the continuous sliding regime only after the second layer of graphene is inserted, whereby each graphene layer becomes attached to one Ni(111) surface. Even more interesting is that inserting more graphene layers between Ni(111) surfaces decreases the friction gradually. On the other hand, the friction between graphene layers sliding over each other are larger and practically independent of the number of layers n in between, when the supporting Ni surfaces are not present. These results reveal the capacity of graphene as a superlubricant leading to nearly frictionless sliding and are explained by a charge exchange mechanism between graphene and Ni slabs.

21.3.1 Model and the Atomic Structure

The frictional properties of graphene layers sandwiched between Ni(111) surfaces and those of bare graphenes are treated using the models described in Fig. 21.9. In all calculations sliding nickel surfaces are represented by slabs consisting of three atomic layers of bulk Ni. We apply periodic boundary conditions along the plane parallel to the surfaces with a primitive unitcell comprising one Ni and two C atoms in each layer. The interaction between periodic images of Ni slabs is hindered by introducing a vacuum spacing of 15 Å. The structure presented in Fig. 21.9a is named as Ni-ABCBA-Ni structure, where A, B, and C correspond to certain in-plane configuration of carbon atoms. To avoid any confusion the atomic layers comprising the Ni slabs are arranged in a mirror symmetry. This arrangement is presented in Fig. 21.9a, while the in-plane configuration of Ni and C atoms in each plane is shown in Fig. 21.9b. Nickel atoms positioned at the bridge sites of graphene structure attracts C atoms and slightly breaks the honeycomb symmetry, as shown in Fig. 21.9b. The nature of this interaction is discussed in forthcoming parts in detail. The equilibrium positions of Ni and C atoms are attained by structure optimization at a given constraint. Details of calculations are given in Sect. 21.2 and in [34].

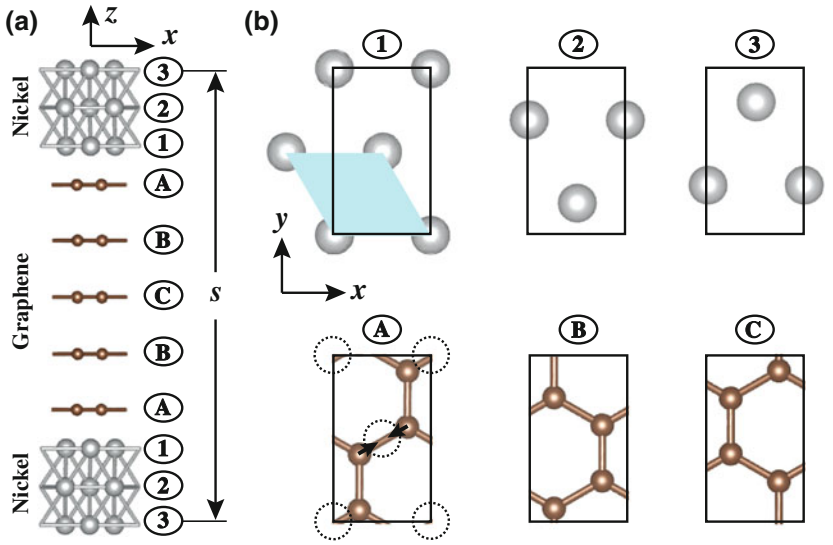


Fig. 21.9 **a** *Side view* of the arrangement of the Ni-ABCBA-Ni structure. The outermost Ni(111) atomic planes are fixed at the separation s. **b** *Top view* of individual layers constructing the Ni-ABCBA-Ni structure. The primitive unitcell is shown by *blue shaded area*. *Dotted circles* represent optimized positions of Ni atoms below the graphene layers in configuration A

21.3.2 Adhesion Hysteresis

We start our analyses by calculation of forces on outermost atoms of Ni slabs when they are kept fixed during the relaxation while the separation, s, between them (see Fig. 21.9a) is gradually varied. We start by two Ni slabs each composed of three atomic layers with no graphene in between. The dashed green curve shown in Fig. 21.10a is obtained when s is gradually decreased. One can observe a slightly attractive region followed by a sudden increase in the attractive force after which the force starts to decrease until the equilibrium distance is reached and the force becomes repulsive. During the sudden increase in attractive force both layers are elongated towards each other and after this stage the distance between the facing atomic layers remain nearly constant until the equilibrium is reached. The red solid line in Fig. 21.10a shows the variation of forces when s is gradually increased. This time, the curve takes a different route right at the point where the sudden increase of the attractive force was observed. When s is increased further the facing two layers detach from the slabs and attach to each other. The observed hysteresis manifests the adhesion and wear phenomena frequently observed between metallic contacts.

Next we insert one graphene layer in the minimum energy configuration A described in Fig. 21.10b between two Ni slabs. This graphene layer screens the interaction between Ni surfaces and significantly decreases the attractive potential between them, however the hysteresis is still present, as seen in Fig. 21.10b. We also

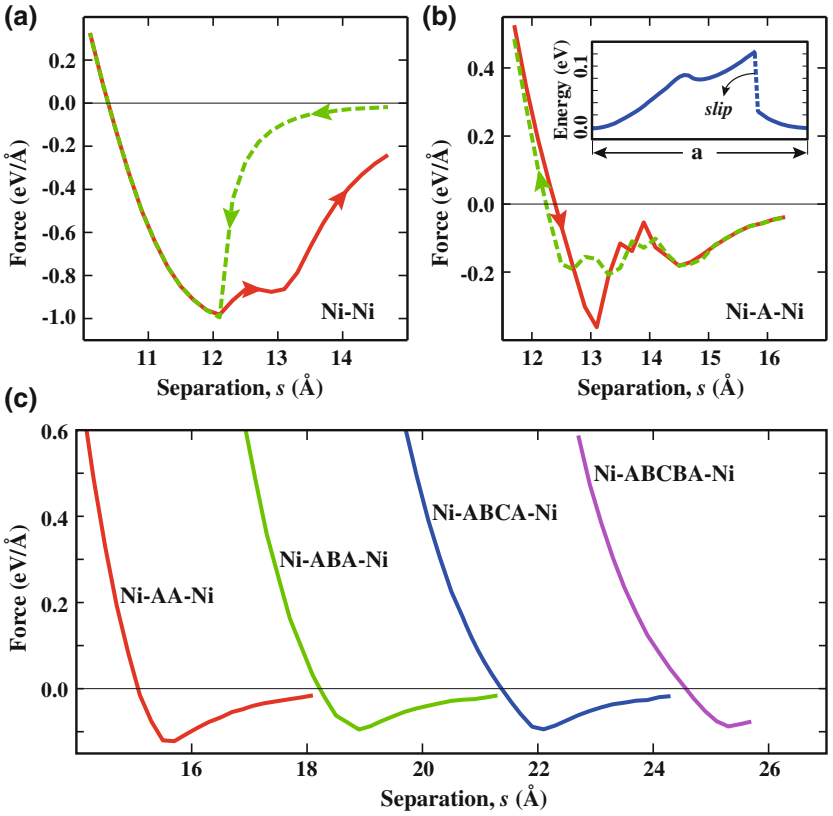


Fig. 21.10 Adhesion hysteresis curves for **a** Ni–Ni and **b** Ni-A-Ni structures and its stick-slip behavior shown by inset. **c** Normal force along z axis as a function of separation s for Ni-graphene-Ni structures with 2–5 graphene layers

observe stick-slip behavior when the Ni slabs are laterally moved relative to each other with one layer graphene in between making sudden jumps. As illustrated by inset in Fig. 21.10b, the sudden variation of energy in constant height mode demonstrates the presence of stick-slip motion causing the dissipation of mechanical energy.

When the second graphene layer is inserted, each layer becomes attached to Ni(111) surface. Under these circumstances the hysteresis is completely removed and the attractive forces are weakened. Further increasing the number of layers shows minor changes as seen in Fig. 21.10c. From these observations we deduce that, inserting single layer graphene can protect the Ni layer from wear during sliding, however one layer is not enough for the onset of the continuous sliding regime. It becomes possible only by including a second layer of graphene whereby each Ni slab is coated by graphene.

21.3.3 Trends in Multilayers

To investigate the effect of including more layers on the potential corrugation during the sliding of the layers under constant pressure, we first calculate the total energies E_T when outermost Ni layers are kept fixed at various relative lateral (x, y) positions and at fixed separation s [33]. These calculations are performed in a 3D grid of x, y, s. The distances between the data points were taken to be ∼0.2 Å in the lateral plane and 0.2 Å in perpendicular axis i.e s, which is then made finer down to ∼0.05 Å by spline interpolation. We also generate F_x, F_y and F_z matrices from the gradient of the total energy $F_{x,y,z} = -\partial E_T(x, y, z)/\partial x, y, z$, which is consistent with Hellmann-Feynman forces calculated on fixed atoms of outermost planes. We then retrieve F_x and F_y corresponding to a given F_z (normal pressure) at each (x, y) in the unit cell and generate the profiles of potential corrugation from $\int F_x dx + F_y dy$, where the minimum of total energy is set to zero. The profiles (contour plots) of potential corrugation calculated for Ni-AA-Ni and AA i.e. two flat graphene layers without Ni(111) are shown in the top and bottom panels of Fig. 21.11a, respectively. We note that the amplitude of the potential corrugation (i.e. the difference between the

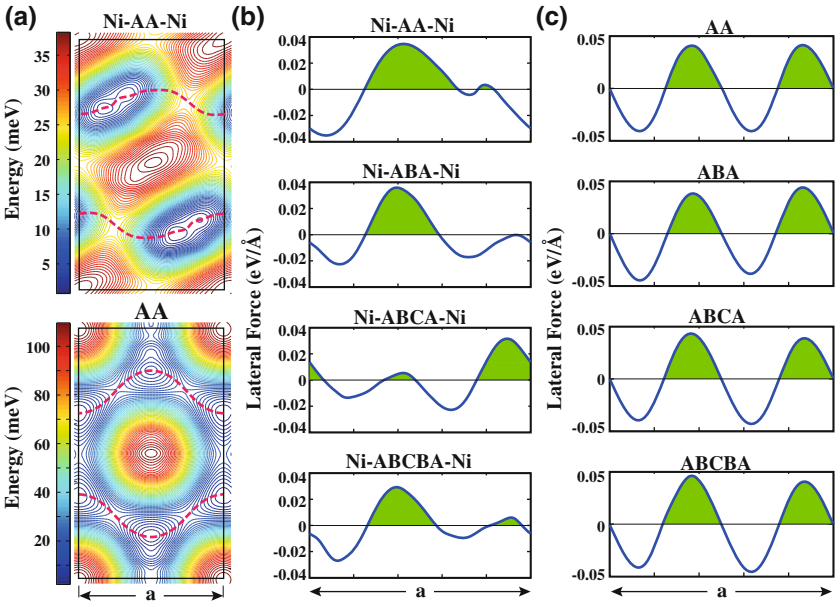


Fig. 21.11 **a** Profiles (contour plots) of potential corrugation for Ni-AA-Ni and AA [without Ni(111) slabs] structures calculated for constant pressure of 7 GPa. The paths along which one slab moves in the course of sliding when pulled along x axis are shown by *red dashed lines*. The lattice constant of the unit cell is indicated by a. **b** Variation of lateral force F_x along x-axis during sliding of Ni-AA-Ni structures over the path shown in (**a**). The sum of areas shaded in green is defined as the corrugation strength W_D (see text). **c** Same as (**b**) for sliding AA structures without Ni(111) slabs

minimum and maximum of energy) is an order of magnitude smaller compared to single-layer honeycomb structures of graphane CH, fluorographene CF, MoS_2 and WO_2 discussed in precious section. On the other hand, the intrinsic stiffness of the present case, which is related to the interaction between Ni and graphene layers is also substantially lower ($k_s = 0.8$ eV/A^2) compared to the intrinsic stiffness of those honeycomb structures [33]. The lower intrinsic stiffness accompanied by low potential corrugation curvature results in a frictional figure of merit of $\sim$10, at constant pressure of 7 GPa, which is enough to keep the system in continuous sliding regime. Comparing the profiles of the potential corrugation of Ni-AA-Ni and AA structures, one can see how the interaction between graphene layers is affected by their interaction with Ni surfaces. The effect of distortion presented in Fig. 21.9b is reflected to the potential corrugation of Ni-AA-Ni, since its symmetry is changed from hexagonal to rectangular. Also note that, the amplitude of the potential corrugation is substantially lowered when Ni slabs are present, which reveals an important effect of substrate (i.e. Ni(111) surface).

To set a measure for the corrugation strength we first derive the path at which the upper slab would slide if it was pulled along x-axis. This path is shown by dashed lines in Fig. 21.11a for the case of Ni-AA-Ni. In the case of structures having more than two graphene layers the path is found directly by starting from the Ni slab positions presented in Fig. 21.9 and moving along the x-axis while minimizing the total energy along y axis. Then we calculate the lateral force F_x along x-axis felt by the slab, as shown in Fig. 21.11b. Here we note that in the sliding of Ni(111) slabs having n graphene layers the dissipation of energy through non-equilibrium phonons generated by sudden processes is hindered for $n \geq 2$ and hence $W = \int_0^a F_x dx$ vanishes. This, however, does not precludes energy dissipation through other mechanisms. With a premise that the maximum of the energy to be dissipated by any mechanism should be smaller than $W_D = \int_0^a F_x^> dx$ i.e. the integral of all positive work done during sliding of one slab over one unitcell shown by the green shaded region in Fig. 21.11b, we took W_D as a measure for the corrugation strength. The result of these calculations are presented in Fig. 21.11b. Note that W_D (is also related to kinetic friction coefficient $\mu_k = (W_D/a)/F_z$) is already very small. To check the effect of the type of stacking we have also calculated the force variation for Ni-ABABA-Ni structure and the result was very close to that of Ni-ABCBA-Ni structure. For comparison, we have performed the same calculations for graphene layers in the same stacking but without Ni slabs above. The results of these calculations are presented in Fig. 21.11c.

Various important trends in the corrugation strength W_D, obtained from above calculations are presented in Fig. 21.12a. As expexted the corrugation strength increases with increasing normal force. Also the corrugation strength is higher in structures composed of only graphene layers (like ABA) compared to the ones having Ni slabs (like Ni-ABA-Ni). This effect is mirrored in the repulsive interaction of graphene layers in the presence and absence of Ni slabs, as shown in Fig. 21.12b. Here one can see that introducing Ni slabs decreases the repulsive interaction between graphene layers, which is consistent with decrease in the corrugation strength discussed above. Another important finding is that corrugation strength of the structures solely

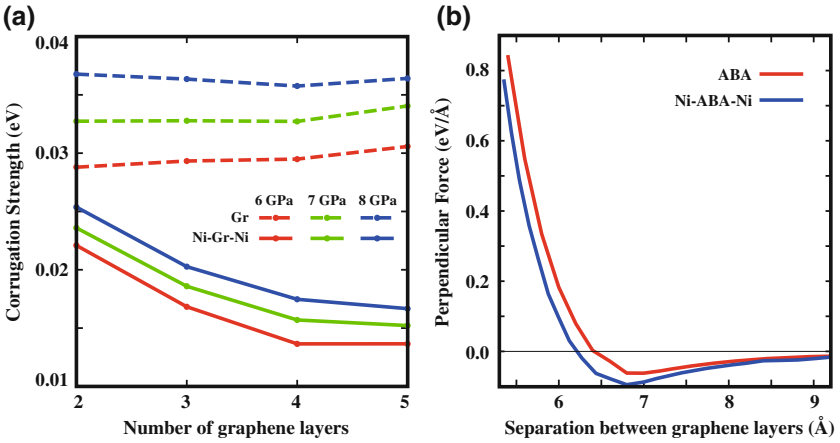


Fig. 21.12 **a** Variation of the corrugation strength with number of layers as a function of applied loading pressure for n number of graphene layers (with and without Ni(111) substrates). **b** Perpendicular force F_z versus the separation distance between outermost graphene layers for Ni-ABA-Ni and ABA structures ($n = 3$). In the repulsive range, the perpendicular force and hence the potential corrugation is larger in the absence of Ni(111) slabs

composed of graphene layers has minor variation with the number of layers. On the other hand, the corrugation strength is significantly decreased when the number of graphene layers in Ni-graphene-Ni structures are increased. This trend is seemingly in accordance with experimental observations [44, 46]. However, as mentioned above, the system at hand is very different from those considered in the experiments and the trends revealed by Fig. 21.12a heralds another important effect.

21.3.4 Analysis of Charge Density

To explain these trends we first examine the effect of Ni slabs on the electronic structure of graphene layers. The self-consistent difference charge density $\Delta\rho$, is obtained by subtracting the charge density of ABA structure and two Ni(111) slabs from that of Ni-ABC-Ni structure. The isosurfaces of $\Delta\rho$ and the variation of its value averaged over (x, y)-planes parallel to graphene layers (called linear density) are presented in Fig. 21.13. The major charge transfer is between Ni and graphene layers attached to each other as seen in top and middle panels of Fig. 21.13. The dangling Ni-d_{z^2} orbitals at the surface of the slab change their character upon coating of graphene layers. This is resulted in the charge depletion denoted by the numerals 1 and 3 in the linear charge density plot. Analysis of the band structure show significant contribution to C-p_z states from s, d_{xz} and d_{yz} orbitals of Ni atoms, while C-p_z orbitals by themselves contribute to d_{xy} and d_{x^2} states of Ni atoms. As a result of these complex mechanism of charge transfer the charge density around the graphene

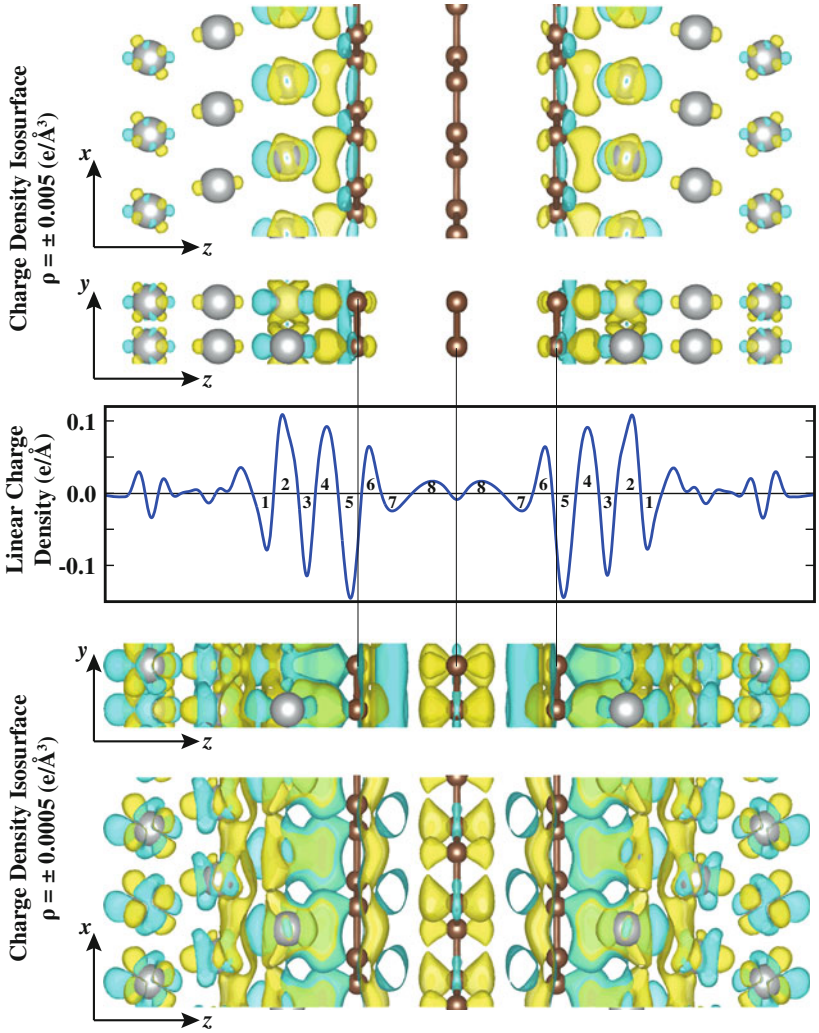


Fig. 21.13 Isosurfaces and variation of linear density of charge density difference along z-axis. The difference charge density is obtained by subtracting the charge densities of Ni slabs and ABA structures from the charge density of Ni-ABA-Ni structure at ~6 GPa. *Yellow* (*blue*) isosurface plots correspond to the charge density accumulation (depletion). Specific regions of depletion and accumulation is denoted by numerals on the linear density of charge density difference plot

layer is shifted towards Ni slab resulting in charge density accumulations (depletions) denoted by numerals 4 and 6 (5 and 7).

The charge density depletion denoted by numeral 7 in the linear density of charge difference may be the key feature to explain the decrease in the corrugation strength between graphene layers due to Ni slabs. The isosurface of charge depletion

corresponding to this region can be seen in the bottom panel of Fig. 21.13. This charge depletion lowers the chemical interaction between graphene layers and results in lowering of corrugation strength as seen in Figs. 21.11 and 21.12. Moreover, similar charge depletions are also observed in Ni-AB-Ni, Ni-ABCA-Ni and Ni-ABCBA-Ni structures and their amplitude asymptotically increases by going from two to five layers. This is in accordance with the decrease in the corrugation strength with increasing number of layers, shown in Fig. 21.12b.

In summary, we find that even in the present model, where graphene layers have negligible puckering, the corrugation strength is decreased upon coating of the sliding Ni surfaces and increasing the number of layers. This is attributed to a complex charge transfer between graphene layers and Ni(111) surfaces, each coated by these graphenes. This transfer results in charge depletion between graphene layers thereby decreasing the corrugation strength. In the absence of Ni slabs each coated by a graphene layer, the corrugation strength is relatively higher and practically independent of the number of graphene layers. Our results demonstrate that graphene attached to sliding surfaces operate as superlubricant by suppressing energy dissipation dramatically.

21.4 Discussions and Conclusions

In conclusion, using a criterion for the transition from stick-slip to dissipationless continuous sliding regime, which is calculated from the first-principles, we showed that a pair of sliding layer of the same nanostructures, such as pairs of CH, CF, MoS_2 and WO_2, execute continuous sliding with ultralow friction. The minute variation of the amplitude of the interaction potential due to the repulsive interaction, as well as stiff C–H(F), Mo–S and W–O bonds underlie the frictionless sliding predicted in the present study. Our predictions put forward an important field of application as ultralow friction coating for the layered honeycomb structures, which can be achieved easily to hinder energy dissipation and wear in sliding friction. Earlier, the sliding motion of the diamond like carbon coatings exposed to hydrogen plasma resulted in a very low friction coefficient [47]. Ultralow friction was attributed to repulsive Coulomb forces between DLC films facing each other in sliding. However, when exposed to open air in ambient conditions, positively charged H atoms was replaced by negatively charged O and hence the uniformity in the charging was destroyed. In the present study, graphane coating is reminiscent of the hydrogenated DLC and accordingly is found to have ultralow friction, but vulnerable to degradation by oxygen atoms. Unlike graphane and DLC coating, WO_2 coating consists of negatively charged oxygens and hence immune to oxidation.

We showed that even if the strong interaction between the sliding surfaces of Ni(111) is dramatically reduced by a single layer graphene placed in between, the bistability between approach and pull-off remains. Also the stick-slip motion still exists and continues to dissipate significant amount of mechanical energy. The stick-slip motion and hence the generation non-equilibrium phonons are eliminated with

the onset of continuous sliding, once each of metal surfaces in relative motion is coated by a single graphene layer. This is attributed to substantial interaction between Ni surface and graphene through complex charge exchange causing to the reduction of the chemical interaction between graphene layers and hence to the decrease of the corrugation strength. The corrugation strength continues to decrease gradually with increasing graphene layer and eventually saturates at a small value. In the absence of metal slabs each coated by a graphene layer, the corrugation strength is relatively higher and practically independent of the number of graphene layers. Our results demonstrate that graphene attached to sliding surfaces operate as superlubricant. One expects to achieve similar lubrication effect but in lesser degree by placing graphene flakes between sliding or rolling Ni(111) surfaces. The interaction between Ni(111) and graphene investigated in this study appears to be important not only for the growth of pristine graphene or for the protection from oxidation, but also for achieving the nearly frictionless friction. Easy growth of graphene on Ni(111) surfaces makes Ni also an attractive substrate for nanotribology applications. Our calculations showed that graphene layers placed between pairs of Al(111) and Cu(111) have the capacity of reducing adhesion and sliding friction. Recent tribological test results for a few layer graphene placed between sliding 440C steel surfaces revealed that wear is decreased by almost 4 orders of magnitude and friction coefficient is decreased by 5 orders of magnitude [48, 49]. These experimental results corroborate the conclusions of our present theoretical work.

Finally, we note that the first-principles calculations of potential corrugations calculated in the constant force mode are achieved by optimizing atomic structure. This way, the elastic deformations of sliding surfaces under perpendicular loading force are taken into account. We believe that this important feature of the present method will be used in future studies dealing with the development of lubricant single layer materials.

Acknowledgments This Chapter is partially based on the doctoral thesis work of S. Cahangirov at Bilkent University and the related research results were initially reported in Phys. Rev. Lett. **108**, 126103 (2012) and Phys. Rev. B. **87**, 205428 (2013). The authors thank C. Ataca, M. Topsakal, H. Şahin and Ongun Özçelik for their contributions to the theoretical research on sliding friction in our group at UNAM, National Nanotechnolgy Research Center at Bilkent University.

References

1. L. Prandtl, Z. Angew, Math. Mech. **8**, 85 (1928)
2. G.A. Tomlinson, Philos. Mag. **7**, 905 (1929)
3. B.N.J. Persson, *Sliding Friction: Physical Principles and Applications* (Springer, Berlin, 1998)
4. M. Urbakh, E. Meyer, Nature Mat. **9**, 8 (2010)
5. C.M. Mate, G.M. McClelland, R. Erlandsson, S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987)
6. D. Tomanék, W. Zhong, H. Thomas, Europhys. Lett. **15**, 887 (1991)
7. A. Buldum, S. Ciraci, Phys. Rev. B **55**, 2606 (1997)
8. M.H. Mueser, M. Urbakh, M.O. Robbins, Advances. Chem. Phys. **126**, 187 (2003)
9. V.L. Gurevich, *Transport in Phonon Systems* (North-Holland, Amsterdam, 1986)

10. A. Buldum, D.M. Leitner, S. Ciraci, Phys. Rev. B **59**, 16042 (1999)
11. H. Sevincli, S. Mukhopadhyay, R.T. Senger, S. Ciraci, Phys. Rev. B **76**, 205430 (2007)
12. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, H.-J. Gntherodt, Phys. Rev. Lett. **84**, 1172 (2000)
13. A. Socoliuc, R. Bennewitz, E. Gnecco, E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004)
14. D.C. Elias, R.R. Nair, T.M.G. Mohiuddin, S.V. Morozov, P. Blake, M.P. Halsall, A.C. Ferrari, D.W. Boukhvalov, M.I. Katsnelson, A.K. Geim, K.S. Novoselov, Science **323**, 610 (2009)
15. H. Şahin, C. Ataca, S. Ciraci, Phys. Rev. B **81**, 205417 (2010)
16. R.R. Nair, W. Ren, R. Jalil, I. Riaz, V.G. Kravets, L. Britnell, P. Blake, F. Schedin, A.S. Mayorov, S. Yuan, M.I. Katsnelson, H.-M. Cheng, W. Strupinski, L.G. Bulusheva, A.V. Okotrub, I.V. Grigorieva, A.N. Grigorenko, K.S. Novoselov, A.K. Geim, Small **6**, 2877 (2010)
17. H. Şahin, M. Topsakal, S. Ciraci, Phys. Rev. B **83**, 115432 (2011)
18. C. Ataca, M. Topsakal, E. Aktürk, S. Ciraci, J. Phys. Chem. C **115**, 16354 (2011)
19. C. Ataca, H. Sahin, E. Aktürk, S. Ciraci, J. Phys. Chem. C **116**, 8983 (2011)
20. P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964)
21. W. Kohn, L.J. Sham, Phys. Rev. **140**, A1133 (1965)
22. M. Topsakal, S. Cahangirov, S. Ciraci, App. Phys. Lett. **96**, 091912 (2010)
23. S. Miyake, R. Kaneko, Y. Kikuya, I. Sugimoto, J. Tribol. **113**, 384 (1991)
24. P. Thomas, K. Delbe, D. Himmel, J.L. Mansot, F. Cadore, K. Guerin, M. Dubois, C. Delabarre, A. Hamwi, J. Phys. Chem. Solids **67**, 1095 (2006)
25. J.M. Martin, C. Donnet, Th. Le Mogne, Th. Epicier, Phys. Rev. B **48**, 10583 (1993)
26. T. Liang, W.G. Sawyer, S.S. Perry, S.B. Sinnott, S.R. Phillpot, Phys. Rev. B **77**, 104105 (2008)
27. S. Chen, L. Brown, M. Levendorf, W. Cai, S.-Y. Ju, J. Edgeworth, X. Li, C.W. Magnuson, A. Velamakanni, R.D. Piner, J. Kang, J. Park, R.S. Ruoff, ACS Nano **5**, 1321 (2011)
28. J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996)
29. S. Grimme, J. Comp. Chem. **27**, 1787 (2006)
30. P.E. Blochl, Phys. Rev. B **50**, 17953 (1994)
31. G. Kresse, J. Hafner, Phys. Rev. B **47**, 558 (1993)
32. G. Kresse, J. Furthmuller, Phys. Rev. B **54**, 11169 (1996)
33. S. Cahangirov, C. Ataca, M. Topsakal, H. Şahin, S. Ciraci, Phys. Rev. Lett. **108**, 126103 (2012)
34. S. Cahangirov, S. Ciraci, V. Ongun, Özçelik. Phys. Rev. B **87**, 205428 (2013)
35. L.C. Lew, Yan Voon, E. Sandberg, R. S. Aga, A. A. Farajian. Appl. Phys. Lett. **97**, 163114 (2010)
36. M. Houssa, E. Scalise, K. Sankaran, G. Pourtois, V.V. Afanas'ev, A. Stesmans, Appl. Phys. Lett. **98**, 223107 (2011)
37. S. Cahangirov, E. Aktürk, M. Topsakal, H. Şahin, S. Ciraci, Phys. Rev. Lett. **102**, 236804 (2009)
38. M. Dienwiebel, G.S. Verhoeven, N. Pradeep, J.W.M. Frenken, Phys. Rev. Lett. **92**, 126101 (2004)
39. J.S. Choi, J.S. Kim, I.S. Byun, D.H. Lee, M.J. Lee, B.H. Park, C. Lee, D. Yoon, H. Cheong, K.H. Lee, Y.W. Son, J.Y. Park, M. Salmeron, Science **333**, 607 (2011)
40. A.E. Filippov, M. Dienwiebel, J.W.M. Frenken, J. Klafter, M. Urbakh, Phys. Rev. Lett. **100**, 046102 (2008)
41. A.S. de Wijn, C. Fusco, A. Fasolino, Phys. Rev. E **81**, 046105 (2010)
42. I.V. Lebedeva, A.A. Knizhnik, A.M. Popov, O.V. Ershova, Y.E. Lozovik, B.V. Potapkin, Phys. Rev. B **82**, 155460 (2010)
43. A.M. Popov, I.V. Lebedeva, A.A. Knizhnik, Y.E. Lozovik, B.V. Potapkin, Phys. Rev. B **84**, 045404 (2011)
44. H. Lee, N. Lee, Y. Seo, J. Eom, S.W. Lee, Nanotechnology **20**, 325701 (2009)
45. T. Filleter, J.L. McChesney, A. Bostwick, E. Rotenberg, K.V. Emtsev, Th. Seyller, K. Horn, R. Bennewitz, Phys. Rev. Lett. **102**, 086102 (2009)
46. C. Lee, Q. Li, W. Kalb, X.Z. Liu, H. Berger, R.W. Carpick, J. Hone, Science **328**, 76 (2010)
47. A. Erdemir, Surf. Coat. Technol. **146**, 292 (2001)
48. D. Berman, A. Erdemir, A.V. Sumant, Carbon **54**, 454 (2013)
49. D. Berman, A. Erdemir, A.V. Sumant, Carbon **59**, 167 (2013)

Chapter 22
Nanoscale Friction of Self-assembled Monolayers

Karine Mougin and Haidara Hamidou

Abstract Self-assembled monolayers (SAMs) have played and still continue playing a major role in the control of the interface behaviors, especially in fluidic and as lubrication coatings. This chapter introduces and focuses on frictional properties of molecular homogeneous thin films; particularly the influence of chain length, terminal group and packing state of SAMs on the one hand, and that of the environmental and experimental conditions on the other hand, on energy dissipation during friction. Then, we will address the important issue of nanoscale heterogeneities and their influence on the frictional behavior and nanotribological performance of SAMs, using binary molecular films of both regular and random distribution of the heterogeneities. Because real surfaces are often heterogeneous in their chemical composition, these nanoscale heterogeneous SAMs provide good model systems to study and tune the frictional properties of tribological coatings.

The rapid development of the microelectromechanical industry (MEMs) has stimulated the need and design of molecular assemblies capable to impart new functional surface properties to materials or improved those existing, for the control of their interface behaviors (adsorption, condensation, wetting, lubrication, adhesion, liquid flow). Self-assembled monolayers (SAMs) of either organosilanes or thiols have played and still continue playing a major role in the achievement of this goal, especially in fluidic and as lubrication coatings. For MEMs devices, for instance, the control of adhesion and friction between the moving surfaces is crucial. Indeed the surface-to-volume ratio in these micro and nanoscale devices are so large that adhesive and friction forces dominate gravitation and inertia. Precise control of the interfacial interactions is thus required to adjust adhesion and friction and prevent wear and stiction. Thus, the basis for molecular design and tailoring of SAMs must include a complete knowledge of interrelationships between the molecular structure

K. Mougin (✉) · H. Hamidou
Institut de Science des Matériaux de Mulhouse, 15, Rue Jean Starcky,
BP 2488, 68057 Mulhouse, France
e-mail: karine.mougin@uha.fr

H. Hamidou
e-mail: hamidou.haidara@uha.fr

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_22

and frictional properties of SAMs, as well as a understanding of friction and wear mechanisms of SAMs at a molecular level.

Until recently, tribological studies have not access to the molecular level detail of interfacial contact. As a consequence, experimental verification of the molecular factors influencing frictional properties has been lacking. Fortunately, development of Atomic Force Microscopy has provided the opportunity to study friction at a nanometer scale, so-called Lateral Force Microscopy (LFM) or Friction Force Microscopy (FFM). LFM provides some images of patterned SAMs but also provides some response for frictional, shear and adhesion properties of monolayers [1].

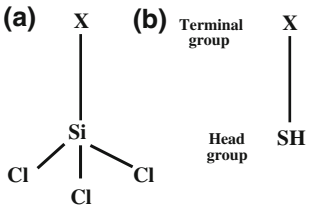
New interdisciplinary ideas and approaches can be effective and physical–chemical techniques based on self-assembling and self-organization principles are useful for fabrication of new nanomaterials and nanostructures [2]. Indeed, SAMs have a great importance in interfacial engineering both for their fundamental and practical interests. For fundamental researches, they provide model systems for studying organized functional molecules and for applications [3], they often serve as template for the bottom-up fabrication of nanostructures Indeed, the nature (polarity, hydrophilicity, …) as well as the structure of an homogeneous SAM coating on a surface has a strong influence on the frictional properties (tribology, drag force to the liquid flow along a wall). Most of the time, SAMs were self-assembled onto smooth substrates (e.g. mica, silicon wafer, …) to minimize contributions to the measurements from surface topography.

The physical state of thin molecular films can be influenced by a number of forces including intermolecular interactions between molecules making up the film, surface interactions in adsorbed films [4]. The structural characteristics of organic thin films (molecular conformation, packing arrangement and chemical composition) directly influence their lubrication properties.

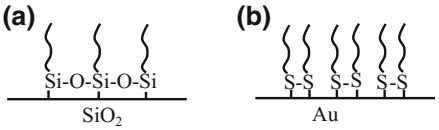
The report is divided in two main parts. The first paragraph will focus on frictional properties of molecular homogeneous thin films; particularly the influence of chain length, terminal group and packing state of SAMs on the one hand, and that of the environmental and experimental conditions on the other hand, on energy dissipation during friction. Then, we will address the important issue of nanoscale heterogeneities and their influence on the frictional behavior and nanotribological performance of SAMs, using binary molecular films of both regular and random distribution of the heterogeneities. Because real surfaces are often heterogeneous in their chemical composition, these nanoscale heterogeneous SAMs provide good model systems to study and tune the frictional properties of tribological coatings.

22.1 Homogeneous Organic Molecular Films

The formation of organized monolayer films on a surface by spontaneous adsorption and covalent grafting of molecules from solution or vapor phase is known as self-assembly. Two kind of organic molecules have been extensively grafted: alkylsilanes



Scheme 22.1 Schematic representation of an **a** alkylsilane, and **b** organosulfur molecule



Scheme 22.2 Schematic representation of **a** siloxane network; **b** dimmer thiols

ontos silicon wafer or glass surfaces and organosulfur onto golden coated substrates (Scheme 22.1).

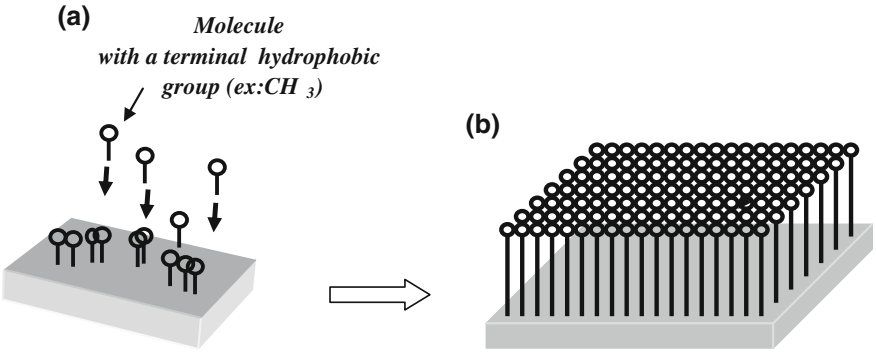
On golden substrate, the anchorage of the organosulfur molecule is provided by a covalent bond between the sulfur end group and the gold atom. On silicon surface, the attachment is less well-defined as previously. From a common point of view, the grafting of the organosilane involves the hydrolysis of the chloro or alkyloxysilyl headgroups with the residual molecular film of water strongly adsorbed onto the silicon substrate (Scheme 22.2).

The silanols groups ensued from this reaction can either form a covalent bond with the silanol group arised from the silicon surface [5] or create hydrogen interaction with trace water physisorbed to the substrate [6–8]. The stable linkage of the SAMs to the substrate is provided by a cross-linking process of the silanol groups to form a siloxane network [9] on the extreme layer of the substrate.

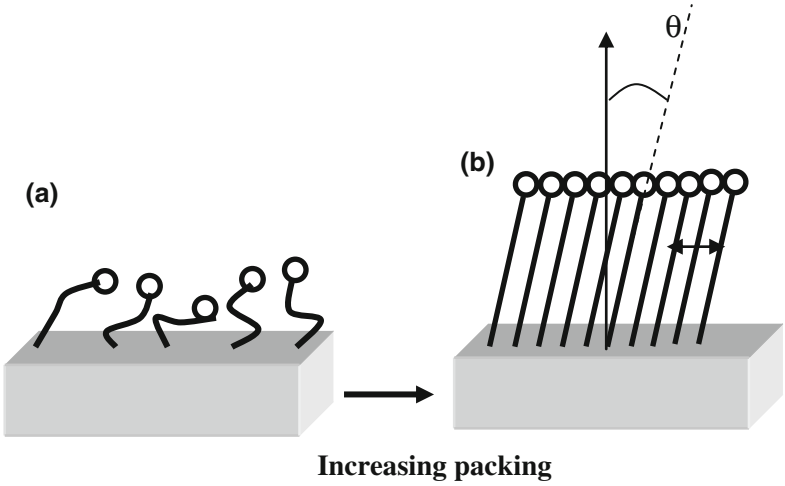
The performance of the self-assembled monolayers is directly dependent on the efficiency of the anchorage to the surface. Granick et al. [10, 11] have shown the importance of surface and density of grafting sites to form well-packed monolayers of SAMs.

Two procedures which have been successful for forming SAMs are available. On the one hand, the organic molecules are dissolved into a solvent, and on the other hand, they are dispersed in a paraffin oil and deposited under vaccum ($\sim 5 \times 10^{-3}$ Torr, 45 min to 1 h). For both processes organic molecules diffuse to the surface and self-assemble into discrete nanodomains [12] which will grow up to the complete confluence to form a thin and continuous organic layer as depicted in Scheme 22.3.

The surface coverage and structure of the resulting SAMs depend on several experimental parameters such as the reaction time, the temperature, the hydration state of the substrate [8, 13], the nature and polarity of the solvent, or the pressure (vaccum) for vapor phase deposition. Particularly, SAMs pioneers, Sagiv et al. [14] have shown the correlation which exists between the adequacy of the solvent and the packing density of thin molecular film. These experimental parameters as well as more intrinsic ones have a strong influence on the structure of the monolayer. Indeed,



Scheme 22.3 Schematic representation of the formation of a self-assembled monolayer. **a** Nucleation and growth of molecular film in discrete nanodomains **b** Elaboration of an organic thin film



Scheme 22.4 Schematic representation of a solid-like structure **a** a liquid-like structure **b** of a thin molecular film

the structural state of the thin film is mainly defined by its uniformity and packing density. The following scheme presents a well-ordered SAM structure referred to as solid-like structure and a less ordered one referred to as liquid-like structure (Scheme 22.4). As shown by Gerber et al. [15], the organic molecules in a well-packed thin film remain in a specific conformation: tilted a few degrees from the surface normal and twisted around their molecular axis to minimize their global free energy at equilibrium.

The possibility of changing the chain length, terminal group and the packing order (disorder) within the molecular film makes self-assembled monolayers attractive model systems to study the nature of frictional interactions at the molecular level.

Along with the intrinsic features of the molecular films, the environment and operating conditions (both experimental and in-service) represent the second class of parameters that can drastically affect the friction, either directly or through the structural changes and response of the molecular film.

The sensitivity of friction to the operating conditions on SAMs arise from the high susceptibility of these molecular films to structural changes which strongly affect the intermolecular surface forces, and hence the threshold lateral force required to slide the contact between two surfaces usually called fiction force F_f is given by the following equation [16],

$$\mathbf{F_f} \sim \boldsymbol{\mu}(\mathbf{F_{ext}} + \mathbf{F_{adh}} + \mathbf{F_c}) \tag{22.1}$$

where and F_f represent the friction force, $\mathbf{F}_{ext}$ corresponds to the applied external load, $\mathbf{F_{adh}}$ is the intermolecular adhesion force, F_c is the capillary condensation force when its exists (22.2) and, $\boldsymbol{\mu}$ is the friction coefficient between surfaces. Obviously, the $\boldsymbol{\mu}$ determined under the action of the sole intermolecular adhesive forces may differ, more or less, from that determined under the combined effect of $(F_c + F_{adh})$, or $(F_{ext} + F_{adh})$ above a certain value of F_{ext}. It also should be noted that even in the absence of any external load, the only adhesive forces can be critical as regards nanotribological performance of micromachines (MEMs, NEMs). This results from the space confinement and the high surface area-to-volume ratio in these systems, to which the adhesion force that tends to stick the different elements is proportional. For vdW interactions between two planar bodies separated by a distance h, for instance, these attracting surface forces scale as $(H/6\,\pi\,h^3)$x *Contact Area*, where H is the Hamaker constant of the interface.

The following section gives a synthetic but critical discussion of the way these chemical, topological and operating parameters affect and control the frictional behaviors of self-assembled molecular films.

22.1.1 Influence of Chain Length and Structure

The films with well-ordered structure exhibited lower friction coefficient than similar films with disordered structure [17, 18]. Contact between tip and sample under a low load is nearly frictionless [19]. Friction is particularly high with short chains of less than eight carbons. Longer chains, stabilized by Van der Waals (vdW) attractions form more compact and rigid layers. This rigidity even results in a much better lubrication [18] effect on frictionless SAMs such as CF_3, CH_3 terminated films. Indeed, friction depends on the intrinsic properties of the grafted organic molecules: on the one hand, the chemical nature of its end-group which is in direct contact with the counterface and exchange molecular interactions, and on the other hand the density and thickness of the monolayer [20]. However, this lubricating action is lost from a certain external load threshold when wear-induced disorder at the molecular level start to occur. In particular, Porter et al. have shown the existence of a bimodal dependence of

friction coefficient versus n, where n represents the number of methylene units for n-alkanethiolate monolayers. The tribological transition is around n $\sim$ 12.

The frictional energy dissipation mechanisms in SAMs has been investigated by different authors and shown to arise from the excitation of the rotational and vibrational modes in the molecules These excitation modes contribute strongly to the energy transfer to the substrate and thus to the frictional properties. Since short chains are shown to be less-well packed than long chain SAMs, they have structural disorder an defects that promote the excitation modes (rotational, vibrationalldots) and energy absorption (dissipation), giving rise to a higher friction. The self-organization and close packing of the molecules ensured by Van der Waals interactions between chains play an essential role in stabilizing thin molecular films and lowering friction. The longer the chains are, the greater the cohesive interactions between chains are. As a result, the frictional behavior of well-packed long chains SAMs might be explained by the strong intermolecular interactions that hold large blocks of molecules together during shear without an effective loss of molecular order, according to Eyring model [21]. The dissipation mechanism in this case appears to involve the motion of molecular domains. Thus, these blocks that effectively tend to retain the molecular order, globally move with respect to the substrate, resisting the shear induced by the friction and minimizing the relative volume change.

It is then reasonable to suggest that the increase in friction for the short chains ($n < 8$) is due to poor packing of the molecules, which makes possible the excitation of numerous defects and energy-dissipating modes [17]. The excitation of these modes is similar to viscoelastic behavior. The fluid-like chains are more compliant and present a smaller resistance to shear.

As a results, the lack of cohesion in thin films formed with short chains, involve a different frictional process but, sliding correspond more to a conformational change of the molecules in the film structure. This process is more energy dissipative due to an inelastic and plastic deformation of surface asperities [22] at a macroscopic scale, but at a microscopic level, energy is dissipated through surface phonons. The molecular dynamics study argues that energy dissipation in a long chain monolayer is associated with vibrational energy through an oscillation around the tilt angle of the chains. The shear-induced disorder of the chains in the shorter monolayers therefore results in a not permanent contact because of energy dissipation through bond rotations and vibrations, leading to a higher microscopic friction. So far, these investigations are still confused as additional information are required to provide quantitative tribological results and correlate friction and energy dissipation in organic thin films.

Moreover, the influence of the preparation method and of the underlying substrate's nature on SAMs friction still remain unclear to date. According to Lio et al. studies [17], the lack of the long range order in the silanes for length chain below n $<$ 10, can be explained by cross-linking of the head groups (Scheme 22.1). In thiols, cross-linking is limited to dimmer formation, while for silanes, there is an extensive siloxane network along the film/substrate interface [9]. Hence, there must be considerable chain distortions near the interface that is amplified along the network, which may be at the origin of the lack of long-range order in organosilane SAMs. This distortion effect is even more amplified for short chains films, leading

to higher disorder and friction coefficient. Finally, Bhushan et al., have proposed to use a C_{18} double grafted films with a higher area density and hence a stronger chain-chain interactions leading to better nanotribological behavior [23].

Other important questions which need to be answered regarding the frictional behavior of SAMs include: the effect of the lateral order within the film, the presence of heteroatoms (O, N, S…) or branches inside the molecular chain and the existence of double or triple bonds that change the rigidity of the molecules.

22.1.2 Influence of Terminal Group

The influence of the terminal and head groups (Scheme 22.1) on friction and wear properties were also investigated by contact mode atomic force microscopy [24].

Ahn et al. have shown that at nanoscale, the frictional behavior of the homogeneous SAMs with different functionality is primarily influenced by the surface energy, which is directly related to the interfacial interactions between the tip and the topmost surface chemistry of the substrate. When using a polar probe as the contacting surface (clean Si_3N_4 AFM tip for instance), friction increases from hydrophobic to hydrophilic homogeneous thin films.

Amongst these nanoscale coating films, fluorocarbon-based SAMs have received a great deal of attention, especially for their singular surface energy (wetting) versus friction behavior [25]. In standard organic chemistry textbooks, the sections devoted to fluorocarbon chemistry are typically brief. Nevertheless, this molecule seems to be very interesting as the size of fluorine [26] allows it to replace hydrogen in many organic molecules. In contrast to hydrocarbon or siloxane coatings, fluorocarbon coatings exhibit repellency toward both water and oil, and yet, frictional investigations on perfluorocarbon-based SAMs have shown that they exhibit a higher friction coefficient than simple hydrocarbon-based SAMs [27, 28]. This singularity can be explained by the real difference of their both lattices compacity. As mentioned previously, Zisman and Timmons [29] have shown that fluorocarbon-based SAMs are less well-packed than hydrocarbon ones. As a result, the space void in a fluorocarbon lattice is larger than in a hydrocarbon one. Indeed, Stoebe et al. [30] have estimated, according to atomic structural data, that CF_3 group occupied 25 % more volume than a regular CH_3 end group. This simple remark shows the real discrepancy of packing density of both lattices. Larger terminal groups in films give rise to increased steric interactions that provide pathways for energy dissipation during sliding. And hence, lateral (and perhaps rotational) motion within the plane of the CF_3 groups is highly cooperative over relatively large distances, involving long-range interactions in the CF_3 terminated films [31]. These long-range interactions might permit the dissipation of energy through phonon modes [32] and thus, induce a higher frictional response for the CF_3-terminated films. These singular nanotribological responses of fluorinated-based SAMs are in total agreement with their wetting properties. Hysteresis of the contact angles of liquids over a fluorinated thin film is actually higher than on a regular hydrocarbon one, given the fact that liquid molecules can penetrate

more easily in a CF_3 self-assembled monolayer, less well packed than a CH_3 one [33]. Indeed, wetting theory specifies that hysteresis of contact angle of a liquid onto a SAMs coated substrate is proportional to the molecular volume of the thin film.

These observations, however, can be rationalized based on earlier models by Yoshizawa et al. [34] assuming (i) that molecular reorientations in the SAMs are included in the definition of interdigitation and (ii) that SAMs of thiols on gold can be classified as 'solid-like' or 'amorphous-like', respectively, regarding their tribological properties.

Finally, Overney et al. [35] have observed that fluorocarbon films were more elastic and displayed a higher friction than hydrocarbon films. They have proposed a correlation between friction and elasticity of organic thin films. Garcia-Parajo, as well as Bhushan et al. have also observed the compression and relaxation of soft and rigid thin film in their loading and unloading tests [36, 37]. At a given normal load, long carbon chain structure such as alkyl chain can be easily compressed with the tip compared to rigid benzene ring structure SAMs. The orientation of the 'molecular springs or brush' under normal load reduces the shearing force at the interface, which in turn reduces the friction force. The possibility of orientation is determined by the spring constant of a single molecule (local stiffness), as well as the interaction between the neighboring molecules, which can be reflected by packing density or packing energy. It should be noted that the orientation can lead to conformational defects along the molecular chains, which lead to energy dissipation.

The previous paragraphs have described the impact of intrinsic structural parameters of the homogeneous molecular film onto its frictional response; and yet, other experimental parameters are also able to drastically modify this behavior. These are especially the environmental conditions and sliding velocity, which we will discuss in the following sections.

22.1.3 Effect of Humidity and Temperature

The environment is a crucial parameter of tribological experiments. The frictional results are directly dependent of humidity and temperature of the surrounding medium. The frictional behaviors of molecular organic thin films relied on the environmental parameters and can completely be reversed if they change drastically. Moreover, humidity and temperature have a direct impact on the structural stability and performance of the SAM coatings. However, only a few studies exist concerning the influence of temperature and water on the monolayers after their formation.

These effects of humidity and temperature on the frictional properties of SAMs were investigated by different authors. The studies of Tian et al. [38] on bare and SAMs coated mica show that the friction decreases on the mica with increasing relative humidity, while increasing on long alkylsilanes SAMs coated mica, when operating at room temperature with low external loads. At low humidity (RH $< 5\,\%$), heating the samples in the temperature range of 20–80 °C was found to induce a negligible effect for both bare mica and SAMs coated mica. On the other hand, for the

same range of temperature but with a high humidity rate (RH > 50 %), a significant variation in friction with temperature was observed. Finally, the influence of humidity in these frictional problems is intrinsically related to the capillary condensation of water that bridges the contacting surfaces. Indeed, the capillary bridging force is given at first order by [39]:

$$\mathbf{F_c} \sim 2\pi \mathbf{R_T} \boldsymbol{\gamma_L} (\mathbf{cos}\boldsymbol{\theta_{SL}} + \mathbf{cos}\boldsymbol{\theta_{TL}}) \tag{22.2}$$

where R_T is the tip radius, θ_{SL} and θ_{TL} are respectively the static contact angles of the liquid on substrate and tip, and γ_L is the liquid (water) surface tension (Scheme 22.5a).

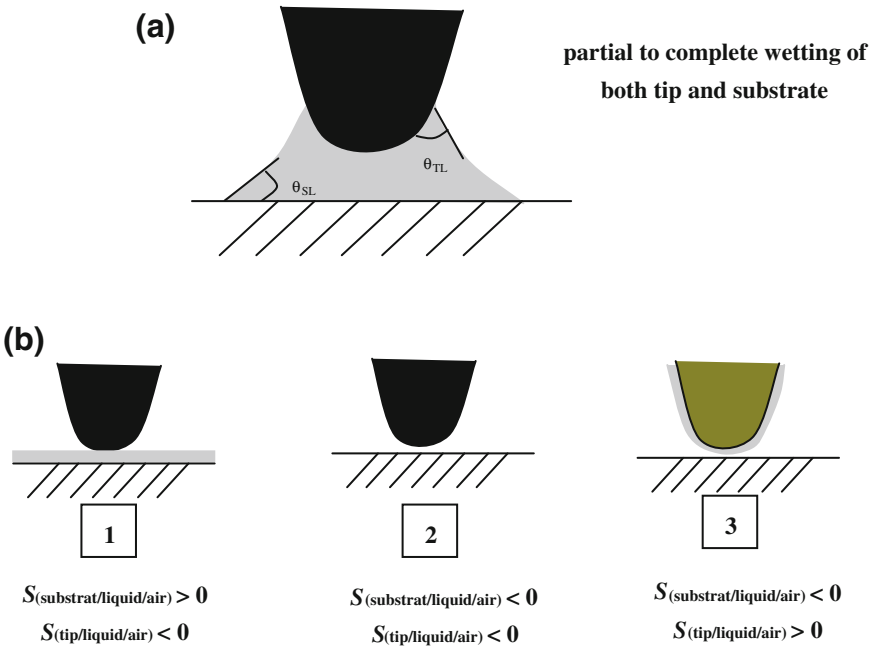
The resulting water meniscus or layer can either enhance friction through increased adhesion in the contact zone (22.1 and 22.2) or reduce it through the lubricating effect of water [40] (Scheme 22.5b).

The existence and local shape of the liquid condensate around the tip/substrate contact depend on the spreading coefficient S of the system [41], $S_{(\text{solid/liquid/air})} = \gamma_S - \gamma_{SL} - \gamma_{LV}$, where "solid" stands for the tip or substrate, γ_S, γ_{SL} and γ_L representing, respectively, the surface (interface) energy of the bare solid, the solid–liquid and liquid condensate. This physical parameter represents the variation of the surface energy by unit area between the dry and wet solid substrate. It is sensitive to both short- and long-range intermolecular interactions between the solid and the liquid in a given environment. A negative value of S indicates that the surface is not wetted by the liquid, and a positive one, the opposite. For the former case, several wetting situations displayed in Scheme 22.5b, can be investigated as a function of $S_{(\text{tip/liquid})}$ and $S_{(\text{substrate/liquid})}$.

As a result, in a humid environment, the magnitude of friction force is strongly dependent of the capillary force which is related to the wetting intrinsic properties of the interfacial system.

Cohen et al. [42] and Brock et al. [43] reported some of the thermally-induced structural changes of self-assembled monolayers. In those experiments, the monolayers were studied up to 200 °C, and no permanent changes in the structure were observed after the monolayers were returned to room temperature.

In addition, it was shown that even for hydrophobic SAMs, water could penetrate the films, altering their structural order and debonding the molecules from the substrate [37]. It should be noted that this process of "ageing" is strongly activated by the temperature and the mechanical action of the friction. As a result, the tribological properties are affected by these phenomena. While heating SAMs coatings at low humidity rate results in no evident changes in their frictional behavior, increasing humidity was shown to generally induce significant structural changes for the same heating conditions. Most of these structural changes of SAMs under the coupled effect of temperature and moisture (T, RH) can be understood and accounted for by the thermal activation of the bonds scission activity of water at SAM/substrate interface.



Scheme 22.5 **a** Schematic representation of a thick water bridge connecting the tip and the substrate (wetted hydrophilic SAM and tip, for instance), **b** Panel of other possible wetting situations showing *1* a thin continuous water film forms only on the hydrophilic substrate, at low RH % for instance; *2* there is no water condensation film on the hydrophobic contacting surfaces bridging ; *3* symmetric case of (*1*), the continuous condensation and lubricating water film forms only on the hydrophilic tip

22.1.4 Influence of Sliding Velocity

The influence of the sliding velocity on friction which accounts, at least partly, for the dynamical response of the boundary layer can be exploited to gain insight on the structure of the investigated monolayers. Because this dynamical response is strongly sensitive to the structure of the topmost surface layer and experimental conditions, the velocity-dependent investigations of friction have often led to contradictory results and discrepancies, even for substrates that were a priori "identical". For instance, molecular films of the same nature, but different structural features (packing, mobility relative to the substrate…) like SAMs and supported Langmuir-Blodgett films won't have the same frictional response to velocity under identical experimental conditions. The same observation may hold for SAMs of identical chemical nature but different packing. For homogeneous SAMs that have compliant long carbon spacer chains, the friction force increases at high (100 μm/s) velocity [24] while for SAMs that have rigid biphenyl chains, the friction force changes in the opposite way. The mechanisms responsible for the variation of the friction forces of

SAMs with velocity are believed to be related to the viscoelastic properties of SAMs. Generally [39, 40, 44], in wearless friction, the nanoscale frictional force (F_f) was found to depend logarithmically on the sliding velocity (v) according to

$$\mathbf{F_f} \sim \mathbf{F_{v=1}} + \boldsymbol{\beta} \mathbf{ln}(\mathbf{v}) \tag{22.3}$$

However, the sign of the slope β (increasing or decreasing F_f with ln**v**) is found to strongly depend on both the velocity range, the humidity and the relative hydrophilicity of the contacting surfaces [39, 45]. These different and apparently ambiguous variations of friction with the sliding velocity, humidity and substrate hydrophilicity are quite meaningful and well accounted for by the time-dependence (contact time $\tau \sim \mathbf{v}^{-1}$) of the magnitude of the capillary adhesive force F_c, which contributes to friction. On hydrophilic substrates and for a given RH %, large contact times (low **v**) allow larger number density and size of liquid bridges at tip/substrate contact, increasing the contribution of F_c to F_f. Since τ decreases in kinetic friction with increasing **v**, one should expect the slope β in 22.3 to be negative on a hydrophilic tip/substrate contact, and positive otherwise [39]: fully hydrophobic and hydrophobic/hydrophilic tip/substrate contacts. Unfortunately, experience shows that the velocity-dependence of friction in these nanoscale contacts are still more complex than what one could predict based on the sole contribution and contact time dependence of capillary liquid bridge adhesive force. Indeed, the thermal energy produced within the tip/substrate contact can induce molecular excitations and structural transitions in the topmost contacting layers (in SAMs for instance), the magnitude of which also increases with the sliding velocity. This velocity dependent conformation and morphological changes can thus affect through the alteration of the energy dissipation and intermolecular interactions the overall frictional force in the nanocontact, regardless of the existence of capillary condensation. An illustration of this complex velocity-dependent behavior of nanoscale friction will be provided and discussed in the section dedicated to "heterogeneous molecular SAMs".

22.1.5 Conclusion

The study of the nanoscale friction of monolayer films is an important issue in the world of nanotechnology as it controls the stability under operation, the performance and reliability of the electromechanical microsystems. Control over the intrinsic properties of the coating (chemical and structural ones) and of the environmental operating conditions such as humidity and temperature are the essential leading parameters for the choice and design of SAM layers useful in nanotribilogical applications. However, a unified and comprehensive picture of the energy dissipation process during friction is still missing, especially at the submicron scale. One obvious reason is that one still has a limited access to what really takes place within the contact during sliding, making quantitative analysis and interpretation rather difficult. The second reason is naturally related to the non-fully understood size effects which show up in nanoscale friction and strongly affect the results.

22.2 Molecular Heterogeneous Thin Films

Nanoscale surface patterns are considered as potential templates and building blocks for nanotechnology [46]. As for nanomaterials in general, these nanoscale surface structures have been of increasing research interest in recent years, due to their unique properties. They are expected to exhibit novel and significantly improved physical, chemical, mechanical and other properties, as well as to offer opportunities for manifestation of new phenomena and processes [47] which, owing to the nanoscale dimensions, are not observed at the macroscopic level. Precise control of nanopatterns is essential to assemble complex two-or three-dimensional structures [48]. As a result, the need for high-precision processes and nonconventional methods of surface patterning has steadily increased. Different techniques such as electrochemical stripping [49], hot embossing lithography [50], nano-imprint lithography [51, 52], edge transfer lithography [53], laser patterning monolayers process [54], dip-pen nanolithography [55, 56] as well as photolithography techniques [57] are commonly used to fabricate accurate and geometrically controlled micro and nanostructures.

However, although these techniques make it possible to produce a variety of patterns, those are often limited to well-defined regular geometries (square, circle, strips), with typical sizes still higher than 100 nm, except for a few time-consuming and rather expensive techniques. And yet, random patterns of nanoscale heterogeneous surfaces either composed of the discrete distribution of nanoscale domains of one molecular compound in the continuous phase of the second [58] or characterized by the formation of a bicontinuous structure of the two molecular phases [59, 60] represent another way to nanostructure surfaces.

The fundamental issue of this variety of nanopatterning is the way the chemistry, topology and surface fraction can affect or be used to adjust nanoscale frictional properties of these thin molecular films. These nano-heterogeneous surfaces also provide model systems for the understanding on the fundamental level recurrent interface phenomena and processes (wetting, de-wetting, nucleation, etc.).

This section will focus onto the influence of the distribution (geometrical controlled or its absence) of nanopatterns onto the frictional behavior of the substrates. As previously, the frictional responses of these heterogeneous molecular thin films will be studied both versus the intrinsic properties of the thin coating and operating parameters.

22.2.1 Influence of Topology

The tribological behavior of micro and nanostructured organic molecular thin films were investigated and interpreted as a function of the topology and organization of the surface patterns: from geometrical and well-organized structures to more random ones. So far, not so many information are available in the literature on this influence of the topological organization. This domain will require further model investigations.

22.2.1.1 Geometrical Structured Patterns

Nanotribological properties of engineered substrates and especially, geometrical patterned surfaces, were studied versus size and shape of the structures. Micro and nanopatterned surfaces were prepared (Fig. 22.1) by microcontact printing [61–63], micromolding in capillaries and nanoimprinting pioneered by Whitesides and Chou [64, 65].

Microcontact printing (μCP) involves the use of an elastomeric stamp with micron or submicron sized relief features on its surface to print molecular inks onto a substrate that has a higher affinity towards the ink molecules than the stamp. The most popular elastomer used is poly(dimethylsiloxane) (PDMS) and the stamps are prepared typically by replica molding against suitable rigid masters such as silicon surfaces microstructured by etching techniques, photoresist patterns, embossed polymer surfaces, etc.

Molecules used are mainly alkanethiols and alkylsilanes with different terminal functional groups (such as NH_2 and CH_3) and will be printed respectively onto coinage metal surfaces (Au, Ag, …) and oxidized surfaces (glass, quartz, silicon wafers…). Self-assembled monolayers are thus patterned with line/spacing and dot array patterns with periodicity down to hundred of nanometers. By using master molds prepared from photolithography and exploiting the same molds as phase shift masks, we expect to be able to pattern lines and dots as narrow as 200 nm [62]. If necessary more complicated patterns, such as for instance asymmetric patterns to study ratchet effects, may be produced by designing new photolithographic masks.

Nanoimprinting is based on the use of a rigid mold (usually made out of silicon, metal or glass) bearing micron or nanosized relief features on its surface. This method is perfectly appropriate for the fabrication of nanopatterning mold. However, the techniques required in this process are more time consuming and heavier than for the previous one.

The model surfaces presented in Fig. 22.1 are heterogeneous molecular films of alkylsilanes compounds self-assembled onto silicon substrates realized by microcontact printing and nanoimprinting. Two organosilanes molecules, hexadecyltrichlorosilanes (referred to as CH_3) and (6-aminohexyl)-aminopropyltrimethoxysilanes (referred to as NH_2) were used to create these patterns. The binary geometrical patterned surfaces were prepared according to a two step printing process. First, different NH_2 SAMs topologies: stripes, squares and stars structures were printed onto the bare and cleaned silicon wafer with different NH_2 surface fraction: from 75 to 25 %; the final and complete heterogeneous surface was obtained by self-assembling a second continuous monolayer of a second molecule (methyl terminated one: CH_3) around the domains in the remaining space. These substrates were then characterized by Atomic Force Microscopy in Contact Mode (LFM). The frictional experiments consisted in scanning laterally the AFM tip under a constant velocity at (i) a constant load to get the friction contrast images of the patterned surfaces and (ii) at different loads to determine their friction coefficient. The tips used are silicon nitride (Si_3N_4) with a radius of curvature $\sim$20 nm. The average size and surface coverage were determined by analysis of AFM pictures (Fig. 22.1).

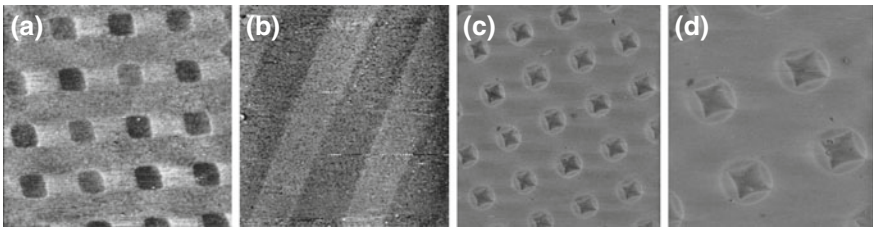


Fig. 22.1 AFM pictures (Contact Mode-Friction) of nano and micropatterned molecular surfaces. **a** hydrophilic (NH_2-terminated SAM) square nanodomains (in *dark*) surrounded by hydrophobic thin molecular film (CH_3-terminated SAM), **b** alternated microstrips of hydrophilic NH_2 (in *dark*) and hydrophobic CH_3 molecular films, **c** hydrophilic (NH_2) micropatterns (in *dark*) surrounded by hydrophobic (CH_3) continuum **d** zoom in on (**c**) patterns

These images show the high sensitivity of Lateral Force Microscopy (LFM) to the chemical contrast of the functional groups grafted onto the surface.

The friction experiment consisted in scanning laterally the AFM tip under a constant velocity at (i) a constant load to get the friction pictures of the patterned surfaces and (ii) at different loads to determine their friction coefficient. The tips used are silicon nitride (Si_3N_4) with a radius of curvature of ~20 nm. The frictional force between tip and surface could be estimated from the voltage signal in the scope loop of LFM and friction coefficient can be calculated by multiplying the average signal of Trace Minus Retrace (TMR) during a scan by a calibration factor.

The results of the external load dependent friction response onto square patterns are plotted in Fig. 22.2. For instance, a 50 % NH_2 surface fraction is composed of 2 μm edge squares separated from the next pattern by 1 μm distance. The varying parameter in the surface fraction is the periodicity of the length patterning. As expected from 22.1, the measured friction force is linearly proportional to the external applied load. This response is characteristic of a Coulombic behavior. Moreover, the averaged frictional response is found to be more sensitive to the chemistry and relative surface fraction of the molecular domains than to the size and distribution of patterns (micro or nanoscopic). This is shown by the regular increase of the tip/substrate friction on these (NH_2–CH_3) patterned surfaces, with the surface fraction of the hydrophilic molecular species (NH_2). The slope of the curves is proportional to the friction coefficient and represented in Fig. 22.3.

The results show that friction coefficient decrease linearly as surface coverage of NH_2 for micropatterned surfaces. The friction forces increase with the magnitude of the adhesive interaction the SAMs and the tip. As a result the greater the surface fraction of NH_2 in the thin film is, the greater the friction response is, in a good agreement with literature [66]. Lopez et al. have also prepared striped micropatterned of methyl and hydroxyl terminated SAMs and studied correlation between friction and grafted mole fraction. The friction seems to increase steadily with mole fraction of one molecule. The dependence of friction on the chemical composition of this geometrical patterned coating seems to be correlated by a Cassie type relation for heterogeneous wetting [67, 68]: the frictional response of the patterned surface composed of two

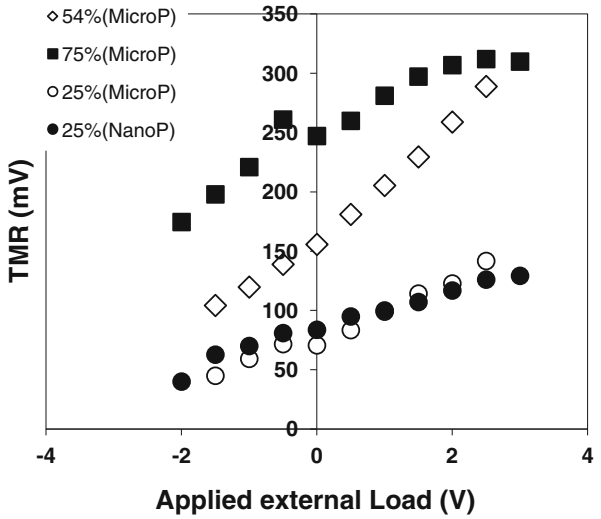


Fig. 22.2 TMR as a function of applied external load for different NH_2 surface coverage on squares Micropatterned (MicroP) and Nanopatterned (NanoP) binary NH_2/CH_3 surfaces

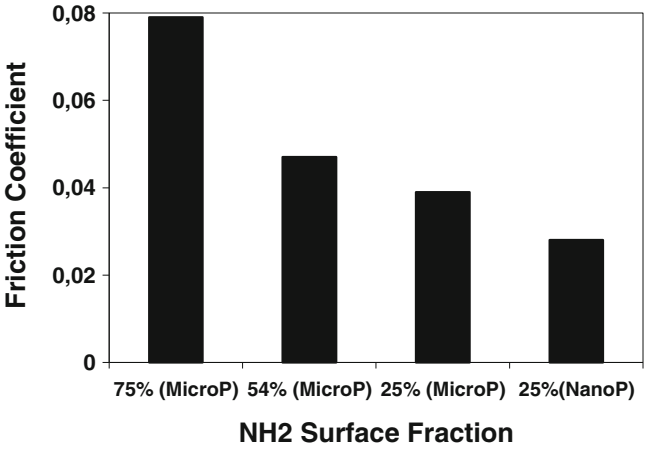


Fig. 22.3 Friction coefficient versus NH_2 surface fraction for Micropatterned (MicroP) and Nanopatterned (NanoP) binary NH_2/CH_3 molecular surfaces at normal ambient conditions

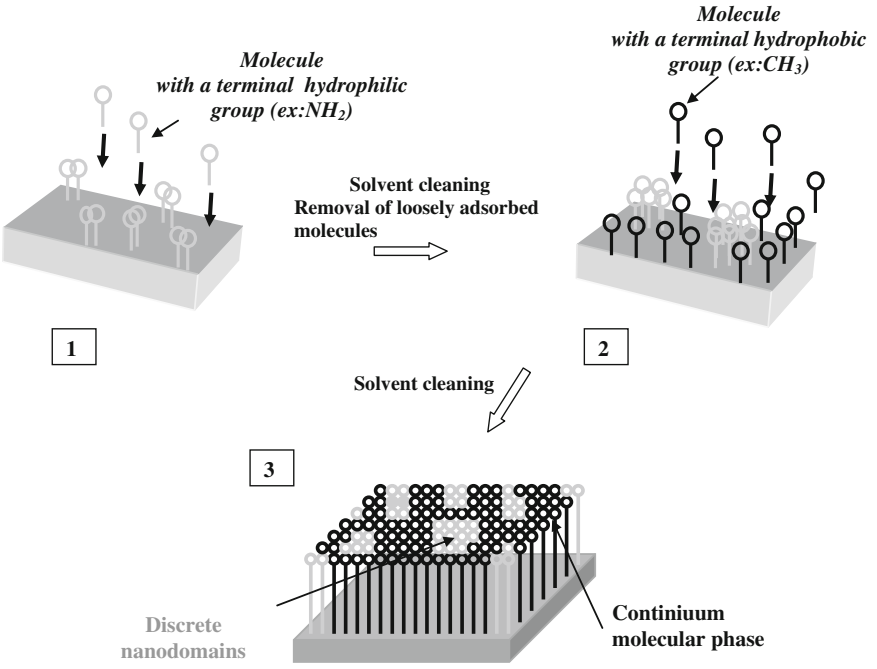
uniformly distributed components is proportional to the surface fraction of each component, $\mu_{\text{binary}} \sim (\phi_{NH_2})\mu_{NH_2} + (1 - \phi_{NH_2})\mu_{CH_3}$. However, at comparable surface fraction, friction coefficient is lower onto nanopatterned surfaces than on micropatterned ones for the same geometrical design. This result may traduce the coupling effects between nanodomains which lead to a collective frictional response on nanoscale patterned surfaces. Indeed, contrary to microsize patterns where the nanosize tip essentially responds to the discrete feature of the microdomains, the

tip/substrate interaction on nanoheterogeneous substrates integrates the force fields emanating from neighboring nanodomains, the extent of which from the tip apex scales with that of long-range forces. This nanosize effect which couples the nearest nanodomains through the lateral overlap of their force fields can involve new properties which are not yet completely understood. Further investigations are required to expand the comprehension of these two-dimensional scale effects on interface phenomena.

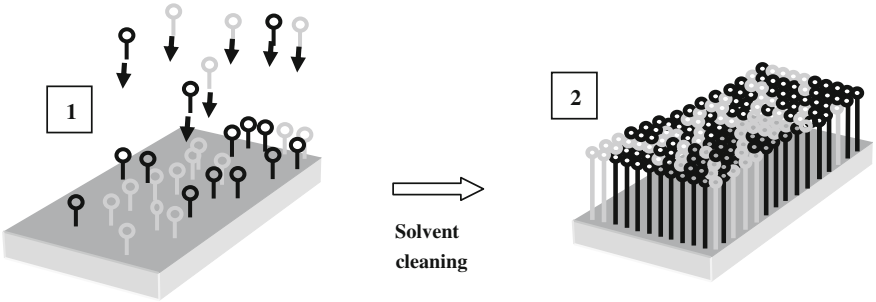
22.2.1.2 Random Nanopatterns

Random chemical nanoscale patterns and features [58, 69] spanning lengthscales from a few nanometers to microns were fabricated using two different self-assembling methods of alkylsilanes on (Si/SiO_2): a ***sequential*** and a ***co-adsorption one,*** and nanoscale friction experiments were performed onto these nano-heterogeneous molecular films.

1. For the random heterogeneous nanopatterns prepared by ***sequential assembling***, the binary surfaces may be composed of a discrete distribution of hydrophilic nanodomains within a continuum of hydrophobic molecules, or vice versa, as illustrated in Scheme 22.6, for CH_3 and NH_2 terminated alkylsilane molecules.



Scheme 22.6 Two step sequential elaboration of binary molecular thin films. Step *1 solvent phase* adsorption-nucleation and growth of discrete molecular nanodomains (nm–μm), *2* self-assembling of the continuum phase by *vapor or solvent phase* adsorption *3* resulting nano-heterogeneous surface



Scheme 22.7 Sketch of the co-adsorption process allowing to prepare mixed SAMs of different chain lengths or terminal groups (hydrophobic CH_3/hydrophilic NH_2): *1* co-adsorption and self-assembling from a solvent or by vapour phase deposition, and *2* resulting SAMs surface

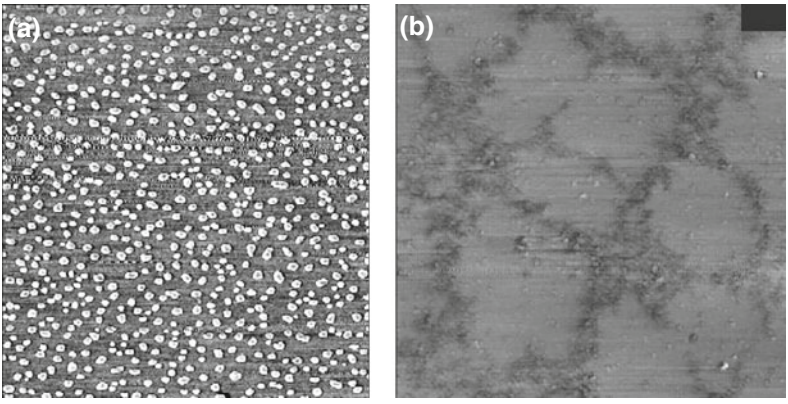


Fig. 22.4 AFM pictures (Tapping Mode) of random nanoheterogeneous molecular surfaces prepared by sequential adsorption. **a** hydrophobic CH_3 nanodomains (*clear dots*), dispersed in a hydrophilic NH_2 continuum grafted by vapor phase adsorption, **b** hydrophilic NH_2 microdomains (in clear) dispersed in a network of hydrophobic CH_3 continuum, both molecular domains realized by solvent phase coating

AFM characterizations were performed onto these binary heterogeneous molecular films and results are displayed in Fig. 22.4. Domains size can be easily tuned from tens of nanometers to micrometer and roughness of the thin coating is of the order of one nanometer as it corresponds to the length difference of both molecules.

As previously, friction experiments were performed onto these substrates, and the results are presented in Fig. 22.5. The external applied load (F_{ext}) dependence of friction exhibits the expected Coulombic behavior of linear increase of friction versus F_{ext}. This result (not plotted here), is similar to the previous one observed onto the regular micropatterned molecular thin coatings. Thus, friction coefficients were determined from the slope of these linear responses and presented in Fig. 22.5.

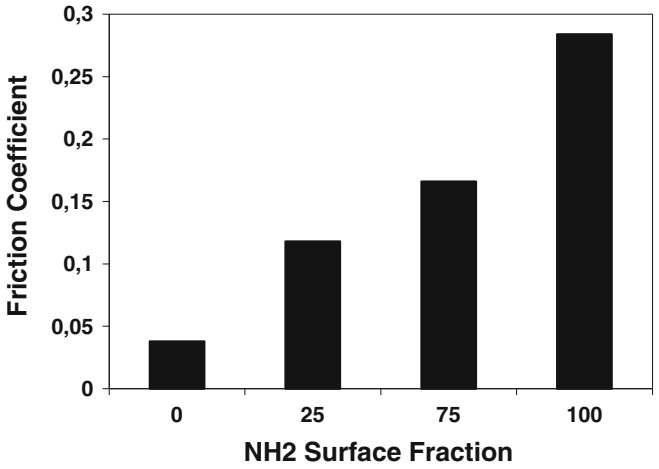


Fig. 22.5 Friction coefficient versus NH_2 surface fraction for nanoscale heterogeneous binary (NH_2/CH_3) molecular surfaces prepared by sequential process

As one could expect from the intermolecular interactions involved at the (hydrophilic tip/substrate) contacts, the magnitude of this friction gradually and significantly increases with the surface fraction of NH_2, when going from the uniform CH_3 to the uniform NH_2 molecular films. On the homogeneous hydrophobic CH_3 SAM, the interactions with the hydrophilic silicon nitride tip mainly involve the London dispersion forces that have a much lower magnitude as compared to the polar, hydrogen and electrostatic bonds involved in adhesion (and hence friction) of the tip with the homogeneous NH_2 terminated substrates.

2. For the heterogeneous surfaces prepared with a *co-adsorption process*, the surfaces are composed of a dissemination of hydrophilic nanodomains in a hydrophobic continuum, and vice versa (Fig. 22.6). The binary heterogeneous surfaces were obtained in this case through a single step process (sketch of Scheme 22.6). This grafting technique is based on the co-adsorption of two molecules (organosilanes or thiols) in the proper solvent, or by vapor-phase deposition.

The AFM images characteristic of such nanoscale heterogeneous SAM film prepared by co-adsorption is shown in Fig. 22.6. The domains are smaller and their contours irregular and less well-defined as compared to those obtained by sequential process.

The nanoscale frictional response of these mixed binary SAMs thus appears to be directly proportional to the surface fraction of their constitutive molecular compounds. This should provide a semi-quantitative tool based on the friction coefficient measurement for the compositional surface analysis [70] of submicron scale heterogeneous substrates

As shown in Fig. 22.8, this assumption is true for mixed SAMs of molecules as diverse as alkyls terminated CH_3, amine NH_2, hydroxide OH, or acidic group COOH

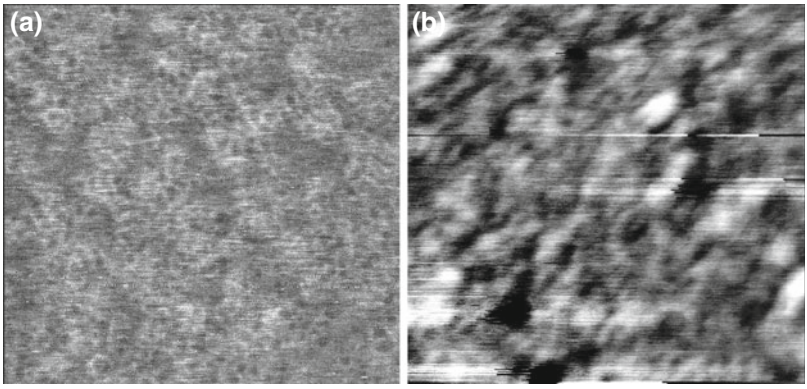


Fig. 22.6 Contact mode AFM pictures of the random nanoheterogeneous molecular surfaces prepared by co-adsorption: **a** hydrophobic CH_3 nanodomains (in clear) dispersed in a hydrophilic NH_2 continuum (both grafted in a single step vapor-phase adsorption from a mixture of the two molecules), **b** nanostructured molecular domains in the binary hydrophilic NH_2 (in clear) and hydrophobic CH_3 heterogeneous SAM (both grafted in a single step solvent-phase adsorption from a mixture of the two molecules)

groups, but no longer holds for fluorocarbon-based (or rich) molecules. Indeed, as compared to thin films of hydrocarbon molecules, the friction is higher on fluorocarbon molecules [71]. The introduction of the fluorocarbon molecule in a mixed SAM leads to an increase in the frictional response, even for a low nominal bulk concentration or surface fraction. For these fluorocarbon-based binary SAMs the overall frictional response does not show any trivial correlation with the chemical composition of the surface (Fig. 22.7b and 22.8). In that case, the high friction that is observed on the mixed SAMs can reasonably be accounted for by additional energy dissipation modes, which arise due to differences in the 'phase state' of the SAMs. For mixed hydrocarbon–fluorocarbon systems for instance, it has been shown that the fluorocarbon chains imposes a nearly perpendicular orientation of the alkyl chains with respect to the surface normal. In that case, the molecular contrast has created a new structural organization of the molecules within the mixed SAMs inducing (creating) some new tribological properties

Based on the nanoscale frictional responses on the different patterns, it seems that the chemistry and the surface fraction of the constitutive molecular species are the leading parameters which control the frictional behavior of these heterogeneous coatings. The size, lengthscales and geometry of the patterns can be used to modulate the fluctuating amplitude of friction from the discrete profile characteristic of each chemical domain on micronsize patterns, to a smoother one characterized by the coupling effects on nanoscale heterogeneous patterns. Unfortunately, there is still no systematic investigation of these scale effects on chemically heterogeneous SAMs and their impact on friction, especially its environment-dependent behavior (dissipation by domain frontiers defects, condensation, etc.). Understanding these scale

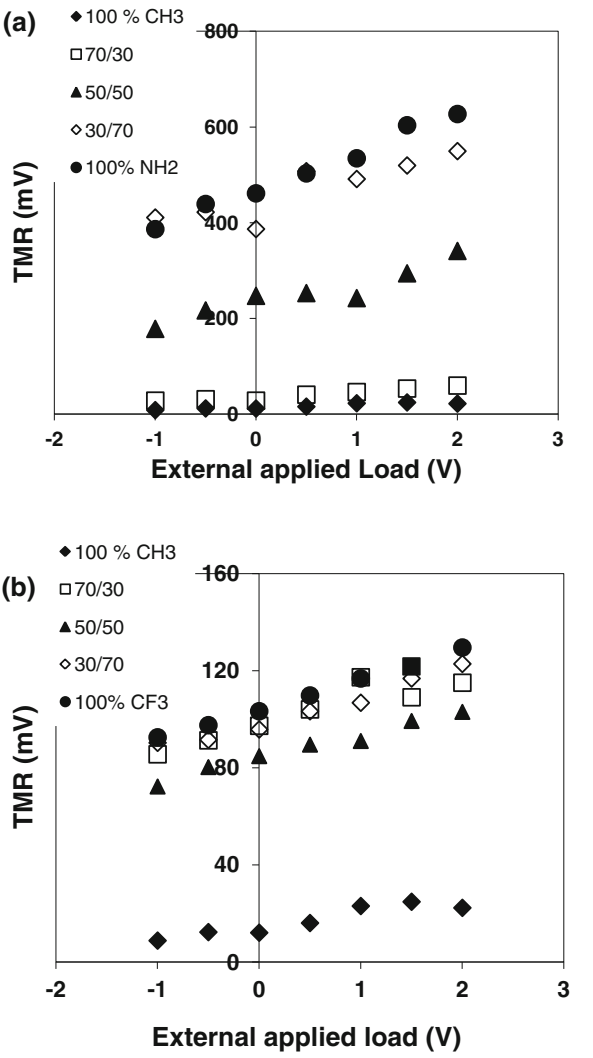


Fig. 22.7 TMR as a function of applied external load for mixed SAMs, for different surface coverage of: **a** NH_2 terminated molecular domains, and **b** CF_3 terminated domains, in the binary NH_2/CH_3 and CF_3/CH_3 terminated SAM, respectively

effects is not only of fundamental importance. This also provides us with a key control parameter for adjusting over different lenghscales (macro, micron to nano) the surface chemistry, patterns and properties for a given application. For heterogeneous SAMs which are being increasingly used in almost all the fields of nanotechnology (MEMs, NEMs, molecular electronics, controlled assembling of nanoparticles arrays, …), this is a challenge that will focus increasing research interests in the near future.

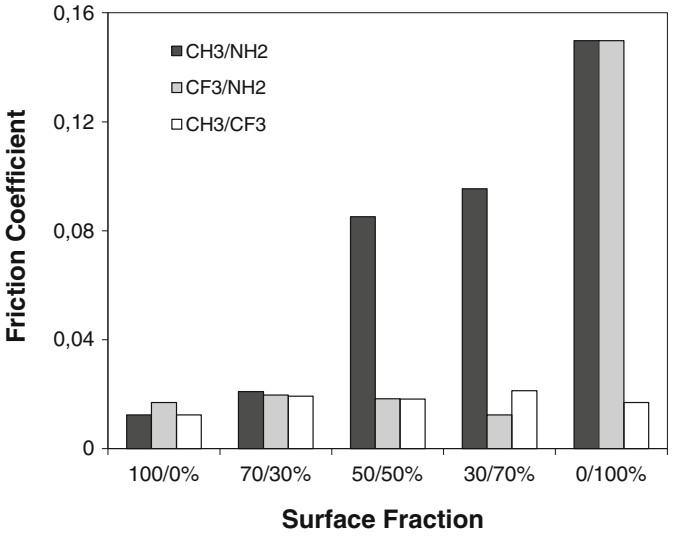


Fig. 22.8 Friction coefficient versus (*1*) NH_2, (*2*) CF_3 and (*3*) CH_3 surface fraction, respectively, on heterogeneous NH_2/CH_3, CH_3/NH_2 and CF_3/CH_3 SAM surfaces prepared by co-adsorption

22.2.2 Influence of Sliding Velocity

Investigating the influence of the tip velocity on the nanoscale friction response of grafted coatings represents the most accessible way to efficiently probe their dynamical behavior and to understand some of the dissipation mechanisms (frequency-dependent ones) involved in the friction process. For random heterogeneous surface patterns [66], either prepared by the sequential or co-adsorption methods, two distinct regimes of frictional response to sliding velocity are observed, depending on the nature of the intermolecular forces and structural sate of the SAM coating.

In these measurements, the sliding velocities were varied between 0.2 and 240 μm/s to cover about 3 orders of magnitude, allowing to depict specific interface phenomena at both low and high molecular stress. It should be noted that these measurements were carried out in normal ambient (20 °C, 30 %RH) and in wearless conditions for the coating (as verified experimentally).

For the more hydrophobic substrates which mainly interact with the tip through London dispersion forces, a regular and low magnitude increase of the friction with the sliding velocity (**v**) toward a plateau [72] is observed. On the other hand, the frictional response on partially to highly hydrophilic substrates clearly displays two distinct regimes. In the first regime (low-velocity region, typically <20 μm/s), a steep increase of the friction is observed up to a maximum ($20\,\mu m/s < \mathbf{v} < 50\,\mu m/s$), followed by a smoother decrease toward a stabilization plateau. The magnitude of the transition between these two regimes directly increases with the hydrophilic character (NH_2, S_iO_2H), or the surface fraction of the polar molecules for binary

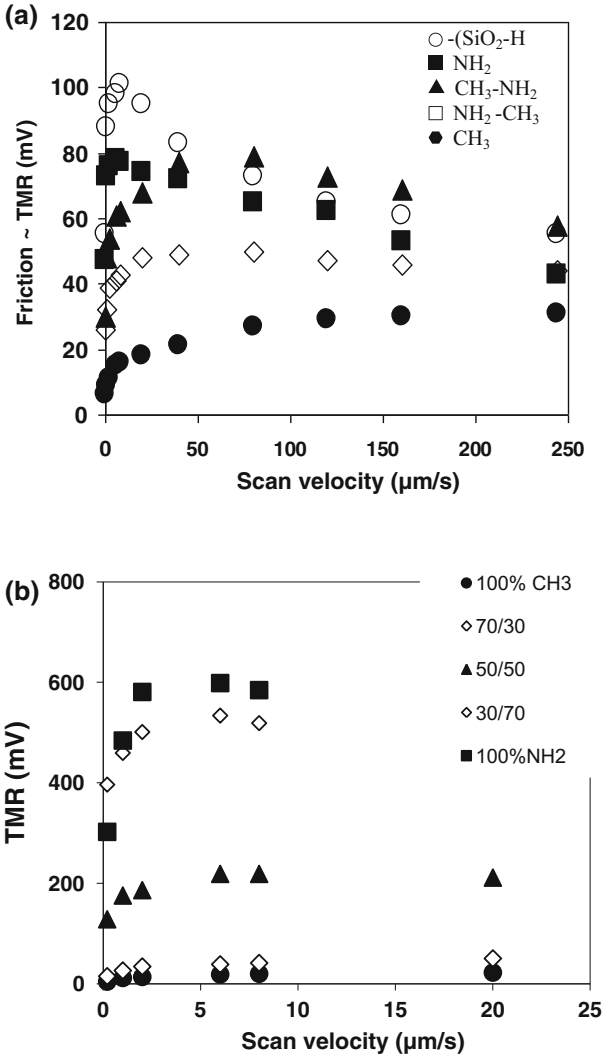


Fig. 22.9 Friction as a function of sliding velocity for random binary nanoheterogeneous NH_2/CH_3 molecular surfaces: **a** prepared by sequential adsorption **b** prepared by co-adsorption

heterogeneous surfaces, as shown in Fig. 22.9. This velocity-dependent frictional transition is thus observed for all hydrophilic substrates, independently of their chemical nature (organic, mineral) or composition (homogeneous, heterogeneous). This transition thus seems to essentially rely on the common and unique feature that is the existence of permanent dipoles and hydrogen bonds at the tip/substrate contact. Indeed, in contrast to the London dispersion forces which predominate the hydrophobic CH_3/tip contact, the directional and orientational polar and hydrogen

bond interactions that are involved at the hydrophilic/tip contact are strongly sensitive to thermal effects. Since these thermal effects increase with the sliding velocity, one equally expects the disorientation over the polar and hydrogen bonds in the molecular film to increase, leading from a certain critical velocity to a decreasing adhesive and friction forces. In addition, it was mentioned that a wetting film could form in a hydrophilic contact, even at a low relative humidity (Fig. 22.6). The destruction of such boundary liquid bridge due to thermal effects (dissipation) can also explain the friction inversion observed at high **v**.

Similar velocity dependence of friction was also reported for other hydrophilic substrates [34, 39, 40]. The slope inversion (friction transition) was explained and attributed to various energy dissipation processes or capillary effects (already discussed above in Sect. I.3). Indeed, during sliding, the frictional energy can be partially and irreversibly dissipated into thermal energy, the so-called thermodynamic energy dissipation $E_{dissip} \sim F_f . \mathbf{v}.dt$, creating an orientational disorder in the end-group dipoles of the molecular film. As known from literature, such thermally induced collective disorder involving the entire topmost molecular population (not limited to a single polar group) can drastically reduce the interaction density and energy within the contact zone(dipolar and H-binding ones), especially on SAMs.

Other theoretical interpretations such as conformation and structural changes in the molecular films have been proposed to explain this velocity-dependent frictional transition: the shear-induced transition from solid-like to liquid-like state, the viscous damping giving rise to molecular interdigitations or entanglements of the molecules [73, 74], and the water condensation effects [39]. However, most of these effects during sliding lead to irreversible interfacial phenomena as dissipation of local heat in organic coatings induce complex film modifications up to severe wear. Nonetheless, this discussion still requires further investigations to acquire a unified picture of this velocity-dependent friction behavior, independently of the nature, chemistry, structure and topology of the contacting substrates (organic, mineral, composites, etc.).

22.3 Wear of SAMs

The frictional properties at nano and micronscales have drawn a lot of attention as they provide critical information regarding the mechanisms and origin of wear on these organic molecular films.

During the sliding of a tip onto a SAMs coated substrate, there are two main factors which create wear: (i) the irreversible disorder induced by the amplication of defects [34] in the thin molecular film and (ii) the rupture of SAMs/substrate bonds in a second step.

The interfacial wear mechanism is mainly governed by a kinetic energy transfer from the tip to the molecular film. This additional energy is dissipated through the film coating by creating modification in the molecules conformations and disorder, leading in extreme case, to plastic deformations up to breaking of covalent bonds

involving real damage in the coatings. In a SAMs coating, the higher the density of the defects is, the faster the wear will occur. Indeed, below a certain critical load, in normal operating conditions (∼20 °C, ∼30 % RH) SAMs can undergo orientational modification without damage. From this critical load (above film cohesion energy), they wear and eventually detach from the substrate, leading to irreversible damage of the molecular film. It is worth noting here that since the intrinsic stability of SAMs (cohesion, structure, molecular mobility) is strongly dependent on the environment conditions (humidity, T°, pH), their wear can be drastically accelerated by the coupling of stress to environment conditions. In order to improve wear resistance, Liu et al. have suggested the grafting of rigid spacer chains [24].

22.4 Conclusion

Self-assembled monolayers composed of organosulfur and organosilanes have found wide-spread applications in surface and interfacial science due to their well-defined and robust structures and the ability to simply modify their chemical composition in order to introduce a variety of modifications to their properties. However, nano and microtribological investigations on these stable thin organic coatings still remains delicate.

First, the tribological performance of the homogeneous organic grafted thins films mostly depend on their intrinsic properties such as the chain length of the molecules, the area density, the chemistry of the terminal group and their method of preparation. Indeed, the frictional behavior of the SAMs with different functionalities is primarily influenced by their surface energy, which is directly related to the interfacial interactions between the tip and the chemistry of the surface. However, the effects of humidity and temperature have been studied and displayed the impact of these environmental conditions on friction on bare and coated surfaces. Particularly, the significant influence of relative humidity on frictional forces could have been explained by the thickness of the adsorbed water layer.

In a second part, random and well-defined nanopatterned heterogeneous molecular thin films have been investigated in term of nanofriction. Depending on their topological patterns, these surface heterogeneities generate a wide variety of frictional responses. Nanostructuration seems to have a fundamental and crucial effect on the tribological properties of the molecular coating; it induces exclusive and specific tribological properties on the thin coating. Then, the velocity-dependent frictional behavior of patterned SAMs seems to be directly related to the polarization ability of the end functionalities of the grafted molecules. Finally, SAMs represent an easy as well as an active and dry lubricant for nanotechnological devices; they have dramatically reduced friction and adhesion and have found use in various MEMS devices. However, alkyl chains do not sustain high compression and shear stresses, which significantly limits their lifetime of use. These limitations are motivating the development of a new generation of thin coating based on polymer nanocomposites that display higher nano and micro-tribological properties, as proposed by Sidorenko et al. [75]

References

1. H.-S. Ahn, P.D. Cuong, S. Park, Y.-W. Kim, J.-C. Lim, Wear **255**, 819 (2003)
2. M.K. Corbierre, N.S. Cameron, R.B. Lennox, Langmuir **20**, 2867 (2004)
3. J.J. Storhoff, S.S. Marla, P. Bao et al., Biosens. Bioelectron. **19**, 875 (2004)
4. H.I. Kim, T. Koin, T.R. Lee, S.S. Perry, Langmuir **13**, 7192 (1997)
5. M.J. Azzopardi, H. Arribart, J. Adhes. **46**, 103 (1994)
6. R.H. Yoon, D.A. Guzonas, Colloid Surf. A: Physicochem. Eng. Aspects **87**, 163 (1994)
7. C.R. Tripp, M.L. Hair, Langmuir **8**, 1120 (1992)
8. P. Silberzan, L. Leger, D. Ausserré, J.J. Benattar, Langmuir **7**, 1647 (1991)
9. D. Angst, G.W. Simmons, Langmuir **7**, 2236 (1991)
10. C.R. Kessel, S. Granick, Langmuir **7**, 532 (1991)
11. A.G. Carson, S. Granick, J. Mat. Res. **5**, 1745 (1990)
12. D.K. Schwartz, S. Stienberg, J. Israelchvili, J. Azasadzinski, Phys. Rev. Lett. **69**, 3354 (1992)
13. J.V. Davidovits, V. Pho, P. Silberzan, M. Goldman, Surf. Sci. **352–354**, 369 (1996)
14. J. Gun, J. Sagiv, J. Colloid Interface Sci. **112**, 457 (1986)
15. E. Delamerche, B. Michel, C. Gerber et al., Langmuir **10**, 2869 (1994)
16. K. Kendall, Nature **319**, 203 (1986)
17. A. Lio, D.H. Charych, M. Salmeron, J. Phys. Chem. B **101**, 3800 (1997)
18. X. Xiao, J. Hu, D.H. Charych, M. Salmeron, Langmuir **12**, 235 (1996)
19. M.T. McDermott, J.-B.D. Green, M.D. Porter, Langmuir **13**, 2504 (1997)
20. D. Dominguez, R.L. Mowery, N.H. Turner, Tribol. Trans. **37**(1), 59 (1994)
21. H.J. Eyring, J. Chem. Phys. **3**, 107 (1935)
22. D. Tabor, in *Fundamentals of Friction: Macroscopic and Microscopic Processes*, ed. by I.L. Singer, H.M. Pollock (Kluwer Academic, Dordrect, The Netherlands, 1992), p. 3
23. B. Bhushan, A.V. Kulkarni, V.N. Koinkar et al., Langmuir **11**, 3189 (1995)
24. H. Liu, B. Bhushan, Ultramicroscopy **91**, 185 (2002)
25. A.J. Kinloch, *Adhesion and Adhesives* (Chapman and Hall, New York, 1987)
26. C.J. Drummond, G. Georgaklis, C. Chan, Langmuir **12**, 2617 (1996)
27. H.I. Kim, T. Koini, T.R. Lee, S.S. Perry, Langmuir **13**, 7192 (1997)
28. M. Graupe, T. Koini, H.I. Kim, N.G. Arg, Y.F. Miura, M. Takenaga, S.S. Perry, T.R. Lee, Langmuir **15**, 3179 (1999)
29. C.O. Timmons, W.A. Zisman, J. Colloid, Interface Sci. **22**, 165 (1966)
30. T. Stoebe, P. Mach, S. Grantz, C.C. Huang, Phys. Rev. **53**, 1662 (1996)
31. M. Graupe, T. Koini, H.I. Kim, N. Garg, Y.F. Miura, M. Takenaga, S.S. Perry, T.R. Lee, Colloids Surf. A: Physicochem. Eng. Aspects **154**, 239 (1999)
32. B.N.J. Person, Phys. Rev. B **44**, 3277 (1991)
33. M.K. Chaudhury, M.J. Owen, Langmuir **9**, 29 (1993)
34. H. Yoshizawa, Y-L. Chen, J. Israelachvili J. Phys. Chem. **97**, 4128 (1993)
35. R.M. Overney, E. Meyer, J. Frommer, H.-J. Giintherodt, M. Fujihira, H. Takano, Y. Gotoh, Langmuir **10**, 1281 (1994)
36. M. Garcia-Parajo, C. Longo, J. Servat, P. Gorostiza, F. Sanz, Langmuir **13**, 2333 (1997)
37. B. Bhushan, H. Liu, Phys. Rev. B **63**, 245412 (2001)
38. F. Tian, X. Xiao, M.M. Moy, C. Wang, C. Bai, Langmuir **15**, 244 (1999)
39. E. Riedo, F. Lévy, H. Brune, Phys. Rev. Lett. **88**, 185505 (2002)
40. Y. Liu, D.F. Evans, Q. Song, D.W. Grainger, Langmuir **12**, 1235 (1996)
41. R. Yerushalmi-Rozen, J. Klein, Langmuir **11**, 2806 (1995)
42. S.R. Cohen, R. Naaman, J. Sagiv, J. Phys. Chem. **90**, 3054 (1986)
43. M.C. Yeh, E.J. Kramer, R. Sharma, W. Zhao, M.H. Rafailovich, J. Sokolov, J.D. Brock, Langmuir **12**, 2747 (1996)
44. E. Gnecco et. al., Phys. Rev. Lett. **84**, 1172 (2000)
45. Y. Liu, T. Wu, D.F. Evans, Langmuir **10**, 2241 (1994)
46. J. Wagner, T. Kirner, G. Mayer, J. Albert, J.M. Köhler, Chem. Eng. J. (Article in Press)

47. G.B. Khomutov, V.V. Kislov, M.N. Antipina, R.V. Gainutdinov et al., Microelectron. Eng. **69**, 373 (2003)
48. W. Li, L. Huo, D. Wang, G. Zeng, S. Xi, B. Zhao, J. Wang, Y. Shen, Z. Lu, Colloids Surf. A: Physicochem. Eng. Aspects **175**, 217 (2000)
49. L.M. Tender, R.L. Worley, H. Fan, G.P. Lopez, Langmuir **12**, 5515 (1996)
50. R. Klauser, M.-L. Huang, S.-C. Wang, C.-H. Chen, T.J. Chuang, A. Terfort, M. Zharnikov, Langmuir **20**, 2050 (2004)
51. M.M. Alkaisi, W. Jayatissa, M. Konijn, Curr. Appl. Phys. **4**, 111 (2004)
52. M. Tormen, L. Businaro, M. Altissimo, F. Romanato, S. Cabrini, F. Perennes, R. Proietti, H-B. Sun, S. Kawata, E.D. Fabrizio, Microelectron. Eng. (2004) (in press)
53. O. Cherniavskaya, A. Adzic, C. Knutson, B.J. Gross, L. Zang, R. Liu, D.M. Adams, Langmuir **18**, 7029 (2002)
54. M.R. Shadnam, S.E. Kirkwood, R. Fedosejevs, A. Amirfazli, Langmuir **20**, 2667 (2004)
55. R.D. Piner, J. Zhu, F. Xu, S. Hong, C.A. Mirkin, Science **283**, 661 (1999)
56. S. Hong, J. Zhu, C.A. Mirkin, Science **286**, 523 (1999)
57. O.J. Schueller, D.C. Duffy, J.A. Rogers, S.T. Brittain, G.M. Whitesides, Sens. Actuators **78**, 149 (1998)
58. H. Haidara, K. Mougin, J. Schultz, Langmuir **16**, 7773 (2000)
59. J.P. Folkers, P.E. Laibnis, G.M. Whitesides, Langmuir **8**, 1330 (1992)
60. M.J. Wirth, R.W. Fairbank, H.O. Fatunmbi, Science **275**, 44 (1997)
61. S.P. Li, A. Lebib, D. Peyrade, M. Natali, Y. Chen, Appl. Phys. Lett. **77**, 2743 (2000)
62. Y. Chen, A. Lebib, S. Li, A. Pépin, D. Peyrade, M. Natali, E. Cambril, Eur. Phys. J. AP **12**, 223 (2000)
63. M. Natali, A. Lebib, E. Cambril, Y. Chen, I.L. Prejbeanu, K. Ounadjela, J. Vac. Sci. Technol. B **19**, 2779 (2001)
64. Y. Xia, G.M. Whitesides, Angew. Chem. Int. Ed. **37**, 550 (1990)
65. S.Y. Chou, P.R. Krauss, P.J. Renstrom, Appl. Phys. Lett. **67**, 3114 (1995)
66. K. Mougin, G. Castelein, H. Haidara, Tribol. Lett. **17**(1), 11 (2004)
67. J. Drelich, J.L. Wilbur, J.D. Miller, G.M. Whitesides, Langmuir **1996**, 12 (1913)
68. F.M. Fowkes, *Contact Angle: Wettability and Adhesion* (American Chemical Society, Washington D.C., 1964)
69. H. Haidara, K. Mougin, J. Schultz, Langmuir **16**, 9121 (2000)
70. Y.G. Zhou, H. Fan, T. Fong, G.P. Lopez, Langmuir **14**, 660 (1998)
71. H. Schonherr, G.J. Vansco, Mat. Sci. Eng. C **8–9**, 243 (1999)
72. S. Gauthier, J.P. Aimé, T. Bouhacina, A.J. Attias, B. Desbat, Langmuir **12**, 5126 (1996)
73. B. Bushan, J.N. Israelachvili, U. Landman, Nature **374**, 607 (1995)
74. M. Schoen, C. Rhykerd, D. Diestler, J. Cushamn, Science **245**, 1223 (1989)
75. A. Sidorenko, H.-S. Ahn, D.-I. Kimb, H. Yanga, V.V. Tsukruk,, H., Wear **252**, 946 (2002)

Part VI
Nanowear

Chapter 23
From Nano and Microcontacts to Wear of Materials

Rogerio Colaço

23.1 Introduction

This chapter is a revised version of the text "Surface damage mechanisms: from nano and microcontacts to wear of materials" published in 2007 Meyer's and Gnecco's book "Fundamentals of Friction and Wear on the Nanoscale" [1]. During these seven years relevant advances in the understanding of the relations between the nanotribological behaviour and the overall tribological macroscale response of materials have been made. Yet, tribology, either at macro, micro or nanoscales is still a wild territory, mostly unknown, in its whole, not only by the common people, but also by the scientific community.

In this revised version the first change that was decided was the title. In fact, although this text deals essentially with wear, and its relations at nano, micro and macroscales, the part "surface damage mechanisms", was removed from the title, by a number of reasons. The most important of that reasons was the fact that wear is not always "damage" in its straight assumption: the several billion euros industries of decorative stones for civil construction or of the Si wafer polishing for electronic applications [2] are just some examples. In fact, wear is simultaneously one of the few causes of components damage and obsolescence, consequently with a high economical and industrial impact, but also a technological issue of growing relevance. Neverthless, historically, not much attention had been paid to wear or, in a wider scope, to tribology. Of course that some important exceptions are worth to mention, but even the world "tribology" did not exist fifty years ago. In fact, it was in the began of 1966 that a report made by an UK government committee presided by H. P. Jost, known as the *Jost Report* found that a large waste of resources—estimated in 515 million sterling pounds per year (approximately 4 % of the 1965 UK GNP)—occurred because of ignorance of surface interaction phenomena [3]. Although, friction,

R. Colaço (✉)
Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001 Lisbon, Portugal
e-mail: rogerio.colaco@ist.utl.pt

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_23

lubrication and wear had been studied for many years before, the fact is that the Jost Report was somehow the kick out for the launch of several R&D and education programs in tribology. Even the introduction in the lexicon of the word "tribology"—the science of rubbing—was one of the (first) consequences of this report.

When the Jost Report was presented and published, it was widely felt that it greatly exaggerated the savings that might result from improved tribological expertise. Later it becomes clear that, on the contrary, the Jost Report underestimated the economical importance of tribology, since it paid small attention to wear, which happens to be, from an economical point of view, the most significant tribological phenomena [4, 5]. Recent studies [6–8] still points for an important economical impact of tribological related phenomena, namely wear, either faced as a damage mechanism or a technology: at least 1 % of the GNP of an industrialized country might be saved with minimal further investment in research, and the potential for even larger savings might exist with further research [9–11]. These amounts should be a strong encouragement for the R&D in tribology all over the world. In what concerns wear studies, an increasing number of recent works as pointed out for the need to understand deeper the atomic origins of wear [11] as well as to establish relations between macro-, micro- and nanowear phenomena [12], i.e., bridging the gap between observation scales.

From an application point of view, wear research at submicrometric scales is not only important because of the growing importance of disk storage, micro and nanotechnologies, but also for the optimization of more common and conventional tribological systems in engineering components. In fact, since the reference work of Greenwood and Williamson [13], it is well known that the contact between nominally flat surfaces (the most frequent in engineering components) occurs between small surface asperities, whose contact areas can be only of a few tens of square nanometers. The point being made is that, even in conventional engineering systems, the interactions at nanometric scales cannot be neglected.

With the development of the atomic force microscopes (AFM) [14], approximately three decades ago, it becomes possible to study wear phenomena at very small scales and loads. The AFM tips can be used to simulate a sharp single asperity travelling over a surface [15, 16] or, by using a stiffer steel cantilever with a sharp diamond tip (or diamond coated tip) mounted on its end, nanowear studies at higher loads and with more resistant materials can also be made [17–21]. AFM's can be also used to measure or monitorize at very fine scales the damage and/or topographical changes in surfaces after local rubbing (e.g. [22–24]) or to measure the work of adhesion between the surface and the single asperity simulated by the tip (e.g. [25, 26]). As a consequence, the appearance of this instrument has opened a wide new field of research for tribologists, enabling to enter in the study of wear and damage mechanisms at the scales which were forbidden before.

This possibility of going to smaller and smaller scales in wear studies have raised a number of questions, such as:

- what are the (or, are there) phenomenological thresholds between nano/microwear and macrowear?

- are the classical wear equations valid for rubbing at nanoscale, or new wear equations are needed for nanowear?

Generally speaking, these questions can be summarized in the general formulae "is the *scale* a characteristic of the tribological system that influences its response (such as speed, load, atmosphere, etc.) or, by the contrary, is the phenomenology kept if the tribological system is homothetically scaled down?"

The results obtained up to now are yet scarce, and road still need to be walked, both from the experimental and theoretical point of view. Nevertheless in the recent years important progresses have been made in this topic, as recently pointed ou in a short and sharp paper by Schirmeisen [27]. This chapter aims to give a brief general overview of the actual state of the art, focusing some of the questions and problems arising from the nano and micro wear studies.

Since wear is essentially a surface and subsurface phenomena, we will start by a brief description of the nature of solid surfaces and than we will advance for the some results and models for macro and nanowear and wear at atomic scales.

23.2 The Nature of Solid Surfaces

The term "surface" can be understood as the transition of a material to its surrounding environment [28], either liquid or gaseous, and this term will be used from now on with the more strictly mean of "solid surface".

From the tribological point of view, surface properties are a key factor for the performance of any moving component since, in contacting rubbing pieces, the work is dissipated in the surface and subsurface region. An "engineering surface" can be quite different from an ideal atomically smooth surface in vacuum conditions, whose equilibrium structure can be described, for instance, by the TLK (terrace, ledge, kink) model [29]. Two important differences should be considered between an engineering surface and a TLK surface. The first is the fact that, in the former case, the transition between the bulk properties and the environment is not as sharp as in the latter. The second is the fact that engineering surfaces have topographical features that can be several orders of magnitude wider than the atomic terrace and ledge scales. The tribological performance of an engineering component is therefore strongly dependent of the physical, chemical and mechanical properties of the surface (and by the gradients of these properties) and of the surface topography. Some aspects of these two topics will be briefly discussed next.

23.2.1 Surface Constitution

On approaching the top, solid surfaces in general can present several layers, with properties that can be quite different from the bulk ones. As a consequence, a solid

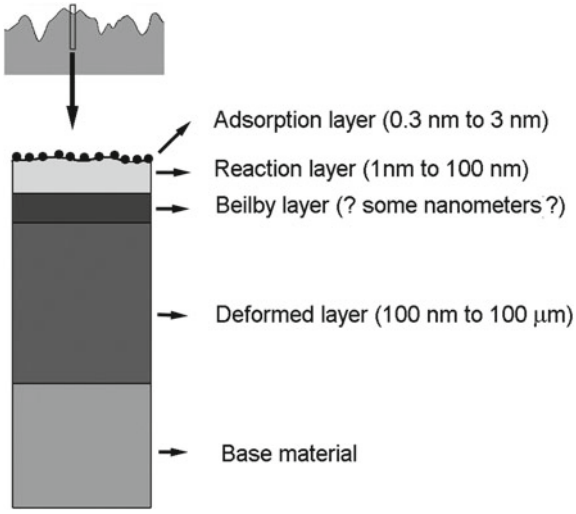


Fig. 23.1 Schematic view of a cross section in an engineering surface

surface is a portion of material that can present a rather complex structure and properties, which, in part depends on the surface preparation method, the nature of the solid and the interactions between the surface and environment (e.g. [10] and see Sect. 23.4).

Usually, in mechanically machined, grinded or polished surfaces, there is a layer of deformed material (formed during the surface preparation process). In metals and metallic alloys, this deformed zone usually is work hardened and can present an hardness higher that of the bulk (see Sect. 23.4). Bowden, Moore and Tabor, in a remarkable set of experiments, were the first to observe clearly this deformed zone, by using taper-section optical microscopy observations [30].

Most of the surfaces are chemically reactive and form native surface oxide layers in air, or other reaction layers depending on the environment (nitrides, sulfides, chlorides, etc.). Besides these native films, adsorbed layers of molecules present in the environment (water, gaseous molecules, organic molecules, etc.) can also be present at the surface.

The presence of all these regions (schematically represented in Fig. 23.1), with their different properties, influences the tribological response of the material. Moreover, if the interaction scale of the damage during contact and rubbing changes between these regions, alterations in the wear response of the material can occur.

23.2.2 Surface Topography

Surfaces are rough. Even the most highly polished component present irregularities (asperities) significantly larger than the atomic scale. Figure 23.2 shows an AFM

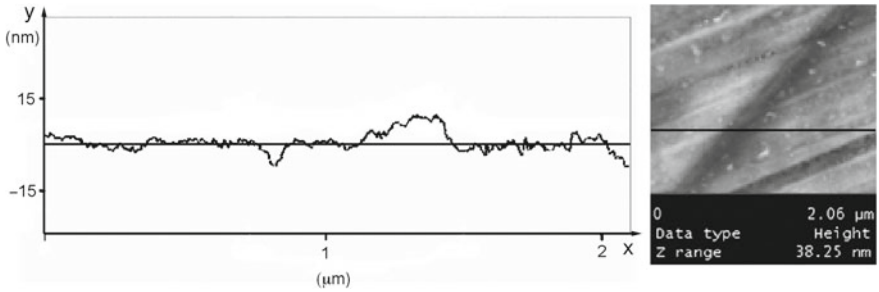


Fig. 23.2 AFM topographic profile of a mirror-like metallic surface, obtained by fine mechanical polishing with a suspension of 1 μm diamond particles

topographic profile (tapping mode) of a finely polished metallic surface (mirror-like finishing). The irregularities, or deviations from the nominal ideal flat surface, form the surface topography. The surface topography can present different features (such as flaws, pores, waviness, etc.). However the contact mechanics of solids depends essentially on the shorter wavelengths fluctuations [13, 31–33] the so-called roughness. Roughness also influences the dynamics (and vibrational modes) of the contact region of rubbing or near-rubbing sliding bodies (such has hard disk drive sliders) [34] and, of course, the wear response of the material [35].

The surface roughness has a fractal nature, in the sense that it is formed of a large number of length scales superimposed on each other [36]. A wide number of roughness parameters and parametric functions can be used to describe and quantify surface roughness (a rather exhaustive presentation can be for instance found in [33]). The most commonly used are roughness amplitude parameters (vertical descriptors), such as the average roughness, R_a, and the average maximum roughness, R_z, defined as:

$$R_a = \frac{\int_0^{l_m} |y| \cdot dx}{l_m}, \tag{23.1}$$

in which l_m is the profile length, and y is the profile function in an axis whose origin is at the centreline of the profile (Fig. 23.2), and:

$$R_z = \frac{1}{5} \sum_{i=1}^{5} y_i, \tag{23.2}$$

with y_i the maximum distance peak to valley in the i^{th} interval of the profile.

23.2.3 Topographic Mechanisms of Wear

Empirical experience shows that, in general, smoother surfaces in rubbing contact present lower wear than rougher ones, either in dry or lubricated contact conditions, and that the running-in period also decreases with increasing smoothness of surfaces (e.g. [35, 37, 38]).

An explanation for this behaviour can be found in the early work of Greenwood-Williamson (GW) model Greenwood and Williamson [13] on the contact mechanics of rough surfaces. The basic idea of the GW model is that, since contact of rough surfaces occurs between the small asperities of opposing surfaces, the real area of contact, A_R, is different from the apparent area of contact, A_N. According with the GW model, A_R is independent of A_N and the ratio $A_R/A_N = f$ can be rather small (typically between 10^{-5} and 10^{-1}).

Since the height of the asperities is not uniform (see Fig. 23.2), when the surface is submitted to a normal load F_N, the touching asperities are under different stresses, but it can be shown that the average stress in the touching asperities, $\langle\sigma\rangle$, is given by [39]:

$$\langle\sigma\rangle = \frac{\sigma_N}{f}, \tag{23.3}$$

in which $\sigma_N = F_N/A_N$, is the normal (compressive) stress in the bulk material, in a plane parallel to the surface.

Equation (23.3) shows that even small elastic compressive stresses in the bulk can lead to high average stresses in the contact spots of the surfaces, since f can be rather small. Moreover, some of the asperities (the higher amplitude ones) can be in plastic regime while the others (the smaller amplitude ones) can be in elastic regime [39].

Although Greenwood and Williamson did not discuss the mechanism of creation of a wear particle, they pointed out that wear is much more probable in touching plastic asperities than in elastic ones. Therefore, they proposed a plasticity index, Ψ, given by:

$$\Psi = \frac{E^*}{H}\sqrt{\frac{R_p}{\rho}}. \tag{23.4}$$

R_p and ρ are topographical parameters: the standard deviation of the asperities height distribution function and the curvature tip radius of the asperities, respectively ($R_p \approx 1.25R_a$, for a gaussian height distribution [33]). E^* is the reduced Young modulus given by:

$$E^* = \left(\frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2}\right)^{-1} \tag{23.5}$$

with ν_i and E_i the Poisson and Young modulus, respectively, of each one of the surfaces. H is the hardness of the softer surface (see points 3 and 4 of this chapter for a more detailed discussion on the concept of hardness and hardness at different scales, respectively).

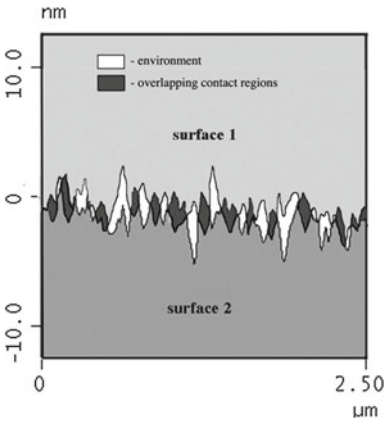


Fig. 23.3 Simulation of the contact regions of two surfaces with a roughness similar to that in Fig. 23.2

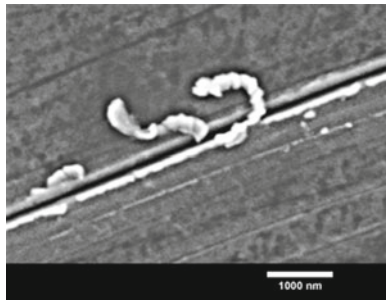


Fig. 23.4 Formation of the wear debris in a Cr tool steel when submitted to abrasion by hard 6 µm diamond particles. Typical wear coefficients obtained in such experiments are between 10^{-12} and 10^{-14} N/m^2 [46]

For $\Psi < 0.6$ the contacts are predominantly elastic and for $\Psi > 1$ the contacts are predominantly plastic. For in-between values of the plasticity index, part of the asperities in contact are in elastic regime and part in plastic regime. The most interesting thing about Ψ is its independence of load: it only depends on mechanical properties of the surface (the ratio E^*/H is inversely proportional to a yield extension) and of its topographic characteristics.

Only very finely polished surfaces have plasticity indexes smaller than 0.6 [40]. As a consequence, most of the surfaces in contact rubbing engineering components have a "mirror-like" finishing, i.e., they are finely polished, in order to decrease Ψ, bringing the contacts as much as possible to the elastic region. The idea is that this will result in a shorter running-in period and in a lower wear rate. It should be pointed out that, in a mirror-like surface, such as the one whose topographical profile is shown in Fig. 23.2, the maximum peak-to-valley amplitude is typically smaller than 50 nm and R_a is typically smaller than 10 nm.

In such a type of surfaces, very common in dry or lubricated moving engineering parts, the contacts will occur at nanometric ranges, as schematically shown in Fig. 23.3. Even in 3-body abrasive wear situations, with free hard micrometric particles moving in-between well polished surfaces, the wear damage and debris can be in the nanometric range. Figure 23.4 shows a scanning electron microscopy (SEM) image of the initial debris formed in a finely polished tempered tool steel surface,

submitted to a common laboratorial ball-cratering abrasive wear test, with 6 μm diamond abrasive particles [41]. It can be observed that the typical dimensions of the scratch caused by the abrasive diamond particles (as well as that of the debris formed) are smaller than 150 nm.

The point being made is that, although up to now the studies concerning wear at nanometric scales has been performed mainly in electronically relevant materials (see Sect. 23.5 of this chapter), the wear mechanisms at this scales are certainly of a much wider relevance than that restricted to the wear of electronical and ultra-precision components.

23.3 Wear Theories

23.3.1 Classical Wear Theories

Most probably Holm [42] was the first to discuss the precursor mechanisms of a wear in his analysis on the relations between contact area and conductance. Later on, J. F. Archard, based on some of Holm's ideas, developed a simple mathematical model for wear caused by adhesion between two opposing asperities [43]. The Archard wear equation for plastic contacts (usually known simply by "Archard equation") is similar to that proposed by Holm, but Archard replace the Holm's concept of "removal of atoms" by the concept of "removal of wear particles". So, in fact, the first model for wear was an atomic wear model. We will come to atomic wear later on. For now let us focus on the "classical" wear theories for elastoplastic dominated wear, the most common form of wear in engineering applications [44].

In spite of its simplicity, the Archard model for "adhesive" wear is still widely used (sometimes outside of its developing context). The model considers two opposing asperities of materials with similar mechanical properties contacting during sliding. At zero time, the location of the two areas forming the contact is fully established (as shown in Fig. 23.5). A short time later the contact area is reduced to zero but, in addition, it is assumed that, at this moment, a new similar contact area as just been

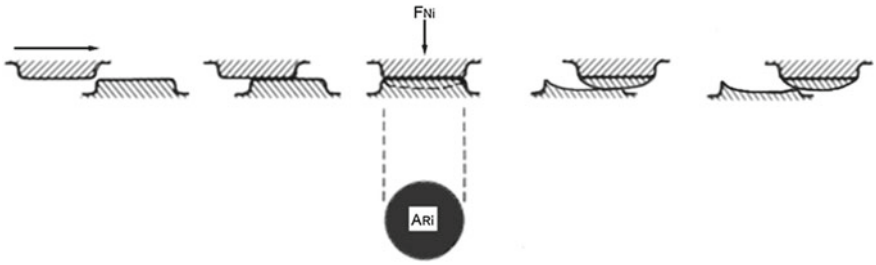


Fig. 23.5 Scheme of the Archard model for adhesive wear

fully established somewhere in the surface. In each asperity contact, chemical interbonding and diffusion could lead to local sticking. The continuous relative motion can, thereof, result in a separation of the asperities at a surface different from the original (Fig. 23.5), leading to surface damage and wear by material transfer.

Archard considered that the volume of material transferred from one asperity to another per unit of sliding length, Q_i, is simply proportional to the contact area when the contact is fully established, A_{Ri}:

$$Q_i = \frac{V_i}{L} = k_{ad} A_{Ri}. \tag{23.6}$$

Although it can be interpreted in various ways (e.g. see [45] for a detailed discussion), the proportionality factor, k_{ad}, is related with the probability of touching asperities adhere and generates a wear particle,

As previously discussed (Sect. 23.2), the most probable wear mechanism is that the contacting asperities are plasticized and wear occurs by the removal of material lumps. If the asperities are fully plasticized, when submitted to a force F_{Ni}, than A_{Ri} results directly from Tabor's definition of hardness [46]:

$$A_{Ri} = \frac{F_{Ni}}{H}, \tag{23.7}$$

with H the hardness of the material.

From (23.5) and (23.6) it can be obtained:

$$Q_i = k_{ad} \frac{F_{Ni}}{H}. \tag{23.8}$$

Assuming that all the contacts are fully plasticized ($\Psi > 1$), the total wear rate is simply:

$$Q = \sum Q_i = \frac{k_{ad}}{H} \sum F_{Ni} = k_{ad} \frac{F_N}{H}. \tag{23.9}$$

Equation (23.8) was deduced for fully plastic contact conditions but Archard proposed the more general relation:

$$Q_i \propto F_N^n, \tag{23.10}$$

with n = 0.6, 0.75, 0.8 or 1, depending if the contacts are elastic or plastic.

However, the subsequent work of Greenwood and Williamson [13] showed that, for multi-asperity contacts, A_R is always directly proportional to F_N, independently if the contacts are elastic or plastic. Assuming this, the direct proportionality between wear rate and load stands not only for fully plastic contact conditions. In this way, a more general form of 23.8, is simply:

$$Q = k_{ad} \frac{F_N}{\phi}. \tag{23.11}$$

ϕ is the hardness, H, if the contacts are fully plasticized, and an "elastic contact hardness" if the contacts are in elastic regime. For gaussian surfaces this elastic contact hardness is given by [13]:

$$\phi = 0.25 E^* \sqrt{\frac{R_d}{\rho}}. \tag{23.12}$$

A slightly different situation is that of the contacting surfaces with much different chemical and mechanical properties. In particular, if the hardness of one material is higher than that of the other (typically $H_1 > 1.2H_2$ as a rule of thumb [9]), the harder asperities will indent the softer ones. The relative motion will lead to the formation of a scratch (as shown in Fig. 23.4), resulting in "abrasive" wear of the softer material. This type of wear damage is not caused by adhesion, but either by mechanical deformation and cutting, i.e., by abrasion. Therefore, the Archard model does not stand for "abrasive" wear.

Rabinowicz model [44] proposed a different approach for abrasive wear. Assuming a conical asperity of the harder surface, infinitely rigid, that, when submitted to a load F_N indents the softer surface, which become fully plasticized (Fig. 23.6), the

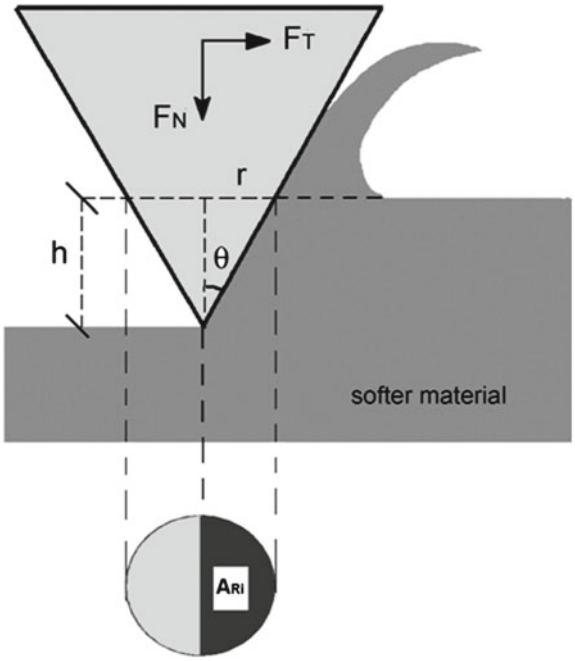


Fig. 23.6 Scheme of the Rabinowicz model for abrasive wear

equilibrium condition in the load application axis is given by:

$$F_N = H \cdot A_R = H\frac{\pi r^2}{2} = H \cdot \frac{\pi}{4} \cdot tg^2\alpha \cdot h^2, \tag{23.13}$$

with r, h and α defined in Fig. 23.6.

If all the material displaced from the groove is removed, than the worn volume, V, is given by:

$$V = L \cdot r \cdot h = L \cdot h^2 tg\alpha. \tag{23.14}$$

From (23.12) and (23.13)

$$Q = \frac{V}{L} = k_{ab}\frac{F_N}{H}, \tag{23.15}$$

with $k_{ab} = 4/(\pi \cdot tg\alpha)$, or less if the material displaced from the groove is not totally removed (e.g. see [28] or [47] for interesting discussions on this topic).

By comparing (23.8), (23.9) and (23.14), it can be concluded that the Rabinowicz model for abrasive wear is formally identical to Archard's model for adhesive wear. Moreover the equations are identical to that obtained by Holm for "atomic wear" [42]. Apparently this can be a quite surprising result since the phenomenology of each type of wear is different. But, in fact, all these approaches for elasto-plastic dominated wear lay on the same basic idea: the worn volume is proportional the real contact area between the tribological pair. This real contact area is proportional to the ratio F_N/H, for fully plastic contacts, or to the ratio F_N/Φ, with Φ defined in (23.11), for GW gaussian multiple elastic contacts. Therefore, a more general equation for elasto-plastic wear could simply be written as:

$$Q = k_{ep}.A_R, \tag{23.16}$$

with k_{ep} a general elasto-plastic wear coefficient, corresponding to a function dependant of the active abrasive or adhesive wear mechanism, i.e. $k_{ep} = f(k_{ad}, k_b)$.

23.3.2 Atomic Wear Theories

From the simple formulation given by equation (23.15), it does not result any discontinuity between wear occurring at atomic/nanometric scales and at larger ones: if the real contact area decreases to nanometric scales the wear rate decreases proportionally (has already mentioned, Holm's "removal of atoms" theory [42] lead to an equivalent equation). However, as the contact scales decreases to near-atomic lengths, the wear mechanisms involving extensive plastic deformation and crack propagation, such as grooving, delamination or fatigue, cannot be predominant if

only a few atomic layers are involved in the contact. Molecular dynamics simulations in contact scales involving a few atomic layers shows that at this small scale lengths, wear is essentially controlled by the dragging of atoms from their initial positions [48, 49]. Bassani and D'Acunto [50] developed a theoretical approach for the atomic transfer between a flat surface and an AFM tip which is interesting to analyse here. The used approach is a double-well potential model, in which the wear volume is quantified in terms of the atomic transition between the two well minima. According with this model, the total wear volume per unit of time, V_t (i.e., the total volume of the atoms jumping from the flat surface to the tip surface in a unit of time) is given by:

$$V_t = n.\omega.V_{at}, \tag{23.17}$$

in which n is the initial atomic population, ω is the transition rate, and V_{at}, the atomic volume. The transition rate is given by an Arrhenius law:

$$\omega = \omega_o \exp\left(-\frac{\Delta U}{kT}\right), \tag{23.18}$$

in which ΔU is the energy barrier height between the two well minima. Later on, D'Acunto [51, 52] extended the model for the quantification of wear mechanisms at the nanoscale. It was proposed that the atomic wear rate (i.e., wear in which the active mechanisms are dragging of atoms (or ionic pairs) from their original positions) could be separated in two basic mechanisms: adhesion and abrasion. At this point it should be emphasized that the concept of "adhesive" and "abrasive" wear at the atomic scale does not have exactly the same meaning as defined previously, when the Archard and Rabinowicz models were presented. At atomic scales, adhesive wear was used for the cases in which the vertical (van der Walls) forces are predominant and the atoms transfer from one surface to another. Conversely, atomic abrasive wear was used for the case in which shear forces are predominant, leading to the dragging of atoms to further positions from the primitive ones (in this case, the atoms does not jump from one surface to another). Figure 23.7a and b shows schematically atomic adhesive and abrasive wear concepts.

The wear volume per unit of time defined in (23.16) can be transformed in a wear rate, simply by dividing by the scanning speed of the asperity / tip, V_s. If that is done, D'Acunto's model lead to the following expression for the wear rate at the atomic scale:

$$Q = \frac{n.\omega.V_{at}}{V_s} \tag{23.19}$$

By noting that n can be given by $A_R/A_{R/n}$, with $A_{R/n}$ the specific area per contact atom, (23.18) can be rewritten in the equivalent form:

$$Q = k_{at} A_R, \tag{23.20}$$

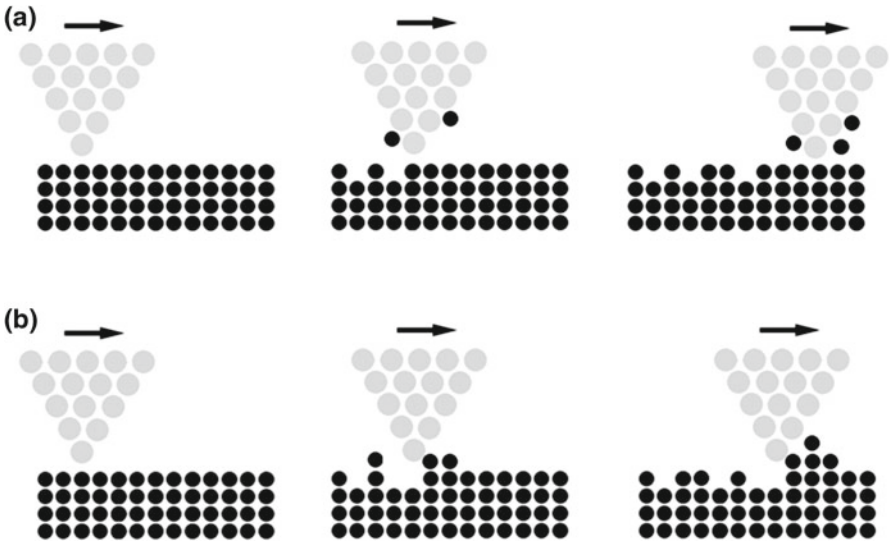


Fig. 23.7 Scheme of atomic adhesive wear **a** and atomic abrasive wear **b**

with

$$k_{at} = \frac{\omega . V_{at}}{A_{R/n} . V_s}. \tag{23.21}$$

Although the formal similarity between (23.19) and (23.15), this result is different from what is obtained for elasto-plastic dominated wear occurring at bulk scales. According with theories for abrasive and adhesive wear at bulk scales that we have previously seen, the wear rate is independent of the scanning speed, if the wear mechanism does not change within a certain range of speeds (e.g. see Lim and Ashby approach to sliding wear [53]). Therefore an important result arising from this model is that atomic scale wear is inversely proportional to the travelling speed of the counterbody. This is caused by the fact that, if speed increases, the interaction time for atom jumping decreases. Experimental observations for atomic wear of ionic crystals [23] do not show a significant velocity dependence of atomic wear as expected from the model. So, at the contact scales involving only a few atomic layers, wear is essentially controlled by the dragging of atoms from their initial positions and, consequently the rate of material removal drops several orders of magnitude as compared with the wear rates in situations involving extensive plastic deformation and/or crack propagation. But, noticeably, the analytical models for "atomic wear", lead again to an expression similar to that of elasto-plastic driven wear (23.15), but this time the proportionality factor, k_{at}, depends on the volumes of the atoms, molecules or ions that are being "worn", on the transition rates of these particles

between energy wells, and on the relative speed of the matching surfaces with, of course and $k_{at} \ll k_{ep}$.

Presently, atomistic methods, that allow for simulations of up to millions of atoms in time scales of microseconds, are still insufficient to provide more extensive links to experimental data. Particularly, in the scale lengths of a few hundreds of nanometers there is a gap of understanding, since it is one too large scale for atomistic simulations and a too small scale for continuum mechanics theories and for conventional wear tests. It is expected that, in practice, the transition from elastoplastic-driven wear mechanisms, described by (23.15), to "atomic wear" mechanisms, described by (23.19), should correspond to the threshold to measurable wear conditions [54]. We will come again to this topic in the following chapter.

23.4 Wear Experiments at Submicrometric Scales Using the AFM

The advent of the AFM invention [14] has provided tribologists with a new tool that enable to perform experiments from atomic scales to (almost-) bulk ones. Up to now, an increasing number of studies have been dealing with this topic showing that that observable tip-induced wear in the sample (and also wear of the tip itself [20, 55, 56]) can occur when the AFM tip slides in contact with the sample, above some threshold load [57].

Tip-induced nanowear experiments have been performed (in UHV and environmental conditions) in a relatively large range of materials, such as thin films of AgBr and C60 deposited on NaCl [58], NaCl [59], KBr [23, 60, 61], PZT's [20], polymeric magnetic tapes [62], silicon [63, 64], SiO_2 [64], mica [65, 66], thermoplastic and thermoset polymers [67–69], among other materials. Fewer experiments have been performed in metals, maybe because, as pointed out by Gnecco et al. [60], metals are not the best candidates to study wear mechanisms by AFM, since the debris tend to accumulate in the tip, leading to a small reproducibility of measurements. Still, some results on gold [70], MnZn ferrites [71], Cu [63] and Cr thin films [72] can be found in the literature and, facing experimental difficulties in nanowear studies of metals and metallic alloys, will be certainly a challenge.

One of the first in situ observations of the transition threshold to detectable wear in AFM experiments was made by Hu et al. [65]. The experiments were performed in mica, with silicon nitride probes, in air, water and ethanol environments. While monitoring the frictional forces, these authors observed that, once a certain critical load threshold was reached, the wear of mica (remotion of an atomic layer) was responsible for a transition from a linear relation between friction and load to a stochastic behaviour (Fig. 23.8). The experiments showed that the wear of this layer of mica occurs both at high loads, in a single scan, or at lower loads, in multiple scans. This result strongly indicates that the load onset for atomic wear of mica has somehow a memory effect to the number of scans of the counterface slider. Hu et al. [65]

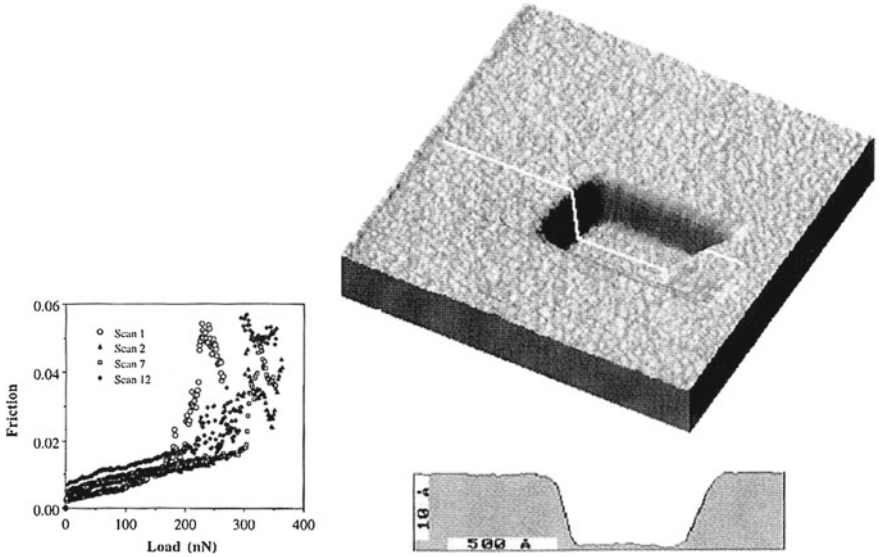


Fig. 23.8 Friction force (au) versus load in the high load regime for multiple scans of a silicon nitride tip in contact with mica. Friction increases smoothly with load until wear occurs. At this time abnormal changes in the friction behaviour are observed [93]

suggested that, in the lower load regime, point defects were accumulating during the each scan. When the number of scans reach a critical value, this accumulation will result in the formation of a small cluster, resulting from the removal of an atomic layer. Helt and Bateas [73] were able to observe the nucleation of defects in muscovite mica under aqueous environments, prior to gross wear, confirming the nanowear mechanism proposed by Hu et al.

Gnecco et al. have performed abrasive wear experiments at nanoscale on (001) KBr [23]. They have shown that, under UHV conditions, the atomic scale wear mechanism is due to the removal of single ion pairs. Also they have shown that this debris is reorganized in regular terraces with similar structure and orientation of the unscratched surface. More recent experiments performed by the same group [61], also confirms the cumulative nature of wear at atomic scales. In fact they observe that, even at very low loads (between 1.7 and 30.1 nN) scratch of KBr with silicon probes in UHV conditions, always result in groove formation, after a sufficient number of scans. Concomitantly, an irregular ripple formation in the periphery of the grooves was observed. This ripple-like structure appears after the onset of wear has been reached and has a periodicity that matches well with the scale length of the tip radii. In this way, the authors attributed the KBr wear mechanism at atomic scale as an accumulation in front of the tip of ionic pairs or small clusters. Ripple arises when the material transported by the tip increases friction and the tip jumps over it as the process restarts again.

The experimental results obtained up to now seam to indicate that, in general, the precursor mechanisms for atomic-scale wear is the formation of point defects that lead to the formation of small clusters, the nanodebris. It is interesting to note that these observations matches quite well with the theoretical double-well approach presented in Sect. 23.3.2. In fact, according to this model, in each interaction between surface and counterface atoms, there is a probability of atomic jumps to occur, to the opposite surface (adhesion) or to another position in the same surface (abrasion). These jumps will result in the formation of point defects.

In this way, both experimental and theoretical results seam to confirm that, in general, the formation of point defects in sliding contacts is the catalyst for atomic wear. This mechanism is predominant in situations in which the contact loads are below the yield onset for plastic deformation or crack nucleation.

In a small, but relatively larger contact scale, Wang and Kato [54, 74] made a set of interesting wear experiments inside the chamber of an E-SEM (environmental scanning electrical microscope). The experiments were performed with a pin-on-disk geometry in carbon nitride coatings and in bare Si(111). A diamond pin with a 10 μm curvature radius and a load range between 10 and 250 mN were used. One of the aims of the experiments was to observe the onset of wear from "no observable wear particles" to "wear particle formation". This transition was defined when the formation of cluster of worn particles, larger than 0.25 μm, could be confirmed by in-situ E-SEM observation. The worn particles were described as "feather-like", for the mild wear regime, and "plate-like", for the higher wear regime.

Figure 23.9 resumes the results obtained for the carbon nitride coatings. It can be observed that the transition from "no observable wear particles" to "wear particle formation" depends on the load (as expected). However, the transition also occurs,

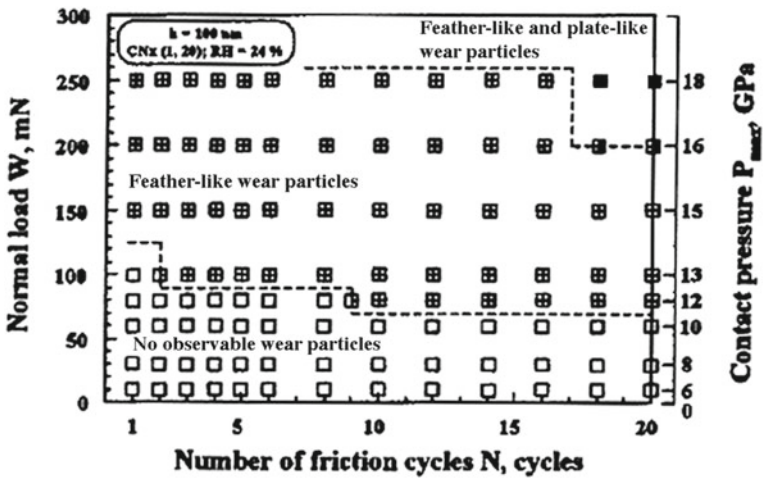


Fig. 23.9 Occurrence of wear particles of a carbon nitride coating, in terms of load and friction cycles, based on E-SEM observations [102]

in certain load conditions, when a critical number of cycles is reached. The authors have attributed this wear transition dependence of the number of cycles to a low cycle fatigue wear mechanism [74]. However, and in spite of the higher loads used in Wang and Kato's experiments, it is worth to mention the similarities with the observations for the nanowear onset of mica previously mentioned [65]. Moreover, Wang and Kato [54] have observed, from AFM measurements, that some wear exists even in the "no observable wear particles" regime.

As previously mentioned, when the load increases, resulting in higher depth interactions, other wear mechanisms become active. Also the chemical and structural composition of the surface interlayers can change (Fig. 23.1). As a consequence, the wear response of the material can be quite different and, moreover, unpredictable, depending on the contact scale. For now, it seems that the only choice for tribologists is to carefully choose their test conditions for them to be representative, in view of the application. The results obtained by Miyake et al. [19] for diamond and N^+-implanted diamond films clearly illustrate this point. These authors have observed that nanowear AFM tests, with loads in the range of μN, lead to the lowest wear resistance in N^+-implanted diamond films, as compared with diamond films. By the contrary, N^+-implanted films present the highest wear resistance in reciprocating wear tests, with loads in the range 0.098 to 4.9 N. This difference was attributed by the authors to the formation of a thin amorphous layer in the N^+-implanted samples (as observed by Raman spectroscopy), whose properties are relevant for AFM wear tests but not for the wear tests at higher loads.

Chung and Kim [63] have presented a very interesting summary of the results obtained for low load wear experiments, in materials for MEMS and ultra-precision electronical applications. The data was obtained by different authors, using low load pin-on-disk and AFM tests in Si-based materials, carbon films and DLC's. The load ranges used changes from nN to mN. Table 23.1 resumes the results for silicon, as reviewed by Chung and Kim. These results show that, although a general perception of the basic small-scale wear mechanisms starts to exist, extreme difficulties in quantifying wear still remain. Clearly, advances can only be made if precise and reproducible experiments are available to test theoretical models [12].

More recently it was shown [75] that there is a wear threshold between atomic wear and elasto-plastic dominared wear that can be easily identified by AFM nanowear experiments. Once the threshold for wear is overcome, the wear rate increases with

Table 23.1 Wear coefficients for Si obtained in tests performed at different loads (after [63])

Material	Load	Test	Wear coefficient (as defined in 23.14)	References
Si (100)	10–30 μN	Specimen-on-disk	10^{-5}	[107]
Si (100)	100 μN	AFM	10^{-1}	[108]
Si (111)	100 mN	Pin-on-disk in SEM	10^{-4}	[54]
Si	100 mN	Pin-on-disk	10^{-3}	[109]

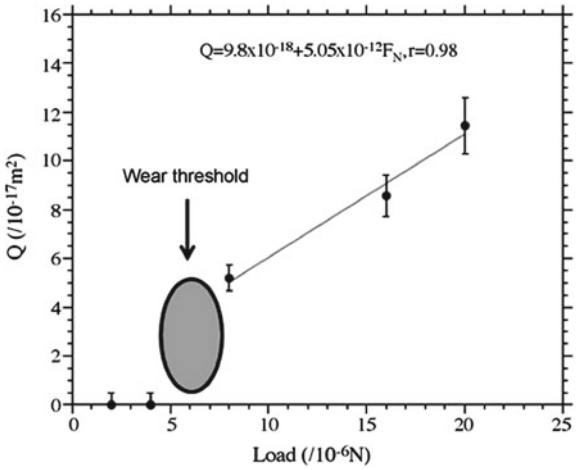


Fig. 23.10 Regression over the wear rates after the wear threshold is overcome (after [75])

load following 23.15 (Fig. 23.10). By using the slope obtained from regression over the experimental points of Fig. 23.10, and the hardness of the material the dimensionless wear coefficient of this material, k_{ab} defined in (23.14), could be determined as 0.01. Such a value is within the range of the dimensionless abrasive wear coefficients for a two-body wear situation obtained in macroscale tests, typically between 0.005 and 0.05 [9]. This result shows that at this single contact scale provided by the AFM tip, once the wear onset is overcome, the Rabinowicz wear equation (23.14), or a more general equation like (23.15), is followed. Moreover, the fact that the values of k_{ab} obtained from AFM abrasive wear tests are consistent with those retrieved from macrowear tests indicate that, most probably, the results obtained form nanoscale wear tests can be scaled up, once the wear mechanism does not change.

These experimental results, however, show clearly that the Rabinowicz equation is followed only after a certain critical value of the load is overcome. Bellow this critical value, changes in the surface can only be detected by LFM observation [1], but these changes never lead to well-defined worn craters. This means that, bellow this critical value, the surface damage mechanism is not elasto-plastically dominated, i.e., there is a load threshold between "macro-wear" and "atomic wear".

At least in metallic materials, what is really interesting in this threshold is that it occurs at a much higher values than those that can be expected. In fact, if the values of the average or maximum Hertzian contact pressures are compared with the measured hardness of the material, or with its indentation yield stress, σ_y, estimated from Tabor relation [46], one can conclude that the plastic yield limit of the material was largely overcome [75]. This result confirms the importance of scale related factors in the tribomechanical properties of materials, in particular at the contact scale of the asperities. Recent works have shown that ductile metallic alloys can

increase their hardness up to 10 times, when the contact depth decreases bellow 100 nm, due to strain gradient plasticity and surface free energy related effects. This will be the topic of the next paragraph.

23.5 Indentation Size Effect

Hardness is a measure of yield stress of the material [46]. The appearance of this property in Archard and Rabinowicz wear models (23.8 and 23.13, respectively) is related with the fact that hardness is a measure of simultaneously two things:

1—the resistance of the asperities to the start of the dissipative plastic deformation processes, which can result in the formation of wear particles and,

2—the depth of the wear damage, e.g., the depth of the groove in the simple case of pure abrasive wear [76].

If the interaction between body and slider is only of a few atomic layers, the elasto-plastic deformation processes are inactive, and the nature of wear is distinct, , i.e., we are in the presence of "atomic wear". Once this threshold is overcome we are in presence of elasto-plastic driven wear as discussed in Sect. 23.4. Hardness as a measure of the elasto-plastic yeld of the material can be a simple measure of this threshold.

However, there is a length scale range, lets say for a few atomic layers to submicrometric lengths in which, although the deformation mechanisms are already still active, the material mechanical response can be different from its bulk mechanical response. In recent years there has been a growing interest in the use of molecular dynamics simulations (MD) methods to investigate nanoindentation [11]. Landman et al. [48], carried out one of the first MD simulations of a Ni tip nanoindentation into a gold surface. After, MD have been used in several studies of nanoindentation in different materials (e.g. [77–80]). Although, at the present, atomistic methods allow for simulations of up to millions of atoms in time scales of microseconds and provide information about the single atom role in the contact [11], the length and time scales are still insufficient to provide more extensive links to experimental data [12]. Specially in the scale lengths of a few hundreds of nanometers there is a gap of understanding, since it is a too large scale for atomistic simulations a too small scale for continuum mechanics theories to be still valid [81].

The fact is that, in this scale range of a few hundreds of nanometers, a large number of experimental works have showed that, in general, an increase of the hardness of the material occurs as compared with its bulk hardness, as measured in conventional micro or macrohardness tests, the so-called *Indentation Size Effect* (ISE). ISE have been observed in a significant variety of metals, metallic alloys and ceramics: Cu [82, 83], Mo and W [84], Ag [85], Al and brass [86], Ni and Co alloys [87], sapphire and MgO [88], TiO_2 and SnO_2 [89], LiF and NaCl [90], among others.

For now, there are various explanations for the observed indentation size effect, including the sample preparation and test methods and the increasing perfection of the materials as the volume is reduced [88]. Of course that, as mention in Sect. 23.2.1, the

constitution of the material changes on approaching the surface (Fig. 23.1). This, most certainly, will lead to different material properties, in particular to different hardness measurements, as the indenter crosses the different layers, schematically represented in Fig. 23.1. In particular, it has been observed that the deformed layer caused by the mechanical polishing of surfaces may contribute to increase the hardness of metallic surfaces [83, 88]. Nevertheless, ISE is still observed in electropolished samples and, in fact, the results strongly indicate that ISE is an intrinsic material length scale related characteristic [91].

Conventional plasticity theories, which are based in continuous mechanics, and in this way do not involve length scale dependency, do not provide an explanation for size dependency of hardness. Up to now, there are two theories that seam to offer the best explanation for ISE: geometrically necessary dislocations (strain gradient plasticity) [92] and surface free energy effects [93, 94]. We will briefly review these two explanations for the indentation size effect based on experimental results recently obtained in our group [95–97].

When a material is deformed dislocations are generated, moved and stored. This storage causes the material to work-harden [98]. Dislocations can be randomly present and stored, forming a network of defects in the crystalline structure of the material and, in this case, they are called statistically stored dislocations (STDs) [99]. In non-uniform deformation, dislocations are necessary for compatibility reasons and, in this case, they are called geometrically necessary dislocations (GNDs) [99, 100].

In the case of indentation experiments, the hypothesis is that plastic strain gradients surrounding the indentation are accommodated by the presence of these geometrically necessary dislocations. Since, in general, strain gradients are inversely proportional to the length scale over which plastic deformation occurs as the indentation depth decreases the density of GNDs must increase, thus leading to an increase of the hardness of the material [101, 102]. Based on this assumptions, Nix and Gao [92], have proposed a model that relates the indentation size effect with the density of GNDs. The Nix-Gao model leads to a simple relation between the variation of the materials hardness caused by the GNDs, ΔH_{GND}, and the indentation depth, h given by [92]:

$$\Delta H_{GND} = H_0 \left(\sqrt{1 + \frac{h^*}{h}} - 1 \right). \tag{23.22}$$

H_0 is the material hardness at an infinite depth (the "bulk hardness"). h^* is a characteristic length that depends on the shape of the indenter and on the tested material, given by:

$$h^* = \frac{3}{2} \frac{\tan^2 \theta}{b \rho_s}, \tag{23.23}$$

where b is the Burgers vector of the dislocations and ρ_s is the density of statistically stored dislocations. θ is the angle between the surface of the material and the indenter.

Surface free energy (SFE) related effects could also contribute to the increase of the hardness with decreasing scale, since the area to volume ratio increases as the scale decreases [93, 94]. According with Jager [94], surface free energy can have a non-negligible contribution to hardness if the three following conditions are simultaneously met: (a) soft ductile samples with high SFE; (b) sharp indenter geometries; (c) shallow indentations. Models accounting for the contribution of SFE to hardness lead to a relation of the type [93, 94]:

$$\Delta H_{SFE} = \kappa \frac{E_s}{h} \tag{23.24}$$

where ΔH_{SFE}, if the hardness variation with depth caused by SFE effects, E_s is the surface free energy, h the indentation depth, and κ a constant which depends of the indenter's geometry (e.g., for a Veeco DNISP diamond AFM tip, $\kappa \approx 3.43$ [87]).

Assuming that both these effects (GNDs and SFE) are independent, and though summative, it results from (23.22) and (23.23), that a general equation for hardness as function of indentation depth, accounting for both effects, has the form [95]:

$$H = H_0 + \Delta H_{GND} + \Delta H_{SFE} = H_0 \sqrt{1 + \frac{h*}{h}} + \left\{ \kappa \frac{E_s}{h} \right\}_{SFE}. \tag{23.25}$$

A typical plot of the hardness of the material according to (23.23) is shown in Fig. 23.11. The plot was calculated with a set of values typical of a metallic alloy, e.g. a steel: $H_0 = 2.5$ GPa, $E_s = 2$ J/m^2. h^* was taken as 1,000 nm [92] and κ as 3.43 [87]. As it can be observed, for indentation depths smaller than some hundreds of nanometers, the combined effect of geometrically necessary dislocations and surface free energy result in a non-negligible increase of hardness. The accuracy of this

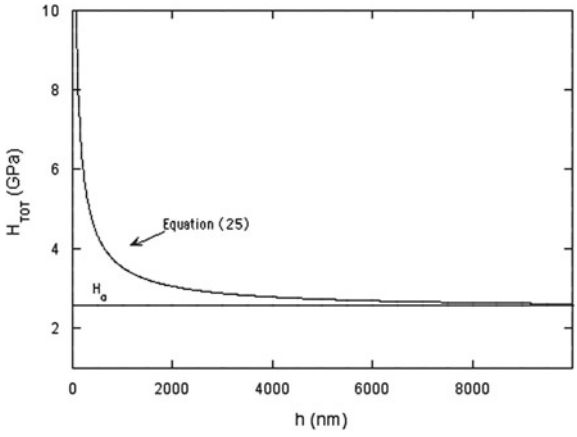


Fig. 23.11 Hardness variation with indentation depth according to (23.25). The plot was calculated with a set of values typical of a metallic alloy, e.g. a plain carbon steel: $H_0 = 2.5$ GPa, $E_s = 2$ J/m^2 and h* = 1,000 nm [75]). κ was taken as 3.43 [70]

therorectical model was subsequently confirmed by Ma et al. [103] and the direct observation of the high density of dislocation pile ups under AFM nanoindentations was recently made [104].

Therefore both experimental and theoretical results seam to lead to the same conclusion: smaller tends to be harder, i.e., as the contact scale decreases to submicrometric ranges the resistance of the material to plastic deformation increases, therefore the threshold from "atomic wear" to "macro" or elasto-plastic driven wear occurs at higher contact loads that the ones that we could expect from hardness based models derived from continuous mechanics theory. But a second factor should also be considered in wear at small contact scales. Is the fact that the parameter H that appears in (23.8) and (23.14) in fact depends on the contact depth, as it results from (23.23). By replacing the depth independent hardness of 23.14 by the expression on 23.23, the contact scale dependence of the hardness can be incorporated into the Rabinowicz equation giving [105]:

$$Q_h = K_{ab} \frac{F_N}{H_0 \sqrt{1 + \frac{h*}{h}} + \kappa \frac{E_s}{h}}, \tag{23.26}$$

Equation (23.24), shows that the depth dependence of hardness in materials presenting ISE (most metallic alloys) results directly in a depth dependence of their wear rate which clearly agrees with published literature experimental results [105, 106].

23.6 Conclusions

As concluding remarks of this chapter, we can say that there has been a great interest in submicrometric wear studies within the last 10–15 years. This studies covers different types of materials, different scale lengths (going from quasi-atomic to quasi-bulk ones) and different research objectives (from generic theoretical studies to well focused application ones). Although the natural dispersion of results, caused by the great complexity of the phenomena involved, and to the youthful nature of the field, significant advances have been made in this recent years.

The models and results obtained up to now suggest that, on approaching atomic scales, a division of mechanisms can be made between "atomic wear" and "bulk wear", i.e., wear involving atomic jumps and point defects formation (the former) and wear involving plastic deformation and crack propagation (the latter).

For atomic wear, experimental results strongly suggest that there is threshold for detectable wear. This threshold is a load threshold but also number of contact cycles threshold. This suggests a cumulative nature for atomic wear, which is in agreement with theoretical models. For higher loads and higher contact interaction depths, the plastic deformation and crack propagation mechanisms will be activated, and the wear can, in principle, be described by classical continuum mechanics based theories. However, it should be noted that there is a scale length interval (from some

nanometers to some hundreds of nanometers) in which, although the deformation and crack propagation mechanisms can be active, important differences as compared with the similar mechanisms occurring at larger scales can occur. These differences can be caused by surface topography related factors, changes in the structural and chemical constitution of the material on approaching the surface and intrinsic effects related with increasing surface to volume ratio (surface free energy, strain gradient plasticity phenomena). Certainly, in the near future, experimental results, MD simulations and the development of scale dependant plasticity theories will bring new insights to this topic, reducing the gap between macro, micro/nanowear.

Acknowledgments The author would like to thank the European Science Foundation and the Portuguese Foundation for Science and Technology for financial support under project ESF Eurocore 07-FANAS-FP-009. Also I thank all the students and colleagues at the Nanolab-IST, namely Dr. Sérgio Graça, for valuable discussions over the years.

References

1. R. Colaço, Surface damage mechanisms: from nano and microcontacts to wear of materials, in *Fundamentals of Friction and Wear on the Nanoscale*, ed. by E. Meyer, E. Gnecco. (Springer, 2007) pp. 453–480
2. J.T. Dickinson, Single asperity nanometer-scale studies of tribochemistry, in *Fundamentals of Friction and Wear on the Nanoscale*, ed.by E. Meyer, E. Gnecco. (Springer, 2007), pp. 480–521
3. H.P. Jost, *Lubrication (Tribology)—A Report of the Present Position and Industry's Needs* (Department of Science and Education, H. M. Stationary Office, London, 1966)
4. H.P. Jost, *Economic Impact of Tribology* (Mechanical Engineering, 1975), pp. 26–33
5. E. Rabinowicz, *Friction and Wear of Materials*, 2nd edn. (Wiley, New York, 1995)
6. Canada, *A Strategy for Tribology in Canada* (National research Council Canada, Canada, 1986)
7. Soc, T.I.C.M.E., *An Investigation on the Application of Tribology in China*. (Tribology Institute of the Chinese Mechanical Engineering Society, China, 1986)
8. J. Krim, Surface science and the atomic scale origins of friction: what once was old is new again. Surf. Sci. **500**, 741–758 (2002)
9. I.M. Hutchings, *Tribology: Friction and Wear of Engineering Materials* (Edward Arnold, London, 1992)
10. B. Bushan, *An Introduction to Tribology* (Wiley, New York, 2002)
11. J.B. Adams et al., Adhesion, lubrication and wear on the atomic scale. Surf. interface Anal. **31**, 619–626 (2001)
12. S.S. Perry, W.T. Tysoe, Frontiers of fundamental tribological research. Tribol. Lett. **19**(3), 151–161 (2005)
13. J.A. Greenwood, Williamson, J.B. Contact of nominally flat surfaces, in *Proceedings of the Royal Society of London*, vol. A295 (1966), pp. 300–319
14. G. Binnig, C.F. Quate, C. Gerber, Atomic force microscope. PRL **56**(9), 930–933 (1986)
15. S. Sundararajan, B. Bhushan, Micro/nanotribology of ultra-thin hard amorphous carbon coatings using atomic force/friction force microscopy. Wear **225–229**, 678–689 (1999)
16. B. Bhushan, Nanos to microscale wear and mechanical characterization using scanning probe microscopy. Wear **251**, 1105–1123 (2001)
17. R. Kaneko, K. Nonaka, K. Yasuda, Scanning tunneling microscopy and atomic force microscopy for microtribology. J. Vac. Sci. Technol. A Vac. Surf. Films **6**(2), 291–292 (1988)

18. Z. Jiang et al., An investigation of the experimental conditions and characteristics of a nanowear test. Wear **181–183**, 777–783 (1995)
19. S. Miyake, T. Miyamoto, R. Kaneko, Increase of nanometer-scale wear of polished chemical-vapor-deposited diamond films due to niytogen ion implantation. Nucl. Instrum. Methods Phys. Res. B Beam Interact. Mater. Atoms **108**, 70–74 (1996)
20. K.-H. Chung et al., Tribological characteristics of probe tip and PZT media for AFM-based recording technology. IEEE Trans. Magn. **41**(2), 849–854 (2005)
21. S. Graça, R. Colaço, R. Vilar, Using atomic force microscopy to retrieve nanomechanical surface properties of materials. Mater. Sci. Forum **514–516**, 1598–1602 (2006)
22. A.R. Machcha, M.H. Azarian, F.E. Talke, An ivestigation of nano-wear during contact recording. Wear **197**, 211–220 (1996)
23. E. Gnecco, R. Bennewitz, E. Meyer, Abrasive wear on the atomic scale. Phys. Rev. Lett. **88**(21), 215501/1-215501/4 (2002)
24. W. Gulbinski, T. Suszko, D. Pailharey, High load AFM friction and wear experiments on V2O5 thin films. Wear **254**, 988–993 (2003)
25. J.Y. Park et al., Friction and adhesion properties of clean and oxidized Al-Ni-Co decagonal quasycrystals: a UHV atomic force microscopy/scanning tunneling microscopy study. Tribol. Lett. **17**(3), 629–636 (2004)
26. J. Drelich, G.W. Tormoen, E.R. Beach, Determination of solid surface tension from particle-substrate pull-off forces measured with the atomic force microscope. J. Colloid Interface Sci. **280**, 484–497 (2004)
27. A. Schirmeisen, Wear: one atom after the other. Nat. Nanotechnol. **8**(2), 81–82 (2013)
28. K.-H.Z. Gahr, *Microstructure and Wear of Materials* (Elsevier Scientific Publishing Company, Amsterdam, 1987), p. 560
29. H. Luth, *Surfaces and Interfaces of Solids* (Springer, Berlin, 1993)
30. F.P. Bowden, A.J.W. Moore, D. Tabor, The ploughing and adhesion of sliding metals. J. Appl. Phys. **14**, 80–91 (1943)
31. K.L. Johnson, K. Kendall, Roberts, D. Surface energy and the contact of elastic solids, in *Proceedings of Royal Society London*, vol. A 324 (1971), pp. 301–313
32. K.N.G. Fuller, D. Tabor, The effect of surface roughness on the adhesion of elastastic solids, in *Proceedings of Royal Society London*, vol. A 345 (1975), pp. 327–342
33. T.R. Thomas (ed.), *Rough Surfaces* (Longman, London and New York, 1982)
34. L. Zhou et al., Slider vibration reduction using slider surface texture. Microsyst. Technol. Micro-Nanosyst. Inf. Storage Process. Syst. **11**(8–10), 857–866 (2005)
35. T. Hisakado, T. Tsukisoe, Effect of surface-roughness on transient wear. J. Jpn. Soc. Lubr. Eng. **21**(4), 228–235 (1976)
36. S. Ganti, B. Bhushan, Generalized fractal analisys and its application to engineering surfaces. Wear **180**, 17–34 (1995)
37. H.H. Gatzen, M. Beck, Wear of single crystal silicon as a function of surface roughness. Wear **254**, 907–910 (2003)
38. X. Wang, K. Kato, K. Adachi, Running-in effect on the load-carrying capacity of a water-lubricated SiC thrust bearing. Proc. Inst. Mech. Eng. Part J J. Eng. Tribol. **219**(J2), 117–124 (2005)
39. M.A. Fortes, R. Colaço, M.F. Vaz, Contact mechanics of cellular solids. Wear **230**, 1–10 (1999)
40. D. Tabor, Lubrication and wear. in *Surface and Colloid Science*, ed. by E. Matijevic (Wiley, New York, 1972), pp. 245–312
41. R. Colaço, R. Vilar, On the influence of retained austenite in the abrasive wear behaviour of a laser surface melted tool steel. Wear **258**(1–4), 225–231 (2005)
42. R. Holm, *Electrical Contacts* (H. Gerber, Stockholm, 1946)
43. J.F. Archard, Contact and rubbing of flat surfaces. J. Appl. Phys. **24**(8), 981–988 (1953)
44. E. Rabinowicz, *Friction and Wear of Materials*, 2nd edn. (Wiley, New York, 1965)
45. D.A. Rigney, Some thoughts on sliding wear. Wear **152**, 187–192 (1992)
46. D. Tabor, *The Hardness of Metals. Oxford Classic Texts* (Clarendon Press, Oxford, 1951)

47. K. Kato, Abrasive wear of metals. Tribol. Int. **30**(5), 333–338 (1997)
48. U. Landman, W.D. Luedtke, E.M. Ringer, Atomistic mechanisms of adhesive contact formation and interfacial processes. Wear **153**(1), 3–30 (1992)
49. J.A. Harrison et al., Effect of atomic-scale surface-roughness on friction—a molecular-dynamics study of diamond surfaces. Wear **168**(1–2), 127–133 (1993)
50. R. Bassani, M. D'Acunto, Nanotribology: tip-sample wear under adhesive contact. Tribol. int. **33**, 443–452 (2000)
51. M. D'Acunto, Wear and diffusive processes. Tribol. int. **36**, 553–558 (2003)
52. M. D'Acunto, Theoretical approach for the quantification of wear mechanisms on the nanoscale. Nanotechnology **15**, 795–801 (2004)
53. S.C. Lim, M.F. Ashby, Wear-mechanism maps. Acta Metallurgica **35**(1), 1–24 (1987)
54. D.F. Wang, K. Kato, Nano-scale fatigue wear of carbon nitride coatings: part I-wear properties. J. Tribol. Trans. ASME **125**, 430–436 (2003)
55. A.G. Khursudov, K. Kato, H. Koide, Wear of the AFM diamond tip sliding against silicon. Wear **203–204**, 22–27 (1997)
56. K.-H. Chung, Y.H. Lee, D.-E. Kim, Characteristics of fracture during the approach process and wear mechanism of a silicon AFM tip. Ultramicroscopy **102**, 161–171 (2005)
57. R.W. Carpick, M. Salmeron, Scratching the surface: fundamental investigations of tribology with atomic force microscopy. Chem. Rev. **97**, 1163–1194 (1997)
58. R. Lüthi et al., Nanotribology: an UHV-SFM study on thin films of C60 and AgBr. Surf. Sci. **338**(1–3), 247–260 (1995)
59. P.E. Sheehan, The wear kinetics of NaCl under dry nitrogen and at low humidities. Chem. Phys. Lett. **410**(1–3), 151–155 (2005)
60. E. Gnecco et al., Friction and wear on the atomic scale. Wear **254**, 859–862 (2003)
61. A. Socoliuc et al., Ripple formation induced in localized abrasion. Phys. Rev. B **68**, 115416/1-115416/4 (2003)
62. B. Bhushan, Nanotribology and nanomechanics. Wear **259**, 1507–1531 (2005)
63. K.-H. Chung, D.-E. Kim, Fundamental investigation of micro wear rate using an atomic force microscope. Tribol. Lett. **15**(2), 135–144 (2003)
64. B. Bhushan, A.V. Kulkarni, Effect of normal load on microscale friction measurements. Thin Solid Films **278**(1–2), 49–56 (1996)
65. J. Hu et al., Atomic scale friction and wear of mica. Surf. Sci. **327**, 358–370 (1995)
66. S. Miyake, 1 nm deep mechanical processing of muscovite mice by atomic-force microscopy. Appl. Phys. Lett. **67**(20), 2925–2927 (1995)
67. D.D. Woodland, W.N. Unertl, Initial wear in nanometer-scale contacts on polystyrene. Wear **203–204**, 685–691 (1997)
68. S.P. Ho et al., Nanotribology of CoCr-UHMWPE TJR prosthesis using atomic force microscopy. Wear **253**, 1145–1155 (2002)
69. J.S.S. Wong et al., Scratch damage of polymers in nanoscale. Acta Materialia **52**(2), 431–443 (2004)
70. Z.G. Jiang et al., An investigation of the experimental conditions and characteristics of a nano-wear test. Wear **181**, 777–783 (1995)
71. Z.G. Jiang et al., Dependence of nano-friction and nano-wear on loading force for sharp diamond tips sliding on SI, Mn-Zn ferrite, and Au. J. Tribol. Trans. ASME **117**(2), 328–333 (1995)
72. W. Lu, K. Komvopoulos, Nanomechanical and nanotribological properties of carbon, chromium, and titanium carbide ultrathin films. J. Tribol. Trans. ASME **123**(4), 717–724 (2001)
73. J.M. Helt, J.D. Batteas, Wear of mica under aqueous environments: direct observation of deffect nucleation by AFM. Langmuir **21**, 633–639 (2005)
74. D.F. Wang, K. Kato, Nano-scale fatigue wear of carbon nitride coatings: part II-wear mechanisms. J. Tribol. Trans. ASME **125**, 437–444 (2003)
75. R. Colaco, An AFM study of single-contact abrasive wear: the Rabinowicz wear equation revisited. Wear **267**(11), 1772–1776 (2009)

76. R. Colaço, R. Vilar, A model for the abrasive wear of metallic matrix particle-reinforced materials. Wear **254**(7–8), 625–634 (2003)
77. C.L. Kelchner, S.J. Plimpton, J.C. Hamilton, Dislocation nucleation and defect structure during surface indentation. Phys. Rev. B **58**(17), 11085–11088 (1998)
78. J.A. Zimmerman et al., Surface step effects on nanoindentation. Phys. Rev. Lett. **87**(16), 165507 (2001)
79. I. Szlufarska et al., Atomistic mechanisms of amorphization during nanoindentation of SiC: a molecular dynamics study. Phys. Rev. B **71**(17), 174113 (2005)
80. E.T. Lilleodden et al., Atomistic simulations of elastic deformation and dislocation nucleation during nanoindentation. J. Mech. Phys. Solids **51**(5), 901–920 (2003)
81. N.A. Fleck, J.W. Hutchinson, A phenomenological theory to for strain gradient effects in plasticity. J. Mech. Phys. Solids **41**(12), 1825–1857 (1993)
82. K.W. McElhaney, J.J. Vlassak, W.D. Nix, Determination of indenter tip geometry and indentation contact area for depth-sensing indentation experiments. J. Mater. Res. **13**(5), 1300–1306 (1998)
83. Y. Liu, A.H.W. Ngan, Depth dependence of hardness in copper single crystals measured by nanoindentation. Scripta Mater. **44**, 237–241 (2001)
84. N.A. Stelmashenko et al., Microindentations on W and MO oriented single-crystals—an STM study. Acta Metallurgica ET Materialia **41**(10), 2855–2865 (1993)
85. Q. Ma, D.R. Clarke, Size dependent hardness of silver single crystals. J. Mater. Res. **10**(4), 853–863 (1995)
86. A.A. Elmoustafa, D.S. Stone, Indentation size effect in polycrystalline FCC metals. Acta Materialia **50**(14), 3641–3650 (2002)
87. S. Graça, R. Colaço, R. Vilar, Indentation size effect in laser clad nickel and cobalt, Surf. Coat. Technol. **202**(3), 538–548, 2007
88. S.J. Bull, On the origins and mechanisms of the indentation size effect. Z. Metallkd **94**(7), 787–792 (2003)
89. H. Li et al., The frictional component of the indentation size effect in low load microhardness testing. J. Mater. Res. **8**(5), 1028–1032 (1993)
90. J.G. Swadener et al., A mechanistic description of combined hardening and size effects. Scripta Materialia **47**(5), 343–348 (2002)
91. H. Gao, Y. Huang, Geometrcally necessary dislocation and size-dependent plasticity. Scripta Materialia **48**, 113–118 (2003)
92. W.D. Nix, H. Gao, Indentation size effects in crystalline materials: a law for strain gradient plasticity. J. Mech. Phys. Sol. **46**(3), 411–425 (1998)
93. T.-Y. Zhang, W.-H. Xu, Surface effects on nanindentation. J. Mater. Res. **17**(7), 1715–1720 (2002)
94. I.L. Jager, Surface free energy—a possible source of error in nanohardness? Surf. Sci. **565**(2–3), 173–179 (2004)
95. S. Graça, R. Colaço, R. Vilar, Indentation size effect in nickel and cobalt laser clad coatings. Surf. Coat. Technol. **202**(3), 538–548 (2007)
96. S. Graça et al., A displacement sensing nanoindentation study of tribo-mechanical properties of the Ni-Co system. Appl. Surf. Sci. **254**, 7306–7313 (2008)
97. S. Graça, R. Vilar, R. Colaço, The role of indentation size effect on the abrasive wear behaviour of ductile metallic materials: a nanotribological study. Wear **268**(7–8), 931–938 (2010)
98. F.R.N. Nabarro, J.P. Hirth, *Dislocations in Solids*, vol. 11, ed. by F.R.N. Nabarro, J.P. Hirth (Elsiever, Amsterdam, 2002)
99. M.F. Ashby, The deformation of plastically non-homogeneous alloys. Phyl. Mag. **21**, 399–424 (1970)
100. J.F. Nye, Some geometrical relations in dislocated crystals. Acta Metallurgica **1**(2), 153–162 (1953)
101. N.A. Fleck et al., Strain gradient plasticity: theory and experiment. Acta Metall. Mater **42**(2), 475–487 (1994)
102. N.A. Fleck, J.W. Hutchinson, Strain gradient plasticity. Adv. Appl. Mech. **33**, 295–361 (1997)

103. Z. Ma et al., Indentation depth dependence of the mechanical strength of Ni films. J. Appl. Phys. **103**(4), 043512 (2008)
104. S. Graça, P.A. Carvalho, R. Colaço, Dislocation structures in nanoindented ductile metals—a TEM direct observation. J. Phys. D Appl. Phys. **44**, 335402 (2011)
105. S. Graça, R. Colaço, On the influence of indentation size effect on the wear of metallic alloy. Int. J. Surf. Sci. Eng. **5**(5/6), 457–462 (2011)
106. A. Misra, I. Finnie, On the size effect in abrasive and erosive wear. Wear **65**, 359–373 (1981)
107. U. Beerschwinger et al., Wear at microscopic scales and light loads for mems applications. Wear **181**, 426–435 (1995)
108. S. Sundararajan, B. Bhushan, Micro/nanotribological studies of polysilicon and SiC films for MEMS applications. Wear **217**, 251–261 (1998)
109. A.R. Krauss et al., Ultrananocrystalline diamond thin films for MEMS and moving mechanical assembly devices. Diam. Relat. Mater. **10**(11), 1952–1961 (2001)

Chapter 24
Nanowear of Polymers

Mario D'Acunto, Franco Dinelli and Pasqualantonio Pingue

Abstract The use of viscoelastic materials, such as polymers, constantly increases in the field of nanotechnology. These materials are softer than metallic and inorganic ones, and, because of that, they are easier to deform and wear off. The wear mechanisms occurring for viscoelastic materials are rather complex, and, generally, present more complications for a direct investigation with respect to metals or ceramics materials. With the advent of Scanning Probe Microscopy (SPM), well characterized forces can be applied to a surface with a nanometer-scale spatial resolution. In particular Atomic Force Microscopy (AFM), working at high contact forces, can significantly modify many surfaces. Polymers are soft enough to be modified by hard AFM tips, such as those of silicon, silicon nitride or diamond. For these reasons, the AFM is today the main tool employed to investigate wear occurrence on polymer surfaces. The wear of a polymer surface caused by an AFM tip in a regime of single asperity contact is an articulate process that depends on conditions such as, namely, the applied forces, the tip shape, size and the relative velocity. Since the influence of all these parameters is in close connection with the sample properties, one can expect a dependence of the wearing process on the mechani-

M. D'Acunto (✉)
Istituto di Struttura della Materia, ISM-CNR, via Fosso del Cavaliere, 100, 00133 Rome, Italy
e-mail: mario.dacunto@ism.cnr.it

M. D'Acunto
Istituto di Scienza e Tecnologie Dell'Informazione, ISTI-CNR, via Moruzzi 1, 56124 Pisa, Italy

M. D'Acunto
NanoICT Laboratory, Area Della Ricerca CNR, via Moruzzi 1, Pisa, Italy

F. Dinelli
Istituto Nazionale di Ottica, INO-CNR, via Moruzzi 1, 56124 Pisa, Italy
e-mail: franco.dinelli@ino.it

P. Pingue
NEST, Scuola Normale Superiore and Istituto Nanoscienze - CNR, Piazza San Silvestro 12, 56127 Pisa, Italy
e-mail: p.pingue@sns.it

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_24

cal properties of the sample surfaces. These properties can vary significantly from the bulk properties, if cross linking is made or, on contrary, residual solvents are present in the specimens. This chapter is divided in three sections following a general introduction. Specifically, the first section deals with wear induced by means of AFM tips to study the mechanical properties of films at the nanoscale; the second one regards the exploitation of wear for the creation of nanolithographic patterns; the last one is finally dedicated to an applicative field such as the characterization ofwear of polymers for biomedical applications at the meso- and nanoscales.

24.1 Introduction

Despite the increasing popularity and technological advances in nanoscience and nanotechnology applications, their weakness related to severe tribological problems, i.e. friction and wear, tend to undermine the performances and reliability of nano-devices. Several studies have shown that the tribology is one limiting factor to the broad-based impact of nanotechnology on our everyday lives [1–3]. Indeed, tribology is the science of surfaces in relative motion with each other and on the nanoscale the surface properties and geometries play a more fundamental role with respect to bulk applications. Miniaturization and the subsequent development of devices for nanotechnology applications require better tribological performances of the system components and a fundamental understanding of the basic phenomena underlying friction and wear on the nanoscale [4, 5]. Moving from the macro- to nanoscale, the surface area-to-volume ratio increases considerably and becomes a cause of serious concern from a tribological point of view. Indeed, surface forces such as meniscus forces, viscous drag, surface tension, adhesion and friction, being proportional to the contact area, significantly increase and can limit the reliability and the life of nano-devices.

Wear, which in a macroscopic sense is defined as the mass or volume loss from a sliding or eroding contact with a sequence of events, can be summarized as follows: mechanical forces, frictional work, impact forces, contact fatigue stress, cavitation force-induced damage in contact surfaces. The surfaces in reciprocal contact lose mechanical cohesion and a debris is produced. Subsequently to its formation, the debris is generally expelled from the contact zone and the process of wear is observed. The profiles of the two counter surfaces can be convoluted or not due to differences in wear resistance of the surfaces involved.

For polymeric systems there is also the very important prospect of marked environmental influences, which will include the role of lubricants as well as the often-dominant consequences of frictional heating. The same complications can be observed in polymer alloys, copolymers and polymer nanocomposites [6–10]. In addition, the knowledge of wear mechanisms on micro/nanoscale could help to quantify the distribution of material loss during the relative motion of the surfaces. In fact, weighting the sample before and after the test has been the dominant wear quantification technique. A precision balance typically has a resolution of

10^{-6} of the maximum load, which puts a limit on the minimum load that may be quantified in relation to the total weight of the component. Moreover, the mass of polymers can change due to water absorption. Further, it is important to note that macro-wear is 10–1,000 times higher than micro-wear on harder materials, even though the mean contact pressure in the micro-test is higher than in the macro-test [11–14]. As an example, the accurate knowledge of two contemporaneous factors, such as chemical degradation and mechanical wearing occurring in vivo, when biostable or biodegradable polymers are used in the human body, could help to improve enormously materials to be used in medicine. Indeed, the formation of nanoparticle debris from biomedical polymers has serious consequences for human health [15, 16].

The mechanisms for the occurrence of wear have been classified in accordance with its phenomenology (abrasion, adhesion, erosion, fretting, etc.) and the damage of rubbing surfaces have been identified and classified as cutting, plowing, cracking, delamination, fatigue and oxidation. The landscape is complicated by the fact the different mechanisms of wear can occur at the same time. In many case, a wear mechanism can initially start promoting damage and in a second moment such phenomenacan overlap or substitute one another making it difficult to individuate the single mechanisms leading to the measured damage and thus to develop adequate mathematical models.

As above mentioned, tribology of polymers is a complex topic, since such materials present specific properties if compared to other materials. These specificities are mainly due to the molecular structure, in a special way chain mobility. Chain mobility allows relaxation movements and energy dissipation, notably by internal friction. In addition the wide range of such movements can be time and temperature dependent. Molecular structures, like chains length or crosslinking degree, are also parameters able to modify the bulk and interfacial dissipation. Moreover, crosslinking reactions are usually incomplete, leading to an uncompleted network. More in detail, some chains are chemically bonded to the network, through both extremities; other chains can be linked to the network only by one extremity; while other chains are not at all chemically bonded to the network acting as free chains. Figure 24.1 summarizes the different wear mechanisms for polymers.

On the nanoscale, wear measurements for polymers, polymers alloy, copolymers, nanocomposites and blends, can be made by using SPMs, where the advantage of a single asperity sliding on a polymer surface sample is clear [16–18]. Scratch test can be carried out on areas or lines, and sample modifications can be observed at the end of wear testing. Again, nanoindentation is a powerful method to study the Young elastic modulus, for areas confined to few nanometers squared. General relationship between the normal load and the elastic response of the sample can be found in the Hertzian and Sneddon theories, with more complex cases better described by the Johnson-Kendal-Roberts (JKR) or Derjaguin-Muller-Toporov (DMT) approach [19–21].

The basic SPM tip-induced test for the analysis on the onset mechanisms of wear on polymer surfaces consists in the repeated scanning of a prefixed area of the film surface at constant force [22, 23]. In SPM the surface topography of the sample is

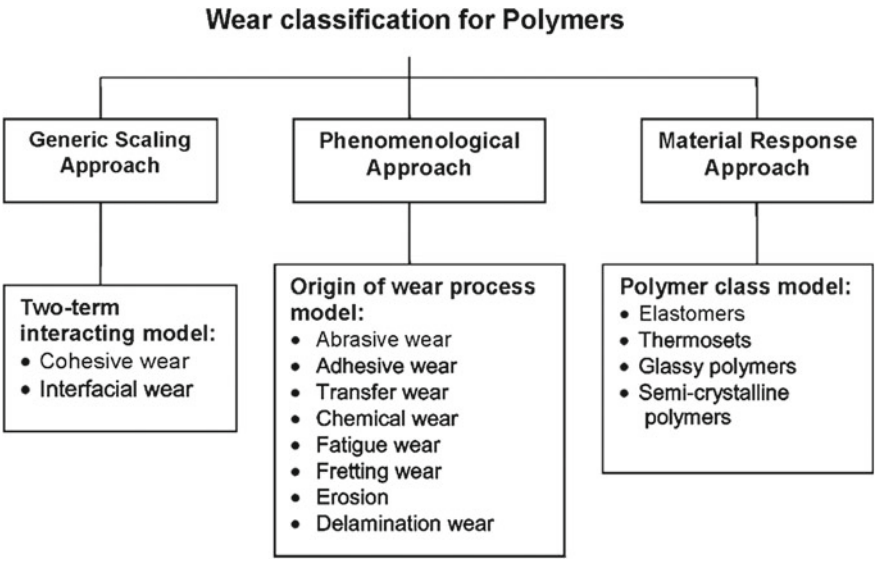


Fig. 24.1 A classification of different wear occurrence for polymers as proposed by Briscoe and Sinha. (Reproduced from [6], with permission)

imaged by scanning a small probing tip in contact mode. The same tip is used to induce wear. Such technique permits the following wear properties to be observed both qualitatively, and, when possible, quantitatively: (1) qualitative evolution of the surface during the test, (2) evaluation of the wear volume, for a comparison with macroscale wear tests, and (3) observation of the AFM tip degradation as a consequence of possible adhesion-wear mechanisms for polymer debris [24–26]. Generally, the tip can be moved along the horizontal axis forward and backward along the same line, then moves perpendicular to it and starts the next line. The motivation behind such abrasive tests can be the recognition of the potential that SPM tip-induced wear experiments offer for the characterization of the molecular organization of polymer surfaces. For example, it is generally recognized that the glass transition temperatures of polymer surfaces differ from those of the bulk, suggesting that viscous behavior may be observed at the surface at temperatures where the bulk of the material is glassy [27]. Studies of tip-induced wear behavior may be valuable in assessing the mechanical properties of polymers surfaces. Moreover, for some classes of thin-film materials, the application of bulk mechanical testing methods is not feasible and other techniques associated with surface mechanical investigation, such as nanoindentation, are inappropriate.

One interesting common phenomenon occurring at the initial stage of wearing test is the formation of ripple structures normal to the fast scanning direction [22, 23, 28–32]. Such ripples structures can be produced both with performing many scans on the same scanning area or with a single scan with a hot probe tip [33–36] or on polymer film initially trapping solvents [37, 38]. This phenomenon

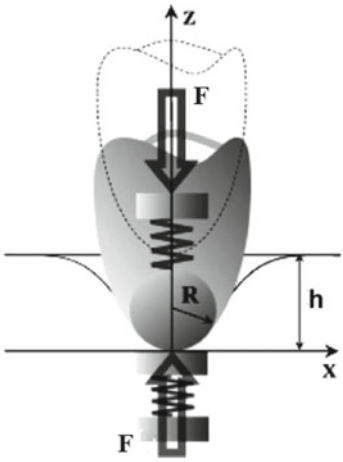


Fig. 24.2 The two-spring model used to describe the tip-surface action. The tip is attached to a cantilever with spring constant k, resulting in a normal load $F = k \cdot d$, and the force interaction is modeled by a spring with a stiffness equal to the force gradient. The force interaction spring is not constant, but depends on the tip-sample distance or penetration depth h. (Adapted from [8], with permission)

is however a complex process and it depends on the contact conditions, namely the applied force, the tip shape and size, and the relative velocity. Since the influence of all these parameters is in close connection with the sample properties, one can also expect a dependence of the wearing process on the sample mechanical properties. These properties can vary significantly due to variation in molecular weight (M_W), cross linking degree, the presence of residual solvents and also confinement effects.

Contact mechanics representing tip-sample interaction can be described by several models [39–41]. The simplest contact-mechanics model used in SPM-based techniques is the Hertzian model, which assumes a hemispherical tip shape and neglects adhesive forces. Such model has been successfully applied to systems with very low surface energy (i.e. very small adhesion forces) and small applied static loads (less than 5–10 nN). These conditions generally prevent wear and plastic deformations both of sample and tip. More detailed contact-mechanics models, such as the DMT [42] and JKR [43] models have been adopted for SPM systems in case where the adhesion forces could be neglected. However, even for the simplest model several parameters must be known, including the tip geometry and modulus, the applied static load, and the adhesive forces present. The overall situation is complicated by the fact that many of the parameters involved in the contact are neither easily measured nor well controlled. For example, the applied force F to the tip is usually calculated from the spring constant k and the deflection d of the cantilever using Hooke's law $F = k \cdot d$. However, determination of k requires separate measurements and the characterization of tip shape requires a separate tip-reconstruction methodology. In addition, indentation depth is a function of F, tip geometry (radius R or parabolic focus distance), as well as the mechanical and the adhesion properties of the contacting bodies (Fig. 24.2).

A non-axial displacement of the AFM cantilevers caused by their tilted orientation does not usually exceed 10 %, thus, it can be generally neglected. In addition, the radius R and spring constant k are initial system parameters, which must be measured

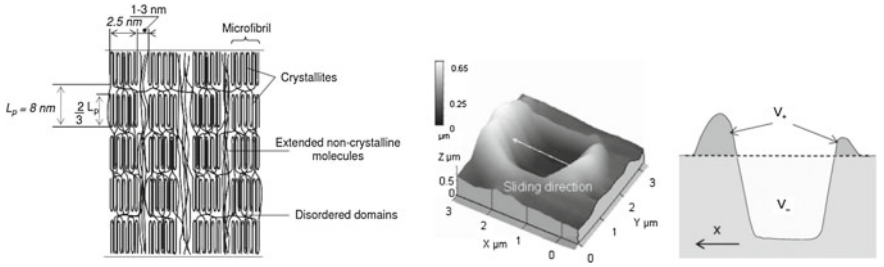


Fig. 24.3 Example of indentation test on a *poly*(*amide*) *6* sample. *Left Prevorsek's swiss-cheese* structural model of the sample [44], where the fibre axis is vertical and crystallites are periodically organized to form fibrils, which are embedded in an amorphous oriented matrix. *Middle* topographic image of the surface after a wear test. *Right* calculation of the wear volume from the surface profile changes, $V_+(V_-)$ represents the volume of material in the frontal pile-up. (the grooved volume adapted from [45], with permission)

and calibrated before the nanomechanical test. The Poisson ratio, v, is used as well as a material parameter and is usually assumed to be known and taken as a bulk value with possible deviations playing a minor role. The most general relation between penetration depth, h, and F in the course of indentation experiment can be presented in the very general form as:

$$F = ah^b \tag{24.1}$$

where a and b are specific, mode-dependent geometrical parameters, (for example, $b = 1.5$ for both Hertzian and parabolic Sneddon's contact theories). In addition, Sneddon model suggests a specific and practical analytical relationship between the surface stiffness, *dF/dh* and Young's modulus E, in the form:

$$\frac{dF}{dh} = 2\sqrt{\frac{A}{\pi}}E \tag{24.2}$$

where A is the contact area and E is the reduced modulus:

$$\frac{1}{E} = \frac{1 - v_1^2}{E_1} + \frac{1 - v_2^2}{E_2} \tag{24.3}$$

and E_1 and E_2 are the elastic moduli of a surface and the tip, whereas $v_{1,2}$ are the corresponding Poisson numbers. Using (24.2) it is possible to evaluate an absolute value of the sample elastic modulus when contact area changes are known as a function of the specific shape of the indenter tip (specific analytic expression are available for circular, pyramidal and parabolic shapes). For small indentation depths, the elastic modulus can be calculated making use of the Hertzian model of a sphere-plane contact type. For larger indentation, the Sneddon model (Fig. 24.3) with a parabolic tip is usually exploited.

Two main contributions cause significant deviations from pure elastic contact mechanical behavior. The first one is a viscoelastic phenomenon that results in additional viscous contribution causing increasing deviations of the loading curves from the expected Hertzian behavior for purely elastic solids. The second one is the presence of a stiff substrate beneath a very thin polymer film contributes in a limited film deformation under a high load resulting in a decreasing deviation from the purely elastic behavior. As a consequence, the loading behavior (i.e. indentation depth versus normal load range) of a wide class of polymers and polymers composites in glassy or rubbery states cannot be reasonably described with Hertzian mode. Johnson proposed a general relationship between A, F and loading time, t, for block copolymers [46]:

$$(A/pgreco)3/2\,(\tau) = \frac{3RF}{4\tau E^*_\infty}\left[\tau - (1-k)\,(1-\exp(-\tau))\right] \tag{24.4}$$

where $\tau = t/T$ is the reduced time and T is the relaxation time of the material, $k = E^*_\infty/E^*_0$ is the reduced modulus with E^*_0 being the initial instantaneous modulus, and E^*_∞ is the equilibrium relaxed modulus for an infinitely slow load, $U = F/t$ is the loading rate. It is shown that by varying two primary variables, E_0 and T, it is possible to fit the experimental data. The presence of the viscous contribution component can principally change the shape of the loading curve from the convex to concave in a penetration depth versus load plot.

One main challenge for tribologists is also the control of wearing processes under some conditions and environments. Recently, Gotsmann et al. focused the attention on controlling the wear characteristic of synthetic polymers for contact mechanical operations in NEMS [47]. They have observed the dependence of wear on segmental relaxation dynamics by tailoring the polymer synthesis with predetermined cross linking levels. As a consequence, a reduction of wear in a polymeric NEMS is possible. Moreover, two wear regimes have been identified: (1) below a critical cross link density, a ductile wear mode exhibits weak dependence on the spacing between cross links, and (2) above a threshold cross link density, a brittle wear mode becomes operative and the wear rate decreases rapidly with additional cross linking (Fig. 24.4). The threshold cross link density that separates the two wear modes occurs at a critical segment weight of approximately 2.15 kDa. This weight corresponds to a critical cross link spacing of approximately 3 nm, suggesting a competition between the cross link spacing and the cooperation length for segmental backbone relaxation. When the spacing between cross links falls below the cooperation length, constraints are imposed on the backbone mobility, essentially stiffening the material. Under these conditions, increased hardness and reduced wear are strongly correlated with cross link spacing. If the concentration of cross links is insufficient to interfere with natural backbone relaxation, the presence of cross links has little impact on the polymer response and the material behaves similarly to the uncross linked native polymer [47].

Under a wide range of conditions, the repeated scanning of a polymer surface leads to the formation of tip-induced wear patterns consisting of ridges oriented

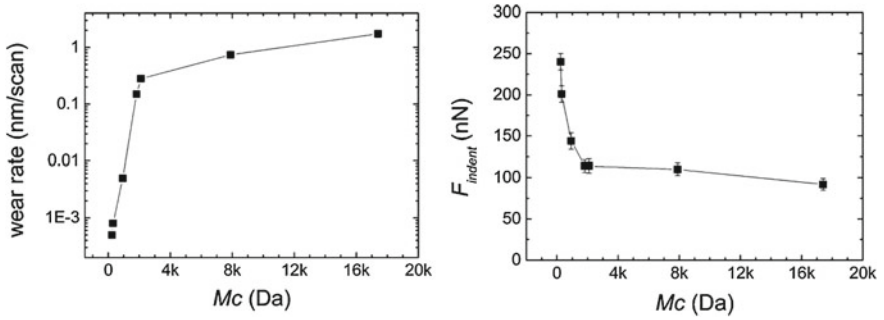


Fig. 24.4 On *left*, the impact of cross link spacing (M_c) on wear rate is shown. Two wear regimes can be individuated by a critical M_c between cross links, M_{crit}. Below $M_{crit} = 2.15 \pm 1.5$ kDa, wear is highly sensitive to the cross link density, while above M_{crit}, there is little wear dependence on M_c. The uncertainty level represents the intersecting domain defined by the root mean square variance of regression fits above and below M_{crit}. On *right*, the impact of cross link spacing on the hardness of the styrene-benzocyclobutene (PS-BCB) random copolymer thin films is shown. The minimum force needed to make a permanent 0.5 nm deep indentation, $F_{\text{indentc,crit}}$, is plotted versus M_c. (*Note* for a given indentation geometry and penetration depth, i.e. equal contact area, the hardness is directly proportional to F_c) (Reprinted from [47], with permission)

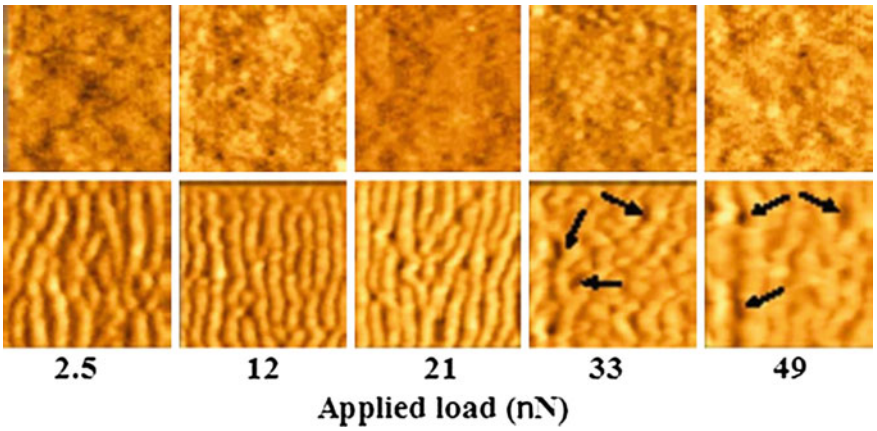


Fig. 24.5 Topographic images for an AFM-tip-induced wear test performed in air (from 1 to 40 scanning cycles) with the indicated applied loads on annealed polystyrene (PS). All experiments were conducted on annealed films with a tip nominal radius of 15 nm and scanning directions is horizontal. Each image is $1 \times 1\,\mu m^2$. The topographic variation is around 2 nm for the *upper row* images (first scanning cycles) and 55 nm for the *last bottom row* image (40th scanning cycle). The *arrows* indicate areas where the film has broken. (Reprinted from [48], with permission)

perpendicular of the scanning directions (Fig. 24.5) [20, 48]. In Sect. 24.3, the ripple formation for polymer materials will be accurately described, here we limit ourselves to present some general features of wear patterns. Surtchev et al., showed that the evolution of the root mean squared roughness follows an exponential saturation law

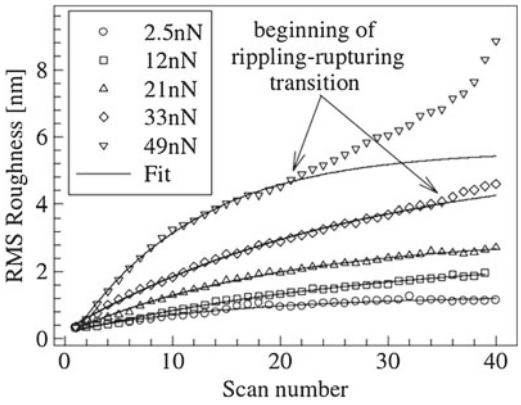


Fig. 24.6 Root mean squared (RMS) roughness as a function of the scan number for the various loads as in Fig. 24.4. The *solid lines* represent the fits of the experimental data. (Reprinted from [48], with permission)

(Fig. 24.6) [48]. In addition, tip-induced wear is more extensive at higher applied loads where transition from rippling to rupturing wear was also observed. And the degree of wear as well as the type of the patterns formed was found to strongly depend on the density of the scan line and an overlap between the successive scan line is necessary to obtain a periodic pattern (Figs. 24.5 and 24.6). This interesting effect of tip-induced rippling on polymer films will be deeply discussed in the next paragraph.

24.2 Wear Tests at the Nanoscale in Polymer Films to Assess Material Properties

The deformation of polymer surfaces upon contact with a tip is a phenomenon known before the inception of Scanning Probe Microscopies (SPM). It has been observed that depending on film properties, contact area, tip shape and trajectory, a surface initially flat can progressively undergo morphological changes. This phenomenon has been considered as a nuisance as the microscopists were interested in finding the best conditions to correctly image the pristine topography of polymer surfaces. Therefore, operational modes avoiding contact or limiting it to simple intermittent contact have been developed. However, later on some researchers have considered that this nuisance could be positively used. They have realized that studying the way the deformation occurs allows one to probe the mechanical properties of the surfaces investigated and hence to deduce their molecular organization. Thus, since the early nineties, a number of experiments has been devoted to this purpose. In this section, we shall consider a range of applications where deformation studies can be useful to investigate and define sample mechanical properties, namely: crystallinity and amorphousness; plasma treatment; solvent presence; molecular organization (entanglement density); and so on.

24.2.1 Schallamach Waves and Ripples

Before the inception of SPM, the phenomenon of polymer wear had been already studied at a macroscopic level [49–55]. The apparatuses employed were constituted by two bodies in contact under a given load. The geometries could vary from pin on flat to sphere on flat or two cylinders rotated by 90° one to the other. Generally, the flat one is the plastic sample under examination. Sliding was performed in a continuous movement at constant velocity. The amount of material worn off from the surface was than quantified in order to assess the wear resistance properties.

Subsequently those instruments were employed in a single scratch configuration at a constant velocity and measuring the lateral force experienced by one of the two bodies [45, 56–62]. The sample surface was observed with optical microscopy. In some cases, these observations could be performed in real time if one of the two bodies was optically transparent. Progressively, the contact area has been reduced moving from a contact radius of mm to μm. Different wear regimes were identified from cutting to abrasive one. In this modality one could determine the sample mechanical properties and correlate them with the pristine molecular configurations of the sample itself.

Schallamach has been one of the pioneers in the macroscale field [49]. In particular the phenomenon known as 'the Schallamach waves' refers to a failure of a continuous sliding motion in which waves of detachment develop perpendicular to the sliding direction. These waves form at once at the front edge of the contact, perpendicularly oriented with respect to the velocity direction, but they may develop also within the contact area. This phenomenon occurs under certain given experimental conditions and can be quite easily observed with an optical microscope in real time (Fig. 24.7).

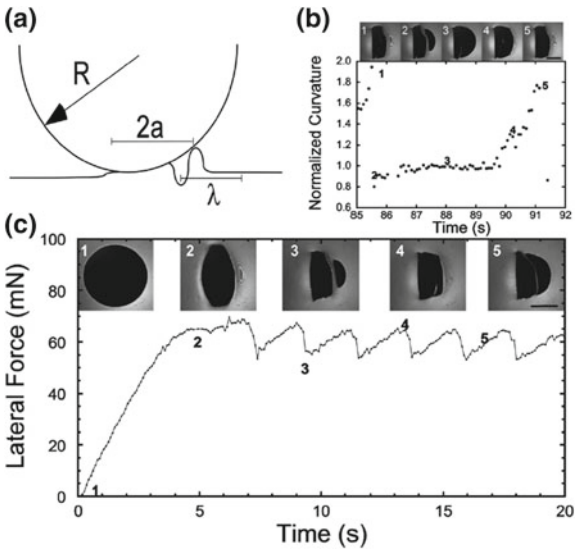


Fig. 24.7 **a** Schematic of Schallamach waves. **b** Normalized curvature of the contact area during one wave cycle, $R = 5\,\text{mm}$, $v = 4.63\,\mu\text{m/s}$, $E^* = 0.9\,\text{MPa}$ at $\omega = 0.3\,\text{Hz}$. **c** Representative lateral force data with associated images of contact area, $R = 5\,\text{mm}$, $v = 20.84\,\mu\text{m/s}$, $E^* = 0.9\,\text{MPa}$ at $\omega = 0.3\,\text{Hz}$. Scale bars in the last image of **b** and **c** represent $250\,\mu\text{m}$. (Reproduced from [57], with permission)

The advent of SPM, in particular Atomic Force Microscopy (AFM), has opened the possibility to study single contact asperity [22]. Thus the contact area can be reduced to a diameter size of the order of the molecules forming the film. Whatever the size of the contact, the aim has always been to deduce the material properties from the way wear occurs as a function of a number of parameters, namely: load, velocity, tip radius, temperature and others. As for the macroscopic case, different wear regimes have been identified including ploughing and morphological rearrangements of the surface, generally named 'ripples'.

When one moves to smaller contact areas the wear features become too small to be below the detection limits of optical microscopy. This is more and more true for the nanoscale case. Grooves or ripples can be observed only after being formed. Additionally the post analysis can be only carried out with the same object producing wear, typically working in a different load regime with lower values compared to those set in wear regimes. The phenomenon of ripple formation is the one, which has attracted most of the attention of the researchers. The first observations were taken by operating in AFM 'imaging mode' at high load values, that is with the probe moved over a given area line by line not or partially superposed [22, 23, 28, 47, 63, 64]. When the ripples form, they are also perpendicular to the velocity direction. The hypotheses proposed for the mechanism of ripple formation on the nanoscale are the same of Schallamachmacroscale case (i.e. crack opening at the edge or within the contact area) or at the front edge of contact area due to a stick slip phenomenon.

If one proceeds in a single scratch mode (Fig. 24.8), the situation above described can be more clearly interpreted [65–67]. It means that for macro contacts, the ripple formation is dependent on a peeling phenomenon within the whole area, whereas for nano-contacts, it is rather due to a stick and slip phenomenon occurring at the front edge only. In fact the ripple periodicity at the nanoscale is larger than the contact area. According to Aoke et al., the friction coefficient measured in the two cases can be equivalent, when one normalizes the normal load values to the contact area. From this observation he infers that the two processes, inducing both plastic deformations, are determined by similar mechanisms [66].

24.2.2 Multiple Line Scratch Test

Although the single line scratch mode can appear more similar to macroscale measurements, as already mentioned, the first studies with AFM were carried out in what it may be considered a multiple line scratch test (Fig. 24.9). This was due to the interest in imaging rather than studying the material properties. Then the analogy with Schallamach waves has immediately attracted attention and the single line scratch test is still not quite popular in the scientific community.

In the multiple line case, the complete formation of ripple patterns is dependent on the tip movement in a raster fashion [22, 23, 28, 47, 63, 64]. Sometime the ripples can form immediately after the first passage. However the analogy with the macroscale case can be only drawn if the line spacing is such that it can be described as the

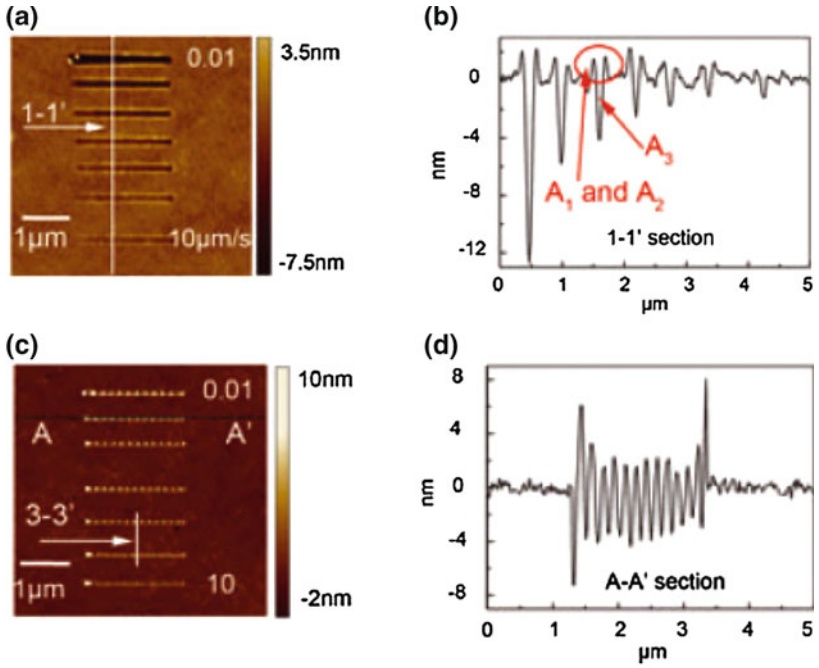


Fig. 24.8 Morphology of the film with $M_W = 1.3$ kDa (**a**), section 1–1 (**b**). Morphology of the film with $M_W = 250$ kDa (**c**), $A - A'$ (**d**). (Reproduced from [67], with permission)

parallel movement of a number of tips moving together along the same direction. In the end, the macroscale phenomenon can be described as a multi-tip one as the contact area is large compared to the single probes.

Thus nowadays wear at the nanoscale is typically studied in the AFM imaging mode. AFM can be operated in a way where the line interspace can be varied. In addition one can vary a range of parameters such as tip shape and chemistry, applied load, velocity, and more. From the patterns that form after one or multiple scan frames the properties of the sample can be qualitatively, and in some instances quantitatively deduced. Thus this technique can be useful to evaluate the wear and scratch resistance for applications like MEMS, nanofabrication and nanopatterning. A wide spectrum of polymers has been investigated including polystyrene (PS), poly (methyl methacrylate) (PMMA), gelatin, polyacetylene (PAc), polycarbonate (PC), poly(ethylene terephthalate) (PET), poly (vynil alcohol) (PVA), poly(4-vinyl pyridine) (PVP) and a range of copolymers. In the following we report some examples such studies, which are mainly meant to give a flavour of a wide spectrum of possible useful applications.

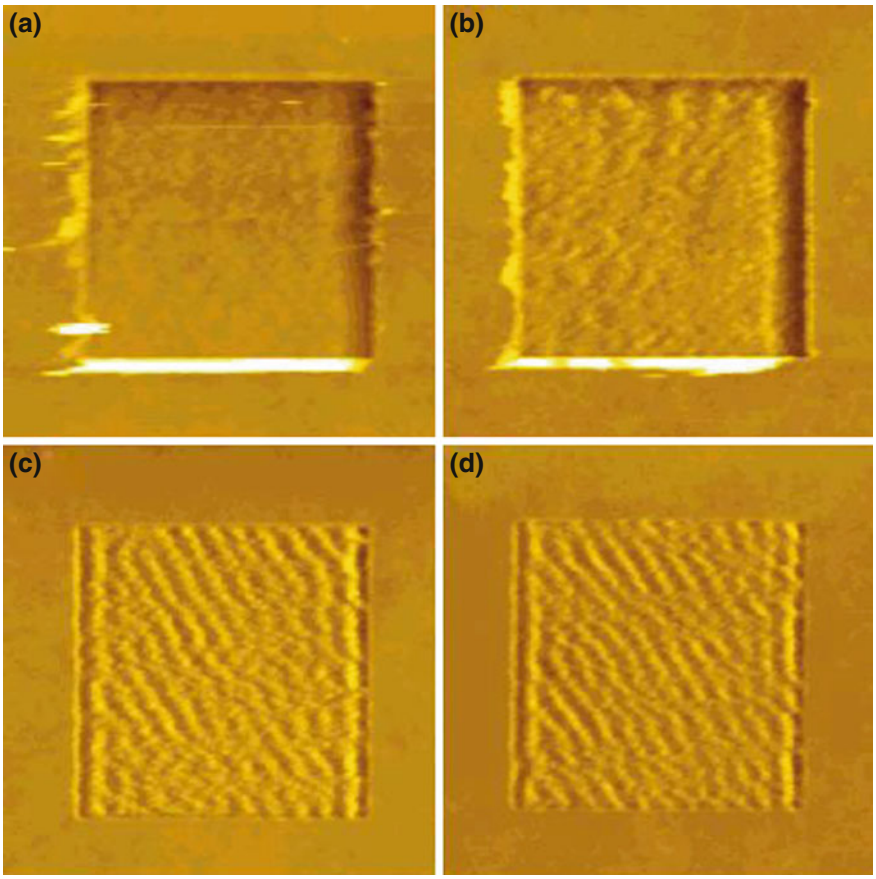


Fig. 24.9 SPM topographical images of scratched surfaces of PS films having different M_W's: **a** 8 kDa, **b** 15.8 kDa, **c** 58 kDa, **d** 164 kDa, and **e** 984 kDa. The scratching was carried out under an applied load of 10 nN. *Scan size* $1.5 \times 1.5\ \mu m^2$. (Reproduced from [64], with permission)

24.2.3 Amorphousness and Crystallinity

The first example of a wear study that we report is represented by the case of amorphousness [20, 64, 68–70] and crystallinity [71–74]. In fact, depending on the aggregation state of the molecules, the sample can present areas where the molecules present a conformational order and areas where the molecular conformations are random and therefore no order is observed. Starting with the case of completely amorphous samples, an important parameter is represented by the average molecular weight (M_W) and the monodispersity index. In general, samples can be an assembly of molecules of different M_W values and thus, in principle, they can have a different wear behavior as those reported in Fig. 24.9 for the case of PS films.

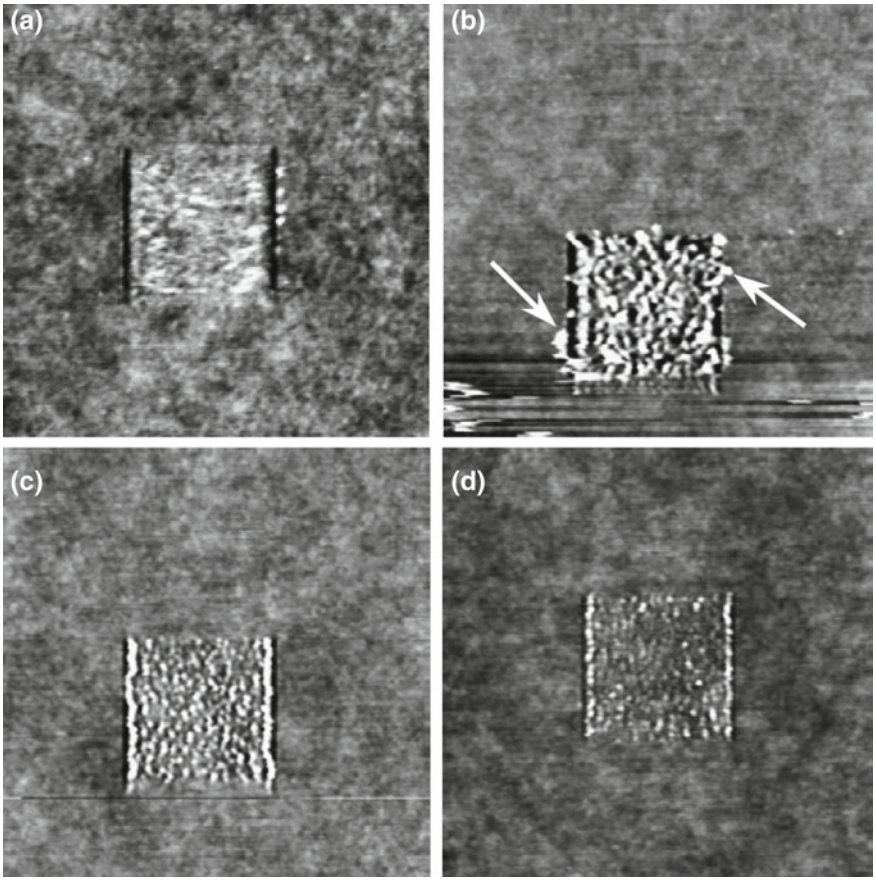


Fig. 24.10 SPM topographical images of wear patterns induced by scanning forthree frames at 12 nN for 13 K and for three frames at 25 nN for the other M_w values: **a** 13 kDa, **b** 30 kDa, **c** 96 kDa and **d** 483 kDa. *Image size* $3 \times 3\,\mu m^2$. *Maximum groove depth* 4 nm for 30 kDa, 2 nm for 96 kDa and 1.4 nm for 483 kDa. (Reproduced from [20] with permission)

It has been suggested that this behavior is correlated to the critical M_W (M_C). In fact if $M_W > M_C$, the molecules are entangled, provided thermal annealing has been performed, solvent has been removed and the sample is close to a thermodynamic stable state. The patterns and their load dependence reveal that for $M_W < M_C$ ripples do not form and bunches of molecules can be moved (Fig. 24.10). For $M_W > M_C$ ripples can form instead and molecules can hardly be removed from the sample. For M_W close to M_C ($\approx$30 kDa), one can pass from a non-wear regime to an abrasive one by increasing the applied load or the number of scans. In Fig. 24.10b (white arrows), it can be observed that lumps of material can be torn off from the surface, corresponding to bundles of entangled molecules.

From macroscopic investigations it was already known that crystalline regions are less prone to wear than amorphous regions. Beake et al. have systematically studied PET samples produced in different fashions so that they might have different crystalline domain size and density [72–74]. For instance they have investigated PET films either realized from simple molding or produced by applying a uniaxial or biaxial stretch to the melt when it is cooled down.

In Fig. 24.11, we report the case of a bi-axially stretched film. In general the applied load needed to form ripples is higher than for amorphous films. Alternatively one can obtain similar characteristic patterns by repeating the scan over the same area. In this case, we can see that changing the scan direction implies a different absolute orientation of the patterns that are always perpendicular to the scan direction. Areas showing different morphologies can show different pattern periodicity. This is possibly due to variations in the degree of crystallinity. In Fig. 24.11, it is finally visible that the periodicity depends on the applied load value.

24.2.4 Plasma Treatment

Plasma treatment of precursor or polymer films is an established technique typically used to produce wear resistant coatings. It is known that an RF radiation (with power P) impinging onto the samples initially produce the activation of chemistry by either exciting monomers or fragmenting long molecules. The reactive fragments then can form chemical bonds among themselves and, in the end the initial set

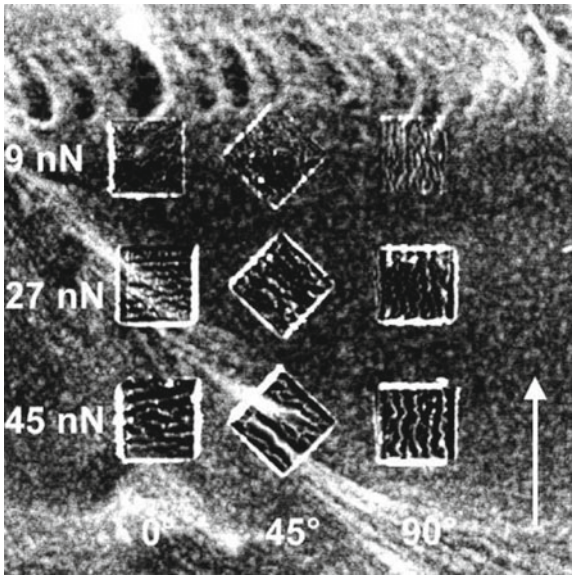


Fig. 24.11 Bi-axially stretched PET film: effects of applied load and scan direction. Topographical image of $8 \times 8\,\mu m^2$. Wear induced by scanning 9 times areas of $1 \times 1\,\mu m^2$. (Reproduced from [72], with permission)

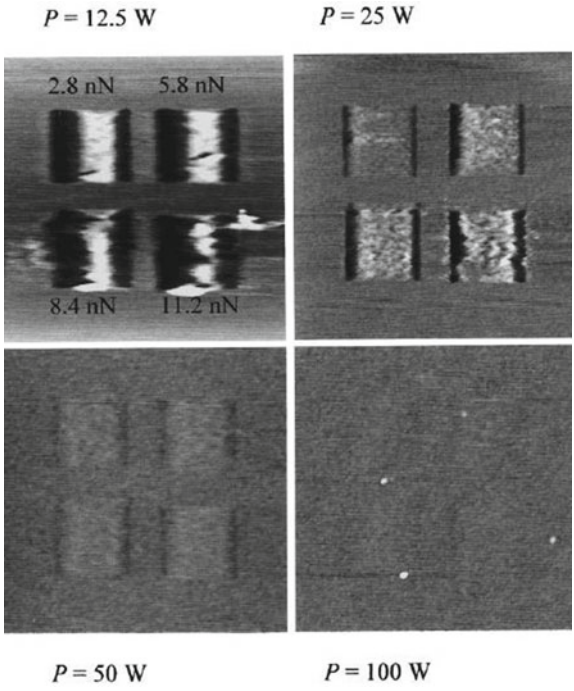


Fig. 24.12 Contact mode AFM images from $4 \times 4\,\mu m^2$ areas (*low force*). The images include $1 \times 1\,\mu m^2$ areas previously scanned at 11 Hz (*two passes*) using the applied loads annotated on $P = 12.5$ W for each sample. *P* stands for power, in this case. (Reproduced from [76], with permission)

of separated molecules evolves in a network more or less interwoven. The networking degree depends on plasma treatment parameters such as time exposure, power intensity, and others.

Wear tests can be carried out by varying the applied load and other scanning parameters [34, 75–79]. The deformation dependence reveal that samples with a higher power treatment or longer exposure time are less prone to deformation, in particular to ploughing (Fig. 24.12). In this example, for plasma polymerised hexane (ppHex) films ripples do not form at any stage. This may be due to the fact that the pristine film was made of precursors of very low M_W. The average molecular size increases with plasma treatment but entanglement may not occur and thus the conditions for the formation of ripples.

24.2.5 Presence of Solvent

The presence of solvent is known to weaken the polymer surface properties. This phenomenon is named 'plasticization' and it depends on the fact that solvent molecules can swell the polymeric film by interposing themselves in the bulk represented by the polymer molecules. Polymer films are often made by spincoating or otherwise by dropcasting. The thermal annealing is a post treatment that is meant to remove

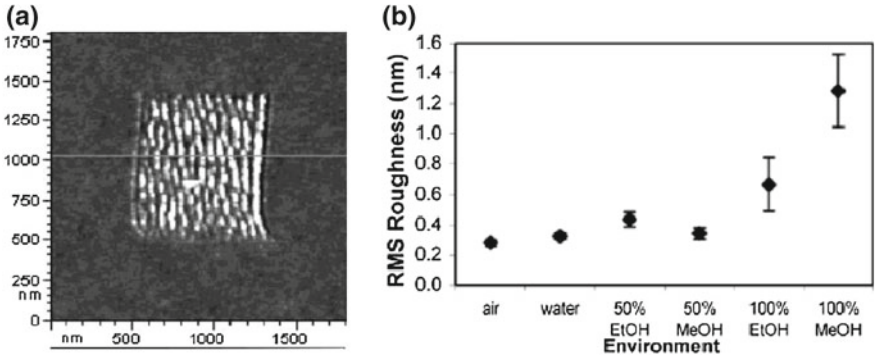


Fig. 24.13 AFM image of a PMMA film scanned in 50 % ethanol solution using a cantilever with a force constant of 0.36 N/m. The central 750 × 750 nm^2 area was scanned once at a higher normal force of 180 nN immediately before this image was acquired. **a** Typical AFM image of ridges; **b** RMS roughness of PMMA films in different solvents. Each data point represents the average and standard deviation of several measurements taken on each of several different samples. Both the average and the scatter increase in the stronger solvents. (Reproduced from [24], with permission)

the solvent molecules. However this process crucially depends on several parameters from the T value to the duration time, from the solvent employed to the interaction with the substrate. It might thus be often important to assess the degree of solvent presence [22–24, 37, 38].

In Fig. 24.13, we report an example of a film exposed to different environmental conditions including solvent vapours. The film more exposed to good solvents deforms at lower load values compared to other conditions. The ripple amplitude (RMS roughness) and periodicity depend on the degree to solvent exposure.

24.2.6 Temperature Dependence

SPM investigation can be extended to the T domain by heating the sample [31, 34, 80–82] or alternatively the probe [36, 83]. In case of amorphous polymers, one can define a second order thermodynamic discontinuity named 'glass transition'. As no polymer sample is ever fully crystalline, this transition can be observed to various extents depending on the crystalline degree and typically occurs at a given T value named glass transition $T(T_g)$. Another transition is represented by the first order transition corresponding to crystalline melting. This is generally not observable by AFM as it occurs at T values too high, where the amorphous regions present are in a rubber state.

In Fig. 24.14, it is shown an experiment in which a heated tip is used to perform scanning over a PS surface. At 100 °C the surface starts to deform as ripples appear. The T_g is thus defined as the value at which the polymer deviates from a glass and starts to show rubber-like behaviour. The fact that at 90 °C it is observed a ironing of

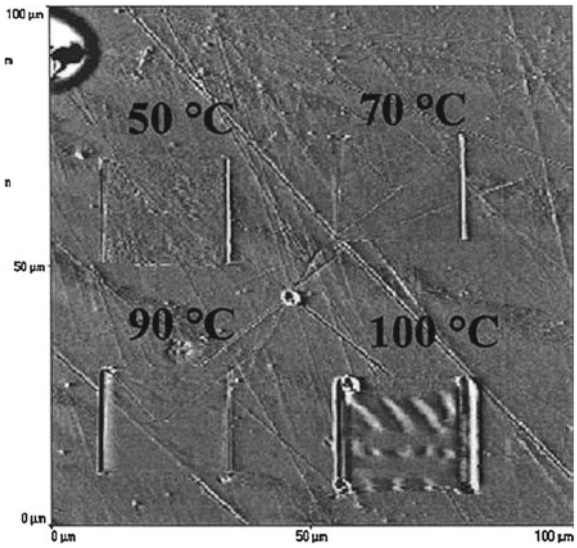


Fig. 24.14 AFM image of a PS film (296 kDa) after scanning regions of $20 \times 20\,\mu m^2$ at 50, 70, 90 and 100 °C. Note the central mark left behind from the hot tip after the scans were performed. The image of the area scanned at 50 °C shows almost no change compared to the surrounding surface. The images of the areas scanned at 70 and at 90 °C show a much flatter surface than before. Finally the image of the area scanned at 100 °C shows a wavy pattern typically for the deformation of a rubbery material. (Reproduced from [83], with permission)

the very superficial layer has induced the authors to state that the film interface with air has a lower T_g value.

Gnecco et al. have investigated the ripple formation in polymer films such as PMMA, PSul and PC [36]. They have clearly shown that the ripple formation varies with T and polymer type (Fig. 24.15a). The T dependence has a characteristic behaviour that can be correlated to the T_g value of the polymer investigated (Fig. 24.15b).

More often the experimental setups are built in a way such that the sample rather than the tip is heated. Several studies have been carried out to study polymer mechanical properties versus T. Incidentally we may cite Schmidt et al. who have investigated the inclination of the ripple patterns deducing the viscoelastic dependence from T [80]. Dinelli et al. have instead performed experiments at various speed and applied load, showing that T_g can be overestimated depending on the scanning parameters, namely the time of tip permanence versus the film viscosity [81].

24.2.7 Composites

The capability of investigating wear at the nanoscale has opened the possibility to investigate a class of material otherwise difficult to explore, the nanocomposites

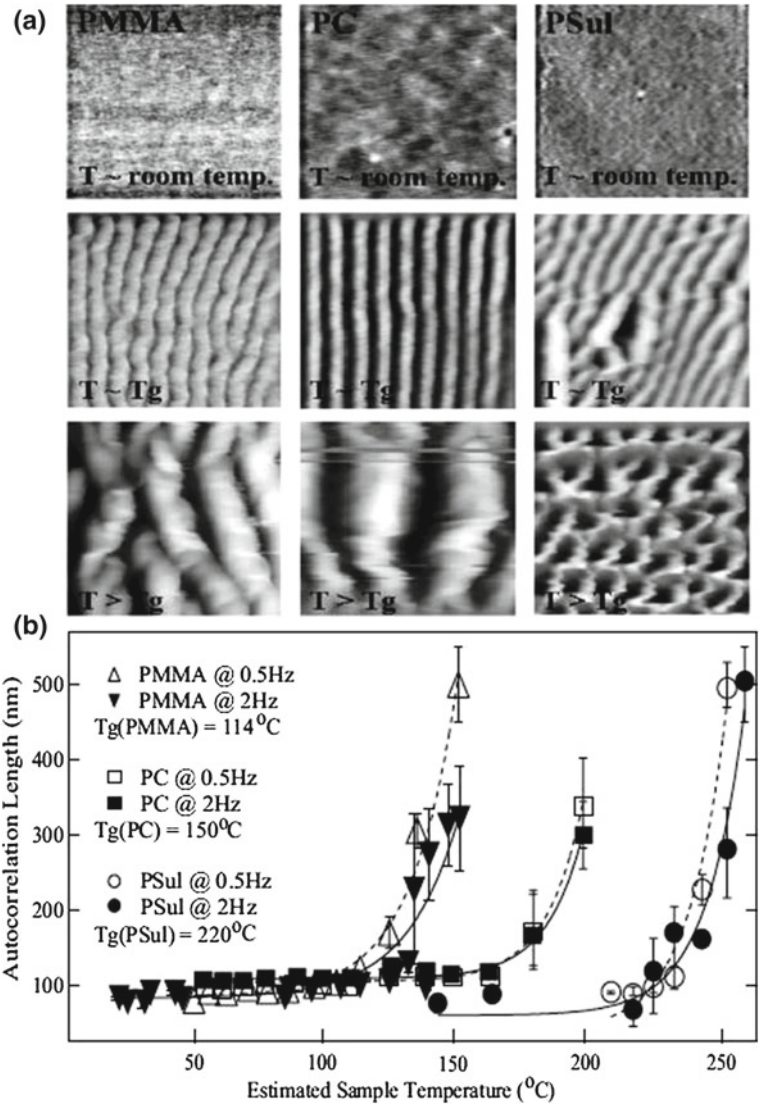


Fig. 24.15 **a** $1 \times 1\mu m^2$ AFM contact-mode images of ripples created by rastering polymer filmsspincoated on glass substrates with a resistively heated AFM probe at several temperatures. **b** A plot of the autocorrelation lengths obtained from **a** reveals an exponential growth of the ripple wavelengths with temperature. Data obtained for 0.5 Hz fitted with *broken lines* and for 2 Hz fitted with *solid lines* scanning speeds are presented for each polymer. (Reproduced from [36], with permission)

[10, 84–89]. Some of these composites have peculiar mechanical properties that may be useful for such applications as MEMS and nano-devices. A composite can be defined as a blend of polymeric molecules with other immiscible molecules, copoly-

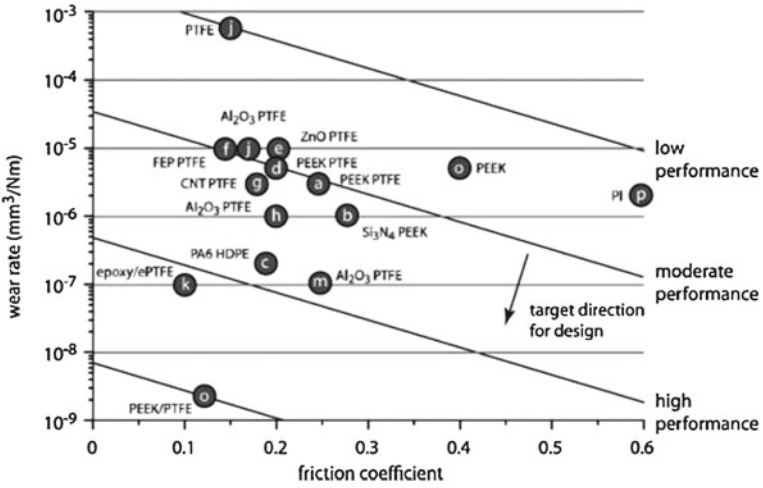


Fig. 24.16 A multivariate plot of wear rate (y axis) versus friction coefficient (x axis) for various polymeric composites, unfilled polymers and polymer blends. All the experimental data were acquired in the same environmental conditions with a geometry pin-on-disk and scanning velocity $v = 50.8$ mm/s, and a load pressure $F = 6.25$ MPa, for further details see [10]. (Reproduced from [10], with permission)

mers where molecules are made of two different components chemically linked or nanoparticles. For a polymeric material which is composed of a single constituent, low wear rate and low friction coefficient usually cannot be achieved simultaneously. For example, polytetrafluoroethylene (PTFE) is a widely used solid lubricant, which is well known for its very low friction coefficient ($\mu < 0.2$ in dry sliding conditions), high melting temperature and chemical inertness, but it wears 10–100 times faster than many other materials in general (Fig. 24.16) [10]. On the contrary, polyetheretherketone (PEEK) has high wear resistance, mechanical strength, operational temperature but high friction coefficient ($\mu > 0.4$ in dry sliding). Such polymer properties can be altered and changed as wanted using composites. The physical properties of a polymer composite can be tuned to satisfy various functional requirements of a target application, including stiffness and strength, thermal and electrical transport, and wear resistance.

One critical key for a deep comprehension of wear mechanisms on polymer nanocomposites is the accurate knowledge of matrix/filler interface. Currently, one need is the quantitative measurements of the transfer film during low wear sliding. Secondly, quantification of the compositional and chemical evolution of these films is crucial. In turn, quantification of fundamental mechanical properties of transfer films is needed to develop models of these interfaces at a more detailed level. All such needs in nanocomposites tribology will require a combination of tools and techniques ranging from materials scientists to tribologists providing a profound impact

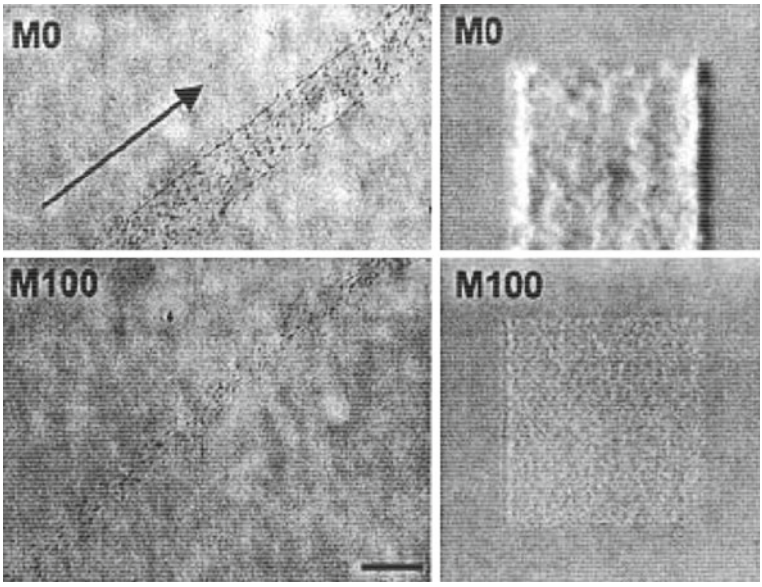


Fig. 24.17 (*left column*) Optical micrographs of wear tracks on copolymer surfaces made by ball sliding. M0 stands for no MMA content, whereas M100 for 100 % MMA content. The arrow and scale bars represent the ball sliding direction and 200μm, respectively. (*right column*) SPM topographical images of copolymer surfaces after multiline scratching under an applied load of 10 nN. The scan dimensions and contrast variations of these images are $1.6 \times 1.6\,\mu m^2$ and 30 nm, respectively. (Reproduced from [87], with permission)

on our knowledge of such complex systems. This new knowledge will help us to design materials with new improved tribological properties.

In Fig. 24.17, we show an example of copolymer samples made of methyl methacrylate (MMA) and *n*-butyl methacrylate.The ratio of the two components was varied. In the Figure, we only report the two extreme cases with maximum (M100) and minimum (M0) MMA content. It can be seen that the wear patterns can vary dramatically. This demonstrates how the mechanical properties can be tuned in copolymer films.

Buenviaje et al. have performed studies on blends made of two immiscible components with different T_g values [86]. In this case there was a clear phase separation and the various regions would undergo a glass transition at different T_g. The authors have shown that friction as well as wear can be useful to evaluate the viscoelastic state of these phases separated. This experiment can be useful to study T_g variations induced by confinement, for instance in cases where the phase separation can be controlled in size. In Fig. 24.18, it is shown the case of PS/PVP copolymer films where the two polymer ratio is varied and wear pattern reveals the phase separation at room T.

Finally in the last example, Pihan et al. [89] have explored the case of a PEMA matrix blended with nanoparticles grafted with PEMA of various M_W (Fig. 24.19).

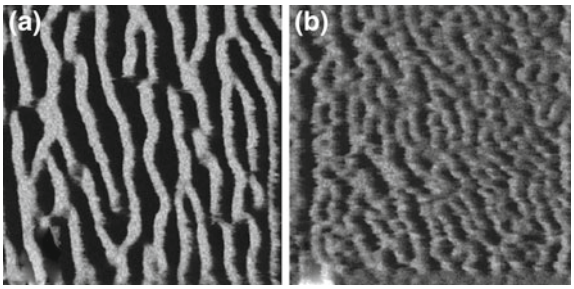


Fig. 24.18 Surface images of dip-coated PS/PVP (from 100 ppm in chloroform) onto silica, captured after the second scan, for PS/PVP 48/21 (**a**) and 194/21 (**b**). Area $3 \times 3\,\mu m^2$, height scale 10 nm and the scanning direction is from *left* to *right*. (Reproduced from [85], with permission)

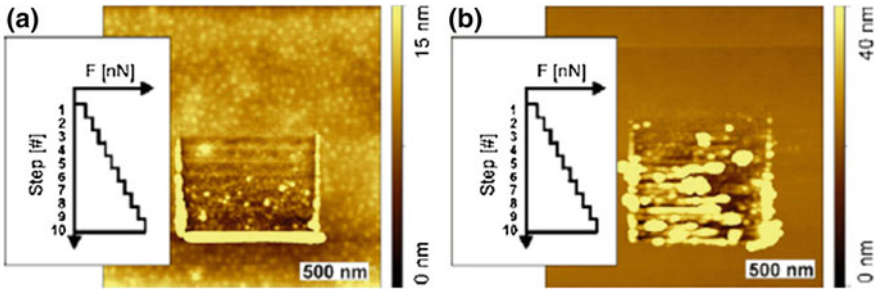


Fig. 24.19 Nanowear test on samples with matrix $M_W = 11.7$ kDa. **a** Topography image of a sample blended with PEMA-g-particles, **b** topography image of a pure homo-polymer sample. In the inner square the force was stepwise increased up to maximal normal force of 150 nN (step 10). (Reproduced from [88], with permission)

They observe that the wear resistance is improved upon adding the nanoparticles with a given ratio between polymeric matrix to nanoparticle density. They have also observed that wear can be correlated to PEMA M_W; in particular the wear behaviour changes for higher M_W values where molecular entanglement occurs. This is in accordance with previous studies on different materials such as for instance PS [20].

24.2.8 Boundary Conditions

To conclude, we show a final example that introduces the following Section on nano-lithography. The wearing properties of polymer surfaces can be exploited to create ripple structures by employing AFM scan with proper boundary conditions that allow fabricating self-assembled and ordered ripples on a nanometer scale. More in detail, in the work done by Napolitano et al. [38] the ripple orientation obtained on PET results significantly modified by the boundary conditions appearing when

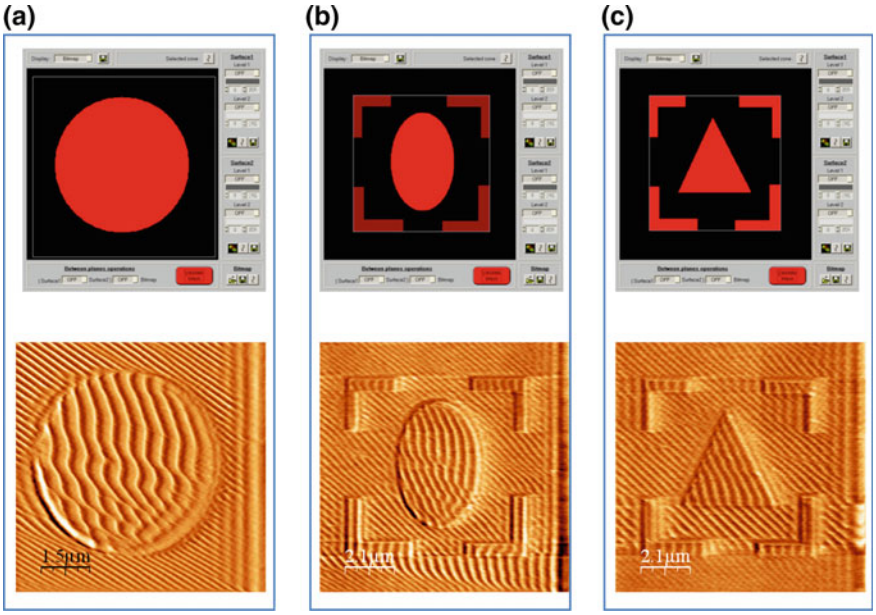


Fig. 24.20 Ripple patterns on **a** a *circular area*, **b** *ellipsoidal* and **c** *triangular regions* surrounded by four L-shaped regions (the color scale covers ∼30 nm on the three images). *Top panels* BMP image in the litho-windows of the AFM. *Bottom panels* AFM images of the pattern obtained on the PET surface. (Reproduced from [38] with permission)

nanolithography is performed on circular, triangular, ellipsoidal and L-shaped areas on the polymer surface, as can be seen in Fig. 24.20.

24.3 Exploiting the Nanowear of Polymers for Lithographic Applications

In this paragraph of the chapter will be shown how the wearing properties of a polymer can be exploited for lithographic purposes, both to directly modify their structure and to induce an ordered self-assembly of their surface. To do this process at the nanoscale, scanning probe microscopy (SPM) results the most appropriate tool because it can be employed both for modification and for imaging of the final patterned structures at high three-dimensional spatial resolution.

The capability of SPM to modify the matter at the nanoscale can be in fact easily employed to exploit the wearing properties of a polymer when a localized load is applied on its surface above a certain "rupture" threshold. More in details, a cantilever tip can locally and permanently modify, in a controlled way, the polymers (bulk structures or thin films) and this capability could be in effect a basic definition of

an SPM-based lithographic process. In this process, the wear characteristics of the polymer play a crucial role, because a nanolithographic process requires very sharp features (cuts, holes, etc.), avoiding as much as possible any detachment of the surrounding polymer (debris) or from the substrate and/or from the bulk polymer during the machining process in a certain location and the wearing of the SPM tip itself must be as much as possible reduced.

Moreover, we could distinguish the localized removal process in two basic wearing mechanisms, the one obtained in literature by scratching the polymer through an SPM tip: "static ploughing" lithography (SPL), where the polymer is removed by applying a static force greater than a certain threshold value, and "dynamic ploughing" lithography (DPL), where an oscillating tips (and force) acts as a sort of "pneumatic hammer" on the polymer surface. We will discuss the results obtained in literature employing both the techniques and the advantage of one respect to the other in terms of wearing effectiveness, tip damages and final spatial resolution of the resulting lithographic process.

A pioneering work in SPL on polymers was the one done by Jin and Unertl (1992) [90] on a polyimide (PI) film employing Si_3N_4 tips and applying a static force the order of 500 nN: in Fig. 24.21A is shown a pattern (University of Maine logo) obtained on the polymer film surface. Later on, the same technique was for the first time successfully applied to fabricate an hybrid semiconductor-superconductor (S-Sm) nanodevice, increasing the spatial resolution of a standard ultra-violet (UV) lithography on a polymeric resist (Shipley S1805): also in this case SPL was employed to remove the resist layer in a very narrow stripe to finally etch the underlying Nb film down to

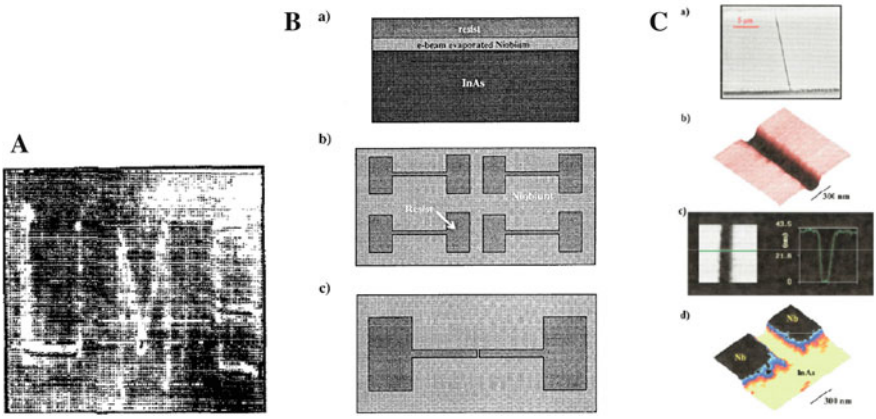


Fig. 24.21 **A** The University of Maine logo "UME" scribed with 110 nm wide grooves on PI film. (Reproduced from [90] with permission) **B** Schematic description of the steps used for the fabrication of weak links: **a** *Cross section* of initial layers; **b** UV lithography-defined resist pattern; **c** final resist pattern after AFM static ploughing of the resist. The pattern in (**c**) is transferred to the Nb film by wet chemical etching; **C** Structural characterization of the weak links fabricated: **a** SEM micrograph of an entire link; **b–d** AFM images of sections of a weak link. (Reproduced from [91], with permission)

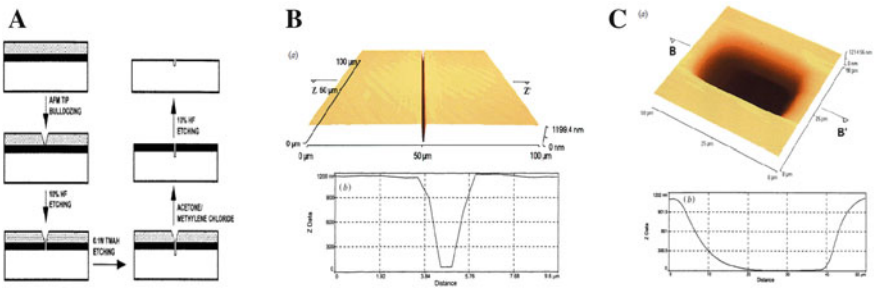


Fig. 24.22 **A** Process flow diagram for the fabrication of a groove (window) on a Si substrate; **B** Stereographical AFM image of a trench of depth 1.080 nm and width 2.500 nm fabricated in the photoresist with a loading of 10 μN, 26 passes of the AFM Si tip and a transverse speed of 10 μm/s. **b** Line profile across ZZ′ of the trench computed by data acquisition software. **C** Three-dimensional AFM image of a $20 \times 20\,\mu m^2$ large window fabricated in the photoresist using xy-raster scanning at a transverse speed of 500 μm/s for 20 min with a loading of 10 μN. **b** Cross-sectional profile across BB' indicating no significant irregular corrugation present at the base of the window. (Reproduced from [92], with permission)

the semiconducting InAs substrate or to create a Nb nanobridge on GaAs substrate by a lift-off lithographic process [91]. The obtained resolution was high enough to measure proximity effect (Cooper-pairs diffusion) from the two S electrodes trough the Sm layer. In Fig. 24.21B it is reported the schematic of the lithographic process and the topography of the fabricated S-Sm-S device (Fig. 24.21c).

A similar procedure was employed by Li et al. [92] to etch the underlying Si substrate by machining trough a Si_3N_4 tip the polymeric photoresist. Cuts and rectangular shapes were obtained on polymer (and finally on Si after HF treatment) and a study of the number of tip scans necessary to completely remove the resist film was accomplished, together with the calculation of the minimum lithographic width achievable, depending on the tip shape and indenting depth. In Fig. 24.22A and B, the schematic and the results obtained by this technique are shown and, in particular, the rectangular shape (Fig. 24.22C) fabricated by xy-raster scanning the tip on a $20 \times 20\,\mu m^2$ large window.

A SPL technique employed to fabricate metallic micro and nano-structures is successfully demonstrated by Porter et al. in 2003 [93] by scratching a polymer resist layer and depositing gold stripes by galvanic displacement. The schematic of the whole process is reported in Fig. 24.23A while the AFM characterization of the different lithographic steps is shown in Fig. 24.23B. More complex structures have also been obtained with the same process and (depending onthe scratching conditions) continuous lines or a series of nanoparticles can be deposited.

In 2004 Lekki and coworkers [94] completed the route toward a standardized SPM lithographic technique by introducing a “color encoded” vector-based nanolithography in SPL mode, where parameters like height, velocity, feedback force, etc. were coded in the colors of a WMF file onto the RGB channels of the image, establishing in this way a distinct relation between a graphical feature (color) and the

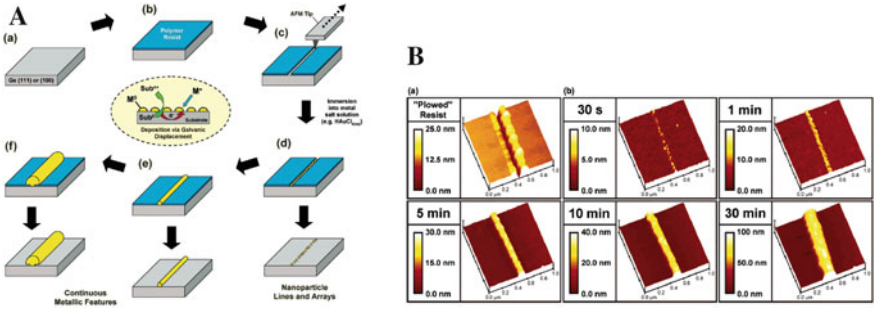


Fig. 24.23 **A** Metallic nanostructures via static plowing lithography: a degreased Ge(111) substrate (**a**) is coated with a thin polymer resist layer (**b**), which is subsequently plowed away by utilizing the tip of an atomic force microscope (**c**). Immersion of the substrate into a dilute, aqueous $HAuCl_4$ solution for a brief time provides for the deposition of discrete nanoparticles onto the exposed Ge(111) surface via galvanic displacement (**d**). As deposition proceeds with longer immersion times, grain growth and nanoparticle coalescence eventually yield continuous metallic structures (**d**), which continue to increase in size with extended periods of electroless plating (**f**). Once the desired features are realized, the resist is completely removed with a solvent rinse. The inset represents a simplified schematic of the galvanic displacement deposition mechanism. **B** Intermittent contact (tapping) mode atomic force micrographs illustrating a resist furrow produced by static plowing (**a**) and the gold nanostructures (**b**) on Ge(111) resulting from increasing immersion times in 25 μM $HAuCl_4$ (aq) at 25 °C following resist removal. (Reproduced from [93], with permission)

used nanolithography scheme (applied voltage, force, tip height, etc.). This approach was successfully employed on PMMA to obtain complex nanopatterns by SPL (see Fig. 24.24). Actually, almost all the modern SPM-controllers have similar software tools to perform lithography by vector scan the cantilever tip on a sample surface.

It can be noticed in all the experimental works employing SPL as lithographic technique that the main features of a static scratching are the presence of debris, ripple-like structures, deposited material along the trench (bulges), and deposited material at the end of the machined region. An interesting work, mainly focused on these aspects of SPL on polymers, is the one of Blach et al. [95], where lateral force and force vs distance (FvsD) information were acquired during and after the lithographic process in order to clarify some aspect of the removal process. In their work, they underline how manipulation of soft surfaces such as polymers is likely to depend on both out-of-plane and in-plane forces acting at the point of contact between the tip and the polymer. Moreover, F vs D curves demonstrated how plastic deformation of the studied polymer (Poly-tert-butylmethacrylate, P(tBuMA)) represent only 10 % of the total dynamic deformation and, therefore, elastic recovery can reach up to 90 % in bulk polymers (see Fig. 24.25A). Finally, lateral force analysis in the removed region of the polymer revealed that the friction signal resulted lower than on the original polymer surface. This lateral force contrast, in their interpretation, represents differentiated surface chemistry, thus the effect of tip-induced surface manipulation is also to alter the surface chemistry, presumably due to chains scission caused by the polymer wearing at the nanometer scale (Fig. 24.25B).

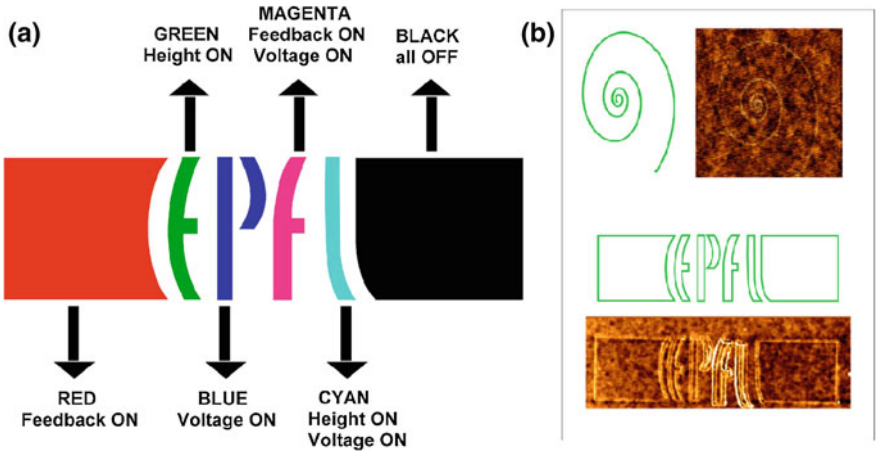


Fig. 24.24 **a** The image shows how different colors will be interpreted by the SPM system. Red corresponds to feedback force, blue to bias voltage and green to tip height; **b** The Figure (an AFM scan) displays a nanolithography pattern obtained using a WMF image: *Green* color corresponds to a constant height mode. *Red* value of color is 0, *green* value is 199, and *blue* is 0. This means feedback off, voltage off and height of 2.133 μm (when the range for height is plus–minus 4.0 μm and plus signifies that the probe moves towards the sample). (Reproduced from [94], with permission)

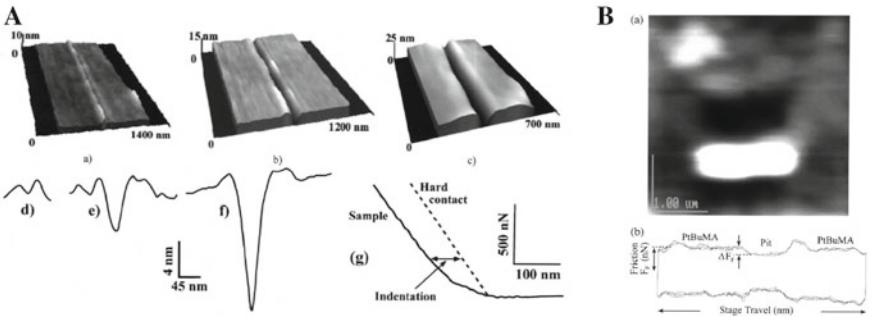


Fig. 24.25 **A** (**a–c** Topographic images of P(tBuMA) surfaces after single line scans (with the long axis of the lever aligned with the scan direction). **d–f** Corresponding line profiles for lever-induced normal force loadings of 60, 260 and 450 nN. The linear scan speed was held constant at 100 μm/s. **g** The F–d curve shows extent of quasi-static tip indentation as a function of force loading (the 'hard' contact calibration curve was obtained from the glass substrate). **B a** Topographic image of pit arising from rastering at a normal force of 1 μN. **b** The trace represents four repeat line scans across the field of view in **a** along the y-axis. A lower lateral force was observed when the tip traversed the region of the pit. (Reproduced from [95] with permission)

Finally, while cutting a trench into the resist by SPL, i.e. with the tip in the AFM contact mode, torsion of the cantilever may lead (as previously underlined) to edge irregularities, which restricts the cutting direction to a certain range around the cantilever axis, we will see how indenting the surface by a vibrating tip in the AFM

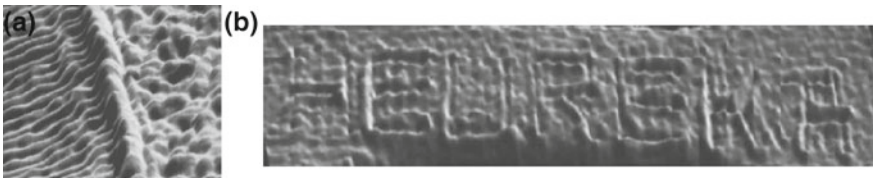


Fig. 24.26 **a** Grooves obtained on a PC surface by DPL technique and **b** the word "HEUREKA" written on the same substrate in lithographic mode: letter height is 700 nm while the indentation depth results 10 nm. (Reproduced from [96], with permission)

tapping mode, so called dynamic plowing lithography (DPL), provides a lithographic technique that is free from directional restrictions and reduces a lot the tip wearing phenomenon.

The first example of this DPL technique applied on polymers can be found in the work of Jung et al. in 1992 [96], where a Si_3N_4 tip under a load of 100nN was employed to plough grooves of 10 nm deep and 70 nm wide on PC and the word "HEUREKA" in lithographic mode (see Fig. 24.26a and b).

In 1994 Wendel et al. [97] performed a similar experiment employing a thin resist layer (Shipley S1805) as a mask for the underlying semiconductor substrate. More in detail, a pattern of holes in the resist was obtained sending a triangular voltage pulse of about 0.1 s duration to the z-piezo in a standard tapping mode configuration, causing the vibrating tip to be pushed nominally about 100 nm against the surface and indenting in this way the polymer. The groove geometry often reflects the triangular shape of the used silicon tip and reaches down to the GaAs/AlGaAs surface. Figure 24.27A shows an AFM image of an array of 16 holes with a period of 55nm. To transfer the AFM generated mask pattern to the two-dimensional electron gas underneath they either used an additional wet chemical etch process or the ion beam irradiation technique. In Fig. 24.27B a transferred array of antidots on a Hall bar is also reported.

As in the case of SPL, also with DPL there was an interest in studying its peculiarities in terms of lithographic capabilities and effects on the sample surface. More in detail, in the work of Klehn and Kunze [98], a vector scan based lithographic software was employed to study the advantages of DPL in terms of tip wearing, piezo nonlinearity, scan velocity, lateral resolution. Controlling of the vibration amplitude and tip movement enables one to plow a narrow furrow along line segments of arbitrary length and direction. Alignment of the designed line pattern is easily achieved in the AFM imaging mode. A vector-scan controlled moving of the vibrating tip enabled a reliable pattern transfer in the thin resist layer which subsequently served as etch mask for semiconductors (Si, GaAs) and metals (Ti, Au). A time delay of a few milliseconds was also introduced in this case between the tip scan from a segment to the other one, demonstrating in this way that is possible to increase the lithographic velocity without any pattern distortion (see Fig. 24.28 for a comparison). We will see in the following how surface acoustic waves could be successfully employed to these purposes.

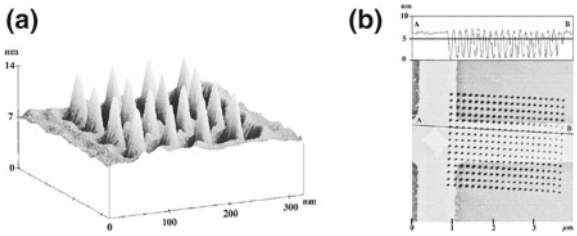


Fig. 24.27 AFM image of a pattern of 16 holes with a period of 55 nm created by AFM lithography in photoresist. **A** To better illustrate the shape of the holes the image is inverted, i.e., the holes appear as mounds. Hole depth as well as resist thickness are 7 nm; the holes thus reach the sample substrate. **B** AFM image of an array of antidots fabricated into a GaAs-AlGaAs heterostructure with AFM lithography and wet etching. Mesa depth and antidot depth are 5 and 6 nm, respectively; the period of the antidots is 145 nm. *Top* section plot along the indicated line. (Reproduced from [97] with permission)

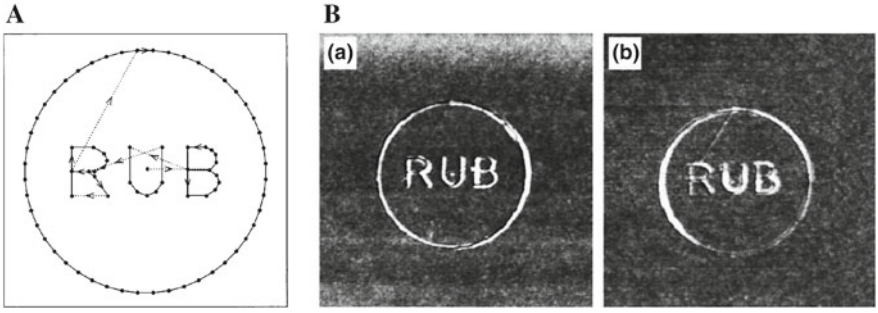


Fig. 24.28 **A** Edited design of characters written as non-contiguous line pattern surrounded by a 48-gon. Arrows indicate the direction of moves beginning at the *center*, *dotted* and *solid* lines respectively represent moves with low and high drive amplitude of the tapping vibration, and solid circles denote "wait" positions. The field dimension is $1.8 \times 1.8\,\mu m^2$. **B** AFM images $\sim 2.8 \times 2.8\,\mu m^2$ of the pertinent resist pattern after plowing with scan speed **a** and **b** $\upsilon = 0.1\,\mu m/s$. In **a** the t_{wait} is zero while in **b** correspond to 10 ms. *Dark lines* represent the plowed furrows, *light lines* embankments due to displaced resist; the total z range is 10 nm. (Reproduced from [98] with permission)

Another example of DPL can be found on Wiesauer and Springholz [99], where the mechanical properties of a thin polymeric UV resist layer (10 nm thick, Shypley MF319) are tuned by a proper hard baking treatment of the sample in order to decrease the reflow and the refilling process during and after the fabrication of holes arrays. They also studied the bulge formation mechanism around the holes (due to the tilt of the cantilever with respect to the sample surface and to the cantilever distortion caused by rather high indentation forces) and its influence on the final spatial resolution achievable. In Fig. 24.29 a schematic of the bulge formation is presented together with a fabricated array of holes, with various patterning parameter (diameter, depth, and width) mapped in function of the applied loads (z-scanner extension). As can be

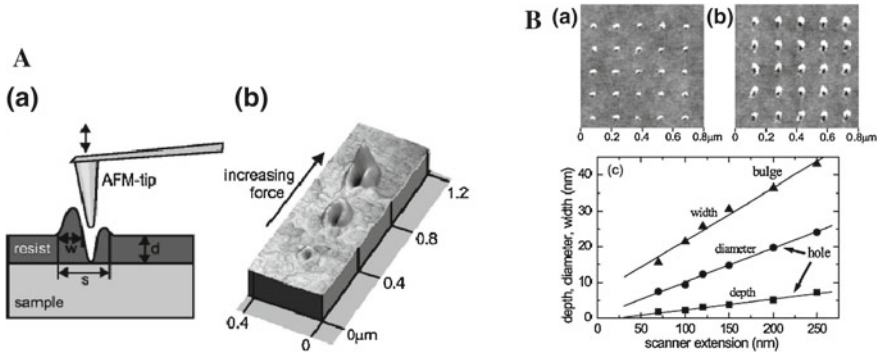


Fig. 24.29 **A** Schematic illustration of the indentation process of an AFM tip into a resist layer with thickness d **a**, showing the formation of hole and bulge; **b** AFM image of holes created with increasing indentation force, leading to larger holes and bulges; **B** *Top panel* Grids of holes produced by nanoindentation of a contact ultralever (spring constant 1.6 N/m) into a 350 nm thick photoresist layer using a scanner extension of **a** 100 nm and **b** 200 nm for indentation. **c** Dependence of the hole depth and diameter, as well as the width of the bulge as a function of the scanner extension. Prior to the AFM modification the resist was subjected to a hard bake step at 130 °C for 30 min. (Reproduced from [99], with permission)

noticed, it results that a linear behavior can be extracted and that in order to obtain high resolution very low forces (and a very thin resist layer) must be employed.

In the work of Cappella et al. [100] the effect of DPL on the surface properties of PMMA and PS is also studied by using FvsD adhesion maps. These Authors found that the structuring of a polymer surface by means of DPL, in this case, produces carved surfaces surrounded by large border walls (bulges), whose volume is bigger than the volume of the carved out regions. They speculate that such a 'creation' of volume is due to changes in the density of the polymer, caused by a loosening of the structure of the polymer provoked by the fast indentation of the AFM tip. The analysis of the border walls by means of force–displacement curves showed in this case that the changes of the physical structure of the polymer are accompanied by changes in its chemical nature (Fig. 24.30). They confirmed these results also through size exclusion chromatography (SEC) measurements and by evaluating the energy required to break the polymer chains, demonstrating in this way that the indentation energy results sufficient to cause the broken of polymer covalent bonds.

Dynamic ploughing technique has been also employed by Balocco et al. [101] on a semiconducting polymer (poly(3-hexylthiophene, P3HT). In Fig. 24.31 cuts, holes and large geometric features are shown, which were obtained by DPL. Authors affirm that the internal tensile strain in the films (deposited on the substrate by spin coating technique) and the long P3HT molecules allow eliminating in this case all major common problems of the nano-indentation methods, namely the refilling of the trenches by debris, tip contamination by debris, and the short AFM tip life time. Successful pattern transfer to the underneath inorganic semiconductor was achieved by a wet chemical etch with the created organic nanostructures employed

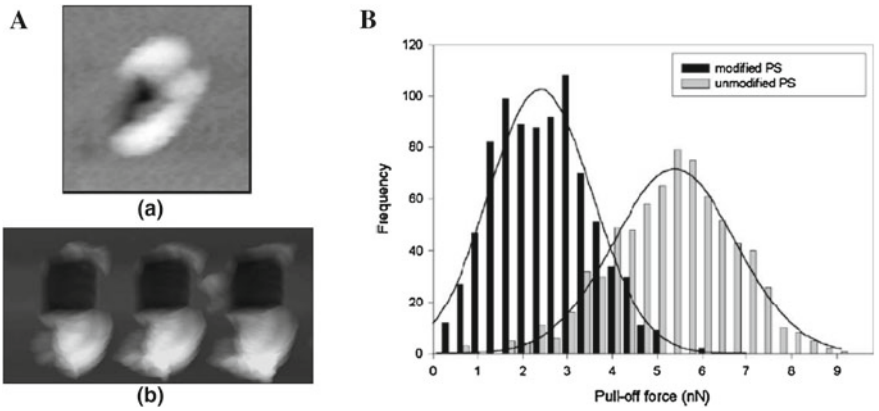


Fig. 24.30 **A** A single hole carved in PMMA through DPL **a** the hole is surrounded by a border wall. **b** *Three squares* carved on PS through overlapping of single holes: the squares are partly surrounded by large border walls, resulting from the deposition of the border walls of the single holes. The amplitude used to carve the squares is smaller than in the case of the single hole shown in (**a**). **B** Histograms of the measured pull-off force on modified (*black bars*) and unmodified (*grey bars*) PS, collecting the pull-off force of 1600 F versus D curves. The mean force is 5.4 nN for unmodified PS and 2.4 nN for modified PS. (Reproduced from [100] with permission)

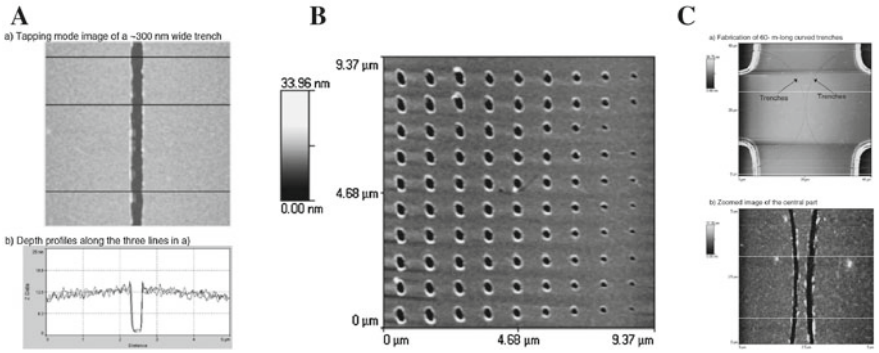


Fig. 24.31 **A** A $5 \times 5\mu m^2$ AFM image of a single trench created by nanoindentation on a thin P3HT film (**a**), and (**b**) the depth profiles along the three *horizontal lines*. The flat bottom of the trench shows that the internal tensile strain effectively pulls the film apart during the lithography, resulting in a trench much wider than the diameter of the AFM tip. **B** An array of holes on a P3HT film fabricated by nanoindentation with different nominal vertical scanner displacements from 0.1 to 0.7 µm. This results in different mechanical impacts that loosen the P3HT film around the holes from the substrate and therefore different hole diameters. **C** Long insulating trenches fabricated by nanoindentation on a cross junction of a thin P3HT film. Each of the semicircular trenches is around 60 µm long, showing good capability of fabricating structures over a long spatial distance. (Reproduced from [101] with permission)

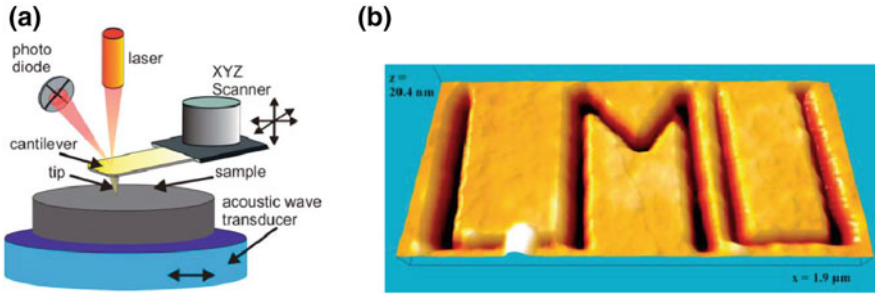


Fig. 24.32 **a** Schematic of the experimental setup for acoustical force nanolithography (AFN). The sample holder consists on an acoustic wave transducer that is used to enhance cantilever flexural vibrations for lithography. **b** A nanostructure generated by acoustical force nanolithography: lithographed emblem of the Ludwig-Maximilians-University. (Reproduced from [102] with permission)

as the etching mask. Furthermore, no obvious degradation of the AFM tip either by debris contamination or mechanical wearing was observed after many days of nanolithography. This allows nanostructures over tens of microns in length to be reproducibly fabricated in large numbers.

An interesting approach to DPL is the one employed by Rubio-Sierra et al. [102] and named "acoustical force nanolithography" (AFN), where the some limitation of plowing are overcome by using surface acoustic waves (SAWs) propagating trough the sample. In standard DPL the fixed end actuation of the cantilever introduces a time-delay in the system between the actuation and the tip. This can limit the bandwidth of high-speed closed-loop control of the machining process, as seen previously [98]. In this AFN technique, the cantilever excitation is therefore obtained by coupling a normal incidence shear acoustic wave to the cantilever tip through sample surface. The acoustic wave is generated by an acoustic transducer below the sample itself (Fig. 24.32a). A thin polymer film (ma-p 1205, based on Novolak, naphthoquinonediazide, and solvents) was machined at various SAW amplitude and frequencies related to the resonances of the surface-coupled cantilever. The relationship between the force feedback set-point and the SAW coupling was also studied. One example of a complex pattern obtained by AFN is presented in Fig. 24.32b.

Another variation on DPL standard approach was the one named "AFM Tip-hammering Nanolithography" (ATHN) employed by Wang et al. [103] for patterning at very high resolution polystyreneblock-poly(ethylene/butylenes)-block-polystyrene (SEBS) triblock copolymer. In this technique the vibrating AFM tip is used as a nanohammer to forge either embossed or imprinted patterns in structure-tailored SEBS thin films with a sub-20 nm linewidth resolution at low cost and high speed. In addition, patterns written by ATHN can be erased by thermal annealing at 70 °C for less than 5 min, and if necessary, the erosion process can be expedited by increasing the annealing temperature. A possible mechanism that allows generating and erasing this very high resolution structures on SEBS is explained in Fig. 24.33A. In the side-view, the tailored structure of hex-spherical PS microdomains embedded

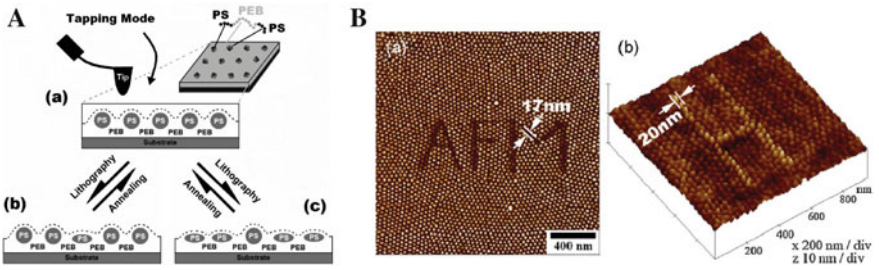


Fig. 24.33 **A** Schematic representation of the mechanism for generating and erasing ATHN patterns. **a** Profile of a vibrating AFM tip over a SEBS monolayer thin film. It can be seen that the tailored hex-spherical PS microdomains are embedded in the PEB matrix. The solid line is an untouched SEBS flat surface. The dashed line is an AFM tip-indented surface under light tapping conditions where the lateral differences in indentation depth between pure PEB and embedded PS microdomains leads to the observed hex-spherical PS arrays as shown in the top-view scheme. **b** Profile of an imprinted line generated by using hard tapping force to deform a single inside PS microdomain in the middle from sphere to ellipsoid. **c** Profile of an embossed line generated in the middle by using hard tapping force to deform PS microdomains on either side. Erasion of the ATHN patterns is done by thermal annealing the SEBS sample to recover the indented structure of either (**b**) or (**c**) to the original (**a**). **B** Two examples of complex ATHN patterns generated manually in SEBS G1650 thin films. **a** AFM phase image of imprinted "AFM". **b** AFM topographic image of embossed "H". (Reproduced from [103] with permission)

in PEB matrix can be seen, and the untouched flat surface is represented by the solid line. After an AFM tip taps the sample surface at an appropriate tapping force level, the lateral difference in tip-indentation depth between pure PEB and embedded PS microdomains leads to the profile of spherical PS microdomains as represented by the dashed line. However, when the oscillating AFM tip is used as a nanohammer to forge the surface of SEBS at a very high tapping force level, the plastic PS microdomains can be deformed from spheres to ellipsoids, and selectively deforming PS spherical arrays leads to the imprinted and embossed patterns generated in SEBS thin films as shown in Fig. 24.33B. Although ATHN and static or dynamic tip plowing are both categorized as mechanical force nanolithography, their working principles, based on forging and plowing, respectively, are quite different and ATHN is limited to plastic deformation-based mechanism on polymers.

More recently, in order to additionally overcome some residual problems related to DPL (presence of debris, tip contamination and trench refill by debris) Lu et al. [104] have introduced during the lithographic process a thermal-annealing process above the glass-transition temperature of the PMMA film on semiconductor sample. The dramatic improvement of the patterned profiles after the thermal treatment, demonstrated in Fig. 24.34 where DPL-fabricated structures with and without thermal treatment are shown, enhances the adhesion of the PMMA film, and hence, also eliminates the underetching problems of the substrate (GaAs in this case). These Authors successfully applied their technique by demonstrating fabrication of working nano-devices: nano-diodes with a high rectification ratio were achieved at room temperature.

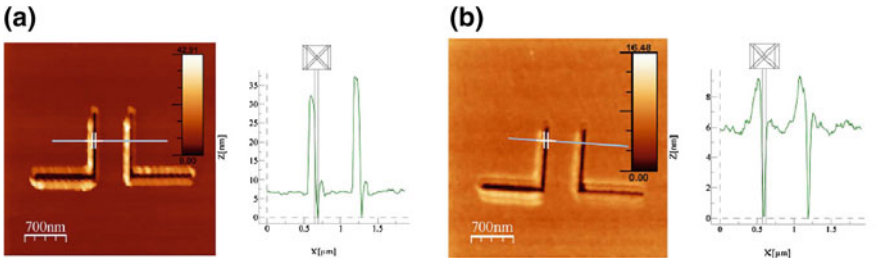


Fig. 24.34 AFM topography images of the same nanoploughed trenches into PMMA **a** before any annealing was carried out and **b** after annealing at 240 °C. The height of the debris was dramatically reduced from about 30–2 nm without undermining the feature definition of nanotrenches. (Reproduced from[104] with permission)

As seen, we discussed SPL and DPL as high resolution lithographic techniques exploiting the wearing properties of polymers at the nanoscale. At the moment, DPL technique still stays on the experimental stage, with all the discussed advantages respect to SPL, but a deeper investigation of the interaction between the tip and the surface is scarce. In the process at the nano-scale involving dynamic ploughing, the force applied on the surface by the tip directly affects the nanomachining process and imaging at the same time, and many other parameters are involved in the nanolithographic process. Liu and coworkers [105], very recently discussed from theoretical point of view how the machining parameters can have influence on lithographic results. Therefore, for better controlling the manufacture process and improving the machining accuracy, they have studied the "tapping" process of an AFM silicon tip, also simulated as a driving oscillator with damping. The factors influencing the tapping force and the polymer sample deformation based on the Hertzian model are also studied. More in details, the effect of driven amplitude, tip radius R, driving frequency, tip-sample separation, cantilever spring constant, and polymer's mechanical and wearing properties are investigated. Regarding these last factors, the focus of our review, they conclude that dull tips may result in failure of the nano dynamic ploughing process. For example, with R of 150 nm the tip cannot plastically deform the polymer surface with a Young's modulus of 2.2 GPa. Simultaneously, in order to ensure consistency of machined structures, it is a good way for selecting a suitable tip radius making the tip wear to keep at a low level. In Fig. 24.35a the relationship between the response amplitude (related to the indentation depth) and the tip radius is shown. It can be noticed that with the increase of R, there is a decrease first and then an increase in the response amplitude (the response amplitude varies from 62.8 to 62.1 nm, a total change of 0.7 nm). For $R < 80$ nm, the response amplitude decreases rapidly. The reason is that the repulsion action between the tip and the surface plays a major role on the oscillator motion at this stage and in the model employed. Moreover, the tapping force increases with the spring constant of the cantilever and the smaller the tip radius, the larger the tapping force is. Therefore, a sharper tip and a harder cantilever should be utilized for a deeper machined

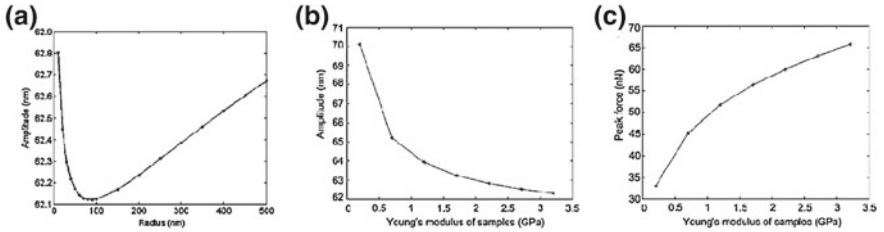


Fig. 24.35 **a** Relationships between the response amplitude and the tip radius. **b** Response amplitude and **c** the peak force applied to polymer samples in function of their Young's modulus. (Reproduced from [105], with permission)

depth in the nano dynamic ploughing process with an AFM vibrating tip. Finally, the local mechanical property of the materials plays a significant role in the AFM-based tapping mode nano machining process. A polymeric sample with a larger Young's modulus will lead to a larger tapping force, a larger maximum contact pressure and a smaller machined depth in order to be locally modified. In the variation range of the Young's modulus of 0.2–3.2 GPa, representing the typical polymer materials used in the field of nanofabrication, simulation results showed that the polymer surface can be successfully machined by the nano DPL technique as expected, but there are some interesting features in the machining process. As shown in Fig. 24.35b, the response amplitude decreases with the Young's modulus of the polymer. It means that with a larger Young's modulus, the deformation of the sample surface is small (the original distance between the tip and sample is set 60 nm in this case). But the corresponding tapping force increases, as shown in Fig. 24.35c. This means that although the tapping force increases when the Young's modulus of the sample is greater the deformation or the machined depth is small, which is therefore not only determined by the tapping force parameter. In the reported theoretical simulations, the hard silicon tip can be looked as a rigid body respect to the polymer but in some experimental conditions the interactions with very hard substrate underlying the polymer film must be also taken into account to avoid tip wearing phenomena.

In the next paragraph we will discuss how wear at the nanoscale results important in some biomedical application.

24.4 Characterization of Meso- and Nanoscale Wear of Polymers in Biomedical Applications

Medical practice today utilizes polymers as biomaterials in a large number of devices and implants. Implants (sutures, bone plates, joint replacements, ligaments, vascular grafts, heart valves, intraocular lenses, dental implants, etc.) and medical devices (pacemakers, biosensors, artificial hearts, blood tubes, etc.) are widely used to replace and/or restore the function of traumatized or degenerated tissues or organs, to improve

function, and thus improve the quality of life of the patients. Surface properties of biomaterials on nanoscale level should be known to have a satisfactory response at the issue/biomaterial interfaces, thereby minimizing or eliminating tissue trauma on macrometer scale [106]. Indeed, thesurface of an implanted polymer, such as any other material, is not perfectlysmooth on a microscopic scale but, rather, has small asperities on the surface. Mechanical contact is localized and supported by the asperities. Thus, a relatively low contact pressure for the entire surface can result in very high local pressures relative to any single asperity. Such localized contact pressures can result in adhesion between the asperities of two surfaces in relative motion.After adhesion, subsequent movements can provoke the formation of debris orsmall fragments. These fragments may react with other chains to form side branches, or react with other chains to form crosslinks increasing the *thrombogenetic* factor, i.e. the partial or total occlusion of vascular *vasi*, and the inflammatory regime.

In this section, we will focus the attention on the incidence of wear on two different kinds of polymers: (1) biodegrabrable polymers (for application of relative low wear), [107], and (2) cross linked Ultra High Molecular weight Polyethylene (UHMWPE) for application where is required high wear resistance [108]. The first group of polymers isused as biomaterials, surgical suture and implants.The second one is the basic material of acetabular cup of hip prostheses and of meniscus in knee prostheses.

As described in the previous sections, with the advent of SPM family it became possible to study dissipation processes in small sliding contact areas down to the atomic and molecular scale. When the SPM instrument is operated in contact mode, several regimes from frictionless sliding to permanent wear are observed, depending on the applied load. Among the SPM family, AFM has become a powerful toll for high resolution imaging of biological objects, because that AFM can operate in aqueous environment [109]. For polymers to be used in biomedical applications the knowledge deriving from AFM tip induced wear tests performed in physiological conditions, (normally, a temperature range of $20 - 40\,^{\circ}\mathrm{C}$, atmospheric pressure of 1 pH of 6–8, glucoseconcentration of 1–20 mM, atmospheric oxygen concentration) may be fully investigated addressing accurate knowledge for their performance in vivo. The first one regards the connection between in vivo chemical biodegradation of polymer and their mechanical degradation induced by contact surfaces in relative motion one each other. The second one is that the macroscale traditional wear tests fail when applied to polymer sample in liquid.

24.4.1 Role of Wear Rates for Biodegradable Polymers

There has been a substantial increase in the use of biodegradable polymers in the last decade [110]. Initially, they were employed as carriers for drug-delivery devices, sutures and as a temporary joint spacer (poly(ε-caprolactone)). However, the tribological characteristics of these polymers have not been fully investigated. In the studies focused on identifying the initiating causes offailures when biomedical polymers are used, two aspects can limit the use or efficiency of such polymers: chemical

degradation and/or wear. Chemical degradation starts breaking the long polymer chain into smaller fragments, or causes chain scission. Alternatively, some enzymes and ionizing radiation are also capable of attacking polymers. Thus, the polymer may be reduced in molecular weight, increasing its solubility or it may become harder and more brittle due to crosslinking. Relative motion between parts can cause mechanical damage and release of small debris due to wear. In addition, the phenomena occurring at the surface of biodegradable polymer due to the interplay of chemical and mechanical stresses could be significantly different with respect the analogous mechanisms at the bulk. Indeed, there has been speculation that the glass transition temperatures of polymer surfaces differ from those of the bulk, suggesting that viscous behavior may be observed at the surface at temperatures where the bulk of the material is glassy.

The AFM tip induced wear test for the analysis on onset mechanisms of wear on polymer surfaces consists in the repeated scanning of a prefixed area of the film surface at constant force. Studies of tip-induced wear behavior may be valuable in assessing the mechanical properties of polymers surfaces. Moreover, for some classes of thin film materials, such as plasma polymers, the application of bulk mechanical testing methods is not feasible and other techniques associated with surface mechanical investigation, such as nanoindentation, are inappropriate.

A particular class of biomedical polymers is composed by co-polymers with non-polymers fillers such as drug molecules. The mechanical properties of such biomedical materials must be accurately known for satisfactory response when used in the human body. PMMA, in fact, is known to be an excellent resistant polymer to wear nearly as PE or UHMWPE polymers. Nevertheless, the copolymer PMMA-co-AA (AA denotes Acrylic Acid monomers) shows a different wear response to the AFM tip wearing test. It has been shown that the plastic deformations in such copolymer in such tests are drastically increased by the presence of low percentages of drug (nearly 1 %) [111].

The connection between wearing and chemical degradation of biodegradable polymers used in health application, in particular in the human body, will require many efforts and specific studies. Many of such knowledge should be addressed in a special way on a scale ranging from micro- to nano-scale due to the specific scales of many biosystems such as cells, proteins, enzymes, etc.

24.4.2 Severe Wear Regime in Biomaterials: Wear of UHMWPE Used in Prostheses

To reduce the clinical implications of wear particles, one needs to reduce the production of these particles by improvement of the design, materials, and surgical technique. UHMWP remains the material of choice for the bearing surface in total joint replacement components. In its conventional form, UHMWPE has exceptionally mechanical integrity owing to its chain entanglements, high tie molecule density,

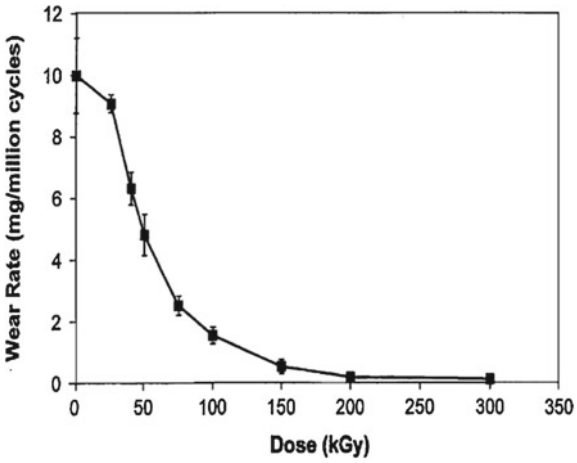


Fig. 24.36 Wear rate of crosslinked UHMWPE as a function of radiation dose. (Reprinted from [108], with permission)

moderate cristallinity and very high molecular weight. Wear debris generated from UHMWPE hip joint replacements has been identified as the major cause of osteolysis and loosening, leading to eventual revision surgery [112]. UHMWPE wear resistance has been improved using crosslinking methods [113–116]. Crosslinking is obtained by exposing the polymer to high doses of gamma radiation or e-beam radiation along with a thermal treatment that can be done above or below the melt temperature (Fig. 24.36). Recent experiments have shown that high degrees of crosslinking in UHMWPE result in a reduction of several important mechanical properties including strength, ductility, elastic modulus, fracture toughness, and crack propagating resistance and relevant reduction of abrasive- and adhesive-wear mechanisms in several in vitro joint simulator studies, for a review of methods and results see [111]. Another critical reason for the revision of orthopaedic implants is the failure of the crosslinked PE components for delamination, which is mainly due to the oxidation of UHMWPE under irradiations [117]. Generally, the interaction between radiation and UHMWPE leads to the scission of C–C (carbon-carbon) and C–H (carbon-hydrogen) bonds. Since during irradiation a number of radicals are formed at a very high rate, oxidation proceeds much faster than in unirradiated UHMWPE doing possible a direct correlation between the rate of macroradicals formation and the oxidation degree of the prosthetic component.

One critical factor for the onset of wear in UHMWPE prosthesis is the cross-shear effect, which refers to local counterface motion transverse to the direction of strain hardening of PEcreated by the joint articulation during sliding [113]. Under linear tracking motion, the molecules of UHMWPE are stretched along the direction of sliding and then orient in that direction, leading to a significant degree of strain hardening in that direction, which results in an increase of wear resistance in that direction [118]. Under conditions of multidirectional cross-shear motion, surface molecules align preferentially in the principal direction of sliding. Strengthening in one particular direction leads to weakening in the transverse direction, this weakening

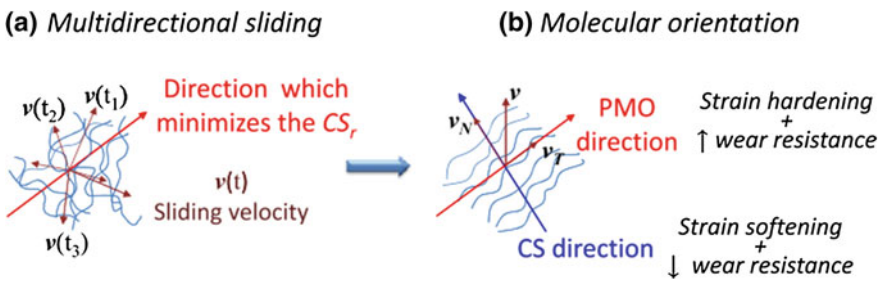


Fig. 24.37 Schematic sketch of the cross-shear effect. **a** UHMWPE polymeric chains, initially randomly oriented, **b** re-orient in the PMO direction because of multi-directional sliding on a harder counter-face. (Reprinted from [120] with permission)

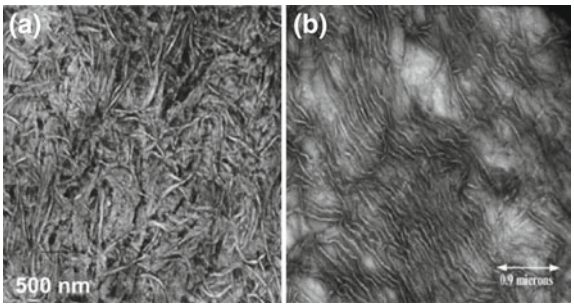


Fig. 24.38 Transmission electron microscopy image of **a** random lamellae organization in UHMWPE without mechanical deformation; **b** texture development of lamellae in direction of sliding. (Reprinted from [108] with permission)

is known as orientation softening, which accelerates the generation of wear debris [119, 120] (Figs. 24.37 and 24.38).

The next developments about the cross-shear effects and multidirectional sliding wear of UHMWPE should be addressed to quantify the correlation of frictional work, mechanical stresses and re-organization of molecular chains at different scales. Such research will require wide theoretical, numerical and experimental efforts which could lead to new high-performance materials to be used as wear resistance in the human body.

24.5 Conclusive Remarks and Future Perspectives

The chapter has been devoted to give a review about experimental data and correspondent modelson the wear mechanisms occurring on nanoscale on polymer materials. Wear is the progressive loss of mass from the operating surface of a body as a result of relative motion at the surface. The occurrence of different wear

mechanisms involves the fundamental mechanics of molecular and supra-molecular scale. As a consequence, a fundamental understanding of surface properties on nanoscale level should be generated to have a satisfactory knowledge of responses of materials and machine components also at macrometer scale. Investigation of the fundamental characteristics of wear at the microscale is complicated by some factors and forces that act on a nanoscale level, which have not yet been addressed in the tribology of macrosystems. Since these forces are sensitive to the environment and surface condition of the specimens, it is quite difficult to determine them accurately. The wear of the polymer surface caused by the tip in regime of single asperity contact (the contact condition reached by a scanning probe microscope) is an articulate process that depends on the contact conditions, as applied forces, tip size, or molecular surface polymer structure. Since the influence of all these parameters is in close connection with the sample properties, one can also expect a dependence of the wearing process on the sample mechanical properties thatgenerally vary significantly from the bulk properties.

Further, quantification of wear is not straightforward since the amount of wear is often too small to be detected by surface-sensitive instruments. The advent of AFM opened a powerful opportunity to study wear mechanisms on nanoscale, making possible to simulate one single asperity surface sliding on other one and controlling many physical parameters during the contact dynamics. The wear rates are generally lower with respect to macroscale, and such rates increase with increasing load. The chapter provided a presentation of nanowear occurrence on polymers induced by means of AFM tips presenting (1) a general introduction to the mechanical properties of polymer films; (2) wear properties at the nanoscale; (3) creation of nanolithographic patterns; and (4) the characterization of meso- and nanowear of polymers used in biomedical applications. In the following years, with the possibility to address new dedicated experiments, it is reasonable to suppose that our knowledge on the basic wear mechanisms and their control can open new technological challenges and opportunity on nanoscale.

Acknowledgments M.D. wishes to acknowledge NanoICT laboratory for hospitality during the review composition.

References

1. B. Bhushan, *Nanotribology and Nanomechanics II*, (Springer, Berlin, 2011)
2. B. Bhushan, *Handbook of Nanotechnology*, 3rd ed. (Springer, New York, 2010)
3. S.K. Sinha, N. Satyanarayana, S.C. Lim, *Nano-Tribology and Materials in MEMS*, (Springer, New York, 2013)
4. R.W. Carpick, M. Salmeron, Chem. Rev. **97**, 1163–1194 (1997)
5. E. Gnecco, E. Meyers, *Fundamentals in Friction and Wear on Nanoscale* (Springer, Berlin, 2007)
6. B.J. Briscoe, S.K. Sinha, *Proceedings of the institution of mechanical engineers* Part J. J. Eng. Tribol. **216**, 401–413 (2002)
7. K. Friedrich, P. Reinicke, Mech. Compos. Mater. Struct. **34**, 503–514 (1998)

8. H. Shulha, A. Kovalev, N.K. Myshkin, V.V. Tsukruk, Eur. Polymer J. **40**, 949–956 (2004)
9. N.K. Myshkin, M.I. Petrokovets, A.V. Kovalev, Tribol. Int. **38**, 910–921 (2005)
10. D.L. Burris, B. Boesl, G.R. Bourne, W.G. Sawyer, Macromol. Mater. Eng. **292**, 387–402 (2007)
11. M. Chandrasekaran, A.W. Batchelor, Wear **211**, 35–43 (1997)
12. N.A. Wright, S.N. Kukureka, Wear **251**, 1567–1578 (2001)
13. R. Prehn, F. Haupert, K. Friedrich, Wear **259**, 693–696 (2005)
14. A. Dasari, Z.Z.Yu, Y.W. Mai, Wear and scratch damage in polymer nanocomposites, in *Tribology of PolymericNanocomposites: Friction and Wear of Bulk Materials and Coatings* 2nd edn. ed. byK. Friedrich, A.K. Schlarb (Butterworth Heinemann, Oxford, UK, 2013). pp. 551–570
15. J. Jagur-Grodzinski, React. Funct. Polym. **39**, 99–138 (1999)
16. S. Ramakrishna, J. Mayer, E. Wintermantel, K.W. Leong, Compos. Sci. Technol. **61**, 1189–1224 (2001)
17. D.D. Woodland, W.N. Unertl, Wear **203–204**, 685–691 (1997)
18. J. Schöfer, E. Santner, Wear **222**, 74–83 (1998)
19. T. Kaule, Ph.D. thesis, Johaness-Gutenber-Universität, Mainz, Germany (2012), http://d-nb.info/1031360751/34
20. F. Dinelli, G.J. Leggett, P.H. Shipway, Nanotechnology **16**, 675–682 (2005)
21. J.L. Gilbert, J.D. Wernle, *UHMWPE Biomaterial Handbook* (2009)
22. O.M. Leung, C. Goh, Science **255**, 2407–2410 (1992)
23. Z. Elkaakour, J.P. Aimé, T. Bouhacina, C. Odim, T. Masuda, Phys. Rev. Lett. **73**, 3231–3234 (1994)
24. R.N. Leach, F. Stevens, C. Seiler, S.C. Langford, J.T. Dickinson, Langmuir **19**, 10225–10229 (2003)
25. B.K.P. Wong, S.K. Sinha, J.P.Y. Tan, K.Y. Zeng, Tribol. Lett. **17**, 609–615 (2004)
26. R.B. Mohamed Sani, S.K. Sinha, J.P. Ying Tan, K.Y. Zeng, Philos. Mag. **85**, 2101–2122 (2005)
27. A. Rubin, C. Gauthier, R. Schirrer, Wear **303**, 40–48 (2013)
28. R. Kaneko, E. Hamada, Wear **162**, 370–377 (1993)
29. A. Khurshudov, K. Kato, J. Vac. Sci. Technol. B. **13**, 1938–1944 (1995)
30. F. Iwata, T. Matsumoto, A. Sasaki, Nanotechnology **11**, 10–15 (2000)
31. X.P. Wang, M.M. Loy, X. Xiao, Nanotechnology **13**, 478–483 (2002)
32. R.H. Schmidt, G. Haugstad, W.L. Gladfelter, Langmuir **19**, 898–909 (2003)
33. R.H. Schmidt, G. Haugstad, W.L. Gladfelter, Langmuir **19**, 10390–98 (2003)
34. B. Götsmann, U. Dürig, Langmuir **20**, 1495–1500 (2004)
35. R. Szoszkiewicz, T. Okada, S.J. Jones, T.-D. Li, W.P. King, S.R. Mader, E. Riedo, Nano Lett. **7**, 1064–69 (2007)
36. E. Gnecco, E. Riedo, W.P. King, S.R. Mader, R. Szoszkiewicz, Phys. Rev. B **79**, 235421 (2009)
37. M. D'Acunto, S. Napolitano, P. Pingue, P. Giusti, P. Rolla, Mater. Lett. **61**, 3305–3309 (2007)
38. S. Napolitano, M. D'Acunto, P. Baschieri, E. Gnecco, P. Pingue, Nanotechnology **30**, 475301 (2012)
39. K.L. Johnson, *Contact Mechanics* (Cambridge University Press, Cambridge, 1985)
40. R.W. Carpick, D.F. Ogletree, M. Salmeron, J. Colloid Interface Sci. **211**, 395–400 (1999)
41. M. Kopycinska-Müller, R.H. Geiss, D.C. Hurley, Ultramicroscopy **106**, 466–474 (2006)
42. B.V. Derjaguin, V.M. Muller, Y.P. Toporov, J. Colloid Interface Sci. **53**, 314–325 (1975)
43. K.L. Johnson, K. Kendall, A.D. Roberts, Proc. R. Soc. London A **324**, 301–313 (1971)
44. V. Bubosek, D.C. Prevorsek, Int. J. Polym. Mater. **47**, 569–572 (2000)
45. J. Cayer-Barrioz, D. Mazuyer, A. Tonck, P. Kapsa, A. Chateauminois, Tribol. Int. **39**, 62–69 (2006)
46. K.L. Johnson, in *Microstructure and Microtribology of Polymer Surfaces*, ed. by V.V. Tsukruk, K. Wahl, ACS Symposium Series, USA (1998)
47. B. Götsmann, U. Dürig, S. Sills, J. Frommer, C.J. Hawker, Nano Lett. **6**, 296–300 (2006)
48. M. Surtchev, N.R. de Souza, B. Jérôme, Nanotechnology **16**, 1213–20 (2005)

49. A. Schallamach, Wear **17**, 301 (1971)
50. G.A.D. Briggs, B.J. Briscoe, Nature **262**, 381–382 (1976)
51. B.J. Briscoe, T.A. Stolarski, Nature **281**, 206–208 (1979)
52. M. Barquins, A.D. Roberts, J. Phys. D: Appl. Phys. **19**, 547–563 (1986)
53. Y.J. Mergler, R.P. Schaake, J. Appl. Polym. Sci. **92**, 2689–2692 (2004)
54. A. Ghorbal, S. Bistac, M. Schmitt, J. Polym. Sci.: Part B: Polym. Phys. **44**, 2449–2454 (2006)
55. P. Samyn, G. Schoukens, P. De Baets, Wear **270**, 57–72 (2010)
56. B.J. Briscoe, S.K. Sinha, Mat.-wiss. u. Werkstofftech, **34**, 989–1001 (2003)
57. C.J. Rand, A.J. Crosby, Appl. Phys. Lett. **89**, 261907 (2006)
58. S. Bistac, A. Ghorbal, M. Schmitt, Prog. Org. Coat. **55**, 345–354 (2006)
59. H.-Y. Nie, M.J. Walzak, N.S. McIntyre, Appl. Surf. Sci. **253**, 2320–2326 (2006)
60. P.G. Whitten, H.R. Brown, Phys. Rev. E **76**, 026101 (2007)
61. A. Dasari, Z.-Z. Yu, Y.-W. Mai, Acta Mater. **55**, 635–646 (2007)
62. W. Brostow, W. Chonkaew, R. Mirshams, A. Srivastava, Polym. Eng. Sci. **48**, 2060–2065 (2008)
63. J. Jing, P.N. Henriksen, H. Wang, J. Mater. Sci. **30**, 5700–5704 (1995)
64. T. Aoike, T. Yamamoto, H. Uehara, T. Yamanobe, T. Komoto, Langmuir **17**, 5688–5692 (2001)
65. A.G. Khurshudov, K. Kam, Wear, **205**, 1–10 (1997)
66. T. Aoike, H. Uehara, T. Yamanobe, T. Komoto, Langmuir **17**, 2153–2159 (2001)
67. Y. Sun, Y. Yan, Z. Hu, X. Zhao, T. Sun, S. Dong, Scanning **35**, 308–315 (2013)
68. G.F. Meyers, B.M. DeKoven, J.T. Seitzj, Langmuir **8**, 2330–2335 (1992)
69. J.P. Pickering, G.J. Vancso, Appl. Surf. Sci. **148**, 147–154 (1999)
70. H. Uehara, T. Asakawa, M. Kakiage, T. Yamanobe, T. Komoto, Langmuir **22**, 4985–4991 (2006)
71. J.S.G. Ling, G.J. Leggett, A.J. Murray, Polymer **39**, 5913–5921 (1998)
72. B.D. Beake, P.H. Shipway, G.J. Leggett, Polymer **42**, 7025–7031 (2001)
73. B.D. Beake, G.J. Leggett, Polymer **43**, 319–327 (2002)
74. B.D. Beake, P.H. Shipway, G.J. Leggett, Wear **256**, 118–125 (2004)
75. M.D. Garrison, R. Luginbühl, R.M. Overney, B.D. Ratner, Thin Solid Films **352**, 13–2 (1999)
76. B.D. Beake, G.J. Leggett, M.R. Alexander, Polymer **42**, 2647–2653 (2001)
77. F. Dinelli, G.J. Leggett, M.R. Alexander, J. Appl. Phys. **91**, 3841–3846 (2002)
78. R. Berger, Y. Cheng, R. Forch, B. Gotsmann, J.S. Gutmann, T. Pakula, U. Rietzler, Langmuir **23**, 3150–3156 (2007)
79. Y. Karade, S.A. Pihan, W.H. Brünger, A. Dietzel, R. Berger, K. Graf, Langmuir **25**, 3108–3114 (2009)
80. R.H. Schmidt, G. Haugstad, W.L. Gladfelter, Langmuir **15**, 317–321 (1999)
81. F. Dinelli, C. Buenviaje, R.M. Overney, J. Chem. Phys. **113**, 2043–2058 (2000)
82. J. Fu, B. Li, Y. Han, J. Chem. Phys. **123**, 064713 (2005)
83. H. Fischer, Macromolecules **35**, 3592–3595 (2002)
84. L. Nick, A. Kindermann, J. Fuhrmann, Colloid Polym. Sci. **272**, 367–371 (1994)
85. J.H. Maas, M.A. Cohen, Stuart, G. J. Fleer Thin Solid Films **358**, 234–240 (2000)
86. C. Buenviaje, F. Dinelli, R. Overney, Macromol. Symposia **166**, 201–212 (2001)
87. T. Aoike, T. Ikeda, H. Uehara, T. Yamanobe, T. Komoto, Langmuir **18**, 2949–2951 (2002)
88. S.A. Pihan, S.G.J. Emmerling, H.-J. Butt, J.S. Gutmann, R. Berger, Wear **271**, 2852–2856 (2011)
89. R.H. Rice, P. Mokarian-Tabari, W.P. King, R. Szoszkiewicz, Langmuir **28**, 13503–13511 (2012)
90. X. Jin, W.N. Unertl, Appl. Phys. Lett. **61**(6), 657–659 (1992)
91. P. Pingue, M. Lazzarino, F. Beltram, C. Cecconi, P. Baschieri, C. Frediani, C. Ascoli, J. Vac. Sci. Technol. B **15**(4), 1398–1401 (1997)
92. S.F.Y. Li, H.T. Ng, P.C. Zhang, P.K.H. Ho, L. Zhou, G.W. Bao, S.L.H. Chan, Nanotechnology **8**, 76–81 (1997)
93. L.A. Porter, A.E. Ribbe, J.M. Buriak. Nano Lett. **3** (8), 1043–1047 (2003)

94. J. Lekki, S. Kumar, S.S. Parihar, S. Grange, C. Baur, R. Foschia, A. Kulik, Rev. Sci. Instr. **75**(11), 4646–4650 (2004)
95. J.A. Blach, G.S. Watson, C.L. Brown, D.K. Pham, J. Wright, D.V. Nicolau, S. Myhra, Thin Solid Films **459**, 95–99 (2004)
96. T.A. Jung, A. Moser, H.J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, U.D. Schwarz, Ultramicroscopy **42**, Part B, 1446 (1992)
97. M. Wendel, S. Kühn, H. Lorenz, J.P. Kotthaus, M. Holland, Appl. Phys. Lett. **65**, 1775 (1994)
98. B. Klehn, U. Kunze, J. Appl. Phys. **85**(7), 3897–3903 (1999)
99. K. Wiesauer, G. Springholz, J. Appl. Phys. **88**(12), 7289–7297 (2000)
100. B. Cappella, H. Sturm, S.M. Weidner, Polymer **43**, 4461–4466 (2002)
101. C. Balocco, A.G. Jones, J.M. Kingsley, J.R. Chan, X.Q. Huang, A.M. Song, Jpn. J. Appl. Phys. **45**(3B), 2095–2098 (2006)
102. F.J. Rubio-Sierra, A. Yurtsever, M. Hennemeyer, W.M. Heckl, R.W. Stark, Phys. Stat. Sol. A **203**(6) (2006)
103. Y. Wang, X. Hong, J. Zeng, B. Liu, B. Guo, H. Yan, Small **5**(4), 477–483 (2009)
104. X. Lu, C. Balocco, F. Yang, A.M. Song, IEEE Trans. Nanotech. **10**(1), 53–57 (2011)
105. W. Liu, Y. Yan, Z. Hub, X. Zhao, J. Yan, S. Dong, Appl. Surf. Sci. **258**, 2620–2626 (2012)
106. A. Buford, T. Goswami, Mater. Des. **25**, 385–393 (2004)
107. B.D. Ulery, L.S. Nair, C.T. Laurecin, J. Polym. Sci. B Polym. Phys. **49**, 832–864 (2011)
108. L.A. Pruitt, Biomaterials **26**, 905–15 (2005)
109. H.J. Butt, B. Cappella, M. Kappl, Surf. Sci. Rep. **59**, 1–152 (2005)
110. L.S. Nair, C.T. Laurecin, Prog. Polym. Sci. **32**, 762–798 (2007)
111. M. D'Acunto, in *Advances in Contact Mechanics: Implications for Materials Science, Engineering and Biology*, ed. by R. Buzio and U. Valbusa, Transworld Research Network (2006)
112. J. Fisher, E. Ingham, *Wear Debris in Encyclopedia of Biomaterials and biomedical Engineering* (M. Dekker edition, New York, 2004)
113. A. Wang, D.C. Sun, S.S. Yau, B. Edwards, M. Sokol, A. Essner, V.K. Polineni, C. Stark, H. Dunbleton, Wear **203**, 230–41 (1997)
114. R.S. Pascaud, W.T. Evans, P.J. McCullagh, D.P. Fitzpatrick, Biomaterials **18**, 727–35 (1997)
115. S.M. Kurtz, L.A. Pruitt, C.W. Jewett, J.R. Foulds, A.A. Edidin, Biomaterials **20**, 1449–62 (1999)
116. A.A. Edidin, L. Pruitt, C.W. Jewett, D.J. Crane, D. Roberts, S.M. Kurtz, J. Arthroplasty **14**, 616–627 (1999)
117. L. Costa, M.P. Luda, L. Trossarelli, E.M. Brach del Prever, M. Crova, P. Gallinaro, Biomaterials **19**, 659 (1998)
118. A. Wang, Wear **248**, 38–47 (2001)
119. L. Kang, A.L. Galvin, T.D. Brown, Z. Jin, J. Fisher, J. Biomechanics **41**, 340–346 (2008)
120. L. Mattei, F. Di Puccio, E. Ciulli, Tribol. Int. **63**, 66–77 (2013)

Part VII
Dissipation Mechanisms at Finite Separations

Chapter 25
Casimir Force and Frictional Drag Between Graphene Sheets

Aleksandr I. Volokitin and Bo Persson

Abstract Some years ago it was predicted that quantum fluctuations of the electromagnetic field produce the Casimir forces between macroscopic bodies. It has recently been shown that two non-contacting bodies moving relative to each other experience a friction due to the same quantum fluctuations of the electromagnetic field. However, until recent time there was no experimental evidence for or against this effect, because the predicted friction forces are very small, and precise measurements of quantum forces are incredibly difficult with present technology. The existence of quantum friction is still debated even among theoreticians. However, the situation drastically changed with the discovery of a new material—graphene. We recently proposed that quantum friction can be observed in experiments studying electrical transport phenomena in nonsuspended graphene on dielectric substrate and by measuring frictional drag between graphene sheets.

We investigate the dependence of the thermal Casimir force and the Casimir friction force between two graphene sheets on the drift velocity of the electrons in one graphene sheet. We show that the drift motion produces a measurable change of the thermal Casimir force due to the Doppler effect. The thermal Casimir force as well as the Casimir friction are strongly enhanced in the case of resonant photon tunneling when the energy of the emitted photon coincides with the energy of electron-hole pair excitations. In the case of resonant photon tunneling, even for temperatures above room temperature the Casimir friction is dominated by quantum friction due to quantum fluctuations.

25.1 Introduction

Friction is usually a very complicated process. It appears in its most elementary form when two flat surfaces, separated by a vacuum gap, are sliding relative to each other at zero Kelvin, where the friction is generated by the relative movement of

A.I. Volokitin (✉)
Samara State Technical University, Samara 443100, Russia
e-mail: alevolokitin@yandex.ru

B. Persson
Peter Grünberg Institut, Forschungszentrum Jülich, 52425 Jülich, Germany

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_25

quantum fluctuations. For several decades, physicists have been intrigued by the idea of quantum friction. It has recently been shown that two non-contacting bodies moving relative to each other experience a friction due to quantum fluctuations inside the bodies [1–4]. However, until recently there was no experimental evidence for or against this effect, because the predicted friction forces are very small, and precise measurements of quantum friction are incredibly difficult with present technology. Recently we proposed [5, 6] that using graphene it should be possible to detect quantum friction. Note the distinction between quantum friction between bodies in parallel relative motion and the dynamic Casimir effect for accelerated relativistic motion in the normal direction [7].

Graphene, isolated monolayer of carbon, which was obtained very recently [8], consists of carbon atoms densely packed into a two-dimensional honeycomb crystal lattice (Fig. 25.1). The unique electronic and mechanical properties of graphene are actively studied both theoretically and experimentally partly because of their importance for fundamental physics, and also because of its possible technological applications [8–11]. In particular, the valence band and conduction band in graphene touch each other at one point named the Dirac point. Near this point the energy spectrum for electrons and holes has a linear dispersion. Due to this linear (or "conical") dispersion relation electrons and holes near this point behave like relativistic particles described by the Dirac equation for massless fermions.

Graphene can also be useful for the detection of quantum friction. Consider graphene located on the surface of, for example, the polar dielectric SiO_2, or nearby of a second graphene sheet. In this case the charge carriers in graphene experience additional friction due to interaction with the optical phonons in the dielectric, or the electrons in other graphene sheet. Due to the high mobility, in a strong electric field the electrons in graphene can move with very high drift velocities ($\sim 10^6$ m/s). At such velocities the main contribution to the friction will arise from quantum fluctuations. Thus, quantum friction can be detected by measuring the high electric field

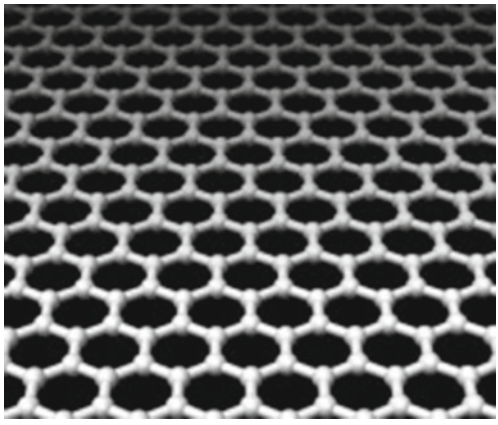


Fig. 25.1 Honeycomb lattice of graphene

transport properties of graphene on a polar dielectric substrate, or by measuring the voltage induced by friction in a second nearby graphene sheet.

25.2 Fluctuations Produce Forces

In the late 1940s Hendrik Casimir predicted [12] that two macroscopic non-magnetic bodies with no net electric charge (or charge moments) can experience an attractive force much stronger than gravity. The existence of this force is one of the few direct macroscopic manifestations of quantum mechanics; others are superfluidity, superconductivity, and the black body radiation spectrum.

Hendrik Casimir based his prediction on a simplified model involving two parallel perfectly conducting plates separated by vacuum. A unified theory of both the van der Waals and Casimir forces between plane parallel material plates, in thermal equilibrium and separated by a vacuum gap, was developed by Lifshitz (1955) [13]. Lifshitz's theory describes dispersion forces between dissipative media as a physical phenomenon caused by the fluctuating electromagnetic field that is always present in both the interior and the exterior of any medium. Outside the medium this field exists partly in the form of the radiative propagating waves and partly in the form of nonradiative evanescent waves whose amplitudes decay exponentially with the distance away from the medium, see Fig. 25.2. To calculate the fluctuating electromagnetic field Lifshitz used Rytov's theory [14]. Rytov's theory is based on the introduction into the Maxwell equation of a "random" field (just as, for example, one introduces a "random" force in the theory of Brownian motion). Both quantum and thermal fluctuations give contributions to the Casimir force. The general theory of Casimir-van der Waals forces was developed in [15] using quantum field theory. This theory confirmed the results

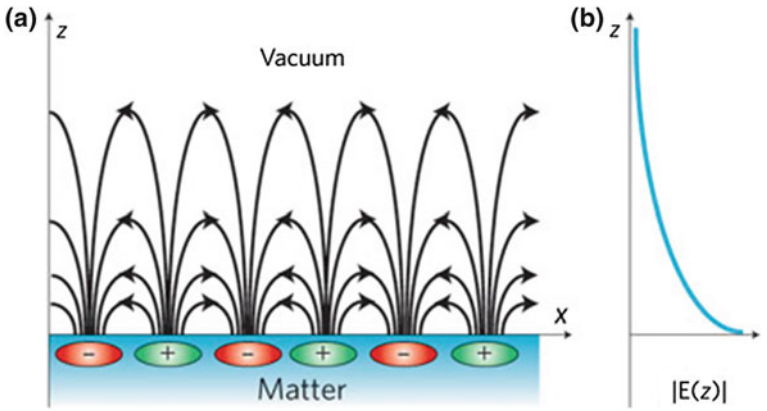


Fig. 25.2 The amplitude of evanescent electromagnetic waves decays exponentially with the distance away from the surface. The intensity of the evanescent waves is especially large close to the surface of a polar dielectric; this is related with the surface phonon polaritons

of Lifshitz's theory. Quantum fluctuations dominate at small separation ($d < \lambda_T = c\hbar/k_BT$) and thermal fluctuations dominate at large separation ($d > \lambda_T$). Casimir forces due to quantum fluctuations have been studied experimentally for a long time [16, 17]. However the Casimir forces due to thermal fluctuations were measured only recently and these measurements confirmed the prediction of the Lifshitz theory [18]. At present the interest to Casimir forces is increasing because they dominate the interaction between nanostructures, and are offen responsible for the adhesion between moving parts in small devices such as micro- and nanoelectromechanical systems [19, 20]. Due to this practical interest and the fast progress in force detection techniques, experimental [18, 21–24] and theoretical [25, 26] investigations of Casimir forces have experienced an extraordinary "renaissance" in the past few years.

The Lifshitz theory was formulated for systems at thermal equilibrium. At present there is an interest in the study of systems out of the thermal equilibrium, in particular in the connection with the possibility of tuning the strength and sign of the interaction. Such systems also present a way to explore the role of thermal fluctuations, which usually are masked at thermal equilibrium by the $T = 0$ K component, which dominates the interaction up to very large distances, where the interaction force is very small. In [27] the Casimir-Polder force was measured at very large distances, and it was shown that the thermal effects on the Casimir-Polder interaction agree with the theoretical prediction. This measurement was done out of thermal equilibrium, where thermal effects are stronger.

Other non-equilibrium thermal effects were investigated by Polder and van Hove [28], who calculated heat flow between two parallel surfaces, separated by a vacuum gap using Rytov's theory. Recently the theoretical predictions were confirmed in experiments [29, 30]. Already for more than 30 years physicists have been interested in how the Casimir forces and the radiative heat transfer are modified for bodies moving relative to each other. A number of researchers have shown that the relative motion of bodies leads to a friction force [1–4, 31–33]. Theory predicts that the Casimir friction acts even at zero temperature, when it is determined by quantum fluctuations. However, in recent years the existence of quantum friction was hotly debated [34–39]. A general theory of the Casimir forces, Casimir friction and the radiation heat transfer between moving bodies was developed by us in [3]. This theory confirmed the correctness of the previous results obtained using quantum mechanical perturbation theory [1, 31], dynamical generalization of the Lifshitz-Rytov's theory [2, 40] and quantum field theory [41].

At present a great deal of attention is devoted to the study of the Casimir forces in graphene systems [42–50]. This is due the unusual electronic properties of graphene, which result in Casimir forces with unusual properties. For normal materials, the contribution to the Casimir force due to thermal fluctuations dominates for $d > \lambda_T = \hbar c/k_BT$, but for two graphene sheets the thermal contribution dominates for much shorter distances [50] $d > \xi_T = \hbar v_F/k_BT$, where $v_F \sim 10^6$ m/s is the Fermi velocity in graphene. At room temperature the parameters ξ_T and λ_T are 25 nm and 7.6 μm, respectively. This property makes it possible to measure the thermal Casimir force using an atomic force microscope, or other force measuring

techniques. Tailoring the thermal Casimir force using Fermi level tuning by gate voltage was discussed in [48].

Alternative method of tailoring the thermal Casimir force consists in driving an electric current in a graphene sheet. It was shown by Pendry [1] that the reflection amplitudes from moving metal surface are modified due to the Doppler effect. The same modification of reflection amplitudes can be obtained if instead of motion of metal plate, a drift motion of charge carriers is induced in it by applied voltage [51]. Due to the high mobility of the carriers in graphene, in a high electric field the electrons (or holes) can move with very high velocities (up to 10^6 m/s). The drift motion of the charge carries in graphene will result in a modification of dielectric properties (and the Casimir force) of graphene due to the Doppler effect [1] (see Fig. 25.3). If in one of two parallel graphene sheets an electric current is induced, then the electromagnetic waves, radiated by the graphene sheet without an electric current, will experience a frequency Doppler shift in the reference frame moving with the drift velocity v of electrons in the other graphene sheet: $\omega' = \omega - q_x v$, where q_x is the parallel to the surface component of the momentum transfer. The same is true for the waves emitted by the other graphene sheet. Due to the frequency dependence of the reflection amplitudes the electromagnetic waves will reflect differently in comparison to the case when there is no drift motion of electrons, and this will give rise to the change of the Casimir force. The effect of the drift motion of charge carriers in one of the graphene sheet, on the thermal Casimir force between graphene sheets, was investigated in [6].

Let us consider two graphene sheets separated by vacuum gap with thickness $d \ll \lambda_T = c\hbar/k_B T$. Assume that the free charge carriers in one graphene sheet move with drift velocity $v \ll c$ along the x-axis (c is the light velocity) relative to the other graphene sheet. Because a drift motion of the free charge carriers produces

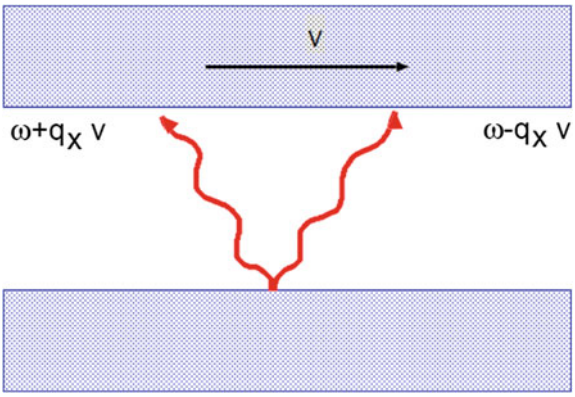


Fig. 25.3 The electromagnetic waves emitted in the opposite direction by the body at the bottom will experience opposite Doppler shift in the reference frame in which the body at the *top* is at rest. Due to the frequency dispersion of the reflection amplitude these electromagnetic waves will reflect differently from the surface of the body at the *top*, which gives rise to the Casimir friction and the modification of the Casimir forces between the bodies

a similar modification of the reflection amplitudes as in the case of moving graphene sheet, the theory of the Casimir forces between moving bodies [3] can be used to calculate the Casimir forces between sheets (both of each are at the rest) in presence of the drift motion of the free charge carriers in one graphene sheet. The force which acts on the surface of the sheet can be calculated from the Maxwell stress tensor σ_{ij}, evaluated at the surface of the sheet at $z = 0$:

$$\sigma_{ij} = \frac{1}{4\pi}\int_0^\infty d\omega \int \frac{d^2q}{(2\pi)^2}\Big[< E_i E_j^* > + < E_i^* E_j > + < B_i B_j^* > + < B_i^* B_j > - \delta_{ij}(< \mathbf{E}\cdot\mathbf{E}^* > + < \mathbf{B}\cdot\mathbf{B}^* >)\Big]_{z=0} \tag{25.1}$$

where $< \cdots >$ denotes statistical average over the random the electric $\mathbf{E}$ and magnetic induction $\mathbf{B}$ field. According to [3] the Casimir force $F_z = \sigma_{zz}$ between moving media is determined by

$$F_z = F_{zT} + F_{z0}, \tag{25.2}$$

where the temperature dependent term F_{zT} and the zero-temperature contribution F_{z0} are given by

$$F_{zT} = \frac{\hbar}{\pi^3}\int_0^\infty dq_y \int_0^\infty dq_x q e^{-2qd} \left\{ \int_0^\infty d\omega \left(\frac{\mathrm{Im}R_1(\omega)\mathrm{Re}R_2(\omega^+)n_1(\omega) + \mathrm{Re}R_1(\omega)\mathrm{Im}R_2(\omega^+)n_2(\omega^+)}{|\,1 - e^{-2qd}R_1(\omega)R_2(\omega^+)\,|^2} + (1 \leftrightarrow 2)\right) + \int_0^{q_x v} d\omega \left(\frac{\mathrm{Re}R_1(\omega^-)\mathrm{Im}R_2(\omega)n_2(\omega)}{|\,1 - e^{-2qd}R_1(\omega^-)R_2(\omega)\,|^2} + (1 \leftrightarrow 2)\right)\right\}, \tag{25.3}$$

$$F_{z0} = \frac{\hbar}{2\pi^3}\int_0^\infty dq_y \int_0^\infty dq_x \left\{ \mathrm{Re}\int_0^\infty d\omega s e^{-2sd}\left(\frac{R_1(i\omega)R_2(i\omega + q_x v)}{1 - e^{-2sd}R_1(i\omega)R_2(i\omega + q_x v)} + (1 \leftrightarrow 2)\right) + \int_0^{q_x v} d\omega q e^{-2qd}\left(\frac{\mathrm{Im}R_1(\omega)\mathrm{Re}R_2(\omega^-)}{|\,1 - e^{-2qd}R_1(i\omega)R_2(\omega^-)\,|^2} + (1 \leftrightarrow 2)\right)\right\}, \tag{25.4}$$

where $n_i(\omega) = [\exp(\hbar\omega/k_B T_i) - 1]^{-1}\,(i = 1, 2), q = \sqrt{q_x^2 + q_y^2}, s = \sqrt{(\omega/c)^2 + q^2}$, T_i is the temperature of i-th graphene sheet, R_i is the reflection amplitude for surface i for p -polarized electromagnetic waves, and $\omega^\pm = \omega \pm q_x v$. The symbol $(1 \leftrightarrow 2)$ denotes the terms that are obtained from the preceding terms by permutation of 1

and 2. In the first term in (25.4) the integration along the real axis was transformed into integration along the imaginary axis.

The reflection amplitude for a 2D-system is determined by [52]

$$R_i = \frac{\epsilon_i - 1}{\epsilon_i + 1}, \quad \epsilon_i = \frac{4\pi p \sigma_i}{\omega \varepsilon} + 1, \tag{25.5}$$

where $p = \sqrt{(\omega/c)^2 - q^2}$, σ_i is the longitudinal conductivity of the sheet which can written in the form $\sigma_i = -i\omega e^2 \Pi_i(\omega, q)/q^2$ where Π_i is the 2D polarizability. The dielectric function of the sheet is determined by $\varepsilon_i(\omega, q) = 1 + v_q \Pi_i(\omega, q)$, $v_q = 2\pi e^2/q$ is the 2D Coulomb interaction. In term of ε_i the reflection amplitude can be written as

$$R_i = \frac{p(\varepsilon_i - 1)}{p(\varepsilon_i - 1) + iq} \tag{25.6}$$

In the integration on the real axis $p \approx iq$ for $d < \lambda_T$. Thus, in this case

$$R_i \approx \frac{\varepsilon_i - 1}{\varepsilon_i}, \tag{25.7}$$

On the imaginary axis $p = is$. In the finite lifetime generalization according to the Mermin approximation [53] the dielectric function is determined by

$$\varepsilon(\omega, q) \approx 1 + \frac{(\omega + i\gamma)(\varepsilon_0(\omega + i\gamma, q) - 1)}{\omega + i\gamma(\varepsilon_0(\omega + i\gamma, q) - 1)/(\varepsilon_0(0, q) - 1)}, \tag{25.8}$$

where $\varepsilon_0(\omega, q)$ is the RPA dielectric function and γ is the damping parameter. In the study below we used the dielectric function of graphene, which was calculated recently within the random-phase approximation (RPA) [54, 55]. The small (and constant) value of the graphene Wigner-Seitz radius r_s indicates that it is a weakly interacting system for all carries densities, making the RPA an excellent approximation for graphene (RPA is asymptotically exact in the $r_s \ll 1$ limit). The dielectric function is an analytical function in the upper half-space of the complex ω-plane:

$$\varepsilon_0(\omega, q) = 1 + \frac{4k_F e^2}{\hbar v_F q} - \frac{e^2 q}{2\hbar\sqrt{\omega^2 - v_F^2 q^2}} \left\{ G\left(\frac{\omega + 2v_F k_F}{v_F q}\right) - G\left(\frac{\omega - 2v_F k_F}{v_F q}\right) - i\pi \right\}, \tag{25.9}$$

where

$$G(x) = x\sqrt{x^2 - 1} - \ln(x + \sqrt{x^2 - 1}), \tag{25.10}$$

where the Fermi wave vector $k_F = (\pi n)^{1/2}$, n is the concentration of charge carriers, the Fermi energy $\epsilon_F = \hbar v_F k_F$, $v_F \approx 10^6$ m/s is the Fermi velocity. The damping parameter γ is due to scattering against impurities and acoustic phonons in graphene sheet, and can be expressed through the low field mobility μ: $\gamma = ev_F/(\hbar k_F \mu)$.

Scattering of the graphene carries by the acoustic phonons of graphene places an intrinsic limits on the low-field room temperature ($T_0 = 300$ K) mobility, given by $\mu_0 = 20$ m^2/Vs at the graphene carriers density 10^{16} m^{-2} (see [56]), which gives $\gamma = 8 \times 10^{11}$ s^{-1}. At other temperatures the mobility can be obtained using the relation $\mu = \mu_0 T_0/T$.

In addition to the intrinsic friction due to scattering against impurities and phonons, on the electrons moving in the graphene sheet acts the extrinsic friction due to the interaction with electrons in the nearby graphene sheet. According to the theory of the Casimir friction [3], the friction force $F_x = \sigma_{xz} = F_{xT} + F_{x0}$, where at $d \ll \lambda_T$ and $v \ll c$ the contributions from thermal (F_{xT}) and quantum (F_{x0}) fluctuations are given by [1, 2, 4, 5]

$$F_{xT} = \frac{\hbar}{\pi^3}\int_0^\infty dq_y \int_0^\infty dq_x q_x e^{-2qd} \left\{ \int_0^\infty d\omega \left(\frac{\mathrm{Im}R_1(\omega)\mathrm{Im}R_2(\omega^+)}{|\,1 - e^{-2qd}R_1(\omega)R_2(\omega^+)\,|^2} \times [n_1(\omega) - n_2(\omega^+)] + (1 \leftrightarrow 2) \right) - \int_0^{q_x v} d\omega \left(\frac{\mathrm{Im}R_1(\omega)\mathrm{Im}R_2(\omega^-)}{|\,1 - e^{-2qd}R_1(\omega)R_2(\omega^-)\,|^2} n_1(\omega) + (1 \leftrightarrow 2) \right) \right\}, \tag{25.11}$$

$$F_{x0} = -\frac{\hbar}{2\pi^3}\int_0^\infty dq_y \int_0^\infty dq_x q_x e^{-2qd} \int_0^{q_x v} d\omega \left(\frac{\mathrm{Im}R_1(\omega)\mathrm{Im}R_2(\omega^-)}{|\,1 - e^{-2qd}R_1(\omega)R_2(\omega^-)\,|^2} + (1 \leftrightarrow 2) \right). \tag{25.12}$$

Equations (25.11) and (25.12) were initially obtained for 3D-systems in [1] at $T = 0$ and in [2] for finite temperatures. However, in [52] it was shown that the same formulas are valid for 2D-systems. For $v < dk_B T/\hbar$ (at $d = 1$ nm and $T = 300$ K for $v < 4 \times 10^4$ m/s) the main contribution to the friction (25.11) depends linearly on the sliding velocity v so that the friction force $F_{\cdot T} = \Gamma v$ where at $T_1 = T_2 = T$ the friction coefficient Γ is given by

$$\Gamma = \frac{\hbar^2}{8\pi^2 k_B T}\int_0^\infty \frac{d\omega}{\sinh^2\left(\frac{\hbar\omega}{2k_B T}\right)} \int_0^\infty dq\, q^3 e^{-2qd} \frac{\mathrm{Im}R_1(\omega)\mathrm{Im}R_2(\omega)}{\left|1 - e^{-2qd}R_1(\omega)R_2(\omega)\right|^2}. \tag{25.13}$$

Due to the presence of an exponential factor in the expression (25.3) for the thermal contribution to the Casimir force, the integration over frequency is effectively limited to $\omega < \omega_T = k_B T/\hbar$. Thus for $q_x v \sim v/d > \omega_T$ (at room temperature and for $d = 1$

nm this condition corresponds to the velocities $v > 10^5$ m/s) the integrand will be modified in the whole range of integration, which will give rise to the significant change of the thermal Casimir force. This change will be especially large in the case of resonant photon tunneling when the integrand has sharp resonances. The integrand in the expression for the zero-temperature contribution to the Casimir force does not contain any sharp cut-off in the frequency integration. Thus the range of integration will be more wide and the change of the zero-temperature contribution will be significant only for much higher velocities than for the thermal contribution.

Figure 25.4a shows the dependence of the Casimir force between two graphene sheets on the separation d between the sheets. The thermal and quantum contributions are shown separately. The thermal contribution was calculated for $T = 600$ K and for the drift velocities $v = 0$ and $v = 2 \times 10^6$ m/s. The thermal contribution becomes larger then the quantum contribution for $d > 50$ nm. For $d < 5$ nm the thermal contribution calculated for $v = 2 \times 10^6$ m/s is significantly larger then the thermal contribution calculated at $v = 0$. For example, at $d \approx 3$ nm the drift motion of the electrons gives rise to the increase of the thermal Casimir force by one order of magnitude, and in this case the thermal contribution is only one order of magnitude smaller then the quantum contribution, and can be measured experimentally. Figure 25.4b shows the dependence of the thermal Casimir force F_{zT} on the drift velocity of the electrons in the graphene sheet at $d = 1$ nm. Note the significant change of the thermal Casimir force for $v/d > \omega_T$ (at room temperature and for $d = 1$ nm this condition corresponds to the velocities $v > 10^5$ m/s). This change will be especially large in the case of resonant photon tunneling. Let us assume that in the rest reference frame in which there is no drift motion of electrons, the electron-hole pair excitation arises with energy $\omega_{eh}(q)$ and momentum $\mathbf{q}$, then in the laboratory reference frame in which the electron system is moving with drift velocity v, due to the Doppler effect the energy of this excitation will be equal to $\omega_{eh}(q) - q_x v$. For $v > \omega_{eh}(q)/q_x$ the excitation energy will be negative thus at the velocities larger than critical ($v_{cr} = \omega_{eh}(q)/q_x$) as a result

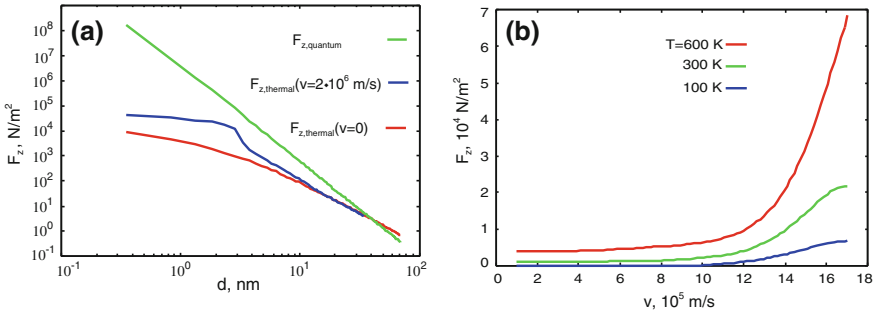


Fig. 25.4 The Casimir forces between two graphene sheets with carrier concentration $n = 10^{16}\text{m}^{-2}$. **a** The dependence of the Casimir force on the separation d between the sheets. The thermal and quantum contributions to the total Casimir force are shown separately. The thermal contribution is shown for $T = 600$ K and for the drift velocities $v = 0$ and $v = 2 \times 10^6$ m/s. **b** The dependence of the thermal Casimir force on the drift velocity of electrons v in one of the graphene sheet at $d = 1$ nm

of such excitation the photon can be created with energy $\omega_{ph}(q) = q_x v - \omega_{eh}(q) > 0$, i.e. the radiation arises. This radiation is reminiscent of the Cherenkov radiation which arises at motion of the electron in the medium with the velocity exceeding the light velocity in the medium. The difference consists in that the Cherenkov radiation is connected with the radiation of the propagating electromagnetic waves, but the radiation which arises at drift motion of the electron in the graphene sheet—with excitation of electron-hole pairs. Resonance arises when the photon emitted by the moving electron system in one graphene sheet with energy $\omega_{ph}(q) = q_x v - \omega_{eh}(q) > 0$ will create excitation with energy $\omega_{eh}(q)$ in other graphene sheet. In the case of graphene the energy of the electron-hole pair excitation $\omega_{eh}(q) \approx v_F q$, where v_F is the Fermi velocity. Thus resonance arises when $q_x v \approx 2v_F q$, which requires that $v > 2v_F \approx 2 \times 10^6$ m/s, in accordance with the numerical calculations.

25.3 Reflection Produces Friction

The origin of the Casimir friction is closely connected with the Casimir forces. The Casimir interaction arises when an atom or a molecule spontaneously develops an electric dipole moment due to quantum fluctuations. The short-lived atomic polarity can induce a dipole moment in a neighboring atom or molecule some distance away. The same is true for extended media, where thermal and quantum fluctuation of the current density in one body induces a current density in other body; the interaction between these current densities is the origin of the Casimir interaction. When two bodies are in relative motion, the induced current will lag slightly behind the fluctuating current inducing it, and this is the origin of the Casimir friction. The Casimir interaction is mostly determined by the exchange of virtual photons between the bodies (connected with quantum fluctuations), and does not vanish even at zero temperature. Thermal fluctuations affect the Casimir forces only at the large distance between the bodies, where the contribution from the quantum fluctuations becomes very small. On the contrary, Casimir friction is related with the exchange of real photons between bodies, which are created in energy conservation processes. Casimir friction at low velocities ($v < dk_BT/\hbar$) is determined by the exchange of the real photons resulting from thermal fluctuations. However, at large velocities and low temperatures ($v > dk_BT/\hbar$), the Casimir friction is determined by the exchange of real photons resulting from the relative motion of quantum fluctuations [1–4].

The origin of the Casimir friction can also be explained by the Doppler effect [1]. Let us consider two flat parallel surfaces, separated by a sufficiently wide vacuum gap, which prevents electrons from tunneling across it. If the surfaces are in relative motion (velocity v) a frictional stress will act between them. This frictional stress is related to an asymmetry of the reflection amplitude along the direction of motion [1, 4], see Fig. 25.3. The electromagnetic waves emitted in the opposite direction by one body will experience opposite Doppler shift in the rest reference frame of the second body. Due to the frequency dispersion of the reflection amplitude these electromagnetic waves will reflect differently from the surface of the second body, which gives rise

to the momentum transfer between the bodies. This momentum transfer is the origin of the Casimir friction.

The Casimir friction originates from two quantum mechanical processes processes [4]. (a) An excitation with the frequency $\omega_\alpha(\mathbf{q})$ in the rest reference frame of the moving (with velocity v) body, in the laboratory reference frame will have frequency $\omega_\alpha(\mathbf{q}) - q_x v$ due to Doppler frequency shift. If $\omega_\alpha(\mathbf{q}) - q_x v < 0$ then in the laboratory reference frame a photon can be created with frequency $\omega_{ph} = q_x v - \omega_\alpha(\mathbf{q})$. (b) An excitation annihilated in the rest reference frame of the moving body will create the photon with frequency $\omega_{ph} = \omega_\alpha(\mathbf{q}) - q_x v$ in the laboratory reference frame. The photons created in processes (a) and (b) can create excitations in the body which is at rest in the laboratory reference frame. This will result in momentum transfer and friction. Thus in process (a) the excitations are created in both bodies, in contrast to the process (b) for which the excitation is annihilated in one body and created in other body. The first process (a) is possible even at zero temperature, when it is associated with quantum friction. The second process (b) is possible only at finite temperatures, when it is associated with thermal radiation. Quantum and thermal friction are associated with quantum and thermal fluctuations, respectively. The process (a) will dominate for $v > dk_BT/\hbar$ and process (b) will dominate for $v < dk_BT/\hbar$. At small velocities thermal and quantum friction forces depend linearly and cubically on sliding velocity, respectively.

It is important to note that only evanescent waves give a contribution to the quantum friction, i.e. it manifest itself only between closely spaced bodies. Relative motion of the bodies produces an excitation of the evanescent modes and the exchange of the corresponding photons will lead to quantum friction. For a body moving in the absolute vacuum, quantum friction (in contrast to the thermal friction) is equal to zero, in agreement with the principle of relativity; according to this principle for a body in the absolute vacuum at $T = 0$ K the friction should be zero.

25.4 Using Graphene to Detect Quantum Friction

Quantum friction determines the ultimate limit to which the friction can be reduced. In order to detect quantum friction it is necessary to reduce the contribution to friction from other mechanisms up to unprecedented levels. However, even in non-contact friction experiments [57, 58], when two bodies are not in direct contact, there are several contribution to the friction [4]. Moreover, quantum friction dominates over thermal friction at velocities $v > dk_BT/\hbar$ (at $d = 1$ nm and room temperature: $v > 10^5$m/s). However, at present even for a hard cantilever the velocity of the tip cannot exceed 1 m/s [58].

We recently proposed [5] that it should be possible to detect quantum friction in graphene adsorbed on an amorphous SiO_2 substrate (Fig. 25.5). The electrons, moving in graphene under the action of electric field, will experience intrinsic friction due to interaction with the acoustic and optical phonons in graphene, and extrinsic friction due to interaction with the optical phonons in the SiO_2-substrate. In a high electric

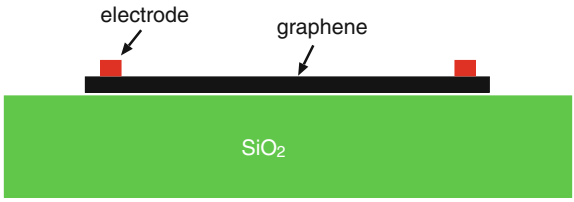


Fig. 25.5 Scheme of the graphene field effect transistor

fields the electrons move with high velocities, and in this case the main contribution to the friction arises from the interaction with the optical phonons in graphene and in SiO_2. However, the frequency of the optical phonons in graphene is approximately four times larger than in SiO_2. Therefore, the main contribution to the friction will result from the interaction with the optical phonons in SiO_2. Thus, this frictional interaction determines electrical conductivity of graphene at high electric field.

Figure 25.6a shows the dependence of the current density on the electric field at the carrier concentration $n = 10^{12}$ cm^{-2}, and for different temperatures. We have found that, in agreement with the experiment [59], the current density saturates at $E \sim 0.5 - 2.0$ V/μm. According to the experiment the saturation current density $J_{sat} = nev_{sat} \approx 1.6$ mA/μm, and using the charge density concentration $n = 10^{12}$ cm^{-2}: $v_{sat} \approx 10^6$ m/s. The saturation current density depends weakly on the temperature. In Fig. 25.6b the contributions to the friction force from quantum and thermal fluctuations are shown separately. In the saturation region the contribution to the friction force from quantum fluctuations dominates.

According to the theory of the Casimir friction [4] (see also above discussion), the quantum friction, which exists even at zero temperature, is determined by the creation of excitations in each of the interacting media. The frequency of the photon which is created as a result of excitation in moving body is determined by $\omega_{ph} = vq_x - \omega_1$,

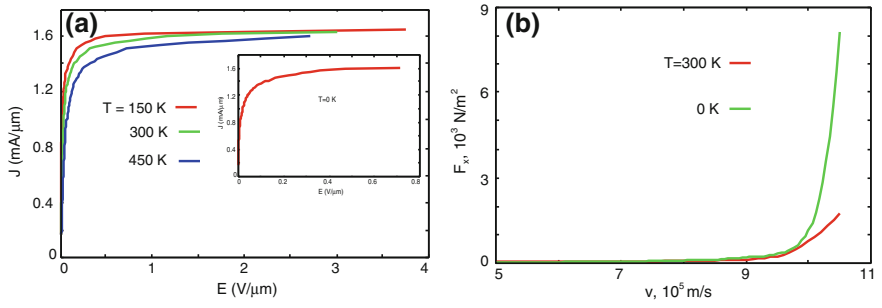


Fig. 25.6 The role of the interaction between phonon polaritons in SiO_2 and free carriers in graphene for graphene field-effect transistor transport. The separation between graphene and SiO_2 is $d = 3.5$ Å, the charge density $n = 10^{12}$ cm^{-12} **a** Current density-electric field dependence for different temperatures. Inset shows the same dependence at $T = 0$ K. **b** Dependence of the quantum and thermal contributions to the friction force between SiO_2 and free carriers in graphene per unit area on the drift velocity of electrons in graphene. The finite temperature curve shows only the thermal contribution

where ω_1 is the excitation frequency in the rest reference frame. This photon will create excitation in the other body with frequency $\omega_2 = \omega_{ph} = vq_x - \omega_1$. The relevant excitations in graphene are the electron-hole pairs with energy $\omega_{eh}(q) \approx v_F q$, while for SiO_2 the frequency of surface phonon polaritons $\omega_0 \approx 60$ meV (9×10^{13} s^{-1}). Resonant photon tunneling occurs at $v > v_F + \omega_0/q_x$. The maximal value of wave vector is determined by $1/d$ thus at $d = 0.35$ nm resonance occurs for $v > v_{sat} = v_F + \omega_0 d \sim 10^6$ m/s, in accordance with numerical calculations. Thus, measurements of the current density-electric field relation of graphene adsorbed on SiO_2 give the possibility to detect quantum friction.

An alternative method of studying of the Casimir friction consists in driving an electric current in one metallic layer and studying the effect of the frictional drag on the electrons in a second (parallel) metallic layer (Fig. 25.7). Such experiments were proposed by Pogrebinskii [60] and Price [61], and were performed for 2D quantum wells [62, 63]. In these experiments a current is driven through layer **1**. Due to the proximity of the layers, the interlayer interactions will induce a current in layer **2** due to a frictional stress acting on the electrons in the layer **2** from layer **1**. If the layer **2** is an open circuit, an electric field E_1 will develop in the layer whose influence cancels the frictional stress σ between the layers. In the experiment [62] the drift velocity $v \sim 10^2$ m/s. According to the theory of the Casimir friction [2, 64], at such velocities the thermal fluctuation give the dominant contribution to the friction, and the theoretical predictions are in agreement with experiment.

Frictional drag between graphene sheets was measured recently in [65, 66]. This study has fueled the recent theoretical investigations of frictional drag between graphene sheets [67–73] mediated by a fluctuating Coulomb field. In all these investigations the current density (or drift velocity v of the charge carries) is linearly related to the driving electric field. Thus only the thermal contribution to the frictional drag was included. In the linear approximation the electric field induced by the frictional drag depends linearly on the current density $J = nev$ (or drift velocity v of the charge carries), $E = \rho_D J = F_{xT}/ne = \Gamma J/(ne)^2$, where Γ is the friction coefficient, $\rho_D = \Gamma/(ne)^2$ is the drag resistivity. For $\omega < v_F q$ and $q < 2k_F$ the dielectric function of graphene has the following form [55]

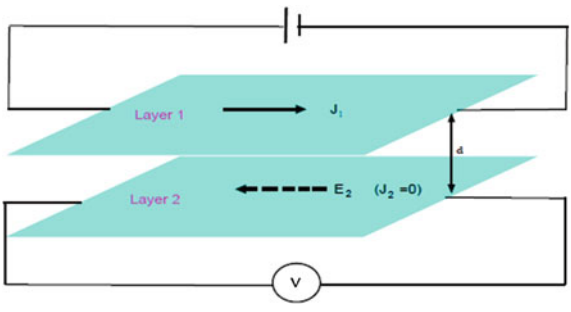


Fig. 25.7 Scheme of experiment for observation of the drag effect

$$\varepsilon_0(\omega, q) \approx 1 + \frac{4e^2 k_F}{\hbar v_F q}\left(1 + i\frac{\omega}{v_F q}\right), \tag{25.14}$$

and the reflection amplitude

$$R_0(\omega, q) = \frac{\varepsilon_0(\omega, q) - 1}{\varepsilon_0(\omega, q)} \approx 1 + i\frac{\hbar\omega}{4e^2 k_F}, \tag{25.15}$$

and (25.15), (25.13) give the known result [67]

$$\rho_D = \frac{\Gamma}{(ne)^2} = \frac{h}{e^2}\frac{\pi\zeta(3)}{32}\left(\frac{k_B T}{\epsilon_F}\right)^2 \frac{1}{(k_F d)^2}\frac{1}{(k_{TF} d)^2}, \tag{25.16}$$

where $k_{TF} = 4e^2 k_F/\hbar v_F$ is the Thomas-Fermi screening wave vector. The frictional drag force is much higher for high drift velocities ($\sim 10^6$ m/s), where it depends nonlinearly on the drift velocity, and is dominated by the quantum friction, existence of which was recently hot debated [34–39]. For $v < v_F$ (25.12), (25.15) give the following result for quantum friction

$$F_{x0} = \frac{\hbar v}{d^4}\frac{15\zeta(5)}{128\pi^2}\left(\frac{v}{v_F}\right)^2 \frac{1}{(k_{TF} d)^2}. \tag{25.17}$$

In linear approximation $E = 5 \times 10^{-4} v$ (SI-units) for $T = 300$ K and $d = 10$ nm. For a graphene sheet of length 1 μm, and with $v = 100$ m/s this electric field will induce the voltage $V = 10$ nV. From (25.16) and (25.17) the ratio of quantum and thermal friction $F_{x0}/F_{xT} = F_{x0}/(ne)^2 \rho_D v \approx (15/8\pi^2)(v/v_T)^2$, where $v_T = \omega_T d$. Thus, for $v > v_T$ the friction is dominated by quantum friction (at $d = 1$ nm and room temperature: $v_T \approx 4 \cdot 10^4$m/s).

Figure 25.8a, b show that much larger electric fields can be induced at $d = 1$ nm (a) and $d = 10$ nm (b) at large velocities. In these figures the contributions to friction from thermal and quantum fluctuations are shown separately. For $v < 10^5$ m/s the frictional drag effect for the graphene sheets strongly depends on temperature, i.e. it is determined mainly by the thermal fluctuations. However, for $v > 10^6$m/s it will be dominated by quantum fluctuations. Strong enhancement of friction occurs in the case of resonant photon tunneling. As discussed above, resonant photon tunneling occurs for $v > 2v_F \approx 2\times 10^6$ m/s. For such velocities and $d = 1$ nm quantum friction dominates over the thermal contribution even at room temperature (see Fig. 25.8a). For $d = 10$ nm quantum friction dominates at low temperatures (see Fig. 25.8b).

The use of graphene in frictional drag experiments has considerable advantages in comparison with quantum wells. Such experiments can be performed in a vacuum where the contribution from the phonon exchange can be excluded. In vacuum one can easily measure the dependence of the frictional drag force on the separation between graphene sheets. Due to the high mobility of the charge carriers in graphene, the charge carriers can move with much higher drift velocity than in quantum wells.

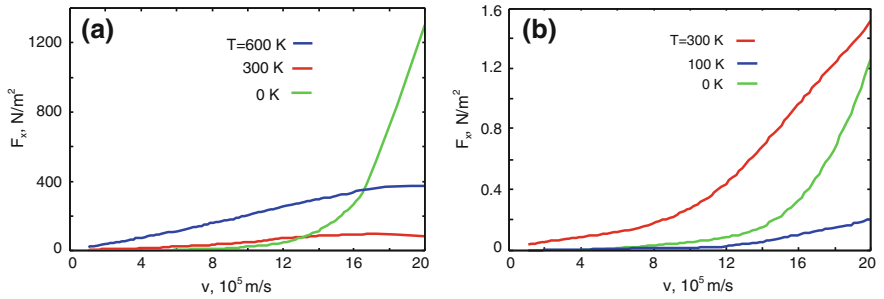


Fig. 25.8 Frictional drag between two graphene sheets at the carrier concentration $n = 10^{12}$ cm^{-2}. The finite temperature *curves* show only the thermal contributions to the friction. **a** Dependence of friction force between graphene sheets on the drift velocity of charge carriers in one graphene sheet at the layer separation $d = 1$ nm. **b** The same as in (**a**) but at $d = 10$ nm

25.5 Conclusion

Quantum friction, like superconductivity and superfluidity, is a macroscopic phenomenon, determined by the laws of quantum physics. The idea of quantum fluctuations of the electromagnetic field has found application in a wide variety of fields of physics. For example, the Lamb shift of atomic spectrum and anomalous magnetic moment of the electron were explained with the help of this idea. Quantum fluctuations of the electromagnetic field are determined by virtual photons-particles that are continuously created and annihilated in the vacuum. Using the metal mirror, moving with acceleration near the light velocity, virtual photons can converted into real, leading to radiation emitted by mirror. This is the dynamic Casimir effect; recently it was observed in superconducting waveguide [7]. There is an obvious similarity between the radiation that causes the dynamic Casimir effect, and Hawking radiation, which is responsible for quantum evaporation of black holes. Besides fundamental, quantum fluctuations have technological significance. Currently, an active work is conducting on the development of the nanoelectromechanical systems that can serve in the different fields of science and technology, like sensing, telecommunications, signal processing, storage, etc. For this reason, interest to the Casimir force has strongly increased in the last decade, as they define the interaction between nanostructures and are responsible for adhesion between the moving parts in the nanoelectromechanical systems.

For practical applications, it is important to learn tailoring of the Casimir forces. Lifshitz theory predicts that if two bodies are placed in the liquid, at the appropriate choice of material of the bodies the attraction between them can changed on the repulsion. The repulsive Casimir-Lifshitz force can provide a quantum levitation of objects in fluid and lead to a new class of customizable nanoscale devices with ultralow static friction.

Furthermore, quantum friction determines the ultimately limit to which the friction force can be reduced and, consequently, also the force fluctuations. The latter follows

from the fluctuation-dissipation theorem, a relation between friction and fluctuations established by Einstein. According to this relationship, the random force that makes a small particle jitter would also cause friction if the particle is dragged through the medium. The force fluctuations (and hence friction) are important for ultrasensitive force detection experiments. Perhaps the most exciting application of these ideas is associated with mechanical detection of nuclear spin resonance [74]. For example, a single spin detection using magnetic resonance force microscopy [75] (which was proposed to obtain images of biological objects, such as proteins, with atomic resolution) and for quantum computer [76] would require reducing the fluctuating forces (and therefore friction) to unprecedented levels. In addition, future measurements of the Casimir forces [18, 21–24], may eventually be limited by non-contact friction effects, one mechanism of which is determined by the Casimir friction with its limiting case—quantum friction. For these applications a better understanding of non-contact friction is only the first step. In the future it will be necessary to learn how to reduce it or, in other words, how to 'lubricate' vacuum.

Ten years ago the Casimir force and friction were academic curiosity. Today it is a technological issue: because broad applications of the nanoelectromechanical devices are predicted in various fields of science and technology, it is needed to learn to control forces, prevailing in the nano world. Quantum mechanics is fastly becoming quantum engineering.

Acknowledgments A.I.V. acknowledges financial support from Russian Foundation for Basic Research (Grant N 12-02-00061-a), European Science Foundation within activity "New Trends and Applications of the Casimir Effect" and COST Action MP1303 "Understanding and Controlling Nano and Mesoscale Friction". A.I.V. also thanks the Condensed Matter group of ICTP for hospitality during the time of working on this article.

References

1. J.B. Pendry, J. Phys. C **9**, 10301 (1997)
2. A.I. Volokitin, B.N.J. Persson, J. Phys.: Condens. Matter **11**, 345 (1999)
3. A.I. Volokitin, B.N.J. Persson, Phys. Rev. B **78**, 155437 (2008)
4. A.I. Volokitin, B.N.J. Persson, Rev. Mod. Phys. **79**, 1291 (2007)
5. A.I. Volokitin, B.N.J. Persson, Phys. Rev. Lett. **106**, 094502 (2011)
6. A.I. Volokitin, B.N.J. Persson, EPL **103**, 24002 (2013)
7. C.M. Wilson, G. Johansson, A. Pourkabirian, M. Simoen, J.R. Johansson, T. Duty, F. Nori, P. Delsing, Nature **479**, 376 (2011)
8. K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, Y. Zhang, S.V. Dubonos, Grigorieva I. V. and A. A. Firsov, Science **306**, 666 (2004)
9. K.S. Novoselov, A.K. Geim, S.V. Morozov, M.I. Katsnelson, I.V. Grigorieva, S.V. Dubonos, A.A. Firsov, Nat. (Lond.) **438**, 197 (2005)
10. A.K. Geim, K.S. Novoselov, Nat. Mater. **6**, 183 (2007)
11. A.K. Geim, Science **324**, 1530 (2009)
12. H.B.G. Casimir, Proc. K. Ned. Akad. Wet. **51**, 793 (1948)
13. E.M. Lifshitz (1955) Zh. Eksp. Teor. Fiz. **29** 94 (Engl. Trnsl. 1956 Sov. Phys.-JETP **2** 73)
14. S.M. Rytov, *Theory of Electrical Fluctuations and Thermal Radiation* (USSR Academy Publishing, Moscow, 1953)

15. I.E. Dzyaloshinskii, E.M. Lifshitz, L.P. Pitaevskii, Adv. Phys. **10**, 165 (1961)
16. S.K. Lamoreaux, Februrary, Phys. Today **60**, 40 (2007)
17. P.W. Milloni, The Quantum Vacuum: An Introduction to Quantum Electrodynamics (Academic) (1993)
18. A.O. Sushkov, W.J. Kim, D.A.R. Dalvit, S.K. Lamoreaux, Nat. Phys. **7**, 230 (2011)
19. F.M. Serry, D. Walliser, G.J. Maclay, J. Appl. Phys. **84**, 2501 (1998)
20. E. Buks, M.L. Roukes, Phys. Rev. B **63**, 033403 (2001)
21. G.L. Klimchitskaya, U. Mohideen, V.M. Mostepanenko, Rev. Mod. Phys. **81**, 1827 (2007)
22. J.N. Munday, F. Capasso, V.A. Parsegian, Nat. (Lond.) **457**, 170 (2007)
23. A.O. Sushkov, W.J. Kim, D.A.R. Dalvit, S.K. Lamoreaux, Phys. Rev. Lett. **107**, 171101 (2011)
24. Y. Bao, R. Guérout, J. Lussange, A. Lambrecht, R.A. Cirelli, F. Klemens, W.M. Mansfield, C.S. Pai, H.B. Chan, Phys. Rev. Lett. **105**, 250402 (2010)
25. A.W. Rodriguez, W.J. Kim, F. Capasso, S.G. Johnson, Nat. Photon. **5**, 211 (2011)
26. R. Zhao, J. Zhou, Th. Koschny, E.N. Economou, C.M. Soukoulis, Phys. Rev. Lett. **103**, 103602 (2009)
27. M. Antezza, L.P. Pitaevskii, S. Stringari, Phys. Rev. Lett. **95**, 113202 (2005)
28. D. Polder, M. Van Hove, Phys. Rev. B **4**, 3303 (1971)
29. E. Rousseau, A. Siria, G. Jourdan, S. Volz, F. Comin, J. Chevrier, J.J. Greffet, Nature Photon. 3 514 (2009)
30. S. Shen, A. Narayanaswamy, G. Chen, Nano Lett. **9**, 2909 (2009)
31. B.N.J. Persson, Z. Zhang, Phys. Rev. B **57**, 7327 (1998)
32. G. Barton, J. Phys.: Condens. Matter **23**, 355004 (2011)
33. J.S. Høye, I. Brevik, EPL **91**, 60003 (2010)
34. T.G. Philbin, U. Leonhardt, New J. Phys. **11**, 033035 (2009)
35. J.B. Pendry, New J. Phys. **12**, 033028 (2010)
36. U. Leonhardt, New J. Phys. **12**, 068001 (2010)
37. J.B. Pendry, New J. Phys. **12**, 068002 (2010)
38. A.I. Volokitin, B.N.J. Persson, New J. Phys. **13**, 068001 (2011)
39. T.G. Philbin, U. Leonhardt, New J. Phys. **13**, 068002 (2011)
40. A.I. Volokitin, B.N.J. Persson, Phys. Rev. B **68**, 155420 (2003)
41. A.I. Volokitin, B.N.J. Persson, Phys. Rev. B **74**, 205413 (2006)
42. M. Bordag, B. Geyer, G.L. Klimchitskaya, V.M. Mostepanenko, Phys. Rev. B **74**, 205431 (2006)
43. M. Bordag, I.V. Fialkovsky, D.M. Gitman, D.V. Vassilevich, Phys. Rev. B **80**, 245406 (2009)
44. D. Drosdoff, L.M. Woods, Phys. Rev. A **84**, 062501 (2011)
45. D. Drosdoff, L.M. Woods, Phys. Rev. B **82**, 155459 (2010)
46. I.V. Fialkovsky, V.N. Marachvsky, D.V. Vassilevich, Phys. Rev. B **84**, 035446 (2011)
47. B.E. Sernelius, EPL **95**, 57003 (2011)
48. V. Svetovoy, Z. Moktadir, M. Elwenspoek, H. Mizuta, EPL **96**, 14006 (2011)
49. J. Sarabadani, A. Naji, R. Asgari, R. Podgornik, Phys. Rev. B **84**, 155407 (2011)
50. G. Gómez-Santos, Phys. Rev. B **95**, 245424 (2009)
51. B. Shapiro, Phys. Rev. B **82**, 075205 (2010)
52. A.I. Volokitin, B.N.J. Persson, J. Phys.: Condens. Matter **13**, 859 (2001)
53. N.D. Mermin, Phys. Rev. B **1**, 2362 (1970)
54. B. Wunscvh, T. Stauber, F. Sols, F. Guinea, New J. Phys. **8**, 318 (2006)
55. E.H. Hwang, S. Das Sarma, Phys. Rev. B **75**, 205418 (2007)
56. J.H. Chen, C. Jang, S. Xiao, M. Ishigami, M.S. Fuhrer, Nat. Nanotechnol. **3**, 206 (2008)
57. M. Kisiel, E. Gnecco, U. Gysin, L. Marot, S. Rast, E. Meyer, Nat. Mater. **10**, 119 (2010)
58. K. Saitoh, K. Hayashi, Y. Shibayama, K. Shirahama, Phys. Rev. Lett. **105**, 236103 (2010)
59. M. Freitag, M. Steiner, Y. Martin, V. Perebeinos, Z. Chen, J.C. Tsang, P. Avouris, Nano lett. **9**, 1883 (2009)
60. M.B. Pogrebinskii (1977) Fiz. Tekh. Poluprov. **11**, 637 (Engl. Transl. Sov. Phys.- Semicond. 1977 **11** 372)
61. P.J. Price, Physica B+C **117**, 750 (1983)

62. T.J. Gramila, J.P. Eisenstein, A.H. MacDonald, L.N. Pfeiffer, K.W. West, Phys. Rev. Lett. **66**, 1216 (1991)
63. U. Sivan, P.M. Solomon, H. Shtrikman, Phys. Rev. Lett. **68**, 1196 (1992)
64. A.I. Volokitin, B.N.J. Persson, J. Phys.: Condens. Matter **13**, 859 (2001)
65. S. Kim, I. Jo, J. Nah, Z. Yao, S.K. Banerjee, E. Tutuc, Phys. Rev. B **83**, 161401 (2011)
66. R.V. Gorbachev, A.K. Geim, M.I. Katsnelson, K.S. Novoselov, T. Tudorovskyiy, T.V. Grigorieva, A.H. MacDonald, K. Watanabe, T. Taniguchi, L.P. Ponamarenko, Nature Phys. **8**, 896 (2012)
67. W.K. Tse, BenYu-Kuang.Hu, S. DasSarma, Phys. Rev. B **76**, 081401 (2007)
68. M.I. Katsnelson, Phys. Rev. B **84**, 041407(R) (2011)
69. N.M.R. Peres, J.M.R. Lopes des Santos, A.H. Castro Neto, Europhys. Lett. **95**, 18001 (2011)
70. E.H. Hwang, R. Sensarma, S. DasSarma, Phys. Rev. B **84**, 245441 (2011)
71. B.N. Narozhny, M. Titov, I.V. Gornyi, P.M. Ostrovsky, Phys. Rev. B **85**, 195421 (2012)
72. M. Carrega, T. Tudorovskiy, A. Principi, M.I. Katsnelson, M. Polini M, New J. Phys. **14**, 063033 (2012)
73. B.Amorin, N.M.R.Peres N M R, J. Phys.:Condens. Matter. **24** 335602 (2012)
74. Rugar D, R.J. Budakian, H.J. Mamin, B.W. Chui, Nature **430** 329 (2004)
75. J.A. Sidles, J.L. Garbini, K.J. Bruland, D. Rugar, O. Zuger, S. Hoen, C.S. Yannoni, Rev. Mod. Phys. **67**, 249 (1995)
76. G.P. Berman, G.D. Doolen, P.C. Hammel, V.I. Tsifrinovich, Phys. Rev. B **61**, 14694 (2000)

Chapter 26
Dissipation at Large Separations

Marcin Kisiel, Markus Langer, Urs Gysin, Simon Rast, E. Meyer and Dong-Weon Lee

Abstract When two macroscopic bodies slide in contact, energy is dissipated due to friction. Sometimes it is desired, like in case brakes in the bicycle, sometimes unwelcome–when you ask yourself why your automated coffee machine broke for the third time. In nanoscale, a tiny friction force is present when bodies in relative motion are separated by few nanometer gap. This non-contact form of friction might be successfully measured by highly sensitive cantilever oscillating like a tiny pendulum over the surface. The elusive non-contact friction might arise due to vdW interaction, which is mediated by the long-range electromagnetic field or in many cases by fluctuations of static surface charges arising from material inhomogeneities. The huge dissipation might also orginate from hysteretic switching of the studied material under the external action of the oscillating probe. In this chapter several experiments reporting on non-contact friction are discussed. First the Joule dissipation channel is discussed. Next we report on non-contact friction measurement over metal–superconductor transition, which allows to distinguish between phononic and electronic contribution to friction. The non-contact friction due to switching of the charge density wave is discused in the last part of this chapter.

M. Kisiel (✉) · M. Langer · U. Gysin · S. Rast · E. Meyer
Institute of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland
e-mail: Marcin.Kisiel@unibas.ch

M. Langer
e-mail: Markus.Langer@unibas.ch

U. Gysin
e-mail: urs.gysin@unibas.ch

D.W. Lee
MEMS and Nanotechnology Laboratory, Chonnam National University, Gwangju, South Korea
e-mail: mems@jun.ac.kr

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_26

26.1 Introduction

In a closed system, friction forces transfer work to heat. Energy dissipation is a non-reversible process, which is well known from everyday life [1, 2]. The mechanisms, which lead to these irreversible processes, are complex and still poorly understood. The vibration amplitude of a damped oscillator [3–5] decays in time that is equivalent to the fact that the kinetic energy is converted to heat. The energy transfer lasts until the cantilever system reaches its thermodynamic equilibrium. In this steady state stationary fluctuations from the mean value $\langle x \rangle$ are observed. Both decay time τ and equilibrium fluctuations $x(t)$ contain information about the dissipative process. The equation motion of a linear damped harmonic oscillator can be described by the knowledge of the spring constant k, the eigenfrequency ω_0 and the damping coefficient Γ, the effective mass m_{eff} and the external force f_{ext}:

$$m_{eff}\frac{d^2x}{dt^2} + \Gamma\frac{dx}{dt} + kx = f_{ext}(t) \tag{26.1}$$

The external force f_{ext} can be regarded as sum of a non-stochastic force and a stochastic force. Since (26.1) is a linear differential equation, both contributions can be treated separately. In both cases the same parameter set (ω_0, Γ, m_{eff}) is required to describe the system. All quantities (fluctuating and non-fluctuating) derived from the cantilever movement depend on the same parameter set. An experimentally accessible quantity to describe the dissipation process of a vibrating cantilever at its resonance frequency is its decay time τ. The quality factor Q is given by:

$$Q = \frac{\tau\omega_0}{2} \tag{26.2}$$

The knowledge of the spring constant k and decay time τ permits us to calculate the damping coefficient or so to say noncontact friction $-\Gamma$.[1] Γ is the proportionality constant between the non-conservative friction force F and the velocity v of the oscillator.

$$F = -\Gamma \cdot v_{rms} \tag{26.3}$$

where the friction coefficient is given by:

$$\Gamma = \frac{k}{\omega_0 Q}. \tag{26.4}$$

For an extremely soft free cantilever ($k = 130\,\mu$N/m) with a resonance frequency of 5.8 kHz and a vibration amplitude of x_0= 20 nm, a Q of 240,000 and a $\Gamma = 10^{-14}$ kg/s we obtain friction force $F = \Gamma\frac{\omega_0 x_0}{\sqrt{2}} = 7.6$ aN. The dissipated power

[1] Although, the term friction is commonly related to the situation of two surfaces in contact, we will stick to it in the rest of this chapter always referring to the noncontact form of friction.

$P = \Gamma \cdot v_{rms}^2 = 4.0 \times 10^{-21}$ Watt (2.5×10^{-2} eV/s or 4.2 μeV/cycle). For non-contact force sensors with a resonance frequency of 300 kHz, a $\Gamma = 10^{-11}$ kg/s and typical oscillation amplitude of 1 nm, the friction force is $F = 1.3 \times 10^{-14}$ N which corresponds to a dissipated power of 1.7×10^{-17} Watt (110 eV/s or 367 μeV/cycle). The experimentally determined friction coefficient Γ might be a superposition of different friction coefficients Γ_i

$$\Gamma = \Gamma_0 + \Gamma_1 + \Gamma_2 \cdots + \Gamma_N \tag{26.5}$$

$\Gamma(T, p, x,)$ represents a quantity which depends on the temperature T, the pressure p, the material properties and external fields (e.g. magnetic or electrical fields). As the cantilever approaches the sample surface the electromagnetic field between tip and surface has an impact on the friction coefficient. This leads to a distance dependent friction coefficient at small tip sample separations [1, 6–8].

The internal friction of the cantilever Γ_0 describes the friction losses which occur by simply bending the cantilever. Internal friction has to be measured under vacuum conditions to minimize the viscous damping due to inelastic scattering between the vibrating cantilever and gas molecules. At pressures below 10^{-6} mbar viscous damping can be neglected [9]. The internal friction depends mainly on the material properties and the sensor geometry. There are two possibilities to lower the kinetic energy of a vibrating cantilever dissipation and sound wave scattering. The cantilever can be regarded as a one side clamped bar. In reality the cantilever is a micro-mechanically etched silicon bar, which is connected to a support. If the junction between bar and support is not ideally constructed, there exists the possibility that sound waves are scattered from the cantilever to the support, which is reducing the kinetic energy of the cantilever. This is not a dissipative process! On the other hand kinetic energy is transferred into heat. It is experimentally a proven fact that a rectangular bar cantilever scatters a minimum energy into the support. The dissipative energy losses can be induced due to stress and strain acting on the cantilever. Experimentally it is hard to distinguish between the two mechanisms. The internal friction Γ_0 is omnipresent and determines if an additional friction coefficient $\Gamma_R = \Gamma_1 + \Gamma_2 + \Gamma_3 \cdots + \Gamma_N$ is detectable at all. Assuming the errors $\delta\Gamma$ and $\delta\Gamma_0$ are Gaussian distributed we find for Γ_R:

$$\Gamma_R = \Gamma - \Gamma_0 \pm (\delta\Gamma + \delta\Gamma_0) \tag{26.6}$$

Γ_R is limited by the sum of the statistical errors $\delta\Gamma + \delta\Gamma_0$. For an ultrasensitve cantilever (cf. Fig. 26.1) an internal friction coefficient of $\Gamma_0 = 1.47 \times 10^{-14}$ kg/s can be obtained under UHV conditions.

Approaching the cantilever close to a surface opens new dissipation channels. Noncontact friction is accurately measured by means of a very soft and sensitive cantilever, which however might jump into contact, when the attractive force gradient is larger than the spring constant. That happens on almost all surfaces because the attractive van der Waals forces and the only exception is the operation under liquid environment, where attractive forces are greatly reduced [11]. An easy way to

overcome this obstacle is to oscillate the cantilever in the so-called pendulum geometry [12, 13]. This way the cantilever is oriented perpendicularly to the surface and oscillates like a tiny pendulum without snapping into the contact with the measured sample (Fig. 26.2).

There are several methods to determine the friction coefficient Γ: ring-down measurements, the measurement of the power spectral density $S(\omega)$ of the cantilever fluctuations [14, 15] and the measurement of the excitation voltage A_{exc} needed to maintain constant amplitude cantilever's oscillation.

The smallest possible friction coefficients Γ can be sensed by using soft cantilever with a smallest possible constant k, a highest possible eigenfrequency and quality factor Q (26.4). Cantilevers with these features are very force sensitive. The minimal detectable force F_{min} is given by:

$$F_{min} = \sqrt{\frac{2k_B T \Gamma \Delta\omega}{\pi}} \tag{26.7}$$

where k_B is the Boltzmann constant, T and $\Delta\omega$ are the temperature and the bandwidth of the measurement. A change of the friction coefficient and the temperature is always accompanied by a change of the minimal detectable force (26.7).

Quantities like the frequency noise $\delta\omega$ of a cantilever which oscillates with an amplitude x_0 at its resonance frequency ω_0 depend on the friction coefficient Γ [16].

$$\delta\omega = \frac{\omega}{x_0 k}\sqrt{\frac{2k_B T \Gamma \Delta\omega}{\pi^3}} \tag{26.8}$$

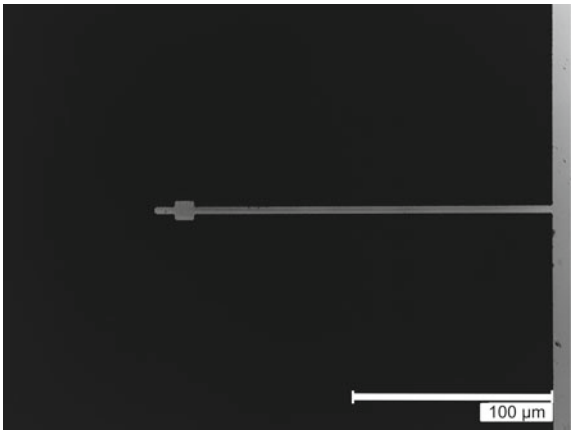


Fig. 26.1 Ultrasensitive single crystalline silicon force sensor with a length of 200 µm, a width of 5 µm and a thickness of 170 nm [10]. The spring constant is $k = 128$ µ N/m and the eigenfrequency $f_0 = 5.8$ kHz. A quality factor Q = 240,000 can be obtained after annealing under UHV conditions for several hours. An internal friction coefficient Γ_0 in the order of $(1.47 \pm 0.27) \times 10^{-14}$ kg/s is achieved. With a vibration amplitude of 20 nm the dissipated power is: 4.0×10^{-21} Watt (2.5×10^{-2} eV/s or 4.2 µeV/cycle)

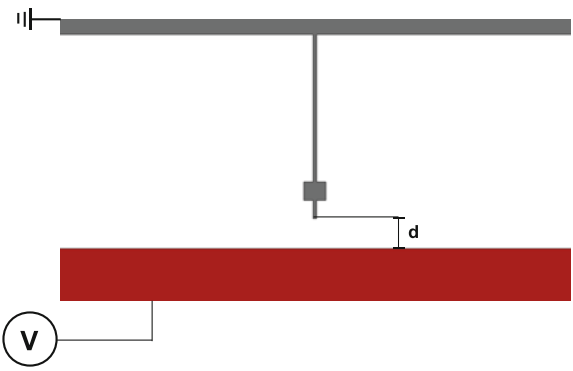


Fig. 26.2 Experimental setup: The cantilever is mounted perpendicularly to the surface to avoid a jump into contact. The distance d between force sensor and surface can be varied and a bias voltage V between cantilever and surface can be applied

To obtain the best possible frequency resolution of a self-driven cantilever the frequency noise $\delta\omega$ has to be minimized. Therefore, the measurement has to be performed at low temperatures, where thermal fluctuations are small and low Γ are found. The oscillator frequency noise $\delta\omega$ is influenced by Γ and the temperature T. With 26.7 the power spectral noise density $S(\omega)^2\Delta\omega = \delta\omega^2$ can be estimated in a narrow bandwidth

$$S(\omega)^2 = \frac{2k_bT\omega_0^2\Gamma}{\pi^3k^2x_0^2} \tag{26.9}$$

This contribution is structured as follows: First we investigate the internal friction of the cantilever Γ_0. In a second part we discuss different dissipation process between moving bodies separated by a nanometer sized distance d. Experiment on superconducting Nb surface allowed to distinguish between different dissipation channels—electronic and phononic. The last part reports on experiment on $NbSe_2$ surface, where the non-contact friction is caused by the hysteretic processes induced in the sample.

26.2 Internal Friction of the Cantilever

The temperature dependence of resonance frequency of cantilevers is rather well understood [17]. Geometry changes due to thermal expansion can be neglected. However, the temperature dependence of the Young's modulus of Silicon, $E(T)$, is given by the Wachtman-formula:

$$E(T) = E_0 - B \cdot T \exp\left(-\frac{T_0}{T}\right), \tag{26.10}$$

where T_0 is related to the Debye-temperature of the sensor material. With (26.10) the temperature dependent resonance frequency can be calculated:

$$\omega_n = \alpha_n^2 \frac{t}{L^2} \sqrt{\frac{E}{12\rho}}. \tag{26.11}$$

where $\alpha_1 = 1.875$ for the first eigenmode. t is the thickness, L the length and ρ the mass density. The experimental frequency versus temperature data are well fitted with $T_0 = 317\,\mathrm{K}$ [17]. According to $T_0 = D/2$ a Debye temperature of $D = 634\,\mathrm{K}$ is determined, which is in good agreement with literature values of $D = 645\,\mathrm{K}$ for silicon.

In contrast, the damping of cantilevers is still rather poorly understood. Several contributions have to be distinguished:

1. Damping due to thermoelastic effects
2. Damping due bulk losses
3. Damping due to surface losses
4. Damping due to acoustic emission into the bulk
5. Losses due to the clamping
6. Viscous damping due to the presence of gases or liquids.

As far as ultra-sensitive measurements under ultrahigh vacuum conditions are concerned, the influence of viscous damping at pressures below $10^{-6}\,\mathrm{mbar}$ can be neglected. The influence of clamping can be optimized by rigid holders and the exclusion of glues with high damping rates [18]. Damping due to acoustic emission is also found to be negligible in most practical cases. Therefore, the first three mechanisms are the most important ones.

26.2.1 Thermo-elastic Damping

The conduction of heat is an important energy loss mechanism. Periodical compression and expansion of oscillating micromechanical elements is associated with heat flow between compressed and expanded areas. The Zener-model is a continuum model of this thermo-elastic damping mechanism [19]. The internal friction is given by

$$Q^{-1} = \frac{\alpha^2 T E}{\rho c_P} \frac{\omega\tau}{1 + (\omega\tau)^2} \tag{26.12}$$

where α is the thermal expansion coefficient, c_p the specific heat capacity and ρ is the mass density. The relaxation time τ is given by:

$$\tau = \frac{t^2}{\pi^2} \frac{\rho c_p}{\kappa}, \tag{26.13}$$

where κ is the thermal conductivity. Typical parameters for silicon at room temperature are $E = 1.68\,\mathrm{GPa}$, $\alpha = 2.54 \times 10^{-6}\,\mathrm{K}^{-1}$, $\rho = 2.33 \times 10^{3}\,\mathrm{kgm}^{-3}$, $c_p = 711$ J $\mathrm{kg}^{-1}\mathrm{K}^{-1}$ and $\kappa = 150\,\mathrm{Wm}^{-1}\mathrm{K}^{-1}$. The temperature dependence of the Young's-modulus is small compared to the strong variations of thermal expansion (zero crossings at 20 and 125 K) [20]. The thermal conductivity in the bulk varies between 100–5,000 $\mathrm{Wm}^{-1}\mathrm{K}^{-1}$. One should also take into account that the thermal conductivity is reduced below $100\,\mathrm{Wm}^{-1}\mathrm{K}^{-1}$ due to phonon-boundary scattering for thickness of the order of microns at temperatures below 30 K [21].

At present, many experimental data indicate that thermo-elastic damping is the dominant loss mechanism at room temperature. At temperature below 200 K other channels start to dominate, which may be related to bulk or surface losses.

26.2.2 Bulk and Surface Losses

The scattering of elastic waves with defects on the surface or in the bulk is an important loss mechanism. The oscillation of the cantilever leads to a time dependent local stress field changing the energy landscape of the defects. Instabilities of these defects may occur, where atoms jump from one equilibrium position to another position. The energy difference between equilibrium positions is the activation energy. Therefore, damping vs. temperature curves show activation peaks, also called Debye peaks. So far, most of the experimental work is limited to crystalline silicon cantilevers which exhibit the highest Q-factor of available cantilevers. Typical Q-factors are between 10,000 up to 900,000. Comparable cantilevers made of Si_3N_4 or SiO_2 show much smaller Q-factors of 100–1,000. Therefore, we conclude that bulk losses are dominant for these amorphous structures. In the case of silicon, bulk or surface losses may become dominant at temperatures below 200 K. At 160 K a peak is observed, which may be related to such an activation peak with an activation energy of 0.25 eV. Unfortunately, the nature of these defects in silicon is still poorly understood. Simple defects, such as vacancies or interstitials are ruled out because of their high activation barriers [17]. Recently, it has been observed that the 160 K peak can be reduced strongly by annealing under vacuum conditions [22]. It is also observed that the peak does not shift with the resonance frequency, which is not in agreement with the simple activation energy model. The authors suggest that the 160 K peak is related to an adsorbate layer. Another peak at 30 K shifts with the resonance frequency and seems to be in better agreement with a Debye peak [22].

Coating of cantilevers leads to a strong increase of dissipation. The polycrystalline nature of these metallic films implies grain boundaries, where increased phonon scattering leads to an increase of damping losses. Also the surface quality, that is the presence of silicon oxide layer or adsorbates, such as H_2O or hydrocarbons, lead to rather large damping losses. Yang et al. annealed extremely small cantilevers (length <80 μm) [23]. They used rather high annealing temperatures (1,000 °C), which was sufficient to remove the oxide layers. It has been shown that annealing at temperatures below 600 °C of rather large silicon cantilevers (length of 400 or 500 μm

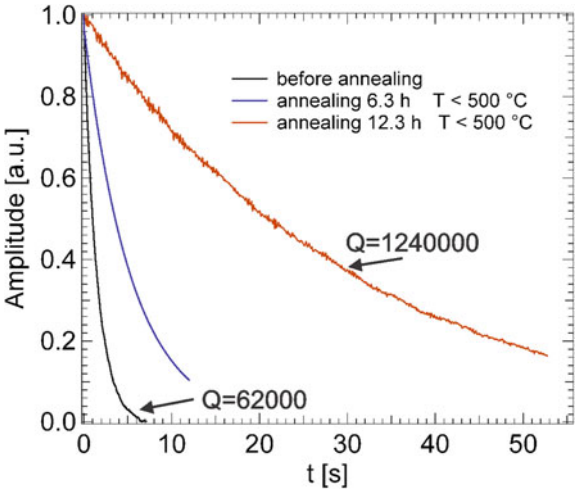


Fig. 26.3 Ring down measurement of a cantilever before and after annealing under UHV-conditions. The initial quality factor of 62,000 could be improved by an order of magnitude after 6 h annealing. Further annealing improved the quality factor to 1,240,000. The annealing temperature was too low to remove the oxide layer. Thus, the removal of weakly bound molecules, such as H_2O, OH or hydrocarbons, improves the quality factor

and thickness of 0.5–1.5 μm) under ultrahigh vacuum (UHV) conditions also leads to a reduction of dissipation [24]. In the case of Fig. 26.3 the quality factor of 62,000 could be improved by an order of magnitude after 6 h annealing at temperatures below 600 °C. Further annealing improved the quality factor to 1,200,000. The annealing temperature was too low to remove the oxide layer and thus results in removal of water or other weakly bound molecules, such as H_2O, OH or hydrocarbons. Alternatively, defects on the surface or in the bulk of the cantilever may be reduced by the annealing procedure. It is also known that long term annealing leads to negligible amount of localized charges at the end of the probing tip.

26.3 Dissipation at Large Separations

In most of the cases the dissipation between two moving bodies separated by a distance d is due to an electromagnetic interaction. The range of the interacting force determines the distance dependency of the friction coefficient Γ. Therefore, it is possible to distinguish between long-range and short-range friction forces. The mechanism of the long-range, electrostatic dissipation known also as joule dissipation is well understood [25]. Static electric fields between two different surfaces can exist without any externally applied voltage due to different work functions of different orientations of the crystallites of a polycrystalline surface. The work function

measured by applying a bias voltage between tip and sample is distance dependent [24]. This is due to the fact that the cantilever senses at large separations an average patch force resulting of several different oriented crystallites.

26.3.1 Dissipation due to Electromagnetic Interaction

The dissipative force can be minimized by compensating the local electric field. Electrostatic dissipation was previously observed by Denk et al. [7]. Later Stipe et al. [8] observed electrostatic dissipation at separations of 1–200 nm by using ultrasensitve force sensors. A gold tip was attached to an ultra-sensitive cantilever in the pendulum geometry. The measurements of Stipe et al. were performed under high vacuum conditions. Friction coefficients of the order of 10^{-13} kg/s between tip and metal substrate were observed. The distance dependence of the friction coefficient was fitted by a power law function $\Gamma \propto d^{-n}$ with an exponent $n = 1.3 \pm 0.2$. Volokitin et al. [26, 27] have calculated that dissipation at these large separations is consistent with the exponent $n = 1.5$ expected for the friction caused (on a spherical tip) by clean metal surface.

Rast et al. performed dissipation measurements under ultra high vacuum conditions at 7 K. They employed an ultrasensitive cantilever with a Co-Sm magnetic tip [24, 28]. The front edge of the tip is flat and has an area of $1.56\,\mu m^2$. As a substrate irradiated quartz was used (Suprasil 300) with 20 nm gold film. The friction coefficient is calculated according to a formula:

$$\Gamma = \Gamma_0 \left(\frac{A_{\mathrm{exc}}(d)}{A_{\mathrm{exc},0}} - \frac{f(d)}{f_0} \right), \tag{26.14}$$

where Γ_0 is the friction due to intrinsic losses of the cantilever, measured at large tip-sample separation, $A_{\mathrm{exc}}(d)$ and $f(d)$ are the distance dependent excitation amplitude (as measured by the excitation voltage needed to excite the cantilever at constant oscillation amplitude A) and frequency of the cantilever and the suffix zero refers to the free cantilever. As shown in Fig. 26.4 both conservative forces and dissipative forces increase with applied bias voltage at separations larger than 250 nm. Both obey to a quadratic power law Γ, $F \propto (V_{bias}^2)$. The quadratic behaviour of Γ versus bias voltage is direct evidence of the electronic origin of friction [26, 27, 29]. The friction coefficient has its minimum, where the conservative force is minimal and changes approximately by 2×10^{-13} kg/s when 1 V bias voltage is applied. Chumak et al. and Volokitin et al. [26, 27, 30] calculated the friction coefficient between a metallic tip and a flat metallic surface. They concluded that time dependent electric field induces local electric currents in the sample and cantilever. Thus the induction of the current is responsible for an energy dissipation.

Compensating the contact potential allows us to investigate the friction force which is not dominated by Joule losses. Figure 26.5 shows the distance dependence of the friction coefficient as a function of distance with compensated contact potential.

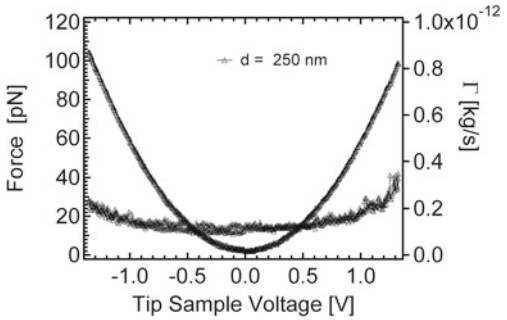


Fig. 26.4 Voltage-dependence of the normal force and friction coefficient at a separation of $d =$ 250 nm of a 20 nm thick gold film. The parabolic force dependence is related to capacitive forces. The friction coefficient increases with applied voltage. The quadratic behavior of the normal force and the quadratic behavior of friction coefficient in respect to bias voltage indicate that an electrical field is involved in the process

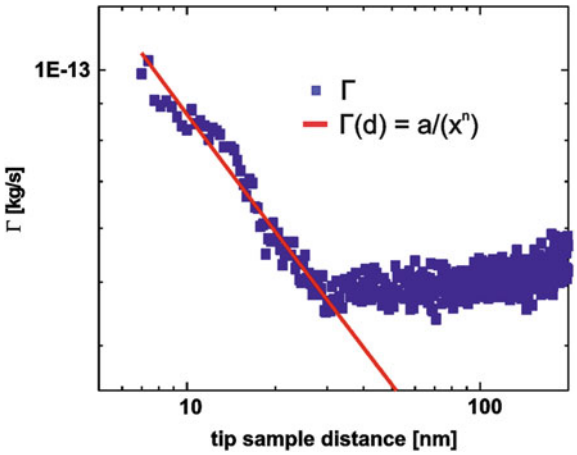


Fig. 26.5 Illustration of the friction coefficient with compensated contact potential at separations between 1 and 100 nm. At a separation larger than 30 nm the friction is governed by the internal friction 0 of the cantilever. Therefore, the friction coefficient of the long-range interaction has to be smaller than $\Gamma \ll 10^{-14}$ kg/s

At tip sample separations which are larger than 30 nm the friction force is dominated by the internal friction of the force sensor. At this separation the friction coefficient is smaller than $< 10^{-14}$ kg/s—the internal friction of the force sensor and the strength of the electromagnetic field between tip and sample is too small to create a measurable friction force. At separations below 30 nm the friction coefficient is larger than the internal friction of the cantilever and an non-reversible energy flow between tip and sample is observed. Volokitin (see previous chapter in this book) calculated the friction coefficient of a tip separated by a distance d generated by a van der Waals

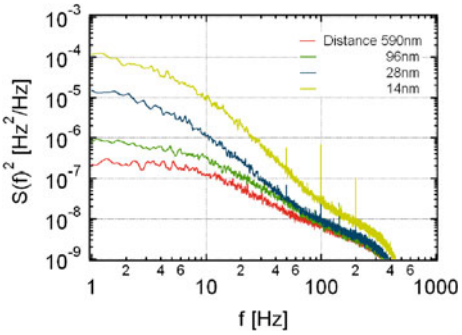


Fig. 26.6 Illustration of the power spectral density S(f) of the frequency noise of a self-oscillating cantilever with an amplitude of 10 nm close to a CaF_2 substrate under ambient conditions. At large tip sample separations the frequency noise corresponds to the frequency noise of the free cantilever. The values of S(f) decrease at higher frequencies since a lowpass filter limits the frequenc band. As smaller the distance between tip and surface is, as higher is the power spectral density. With decreasing distance between tip and surface, Γ increases which is reflected in an increase of $S(\omega)$ (26.9)

friction force in dielectrics. For a cylindrical tip with radius R and width w the calculated friction coefficient scales with $d^{-1/2}$, which fits to our data.

Usually the cantilever oscillations are controlled by means of phase-locked loop (PLL) feedback system, where the frequency ω caused by the tip sample interaction is measured. The frequency noise determines frequency resolution of the measurement (26.8). Figure 26.6 illustrates the power spectral density $S(\omega)$ of the frequency noise of a self-oscillating cantilever with an amplitude of 10 nm close to a CaF_2 substrate at room temperature under high vacuum conditions. The power spectral density is measured by sampling the frequency noise of the FM-demodulated signal of the cantilever vibration amplitude. At smaller tip sample separations the noise floor is increasing several orders of magnitude due to the distance dependent change of the friction coefficient Γ (26.9). The noise floor $S(\omega)$ is changing more at low frequencies (1–10 Hz) than at frequencies above 10 Hz.

26.3.2 Suppression of Electronic Friction in the Superconducting State

The pendulum AFM tip oscillating in close proximity to a substrate might induce phononic excitations. To estimate the phononic and electronic contributions to friction, Kisiel et al. [31] measured non-contact friction across the superconducting transition of a Nb sample. A 150 nm thick Nb film was deposited on a Si(100) substrate and was subsequently cleaned by a few cycles of Ar sputtering and annealing under UHV conditions. A ultrasensitive cantilever with spring constant $k = 30$ mN/m and

resonance frequency $f = 5.3\,\text{kHz}$ was used. Afer long term annealing under UHV condition the quality factor was improved up to $Q = 500{,}000$. The corresponding internal dissipation was equal $t_0 = 2.0 \times 10^{12}\,\text{kg/s}$ at temperature $\text{T} = 6\,\text{K}$. The cantilever end was exposed to focus ion beam (FIB) to form a sharp tip with spherical apex, approximately 50 nm in diameter (inset in Fig. 26.7). The oscillation amplitude $A = 5\,\text{nm}$ was constant during the measurement. The temperature dependence of the friction coefficient over critical temperature of Nb ($T_c = 9.2\,\text{K}$) is shown in Fig. 26.7. Here the tip sample distance of $\text{f} = 0.5\,\text{nm}$ was constant and the contact potential difference was compensated. The friction coefficient was measured by means of ring down method (see (26.2) and (26.4)). The dissipated power rise by a factor of three when the critical temperature T_c is approached from below and levels off in the normal metal state. It is equal about 25 μeV/cycle and 80 μeV/cycle for the superconducting and metallic phase, respectively. The rise of friction coefficient in the vicinity of T_c has a smooth character, which is in good agreement with BCS (Bardeen-Cooper-Schrieffer) theory of superconductivity [32]. The interpretation of the experiment is based on the theoretical analysis proposed by Persson [33, 34], who realized that friction over superconducting phase transition is analog to the acoustic attenuation of the longitudinal acoustic phonons known from BCS theory. The temperature dependence of surface contribution to the electronic friction has the form:

$$\frac{\Gamma_{surf}(T)}{\Gamma_{surf}(T_c)} \simeq \frac{2}{exp\left(\Delta(T)/k_B T\right) + 1}, \tag{26.15}$$

where $\Delta(T)$ is the temperature dependent energy band gap:

$$\Delta(T) = C \cdot k_B T_c \left(1 - (T/T_c)\right)^{0.5} \tag{26.16}$$

The factor $C = 3.52$ is the same for all BCS superconductors and determines the ratio of the energy gap at $T = 0\,\text{K}$ to the thermal energy at T_c. The data shown in Fig. 26.7 agree with (26.15), with the fit parameter $C = 3.8 \pm 0.7$. The cantilever tip oscillating very close to the Nb-film produce the surface acoustic longitudinal waves due to time dependent mechanical stresses acting on the sample. The energy is lost to the emission of phonons. The acoustic wave can interact only with the normal electrons near the Fermi surface. Close to T_C the electron population is gradually growing and the acoustic wave attenuation rises rapidly. Thus the electronic friction caused by electron—phonon interaction vary with the temperature according to the formula (26.15).

Further support to this hypothesis comes from voltage V and distance d dependent friction. In Fig. 26.8 the $\Gamma(V)$ is shown in normal, metallic state ($\text{T} = 13\,\text{K}$) and superconducting state ($\text{T} = 5.8\,\text{K}$), both acquired at the same tip sample distance $d = 0.5\,\text{nm}$. In the normal state data follow a quadratic dependence $\Gamma(V) \propto V^2$, as expected for the electronic friction [26, 27], whereas in the superconducting state $\Gamma(V) \propto V^4$. Volokitin et al. [26, 27] calculated that, when an spherical tip is oscillating laterally above the elastic surface the friction $\Gamma \propto F^2$. Since $F \propto V^2$,

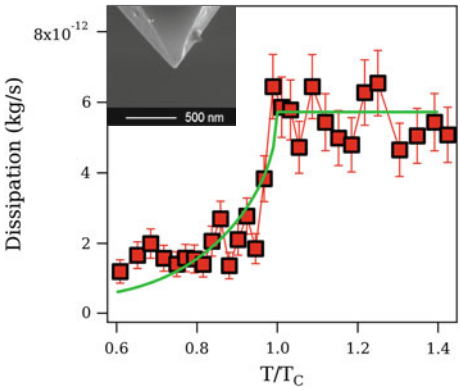


Fig. 26.7 Temperature dependence of the non-contact friction coefficient of the Nb surface. The friction coefficient increases as the critical temperature $T_c = 9.2\,\mathrm{K}$ is approached and levels off in the Nb normal metal state. The green line is a fit of (26.15) to the measured data. Inset: the cantilever tip with spherical apex prepared by means of focused ion beam

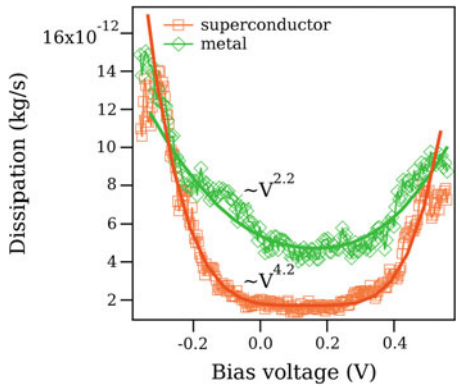


Fig. 26.8 The voltage dependence of the non-contact friction coefficient between oscillating cantilever tip and Nb surface, measured in superconducting (*red*) and normal (*green*) state. The dessipation has a parabolic dependence in metallic state, while $\Gamma \propto V^4$ for superconducting state of Nb. The tip sample distance is constant and equal $d = 0.5\,\mathrm{nm}$

the friction coefficient has to vary as a fourth power of the voltage, which is indeed the case in the superconducting state. The dependence of the friction coefficient Γ on distance d is shown in Fig. 26.9, again for normal and superconducting state of Nb film. In both cases the contact potential was compensated and equal about $V_{CPD} = 150 \pm 30\,\mathrm{mV}$. The distance was swept in between $0 < d < 3\,\mathrm{nm}$. As expected, the friction coefficient has much steeper distance dependence in the superconducting state as compared to the metallic state. The fit of the negative power function $\Gamma \propto d^{-n}$ to the experimental data gives $n = 1.0 \pm 0.1$ and $n = 3.8 \pm 0.3$ for normal and superconducting state, respectively. In normal state the distance dependence is in

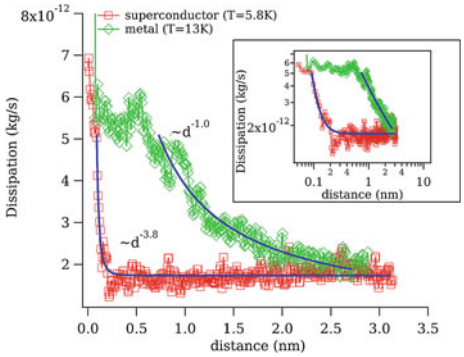


Fig. 26.9 Distance dependence of the non-contact friction coefficient for Nb superconducting (*red*) and Nb normal metal (*green*) state. In normal state clear rise of dissipation is visible few nanometers away from the surface, otherwise as in the case of superconducting Nb state. The data are well fitted by inverse power law dependence $\Gamma \propto d^{-1}$ and $\Gamma \propto d^{-4}$ respectively for normal and superconducting Nb state

good agreement with the results obtained by Stipe et.al ($n = 1.3$) [29]. In the case of phononic friction, Lifshitz theory of van der Waals interaction [35] predicts that elastic stress leads to a vdW force $F(d) \propto d^{-2}$. Thus the phononic friction should vary as $\Gamma_{ph} \propto d^{-4}$ and the experimental value is in excellent agreement with this prediction.

26.3.3 The Noncontact Friction due to Phase Slips of the Charge Density Wave (CDW) in $NbSe_2$ Sample

So far we deal with friction having the form of viscous drag as given by the 26.3. It this section we report on data when dissipation is produced by cycle of hysteretic processes induced in the sample by oscillating cantilever tip [36]. The sample of interest is $NbSe_2$—an intercalated dichalcogenide compound with bulk charge density wave (CDW) state accompanied by periodic lattice distortion (PLD). The CDW materials have been an object of scientific interest since mid-1970s because their unusual properties; namely nonlinear conductivity, huge dielectric constants, elastic and thermal properties, etc. [37]. Most of them are related to the CDW motion under the external force. The probe consisted of a soft cantilever (ATEC-CONT from Nanosensors) with spring constant $k = 120$ mN/m, the resonance frequency $f = 12$ kHz, quality factor $Q = 9.0 \times 10^5$ and a friction coefficient $\Gamma_0 = 1.7 \times 10^{-12}$ kg/s (at 6 K). The probe was operating in pendulum geometry, however otherwise as in the case of conventional pendulum AFM experiments, the cantilever tip was chosen to be asymmetric (see inset Fig. 26.10). Due to the tip design the cantilever lateral oscillatory motion also implies a normal action. The lateral oscillation amplitude

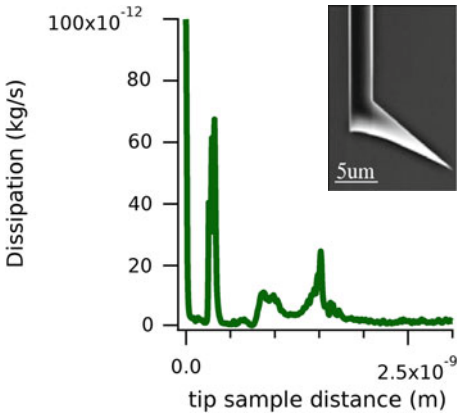


Fig. 26.10 The energy dissipation as a function of tip $NbSe_2$ sample distance for compensated contact potential difference. Three dissipation spikes positioned few nanometers above the surface are clearly visible. Inset: the pendulum AFM cantilever tip used in the experiment

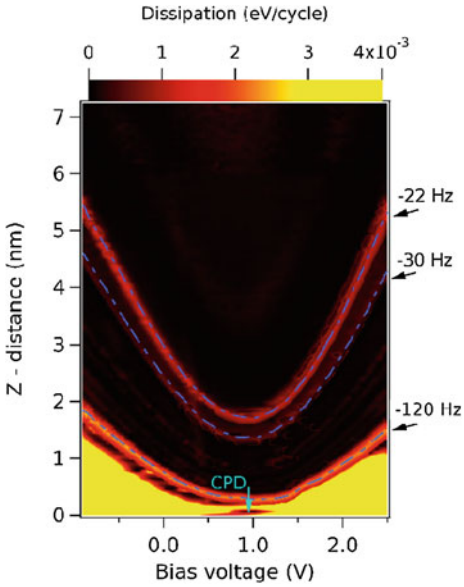


Fig. 26.11 Energy dissipation between $NbSe_2$ surface and oscillating pendulum AFM cantilever tip versus tip sample voltage and tip sample distance. Bright features correspond to multiplet of dissipation spikes observed few nanometers above the surface. The *peaks* always follow the same cantilever frequency $f = f_0 - 22\,Hz$, $f = f_0 - 30\,Hz$, $f = f_0 - 120\,Hz$ and constant frequency contours are shown with *dashed lines*

$A = 5nm$ was kept constant by means of PLL. Due to tip asymmetry the normal oscillation amplitude was equal to $A_{norm} = 180\,pm$. The measurement of friction versus distance $\Gamma(d)$ shows striking multiple of dissipation peaks arising at few nanometer distance from $NbSe_2$ surface, as shown in Fig. 26.10. Here the CPD was compensated. The friction coefficient was calculated according to 26.14. The presented result is contrary to the most of the non contact friction experiments, discussed so far in this chapter. Typically Γ increases smoothly as the tip approaches the sample and eventually saturates when the tip is close to contact with the surface. The amount of dissipated power (up to $P = 2\,meV/cycle$) is much larger as compared to the experiment on Nb surface, for instance.

The dissipated power $P(z, V)$ versus tip sample distance z and tip sample bias voltage V is shown in Fig. 26.11. The bright features correspond to high dissipation maxima. Three dominant of them are superimposed with dashed lines which correspond to contours of constant cantilever oscillation frequencies—$f_0 - 22\,\text{Hz}$, $f_0 - 30\,\text{Hz}$ and $f_0 - 120\,\text{Hz}$—as tip approaches the sample's surface. Within each dissipation branch the amount of energy loss stays constant, meaning it doesn't depend on bias voltage V. The huge non contact friction increase is observed even after careful compensation of CPD between tip and sample. In other words no matter the character of the interaction force is van der Waals ($V = V_{CPD}$) or electrostatic ($V \neq V_{CPD}$). Moreover, $z(V)$ dependence of the particular dissipation branch has a parabolic behaviour. In fact, for capacitively coupled conical tip and sample the force varies as $F \propto V^2/z$, meaning that the each dissipation peak always follows the same tip-sample interaction force. The above observations mean that the effect is force controlled rather than voltage controlled.

At this stage, it is necessary to mention that the similar dissipation peaks might occur if trapped charges are present at the end of the oxidized tip and they do not remain constant on the timescale of the experiment. That might be for instance due to charge leakage across the oxide layer [38]. In order to exclude this process the whole cantilever bar was annealed up to 700 °C, which results in removal of static charges from the probing tip. Moreover, to confirm further the universal character of the effect the experiment was repeated with an alternative instrument—the tuning fork AFM furnished with the metallic tip made of tungsten. In this setup the oscillation direction was perpendicular to the surface. The frequency, spring constant, quality factor and oscillation amplitude were equal to $f_0 = 25\,\text{kHz}$, $\text{k} = 2{,}000\,\text{N/m}$, $\text{Q} = 25{,}000$ and $\text{A} = 200\,\text{pm}$, respectively. Figure 26.12 shows dependence of non contact friction versus tuning fork tip-sample interaction force for several values of bias voltage ($-0.4\,\text{V} \leq \text{V} \leq 1.5\,\text{V}$). Again the friction measurements systematically show the existence of three dissipation maxima positioned at different interaction forces $\text{F}_{int} = -6.4, -8.2, -11.8\,\text{nN}$, as the tip approaches the surface.

In order to understand the origin of dissipation spikes a theoretical model was proposed, where the CDW is considered to be an elastic medium, perturbed locally by the attractive potential of the cantilever tip [39, 40]. The total energy is estimated according to equation:

$$E[\phi(x)] = \int [(\nabla\phi(x))^2 + V(x)\rho(x)]\mathrm{d}x. \tag{26.17}$$

where first term is an elastic energy and the second stands for perturbation. The $\rho(x) = \rho_0 \cos(Qx + \phi(x))$ is an electron charge, where $\rho_0 = const$ and $\phi(x)$ are CDW amplitude and CDW phase. The energy given by 26.17 is next minimized in order to find the preferential shape of the phase. This is shown in Fig. 26.13. For a given perturbation the charge peak below the tip displaces (red solid line) and the resulting CDW phase is getting distorted (red dashed line). As the force reaches critical value the central peak disappears and an extra 2π phase is pumped locally below the cantilever tip (green dashed line). This mechanism is provided

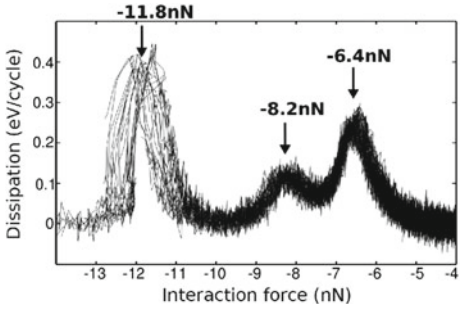


Fig. 26.12 The dissipation versus tip sample interaction force measured by means of tuning fork sensor. Many *curves* for different bias voltages are shown and the dissipation maxima, positioned always at the same interaction force are clearly visible

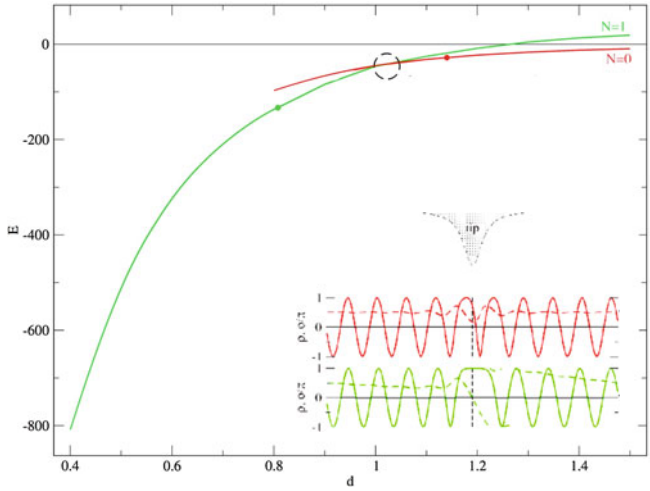


Fig. 26.13 The calculated energy E as a function of distance d for locally perturbed elastic charge density wave. The tip oscillation around crossover point (marked by a *circle*) causes hysteresis in tip dynamics. Inset shows the charge density (*solid lines*) and phase (*dashed lines*) of the CDW under the tip perturbation respectively for two different energy-distance curves

by the "phase slip". The energies for the two discussed phase deformations are shown in Fig. 26.13 and a crossover energy after which the second phase deformation becomes more favourable then the first is marked by a circle. The tip oscillation near the threshold causes that phase motion is quasi-periodic, alternating between rapid advances by 2π and a slow creep of the phase. Since the process is hysteretic, it implies dissipation. The discussed mechanism works in broad frequency range, even if the perturbation frequency is small. Approaching the tip further down, there can be another phase slip explaining multiple dissipation peaks.

26.4 Summary and Conclusions

Thermal fluctuations and fluctuating electromagnetic fields are the limiting factors for ultrasensitive measurements close to the surface. Annealing the cantilever under UHV-conditions removes adsorbates on the cantilever surface and increases the force sensitivity about an order of magnitude. Cooling to cryogenic temperatures reduces the thermal fluctuations and lowers the internal friction further down. The temperature dependent quality factor can be described by the Zener's theory of thermoelastic damping. Varying the distance between tip and sample opens new dissipation channels. Up to tip sample separations of 250 nm the long-range Joule dissipation is observed. Joule dissipation is relatively well understood, but the quantitative analysis indicates that both the probing tip and the cantilever contribute to dissipation. By compensating the local electric field Joule dissipation can be minimized and attonewton sensitivity can be achieved at separations down to 30 nm. Below 30 nm and for compensated contact potential the energy losses might be caused by short-range forces such as van der Waals forces. The measurement across the superconducting phase transition allows to distinguish between electronic and phononic contribution to friction. The experiment on Nb surface showed suppression of electronic dissipation in the superconducting state by factor of three. The friction coefficient increases as the critical temperature is approached from below. The rise in friction has a form of longitudinal acoustic phonon attenuation known from BCS theory. The distance and voltage dependence of the friction coefficient showed that friction has a phononic and electronic character respectively in superconducting and normal state. Most experiments on non-contact friction deal with linear frictional response of the system under study. In that case the energy looses could be accessed only by means of ultra-sensitive probes characterized by very small spring constants. The recent experient on $NbSe_2$ surface reports on friction produced by hysteretic effects induced in the sample. The experiment showed that it is possible to couple phase of charge density wave of $NbSe_2$ to the oscillating tip at few nanometer distance and the observed dissipation peaks are due to series of CDW phase slip events. For nonlinear frictional response the dissipated power is few order of magnitude larger as compared to friction losses measured on γ-irradiated quartz crystal or Nb surface.

Acknowledgments We express our gratitude to E. Gnecco, L. Marot, R. Pawlak, F. Pellegrini, G.E. Santoro, R. Buzio, A. Gerbi, G. Balakrishnan, A. Baratoff and E. Tosatti for collaboration and helpful discussions. This work was supported by the Swiss National Science Foundation.

References

1. B. Gottsmann, H. Fuchs, Phys. Rev. Lett. **86**, 2597 (2001)
2. A. Vanossi, N. Manini, M. Urbakh, S. Zapperi, E. Tosatti, Rev. Mod. Phys. **85**, 529–552 (2013)
3. G. Binnig, C. Quate, Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986)
4. S. Morita, R. Wiesendanger, E. Meyer, *Noncontact Atomic Force Microscopy* (Springer, Berlin, Germany, 2002)

5. F.J. Giessibl, Rev. Mod. Phys. **75**, 949 (2003)
6. I. Dorofeyef, H. Fuchs, G. Wenning, B. Gotsmann, Phys. Rev. Lett **83**, 2402 (1999)
7. W. Denk, D.W. Pohl, Appl. Phys. Lett. **59**, 2173 (1991)
8. B.C. Stipe, H.J. Mamin, T.D. Stowe, T.W. Kenny, D. Rugar, Phys. Rev. Lett. **86**, 2874 (2001)
9. S. Rast, Sensoren mit geringer Dissipation zur Messung kleiner Krafte, Diss. Phil.-naturwiss. Fak. Basel, 1999
10. D.W. Lee et al., J. Magn. Magn. Mater. **18–19**, 2413–2417 (2006)
11. F. Ohnesorge, G. Binnig, Science **260**, 1451 (1993)
12. D. Rugar, R. Budakian, H.J. Mamin, B.W. Chui, Nature **43**, 329 (2004)
13. U. Gysin, S. Rast, M. Kisiel, C. Werle, and E. Meyer Low temperature ultrahigh vacuum noncontact atomic force microscope in the pendulum geometry. Rev. Sci. Instrum. **82**, 023705 (2011)
14. S. Rast, C. Wattinger, U. Gysin, E. Meyer, Nanotechnology **11**, 169 (2000)
15. M. Roukes et al., J. Appl. Phys. **95**, 2682 (2004)
16. T. Albrecht et al., T.R. Albrecht, P. Grutter, H.K. Horne, D. Rugar, J. Appl. Phys. **69**, 668–673 (1991)
17. U. Gysin, S. Rast, P. Ruff, D.W. Lee, P. Vettiger, C. Gerber, Phys. Rev. B **69**, 045403 (2004)
18. J. Lübbe, L. Tröger, S. Torbrügge, R. Bechstein, Ch. Richter, A. Kühnle, M. Reichling, Achieving high effective Q - factors in ultra-high vacuum dynamic force microscopy. Meas. Sci. Technol, **21**, 125501 (2010)
19. C. Zener, Phys. Rev. **53**, 100 (1938)
20. K.G. Lyon, G.L. Salinger, C.A. Swenson, G.K. White, Linear thermal expansion measurements on silicon from 6 to 340 K. J. Appl. Phys. **48**, 865–868 (1977)
21. M. Ashegi, Y.K. Leung, S.S. Wong, K.E. Goodson, Appl. Phys. Lett. **71**, 1798 (1997)
22. H. Haucke, X. Liu, J.F. Vignola, B.H. Houston, M.H. Marcus, J.W. Baldwin, Appl. Phys. Lett. **86**, 191903 (2005)
23. J. Yang et al., Appl. Phys. Lett. **77**, 3860 (2000)
24. S Rast, U. Gysin, P. Ruff, Ch. Gerber, E. Meyer D.W. Lee, Nanotechnology **17**(7), 189 (2006)
25. Denk, Winfried, Pohl, W. Dieter, Local electrical dissipation imaged by scanning force microscopy. Appl. Phys. Lett. **59**, 2171–2173 (1991)
26. A. Volokitin, B. Persson, Phys. Rev. Lett **94**, 086104 (2005)
27. A.I. Volokitin, B.N.J. Persson, H. Ueba, Phys Rev. B **73**, 165423 (2006)
28. M.J. Naughton et al., Rev. Sci. Instr. **68**, 4061 (1997)
29. B.C. Stipe, H.J. Mamin, T.D. Stowe, T.W. Kenny, D. Rugar, Phys. Rev. Lett. **87**, 096801–1 (2001)
30. A.A. Chumak, P.W. Milonni, G.P. Berman, Phys. Rev. B **70**, 085407 (2004)
31. M. Kisiel, E. Gnecco, U. Gysin, L. Marot, S. Rast, E. Meyer, Nat. Mater. **10**, 119 (2011)
32. J. Bardeen, L.N. Cooper, J.R. Schrieffer, Phys. Rev. **108**, 1175 (1957)
33. B.N.J. Persson, *Sliding Friction* (Springer, Berlin, 2000)
34. B.N.J. Persson, Solid State Commun. **115**, 145 (2000)
35. R. Lifshitz, M. Roukes, Phys. Rev. **B61**, 5600 (2000)
36. M. Langer, M. Kisiel, R. Pawlak, F. Pellegrini, G.E. Santoro, R. Buzio, A. Gerbi, G. Balakrishnan, A. Baratoff, E. Tosatti, E. Meyer, Giant frictional dissipation peaks and charge-density-wave slips at the NbSe2 surface. Nat. Mater. **13**, 173–177 (2014)
37. C. Schlenker, in *Physics and Chemistry of Low-Dimensional Inorganic Conductors, North Atlantic Treaty Organization*. NATO Advanced Study Institute on Physics and Chemistry of Low-Dimensional Inorganic Conductors (Plenum Press, New York, 1996)
38. J. Polesel-Maris, A. Piednoir, T. Zambelli, X. Bouju, S. Gauthier, Nanotechnology **15**, S24–S29 (2004)
39. H. Fukuyama, P.A. Lee, Dynamics of the charge-density wave. I. Impurity pinning in a single chain. Phys. Rev. B **17**, 535–541 (1978)
40. P.A. Lee, T.M. Rice, Electric field depinning of charge density waves. Phys. Rev. B **19**, 3970–3980 (1979)

Part VIII
Applications

Chapter 27
Nanotribology of MEMS/NEMS

Satish Achanta and Jean-Pierre Celis

Abstract Micro-/Nano- electromechanical systems (MEMS/NEMS) are future devices that have a spectrum of applications ranging from rocket technology to biological sciences. Although MEMS devices are known for over two decades, very few categories of them are used in commercially applications due to their poor reliability. Tribological phenomena like stiction, friction, and wear are major issues affecting the reliability of contact MEMS/NEMS devices and micromotors, microgears, nanosliders, etc., are some examples in which reliability is greatly hampered by such dissipation processes. In recent years, lot of research was dedicated for improving the reliability of MEMS/NEMS through lab scale tribological studies e.g., nanotribological studies. Similar tribological studies were earlier carried out on components like magnetic storage devices, electrical connectors, etc., and were successfully tackled. Present chapter is an overview of the research done over the years to understand the phenomena like stiction, friction and wear in such devices. This chapter addresses the variety of tribological problems associated with MEMS/NEMS, tribological characterization through lab scale and in-situ techniques, and various solutions that are used for improving the reliability of such devices with respect to tribological problems.

List of Symbols

Symbol	Definition	Units
r	Radius of a meniscus	m
R	Radius of a ball	m

S. Achanta (✉) · J.-P. Celis
Department of MTM, Katholieke Universiteit Leuven, Kasteelpark Arenberg 44,
3001 Leuven, Belgium
e-mail: achanta.satish@gmail.com

J.-P. Celis
e-mail: jean-pierre.celis@mtm.kuleuven.be

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_27

γ_l	Surface tension of a fluid	J/m^2
F_{cap}	Capillary force between ball and counterbody	N
z	Distance of separation between two surfaces	m
d_{cap}	Characteristic distance for capillary condensation	m
θ	Contact angle	–
$e_{cap}(z)$	Interaction energy function	J/m^2
A_{Ham}	Hamaker constant	J
d_{ret}	Maximum distance between molecules for Van der Waals forces	m
d_{co}	Cut-off distance between molecules for repulsive interactions	m
e_{vdW}	van der Waals interaction function	J/m^2
F_{total}	Total normal force	N
$F_{applied}$	External applied normal force	N
$F_{adhesion}$	Adhesion force	N
F_a	Pull-off force	N
γ	Work of adhesion	J/m^2
Γ	Surface interaction energy	J/m^2
t	Thickness of beam	m
h	Initial separation between substrate and cantilever	m
s	Deflection length of a cantilever	m
F_f	Friction force	N
a	Modified Hertzian contact length	m
μ	Coefficient of friction	–
L	Applied normal force	N
K	Effective elastic modulus of a material couple	N/m^2

27.1 MEMS/NEMS Devices, Applications, and Their Reliability Issues

The miniaturization of devices continues to fuel the integrated circuit industry. The fabrication of micro-/nano-sized complex structures and devices have been made possible thanks to the advancement in lithographic techniques. Micro-/nano-electromechanical systems (MEMS/NEMS) form an integral part of the microelectronic industry. In 2000, the MEMS/NEMS industry was worth approximately $15 billion and with a projected 10–20 % annual growth rate, it is expected to worth more than $100 billion by the end of this decade [1]. MEMS/NEMS devices are group of products ranging in size from a micron to a centimetre that combine mechanical and electrical structures. They may also consist of micromechanical components such as comb drives, microgears, microlevers, etc., which move to perform certain tasks, and microelectronic components to control motion or to obtain information from that motion [2]. To give an idea of their size and complex structure, a spider mite next to a MEMS device with series of comb drives is shown in Fig. 27.1 [3]. MEMS technology is a general term used for materials and processes required to make MEMS components, the integration of such components to make devices (e.g., sensors, actuators), and their applications. MEMS/NEMS find a wide range of applications in engineering fields ranging from electronic devices, space technology to biological sciences due to their existence in many forms and the ability to engineer

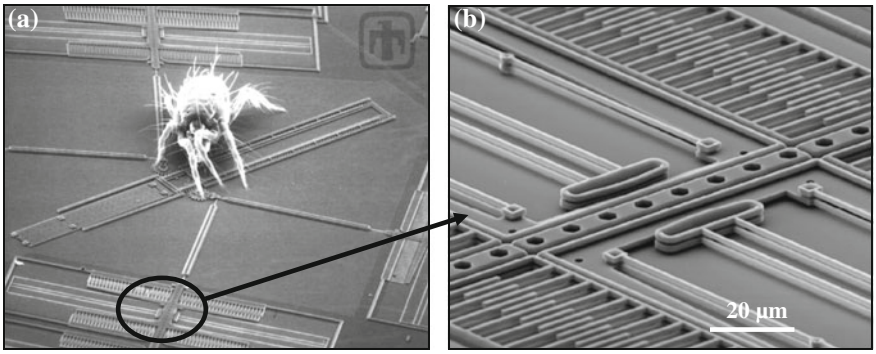


Fig. 27.1 **a** MEMS device along with a spider mite indicating the size of these microcomponents. **b** Zoom-in of the picture showing comb drives (series of aligned cantilevers) used in MEMS devices [3]

these devices as necessary for given applications [4]. MEMS are commercially used in inkjet printer heads, microwave switches, accelerometers, and sensors.

The future of MEMS/NEMS looks bright as new types of microsystems emerge in drug delivery systems, optical switches, chemical lab-on-a-chip systems, gas turbine engines, microgears, chemical sensors, infrared imagers, etc. [5]. Nano-electromechanical systems (NEMS) are future devices that are further miniatures of MEMS in which typical dimensions of the device structures range in nanometers. Recent examples of NEMS devices are shown in Fig. 27.2. Sensitive sensors (Fig. 27.2a) are being developed like force and mass detection down to molecular level, high-frequency resonators (GHz range), and ultra-sensitive low-power switches [6]. Molecular gears (Fig. 27.2b), motors and nano bearings and engines are future devices that will be used in medical treatments inside the human body [7]. Even though MEMS/NEMS offer a wide spectrum of applications, their poor reliability is a major hindrance to their commercialization. Most of these devices suffer from a range of reliability issues depending on conditions under which they are manufactured, and operated. In recent years, the reliability of MEMS has gained importance and researchers around the world are putting lot of efforts in unraveling these failure issues. The MEMS technology is still in its infancy, and not much has been done concerning their reliability [8].

Hereafter, tribological issues like adhesion, friction and wear are addressed that are hampering the commercialization of MEMS/NEMS devices. Recent advancements made in evaluating MEMS reliability and ways to mitigate these disastrous effects are reviewed. On miniaturizing any device or system, it is critical to have a good understanding of the scaling properties of the overall design, materials, and the fabrication processes involved. The scaling properties related to any one of these components could present a formidable barrier to achieving adequate performance or economic feasibility [9]. MEMS are designed with some basic parts which appear repeatedly in the structure like cantilever beams, membranes, springs, hinges, gears,

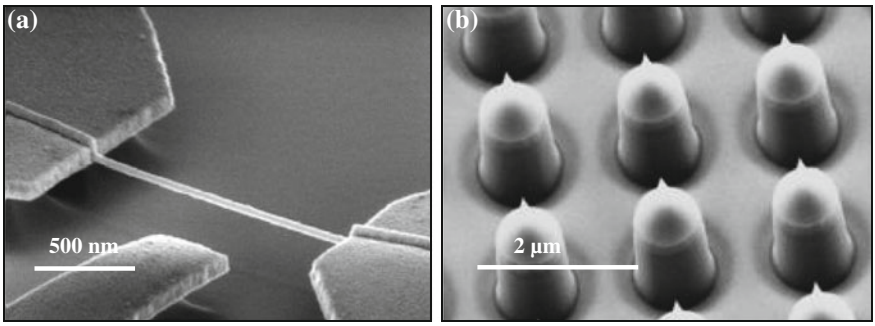


Fig. 27.2 **a** NEMS based ultra sensitive sensor [6] and **b** Protein motors attached to a substrate with precision [7]

Table 27.1 Types of MEMS devices and their description and operating conditions [10]

Type	Description	Examples
I	Devices with no moving parts	Accelerometers, pressure sensors, ink jet heads, etc.
II	Devices with moving parts without rubbing or impacting surfaces	Gyros, comb drives, resonators, RF MEMS
III	Devices with impacting surfaces	Relays, valve pump
IV	Moving parts with impacting and rubbing surfaces	Shutters, scanners, microgears, electrostatic rotors, optical switches

etc. A variety of problems arise from the environment and the contact or design under which these devices function. Hence, material related problems and their failure modes depending on factors like design and operating conditions have always been important for evaluating reliability and performance. MEMS/NEMS devices can be broadly classified into 4 types depending on factors like design and contact mode, as shown in Table 27.1 [10]. Based on this classification, material failure mechanisms can be fracture, creep, tribological issues like stiction, friction, and wear, delamination, and other factors like electromigration, corrosion, pitting, and contamination [8]. In the following section only components operating in intermittent and continuous contact conditions are addressed (Type III & IV) where tribological issues like stiction, friction, and wear are of prime concern.

27.2 Tribological Problems in MEMS/NEMS

A major challenge for MEMS designers is to overcome the effects of stiction. As the name suggests, stiction is the effect where microscopic structures tend to adhere with each other when they come into contact [11]. MEMS engineers employing surface

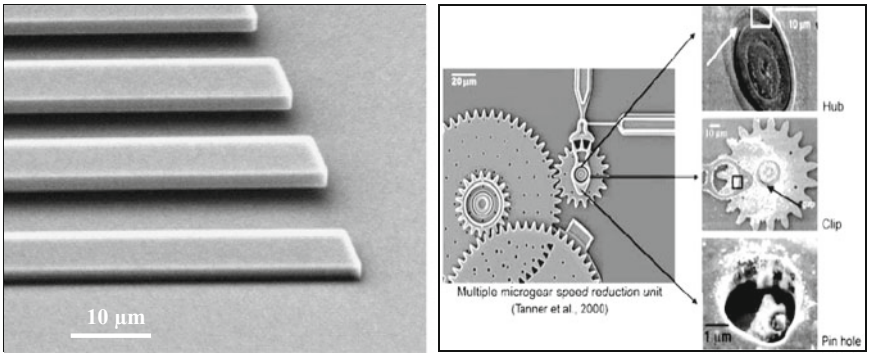


Fig. 27.3 Tribological problems encountered in MEMS: **a** Stiction problem resulting in the collapse of freestanding cantilever structure in a comb drive [15] and **b** Friction and wear problems resulting in severe damage microgears that operate at 2,50,000 rpm [16]

micromachining frequently encounter a fatal stiction effect when they attempt to release the structures in the final step of processing. A practical example of stiction faced in comb drives is shown in Fig. 27.3a: a freestanding cantilever sticks to the substrate due to stiction in the final release step [12]. With the increasing complexity of devices, an in-depth investigation of this phenomenon is essential to counteract it. Stiction is also a leading cause of failure in many MEMS/NEMS applications including accelerometers used in air bag devices in automobiles [13] and digital micromirror devices (DMDs) used in commercial digital light processing (DLP) equipment [14].

The friction force is yet the limiting factor to a successful operation and the missing reliability of MEMS having parts in relative motion to each other. Micromotors, microgears, and microturbines are examples of MEMS that operate in contact mode. For example a typical microgear unit may rotate at a very high speed up to 2,50,000 rpm. The damage resulting from friction and wear between contacting surfaces at various locations of such a microgear after its use is illustrated in Fig. 27.3b. Friction and wear are the dominant degradation mechanisms noticed in microgears [17]. These devices are not yet commercialised due to their poor reliability as a result of friction and wear. Since the advent of the first micromotors in the late 1980s, much work has been dedicated to the reliability of electrostatic motors (Fig. 27.4) but the understanding of friction and wear in such devices is still insufficient. A failure is said to occur when a micromotor or a system no longer performs the required functions under the special conditions within the stated period of time. Electrostatic motors suffer from a variety of tribological problems. The intermittent contact at the rotor–stator interface and physical contact at the hub flange interface result in friction and wear. In a micromotor, the rotor is driven electrostatically in the stator. Adhesive wear and abrasive wear often occur between the rotor and the ground plane. Stiction is a problem often occurring between the rotor and stator, rotor and the substrate.

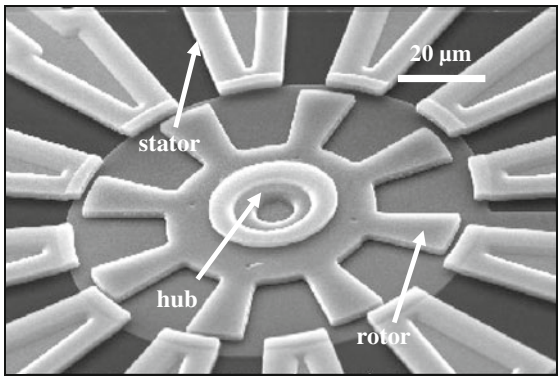


Fig. 27.4 Electrostatic motor [19]

Moreover, a coefficient of friction as high as 0.5–1.1 has been recorded in linear stepper motors [18].

To date, most of the micro-electromechanical system (MEMS) devices are based on silicon technology. This is due to the large technological view accumulated on manipulating, machining, and manufacturing of silicon. Silicon has a poor fracture toughness (0.8 MPa/m$^{-1/2}$) and undergoes a severe wear and degradation in contact MEMS [20]. This has opened a new area of research in tribology namely special coatings that can be used to protect silicon. As the devices range in micrometers, the coatings or the protective layers on silicon should be only a few monolayers or nanometers thick. Notwithstanding that, the tribological characterization of coatings and materials using special techniques and the fundamental investigation of tribological mechanisms in such materials have gained importance. In the following chapter, the various available techniques are discussed, and materials that were developed successfully to mitigate various tribological issues of MEMS/NEMS devices are reviewed.

27.3 Tribological Evaluation of Materials for MEMS/NEMS

27.3.1 Background on Adhesion, Friction and Wear at Nano-/Micro- Scales

Friction and wear are the two system properties that significantly influence material losses in contacting materials. In earlier years, friction was considered to be a dissipation process associated with large machinery and moving components. With the advancements in technology it was realized that friction could be havoc at micro/nano-scale. Wear resistance depends mostly on the mechanical properties of materials whereas friction is a complex phenomenon that depends on parameters

belonging to the material and to the system. Surface, physical, mechanical, and chemical properties influence friction. Friction at high normal forces is dominated by mechanical aspects like plastic deformation, ploughing and chemical aspects like wear debris chemistry, triboreactions, etc. At micro-/nano- Newton normal forces, these properties are of different relevance because the contact pressures and the size of the contact are small. Adhesion force arises from van der Waals forces, and capillary forces acting between two contacting surfaces. Adhesion force depends on surface reactivity, and affinity to water (hydrophobic/hydrophilic behavior) of the interacting surfaces [21]. Adhesion force in turn affects friction at micro/nano-scales. As the size of component decreases down to micrometers and nanometers, surface related properties like roughness, adhesion, capillary, and electrostatic forces play a major role in affecting friction force over mechanical properties unlike in large components [22]. Conventional tribological tests as pin-on-disk or reciprocating fretting equipment, are associated with large quantities of wear that are unrealistic for real engineering applications. They are no more suitable to evaluate materials in microelectronic applications where the contact areas involved are hundreds of nm^2 and contact loads are in μN or mN range [23].

Mechanisms and dynamics of interactions between contacting solids in relative motion, ranging from atomic- to micro- load and length scales must be understood to develop fundamental understanding of adhesion, friction, wear, indentation, and lubrication processes occurring in miniature devices. Another way to understand tribological phenomena is by analytical models that help in predicting the sequence of events. Analytical modeling in tribology is still at its infancy but promising. Efforts are being put to model tribological phenomena by molecular dynamics and finite elements. The biggest challenge in atom-based computational techniques is to simulate friction and wear in traditional engineering applications with respect to both size and time scales. Explicit atomistic simulations on rough surfaces are still intractable, and investigation carried out with interfacial velocities below 1 m/s requires prohibitively long runtimes. The simulation cell size in analytical models is barely large enough to adequately model a single asperity contact [24]. Because of these constraints, experimental methods and appropriate techniques are still of prime interest to evaluate materials at low normal forces and small contact areas as prevailing in MEMS/NEMS components.

27.3.2 Techniques for Tribological Characterization of Materials

Scanning probe microscopy (SPM) has opened a new era of science called nanotechnology. High-resolution microscopes enable literally to see atoms and molecules. Scanning tunnelling microscope (STM) and atomic force microscope (AFM) are broadly used microscopes belonging to this branch of SPMs. With the introduction of the first friction measurement technique using an AFM by Mate et al. [25] in 1985 known as friction force microscopy (FFM) or lateral force microscopy (LFM) using an atomic force microscope (AFM), a new branch of science known as nanotribol-

ogy emerged. This branch aims at studying tribological properties like friction, wear, adhesion, and lubrication at nanocales. AFM simulates a single asperity contact and thereby enables fundamental understanding of tribological phenomena at nanoscales. AFM is also being used for quantifying adhesion, friction, scratch resistance, wear resistance, and nanohardness. Its ability to unveil topographical information with subnanometer resolution makes it perfect surface characterization equipment for materials used in microdevices. Apart from AFM, surface force apparatus (SFA), developed in 1968, is commonly employed to study both static and dynamic properties of molecularly thin films sandwiched between two molecularly smooth surfaces [26]. Recent developments on pin-on-disk and fretting equipments are also used for the tribological characterisation of surfaces and coatings which are operated at low normal force and contact pressures simulating contact pressures prevailing in real applications. Apart from laboratory level tribological tools, characterization tools like cantilever beam array (CBA), inchworm, side-wall friction measurement devices, etc., are also used to evaluate tribological properties of devices [27].

27.3.3 Tribological Evaluation of Materials

27.3.3.1 Quantitative Measurement of Adhesion/Stiction

Adhesion is an important cause of failure in MEMS/NEMS with compliant structural members. When the surface to volume ratios of structures become large, surface forces dominate inertial forces and adhesion becomes important. Adhesion can be classified into 'in-process adhesion' and 'in-use adhesion'. The 'in-process adhesion' refers to the sticking of movable elements typically due to drying after wet chemical etching of the sacrificial layers. The 'in-use adhesion' refers to the collapse of initially free-movable structures during handling or in operation [8]. The adhesion force is the source of stiction. Therefore adhesion by far is the most important parameter that must be minimized to improve the reliability of MEMS/NEMS operating in intermittent or continuous contact modes. A systematic study on adhesion between two surfaces was first done by Bowden and Tabor in 1950 [28]. In a dry environment no adhesive force was measured, but in a humid environment a measurable force holds the two surfaces. Water vapor has a tendency to condense in small cracks and pores to form a thin layer due to capillary condensation. This water generates a force that one encounters when trying to separate two macroscopic wet glass plates. Hence, hydrophilic surfaces suffer badly from capillary forces in humid environments. Also, high free surface energies lead to high adhesive forces. When a wet ball is pressed against a flat surface and the radius of the meniscus 'r' is smaller than the radius of the ball 'R', the pressure inside the liquid is smaller than the atmospheric pressure by the ratio γ_l/R, where γ_l is surface tension of the water. The adhesive capillary force, F_{cap}, is then:

$$F_{cap} = 4R\pi\gamma_l \tag{27.1}$$

A roughening of the ball surface reduces the adhesive force. However, this is applicable only up to a certain level of relative humidity after which adhesion remains the same for both rough and smooth balls due to attainment of saturation. Water at atmospheric pressure is highly responsible for stiction and silicon as well as its oxide are hydrophilic in nature and hence little traces of water on the surface after etch release can lead to stiction. The capillary condensation of water vapor occurs when the surfaces come closer than a characteristic distance z = d_{cap}. z being the separation distance between the surfaces. Maboudian and Howe [12] calculated the energy required to pull the two surfaces apart against the capillary forces between them. The surface interaction energy due to capillary forces, $e_{cap}(z)$, that equals work of adhesion W, can be expressed as:

$$\left\langle e_{cap}(z) = 2\gamma_l \cos\theta \right|_{z \le d_{cap}} \tag{27.2}$$

$$\left\langle e_{cap}(z) = 0 \right|_{z > d_{cap}} \tag{27.3}$$

The function $e(z)$ does not depend on the exact distance as long as the surfaces are closer than d_{cap} and the contact angle θ is a constant for a given surface. Both γ_l and d_{cap} are function of temperature and d_{cap} in turn is also a function of relative humidity. The surface interaction energy can thus be manipulated by changing temperature and relative humidity. The above expressions were deduced assuming surfaces to be flat. Hence calculated forces from these equations may vary for real situations. The van der Waals interaction forces are short-range attractive forces that operate when two surfaces are brought extremely close to each other. The molecular van der Waals forces are caused by mutually induced dipoles in molecules. The interaction energy due to van der Waals forces depends on the medium between the surfaces and the distance between them. The force description becomes complex in a media other than air and vacuum. The energy due to van der Waals, e_{vdW}, forces is expressed as:

$$\left\langle e_{vdW}(z) = \frac{A_{Ham}}{12\pi z^2} \right|_{d_{co} < z < d_{ret}} \tag{27.4}$$

$$\langle e_{vdW}(z) = 0 |_{z > d_{ret}, z < d_{co}} \tag{27.5}$$

with A_{Ham} the Hamaker constant of the molecule. d_{ret} the retarded distance after which force does not act, and d_{co} the cut–off distance close to the molecule below which repulsive forces start to operate. The Hamaker constant depends on the molecular surface termination of the approaching surfaces. The Hamaker constant for most non-polar molecules lies in the range of $0.4–4.10^{-19}$ J [29]. Stifter et al. [30] studied the distance dependence of the meniscus and van der Waals forces to determine their relative importance at various operating conditions. At larger distances, the meniscus force is always stronger than the van der Waals force. Adhesion contributes to friction between surfaces when operated under low normal forces as in many MEMS. Adhesion force acts as an additional force to the applied external normal force and

thereby results in higher friction [31]. This can be mathematically expressed as

$$F_{total} = F_{applied} + F_{adhesion} \tag{27.6}$$

with F_{total} the total normal force, $F_{applied}$ the external applied normal force and $F_{adhesion}$ the adhesion force.

AFM is a commonly used tool to determine adhesion force on surfaces at nanoscale with a force resolution of in the range of pico-N. The sensing probe has a nanometer dimension tip usually made of silicon nitride. The adhesion force is calculated from an approach-retraction cycle between tip and a surface, known as 'force calibration curve'. The adhesion force is calculated by multiplying the maximum vertical deflection of the cantilever with the cantilever stiffness. The measured adhesion force is also known as the pull-off force. Precise knowledge of the cantilever stiffness is thereby essential to obtain reliable results. According to the Johnson Kendall Roberts (JKR) theory [32], the pull-off force, F_a, is related to the work of adhesion as:

$$F_a = \frac{3}{2}\pi\gamma R \tag{27.7}$$

with γ the work of adhesion expressed in J/m^2, and R the radius of the AFM tip. Adhesion measurements with an AFM are a versatile technique because the tip can be coated with virtually any material of interest. Biomolecules, self-assembled monolayers were coated on AFM tips to investigate the interaction forces between various molecules [33]. The adhesion force in general decreases with increasing surface roughness because the true contact area of interaction between the surfaces decreases [34]. Using colloidal probes of different radii and roughness, Yang et al. [35] proved that the adhesion force does not vary linearly with the probe radius contradictory to the JKR theory due to surface roughness. This suggests that a quantitative adhesion calculated from AFM does not take into account surface roughness. It is unlikely that microcomponents possess a single asperity contact (i.e., no surface roughness). De Boer et al. [36] estimated that for planar polysilicon surfaces, the actual contact area is approximately eight orders of magnitude smaller than the apparent area, and individual contacts are tens of micrometers apart. Therefore adhesion measurements by AFM are limited to the fundamental understanding and comparison of different materials. Alternative measurement tools have been developed to take into account surface roughness in quantifying adhesion in real microsystems as MEMS.

Mastrangelo and Hsu [37] developed the cantilever beam array (CBA) technique to measure adhesion taking into account surface roughness. The set-up consists of 20 μm wide array of cantilever beams that extend from 150–1,700 μm in length, and are 2 μm above the surface. The beam lengths are incremented by 50 μm. These beams are brought in contact with the underlying surface by electrostatic actuation or by pushing it down with a probe needle. From a measurement of the length over which the beam is not adhering to the surface, the interaction energy can be calculated. Two modes of cantilever deflection were described in the literature, arc-shaped or S-shaped. S-shaped beam deflections give adhesion measurements

Table 27.2 Adhesion measurements on various materials measured using AFM and cantilever beam array method [39–41]

AFM technique			CBA technique	
Material couple	Adhesion force (nN)	Work of adhesion (mJ/m^2)	Material tested	Work of adhesion (mJ/m^2)
Si_3N_4 tip/Si	50, 80, 52, 33	530	Silicon oxide (SiO_2)	>8
Si_3N_4 tip/SiO_2	35	370		
Si_3N_4 tip/carbon	33	350	Octyltrichorosilane (OTS)	0.012
Si_3N_4 tip/Au	14	148	Octadimethylsilane (ODMS)	0.045
diamond tip/Si	17.4	36.9	Perfluorodecyltri-chlorosilane	0.005
diamond/Carbon	10.9	23.1	(FDTS)	

that are statistically more reliable than arc shaped ones due to a substantial contact area. The contact length of the beams that adhere to the substrate is determined by differential interference contrast microscopy. The surface energy of the S-shaped beam is calculated from (27.8):

$$\Gamma = \frac{3}{8} E \frac{t^3 h^2}{s^4} \tag{27.8}$$

with Γ the surface interaction energy, t thickness of the beam, h the initial separation between the substrate and the cantilever, and s the deflection length of the cantilever. Using the CBA technique, De Boer et al. [38] proved that theoretically calculated surface interaction energies are much higher than experimental values. This was attributed to surface roughness, not included in most adhesion models. Recently, a comprehensive stiction model was proposed by Van Spengen et al. [8] taking into account surface roughness. The theory assumes surface asperities with a Gaussian distribution. However, the applicability of theoretical models to adhesion is still limited and a lot has still to be done to formulate a unified model because surface asperities are rarely Gaussian, and the contact mode can be elastic, plastic, or as in most cases a combined elastic and plastic contact mode depending on the geometry, applied load, and distribution of asperities. Adhesion force and work of adhesion of some MEMS/NEMS materials calculated from AFM pull-off curves and CBA method are summarized in Table 27.2.

Four different adhesion values reported in the literature on Si (001) are mentioned in Table 27.2. That scatter is understandable because most surfaces are non-homogeneous at nanoscale, and since surface preparation can vary from one laboratory to another. Such a scatter in adhesion data obtained with AFM is associated to differences in surface condition, calibration technique used for cantilever stiffness and piezo, etc.

27.3.3.2 Quantification of Friction at Micro-/Nano- Scales

Friction is hard to characterize due to its dependence on chemical, mechanical, environmental and physical aspects of a material couple. Control of friction is however essential for the development of contact MEMS. For example, in-situ friction measurements on a microengine running at 18,300 rpm revealed a coefficient of friction of 0.5 for normal forces less than 4 μN [42]. Such a high coefficient of friction must be minimized to increase the reliability. In nanoscopic contacts under low normal forces, adhesion dominates the friction force. Santer et al. [43] conducted FFM measurements on Au (001) surface with Si tip and measured the nanoscopic frictional forces as a function of the applied normal force. A non-zero friction force was noticed even in the absence of any external applied normal force. This non-zero frictional force (finite coefficient of friction) in the contact arises from adhesion. These observations are in contradiction to the empirical Amontons' law that predicts an infinite coefficient of friction. Carpick et al. observed a finite friction force for negative applied load between a Pt-coated AFM tip and a mica surface [44]. Friction force between a Pt-coated AFM tip and mica in ultrahigh vacuum (UHV) varies with load in proportion to the contact area as predicted by the JKR theory. Above observations revealed that adhesion force in the contact acts as an additional normal force. The contributions to the contact load from adhesion are described in detail by theories like the JKR [32] for short-range adhesion forces between compliant materials, and Derjaguin, Muller, and Toporov (DMT) [45] for long-range forces acting on stiffer materials. The JKR theory was found to fit with the measured friction versus normal force curves at nanoscales. According to the JKR theory, the total friction force, F_f ,in the contact and the modified Hertzian contact radius, a, are:

$$F_f = \mu\left(L + 3\pi R\gamma + \sqrt{6\pi RL\gamma + (3\pi R\gamma)^2}\right) \tag{27.9}$$

$$a^3 = \frac{R}{K}\left\{L + 3\pi R\gamma + \sqrt{6\pi RL\gamma + (3\pi R\gamma)^2}\right\} \tag{27.10}$$

$$\left\langle \begin{matrix} F_f = \mu\,(6\pi R\gamma) \\ a = \left(\frac{6\pi R^2\gamma}{K}\right)^{1/3} \end{matrix} \right|_{L=0} \tag{27.11}$$

with μ the coefficient of friction, L the applied normal force, R the radius of the contacting body, K the effective elastic modulus of the material couple, and γ the adhesion energy per unit area also known as the surface interaction energy. This γ is equivalent to Dupre's energy of adhesion [8]: $\gamma = \gamma_1 + \gamma_2 - \gamma_{12}$. At $\gamma = 0$ then 27.9 reduces to the Hertzian expression. At an external load equal to zero, 27.9 and 27.10 reduce to 27.11. The existence of adhesion was also confirmed at micro-Newton (μN) normal forces by Ando et al. for steel against steel at normal loads ranging from 0.8 to 3,000 μN [46]. The coefficient of friction increased with decreasing normal force but, when the same friction force is divided with the applied normal force plus the adhesion force, a constant coefficient of friction was obtained.

Therefore at low normal forces, the coefficient of friction can be apparent (calculated from the applied normal force alone) and real (calculated from adhesion + applied normal force). Thus adhesion is a significant factor that must be controlled to reduce both stiction and friction force in microcomponents. The true contact area between surfaces affects both adhesion and friction. The true contact area can be defined as the sum of individual microscopic areas at contact points between surfaces. The friction force increases with increasing true contact area. As the true contact area decreases, the interaction zone between two surfaces becomes small and the adhesion reduces. At macroscale, surface asperities are leveled after the first few initial contact events due to high contact pressures. On the other hand at low contact pressures as in the case of microcomponents, the roughness greatly affects adhesion and therefore friction.

The measurement of friction with an AFM is known as lateral/friction force microscopy (LFM/FFM). The force sensor in FFM is a microfabricated cantilever ending with a sharp tip. The tip is brought into repulsive contact with the counterpart. The load is derived from the normal deflection of the cantilever. On sliding, the lateral force acting on the tip is derived from the twist on the cantilever. Bending and twisting are usually measured by sensing with a laser beam reflected on the backside of the cantilever. The applied normal force is derived from the cantilever stiffness. Hence, accurate cantilever stiffness values are desired for reliable friction data. Apart from AFM, surface force apparatus (SFA) is used for characterizing surface interactions at atomic scale. Quartz crystal microbalance (QCM) is another technique used to measure frictional forces between surfaces and thin adsorbed layers [47]. SFA and QCM are suitable for fundamental investigation of tribological phenomena and hence seldom used for materials characterization on a regular basis. Sundararajan et al. [48] successfully determined by AFM the static friction forces encountered in a surface micromachined micromotor. The AFM tip was pushed against a rotor arm of the micromotor so as to generate a lateral deflection of the tip measured by the AFM lateral deflection signal. The maximum lateral deflection obtained prior to the rotor movement is a measure of the static friction of the micromotor. The same technique was used to investigate the effect of humidity and rest time on the static friction force of polysilicon motors using polyfluoropolyether as a solid lubricant.

Using FFM Liu et al. [49] successfully characterized the stiction and friction in a with digital micromirror device (DMD) used in lightweight projection displays. The DMD chip is an array of two million independently controlled reflective aluminum alloy micromirrors (14 μm square and 15 μm pitch) that switch forward and backward at a frequency of 5–7 kHz integrated onto a silicon CMOS static random access memory (SRAM) array. The micromirror rotates as a result of electrostatic attraction between the micromirror structure and the underlying electrodes. A schematic view of two pixels of a DMD is shown in Fig. 27.5. The micromirror is attracted to the side towards the landing site with the largest electrostatic field differential. To release the micromirror from the landing site, a short reset pulse is applied and the bias voltage is removed. The lifetime of micromirrors is a challenge for the commercial success of DMD's. At an operating frequency of 7 kHz, each micromirror element has to switch about 2.5 trillion cycles. Stiction and wear between the spring

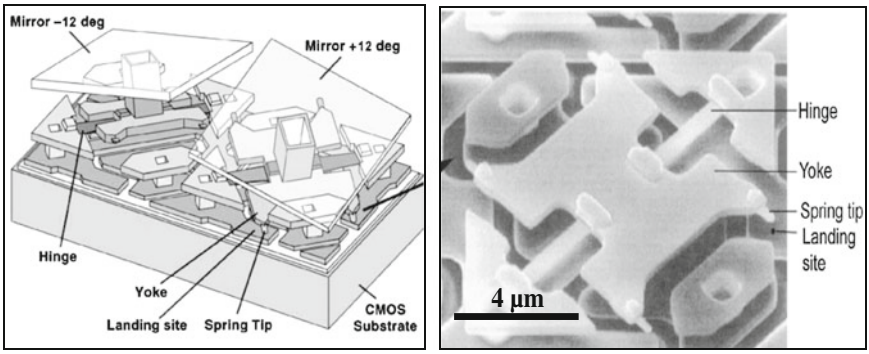


Fig. 27.5 Schematic picture of **a** 2 micromirror pixels and **b** SEM image of a pixel in digital micromirror device (DMD) [49]

tips and the landing sites are important issues that affect its reliability. Liu et al. [49] presented a method to analyze and compare the tribological properties of a defective mirror with normally functioning mirrors. Using AFM adhesion and friction maps, they concluded that the adhesion and friction on the landing site of stuck mirrors (defective) are higher than of properly functioning mirrors. They found a significant influence of capillary forces on stiction between mirrors and the landing sites.

27.3.3.3 Limitations of Tribological Lab Scale Tests

The quantitative data obtained from FFM measurements do not represent data during real applications even though FFM is a good technique for a fundamental study and a comparison of materials. The contact pressures in FFM are high namely in the range of GPa whereas, MEMS have maximum operating pressure of 400 MPa. Single asperity contact in FFM eliminates the effect of roughness and oversimplifies the real contact situation in MEMS. LFM/FFM measurements are extremely sensitive, and a meticulous calibration of the cantilever stiffness and force conversion factor is essential. It was proved by Karuppiah et al. [50] that the same material tested with different lateral force calibration techniques gives rise to different results. Moreover, nanotribology techniques like AFM, QCM and SFA operate at sliding speeds that are too slow compared to the typical micromachine operating speeds. From Fig. 27.6 it appears that the conventional tribological equipments can better simulate the contact pressures and velocities existing in MEMS, but this is not the case. Indeed, large contact size and non-isothermal heat generation unlike in MEMS limit the usage of conventional equipment. Coming to analytical models and atomistic simulations, there is no means to extrapolate AFM data and to simulate tribological events at the dimensions and operating conditions of real devices [27]. On the other hand, atomistic simulations are normally performed at sliding speeds above operating speeds of MEMS devices in the range of thousands of mm/s [51, 52].

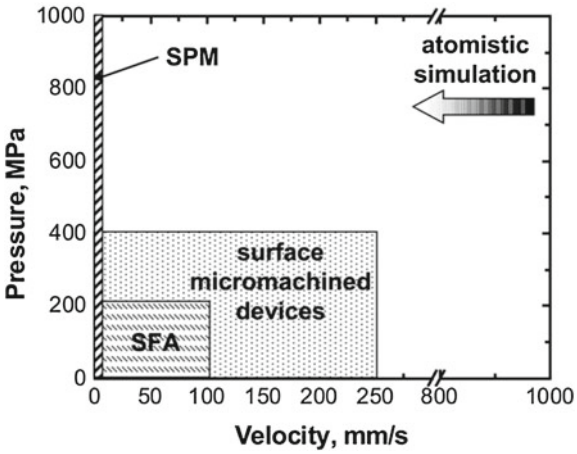


Fig. 27.6 Contact pressure and sliding speed of nanotribological techniques or atomistic simulations compared to surface micromachined devices [27]

Notwithstanding this, new tribological tools are developed using microfabricated devices with a bulk microstructure, surface topography, and a surface chemistry that closely represent real contact situations in MEMS. Lim et al. [53] developed a planar surface friction apparatus for measuring friction at microscales. The device uses an electrostatic comb drive to move a shuttle laterally. The contact occurs between a planar surface and dimples on the bottom of the shuttle. They obtained a very high coefficient of friction of 4.9 ± 1.0 between polysilicon surfaces, and 2.5 ± 1.0 between polysilicon and silicon nitride tested in ambient air. A continuous sidewall tribometer was developed by Senft et al. [54], which allow in-situ measurement of the kinetic friction between MEMS sidewall contacts while operating at realistic speeds for over millions of sliding cycles. The latest addition to characterize friction between micromachined surfaces is an inchworm tribometer developed by de Boer et al. [55]. The device consists of planar polysilicon rails that make contact with a planar track. Parallel plate capacitors with large area can be used to tailor the size of the contact. This instrument can operate under loads in the range of μN, and therefore simulate contact conditions prevailing in microdevices.

27.3.3.4 Evaluation of Wear Resistance by Commonly Used and In-situ Measurements

Materials or coatings used in contact MEMS/NEMS applications must have a good wear resistance apart from low friction. Wear resistant materials form a basis for achieving reliable contact MEMS like microgears, micromotors, etc. In macroscale tribological tests, it is customary to quantify wear loss by measuring the wear volume. A novel way of representing the wear rate is by expressing the wear volume generated per unit dissipated energy which predicts the wear rate of any given material

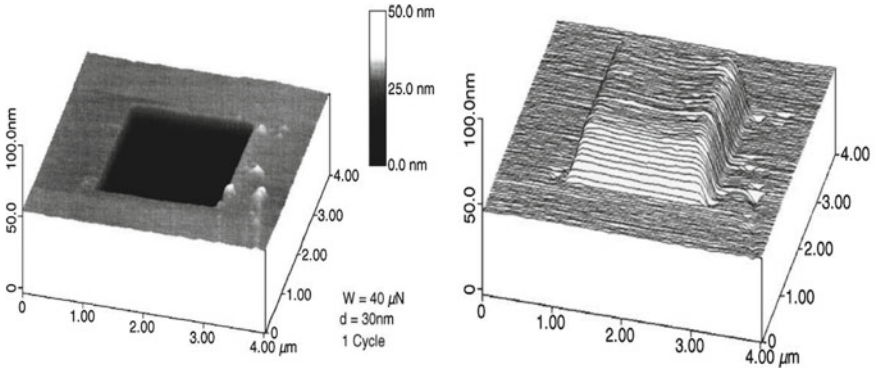


Fig. 27.7 **a** AFM image of wear scar on Si (001) surface after scanning wear test with 30 nm diameter diamond tip at 40 μN normal force after one scan i.e., 1 cycle. **b** Inverted image of the wear track [58]

couple [56]. However at micro-/nano- normal loads, it is hard to accurately determine the wear volume because wear occurs only on few atomic layers. Hence, wear is characterized and quantified using topographical changes. The atomic force microscope (AFM) is undisputedly the best available equipment for an in-depth investigation of wear mechanisms. There are various methods to characterize the wear resistance of a material like measuring the critical load, recording dimensional changes of a reference point with time, measuring the maximum wear depth after certain duration of a test, and time required for complete removal of a coating. AFM is highly used for characterizing wear at micro-/nano-scales. AFM is operated under high normal forces using stiff cantilevers with a diamond tip. The wear tests conducted using an AFM are thus accelerated ones due to high contact pressures. Presently, a new generation of tribometers, like e.g., MUST microtribometer, (Producer Falex Tribology N.V. Belgium) that can operate in milli-/micro- N normal forces are also used for evaluating frictional wear of thin films [57].

In wear tests performed with AFM, the diamond tip of the AFM cantilever is brought into contact with the counter material at a high normal force and the surface is scratched for a given period of time. Afterwards, the profile of the scratch mark is scanned with the same tip but at a light load. The wear depth is then determined using a 2-D profile of the wear scar (Fig. 27.7 [58]). There are two scratching methods: line-scratching (scanning repeatedly over a line) and scanning-scratching (scanning over an area) [59]. Using this technique, Miyake et al. [60] showed that a fluorinated Si-C film has a 30 times better wear durability than conventional amorphous carbon films. Sundararajan et al. [61] used AFM for studying the dependence of wear resistance of DLC films on thickness. They found that the wear resistance of DLC films is directly proportional to their thickness. AFM is now used to characterize the wear resistance of SAM monolayers for MEMS/NEMS applications. Using this technique, it is possible to quantify determine the critical load at which a SAM coating fails.

Machcha [62] proposed another method for measuring accurately the wear rate at nanoscale. This method relies on 20–30 nm deep, micron-sized indentations made on the sliding surface that act as reference points. Changes in the size and depth of these indentations as a function of sliding time are monitored using an AFM. This method was used to quantify wear on a contact recording head consisting of diamond-like carbon pads used in magnetic hard disks. Wear rate of 4.5 nm per week were reported for constant speed test at 7.84 ms^{-1} with 350 μN load and wear coefficient using Archard's law was found to be 0.2×10^{-10}.

The wear resistance of coatings can be measured in terms of the coating lifetime by observing the evolution of friction force with time. When a coated substrate is tested, a sudden change in the friction force indicates coating wear-through. The coating durability is reported in terms of the sliding distance or by total sliding time before wear-through [63]. This method is appropriate for comparing thin films. Conventional sliding equipments like pin-on-disk and fretting equipment are used for this method and operated under low contact pressures and sliding speeds as existing in microdevices. Bhushan et al. [64] successfully used this technique for determining the wear resistance of polyfluoropolyethylene (PFPE) solid lubricants in different environments for MEMS/NEMS applications. Beerschwinger et al. [65] showed the possibility of using conventional pin-on-disk equipment for simulating wear in microdevices. They used specially designed triangular macroscopic specimens with 1.5 mm side length on which circular shaped microstructures were micromachined and coated with different coatings. The contact pressures can be varied from 10^{-6} to 2 MPa depending on the contact area like in MEMS. The wear rate was determined by measuring the decrease in the length of circular projections. Achanta et al. [66] elaborated a method for determining nanowear by first performing a wear test on coatings at low normal forces using a MUST microtribometer and then characterizing wear by observing changes in surface topography using an AFM.

Apart from laboratory tests, in-situ wear measurements are carried out on microdevices directly. For example, in-situ wear measurements on polysilicon micro turbines were carried out by Gabriel et al. [67] that were spun at 10,000 rev/s. The reported typical lifetime of these components was nearly one million cycles at 5,000 rev/sec. Mehregany et al. [68] carried out wear studies on a wobble micromotor by comparing its gear ratio before and after wear test. Indeed, an increase in bearing clearance due to wear will lead to a decrease in gear ratio. Nearly 100 million cycles, equivalent to a rolling distance of 11 km, were achieved with these micromotors. Although in-situ measurements give a direct insight on the performance of a device, such tests are expensive and time-consuming. Therefore laboratory scale tribological tests act as a preliminary materials selection process and further on in-situ measurements can be performed to verify their real time performance.

27.4 Prospective Materials

In recent years, many methods were addressed in order to counteract tribological problems in microdevices. Modifying the surface properties of materials can mitigate adhesion. Common ways to reduce adhesion are: (i) surface roughening or surface texturing, (ii) hydrophobization e.g., with surface coatings, and (iii) special drying techniques after etch release step e.g., critical CO_2 drying. Inducing roughness on a surface by chemical means was one of the solutions used for solving adhesion problems. When a surface is roughened, the true contact area between the contacting surfaces lowers and the surfaces have less chance to interact. Roming et al. [27] reported that surface roughening of silicon oxide resulted in a decrease of the work of adhesion from 20 down to 0.3 mJ/m^2. Houston et al. [69] roughened the silicon using ammonium fluoride and terminated the surface with hydrogen bonds. This surface modification resulted in a work of adhesion less than 0.3 mJ/m^2. A more recent method of surface roughening was adapted from a lotus leaf [70]. A lotus leaf is waxy and consists of micro-bumps (Fig. 27.8a [71]) well dispersed on its surface. If a water droplet falls on its surface, the water drop rolls off because the bumps are much smaller than the droplet and so close together that they cannot puncture the droplet. Scientists are modeling the locations of such bumps so that they can be fabricated onto the surface of a micromachined device to obtain maximum hydrophobicity. Recent experiments on such lotus-like surfaces have shown promising results. Yoon et al. [71] fabricated nano/micro-patterns made of polymethyl methacrylate (PMMA) on a silicon wafer using capillary force lithography (Fig. 27.8b). The patterned samples with certain aspect ratio of the bumps exhibited superior tribological properties at nano and micro-scales when compared to the non-patterned PMMA film.

Surface roughening is not a complete solution for addressing adhesion issues because most of the chemically roughened surfaces react with water vapor and become smooth on a long term [69]. Apart from this, rough surfaces give rise to high local contact pressures thereby leading to wear in the contact area during oper-

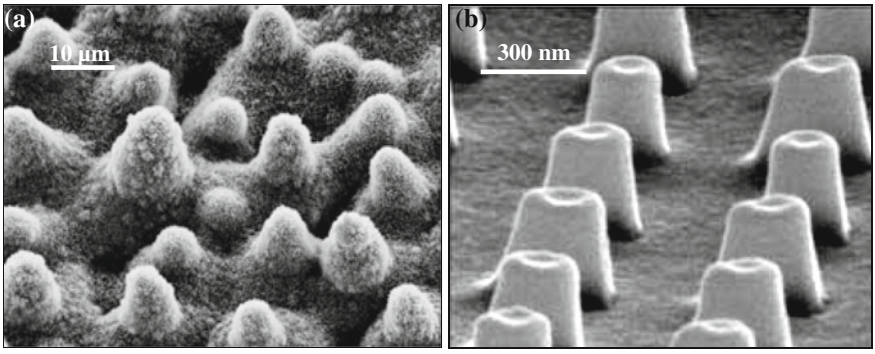


Fig. 27.8 **a** SEM image of a lotus leaf [71] and **b** Artificially grown PMMA SAM bumps on silicon wafer [72]

ation. Hence, it is extremely important to know the optimum roughness for a given material couple operating under certain contact conditions.

Another way to reduce problems due to adhesion, friction, and wear is by surface modification of materials using special coatings or solid lubricants with low surface energy, low friction, and good wear resistance. Surface modification is an extensively investigated area in MEMS technology [27]. The coatings applied are typically a few molecular layer thick and specially selected such that they do not modify the stress state of the polysilicon substrate. An ideal solid lubricant for MEMS/NEMS must be easy to apply, must chemically bond to micro/nano- devices surfaces, and must be insensitive to the environment.

Much of the earlier attention was diverted towards hard ceramic-based materials like silicon nitride (Si_3N_4), silicon carbide (SiC), titanium carbide (TiC), amorphous carbon (a-C:H), diamond-like carbon (DLC) and diamond coatings [73, 74]. Diamond has recently emerged as a promising coating thanks to its highest hardness ~100 GPa and high elastic modulus ~1,100 GPa. DLC coated surfaces sliding against each other have very low coefficient of friction between 0.02 and 0.1 compared to silicon surfaces with coefficient of friction of 0.45–0.6. DLC has shown promising results as a bearing material when used as a coating material on single crystal silicon (SCS) and polysilicon surfaces under large contact stresses. DLC coatings have the disadvantage that their tribological properties are highly dependent on thickness. It was reported by Chen et al. [71] that DLC coatings with a thickness above 15 nm have good tribological properties. Thus DLC coatings can only be applied on components where a large coating thickness can be accommodated. The disadvantage of CVD polycrystalline diamond coating is its high surface roughness (RMS $\sim 1\,\mu$m) [72]. Most of the ceramic-based are deposited using PVD techniques. Intricate parts of a device remain inaccessible for line-of-sight deposition techniques [74]. Some typical tribological results of ceramic coatings are listed in Table. 27.3.

To overcome the problems encountered with hard coatings, research was diverted towards super hydrophobic organic coatings. Typically, these are very thin organic monolayer coatings, either physisorbed or covalently bound to the surface with water contact angles greater than 90°. These films are self-limiting in thickness, and can penetrate through the liquid or vapor phase to deeply hidden interfaces. There are two types of organic molecules commonly applied on solid substrates namely Langmuir-Blodgett (LB) films and self-assembled monolayers (SAM). LB film contains amphilic molecules, i.e., molecules with a hydrophilic head and a hydrophobic tail. SAM contains organic molecules in a dilute solution that spontaneously adsorb and organize onto a solid substrate to form a self-assembled monolayer. LB-films require a flat sample surface, which limits its applicability and SAM on the other hand can be easily deposited irrespectively of the surface structure. Examples of coatings that were successfully applied in commercial MEMS products include a perfluorodecanoic acid coating on structural aluminum of a digital micromirror device (DMD Texas Ins.) and phenylsiloxane coatings for accelerometers (Analog Devices) that resist a charge build-up and also survive packaging temperatures as high as 500 °C. The tribological behavior of self-assembled monolayers is significantly influenced by their functionality and chain length. Friction tests carried out at nano- and microscale

Table 27.3 Hard coatings used for tribological applications in MEMS [74–79]

Material	Adhesion (nN)	Microscale coefficient of friction using AFM/FFM	Macroscale coefficient of friction using pin-on-disk, fretting equipment	Hardness (GPa)
Silicon Si(001)	50–80	0.04–0.07	0.45–0.6	12
Silicon oxide	35	0.087	0.65–1	–
Silicon nitride*	50	0.06	0.66–0.45	30-50
Silicon carbide*	–	0.02, 0.06–0.08	0.20	25
Titanium carbide*	–	–	0.16	35
Diamond-like-carbon**	–	0.02–0.04	0.08–0.1	90(max)
Diamond**	15	0.01–0.02, 0.05, 0.11	0.20	100
Amorphous carbon**	33	0.02–0.14	0.05–0.20	90 (max)

* deposited by PVD techniques, ** deposited by CVD techniques

revealed that SAMs with $-CH_3$ functional group exhibit a lower kinetic coefficient of friction and a superior wear resistance than SAM with $-COOH$ functional groups. Hence proper knowledge on SAM chemistry is essential to obtain good tribological properties [80].

Presently, organic trichlorosilane ($RSiCl_3$) SAM monolayers like octyltrichlorosilane (OTS) and octadecyltrichlorosilane ($C_{18}H_{37}SiCl_3$, ODTS) are examples of SAM coatings that are highly investigated. The work of adhesion on silicon oxide is 8 mJ/m^2 whereas OTS coated silicon has a work of adhesion of 0.012 mJ/m^2. Similarly, the coefficient of friction on OTS was 0.073 compared to 1.1 in the case of SiO_2 [40]. The only disadvantage of OTS is its poor thermal stability above 225 °C whereas in packaging applications the required temperature stability exceeds 400 °C. Therefore, in addition to low surface energy, coatings must be compatible with subsequent device processing, including packaging thermal treatments at 400–500 °C. Perfluorinated alkyltrichlorosilanes ($C_8F_{17}C_2H_4SiCl_3$, FDTS), dichlorodimethylsilane ($(CH_3)_2SiCl_2$, DDMS) are examples of monolayers that possess good thermal stability and maintain hydrophobic behavior without dissociating up to 450 °C. FDTS coated surfaces sliding against each other exhibit a low coefficient of friction at 0.02 compared to OTS coated surfaces with a coefficient of friction at 0.073. However, FDTS causes a larger particulate contamination and release of harmful chemicals than OTS SAM. On the other hand, DDMS has a good thermal stability but surface properties like high water contact angle (103°), high work of adhesion at 0.045 mJ/m^2, and high static coefficient of friction of 0.28, make it inferior to OTS [81]. Bhushan et al. [41] carried out extensive investigation of the tribological properties of fluoroalkylsilane and alkylsilane groups using friction force microscope. They reported that flouroalkylsilanes SAM are superior in terms of wear resistance over alkylsilane SAM. For example, perfluoroalkylsilane (PFTS) has a

Table 27.4 Typical surface properties of on various self assembled monolayer coatings and polymer coatings [40, 41, 81]

Material/coating	Adhesion force (nN)	COF nano/micro-scale	COF macroscale	Water contact angle (°)	Surface free energy using CBA (mJ/m^2)	Side wall coefficient of friction
Si	50–80	0.04–0.07	0.45–0.6	30	–	0.1–0.27
SiO_2	35	0.087	0.65–1	38	>8	1.1
Octyltrichorosilane (OTS)	–	–	0.14	109	0.012	0.073
Octadimethylsilane (ODMS)	26	0.017	0.14	103	0.045	0.28
Octadecylydimethyl-silane (ODDMS)	29	0.018	0.13	103	–	–
Perfluoroalkyltrichloro-silane (FDTS)	–	–	0.12	108	0.005	0.02
PFTS (perfluoroalkylsilane)	19	0.024	0.12	108	–	–
Polyfluoropolyethylene (PFPE) Z-DOL bonded	34	0.04	0.25	97	–	–
Z-15 PFPE lubricant	91	0.09	0.2	52	–	–
Polydimethylsiloxane (PDMS)	37	0.04–0.06	0.2–0.3	105	–	–
Polymethylmetha-crylate (PMMA)	25	0.03–0.05	0.3–0.4	75	–	–

critical failure load of 56 μN compared to 17 μN for ODMS alkylsilane monolayers. More details on adhesion and wear behaviour of different SAMs are given in Table 27.4. SAM coatings suffer from a poor wear resistance since they tend to wear away easily at moderate contact pressures. Each SAM coating has a critical contact load at which it wears off completely [49]. The importance of good wear resistance was illustrated by Liu et al. [49] during LFM/FFM measurements on perfluorodecanoic acid ($CF_3(CF_2)_8COOH$) used in digital micromirrors. Wear on the coating leads to a sudden rise in capillary forces (stiction) and friction, and results in the failure of the micromirror. Polymer coatings like polydimethylsiloxane (PDMS) and polymethylmethacrylate (PMMA) and perfluoropolyether (PFPE) [82] are under investigation as potential candidates in MEMS/NEMS devices (Table 27.4). PFPE is a commonly used lubricant in magnetic disk drive industry to reduce friction and wear at the head–disk interface [82]. Hybrid coatings like PTFE+Si_3N_4 are being considered because these coatings offer a good combination of low friction and high wear resistance [83]. Further, novel coatings with low friction along with good wear resistance must be introduced. Nanostructured coatings are such materials and can be deposited with desired thickness on surfaces. For instance, 'chameleon coating' is a type of coating that contains nanoparticles of hard material like DLC and MoS_x solid lubricant that are embedded in Au matrix. This coating reacts with the environment and forms lubricious layers in the tribocontact and reduces friction. These coatings are now being used in space applications [84].

27.5 Conclusions

Tribological problems are still hindering the future development of contact MEMS/NEMS. Despite the research effort dedicated on reliability of NEMS/MEMS, many tribological issues still remain unsolved. The true contact area between the interacting surfaces plays a major role in affecting adhesion and friction. A good understanding of surface roughness and true contact area between surfaces is necessary to optimize surfaces with less interactions. Most of the theoretical models on adhesion and friction are based on the assumption that surfaces remain flat. However, in reality this is not true and new theories taking into surface roughness are necessary.

Analytical solutions and atomistic simulations in tribology are far from real engineering contact situations. There is no analytic tool yet to extrapolate the data obtained at nanoscale (using AFM or SFA) to any contact size of interest. Further development in this domain is necessary to predict tribological phenomena occurring at nano-/micro- scales.

An improvement in laboratory testing tools is also needed. Tools that can operate at contact pressures, thermal conditions, and sliding speeds corresponding to real application conditions are necessary for an accurate evaluation of materials. Recently, new test equipment has been developed in that respect.

Limitations from a material point of view must be bypassed. It is necessary to either develop new ways of depositing ceramic coatings efficiently or to improve the chemistry of organic layers so that they offer a better wear resistance. Hybrid coatings as multilayers, nanocomposites, nanostructured, etc., are essential that can offer both low friction and good wear resistance.

A major challenge that still has to be tackled is the in-depth understanding of the evolution of material surface properties when two surfaces interact and rub against each other. Tribo(electro)chemical reactions at contacting asperities may generate surface layers with totally different properties. Physical, chemical, electrochemical, and mechanical properties of material change in a tribocontact depending on the contacting conditions and the time in-between successive contact events, and evolve with testing time. As long as that aspect will not be completely understood, the development of optimised materials and coatings for MEMS/NEMS applications will be a trail-and-error approach rather than a scientifically-based one.

References

1. M.A. Huff, *A distributed MEMS processing environment, The MEMS Exchange: A Distributed MEMS Processing Environment* (Corporation for National Research Initiatives Commercialization of Microsystems, MST, Michael Huff, 1998)
2. B. Bhushan, *CRC Handbook of Micro/nanotribology*, 2nd edn. (CRC Press, Boca Raton, 1999)
3. Courtesy of Sandia National Laboratories.
4. B. Stark, *MEMS Reliability Assurance Guidelines for Space Applications* (Jet Propulsion Laboratory, NASA, JPL Publication, Cafifomia, 1999)

5. A. White, *A Review of Some Current Research in Microelectromechanical Systems (MEMS) with Defence Applications* (Weapons systems division aeronautical and maritime research laboratory, DSTO, Australia, 2000)
6. Y. Blanter, Nano-electromechanical systems (NEMS), Work sponsored by FOM (o.a. Physics for Technology), NanoNed and by an EU project (CANEL), T.U. Delft, Nederlands
7. P. Sharke, Eye on the future: nanotechnology: hybrid NEMS. ASME Mech. Eng. Mag. Online (2001)
8. M.W. van Spengen, MEMS Reliability stiction, charging and RF MEMS, Ph.D Dissertation, IMEC-K.U. Leuven, 2005
9. J.W. Judy, Microelectromechanical systems (MEMS): fabrication, design and applications. Smart Mater. Struct. **10**, 1115–1134 (2001)
10. J.A. Walraven, Failure mechanisms in MEMS. Paper presented at ITC internernational test conference, vol. 33, issue 1 (2003), p. 828
11. M.W. van Spengen, MEMS reliability from a failure mechanisms perspective. Microelectron. Reliab. **43**, 1049–1060 (2003)
12. R. Maboudian, R.T. Howe, Critical review: adhesion in surface micromechanical structures. J. Vac. Sci. Technol. B. **15**, 1 (1997)
13. R.E. Sulouff, MEMS opportunities in accelerometers and gyros and the microtribological problems limiting commercialization, in *Tribology Issues and Opportunities in MEMS*, ed. by B. Bhushan (Kluwer Academic Publishers, Dordrecht, Netherlands, 1998), pp. 109–120
14. M.R. Douglass, in *36th IEEE International Reliability Physics Symposium Proceedings* (IEEE, New York, 1998), pp. 9–16
15. M.W. van Spengen, R. Puers, I. De Wolf, A physical model to predict stiction in MEMS. J. Micromech. Microeng. **12**, 702–713 (2002)
16. B. Bhushan, in *Plenary lecture 1st Vienna conference on micro/nano technology*, March 2005
17. S.L. Miller, G. LaVine, M.S. Rodgers, J.J. Sniegowski, J.P. Waters, P.J. McWhorter, Routes to failure in rotating MEMS devices experiencing sliding friction. Proceedings of SPIE Micromachined devices and Components III, vol. 3224 (1997), pp. 24–30
18. X. Lafontan, F. Pressecq, F. Beaudoin, S. Rigo, M. Dardalhon, The advent of MEMS in space. Microelectron. Reliab. **43**, 1061–1083 (2003)
19. Courtesy of TIMA micro and nanosystems group, France
20. N.R. Tas, A.H. Sonnenberg, A.F.M. Sander, M.C. Elwenspoek, in *MEMS. Proceedings of IEEE, Tenth Annual International Workshop on MEMS_97*, vol. 26–30 (1997), pp. 215–220
21. M. Urbakh, J. Klafter, D. Gourdon, J. Israelachvili, The nonlinear nature of friction. Nature **430**, 525–528 (2004)
22. B. Bhushan, *Handbook of Nanotechnology* (Springer, Berlin/Heidelberg/New York, 2004)
23. B. Bhushan, Applications of micro/nanotribology to magnetic storage devices and MEMS. Trib. Int. **28**, 85–96 (1995)
24. D.W. Brenner, Mysteries of friction and wear unfolding: CMS advances the field of tribology. Amptiac Newslettr. **5**, 1–14 (2001)
25. C.M. Mate, Force microscopy studies of the molecular origins of friction and lubrication. IBM J. Res. Dev. **39**, 617–627 (1995)
26. J.N. Israelachvili, Techniques for direct measurements of forces between surfaces in liquids at the atomic scale. Chemtracts Anal. Phys. Chem. **1**, 1 (1989)
27. A.D. Romig, M.T. Dugger, P.J. McWhorter, Materials issues in microelectromechanical devices: science, engineering, manufacturability and reliability. Acta Mater. **51**, 5837–5866 (2003)
28. F. Bowden, D. Tabor, *The Friction and Lubrication of solids* (Clarendon press, Oxford, 1968)
29. N. Tas, T. Sonnenberg, H. Jansen, R. Legtenberg, M. Elwenspoek, Stiction in surface micromachining. J. Micromech. Microeng. **6**, 385 (1996)
30. T. Stifter, O. Marti, B. Bhushan, Phys. Rev. B **62**, 13667–13673 (2000)
31. Y. Ando, J. Ino, Friction and pull-off forces on sub-micron size asperities. Wear **216**, 115–122 (1998)
32. K.L. Johnson, K. Kendall, A.D. Roberts, Proc. R. Soc. London A **324**, 301 (1971)

33. B. Bhushan, in *Self-Assembled Monolayers for Controlling Hydrophobicity and/or Friction and Wear*, ed. by B. Bhushan. Modern Tribology Handbook: Materials Coatings and Industrial Applications, vol. 2 (CRC Press, Boca Raton, FL, 2001) pp. 909–929
34. B.N.J. Persson, E. Tosatti, The effect of surface roughness on the adhesion of elastic solids. J. Chem. Phys. **115**, 5597–5610 (2001)
35. S.-H. Yang, S.M. Hsu, Effect of colloidal probe random surface features on adhesion. Proceedings of World tribology Congress III, Sep 2005
36. M.P. de Boer, J.A. Knapp, T.A. Michalske, U. Srinivasan, R. Maboudian, Adhesion hysteresis of silane coated microcantilevers. Acta Mater. **48**, 4531 (2000)
37. C.H. Mastrangelo, C.H. Hsu, *A Simple Experimental Technique for the Measurement of the Work of Adhesion of Microstructures* (IEEE Solid-State Sensor and Actuator Workshop, New York, USA, 1992)
38. M.P. de Boer, J.A. Knapp, T.M. Mayer, T.A. Michalske, The role of interfacial properties on MEMS performance and reliability. Proc. SPIE **3825**, 2 (1999)
39. J.J. Rha, S.C. Kwon, J.R. Cho, S. Yim, N. Saka, Creation of ultra-low friction and wear surfaces micro-devices using carbon films. Wear **259**, 765–770 (2005)
40. W.R. Ashurst, C. Yau, C. Carraro, R. Maboudian, M.T. Dugger, Dichlorodimethylsilane as an anti-stiction monolayer for MEMS: a comparison to the Octadecyltrichlosilane self-assembled monolayer. J. MEMS **10**, 41–49 (2001)
41. B. Bhushan, T. Kasai, G. Kulik, L. Barbieri, P. Hoffmann, AFM study of perfluoroalkylsilane and alkylsilane selfassembled monolayers for anti-stiction in MEMS/NEMS. Ultramicroscopy **105**, 176–188 (2005)
42. S.L. Miller, J.J. Sniegowski, G. LaVigne, P.J. McWhorter, *Friction in Surface Micromachined Microengines*, Proceedings SPIE Smart Electronics and MEMS, vol. 2722 (San Diego, 1996) pp. 197–204, 28–29 Feb 1996
43. Spaltmann Santer, Adhesion of cleaned nanoscopic metal contacts. Tribotest **9–4**, 333–344 (2003)
44. R.W. Carpick, N. Agrait, D.F. Ogletree, M.Salmeron, Variation of the interfacial shear strength and adhesion of a nanometer-sized contact. Langmuir **12**, 3334–3340 (1996)
45. B.V. Derjaguin, V.M. Muller, Y.P.Toporov, J. Colloids Interf. Sci. **53**, 314 (1975)
46. Y. Ando, Y. Ishikawa, T. Kitahara, Friction characteristics and adhesion force under low normal load. Trans. ASME J. Tribol. **117**, 569–574 (1995)
47. J. Krim, R.A. Widom, Damping of a crystal oscillator by an adsorbed monolayer and its relation to interfacial viscosity. Phys. Rev. B **38**, 12184 (1988)
48. S. Sundararajan, B. Bhushan, Static friction and surface roughness studies of surface micromachined electrostatic micromotors using an atomic force/ friction force microscope. J. Vac. Sci. Technol. A **19**(4), 1777–1785 (2001)
49. H. Liu, B. Bhushan, Nanotribological characterization of digital micromirror devices using an atomic force microscope. Ultramicroscopy **100**, 391–412 (2004)
50. K.S.K. Karuppiah, S. Sundararajan, A comparison of lateral calibration techniques for quantitative friction force microscopy. Proceedings of World Tribology Congress III, sep 2005
51. S. Zhang, G. Wagner, S.N. Medyanik, W.-K. Liu, Y.-H. Yu, Y.-W. Chung, Experimental and molecular dynamics simulation studies of friction behavior of hydrogenated carbon films. Surf. Coat. Technol. **177–178**, 818–823 (2004)
52. H. Wang, Y.-Z. Hu, T. Zhang, Simulations on atomic-scale friction between self-assembled monolayers: phononic energy dissipation. Tribol. Int. **21**, 3926–3932 (2005)
53. M.G. Lim, J.C. Chang, D.P. Schultz, R.T. Howe, R.M. White, Polysilicon microstructures to characterize static friction. Proceedings of IEEE workshop on micro electro mechanical systems, vol. 82 (1990), pp. 11–14
54. D.C. Senft, M.T. Dugger, Friction and wear in surface micromachined tribological test devices. Proc. SPIE- Opt. Eng. **3224**, 31 (1997)
55. M.P. de Boer, D.L. Luck, J.A. Walraven, J.M. Redmond, Characterization of an inchworm actuator fabricated by polysilicon surface micromachining. Proc. SPIE-Opt. Eng. **4558**, 169 (2001)

56. M.Z. Huq, J.-P. Celis, Expressing wear rate in sliding contacts based on dissipated energy. Wear **252**, 375–383 (2002)
57. D. Drees, S. Achanta, J.-P. Celis, Surface testing into the 21st century, Bridging the gap between nano and macro. Paper presented at 1st Vienna international conference on micro- and nanotechnology, Vienna, Austria, 9–11 March 2005
58. B. Bhushan, Nanotribology and nanomechanics. Wear **259**, 1507–1531 (2005)
59. R. Kaneko, T. Miyamoto, Y. Andoh, E. Hamada, Microwear. Thin Solid Films **273**, 105-111 (1996)
60. S. Miyake, R. Kaneko, Thin Solid Films **212**, 256–261 (1992)
61. S. Sundararajan, B. Bhushan, Micro/nanotribology of ultra-thin hard amorphous carbon coating using atomic force /friction force microscopy. Wear **225–229**, 678–689 (1999)
62. A.R. Machcha, An investigation of nanowear in contact recording. Wear **197**, 211–220 (1996)
63. A. Ramalho, J.-P. Celis, High temperature fretting behaviour of plasma vapour deposition TiN coatings. Surf. Coat. Technol **155**, 169–175 (2002)
64. Z. Tao, B. Bhushan, Bonding, degradation, and environmental effects on novel perfluoropolyether lubricants. Wear **259**, 1352–1361 (2005)
65. U. Beerschwinger, T. Albrecht, D. Mathieson, R.L. Reuben, S.J. Yang, Wear at microscopic scales and light loads for MEMS applications. Wear **181–183**, 42–35 (1995)
66. S. Achanta, D. Drees, J.-P. Celis, Friction and nanowear of hard coatings in reciprocating sliding at milli-Newton loads. Wear **259**, 719–729 (2005)
67. K.J. Gabriel, F. Behl, In-situ friction and wear measurements in polysilicon mechanisms. Sens. Actuators **A21–23**, 184 (1990)
68. M. Mehregany, S.D. Senturia, J.H. Lang, Measurements of wear in polysilicon micromotors. IEEE Trans. Electron. Dev. **39**, 1136–1143 (1992)
69. M.R. Houston, R.T. Howe, R. Maboudian, in *Proceedings of Solid-State Sensors and Actuators-Transducers 95*, Stockholm, Sweden (1995), pp. 210
70. E. D. Flinn, Lotus leaf yields slick idea for MEMS. Aerosp. Am. **5**, 25 (2005)
71. K. Glasmästar, S. Sjödin (eds.), *Final Report for the Research Program, Biocompatible Materials, SSF Research Programme: Biocompatible Materials* (Chalmers University of Technology and Göteborg University, Sweden, 2004), pp. 19–39
72. E.-S. Yoon, S.R. Arvind, K. Hosung, B. Kim, D.-H. Suh, K.Y. Jeong, H. Eui, Tribological properties of nano/micro-patterned PMMA surfaces on silicon wafer. Proceedings of world tribology congress III, sep 2005
73. D. Liu, G. Benstetter, E. Lodermeier, Surface roughness, mechanical and tribological properties of ultrathin tetrahedral amorphous carbon coatings from atomic force measurements. Thin Solid Films **436**, 244–249 (2003)
74. R. Huber, N. Singer, Out with the old and in with the new. Mater. Today **191**, 36–43 (2002)
75. A.R. Kraussa, Ultrananocrystalline diamond thin films for MEMS and moving mechanical assembly devices. Diam. Relat. Mater. **10**, 1952–1961 (2001)
76. X. Li, B. Bhushan, Micro/nanomechanical characterization of ceramic films for microdevices. Thin Solid Films **340**, 210–217 (1999)
77. G. Radhakrishnan, R.E. Robertson, P.M. Adams, R.C. Cole, Integrated TiC coatings for moving MEMS. Thin Solid Films **420–421**, 553–564 (2002)
78. S. Sundararajan, B. Bhushan, Micro/nanotribological studies of polysilicon and SiC films for MEMS applications. Wear **217**, 251–261 (1998)
79. D. Liu, G. Benstetter, E. Lodermeier, Surface roughness, mechanical and tribological properties of ultrathin tetrahedral amorphous carbon coatings from atomic force measurements. Thin Solid Films **436**, 244–249 (2003)
80. H.-S. Ahna, P.D. Cuonga, S. Park, Y.-W. Kim, J.-C. Lim, Effect of molecular structure of self-assembled monolayers on their tribological behaviors in nano- and microscales. Wear **255**, 819–825 (2003)
81. R. Maboudian, W.R. Ashurst, C. Carraro, Self-assembled monolayers as anti-stiction coatings for MEMS: characteristics and recent developments. Sens. Actuators **82**, 219–223 (2000)

82. N.S. Tambe, B. Bhushan, Micro/nanotribological characterization of PDMS and PMMA used for BioMEMS/NEMS applications. Ultramicroscopy **105**, 238–247 (2005)
83. X.-C. Lu, B. Shi, L.K.Y. Li, J. Luo, J. Wang, H. Li, Investigation on microtribological behavior of thin films using friction force microscopy. Surf. Coat. Technol. **128–129**, 341–345 (2000)
84. A.A. Voevodin, J.S. Zabinski, Supertough wear-resistant coatings with 'chameleon' surface adaptation. Thin Solid Films **370**, 223–231 (2000)

Chapter 28
Nanotribology in Automotive Industry

Martin Dienwiebel and Matthias Scherge

Abstract Nanotribology in Automotive Industry The chapter describes and reviews nanoscale effects that occur at the sliding interface of macroscopic tribological systems such as a passenger car. It is shown that the wear rates in such a system are in the order of few nanometers per hour and that it is necessary to introduce the third body concept in order to understand running-in phenomena. Moreover, the chapter reviews some simulation results that demonstrate the effect of mechanical mixing. Finally atomic-scale wear studies, performed using Friction Force Microscopy, are discussed.

28.1 Introduction

The automotive industry, although commonly considered to be "old technology" is one of the industrial sectors where nanotechnology is being broadly applied, for example in paintings, windows, tyres and coatings.
In engine development, the importance of small length scales below one micrometer was long underestimated. It is becoming more important as downsizing of the engines continues; The engine that powered the first "automobile", built by Carl Benz in 1885 developed a power of 0.6 kW with a displacement of about 1 liter. Present day passenger car engine have a power density of 60–110 kW/l. As a consequence the contact pressures and the thermal load of the material increased significantly and sliding surfaces in modern engines have to withstand much harsher conditions.
That they actually do is less the result of new materials but was mainly made possible by significant improvements in manufacturing skills. One example of this is

M. Dienwiebel (✉)
Karlsruhe Institute of Technology, Institute for Applied Materials—Reliability of Systems and Components, Microtribology Center μTC, Kaiserstr. 12, 76131 Karlsruhe, Germany
e-mail: martin.dienwiebel@iwm.fraunhofer.de

M. Scherge
Fraunhofer IWM, Microtribology Center, Joseph-von-Fraunhofer Street 6, 76427 Pfinztal, Germany
e-mail: matthias.scherge@iwm.fraunhofer.de

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_28

Table 28.1 Typical average wear rates of engine components. From [1]

Engine component	Wear rate (nm/h)
Piston ring	5–15
Small conrod bearing	Max. 8
Large conrod bearing	2–10
Tappet	10
Cam	5–10

the finishing quality of sliding surfaces. The rms roughness of bearing surfaces is nowadays often well below 100 nm. Hence traditional methods used in engineering to determine roughness, like tactile profilometry are approaching their limit. Therefore new techniques such as optical profilometry or scanning probe methods are recently being adopted from the nano sciences and semiconductor industry also in engine engineering.

28.1.1 Wear and Length Scales

Engine wear can lead to unwanted noise and vibrations, to an increase in emissions and fuel consumption, to a loss of power or eventually to the end-of-life. A rough back-of-the-envelope calculation illustrates that the wear rates in technical systems are in the nanometer regime as well: A conrod bearing will produce unwanted noise if the clearance is larger than approximately 20 μm. If we aim at a service life of the bearing of 200,000 km and assume an average speed of the car of 80 km/h this would demand a wear rate of less than 8 nm/h. Table 28.1 lists typical wear rates of sliding friction systems in an internal combustion engine (given as average values over the contact area).

Wear rates and friction losses of an automotive component are influenced by factors that play on very different length scales from centimeters to a few nanometers. If e.g. due to a design error, the clearance of a crankshaft bearing is too low or to high, processes which lead to a reduction of the friction coefficient and the wear rate during running-in will not take place. Bulk hardness and hardness depth curves as well as e.g. the adhesion properties of coatings are the foundation of the tribo-system. The development of a so-called “third-body” between the tribo-partners is crucial as well for low wear as for a high stability of the system against external disturbances. We will discuss this point in detail in Sect. 28.3. Finally lubricant and additive properties determine friction and wear at the molecular level. This chapter considers some experimental and theoretical studies of friction-induced changes of the near-surface volume, that take place when the wear rates are in the range of nanometers per hour.

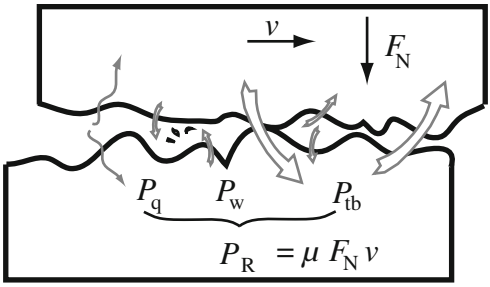


Fig. 28.1 Illustration of the energy dissipation channels in a tribological system. From *left* to *right* the arrows denote energy dissipation by heat, wear and change of material. These three contribution form the largest part of total dissipated power of the system

28.2 Energetic View of Friction and Wear

While atomic scale studies on friction using tools like atomic force microscopy (AFM) or the surface forces apparatus (SFA) usually focus on clean surfaces and model fluids, the technical tribosystem usually consists of "dirty" metal surfaces, formulated oil (often "contaminated" by wear particles and soot), and an ill-defined atmosphere. Because of the complexity that allows many different microscopic processes to occur, it is difficult to generalize mechanisms. Therefore, a very universal approach which is valid at the first hand, is to evaluate the tribological system in terms of the friction power P_F [2] generated during sliding:

$$P_F = \frac{d}{dt} \int \mathbf{F}_F \cdot d\mathbf{s} = \mu \cdot F_N \cdot \mathbf{v}, \tag{28.1}$$

with F_F the friction force, μ the friction coefficient, F_N the normal force and v the sliding speed.

We have to add that a more suitable quantity to describe the stressing of the tribosystem is the friction power density $\rho_F = P_F / V_R$, where V_R is the real volume, in which the frictional energy is dissipated. Unfortunatly, V_R is not measurable during the experiment. It is also not possible to determine it *ex-situ*, since the real area of contact is changing during running (see Sect. 28.4.4).

The energy which is fed into the materials is assumed to be dissipated in three main channels [3]. A part P_q is consumed for the generation of heat and a part P_w for wear particle generation. The remaining friction power P_{tb} will induce changes in the material, leading to the power balance:

$$P_F = P_q + P_w + P_{tb}. \tag{28.2}$$

A similar idea, called the "forced alloy concept" was introduced by G. Martin et al. [4]. He proposed it first to describe phase transformations which occur under irradiation of high energy electrons, during ball milling or during shearing of material. The material is influenced by atomic movements which depend on temperature ("thermal jumps") and atomic movements which depend on external forcing

("ballistic jumps"). In the case of sliding, the magnitude of ballistic jumps will depend on the friction generated in the contact. The competition between thermal and ballistic jumps governs the existence of new phases. One can define an effective temperature $T_{eff} = T(1 + D_{ball}/D_{chem})$, where D_{ball} and D_{chem} are the ballistic and chemical diffusion coefficients. This temperature takes a similar role as the flash temperatures proposed by Blok [5].

28.3 The "Third Body"

The experimental evidence for mechanical intermixing is vast and its existence was already pointed out more than 60 years ago by Bowden and Tabor [6, 7] when they introduced their often quoted friction law $F_F = \tau A_R$.
Here, F_F is the friction, τ is the shear strength and A_R the real area of contact, constituted of many individual microcontacts. They observed that the values of τ which was needed to fit their friction data were close to the bulk shear strength of several bearing materials. Consequently, most of the energy dissipation has to be due to plastic deformation.

Later, Kragelski [8] and also Godet [9] introduced the concept of the "third body". A third body can be a particle in between the sliding surfaces but also a continuous layer whose material properties differ significantly from the first and the second body. Since the early 1980s Rigney and coworkers studied the microstructural changes of the third body of several materials [10, 11]. They observed that in copper, nanocrystalline material is formed near the surface. Theses structures were found to be similar to copper samples that had been created by severe plastic deformation (SPD) [12]. The thickness of the third body was found to differ strongly depending on the materials. Gervé [13] observed that implantation of sliding surfaces by carbon or nitrogen ions lead to the development of a sub-surface material composition similar to what is observed on worn engine parts (see Fig. 28.4).

The process of mixing was also studied using computer simulations. The simulations allow to observe the creation of the third body but are still far to simple to show all effects that might occur in a lubricated tribocontact. Popov et al. performed mesoscale 2D simulations using the method of movable cellular automata (MCA) [14, 15]. The method employs discrete entities with a diameter ranging from 2.5 to 10 nm [15] which have a given density, plastic and elastic properties, yield stress and viscosity of the bulk material and interaction with their neighbors. The MCA simulations showed mechanical mixing and was able to produce depth profiles of the element concentration similar to those found in the cylinder of a combustion engine [16]. Mechanical mixing was also observed in simulations using simple 2D molecular dynamics (MD) calculations by Rigney and Hammerberg [17] and also by Fu et al. [18]. The simulations showed that the friction decreased with running time while the thickness of the mixed layer increased with $t^{-\frac{1}{2}}$. During the mixing the MD simulations show many effects which are known from fluid dynamics, such

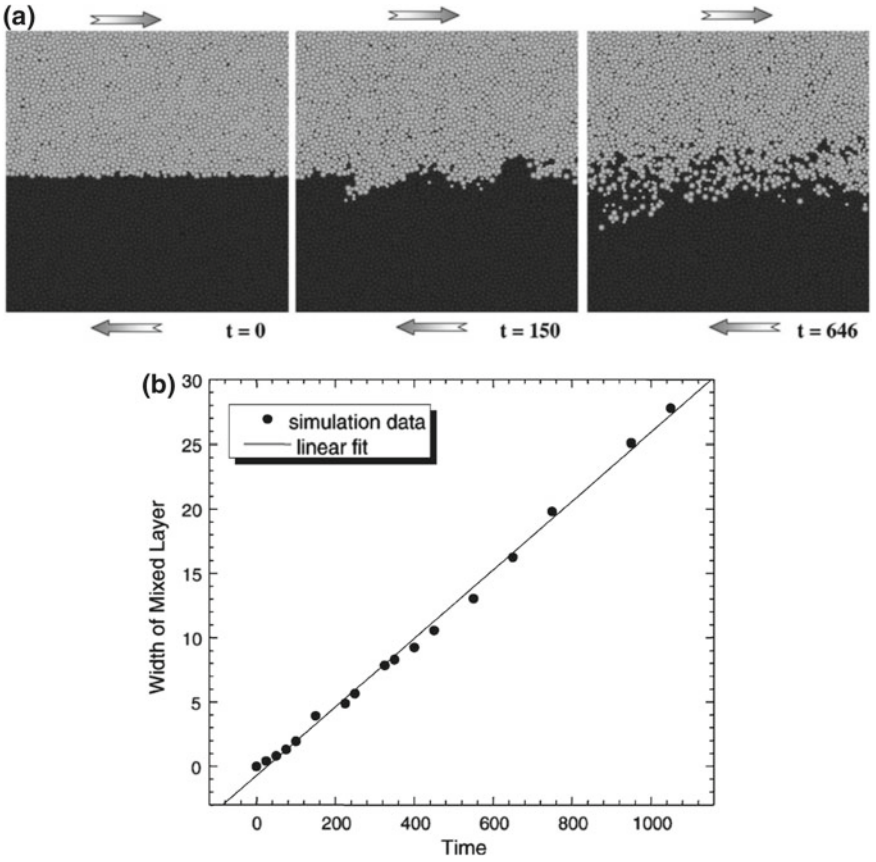


Fig. 28.2 **a** Snapshots from a MD Simulation by Fu et al. showing the mixing of two sliding bodies. **b** growth of the mixed layer thickness versus time. Reprinted from [1] with permission from D. Rigney

as convection cells, laminar but also turbulent flow. Therefore the third body might be envisioned as a viscous fluid which is flowing during the tribological interaction. The analogy with a fluid made Popov assume that wear particles is material of the quasi-liquid layer which is extruded from the contact. Using classical hydrodynamics equations he derived that the wear rate $\dot{W}$ depends quadratically on the thickness of the mixed zone. This would mean that wear should be zero at zero thickness of the third body which is certainly not the case. Therefore this simple assumption might be valid only for a sufficient thickness of the quasi-liquid layer.

28.4 Nanowear

Scherge et al. [19] performed a series of tribometer experiments of a steel (100Cr6) pin resembling a typical piston ring sliding against a gray cast iron disk resembling the cylinder surface of an internal combustion engine. The tribometer was coupled to a high-resolution wear measurement system using the radionuclide technique (RNT). For this technique the pin was labeled with radionuclides and the level of radioactivity in the oil corresponds to the amount of material worn off. The resolution of RNT is usually a few micrograms per liter of oil. Knowing the density of the material and the activity of the sample, the wear can be given as average wear depth over time. A maximum resolution of 0.5 nm/h can be achieved by choosing the appropriate activation parameters. Besides the high resolution, the strength of this method lies in the possibility of measuring the wear continuously during the test.

An example of an RNT wear experiment is shown in Fig. 28.3. The two experiments differed only in the normal force that was applied during the first 6 h. But interestingly the different running-in conditions yielded clearly distinguishable wear rates after even after 200 h of testing. The reason for this is that the difference in the energy dissipation leads to variations of the third body, whose influence is observable even after many hundreds of hours of running time. These variations can be observed in the composition, the topography and the microstructure of the third body as we will show in the remainder of this chapter.

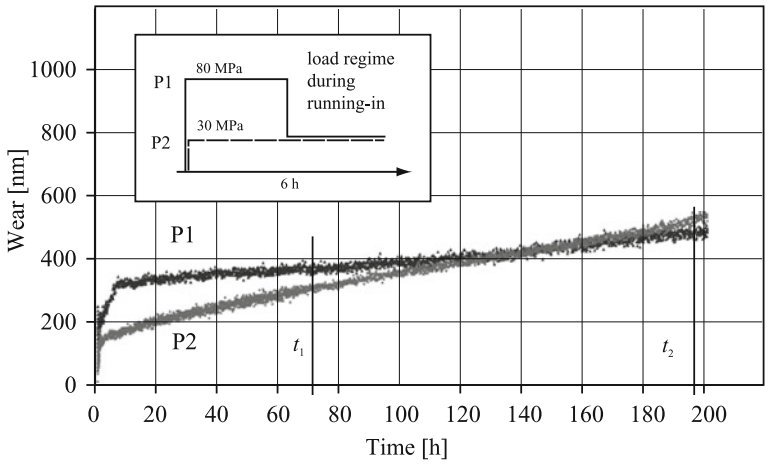


Fig. 28.3 Wear as a function of time for two different running-in procedures. The inset shows the running-in regime. The high load was maintained over a period of 6 h. Then both systems were run at 30 MPa at a sliding velocity of 5 m/s. From [19]

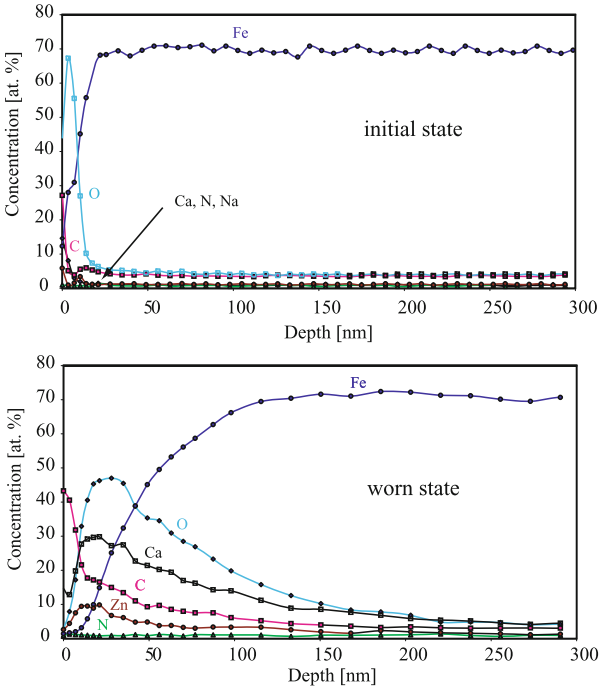


Fig. 28.4 Depth profile by Auger electron spectroscopy of gray cast iron. *Upper graph* initial state; *lower graph* worn state

28.4.1 Composition of the Near-surface Material

Changes in the elemental composition are shown in Fig. 28.4, which were obtained by Auger electron spectroscopy (AES) depth profiling. The upper graph represents the initial state. Due to finishing, the near-surface area has already been changed. The detected oxygen points to oxidation processes. Carbon, zinc and calcium originate from the cooling fluid during machining. At a depth of approx. 20 nm the bulk concentrations are reached. After tribological interaction the distribution of elements changed considerably. The modified zone now exhibits an extension of more than 150 nm. In addition we find a high degree concetration of calcium, zinc and nitrogen. These elements originate from the additive package of the engine oil. Antiwear additive molecules therefore do not only adhere at metal surfaces to form protective layers, moreover they influence friction and wear by becoming an integral part of the near-surface zone (see also Sect. 28.4.4). The change of chemical composition of the near-surface volume can be considered as a measure for the degree of mechanical intermixing.

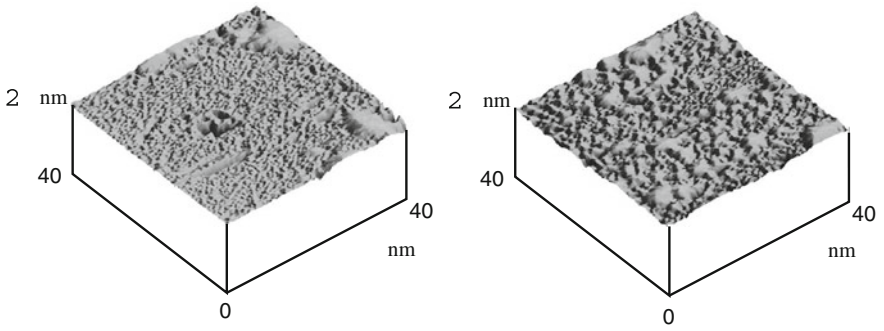


Fig. 28.5 AFM images of two worn gray cast iron surfaces. The normal force was 30 N (*left hand side*) and 90 N (*right hand side*). From [19]

28.4.2 Friction- and Wear-Induced Changes of the Surface

As the near-surface volume changes, the surface itself is also subject to permanent change. Very often surfaces in tribological interaction develop a wave-like microstructure as shown in Fig. 28.5. The wavelength of the microstructure was found to correspond to the applied load meaning that the long-wave portions in the profile of the surface increased with load. One might imagine tribological interaction at contacting surfaces like the impacts of water drops on a sandy beach. A wave-like topography emerges as a result, shown in Fig. 28.5. The size of the impact crater corresponds to the local contact pressure in this very mechanistic sketch. The higher the pressure is, the larger are the waves.

28.4.3 Structural Changes of the Mixed Zone

The focused ion beam (FIB) technique is an ideal tool to study the grain structure of the mixed zone. Cross-section preparation by ion beam milling minimizes the impact of the sample preparation on the grain structure. The imaging capability of the FIB allows the use of either the secondary electrons or the secondary ions for image formation. Figure 28.6 shows cross-sections perpendicular to the surface of three brass alloy samples. Brass is used e.g. for piston pin bushings. The images were recorded using ion channeling contrast, which is a function of the grain orientation. Figure 28.6a shows a sample before the test. In order to reduce the effect of finishing on the microstructure, all samples were prepared by electropolishing. Therefore, the microstructure is nearly undisturbed up to the platinum layer, which was deposited prior to the ion milling to protect part of the surface from beam damage. The samples were tested for 48 h against a stainless steel disk using formulated engine oil as a lubricant. At a contact pressure of 4 MPa, a 170 nm thick zone has developed where

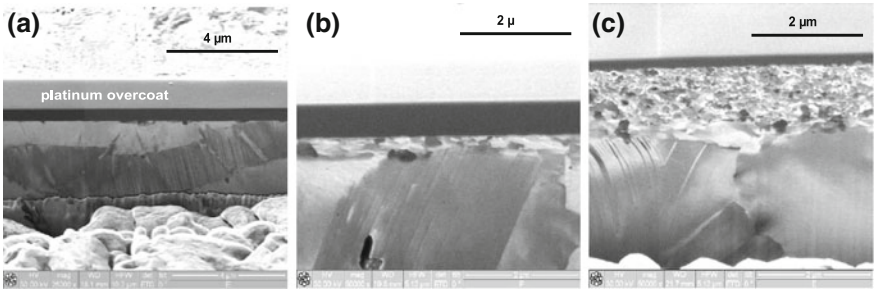


Fig. 28.6 Scanning ion micrographs of brass cross-sections. **a** sample after electropolishing, **b** after tribological stressing with a contact pressure of 4 MPa, **c** after tribological stressing at a contact pressure of 16 MPa

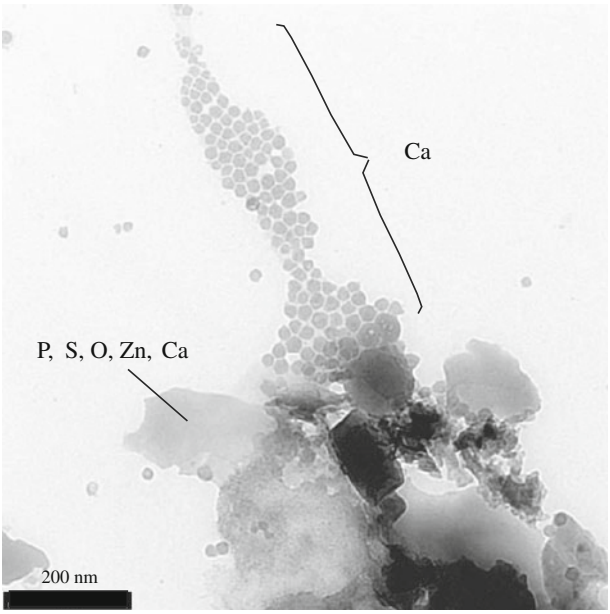


Fig. 28.7 Chemical composition of wear particles at selected locations

the size of the grains are reduced from approximately 1.5 ?m to about 50–150 nm, with the smallest grains found directly underneath the protective platinum overcoat. Interestingly, the transition to the initial microstructure appears to be very sharp. At a contact pressure of 16 MPa the thickness of the modified zone is 780 nm. Directly under the surface the grain size is around 30 nm. The grain size increases slightly with depth to around 400 nm. The thickness of the nanocrystalline zone increases nonlinearly with the dissipated friction power which corresponds with the observed changes of the composition of gray cast iron surfaces.

28.4.4 Wear Debris

Wear particles generated during the tribometer experiments were taken from the oil bath. The particles were extracted from the oil by centrifugation, rinsed in n-heptane and deposited on a copper grid, covered by a very thin carbon film (approximately 5 nm thick) for Transmission Electron Microscopy (TEM) observations. TEM micrographs showed that the particles of different shapes and sizes. A common feature of all particles is their very small thickness, mainly less than 30 nm.

All elements occurring in either the two interacting solids or the oil are present within single wear particles. The element spectrum shows a high content of carbon and oxygen as the result of intense interaction with the lubricant and/or air. P, S and Zn are the fingerprints of additives, especially of antiwear additives, whereas particles with a high concentration of Ca originates from a detergent. The concentration of iron and chromium is extremely low, since the intermixed near-surface zone posesses a low concentration of the original elements (either chromium or iron, see Fig. 28.4).

In 1973 Suh [20] proposed a new theory for wear of metals. The theory focuses on plastic deformation and dislocations at the surface, subsurface cracks and void formation. The subsequent joining of cracks by shear deformation as well as the crack propagation lead to particle generation. The theory predicts flakelike wear particles. In continuous wear measurement using radionuclide technique (RNT) wear curves show a strong increase of total wear during running-in. For optimized running-in conditions [19] the increase in total wear becomes gradually smaller and adopts constant increments (stationary conditions). After the running-in, the wear rate is often low but not zero. Assuming that the interaction of both solids with the additivated oil forms a protective film and that the tribological interaction is confined to that film, then the RNT should not be able to detect any wear, i.e., the wear curve should become horizontal. Interestingly, this effect has never been observed so far. Therefore, the delamination theory applied to the formed additive layer alone does not hold. Only when we assume that, due to delamination, also base material (carrying nuclides) is removed, then low but increasing total wear can be achieved. The TEM images have shown that the majority of the wear particles have a thickness in the range of typically less than 30 nm suggesting that the formed film should not be thicker.

28.4.5 Atomic-Scale Wear Studies

Very few experimental studies on wear phenomena have been carried out on the atomic level. Gnecco et al. [21] used an AFM tip to wear down a KBr crystal in UHV. During the experiment ions were removed from and redeposited on the sides of the wear track. On much smaller scale this is also a mixing process as it is observed in macroscopic experiments, here the mixing process occurs only between KBr atoms because of the large difference in hardness between the Si tip and the

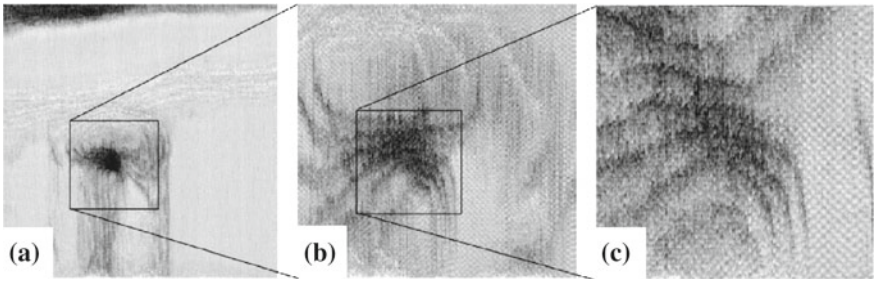


Fig. 28.8 FFM images of a wear groove on KBr after 256 sliding cycles with a normal load of 20.9 nN. Image size **a** 60 nm×60 nm, **b** 25 nm×25 nm, **c** 15 nm×15 nm. From [21]. Reprinted with permission from E. Gnecco.

surface. Repeated scanning over a square area of the surface eventually leads to a wavy surface [22], looking similar to surfaces which were observed by AFM after pin-on-disk experiments. In the macroscopic experiment, processes of removal and redeposition will take place at many asperities in parallel.

28.5 Conclusions

Due to the steady increase of the power density of engines and surface finishing quality, we find that nowadays tribological problems frequently have their origin at the first few hundreds of nanometers of the sliding surfaces. As this trend continues the understanding of wear mechanisms at the nanoscale will become increasingly important in automotive engineering.

A high service life of an internal combustion engine can only be achieved when the wear rates are just a few nanometers per hours. This requires that a third body forms at contacting spots within the first seconds to hours of running. Many experimental and theoretical studies have shown that these plastic processes are alike in a turbulent fluid in which atoms are transported laterally and also perpendicular to the surfaces. This has the consequence that the shape and distribution of asperities is not static but the landscape of asperities is changing constantly during sliding due to plastic flow processes. The composition, the crystal structure and the surface topography are very different from the original material, leading to a reduction of the friction coefficient and the wear rate. The thickness of the third body depends on the friction power dissipated within the contact.

Acknowledgments We are grateful to the contributions to the results shown in Sect. 24.4 by Dmitry Shakhvorostov, Klaus Pöhlmann, Patrick Thomas, Susan Enders, Oliver Kraft and Jean-Michel Martin. The work was made possible by financial support from IAVF Antriebstechnik AG.

References

1. M. Scherge, J.M. Martin, K. Pöhlmann, Wear **260**, 458–461 (2006)
2. G. Fleischer, Schmierungstechnik **16**, 385 (1985)
3. D. Shakhvorostov, K.Pöhlmann, M. Scherge, Wear **257**, 124 (2004)
4. L. Chaffron, Y.L. Bouar, G. Martin, C. R. Acad. Sci. Paris t.2, Serie IV, 749 (2001)
5. H. Blok, Wear **6**, 483 (1963)
6. F.P. Bowden, D. Tabor, Proc. R. Soc. Lond. A **169**, 391 (1938)
7. F.P. Bowden, D. Tabor, J. Appl. Phys. **14**, 141 (1943)
8. I.V. Kragelski, M.N. Dobycin, *Grundlagen der Berechnung von Reibung und Verschleiß*, 1st edn. (VEB-Verlag Technik, Berlin, 1982)
9. M. Godet, Wear **100**, 437 (1984)
10. P. Heilman et al., Wear **91**, 171 (1983)
11. D.A. Rigney, in *Proceedings of Materials Week '97* (ASM International, Indianapolis, Indiana, 1997)
12. D.A. Rigney, L.H. Chen, M.G.S. Naylor, A.R. Rosenfield, Wear **100**, 195 (1984)
13. A. Gervé, Surf. Coat. Technol. **60**, 521 (1993)
14. V.L. Popov, S.G. Psakhie, A. Gervé, Physical Mesomechanics **4**, 73 (2001)
15. V.L. Popov, S.G. Psakhie, A.D. Amde, E. Shilko, Wear **254**, 901 (2003)
16. B. Kehrwald, Ph.D. thesis, University of Karlsruhe (1998)
17. D.A. Rigney, J.E. Hammerberg, MRS Bull. **23**, 32 (1998)
18. X.Y. Fu, M.L. Falk, D.A. Rigney, J. Non-Cryst. Solids **317**, 206 (2003)
19. M. Scherge, D. Shakhvorostov, K. Pöhlmann, Wear **255**, 395 (2003)
20. N.P. Suh, Wear **25**, 111 (1973)
21. E. Gnecco, R. Bennewitz, E. Meyer, Phys. Rev. Lett. **88**, 215501 (2002)
22. A. Socoliuc, E. Gnecco, R. Bennewitz, E. Meyer, Phys. Rev. B **68**, 115416 (2003)

Chapter 29
Adhesion and Friction Contributions to Cell Motility

Mario D'Acunto, Serena Danti and Ovidio Salvetti

Abstract One challenge in biotribology is a complete explanation of cell motility. As the basic unit of life, cells are complex biological systems. Cells must express genetic information to perform their specialized functions: synthesize, modify, sort, store and transport biomolecule, covert different forms of energy, transduce signals, maintain internal structures and respond to external environments. All of these processes involve mechanical, chemical and physical processes. Mechanical forces play a fundamental role in cell migration, where contractile forces are generated within the cell and pull the cell body forward. On the other side, mechanical forces and deformations induce biological response in cells, and many normal and diseased conditions of cells are dependent upon or regulated by their mechanical environment. The effects of applied forces depend on the type of cells and how the forces are applied on, transmitted into, and distributed within cells. Traction forces exerted by cells on substrates can now be determined with a good degree of accuracy, but the intimate relation between cell shape and traction mechanics requires further qualitative investigation. In this chapter, we overview measurements strategies and the models for quantifying adhesion forces and friction developed during cell-cell and cell-substrate interaction during migration.

M. D'Acunto (✉)
Istituto di Struttura Della Materia, Consiglio Nazionale Delle Ricerche (ISM-CNR),
Via Fosso Del Cavaliere 100, 00133 Roma, Italy
e-mail: mario.dacunto@ism.cnr.it

M. D'Acunto · O. Salvetti
Istituto di Scienza E Tecnologie Dell'Informazione, Consiglio Nazionale Delle Ricerche,
ISTI-CNR, Via Moruzzi 1, 56124 Pisa, Italy
e-mail: ovidio.salvetti@isti.cnr.it

S. Danti
Department of Surgical, Medical, Molecular Pathology and Emergency Medicine,
University of Pisa, Via Paradisa 2, 56124 Pisa, Italy
e-mail: s.danti@med.unipi.it

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*,
NanoScience and Technology, DOI 10.1007/978-3-319-10560-4_29

29.1 Introduction

Normal tissue cells are generally not viable when suspended in a fluid and are therefore said to be anchorage dependent. Such cells must adhere to a solid, where the mechanical elasticity of the solid can vary in a wide range, from rigid to soft materials. It has been widely documented that the behavior of some cells on soft materials involves some characteristic phenotypes, for example cell growth on soft agar gels is used to identify cancer cells [1]. An understanding of how tissue cells-including, principally but not only, fibroblasts (most common cells of connective tissue in animals), myocites (muscle cells), neurons, keratocytes (erythrocytes with a blister-like vesicle)—sense matrix stiffness is emerging with quantitative studies of cells adhering to gels, or to other cells, with which elasticity can be tuned to approximate that of tissues [2–5]. Key roles in molecular pathways are played by adhesion complexes and the actin-myosin cytoskeleton, whose contractile forces are transmitted through trans-cellular structures, Fig. 29.1. The feedback of local matrix stiffness on cell state likely has important implications for development, differentiation, disease and regeneration [6, 7].

The foundation of the research of cell motility as a distinct discipline is given by the group of M. Abercrombie in the early '70 decade of the last century [8]. He was the first to divide the motion of fibroblasts into three phases: (i) Extension, (ii) Adhesion and (iii) Contraction, that is the sequence commonly recognized today. In such general framework for the motility, movements with velocity $\sim 1\,\mu m/min^{-1}$ are generated by the extension of flat membrane sheets, lamellipodia, into the direction of movement, Fig. 29.2. However, there exist variations of this scheme. For example, fish keratocytes perform more rapid continuous motion, velocity $\sim 10\,\mu m/min^{-1}$, with a constant shape. They almost seem to glide over the surface and form only transient focal contacts with the substratum with a much shorter lifetime than the focal adhesions formed in fibroblasts [9]. Another example is given by the analogous rapid motion ($\sim 10\,\mu m/min^{-1}$) of the slime mold *Dyctosteliumdiscodeum*, which moves in an amoeboid fashion. During this amoeboid motion non-specific contacts with the substratum are formed and actin stress fibers absent [7, 10]. Anyway, independently by the specific nature of cell, cell motion requires first of all the self-organization of the cell into the an advancing and receding edge. This manifests itself by different molecular concentrations or activation levels of enzymes at the two poles. The polarization can be guided external signals, principally chemical gradients or variation of local mechanical properties of the guide on which the cell is moving. The chapter is devoted to present how adhesive and frictional forces connected to biochemical processes play a key role for cell motility. The chapter is organized as follows: in Sect. 29.2, a general overview about cell motility will be discussed focusing the attention on actin based motility and traction force microscopy topics. In Sect. 29.3, the role of mechanotaxis on contact guidance will be analyzed in a special way for cell-scaffold system motility and correspondent effects due to substrate roughness. Finally, Sect. 29.4 will present details on some recent models on the role of adhesion and friction on cell motility in various contexts.

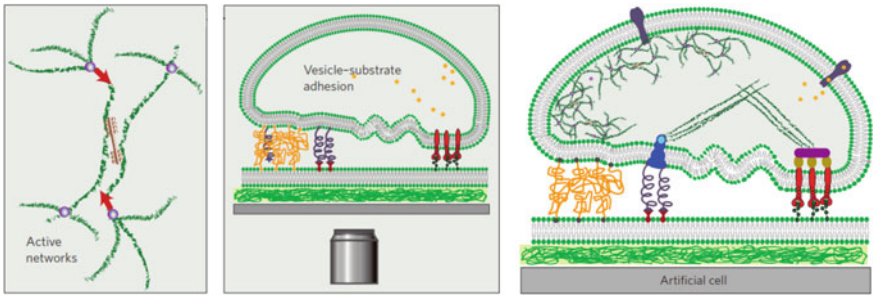


Fig. 29.1 Schematic sketch for cellular structures involved in traction and adhesion. *Left* Active networks composed of actin filaments (*green*) crosslinked passively (*purple*) and actively by motor bundles (*brown*). The motor bundles exert forces on filaments in the direction of the *red arrows*. The elastic properties of these networks are probed with a variety of rheological techniques. *Middle* Phospholipid vesicles interacting with a supported membrane. Both are decorated by functional molecules such as glycolipids (*yellow*), glycoproteins (*purple*) and adhesion proteins (*red*). If the adhesion proteins have counterparts in the opposing membrane, adhesion domains composed of numerous bonds form spontaneously. The contact zone between the two membranes can be observed by a variety of techniques involving microscopy. *Right* The combination of both active networks and vesicles with the addition of coupling proteins, and active control of the whole, will leads to a more realistic model for the mechano-response and the first artificial cells. (Adapted from [14], with permission from Nature)

29.2 Cell Motility: A General Overview

Identifying and testing simple mechanisms of cell motion is an interdisciplinary challenge of interest involving biology, mathematical model physical sciences, with suitable applications to medical and bio-technological sciences. Cells often move within a complex environment such as living tissue. In general, such environments have low Reynolds number, meaning viscous forces dominate over inertial forces. Since inertia can be neglected, motion needs a constant input of energy. Systems driven out of equilibrium by an internal or external energy source can be considered active systems. Examples of active systems are represented by animal flocks [11] and bacterial colonies [12]. Also the cell cytoskeleton is a striking example of such an active system, composed by a network of long semi-flexible filaments made up of protein subunits, interacting with other proteins such as motor proteins, Fig. 29.1. The filaments, such as actin, use the chemical energy of ATP hydrolysis to polymerise and the motor proteins use this chemical energy to exert active stresses that deform the network [13].

Two main different mechanisms for cell motility have now been distinguished: polymerization and contractility due to molecular motors. Specifically, polymerization is referred to polymerization of a filament at one end and depolymerization at the other end. Motion due to polymerization [15, 16] and to contractility [17] have been observed experimentally and studied theoretically [18–20] and numerically [21]. In all the models describing cell motility the essential ingredients for motion are an

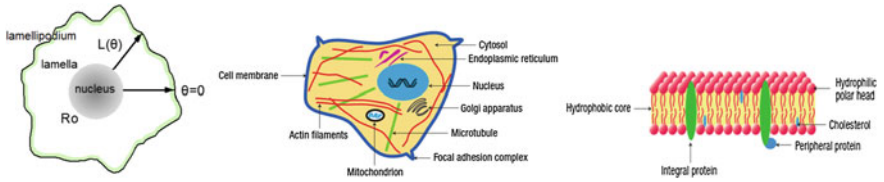


Fig. 29.2 Schematic sketches of cell organs and components. *Left* other the cell nucleus (in *gray*), it is possible to distinguish two zones, one is the lamellipodium which is a narrow peripheral region containing a dense network of interconnected actin filaments, and the other one is the lamella which is a zone located between the lamellipodium and the cell nucleus. *Middle* a plot of cellular components that will be recall during the description of cell motility mechanisms. *Right* molecular composition of cell membrane composed by lipid bilayersembedding transmembrane proteins

energy input to overcome dissipation and sufficient adhesion or friction with a substrate to transfer momentum. The picture of a cell crawling on a surface is as follows: the thin protrusion at the leading edge of the cell (called lamellipodium, Fig. 29.2) builds strong adhesion points with the substrate and pushes forward its membrane by polymerization actin. At the back of the cell, the cell body contracts and breaks the adhesion points. The resulting overall cell velocity is then limited by the actin polymerization rate [22].

29.2.1 Actin Based Motility

Eukaryotic cells and intracellular pathogens such bacteria or viruses utilize the actin polymerization machinery to propel themselves forward [23]. As a consequence, it is very important to identify the general conditions for onset motion and choice of direction that can be the result of a spontaneous symmetry-breaking or can be triggered by external signals and preexisting symmetries. A key feature of cellular motility is the ability of actin to form dense polymeric networks. Such networks produce the forces necessary for propulsion but may also be at the origin of a spontaneous symmetry-breaking. The accurate understanding of the exact role of actin dynamics in cell motility requires multiscale approaches which capture at the same time the polymer network structure and dynamics on the scale of a few nanometers and the macroscopic distribution of elastic stresses on the scale of the whole cell.

Actin is a small globular protein of 42 kDa present in all eukaryotic cells. Under physiological conditions actin monomers (G-actin) polymerize into long helical filaments (F-actin). The filament growth process typically starts with a nucleation process, since actin dimers and trimmers are unstable. Actin monomers at a concentration c may bind to a filament end with a rate $\sim k_+ c$ and dissociate with a rate $\sim k_-$. The concentration $c = k_+/k_-$ associated with the stationary solution is called the critical concentration. In the presence of Mg-ATP (Magnesium-Adenosine triphosphate) the structural difference translates into a difference of the

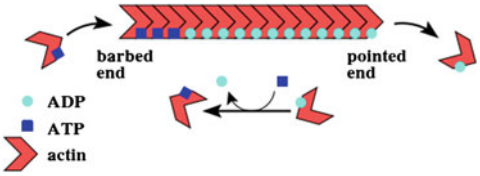


Fig. 29.3 Schematic sketch on how actin subunits treadmill through a filament. ATP-actin polymerizes at the *barbed end*, ATP hydrolyzes and ADP-actin depolymerizes at the *pointed end*. (Reprinted from [13], with permission)

critical concentrations and rate constants between the barbed and pointed ends and causes a treadmilling of subunits through the filament. In Fig. 29.4, it is presented a schematic sketch of actin polymerization [13]. Most actin monomers are bound to Mg-ATP, Fig. 29.3, the critical concentration for this ATP bound species is about 6 times lower for the barbed end than for the pointed end. For ADP (adenosine diphosphate) actin are about the same for both the ends but about ten times higher than for ATP actin at the barbed end. As a consequence, in the steady state the ATP actin concentration is above the critical concentration of the barbed end and below the critical concentration of the pointed end. Polymerized actin subunits are still bound to ATP but as soon ATP hydrolyzes irreversibly into ADP + P and, later on, the anorganic phosphate P dissociates from the filament with a half-time of several minutes. Therefore, ATP actin polymerizes at the barbed end, travels along the filament where by ATP is hydrolyzed and finally ADP actin depolymerizes at the pointed end. Finally, the process of irreversible ATP hydrolysis and P release keeps the system out of equilibrium and allows for a constant flux of monomers through the filament at constant filament length, which forms the base of cellular motility. In Sect. 29.4, we will briefly present some actin based motility quantitative models accounting the complex dynamics involved.

29.2.2 Traction Force Microscopy

In 1995 Oliver et al. and Dembo et al. developed a quantitative technique called traction force microscopy (TFM) to study fibroblast migration on two-dimensional substrate surfaces [24, 25]. While other experimental techniques, such as micropillars and embedded force sensors have made significant contributions in quantifying cell-matrix interaction (see the next section) [26], TFM remains the most widely used methods for measuring cellular traction forces. TFM uses optical phase and wide-field microscopy to track substrate surface displacements due to cellular traction forces through the spatial correlation of fluorescent particles embedded in the substrate. Polyacrylamide gels are among the most commonly used substrate materials in studying cell force responses due to their mechanical tunability, optical translucency and elastic material behavior [27]. By controlling the mole fraction of

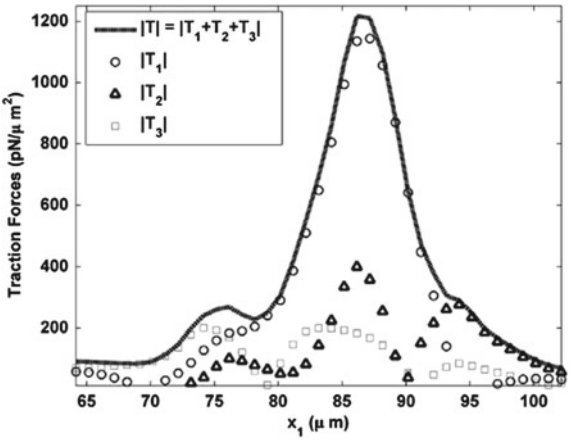


Fig. 29.4 Cell surface traction line profile after fibroblast cell are deposed on a polyacrylamide substrate 35 min. T_1 and T_2 are in plane traction vector, and T_3 is the normal component. (Reprinted from [26], with permission)

added cross-linker to polyacrylamide, the Young's modulus of each polyacrylamide gel can be controlled, with typical moduli ranging from ∼1 to 30 kPa [28]. To record cell surface deformations, cells are initially seeded on the substrate material, after some time, a first image is captured optically checking that cell and tracker particles are recorded simultaneously. Then, cells are detached from the surface through trypsinization or similar treatment. Hence, a second image is captured as a reference configuration. Cell-induced substrate displacements are then determined from the two images by using either a single particle tracking or a digital image correlation algorithm. The resulting gel displacements are converted into traction forces using the inverse Boussinesq formulation, where the Boussinesq theory is used to describe the displacement equilibrium solutions inside a semi-infinite elastic half-space with applied forces as its free boundary [29]. Boussinesq formulation require accurate computational efforts because such theory needs to be utilized inversely to compute cell traction forces, it has the complication that the solutions is no longer unique and the computation itself can be time expensive. Recently, a 3D TFM technique capable of measuring cellular deformations in three dimensions with submicron accuracy [30]. In such method, cell-induced three-dimensional displacement and strain fields are experimentally determined by tracking the motion of submicron fluorescent markers embedded in hydrogels such as polyacrylamide using laser scanning confocal microscopy and digital volume correlation. The traction force components (T_1, T_2, T_3) can be decoupled and quantified, Fig. 29.4.

In addition, 3D TFM method showed that cells can actively modulate matrix stresses in-plane (T_1, T_2) and out of plane (T_3) while exploring their local environment, Fig. 29.5. These results show that the normal traction force component, (T_3) is equally important in cell-matrix interactions during locomotion as the in plane $x - y$, (T_1, T_2) traction force components. In Fig. 29.5, the color contour displays

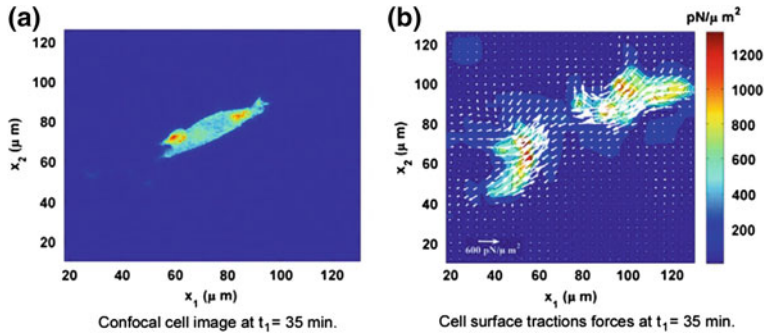


Fig. 29.5 Laser scanning confocal cell image (*left*) and traction force contour (*right*) on the surface during cell migration of a single cell. The cell image on the *left* represent two-dimensional projections of the confocal volumetric data set showing GFP-actin. The cell-applied surface traction force contours display the magnitude of the three-dimensional traction force vector ($|\mathbf{T}|$). The *white arrows* represent the in-plane traction force components (T_1 and T_2) only. (Adapted from [26], with permission)

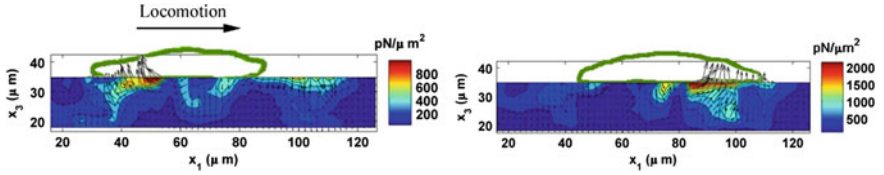


Fig. 29.6 Time evolution of cell induced traction forces as a function of depth (x3) over 70 min along an arbitrary slice below the cell's long axis. The contour plots show the magnitude of the three-dimensional traction force vector ($|\mathbf{T}|$) for a single 3T3 fibroblast measured in pN/μm^2. The *black arrows* represent the in-plane traction forces (T_1) and normal traction force (T_3). The direction of cell migration is from *left* to *right* and the time increment 70 min from the *left*- to *right* image. (Adapted from [26], with permission)

the magnitude of the three-dimensional traction force vector, while the in-plane vectors display the absolute valor of the traction force in plane and normal components ($|T_1|$, $|T_3|$).

The extraction of cell displacements and traction forces in the 3D TFM method as propose by Franck et al., [26], needs some interesting remarks. Digital volume correlation (DVC) is applied to recorded laser scanning confocal microscopy volume images to achieve three-dimensional full field deformation measurement technique, commonly known as digital image correlation. The DVC principle is based on the cross-correlation of individual cubic subsets that comprise each image per given time or deformation increment. In such method, it is usual to work with a function pair, $f(\boldsymbol{x})$, and $g(\boldsymbol{x})$ where $f(\boldsymbol{x})$ refers to the reference configuration and $g(\boldsymbol{x})$ to the deformed or time incremented time. A displacement vector $\mathbf{u}$ can then be calculated by determining the spatial location of maximum peak of the correlation function on the volume configuration V. In order to calculate the traction stresses including

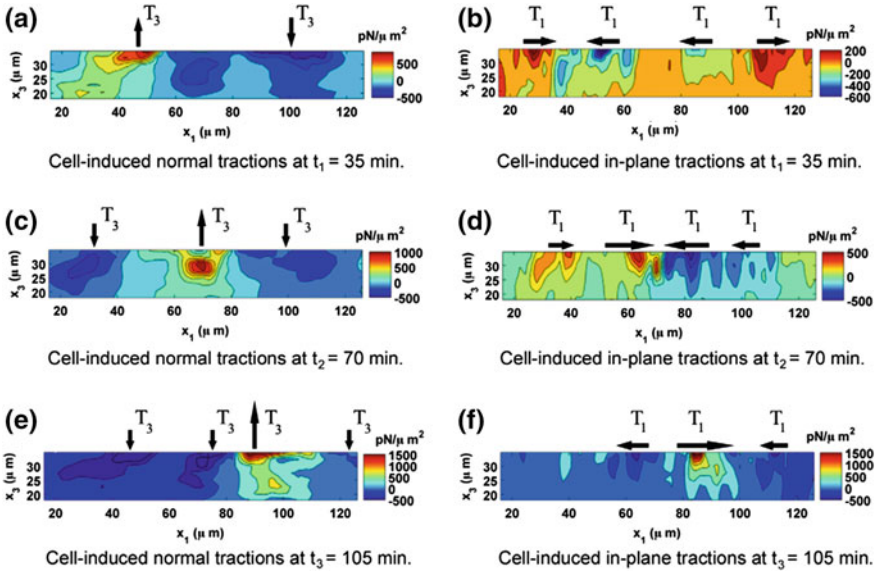


Fig. 29.7 Decomposition of the local 3D cell tractions during locomotion. The contour plots show the magnitude of the shear traction force components. The *black arrows* on the *top* of each plot give the general direction of cell-induced traction forces. Time increment between successive frame is 35 min, and the direction motion is from *left* to *right*. (Reproduced from [26], with permission)

surface tractions, the stress tensor $\boldsymbol{\sigma}$ to be determined first, and is calculated based on the relation $\boldsymbol{\sigma} = 2\,\boldsymbol{\mu\varepsilon}$, where ε is the strain tensor and μ is the shear modulus, which can be related to Young's modulus E and Poisson's ration by $E = 2\mu(1+\upsilon)$. Then the calculation of the traction forces is made using the well-known Cauchy-relation $\mathbf{T} = \boldsymbol{\sigma} \cdot \mathbf{n}$, where $\mathbf{T}$ (components T_1, T_2, T_3) is defined as the three-dimensional traction force vector, and $\mathbf{n}$ (components, n_1, n_2, n_3) is the surface normal of an arbitrary plane on which $\mathbf{T}$ acts. Analogously, the magnitude of the Three-dimensional traction force vector is defined as $|T| = \sqrt{T_1^2 + T_2^2 + T_3^2}$.

29.3 Mechanotaxis and Scaffold Surfaces

In the early 20th century the biologist Peter Weiss observed that cells preferentially orient along ECM-fibers, an organization principle that today is named *contact guidance* [7]. Moreover, Weiss observed that two tissue explants re-organize the collagen gel between them into aligned parallel fiber bundles and that cells leaving the explants migrate and orient along the aligned fibers [31]. Contact guidance therefore could serve both as a cue for organization on cellular scales and as a large-scale organization principle in tissue development by guiding motile cells along aligned ECM-bundles. The term contact guidance has now gained strong topographical connotation, since early '70 decade of the last century, it has been demonstrated that cell

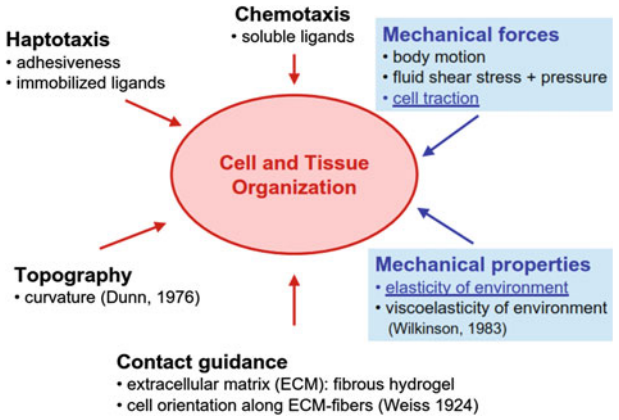


Fig. 29.8 A schematic sketch resuming the self-organization principles of adherent cells. In this section we will focus the attention on the role of mechanical forces on cell motility in different environments. (Adapted from [31], with permission)

react to surface curvature and prefer to align along the axis of minimal curvature, where minimal distortion of the cytoskeleton occurs [32, 33]. This mechanism is supposed to favor orientation of cells along thick fiber bundles. While contact guidance provides only a bidirectional cue for cell migration, a unidirectional cue is supposed to be generated from spatial variations in adhesiveness (motility phenomenon called haptotaxis) [34]. Haptotaxis could be supplemented by specific biochemical information encoded in an inhomogeneous spatial distribution of immobilized ligands. In addition, it has long implied that adherent cells could also respond to mechanical properties of their environment (mechanotaxis) [35–37]. However, for a long time this idea received little attention, on the contrary, during recent years the sophisticated use of elastic substrates has provided strong evidence that cells respond to purely elastic features in their environment, including rigidity, rigidity gradients and strain maps in the environment. Today, it is clearly recognized that mechanical forces play a major role in tissue development, remodeling and reconstruction [38, 39]. Indeed, cells in the body are constantly subject to external forces like those induced by body motion or fluid shear stress. These forces are extremely important to ensure proper functioning for many different tissues including bone, lung and blood capillaries. For example, anchorage-dependent cells like fibroblasts in the connective tissue show a remarkable degree of mechanical activity. Figure 29.8 shows a schematic sketch resuming the self-organization principles of adherent cells. The various properties and taxis component form a complex interaction one each other, making difficult to distinguish them as showed in the plot of Fig. 29.8. However, in some cases it is possible to evidence the incidence on cell motility of one component with respect to others.

Indeed, for example, Wang et al. [36] measured the effects of focal adhesion kinase (FAK) on mechanosensing during fibroblast migration. FAK is a non-receptor

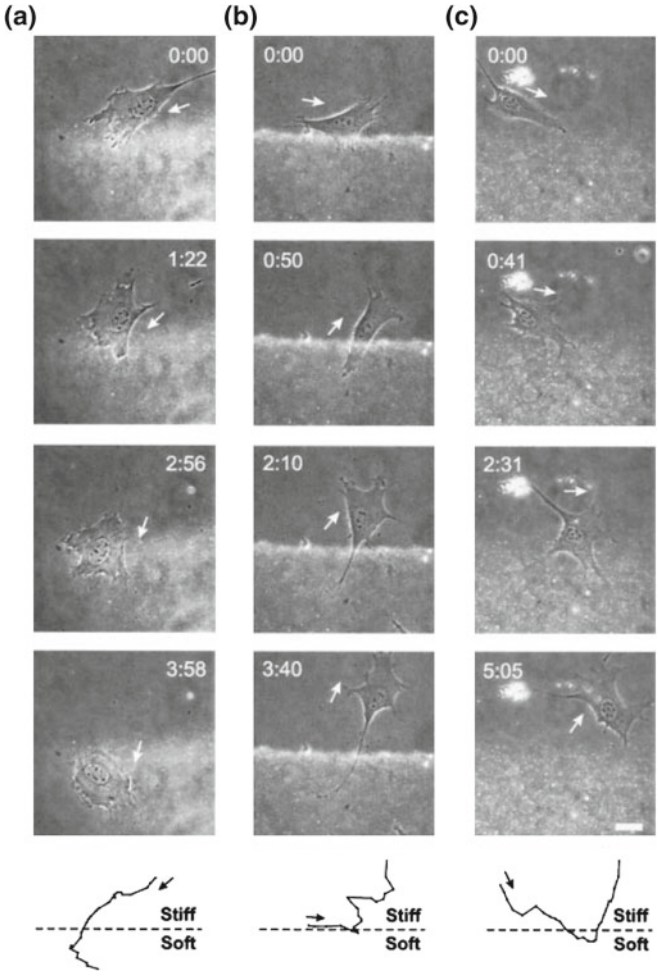


Fig. 29.9 Effects of the FAK on the movements of cellson substrates with a rigidity gradient. A FAK-null cell (**a**) approached the rigidity boundary from the hard side and moved into the soft side of the substrate. In contrast, cells expressing wild-type WT-FAK (**b**) or F397-FAK (**c**) reversed the direction of migration as part of the cell entered the soft substrate. *Bottom* paths of these cells plotted every 10 min over a period of 3 h. Bar = 40 μm. (Adapted fom [36], with permission)

protein tyrosine kinase localized at focal adhesions and is considered to mediate adhesion-stimulated effects. The experiments were performed seeking 3T3 mouse fibroblasts on flexible substrates (polyacrylamide coated with collagen I) so that both mechanical forces and substrate rigidity effects can be evidenced. A synthetic description of results as obtained by Wand et al., are represented in Fig. 29.9.

One biomedical application where cell mobility plays a central role is the tissue engineering [3, 7]. Tissue engineering is aimed at providing the ex vivo cell

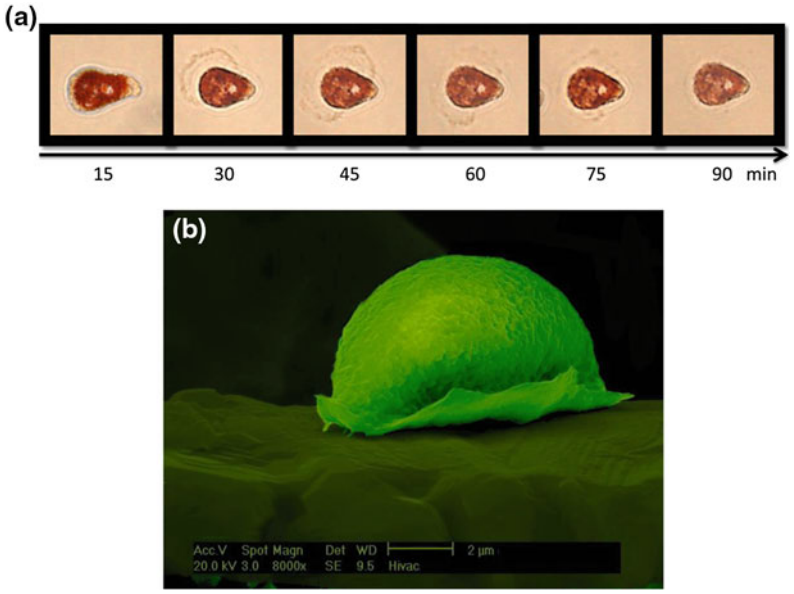


Fig. 29.10 Mesenchymal stem cell (MSC) motion motifs on different surfaces. **a** Light microscopy image of a rat MSC stained with Neutral Red dye on tissue culture polystyrene extending its pseudopodia at different times after seeding. **b** SEM micrograph of a human MSC seeded on a titanium surface. (Unpublished data from the authors)

growth with a third dimension by means of biomaterial-based porous scaffolds, thus replicating the constitutive dimensionality of tissues. Three-dimensional (3D) environments offer mechanical, chemical, architectural and topographic cues influencing fundamental cellular processes, such as adhesion, migration and extracellular matrix (ECM) deposition and remodeling. Indeed, the architectural dimensionality of scaffolds is well-known to affect the cell behavior ex vivo [40]. In 3D contexts, cells can sense the surrounding microenvironment according to cell-perceived dimension and to cell type-specific morphology (e.g., round-shaped, such as the chondrocyte, or elongated morphology, such as the fibroblast) [41]. If pore size is much larger than cell size, single cells will interpret the surrounding surface almost bidimensionally. Anyhow, cells, at a higher hierarchical order (e.g., cell aggregates or tissue-like structures), will be able to cross-talk three-dimensionally, as it occurs in the native tissues. Differently, if pore size is approachable by single cells (e.g., they can organize with a sufficiently stretched morphology), any cell can potentially interact with the biomaterial surfaces in a 3D fashion. After seeding and adhesion, cell migration within the scaffold is driven by typical motility processes, Fig. 29.10. Here, cell protrusion forces are being generated by the cytoskeleton and controlled by integrin receptors [42].

Depending on dimensionality factors, scaffold chemistry, mechanics and architecture, cells can differently interact with ECM molecules present in the

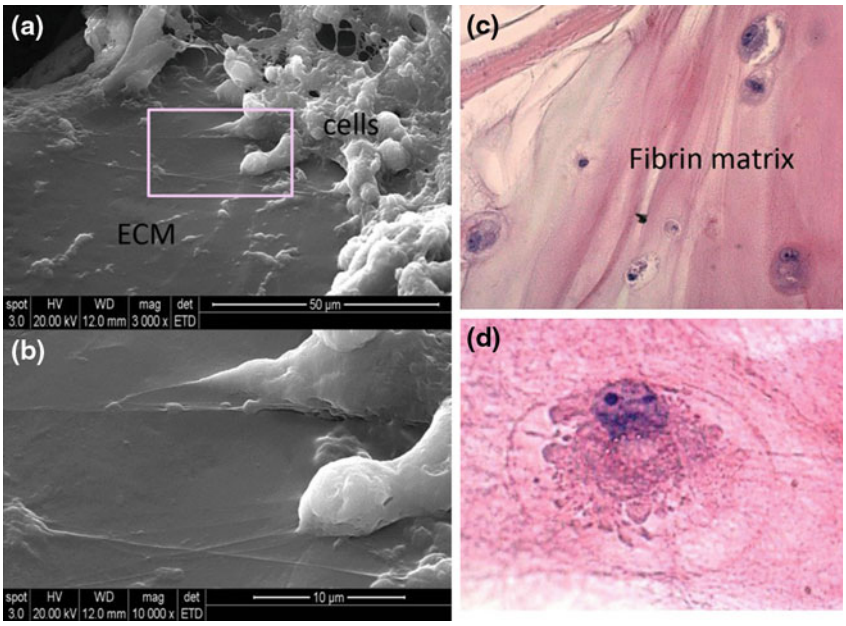


Fig. 29.11 Human MSCs interacting with ECM molecules in 3D contexts. **a**, **b** SEM micrograph of cells migrating on the surface of a titanium scaffold covered by osteoblast-pregenerated ECM. **c**, **d** Light microscopy images of cells embedded in allogenic fibrin showing 3D remodeling (hematoxylin and eosin stain). (Unpublished data from the authors)

microenvironment, the latter being either self-generated or artificially applied, Fig. 29.11. In tissue engineering, cell-ECM interactions have been invoked in guiding pivotal biological phenomena, including the differentiation processes of stem cells for both regenerative and cancer-modeling purposes [43]. Cell adhesion, motion and ECM remodeling are additional processes involving cell-ECM signaling forces that have shown differences between 2D and 3D cultures [44].

In addition, in tissue engineering, mechanotaxis assumes particular relevance using stem cells, as it is often related to the control of their differentiation, which is hypothesized to lead, in the end, to more biomimetic and functional tissue substitutes [45, 46]. Mechano-transduction processes in 3D cultures can be induced by stiffness-gradient substrates as well as by fluidic strain-stress fields generated via bioreactors, Fig. 29.12.

29.3.1 Role of Roughness

Surface roughness or topography is another important factor influencing the cell adhesion, motility and proliferation. Indeed, roughness modulates the biological

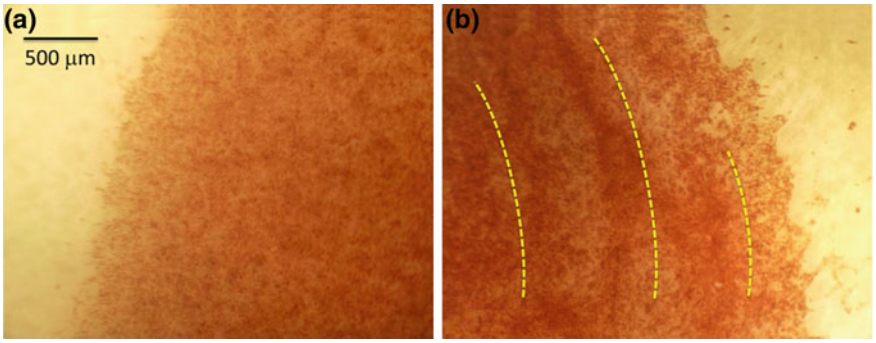


Fig. 29.12 Stereomicroscopy images of human MSCs cultured on electrospun polycaprolactone (PCL) scaffolds stained with Neutral *Red dye*. **a** Sample cultured under static conditions. **b** Sample cultured under dynamic conditions. Viable cells appear concentrated along circumferential lines. (Unpublished data from the authors)

response of tissues in contact with the implant. In the case of scaffolds or ECM, the materials surface roughness has a direct influence in vitro and as well as in vivo on cellular morphology, proliferation and phenotype expression. Literature papers have been reported that cells grown on microrough surfaces, were stimulated towards differentiation; as shown by their gene expression in comparison with cells growing on smooth surfaces. Depending on the scale of mount-valleys distribution of the material surface and on the cell dimension, it is possible to divide surface roughness into macroroughness (100 μm-millimeters), microroughness (100 nm–100 μm) and nanorughness (less than 100 nm), each of such regime with its specific influence [47]. In addition, the response of cells to roughness is different depending on the cell type. For large cells, such as osteoblasts and neurons, macroscopic descriptions of the surface roughness could be reasonable [48–50]. Bartolo et al. investigated neuronal cells behavior on surfaces with nanoscale (less than 10 nm) to mesoscale (200 nm) roughness [49]. The axonal length increased and the neuritis becomes highly branched on the nanoscale rough surfaces (10–50 nm). In the case of higher scale membranes (50–200 nm), the neurons were less developed as demonstrated by the round-shaped soma and poorly branched processes. Therefore, the nanoscale rough membranes seem to be more supportive of neurite outgrowth modulating the development process of the neurons. For smaller cells, such as human vein endothelial cells, increasing surface roughness of biomaterial surfaces at nanometer scale (10–100 nm) could enhance cell adhesion and growth on roughness surfaces [49]. In addition, Kim et al. used the dendrimer-immobilized surfaces to study nanoscale to study nanoscale modifications and discovered that the human mammary epithelial cells cultured on the naked dendrimer surface (4 nm) were abundant in F-actin filaments of peripheral stress fibers and filopodia, compared with those cultured on the plain surface [51]. However, when the surface roughness was larger than 4 nm, such cell stretching was inhibited, resulting in the predominant existence of round-shaped cells. Similar investigation was also reported by Dalby in the development of F-actin filaments in fibroblast [52].

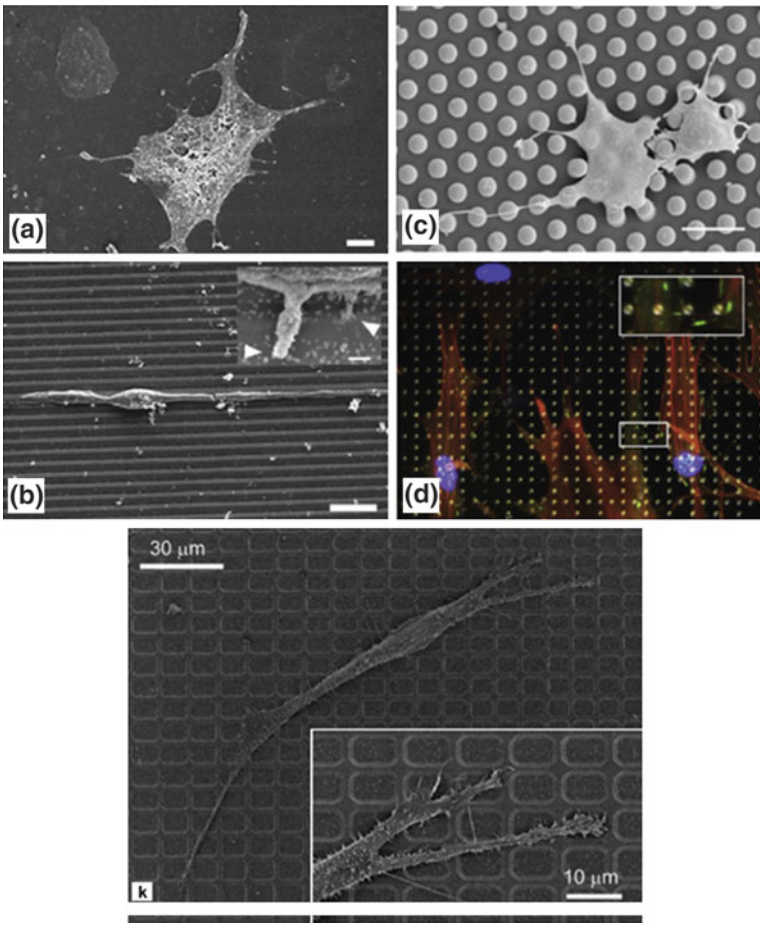


Fig. 29.13 Fabricated topographies for cell adhesion and motility analysis. SEM images of endothelial cells cultures on a flat polyglycerol-sebacate surface (**a**) and a microgrooved surface (**b**) showing cellular alignment on microgrooves. Scale bars, **a** 10 μm, **b** 10 and 1 μm for inset image. **c** SEM image of the induced of the induced branched morphology of fibroblasts on array of PDMS micropillars, scale bar, 20 μm. **d** Human skin fibroblasts on tantalum-coated silicon micropillars showing diffuse actin fibers (*red* fluorescent staining) and fewer Fas (*green*, vinculin fluorescent staining). *k* Directional fibroblast migration on poly(urethane acrylate), rectangular lattice patterns. (All the images are adpated from [47], with permission)

As a consequence, the selectivity of cells on surface roughness could be highly advantage on the development of implanted devices [53, 54]. In Figs. 29.13 and 29.14, some explicit examples on the roughness-induced effects on cells morphology and locomotion are shown.

Cellular adhesion involves the activation and recruitment of proteins named integrins. These receptors bind specifically to motifs located on the ECM molecules (for

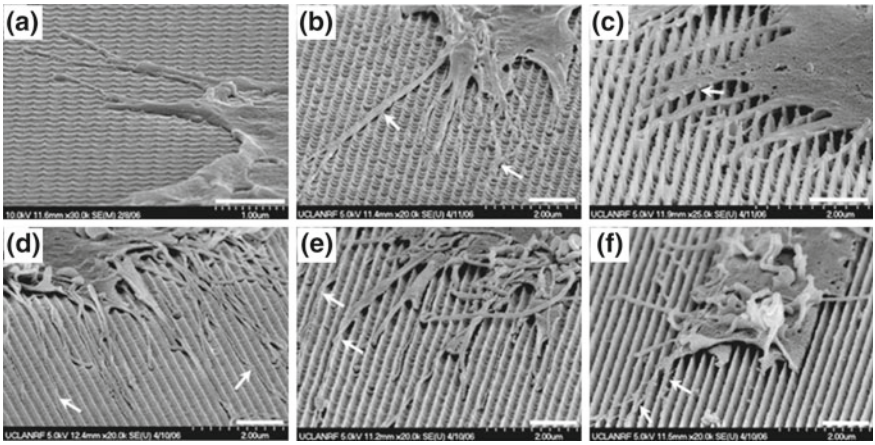


Fig. 29.14 Filopodial interactions with 3D sharp-tip nanotopography. The SEM images of cells filopodia were taken at the culture period of 3 days on nanopost (**a** low, **b** mid, **c** high) and nanograte (**d** low, **e** mid, **f** high) samples. (The scale bar in ech image indicates 1 μm). (Reproduced from [48], with permission from Elsevier)

example, the RGD tripepdide motif found in fibronectin, vitronectin and laminin), see as example Fig. 29.15. It is now commonly recognized that topography plays a fundamental role for the formation of focal adhesion sites. Of particular interest for the community of biotribologists is the spatio-temporal reorganization of the cell cytoskeleton, and of focal adhesion formation in response to nanofeatures [50]. Today, it has been suggested that initiation of adhesive process is dependent on integrin interactions with the substratum and the topographical regulation of cell adhesion, a process that seems to be dependent on the symmetry and spacing as well as the x, y and z dimensions of the topographical nanofeatures. For example, it has be inferred that decreasing the nanofeature spacing to less than 60–70 nm or increasing this distance to the submicron range facilitates integrin clustering, eventually restoring focal adhesion formation, ([50], and there in references), Fig. 29.15.

A series ofAFM-based experiments on the role of cell membrane on cell motility and proliferation were made by D'Acunto et al., [55, 56]. In such experiments, lecithin of soya simulating a bilayer cell membrane was attached to a AFM probe tip and force-distance curves were acquired on a series of tri-block polyurethane copolymers commonly used as scaffolds for tissue engineering. The adhesion forces measured during the force-distance curves were normally more higher on the polyurethane samples were the proliferation testes gave the best fibroblast cells population increment. This result seems to confirm the idea that the role of cell membrane is to stabilize the cell structure during the generation of actin polymerization internal stresses. However, new similar experiments need to be carried out in order to clarify the passive or active role of the cell membrane on cell motility.

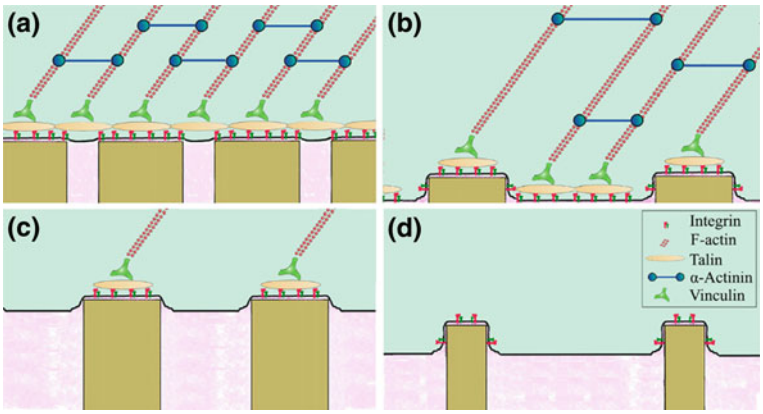


Fig. 29.15 A cartoon describing nanoscale protrusions on focal adhesion formation and reinforcement. **a** Integrin clustering and focal adhesion reinforcement is unaffected on nanoscale protrusions with critical spacing of <70 nm and a protrusion diameter of >70 nm. **b** Increasing the inter-feature spacing to the submicron scale facilities cell-basal substratum interactions below a feature height of ~ 70 nm. **c** Conversely, increasing the feature height restricts integrin binding to the planar basal substrate and restricts focal adhesion formation to the feature apexes. **d** Integrin clustering and cellular adhesion is greatly perturbed on nanoscale protrusion with a feature diameter of <70 nm and inter-feature distance >70 nm. (Reprinted from [50], with permission)

29.4 Adhesion and Friction Models for Cell Motility

Theoretical explanation and quantitative modeling of adhesion and/or friction occurrences for cell motility have been subject of many papers ([7, 57–62], and therein references). Since, many basic mechanisms on the effective role of adhesion and frictional force on the cell motility are still not well defined and understood, here, we limit ourselves to focus the attention on few examples describing more in detail the basic mechanisms proposed for describing the cell motility. After the presentation of a recent model introduced by Recho et al. [59], we have subdivided the section describing some actin-based motility models (Sect. 29.4.1) or describing active gels models (Sects. 29.4.2 and 29.4.3).

Recently, Recho et al. proposed a mechanism for mechanotaxis based on contraction and does not require protrusion [59]. In such model, the initiation of motility is controlled by the average concentration of motor proteins. The increase of motor concentration beyond a particular threshold leads to a bifurcation from a static symmetric regime to an asymmetric traveling wave regime describing a moving cell. In details, the force balance equation in a 1D layer of an active gel in viscous contact with rigid background $\partial_x \sigma = \xi v$, where $\sigma(x,t)$ is the stress, $v(x,t)$ is the velocity and ξ is the friction coefficient. Several authors suggested the following expression for the stress field $\sigma = \eta \partial_x v + \chi c$, where η is the bulk viscosity, c is the concentration of motors and $\chi > 0$ is a contractile stress function. In addition, the function $c(x,t)$ satisfies advection-diffusion equation $\partial_t c + \partial_x (cv) = D\partial_{xx} c$, where D is the

diffusion coefficient. Introducing $l_{\pm}(t)$ as the unknown boundaries of the cell, the cell length is expressed as $L(t) = l_{+}(t) - l_{-}(t)$, and the mechanical boundary condition for the stress is $\sigma(l_{\pm}(t), t) = -k(L - L_0)/L_0$, where k is the effective elastic stiffness and L_0 is the reference length, while the kinematic boundary conditions are in the form $\dot{l}_{\pm} = v(l_{\pm})$ and $\partial_x c(l_{\pm}(t), t) = 0$ which implies that the average concentration, defined as $L_0^{-1} \int_{l_-}^{l_+} c(x, t)dx$, is conserved. The model introduced by Recho et al. is able to calculate the motion of the center of the cell $G(t) = (l_{+}(t) + l_{-}(t))/2$ governed by the following equation [59]

$$\dot{G}(t) = \frac{\Lambda \wp}{2Zc_0} \int_{l_-(t)}^{l_+(t)} \frac{sh\left((G - x)/\sqrt{Z}\right)}{sh\left(L/\left(2\sqrt{Z}\right)\right)} c(x, t)\, dx \tag{29.1}$$

where $\Lambda = k/(\xi \mathrm{D})$, $\wp = c_0 \chi / k$ and $\mathrm{Z} = \eta/(\xi L_0^2)$. From (29.1) is possible to evaluate the maximum cell velocity as $\chi L_0 c_0/(2\eta)$ that gives approximately $10\,\mu\mathrm{m/min}$ that is realistic value. When the cell starts to move, both stress and myosin concentration can be expressed as a function of the moving coordinate $y = x - Vt$, where V is the unknown cell velocity, and $\dot{l}_{\pm} = V$. Under the news variables, the combined effects of stress and myosin concentration can be written in the following manner [59]

$$-Z\frac{d^2 s}{dy^2} + s - \Lambda(L - 1) = \Lambda \wp \frac{\exp(s - Vy)}{\int_0^L \exp(s - Vy)\, dy} \tag{29.2}$$

where $s(y) = \Lambda[\sigma(y) + (L - 1)]$, and the boundary conditions are now $s(0) = s(L) = 0$ and $s'(0) = s'(L) = V$. The initiation of motility is associated with an instability of a static solution of (29.2) with $\mathrm{V} = 0$. Analogously, motile branches with $\mathrm{V} \neq 0$ can bifurcate only from homogeneous static solutions. Figures 29.16 and 29.17 represent the bifurcation diagram as a function of specific parameters.

29.4.1 Actin-Based Motility Models

Actin based motility exhibits a rich spectrum of behavior including spontaneous symmetry-breaking and the interplay of various regime of motion. Two theoretical models dominate current understating of actin-based propulsion: microscopic polymerization ratchet model predicts that growing and writhing actin filament generate forces and movements, while macroscopic elastic propulsion model suggests that deformation and stress of growing actin gel are responsible for the propulsion. Nevertheless, neither of such two models can explain the observed instability of the orientation of the beads. A proposed combination of such two models has been recently proposed [60]. In such combination individual actin filaments undergoing nucleation, elongation, attachment-detachment and capping are embedded into the

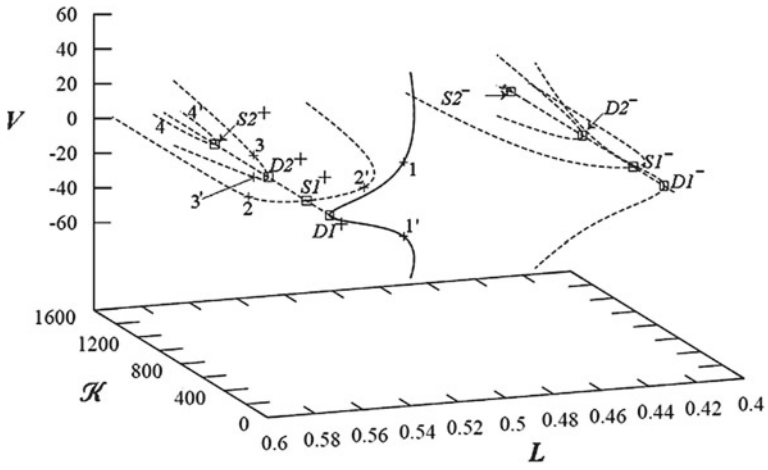


Fig. 29.16 Bifurcation diagram with Λ as a parameter showing non-trivial solutions branching from families of homogeneous static solutions. The value $\wp = 0.245$ is fixed. *Solid lines* show stable motile branches while all the *dotted lines* correspond to unstable solutions. The internal configurations corresponding to branches are indicated by numbers (1, 1′, 2, 2′, etc.). (Adapted from [59], with permission)

boundary of a node-spring viscoelastic network representing the macroscopic actin gel, Fig. 29.18. Such hybrid model is able to capture the observed bistable orientation of the actin-propelled beads described as the combined effects of a macroscopic elastic deformation and microscopic ratchet. In addition, the model suggested by Zhu and Mogilner is able to fit both concave-up and concave-down force-velocity relations for growing actin networks depending on the characteristic time scale and network recoil, confirming that both microscopic ratchets and macroscopic stresses of the deformable actin network are responsible for the force and movement generation. Figures 29.18, 29.19 and 29.20 summarize the basic performances of the model proposed by Zhu and Mogilner [60].

Another recent model for the motility of a bead driven by the polymerization of actin filaments has been proposed [61]. Such model assumes the presence of two populations of filaments: (1) attached filaments that pull the bead, and (2) detached filaments that polymerize and generate a pushing force. The filaments grow in the vicinity of a spherical bead which is propelled by the formation of a filament tail, figure. The dynamics of the both filaments and the bead are treated via the Langevin picture. Briefly, the model considers a bead of radius $\mathbf{R}_0$, at each time step of the calculations the filaments are nucleated uniformly in a ring of radius $\mathbf{R}_0 + \Delta\mathbf{R}$ around the bead with probability $p_g \Delta t$, the points of nucleations are denoted $\mathbf{r}_i$, where the index i is a filament number and at any point $\mathbf{r}_i$ is attached the cross-linked actin network. Filaments which touch the bead are attached to it with probability $p_a \Delta t$, while the other filaments are treated as detached. The elongation of filaments can be described by the following Langevin equation [61]:

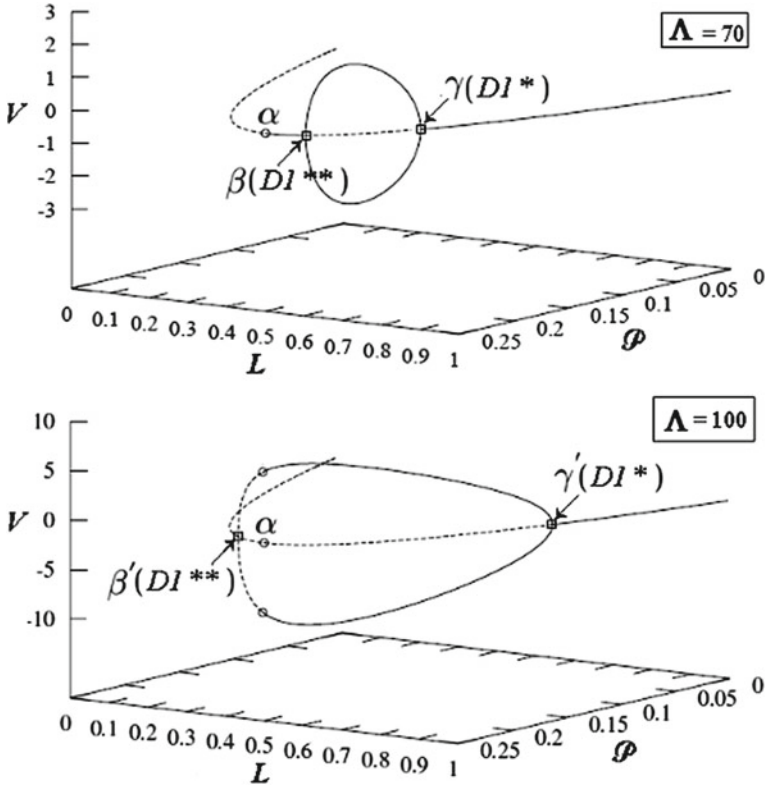


Fig. 29.17 Bifurcation diagram with $\wp$ as a parameter showing motile branches connecting points $D1*$ and $D1**$. Parameter Λ is fixed in each graph ($\Lambda = 70$ and $\Lambda = 100$). (Adapted from [59])

$$\delta \mathbf{r}_i = \left[q_i \dot{\mathbf{R}} + \frac{1}{\gamma} (1 - q_i) \left(\mathbf{f} (\mathbf{r}_i - \mathbf{R}) + \xi_i (t) \right) \theta \left(|\mathbf{r}_i - \mathbf{R}| - R_0 \right) \right] \Delta t \qquad (29.3)$$

where the vectors $\mathbf{R}$ and $\mathbf{r}_i$ define the positions of the bead and the end of the ith filaments on its side close to the bead (barbed end, see Fig. 29.22), respectively; the parameters q_i characterizes the state of an individual filament: $q_i = 1$ for the attached filaments and $q_i = 0$ for the detached ones. $\mathbf{f} (\mathbf{r}_i - \mathbf{R})$ denote the rate of polymerization, ξ is a random force δ-correlated in time accounting the effects of thermal fluctuations on the rate of polymerization, $< \dot{\xi}_i (t) \xi_j (0) >= 2 k_B T \gamma \delta (t) \delta_{ij}$, finally $\theta(z)$ is the Heavised step function. In addition, Filippov et al. assumed that an attached filament is stretched with a velocity which is equal to the velocity of the bead, $\dot{\mathbf{R}}$, while a detached filament polymerizes in the direction towards the centre of the bead and the rate of polymerization, $\mathbf{f} (\mathbf{r}_i - \mathbf{R})$, is described as [61]:

$$\mathbf{f} (|\mathbf{r}_i - \mathbf{R}|) = (\mathbf{r}_i - \mathbf{R}) A \exp \left(- |\mathbf{r}_i - \mathbf{R}|^2 / \lambda^2 \right) \qquad (29.4)$$

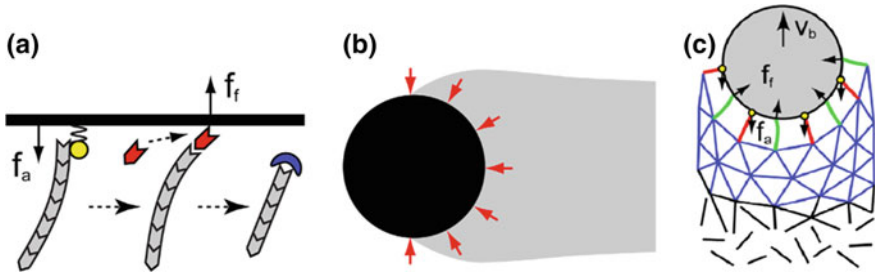


Fig. 29.18 Cartoons describing the actin motility model as proposed by Zhu and Mogilner. **a** Tethere ratchet model. Actin filaments (*gray*) can attach to the obstacle surface (*black line*) via attachment sites (*yellow*) and exert pulling forces (f_a). Detached filaments can elongate by assembling actin monomers (*red*) onto their barbed ends and exert pushing forces (f_f) via a Brownian ratchet mechanism. Detached filaments can be capped by capping proteins (*blue*) and stop interacting with the obstacle. **b** Elastic theory. An elastic actin network (*gray*) propels a curved obstacle (*black*) with squeezing forces (*red arrows*). **c** Hybrid model incorporating both discrete filaments (*green lines* free filaments; *red lines* attached filaments; *yellow circles* attachment sites) and deformable network (*blue lines*),which is treated as a node-spring meshwork. Filaments are created along the surface of the bead (*gray*) and immediately anchor to the network can be ruptured by a high stretching force. The network's nodes, together with connected springs, are removed from the network at a constant rate to represent the disassembly of the network (*black lines*). (Adapted from [60], with permission)

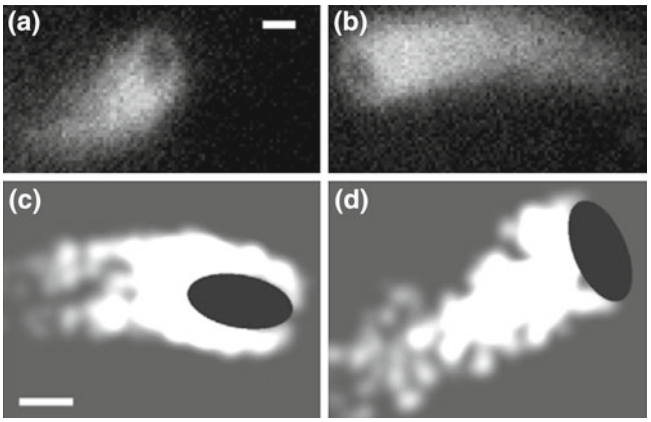


Fig. 29.19 Motion of actin-propelled ellipsoidal beads. **a**–**b** Fluorescent images show actin tails of the motile beads. The dark ellipsoidal shapes at the fronts of the tails illustrate bead's propulsion along its **a** long-axis and **b** short-axis. **c**–**d** Simulation snapshots of the same bead moving along its **c** long-axis and **d** short-axis at different time moments. *Black circle* represents the bead, the white zone denotes the actin networks. All the bars are 1 μm. (Adapted from [60], with permission)

where A is the amplitude of the rate and λ is the characterization of the region around the bead where the polymerization occurs. Each of the attached filaments is detached with probability $p_d^0 \exp(K l_i a / k_B T)$, where p_d^0 is the spontaneous probability of detachment in a stress free state, K is the elasticity constant of a filament, l_i is the

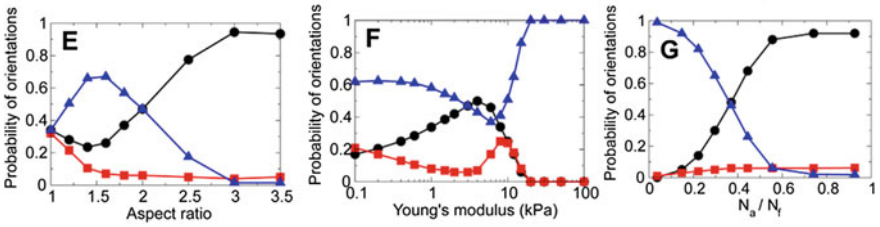


Fig. 29.20 **e–g** Probability distribution of bead's orientation as a function of **e** Bead's aspect ratio, **f** Young's modulus of actin network, and **g** Ratio of the numbers of attached and pushing filaments. *Black circles* denote a bead moving along the long-axis ($0 \leq \theta \leq 30°$); *red squares* represent a bead moving at a skewed orientation ($30 \leq \theta \leq 60°$), finally, *blue triangles* indicate a bead moving along a short axis ($30 \leq \theta \leq 90°$). (Adapted from [60], with permission)

filament elongation after attachment. The filament detachment rate is described by an expression similar to equation but with a different amplitude. The macroscopic motion of the bead is described by an overdampedLangevin equation [61]

$$\eta \dot{\mathbf{R}} + F + \xi(t) = 0 \tag{29.5}$$

where η is a factor related to a viscosity of surrounding medium through the Stoker equation, ξ is a δ-correlated random function acting on the bead and F is the force due to the interaction between the filaments and the bead

$$F = \sum_{i=1} \left[q_i F_2^{(i)} + (1 - q_i) F_1^{(i)} \right] \tag{29.6}$$

The motion of the bead is characterized by three different regimes depending on the bead radius: (a) small beads move steadily with an almost constant velocity, (b) larger beads move in a hopping-like manner and (c) a regime of localized-like fluctuations of the bead within the actin cloud when the bead radius is further increased. Moreover, in the transition regime between steady and hopping motions the movement of the bead becomes erratic. The simulation has been done under the assumption that there are two populations of filaments: attached, which pull the beads, and detached, which push it, and it is assumed that the number of attached filaments is smaller than the number of detached ones. It has been shown that there is a correlation between the bead velocity and the number of attached filaments, i.e., the peaks of the velocity are followed by the minima of the number of attached filaments. The motion of the bead in the hopping regime seems to be in close connection with a stick-slip behavior observed in many frictional problems. It has been suggested by Filippov et al. that such stick–slip motion could originate from a collective rupture of adhesion bonds from the attached filaments. In the hopping regime, the bead motility is characterized by a stochastic interplay of free motion and trapping events. The bead dynamics is characterized by two timescales: a long timescale corresponding to large jumps a dn a shorter timescale corresponding to velocity fluctuations. If all filaments are assumed

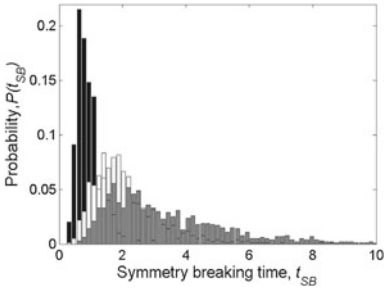


Fig. 29.21 Probability distribution function histograms of symmetry-breaking times calculated for three bead radii, R_0/λ: 0.35- *black*, 0.60- *white* and 0.70-*grey*. (Reproduced from [61], with permission)

to push the bead, an analogous scenario for the bead motility can be observed: three regimes of motion depending by the bead radius.

In a similar way to microscopic model, Fillipov et al. introduced a continuum approach in which the filament subsystem is given in terms of two fields with densities $w_1(r)$ and $w_2(r)$, corresponding to the barbed ends of the detached and attached filament populations, respectively, which time evolution is described by dynamical equations [61]. The motion of the bead is always described by the (29.5) where now the effects of the filaments on the bead is given by the integral force

$$F = \int d\mathbf{r}' F_1 \left(\mathbf{R} - \mathbf{r}'\right) w_1 \left(\mathbf{r}'\right) \tag{29.7}$$

where the kernel force F_1 has the same form as in (29.6). The continuum model also exhibits the phenomenon of spontaneous symmetry-breaking and three different regime of motion depending on the size of the bead, as in the case of microscopic model. In addition, two phenomena are well described by the model: the symmetry-breaking, Fig. 29.21, and viscosity dependence of the bead velocity, Fig. 29.22. In such model symmetry-breaking is caused mainly by the stochastic nature of the nucleation and polymerization that lead to a spontaneous asymmetric fluctuation in the spatial distribution of filaments density. The instantaneous asymmetry of the density distribution results in a net force that pushes the bead out of the actin cloud, followed by a stable directed motion with an almost constant velocity for a broad range of parameters. Figure 29.21 shows the probability distribution functions for symmetry breaking times calculated for three bead radii. With an increase in the bead radius not only does the mean value of symmetry-breaking time increases but also the width of the time distribution, clarifying that is the whole probability distribution function of the symmetry-breaking time that contributes to the understating the mechanisms of motility.

Another important quantity simulated by Filippov et al. is the time spent by the beads in a slippage state (in analogy to stick-slip phenomenon) as a function of the viscosity. In the regime of steady motion, at small viscosities the mean velocity seems

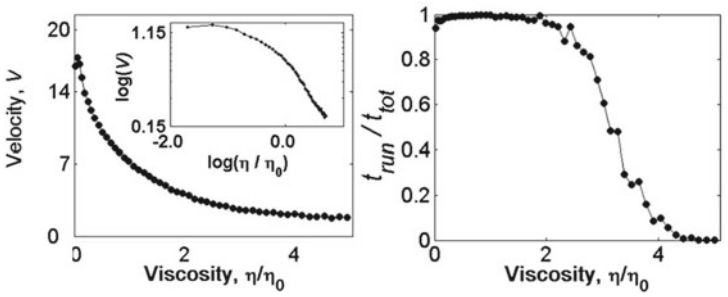


Fig. 29.22 Time averaged bead velocity, and fraction of time that the bead spends in a slippage state as a function of viscosity of the surrounding medium. (Reproduced from [61], with permission)

to be independent of η/η_0, while for larger η/η_0 there is significant decrease. This behavior is justified taking in account that for small η/η_0 the external resistive force due to a viscous friction is smaller than the internal resistive force due to pulling attached filaments [61]. While for higher values of η/η_0, i.e. when the resistance is dominated by viscous force, the mean velocity decreases essentially with η. In the hopping regime, he velocity again decreases slowly with the viscosity. This is because the velocity with which a bead leaves a trapped state is determined by the height of the barrier created by the symmetric actin cloud and does not depend on viscosity. As a consequence, the mean bead velocity during slippage intervals only slightly depends on η/η_0, but the time fraction of the slippage intervals decreases sharply with η/η_0 as shown in Fig. 29.22 (right).

29.4.2 Active Gels Model

In active gel model, the cell is modeled as an active gel described in terms of a density, $\rho(\mathbf{r}, t)$, and a displacement field, $u(\mathbf{r}, t)$, characterizing local deformations [63–65]. In addition, the possibility of cell polarization, for example induced by directed myosin motion or filament treadmilling, can be introduced taking in account a polar orientational order parameter field $\mathbf{P}(\mathbf{r}, t)$. Recently, Banerjee and Marchetti [64], suggested a formulation of active gels out of equilibrium dynamics based on local stress energy density $f = f_{el} + f_P + f_w$ where the three terms are expressed by the following set of equations:

$$f_{el} = \frac{B}{2} u_{kk}^2 + G u_{ij}^2 \tag{29.8a}$$

$$f_P = \frac{a}{2} |\mathbf{P}|^2 + \frac{b}{4} |\mathbf{P}|^4 + \frac{K}{2} \left(\partial_i P_j\right) \left(\partial_j P_i\right) \tag{29.8b}$$

$$f_w = \frac{\omega}{2} \left(\partial_i P_j + \partial_j P_i\right) u_{ij} + \omega' \left(\nabla \cdot \mathbf{P}\right) u_{kk} \tag{29.8c}$$

where f_{el} is the elastic deformations, with B and G the compressional and shear elastic moduli of the gel, respectively, $u_{ij} = 1/2(\partial_i u_j + \partial_j u_i)$ is the strain tensor. In (29.8b), the first two terms allow the onset of a homogeneous polarized state when $a < 0$ and $b > 0$, the last term represents the energy cost for spatially inhomogeneous deformations of the polarization, with K representing a stiffness parameter. Hence, f_w couples strain and polarization and describes the fact that in the active polar system a local strain is associated with a local gradient in polarization. Such gradients can align or oppose each other depending on the sign. Active stresses arise because the gel is driven out of equilibrium by continuous input of energy from the hydrolysis of ATP, characterized by the chemical potential difference between ATP and its products. In polar gels, for example, the active stresses are proportional to $\Delta\mu(\partial_i P_j + \partial_j P_i)$, where $\Delta\mu$ is the chemical potential. Force balances requires $\partial_j \sigma_{ij} = 0$, while the interaction with the substrate can be expressed as $[\sigma_{ij} n_k]_{\text{int}} = E u_i(\mathbf{r}_{\text{int}})$, where n_k is the component normal to cell-substrate interface and $\mathbf{r}_{\text{int}}$ is the point at the interface, E is an effective Young modulus of the substrate to which cell-substrate adhesions contribution must be added. In the case of isotropic cells, the active stress can be expressed by the expression [64]:

$$\sigma(x) = \varsigma \Delta\mu \left(1 - \frac{\cosh\left[(L - 2x)/2\lambda\right]}{\cosh\left(L/2\lambda\right)}\right) \tag{29.9}$$

and the deformation field given by

$$u(x) = \frac{\varsigma \Delta\mu \lambda}{B} \frac{\sinh\left[(L - 2x)/2\lambda\right]}{\cosh\left(L/2\lambda\right)} \tag{29.10}$$

Banejee and Marchetti adapted the (29.10) to substrates with varying stiffness, with an effective Young modulus as $E(x) = E_0 x/L$, then the stresses are described by an equation like [64]

$$\sigma = \frac{\lambda^2 L}{x}\left(\frac{d^2\sigma}{dx^2} - \frac{1}{x}\frac{d\sigma}{dx}\right) + \varsigma\Delta\mu \tag{29.11}$$

which solution can be obtained in terms of hypergeometric functions. The results are summarized in Fig. 29.23

The model proposed by Banejee and Marchetti provides a simple but powerful continuum formulation for the description of cell-substrate interactions and can be extended in various directions by considering more realistic two-dimensional cell geometriesand anisotropic or deformable substrates. Cell migration dynamics could be incorporated in such model, and probably the authors are moving towards such direction. Finally, the model can be used for a direct comparison with experiment describing the stress distribution in migrating cells.

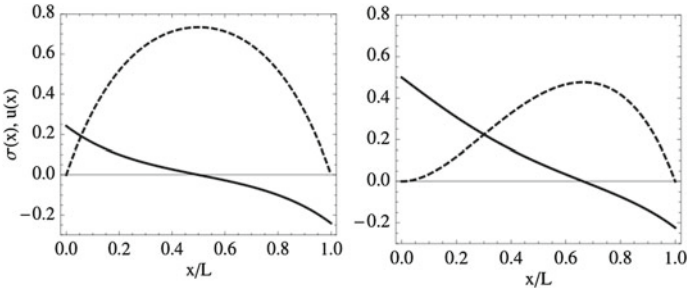


Fig. 29.23 Stress, $\sigma(x)$, and deformation, $u(x)$ as a function of the position x inside a cell of length L. *Left*, stress and deformation changes in the case of isotropic cells, (29.11). *Right*, stress and deformation changes in the case of varying stiffness, (29.10). (Adapted from [64], with permission)

29.4.3 Polymerization of Viscoelastic Gel Confined in a Channel

Active gel models can be able to describe cell motility in confined channels, where actin polymerization at cell membrane are coupled to geometric confinement [66–68]. Hawkins and Voituriez proposed recently a model where the active gel is polymerized at the gel/wall interface with speed velocity v_p, with conditions $v_x(x, z = 0) = -v_z(x, z = b) = v_p$, see Fig. 29.24, right [68]. In this subsection, we limit ourselves to describe the relation between pressure and friction for cell motility constrained along narrow channels, Fig. 29.24, left. The average velocity along the channel can be written as

$$\frac{d}{dx}(bv) = 2v_p - k_d b \tag{29.12}$$

Defining the average velocity along the channel $v(x) = (1/b)\int_0^b v_x(x, z)\,dz$ and introducing the Darcy's law between a fluid flux and pressure gradient, (29.12) becomes

$$\frac{d}{dx}\left(b^3\left(1+\mu^{-1}\right)\frac{dP}{dx}\right) = -12\eta\left(2v_p - k_d b\right) \tag{29.13}$$

where v_p and the non-dimensional friction $\mu = \xi b/6\eta$ are functions of the pressure P and x. Note that the friction coefficient ξ is defined by the relation $\sigma_{xz} = \xi v(x)$. In such model, as introduced by Prost et al. the friction coefficient μ for a polymeric gel depends on the normal constraint. A high normal constraint increases the attachment rate of polymers onto the channel walls by lowering the entropic barrier, and decreases the detachment rate [68]. In a gel moving at velocity v_0, a single filament attached

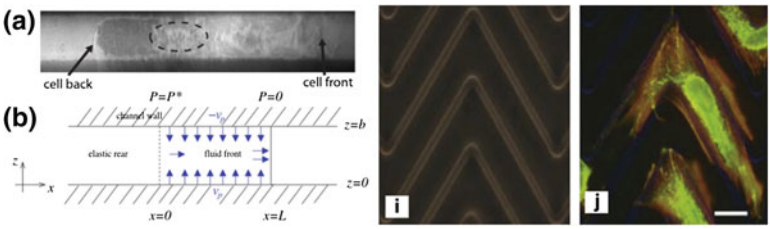


Fig. 29.24 *Left upper* Image of a dendritic cell moving towards right in a channel of 4 μm width. Typical velocity reaches 12–15 μm/min in channels and 4–6 μm/min on a flat surface. *Left bottom* Example of a channel geometry with boundary conditions. The arrows indicate the flow direction. (Adapted from [68], with permission). *Middle and right* Image of a polydimethylsiloxanezig-zagpattern (*i*) on which a smooth muscle cell was cultured adapting morphology and locomotion direction, (*j*). (Adapted from [47], with permission)

to the surface for a typical time τ_a will experience aforce $F = E v_0 \tau_a$, where E is the elastic modulus. If there are n filaments and the probability of a filament p_a, the total frictional force will be given by

$$F_{fric} = \frac{n E v_0 \tau_a^2}{\tau_a + \tau_d} \tag{29.14}$$

where it has been assumed that the number of connected filaments changes slowly and detach with a rate τ_d. For $\tau_a \gg \tau_d$ the friction force reduces to $F_{fric} \approx n E v_0 \tau_a$. If the speed of the gel is relatively small, mechanical breakage of filaments can be neglected and the detachment rate of filaments is given by the chemical off rate which is proportional to the normal constrain $\exp(-\beta(P - \sigma_{nn}))$, where $\beta = 1/k_B T$, and the normal stress $\sigma_{nn} = \sigma_{zz}$ for both the walls at $z = 0$ and $z = b$. Analogously, the friction coefficient can be written as $\mu = \mu_0 \exp(-\beta(P - \sigma_{nn}))$.

The other quantity that must be described now is the relation between pressure and polymerization. Based on the ratchet model, the polymerization speed of a filament at the cell membrane is given by the bare attachment rate k_{on} in the absence of any force multiplied by the probability of thermal fluctuations producing a gap large enough for a new subunit to attach. Since the energy required for a gap the size of a subunit δ is given by $F\delta$, where F is the load, therefore the polymerization speed is given by $v_p = k_{on}\delta \exp(-\beta(P - \sigma_{nn})$. Such simple relation is obtained under the assumption that the diffusion over the distance δ is fast compared to the rate of addition of subunits k_{on} and therefore does not influence the polymerization speed.

As mentioned, the model introduced in [67, 68] does not require specific adhesion proteins since the mechanisms relies on a pressure build-up in confinement which enhances friction. The cell motility is based principally on polymerization against the confining walls, however, also if contractility due to myosin molecular motors is not essential to motion it can increase the speed, as shown by experiments [69].

29.5 Conclusions and Future Perspectives

One basic challenge in biotribology is to identify and to test simple mechanisms of cell motion that is also an interdisciplinary challenge of great interest involving biology, mathematical model physical sciences, with suitable applications to medical and bio-technological sciences. Cells often move within a complex environment such as living tissue, or human being made scaffolds. In general, such environments have low Reynolds number, meaning viscous forces dominate over inertial forces. Since inertia can be neglected, motion needs a constant input of energy. Systems driven out of equilibrium by an internal or external energy source can be considered active systems. The cell cytoskeleton is a striking example of such an active system, composed by a network of long semi-flexible filaments made up of protein subunits, interacting with other proteins such as motor proteins, for example, the filaments, such as actin, use the chemical energy of ATP hydrolysis to polymerize and the motor proteins use this chemical energy to exert active stresses that deform the network.

Currently, two main different mechanisms for cell motility have now been distinguished: polymerization and contractility due to molecular motors. Specifically, polymerization is referred to polymerization of a filament at one end and depolymerization at the other end.

In this chapter we have dealt out some scientific achievement on the role of adhesion and frictional force on cell motility, in a special way, on the interaction of a cell with an non-cell surface (extracellular matrix, or polymericscaffold). We have described experiments in Force Traction Microscopy, on the role of substrate stiffness on cell motility and carried out some recent friction-based cell motility models. In all the models describing cell motility the essential ingredients for motion are an energy input to overcome dissipation and sufficient adhesion or friction with a substrate to transfer momentum. Many open problems inherent the cell motility, inparticular connected to the tribology aspects, are an interesting and exciting challenge for a wide range of scientists operating in physics, chemistry, biology, medicine and engineering.

Acknowledgments MD and OS like to thank the NanoICT Project at the Area della Ricerca di Pisa, CNR.SD greatly thanks Dr. Luisa Trombi for her valuable experimental contribute (stem cell isolation and culture), and the Tuscany Region for financial support (CUCCS/RRMR Project-University of Pisa and Salute 2009 grant).

References

1. D.E. Discher, P. Janmey, Y.-L. Wang, Science, **310**, 1139–1143 (2005)
2. D. Bray, *Cell Movements: From Molecules to Motility*, 2nd edn. (Garland Science, New York, 2001)
3. B. Palsson, J.A. Hubbell, R. Plonsey, J.D. Bronzino, *Tissue Engineering* (CRC Press, London, 2003)
4. T. Shemesh, B. Geiger, A.D. Bershadsky, M. Kozlov, PNAS **102**, 12383–12388 (2005)

5. E.L. Barnhart, K.-C. K-C, K. Keren, A. Mogilner, J.A. Theriot, Plos Biol. **9**, e1001059 (2011)
6. A. Curtis, M. Riehle, Phys. Med. Biol. **46**, R47–R65 (2001)
7. M. D'Acunto, G.Ciardelli, A. Rechichi, F.M. Montevecchi, P. Giusti, in *Fundamentals on Friction and Wear on the Nanoscale*, ed. by E. Gnecco, E. Meyer (Springer, 2007)
8. M. Abercrombie, Proc. Roy. Soc. London B, **207**, 129–149 (1980)
9. K.I. Anderson, R. Cross, Curr. Biol. **10**, 253–260 (2000)
10. M. Buenemann, H. Levine, W.J. Rappel, L.M. Sander, Biophys. J. **99**, 50–58 (2010)
11. J. Toner, Y. Tu, S. Ramaswamy, Ann. Phys. **318**, 170–244 (2005)
12. O.D. Wang, Enrion. Microbiol. **12**, 1363–73 (2010)
13. K. John, P. Peyla, M. Ismail, C. Misbah, D. Caillerie, A. Raoult, J. Prost. http://arxiv.org/pdf/0909.0630v1.pdf
14. A.S. Smith, Nat. Phys. **6**, 726–729 (2010)
15. A. Mogilner, G. Oster, Biophys. J. **71**, 3030–3045 (1996)
16. A. Bernheim-Groswasser, S. Wiesner, R.M. Golsteyn, M.F. Carlier, C. Sykes, Nature **417**, 308–311 (2002)
17. F.J. Nadelec, T. Surrey, A.C. Maggs, S. Leibler, Nature **389**, 305–308 (1997)
18. K. Kruse, J.F. Joanny, F. Julicher, J. Prost, K. Sekimoto, Phys. Rev. Lett. **92**, 078101 (2004)
19. F. Julicher, K. Kruse, J. Prost, J.F. Joanny, Phys. Rep. **449**, 3–28 (2007)
20. R. Voitureiz, J.F. Joanny, J. Prost, Phys. Rev. Lett. **96**, 028102 (2006)
21. D. Marenduzzo, E. Orlandini, M.E. Cates, J.M. Yeomans, Phys. Rev. E **76**, 031921 (2007)
22. F.M. Carlier, D. Pantaloni, J. Mol. Biol. **269**, 459–467 (1997)
23. J.A. Theriot, T.J. Mitchinson, Nature **352**, 126–131 (1991)
24. T. Oliver, M. Dembo, K. Jacobson, Cell Motil. Cytoskelet. **31**, 225–240 (1995)
25. T. Oliver, M. Dembo, A. Ishibara, K. Jacobson, Biophys. J. **70**, 2008–2022 (1996)
26. C. Franck, S.A. Maskarinec, D.D.A. Tirrell, G. Ravichandran, PloS One **6**, e17833 (2011)
27. R. Pelham, Y. Wang, PNAS **94**, 13661–13665 (1997)
28. I. Levental, P. Georges, P. Janmey, Soft Matter **3**, 299–306 (2007)
29. L. Landau, E. Lifshitz, *Theory of Elasticity*, 3rd edn. (Pergamon Press, Oxford, 1986)
30. S. Maskarinec, C. Franck, D. Tirrell, G. Ravichandran, PNAS **106**, 22108–22113 (2009)
31. I.B. Bischofs, Elastic Interactions of Cellular Force Patterns, PhD Thesis, Postadam University, 2004
32. G.A. Dunn, J.P. Hearth, Exp. Cell Res. **101**, 1–14 (1976)
33. G.A. Dunn, T. Ebendal, Exp. Cell Res. **111**, 475–479 (1978)
34. S.B. Carter, Nature **213**, 256–260 (1967)
35. W.S. Haston, J.M. Shields, P.C. Wilkinson, Exp. Cell Res. **146**, 117–126 (1983)
36. C.M. Lo, H.B. Wang, M. Dembo, Y.L. Yang, Biophys. J. **79**, 144–152 (2000)
37. N.Q. Balaban, U.S. Schwarz, D. Riveline, P. Goichberg, G. Tzur, I. Sabanay, D. Mahalu, S. Safran, A. Bershadsky, L. Addadi, B. Geiger, Nat. Cell Biol. **3**, 466–472 (2001)
38. A.S.G. Curtis, C.D. Wilkinson, Biomaterials **18**, 1573–1583 (1997)
39. A.S.G. Curtis, C.D. Wilkinson, J. Biomat. Sci. Polym. Ed. **9**, 1313–1329 (1998)
40. S. Bajpai, N.Y. Kim, C.A. Reinhart-King, Rev. Int. J. Mol. Sci. **12**, 8596–8609 (2011)
41. S.S. Chang, W.H. Guo, Y. Kim, Y.L. Wang, Biophys. J. **104**, 313–321 (2013)
42. S. Li, J.L. Guan, S. Chien, Ann. Rev. Biomed. Eng. **7**, 105–150 (2005)
43. K. von der Mark, J. Park, S. Bauer, P. Schmuki, Cell Tissue Res. **339**, 131–53 (2010)
44. A.L. Berrier, K.M. Yamada, J. Cell Physiol. **213**(3), 565–573 (2007)
45. C.S. Shin, B. Kwak, B. Han, K. Park, Mol. Pharm. **10**, 2167–75 (2013)
46. B.D. Riehl, J.H. Park, I.K. Kwon, J.Y. Lim, Tissue Eng. Part B Rev. **18**, 288–300 (2012)
47. M. Nikkhah, F. Edalat, S. Manoucheri, A. Khademhosseini, Biomaterials **33**, 5230–5246 (2012)
48. C.H. Choi, S.H. Hagvall, B.M. Wu, J.C.Y. Dunn, R.E. Beygui, C.J. Kim, Biomaterials **28**, 1672–1679 (2007)
49. L. De Bartolo, M. Rende, S. Morelli, G. Giusi, S. Salerno, A. Piscioneri, A. Gordano, A. Di Vito, M. Canonaco, E. Drioli, J. Membr. Sci. **325**, 139–149 (2008)
50. M.J.P. Biggs, R.G. Richards, M.J. Dalby, Nanomed. Nanotechnol. Biol. Med. **6**, 619–633 (2010)

51. M.H. Kim, M. kino-Oka, M. Kawase, K. Yagi, M. Taya, J. Biosci. Bioeng. **103**, 192–199 (2007)
52. M.J. Dalby, Med. Eng. Phys. **27**, 730–742 (2005)
53. H-I. Chang, Y. Wang, in *Regenerative Medicine and Tissue Engineering- Cells and Biomaterials*, ed. by D. Eberli (2011) http://www.intechopen.com
54. G. Guérin, D. Ambard, P. Swider, J. Biomech. **43**, 1287–1291 (2009)
55. M. D'Acunto, G. Ciardelli, P. Narducci, A. Rechichi, P. Giusti, Mater. Lett. **59**, 1627–1633 (2005)
56. A. Rechichi, G. Ciardelli, M. D'Acunto, G. Vozzi, P. Giusti, J. Biomed. Mater. Res. Part A, **84**, 847–855 (2008)
57. B. Harland, S. Walcott, S.X. Sun, Phys. Biol. **8**, 015011 (2011)
58. S. Reboux, G. Richardson, O.E. Jensen, Proc. R. Soc. A, **464**, 447–467 (2008)
59. P. Recho, T. Putilat, L. Truskinovsky, http://arxiv.org/abs/1302.3999 (2013)
60. J. Zhu, A. Mogilner, PloS Biol. **8**, e1002764 (2012)
61. A.E. Filippov, J. Klafter, M. Urbakh, J. Phys.: Condens. Matter **17**, S3929–S3944 (2005)
62. F. Ziebert, I.S. Aranson, Plos One **8**, e64511 (2013)
63. S. Banerjee, M.C. Marchetti, Europhys. Lett. **96**, 28003 (2011)
64. S. Banerjee, M.C. Marchetti, Phys. Rev. Lett. **109**, 108101 (2012)
65. S. Banerjee, M.C. Marchetti, New J. Phys. **15**, 035015 (2013)
66. G. Faure-André, P. Vargas, M.I. Yuseff, M. Heuzé, J. Diaz, D. Lankar, V. Steri, J. Manry, S. Hugues, F. Vascotto, J. Boulanger, G. Raposo, M.R. Bono, M. Rosenblatt, M. Piel, A.M. Lennon-Duménil, Science **322**, 1705–1710 (2008)
67. R.J. Hawkins, M. Piel, G. Faure-André, A.M. Lennon-Duménil, J.F. Joanny, J. Prost, R. Voituriez, Phys. Rev. Lett. **102**, 058103 (2009)
68. R.J. Hawkins, R. Voituriez, Math. Model. Nat. Phenom. **5**, 84–105 (2010)
69. T. Lämmermann, B.L. Bader, S.L. Monkley, T. Worbs, R. Wedlich-Söldner, K. Hirsch, M. Keller, R. Förster, D.R. Crithley, R. Fässler, M. Sixt, Nature **453**, 51–55 (2008)

Index

E. Gnecco and E. Meyer (eds.), *Fundamentals of Friction and Wear on the Nanoscale*, NanoScience and Technology, DOI 10.1007/978-3-319-10560-4