

Bharat Bhushan *Editor*

Nanotribology and Nanomechanics

An Introduction

Fourth Edition

Nanotribology and Nanomechanics

Bharat Bhushan
Editor

Nanotribology and Nanomechanics

An Introduction

Fourth Edition



Editor

Bharat Bhushan
Nanoprobe Laboratory for
Bio- & Nanotechnology and
Biomimetics (NLBB)

Battelle Center for Science and Technology
Policy, The John Glenn College of Public
Affairs,
Ohio State University
Columbus, OH
USA

ISBN 978-3-319-51432-1 ISBN 978-3-319-51433-8 (eBook)
DOI 10.1007/978-3-319-51433-8

Library of Congress Control Number: 2016963428

1st–3rd edition: © Springer-Verlag Berlin Heidelberg 2005, 2008, 2011

4th edition: © Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Foreword



The invention of the scanning tunneling microscope in 1981 led to an explosion of a family of instruments called scanning probe microscopes (SPMs). One of the most popular instruments in this family is the atomic force microscope (AFM), which was introduced to the scientific community in 1986. The application of SPMs has penetrated numerous science and engineering fields. Proliferation of SPMs in science and technology laboratories is similar to optical microscopes fifty years ago. SPMs have even made it into some high school science laboratories.

Evolution of nanotechnology has accelerated the use of SPMs and vice versa. The scientific and industrial applications include quality control in the semiconductor industry and related research, molecular biology and chemistry, medical studies, materials science, and information storage systems.

AFMs were developed initially for imaging with atomic or near-atomic resolution. After their invention, they were modified for tribological studies. AFMs are now intensively used in this field and have lead to the development of the field of nanotribology. Researchers can image single lubricant molecules and their agglomeration and measure surface topography, adhesion, friction, wear, lubricant film thickness, and mechanical/electrical/magnetic/thermal properties on the nanometer scale. SPMs also are used for nanofabrication and nanomachining. Beyond their use as an analytical instrument, SPMs are being developed as industrial tools for processes such as data storage.

With the advent of more powerful computers, atomic-scale simulations of tribological phenomena have been conducted. Simulations have been able to predict the observed phenomena. The development in the field of nanotribology and nanomechanics has attracted numerous physicists and chemists, in addition to engineers and materials scientists. I am very excited that SPMs have had such an immense impact on the field of tribology.

I congratulate Prof. Bharat Bhushan in helping to develop this field of nanotribology and nanomechanics. Professor Bhushan has harnessed his own knowledge and experience, gained in several industries and universities, and has assembled a large number of internationally recognized authors from

both academia and industry. The depth and breadth of their combined experience is impressive.

Professor Bharat Bhushan's comprehensive book is intended to serve as both a textbook for university courses as well as a reference for researchers. It is a timely addition to the literature on nanotribology and nanomechanics, which I anticipate will stimulate further interest in these important fields. I expect that it will be well received by the international scientific community.

Prof. Gerd Binnig

Nobel Laureate Physics, 1986, Kavli Award Winner Nanoscience, 2016

Definiens AG Munich, Germany

Preface for the Fourth Edition

Tribology is the science and technology of interacting surfaces in relative motion and of related subjects and practices. The nature and consequences of interactions that take place at a moving interface control its friction, wear, and lubrication behavior. Understanding the nature of these interactions and solving the technological problems associated with interfacial phenomena constitute the essence of tribology. The importance of friction and wear control cannot be overemphasized for economic reasons and long-term reliability.

There has been an emergence and proliferation of proximal probes, in particular tip-based microscopies and the surface force apparatus, and of computational techniques, for simulating tip–surface interactions and interfacial properties. These have allowed systematic investigations of interfacial problems with high resolution, as well as ways and means of modifying and manipulating nanostructures. These advances provide the impetus for research aimed at developing a fundamental understanding of the nature and consequences of the interactions between materials on the atomic scale, and they guide the rational design of material for technological applications. In short, they have led to the appearance of the new fields of nanotribology and nanomechanics.

The field of tribology is truly interdisciplinary. Until the 1980s, it had been dominated by mechanical and chemical engineers who conducted macroscale friction and wear tests to predict the useful life of machine components. They devised new lubricants to minimize the effects of friction and wear, and extend that useful life. The development in the field of nanotribology has attracted many more physicists, chemists, and material scientists who have contributed significantly to the fundamental understanding of friction and wear processes and lubrication on an atomic scale. Thus, tribology and mechanics are now studied by both engineers and scientists. The nanotribology and nanomechanics fields have grown rapidly, and it has become fashionable to call oneself a “tribologist.” Tip-based microscopies have also been used for materials’ characterization, as well as for the measurement of mechanical and electrical properties, all on the nanoscale. Since 1991, international conferences and courses have been organized regularly by various professional societies and other organizations on the new fields of nanotribology, nanomechanics, and nanomaterials characterization.

There are many applications that require detailed understanding of tribological and mechanical processes on the macro- and nanoscales. In the early 1980s, tribology of magnetic storage systems (rigid disk drives and tape drives) became one of the important parts of tribology. Microelectromechanical Systems (MEMS)/Nanoelectromechanical Systems (NEMS) and bioMEMS/NEMS, all part of nanotechnology, have appeared in the marketplace since the 1990s. These, of course, present new tribological challenges. Tribology of processing systems such as copiers, printers, scanners, and cameras is important also, although it has not received much attention. Along with many industrial applications, there has been development of new materials, coatings, and treatments such as synthetic diamond, diamond-like carbon films, self-assembled monolayers, and chemically grafted films, to name a few with nanoscale thicknesses.

It is clear that the general field of tribology has grown rapidly since the mid-1980s. Conventional tribology is well established, but nanotribology and nanomechanics also have taken the center stage. Nanomaterials also have been developed and are used in some applications. Furthermore, new industrial applications, particularly biomedical applications, continue to evolve with their unique challenges.

Very few tribology handbooks exist, and those that do are dated. They have focused on conventional tribology, traditional materials, and matured industrial applications. No mechanics handbooks exist. Nanotribology, nanomechanics, and nanomaterial characterization have become important in many nanotechnology applications. A primer to these subjects is needed. The purpose of this revised 4th Edition of Nanotribology and Nanomechanics is to present the principles of nanotribology and nanomechanics and their relevance to various applications. The appeal of the subject book is expected to be broad. The first edition was published in 2005, the second in 2008, and the third in 2011. This fourth edition is an update based on recent developments.

The chapters in the book have been written by internationally recognized experts in the field, from academia, national research laboratories and industry, and from all over the world. The book integrates the knowledge of the field from mechanics and materials science points of view. In each chapter, we start with macroconcepts leading to microconcepts. We assume that the reader is not expert in the field of nanotribology and nanomechanics, but has some knowledge of macrotribology/mechanics. It covers various measurement techniques and their applications, and theoretical modeling of interfaces. The organization of the book is straightforward. The first part of the book covers fundamental experimental and theoretical studies. The latter part covers applications.

The book is intended for three types of readers: graduate students of nanotribology/nanomechanics/nanotechnology, research workers who are active or intend to become active in this field, and practicing engineers who have encountered a tribology and mechanics problem and hope to solve it as expeditiously as possible. The book should serve as an excellent text for one or two semester graduate courses in scanning probe microscopy/applied

scanning probe methods, and nanotribology/nanomechanics/nanotechnology in mechanical engineering, materials science, applied physics, or applied chemistry.

I would like to thank the authors for their excellent contributions in a timely manner. I also wish to thank my wife, Sudha, my son, Ankur, and my daughter, Noopur, who have been very forbearing during the preparation of this book.

Powell, OH, USA
November 2016

Bharat Bhushan
bhushan.2@osu.edu

Contents

1 Introduction—Measurement Techniques and Applications	1
Bharat Bhushan	
Part I Scanning Probe Microscopy	
2 Scanning Probe Microscopy—Principle of Operation, Instrumentation, and Probes	33
Bharat Bhushan and Othmar Marti	
3 Calibration of Normal and Lateral Forces in Cantilevers Used in Atomic Force Microscopy	95
Manuel L.B. Palacio and Bharat Bhushan	
4 Biomedical Sensing with the Atomic Force Microscope	135
Constanze Lamprecht, Jürgen Strasser, Melanie Köhler, Sandra Posch, Yoo Jin Oh, Rong Zhu, Lilia A. Chtcheglova, Andreas Ebner and Peter Hinterdorfer	
Part II Nanomechanics	
5 Depth-sensing Nanoindentation Techniques for Nanomechanical Characterization	177
Bharat Bhushan	
6 Nanomechanical Properties of Nanostructures and Scale Effects	253
Bharat Bhushan	
7 Computer Simulations of Nanometer-Scale Indentation and Friction	301
Susan B. Sinnott, Seong-Jun Heo, Donald W. Brenner, Judith A. Harrison and Douglas L. Irving	

Part III Nanotribology

8 Nanotribology, Nanomechanics and Materials Characterization Studies Using Scanning Probe Microscopy	373
Bharat Bhushan	
9 Surface Forces and Nanorheology of Molecularly Thin Films	457
Dong Woog Lee, Marina Ruths and Jacob N. Israelachvili	
10 Atomic Scale Friction Phenomena	519
Enrico Gnecco, Rémy Pawlak, Marcin Kisiel, Thilo Glatzel and Ernst Meyer	
11 Scale Effect in Mechanical Properties and Tribology	549
Bharat Bhushan and Michael Nosonovsky	

Part IV Molecularly-Thick Films for Lubrication

12 Nanotribology of Ultrathin and Hard Amorphous Carbon Films	593
Bharat Bhushan	
13 Self-assembled Monolayers (SAMs) for Nanotribology and Surface Protection	641
Bharat Bhushan	
14 Nanoscale Boundary Lubrication Studies	689
Bharat Bhushan	

Part V Industrial Applications

15 Micro/Nanotribology and Micro/Nanomechanics of Magnetic Storage Devices	749
Bharat Bhushan	
16 Nanotribology and Nanomechanics of MEMS/NEMS and BioMEMS/BioNEMS Materials and Devices	797
Bharat Bhushan	
Index	909

Editors and Contributors

About the Editors



Dr. Bharat Bhushan received an M.S. in mechanical engineering from the Massachusetts Institute of Technology in 1971, an M.S. in mechanics, and a Ph.D. in mechanical engineering from the University of Colorado at Boulder in 1973 and 1976, respectively, an MBA from Rensselaer Polytechnic Institute at Troy, NY, in 1980, Doctor Technicae from the University of Trondheim at Trondheim, Norway, in 1990, a Doctor of Technical Sciences from the Warsaw University of Technology at Warsaw, Poland, in 1996, and Doctor Honouris

Causa from the National Academy of Sciences at Gomel, Belarus, in 2000 and University of Kragujevac, Serbia, in 2011. He is a registered professional engineer. He is presently an Ohio Eminent Scholar and The Howard D. Winbigler Professor in the College of Engineering, and the director of the Nanoprobe Laboratory for Bio- & Nanotechnology and Biomimetics (NLBB) and affiliated faculty in John Glenn College of Public Affairs at the Ohio State University, Columbus, Ohio. In 2013–14, he served as an ASME/AAAS Science & Technology Policy Fellow, House Committee on Science, Space & Technology, United States Congress, Washington, DC. His research interests include fundamental studies with a focus on scanning probe techniques in the interdisciplinary areas of bio/nanotribology, bio/nanomechanics and bio/nanomaterials characterization and applications to bio/nanotechnology, and biomimetics. He is an internationally recognized expert of bio/nanotribology and bio/nanomechanics using scanning probe microscopy, and is one of the most prolific authors. He is considered by some a pioneer of the tribology and mechanics of magnetic storage devices. He has authored 8 scientific books, 90+ handbook chapters, 800+ scientific papers (Google Scholar h-index—110+ with 60k+ citations; Web of Science h-index—80+; ISI Highly Cited Researcher in Materials Science since 2007 and in Biology and Biochemistry, 2013; ISI Top 5% Cited Authors for Journals in Chemistry, 2011), and 60+ technical reports. He has also edited 50+ books and holds more than 25 US and foreign patents. He is coeditor of

Springer NanoScience and Technology Series and coeditor of Microsystem Technologies, and Member of Editorial Board of PNAS. He has given more than 400 invited presentations on six continents and more than 200 keynote/plenary addresses at major international conferences.

Dr. Bhushan is an accomplished organizer. He organized the first symposium on Tribology and Mechanics of Magnetic Storage Systems in 1984 and the first international symposium on Advances in Information Storage Systems in 1990, both of which are now held annually. He organized two international NATO institutes in Europe. He is the founder of an ASME Information Storage and Processing Systems Division founded in 1993 and served as the founding chair during 1993–1998. His biography has been listed in over two dozen Who's Who books including Who's Who in the World. He has received more than two dozen awards for his contributions to science and technology from professional societies, industry, and US government agencies including Life Achievement Tribology Award and Institution of Chemical Engineers (UK) Global Award. His research was listed as the top ten science stories of 2015. He is also the recipient of various international fellowships, including the Alexander von Humboldt Research Prize for Senior Scientists, Max Planck Foundation Research Award for Outstanding Foreign Scientists, and Fulbright Senior Scholar Award. He is a foreign member of the International Academy of Engineering (Russia), Byelorussian Academy of Engineering and Technology and the Academy of Tribotechnology of Ukraine, an honorary member of the Society of Tribologists of Belarus and STLE, a fellow of ASME, IEEE, and the New York Academy of Sciences, and a member of ASEE, Sigma Xi and Tau Beta Pi.

Dr. Bhushan has previously worked for Mechanical Technology Inc., Latham, NY; SKF Industries Inc., King of Prussia, PA; IBM, Tucson, AZ; and IBM Almaden Research Center, San Jose, CA. He has held visiting professorship at University of California at Berkeley, University of Cambridge, UK, Technical University Vienna, Austria, University of Paris, Orsay, ETH Zurich, EPFL Lausanne, Univ. of Southampton, UK, Univ. of Kragujevac, Serbia, Tsinghua Univ., China, Harbin Inst., China, and KFUPM, Saudi Arabia.

Contributors

Bharat Bhushan Nanoprobe Laboratory for Bio- and Nanotechnology and Biomimetics, Battelle Center for Science and Technology Policy, The John Glenn College of Public Affairs, The Ohio State University, Columbus, OH, USA

Donald W. Brenner Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC, USA

Lilia A. Chtcheglova Center for Advanced Bioanalysis GmbH, Linz, Austria

Andreas Ebner Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Ohio Eminent Battelle Center for Science and Technology Policy, The John Glenn College of Public Affairs, Columbus, USA

Thilo Glatzel Department of Physics, University of Basel, Basel, Switzerland

Enrico Gnecco Otto Schott Institute of Materials Research (OSIM), Friedrich Schiller University Jena, Jena, Germany

Judith A. Harrison Department of Chemistry, U.S. Naval Academy, Annapolis, MD, USA

Seong-Jun Heo Department of Materials Science and Engineering, University of Florida, Gainesville, FL, USA

Peter Hinterdorfer Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Douglas L. Irving Department of Materials Science and Engineering, North Carolina State University, Raleigh, NC, USA

Jacob N. Israelachvili Department of Chemical Engineering and Materials Department, University of California, Santa Barbara, CA, USA

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

Melanie Köhler Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Constanze Lamprecht Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Dong Woog Lee School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

Othmar Marti Abteilung Experimentelle Physik, Universitaet Ulm, Ulm, Germany

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

Michael Nosonovsky Department of Mechanical Engineering, University of Wisconsin at Milwaukee, Milwaukee, WI, USA

Yoo Jin Oh Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Manuel L.B. Palacio Laboratory for Bio- and Nanotechnology and Biomimetics, The Ohio State University, Columbus, OH, USA

Rémy Pawlak Department of Physics, University of Basel, Basel, Switzerland

Sandra Posch Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Marina Ruths Department of Chemistry, University of Massachusetts Lowell, Lowell, MA, USA

Susan B. Sinnott Department of Materials Science and Engineering, University of Florida, Gainesville, FL, USA

Jürgen Strasser Center for Advanced Bioanalysis GmbH, Linz, Austria

Howard D. Winbigler The Ohio State University, Columbus, OH, USA

Rong Zhu Institute of Biophysics, Johannes Kepler University Linz, Linz, Austria

Introduction—Measurement Techniques and Applications

Bharat Bhushan

Abstract

In this introduction chapter, the definition and history of tribology, its industrial significance and origins, and the significance of nanotribology and nanomechanics are described. Next, various measurement techniques used in nanotribological and nanomechanical studies are described. Interest in the nanotribology field grew from magnetic storage devices and, later, the applicability to the general field of micro/nanoelectromechanical systems (MEMS/NEMS) became clear. A few examples of magnetic storage devices and MEMS/NEMS and BioMEMS/NEMS are presented where nanotribological and nanomechanical tools and techniques are essential for interfacial studies. Finally, reasons why nanotribological and nanomechanical studies may be important in some applications are presented. In the last section, the organization of the book is presented.

Keywords

Nanotribology • Molecular tribology • Nanomechanics • Tribology • Surface roughness • Adhesion • Friction • Wear • Lubrication • Atomic force microscopy • Friction force microscopy • Surface force apparatus • Molecular dynamic simulation

1.1 Definition and History of Tribology

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: bhushan.2@osu.edu

The word tribology first was reported in a landmark report by Peter Jost (1966). The word is derived from the Greek word *tribos* meaning rubbing, so the literal translation would be “the

science of rubbing.” Its popular English language equivalent is friction and wear or lubrication science, alternatively used. The latter term is hardly all-inclusive. Dictionaries define tribology as the science and technology of interacting surfaces in relative motion and of related subjects and practices. Tribology is the art of applying operational analysis to problems of great economic significance, such as reliability, maintenance, and wear of technical equipment, ranging from spacecraft to household appliances. Surface interactions in a tribological interface are highly complex, and their understanding requires knowledge of various disciplines including physics, chemistry, applied mathematics, solid mechanics, fluid mechanics, thermodynamics, heat transfer, materials science, rheology, lubrication, machine design, performance and reliability (Bhushan 2013a, b).

It is only the name tribology that is relatively new as interest in the constituent parts of tribology is older than recorded history (Dowson 1998). It is known that drills made during the Paleolithic period for drilling holes or producing fire were fitted with bearings made from antlers or bones, and potters’ wheels or stones for grinding cereals, etc., clearly had a requirement for some form of bearings (Davidson 1957). A ball thrust bearing dated about CE 40 was found in Lake Nimi near Rome.

Records show the use of wheels from 3500 BCE, which illustrates our ancestors’ concern

with reducing friction in translatory motion. The transportation of large stone building blocks and monuments required the know-how of frictional devices and lubricants, such as water-lubricated sleds. Figure 1.1 illustrates the use of a sledge to transport a heavy statue by Egyptians Circa 1880 BCE (Layard 1853). In this transportation, 172 slaves are being used to drag a large statue weighing about 600 kN along a wooden track. One man, standing on the sledge supporting the statue, is seen pouring a liquid (most likely water) into the path of motion; perhaps he was one of the earliest lubrication engineers. (Dowson (1998) has estimated that each man exerted a pull of about 800 N. On this basis, the total effort, which must at least equal the friction force, becomes 172×800 N. Thus, the coefficient of friction is about 0.23.) A tomb in Egypt that was dated several thousand years BC provides the evidence of use of lubricants. A chariot in this tomb still contained some of the original animal-fat lubricant in its wheel bearings.

During and after the glory of the Roman empire, military engineers rose to prominence by devising both war machinery and methods of fortification, using tribological principles. It was the renaissance engineer-artist Leonardo da Vinci (1452–1519), celebrated in his days for his genius in military construction as well as for his painting and sculpture, who first postulated a scientific approach to friction. Da Vinci deduced the rules governing the motion of a rectangular

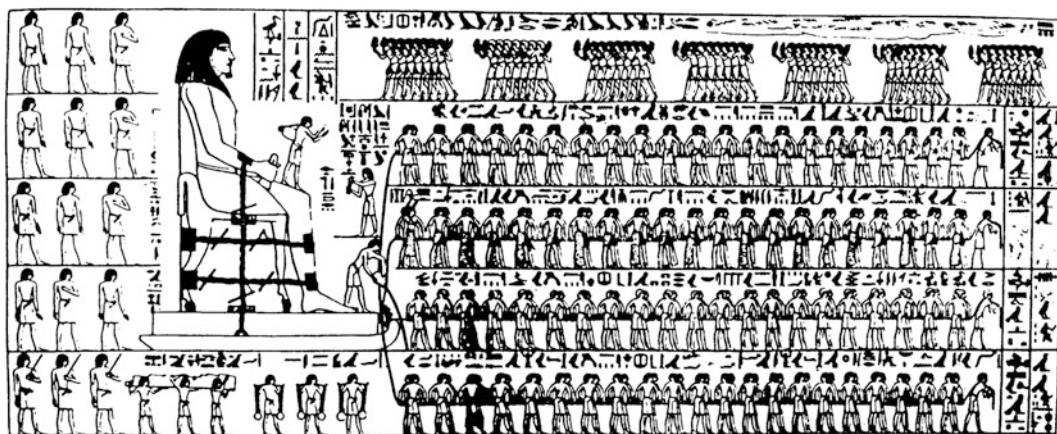


Fig. 1.1 Egyptians using lubricant to aid movement of colossus, El-Bersheh, circa 1800 BC

block sliding over a flat surface. He introduced for the first time the concept of the coefficient of friction as the ratio of the friction force to normal load. However, his work had no historical influence because his notebooks remained unpublished for hundreds of years. In 1699, the French physicist Guillaume Amontons rediscovered the rules of friction after he studied dry sliding between two flat surfaces (Amontons 1699). First, the friction force that resists sliding at an interface is directly proportional to the normal load. Second, the amount of friction force does not depend on the apparent area of contact. These observations were verified by French physicist Charles-Augustin Coulomb [better known for his work on electrostatics (Coulomb 1785)]. He added a third rule that the friction force is independent of velocity once motion starts. He also made a clear distinction between static friction and kinetic friction.

Many other developments occurred during the 1500s, particularly in the use of improved bearing materials. In 1684, Robert Hooke suggested the combination of steel shafts and bell-metal bushes as preferable to wood shod with iron for wheel bearings. Further developments were associated with the growth of industrialization in the latter part of the eighteenth century. Early developments in the petroleum industry started in Scotland, Canada, and the United States in the 1850s (Parish 1935; Dowson 1998).

Though essential laws of viscous flow were postulated by Sir Isaac Newton in 1668; scientific understanding of lubricated bearing operations did not occur until the end of the nineteenth century. Indeed, the beginning of our understanding of the principle of hydrodynamic lubrication was made possible by the experimental studies of Beauchamp Tower (1884) and the theoretical interpretations of Osborne Reynolds (1886) and related work by Petroff (1883). Since then developments in hydrodynamic bearing theory and practice have been extremely rapid in meeting the demand for reliable bearings in new machinery.

Wear is a much younger subject than friction and bearing development, and it was initiated on a largely empirical basis. Scientific studies of wear developed little until the mid-twentieth

century. Ragnar Holm made one of the earliest substantial contributions to the study of wear (Holm 1946).

The industrial revolution (1750–1850 A.D.) is recognized as a period of rapid and impressive development of the machinery of production. The use of steam power and the subsequent development of the railways in the 1830s led to promotion of manufacturing skills. Since the beginning of the twentieth century, from enormous industrial growth leading to demand for better tribology, knowledge in all areas of tribology has expanded tremendously (Holm 1946; Bowden and Tabor 1950, 1964; Bhushan 1996, 2000, 2013a, b, 2016, 2017a; Bhushan and Gupta 1997).

1.2 Industrial Significance of Tribology

Tribology is crucial to modern machinery that uses sliding and rolling surfaces. Examples of productive friction are brakes, clutches, driving wheels on trains and automobiles, bolts, and nuts. Examples of productive wear are writing with a pencil, machining, polishing, and shaving. Examples of unproductive friction and wear are internal combustion and aircraft engines, gears, cams, bearings, and seals.

According to some reports in 1966, losses resulting from ignorance of tribology in the United States were estimated to be about 4% of the gross national product (or about \$700 billion dollars per year in 2015). It also is estimated that approximately one-third of the world's energy resources in present use appear as friction in one form or another. Thus, the importance of friction reduction and wear control cannot be overemphasized for economic reasons and long-term reliability. According to Jost (1966, 1976), savings of about 1% of gross national product of an industrialized nation can be realized by research and better tribological practices. According to recent studies, expected savings are expected to be on the order of 50 times the research costs. The savings are both substantial and significant, and these savings can be obtained without the deployment of large capital investment.

The purpose of research in tribology is the minimization and elimination of losses resulting from friction and wear at all levels of technology where the rubbing of surfaces is involved. Research in tribology leads to greater manufacturing efficiency, better performance, fewer breakdowns, and significant savings.

Tribology is not only important to industry, it also affects day-to-day life (Bhushan 2013a, b). For example, writing is a tribological process. Writing is accomplished by a controlled transfer of lead (pencil) or ink (pen) to the paper. During writing with a pencil there should be good adhesion between the lead and paper so that a small quantity of lead transfers to the paper and the lead should have adequate toughness/hardness so that it does not fracture/break. The objective during shaving is to remove hair from the body as efficiently as possible with minimum discomfort to the skin. Shaving cream is used as a lubricant to minimize friction between a razor and the skin. Friction is helpful during walking and driving. Without adequate friction, we would slip and a car would skid! Tribology is also important in sports. For example, low friction between ski and snow is desirable during skiing. There should be low friction between fabric fibers and human skin.

Body joints need to be lubricated for low friction and wear to avoid osteoarthritis and joint replacement. The surface layer of cartilage present in the joint provides the bearing surface, and is lubricated with a joint fluid consisting of lubricin, hyaluronic acid (HA), and lipids. Hair conditioner coats hair in order to repair damage and lubricate it (Bhushan 2010). It contains silicone and fatty alcohols. Low friction and adhesion provide smooth feel in wet and dry environments, reduce friction between hair fibers during shaking and bouncing, and provide easy combing and styling. Skin creams and lotions are used to reduce friction between the fingers and body skin (Bhushan 2017b). Saliva and other mucous biofluids lubricate and facilitate transport of food and soft liquids through the body. The saliva in the mouth interacts with food and influences taste—mouth feel.

1.3 Origins and Significance of Nanotribology and Nanomechanics

At most interfaces of technological relevance, contact occurs at numerous asperities. Consequently, the importance of investigating a single-asperity contact in studies of the fundamental tribological and mechanical properties of surfaces has been long recognized. The proliferation of proximal probes, in particular tip-based microscopies (e.g., the scanning tunneling microscope and the atomic force microscope) and of computational techniques for simulating tip-surface interactions and interfacial properties, has allowed systematic investigations of interfacial problems with high resolution as well as ways and means for modifying and manipulating nanoscale structures. These advances have led to the development of the field of nanotribology, also known as molecular tribology or atomic-scale tribology (Bhushan et al. 1995; Bhushan 1997, 1999, 2001a, 2005, 2008a, b, 2013a, b, 2017a). This field is concerned with experimental and theoretical investigations of processes ranging from atomic and molecular scales to micro-scales occurring during adhesion, friction, wear, and thin-film lubrication at sliding surfaces.

The differences between the conventional or macrotribology and nanotribology are contrasted in Fig. 1.2. In macrotribology, tests are conducted on components with relatively large mass under heavily loaded conditions. In these tests, wear is inevitable and the bulk properties of

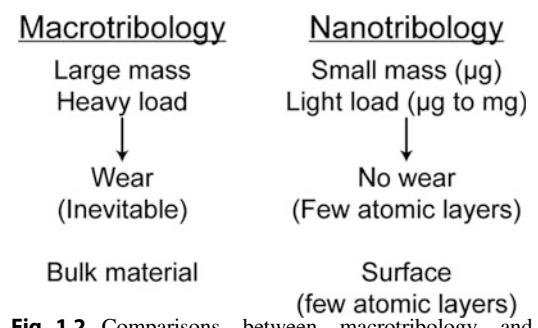


Fig. 1.2 Comparisons between macrotribology and nanotribology

mating components dominate the tribological performance. In nanotribology, measurements are made on components with at least one of the mating components having relatively small mass under lightly loaded conditions. In this situation, negligible wear occurs and the surface properties dominate the tribological performance.

Nanotribological studies are needed to develop a fundamental understanding of interfacial phenomena on a small scale, and to study interfacial phenomena of ultrathin films (as low as 1–2 nm) and micro/nanostructures, both used in magnetic storage systems, micro/nanoelectromechanical systems (MEMS/NEMS), and other industrial applications. The components used in micro/nanostructures are very light (on the order of few micrograms) and operate under very light loads (smaller than one microgram to a few milligrams). As a result, friction and wear (on a nanoscale) of lightly loaded micro/nanocomponents are highly dependent on the surface interactions (few atomic layers). These structures generally are lubricated with molecularly thin films. Nanotribological techniques are ideal to study the friction and wear processes of ultrathin micro/nanostructures, and although these are critical for this purpose, they are also valuable in a fundamental understanding of interfacial phenomena in macrostructures to provide a bridge between science and engineering.

Scanning probe microscopies (SPMs) and the surface force apparatus (SFA) are used widely for nanotribological studies (Bhushan et al. 1995, 2004; Bhushan 1997, 1999, 2001a, 2005, 2008a, b, 2017a). The most widely-used SPMs are the scanning tunneling microscope (STM) and the atomic force microscope (AFM). To give a historical perspective of the field, the scanning tunneling microscope (STM) developed by Drs. Gerd Binnig and Heinrich Rohrer and their colleagues in 1981 at the IBM Zurich Research Laboratory, Forschungslabor, was the first instrument capable of directly obtaining three-dimensional (3D) images of solid surfaces with atomic resolution (Binnig et al. 1982). STMs can only be used to study surfaces that are electrically conductive to some degree. Based on their design of STM, in 1985, Binnig et al. (1986, 1987)

developed an atomic force microscope (AFM) to measure ultrasmall forces (less than 1 μN) present between the AFM tip surface and the sample surface. AFMs can be used for measurement of all engineering surfaces that may be either electrically conducting or insulating. AFM has become a popular surface profiler for topographic measurements on micro- to nanoscale. AFMs modified to measure both normal and friction forces, generally called friction force microscopes (FFMs) or lateral force microscopes (LFMs), are used to measure friction on micro- and nanoscales. AFMs now are used for a larger number of physical and chemical characterizations in the field of nanotribology and beyond.

SFAs, first developed in 1968, are used to study both static and dynamic properties of the molecularly thin liquid films sandwiched between two molecularly smooth surfaces (Tabor and Winterton 1969; Bhushan 1997, 1999, 2001a, 2017a). However, the liquid under study has to be confined between molecularly-smooth optically-transparent or sometimes opaque surfaces with radii of curvature on the order of 1 mm (leading to poorer lateral resolution as compared to AFMs). Only AFMs/FFMs can be used to study engineering surfaces in the dry and wet conditions with atomic resolution.

Significant progress in understanding the fundamental nature of bonding and interactions in materials, combined with advances in computer-based modeling and simulation methods, have allowed theoretical studies of complex interfacial phenomena with high resolution in space and time (Bhushan et al. 1995; Bhushan 1997, 1999, 2001a, 2005, 2008a, b, 2017a). Such simulations provide insights into atomic-scale energetics, structure, dynamics, thermodynamics, transport, and rheological aspects of tribological processes. Furthermore, these theoretical approaches guide the interpretation of experimental data and the design of new experiments, and enable the prediction of new phenomena based on atomistic principles.

Mechanical properties are scale dependent. Hardness is known to increase with a decrease in the indentation depth (Bhushan et al. 1996). These should be measured at relevant length

scales. Ultrathin films as thin as on the order of a couple of nm are used in various applications and their mechanical properties need to be measured at very shallow indentation depths (Bhushan 1996, 2000, 2017a). Compared to their bulk material counterparts, nano-objects including nanoparticles, nanorods, and nanowires exhibit enhanced mechanical, electrical, magnetic, chemical, friction, and wear-reducing properties (Maharaj and Bhushan 2015). Structural integrity of various MEMS/NEMS is of paramount importance. Some applications require measurement of mechanical properties of these structures (Bhushan 1999, 2011, 2017a). The field of nanomechanics encompasses mechanical characterization on the nanoscale. SPMs and specialized nanoindentation techniques are used for nanomechanics studies (Bhushan et al. 1996; Bhushan 1999, 2011, 2017a; Bhushan and Li 2003).

1.4 Measurement Techniques

The popularity and ease of imaging at the molecular level and force measurements can be traced to the development of the SFA and the STM and AFM techniques (Bhushan 1999, 2011, 2017a). Schematics showing the principle behind these experimental tools are shown in Fig. 1.3.

The SFA enables the measurement of the normal forces between two curved molecularly smooth surfaces (such as mica and silica) immersed in liquid or vapor. Aside from normal forces, friction forces also can be measured at varying sliding speeds or oscillating frequencies with the use of attachments. STM and AFM techniques are used to obtain high-resolution, three-dimensional (3-D) images and to measure forces. These techniques rely on the use of a probe tip, which scans a specimen in order to generate a high-resolution image. In STM, a bias voltage is applied between a metallic probe and the sample, allowing for tunneling current flow. This tunneling current is monitored while the tip is scanned over the sample in order to generate a topographic image. In AFM, the force between the tip and the sample is used (instead of the

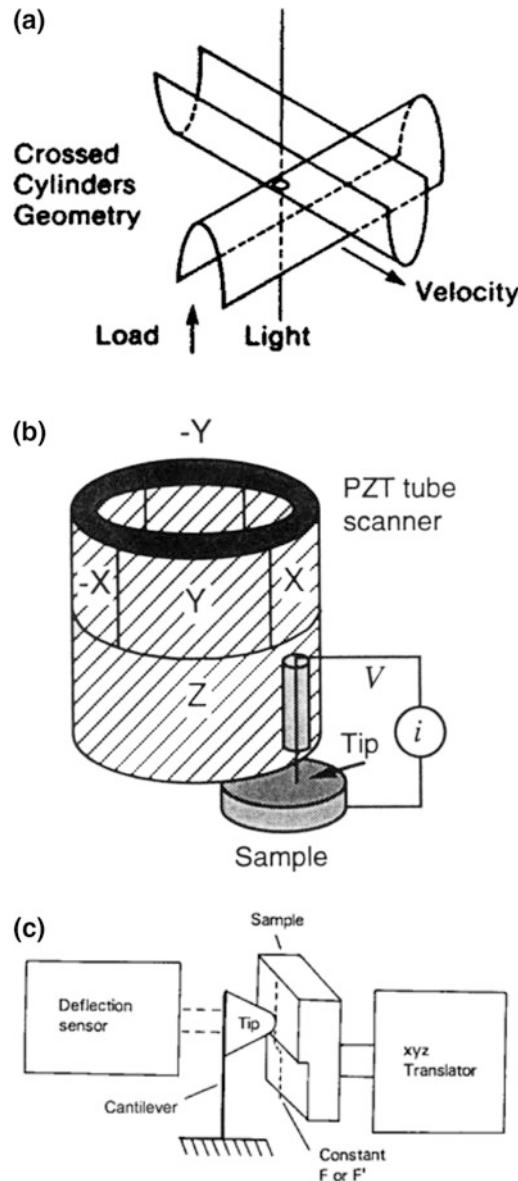


Fig. 1.3 Schematics of the **a** surface force apparatus, **b** scanning tunneling microscope, and **c** atomic force microscope

tunneling current) to produce an image. The tip can be in contact with the sample during imaging in AFM, whereas in STM, the tip is not in contact. As a consequence, the AFM can be used for a variety of physical measurements. In addition, while STMs can only be used to investigate surfaces that are electrically conductive to some degree, AFMs do not have this limitation, and

can be used to study any material type (Bhushan 1999, 2011, 2017a). Over the years, the use of AFMs has expanded beyond surface profiling to other capabilities such as the measurement of adhesion, friction, scratching, wear, lubrication, elastic/plastic mechanical properties, electrical, magnetic, and thermal properties, as well as *in situ* nanofeature creation and nanomanipulation (Bhushan et al. 1995; Bhushan 1999, 2011, 2013a, b, 2017a).

An introduction to SFAS, STMs, and AFMS, as well as various commercial instruments, is presented in this section.

1.4.1 Scanning Probe Microscopy

SPMs have been developed for various applications of scientific and industrial interest. These include—STM, AFM, FFM (or LFM), scanning electrostatic force microscopy (SEFM), scanning force acoustic microscopy (SFAM) [or atomic force acoustic microscopy (AFAM)], scanning magnetic microscopy (SMM) (or magnetic force microscopy (MFM)), scanning near field optical microscopy (SNOM), scanning thermal microscopy (SThM) scanning electrochemical microscopy (SEcM), scanning Kelvin Probe microscopy (SKPM), scanning chemical potential microscopy (SCPM), scanning ion conductance microscopy (SICM), and scanning capacitance microscopy (SCM). The family of instruments that measure forces (e.g. AFM, FFM, SEFM, SFAM, and SMM) are also referred to as scanning force microscopies (SFM). Although these instruments offer atomic resolution and are ideal for basic research, they also are used for cutting edge industrial applications that do not require atomic resolution.

STMs, AFMs, and their modifications can be used at extreme magnifications ranging from $10^3 \times$ to $10^9 \times$ in x, y, and z directions for imaging macro to atomic dimensions with high-resolution information and for spectroscopy. These instruments can be used in any environment such as ambient air, various gases, liquid, vacuum, and low and high temperatures. Imaging in liquid allows the study of live biological samples and eliminates

water capillary forces present in ambient air at the tip-sample interface. Low-temperature imaging is useful for the study of biological and organic materials and low-temperature phenomena such as superconductivity or charge-density waves. Low-temperature operation is also advantageous for high-sensitivity force mapping due to the reduction in thermal vibration. These instruments also have been used to image liquids such as liquid crystals and lubricant molecules on graphite surfaces. While the pure imaging capabilities of SPM techniques dominated the application of these methods at their early development stages, the physics and chemistry of probe-sample interactions and the quantitative analyses of tribological, electronic, magnetic, biological, and chemical surfaces are commonly carried out. Nanoscale science and technology are strongly driven by SPMs which allow investigation and manipulation of surfaces down to the atomic scale. With growing understanding of the underlying interaction mechanisms, SPMs have found applications in many fields outside basic research fields. In addition, various derivatives of all these methods have been developed for special applications, some of them targeting far beyond microscopy.

A detailed overview of scanning probe microscopy—principle of operation, instrumentation, and probes—is presented in a later chapter. (Also see, Bhushan 1997, 1999, 2001a, 2005, 2008a, b, 2011, 2017a; Bhushan et al. 2004.) A brief description of commercial STMs and AFMs follows.

1.4.1.1 Commercial STMs

There are a number of commercial STMs available on the market. Digital Instruments, Inc. (now Bruker Instruments, Santa Barbara, CA) introduced the first commercial STM, the Nanoscope I, in 1987. In a recent Nanoscope IV STM for operation in ambient air, the sample is held in position while a piezoelectric crystal in the form of a cylindrical tube (referred to as PZT tube scanner) scans the sharp metallic probe over the surface in a raster pattern, while sensing and outputting the tunneling current to the control station, Fig. 1.4. The digital signal processor

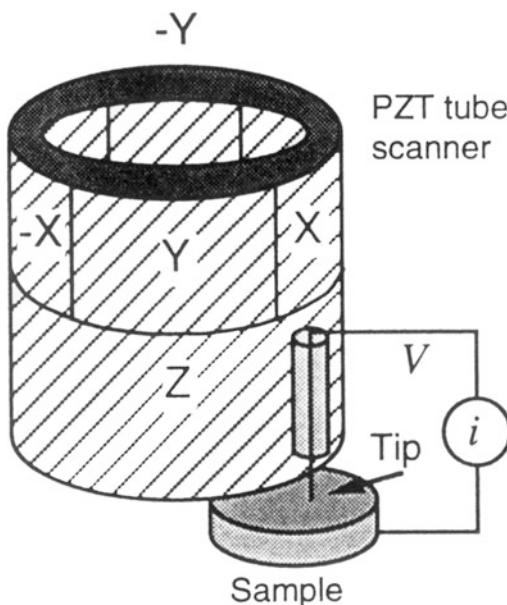


Fig. 1.4 Principle of operation of a commercial STM, a sharp tip attached to a piezoelectric tube scanner is scanned on a sample

(DSP) calculates the desired separation of the tip from the sample by sensing the tunneling current flowing between the sample and the tip. The bias voltage applied between the sample and the tip encourages the tunneling current to flow. The DSP completes the digital feedback loop by outputting the desired voltage to the piezoelectric tube. The STM operates in both the “constant height” and “constant current” modes depending on a parameter selection in the control panel. In the constant current mode, the feedback gains are set high, the tunneling tip closely tracks the sample surface, and the variation in the tip height required to maintain constant tunneling current is measured by the change in the voltage applied to the piezo tube. In the constant height mode, the feedback gains are set low, the tip remains at a nearly constant height as it sweeps over the sample surface, and the tunneling current is imaged.

Physically, the Nanoscope STM consists of three main parts: the head which houses the piezoelectric tube scanner for three dimensional motion of the tip and the preamplifier circuit (FET input amplifier) mounted on top of the head

for the tunneling current, the base on which the sample is mounted, and the base support, which supports the base and head (Bhushan 1999). The base accommodates samples up to 10 mm \times 20 mm and 10 mm in thickness. Scan sizes available for the STM are 0.7 μm (for atomic resolution), 12, 75, and 125 μm^2 .

The scanning head controls the three dimensional motion of tip. The removable head consists of a piezo tube scanner, about 12.7 mm in diameter, mounted into an invar shell used to minimize vertical thermal drifts because of good thermal match between the piezo tube and the Invar. The piezo tube has separate electrodes for X, Y, and Z which are driven by separate drive circuits. The electrode configuration (Fig. 1.4) provides x and y motions, which are perpendicular to each other, minimizes horizontal and vertical coupling, and provides good sensitivity. The vertical motion of the tube is controlled by the Z electrode that is driven by the feedback loop. The x and y scanning motions are each controlled by two electrodes that are driven by voltages of same magnitudes, but opposite signs. These electrodes are called $-Y$, $-X$, $+Y$, and $+X$. Applying complimentary voltages allows a short, stiff tube to provide a good scan range without large voltages. The motion of the tip due to external vibrations is proportional to the square of the ratio of vibration frequency to the resonant frequency of the tube. Therefore, to minimize the tip vibrations, the resonant frequencies of the tube are high about 60 kHz in the vertical direction and about 40 kHz in the horizontal direction. The tip holder is a stainless steel tube with a 300 μm inner diameter for 250 μm diameter tips, mounted in ceramic in order to keep the mass on the end of the tube low. The tip is mounted either on the front edge of the tube (to keep mounting mass low and resonant frequency high) (Fig. 1.4) or the center of the tube for large range scanners, namely 75 and 125 μm (to preserve the symmetry of the scanning.) This commercial STM accepts any tip with a 250 μm diameter shaft. The piezotube requires X-Y calibration, which is carried out by imaging an appropriate calibration standard. Cleaved graphite is used for the small-scan length head while

two dimensional grids (a gold plated ruling) can be used for longer range heads.

The Invar base holds the sample in position, supports the head, and provides coarse x-y motion for the sample. A spring-steel sample clip with two thumb screws holds the sample in place. An x-y translation stage built into the base allows the sample to be repositioned under the tip. Three precision screws arranged in a triangular pattern support the head and provide coarse and fine adjustment of the tip height. The base support consists of the base support ring and the motor housing. The stepper motor enclosed in the motor housing allows the tip to be engaged and withdrawn from the surface automatically.

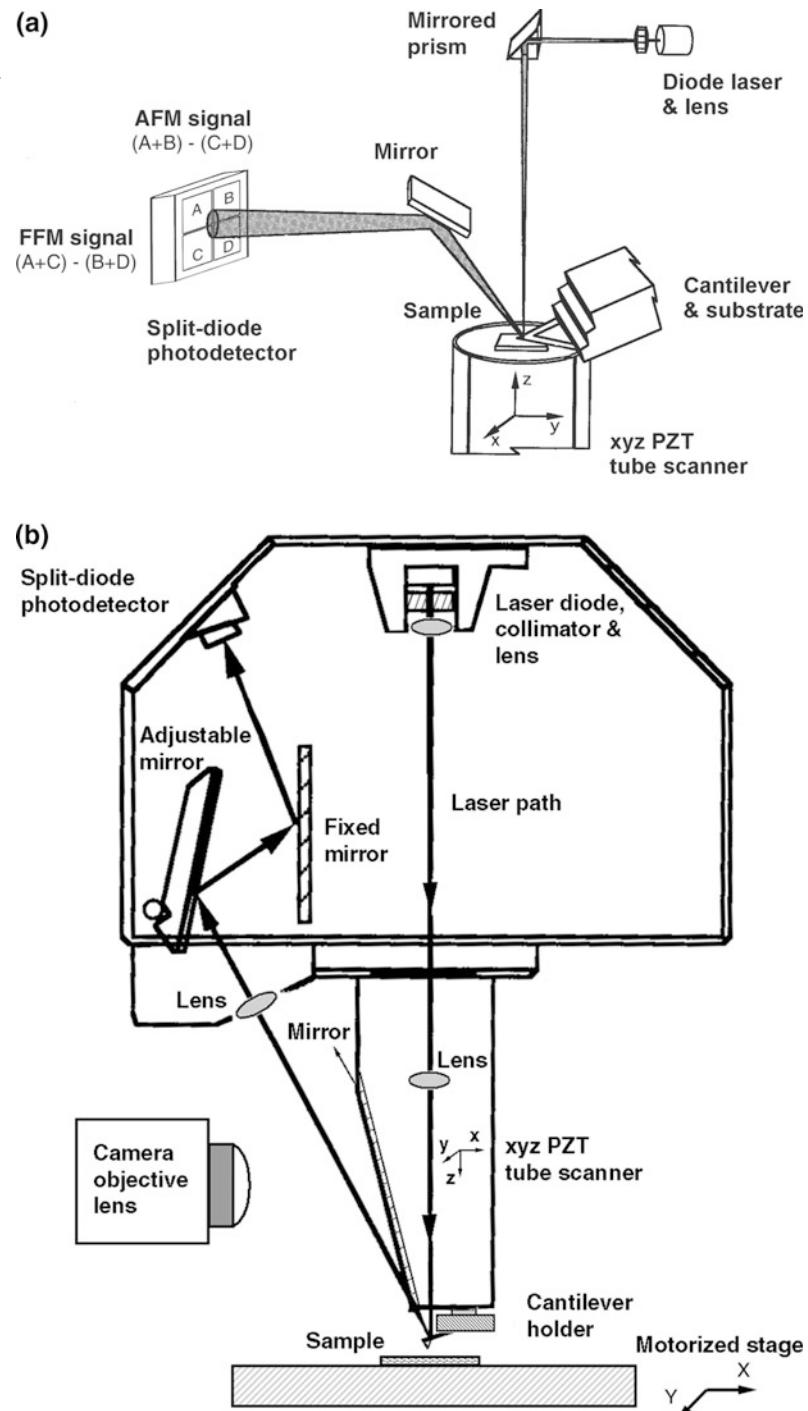
Samples to be imaged with STM must be conductive enough to allow a few nanoamperes of current to flow from the bias voltage source to the area to be scanned. In many cases, nonconductive samples can be coated with a thin layer of a conductive material to facilitate imaging. The bias voltage and the tunneling current depend on the sample. Usually, they are set at a standard value for engagement and fine-tuned to enhance the quality of the image. The scan size depends on the sample and the features of interest. Maximum scan rate of 122 Hz can be used. The maximum scan rate usually is related to the scan size. Scan rate above 10 Hz is used for small scans (typically 60 Hz for atomic-scale imaging with a 0.7 μm scanner). The scan rate should be lowered for large scans, especially if the sample surfaces are rough or contain large steps. Moving the tip quickly along the sample surface at high scan rates with large scan sizes will usually lead to a tip crash. Essentially, the scan rate should be inversely proportional to the scan size (typically 2–4 Hz for 1 μm , 0.5–1 Hz for 12 μm , and 0.2 Hz for 125 μm scan sizes). Scan rate in length/time, is equal to scan length divided by the scan rate in Hz. For example, for 10 $\mu\text{m} \times$ 10 μm scan size scanned at 0.5 Hz, the scan rate is 10 $\mu\text{m}/\text{s}$. The 256 \times 256 data formats are most commonly used. The lateral resolution at larger scans is approximately equal to scan length divided by 256.

1.4.1.2 Commercial AFM

A review of early designs of AFMs is presented by Bhushan (1999). There are a number of commercial AFMs available on the market. Major manufacturers of AFMs for use in the ambient environment are: Bruker Instruments, Santa Barbara, CA (formerly Digital Instruments Inc.; later Veeco Instruments, Inc.); Keysight Technologies, Santa Rosa, CA (formerly Molecular Imaging Corp., Phoenix, Arizona and later Agilent Technologies, Santa Clara, CA); Park Systems Corp, Suwon, South Korea; Asylum Research, Santa Barbara, CA; NT-M DT Spectrum Instruments, Moscow, Russia; Nanosurf AG, Liestal, Switzerland; Seiko Instruments, Japan; and Olympus, Japan. AFM/STMs for use in UHV and low temperature (LT) environments are primarily manufactured by Scienta Omicron GMBH, Taunusstein, Germany.

Here, we describe two commercial AFMs—a small sample and a large sample AFM—for operation in the contact mode produced by Bruker Instruments, Santa Barbara, CA, with scanning lengths ranging from about 0.7 μm (for atomic resolution) to about 125 μm (Alexander et al. 1989; Bhushan and Ruan 1994; Ruan and Bhushan 1994a, b). The original design of these AFMs comes from Meyer and Amer (1988). Basically, the AFM scans the sample in a raster pattern while outputting the cantilever deflection error signal to the control station. The cantilever deflection (or the force) is measured using laser deflection technique, Fig. 1.5. The DSP in the workstation controls the z position of the piezo based on the cantilever deflection error signal. The AFM operates in both the “constant height” and “constant force” modes. The DSP always adjusts the height of the sample under the tip based on the cantilever deflection error signal, but if the feedback gains are low, the piezo remains at a nearly “constant height” and the cantilever deflection data is collected. With the high gains, the piezo height changes to keep the cantilever deflection nearly constant (therefore the force is constant) and the change in piezo height is collected by the system.

Fig. 1.5 Principles of operation of **a** a commercial small sample AFM/FFM, and **b** a large sample AFM/FFM



To describe further the principle of operation of the commercial small sample AFM shown in Fig. 1.5a, the sample, generally no larger than

10 mm × 10 mm, is mounted on a PZT tube scanner which consists of separate electrodes to scan precisely the sample in the x-y plane in a

raster pattern and to move the sample in the vertical (z) direction. A sharp tip at the free end of a flexible cantilever is brought in contact with the sample. Features on the sample surface cause the cantilever to deflect in the vertical and lateral directions as the sample moves under the tip. A laser beam from a diode laser (5 mW max peak output at 670 nm) is directed by a prism onto the back of a cantilever near its free end, tilted downward at about 10° with respect to the horizontal plane. The reflected beam from the vertex of the cantilever is directed through a mirror onto a quad photodetector (split photodetector with four quadrants, commonly called position-sensitive detector or PSD, produced by Silicon Detector Corp., Camarillo, California). The differential signal from the top and bottom photodiodes provides the AFM signal which is a sensitive measure of the cantilever vertical deflection. Topographic features of the sample cause the tip to deflect in the vertical direction as the sample is scanned under the tip. This tip deflection will change the direction of the reflected laser beam, changing the intensity difference between the top and bottom sets of photodetectors (AFM signal). The height mode of operation for the AFM is used for topographic imaging or for any other operation in which the applied normal force is to be kept a constant. In this mode, a feedback circuit is used to modulate the voltage applied to the PZT scanner to adjust the height of the PZT, so that the cantilever vertical deflection (given by the intensity difference between the top and bottom detector) will remain constant during scanning. The PZT height variation is thus a direct measure of the surface roughness of the sample.

In a large sample AFM, both force sensors using optical deflection method and scanning unit are mounted on the microscope head, Fig. 1.5b. Because of vibrations added by cantilever movement, lateral resolution of this design is somewhat poorer than the design in Fig. 1.5a in which the sample is scanned instead of cantilever beam. The advantage of the large sample AFM is that large samples can be measured readily.

Most AFMs can be used for topography measurements in the so-called tapping mode (intermittent contact mode), also referred to as dynamic force microscopy. In the tapping mode, during scanning over the surface, the cantilever/tip assembly is sinusoidally vibrated by a piezo mounted above it, and the oscillating tip slightly taps the surface at the resonant frequency of the cantilever (70–400 Hz) with a constant (20–100 nm) oscillating amplitude introduced in the vertical direction with a feedback loop keeping the average normal force constant, Fig. 1.6. The oscillating amplitude is kept large enough so that the tip does not get stuck to the sample because of adhesive attractions. The tapping mode is used in topography measurements to minimize effects of friction and other lateral forces and/or to measure topography of soft surfaces.

Topographic measurements are made at any scanning angle. At a first glance, scanning angle may not appear to be an important parameter. However, the friction force between the tip and the sample will affect the topographic measurements in a parallel scan (scanning along the long

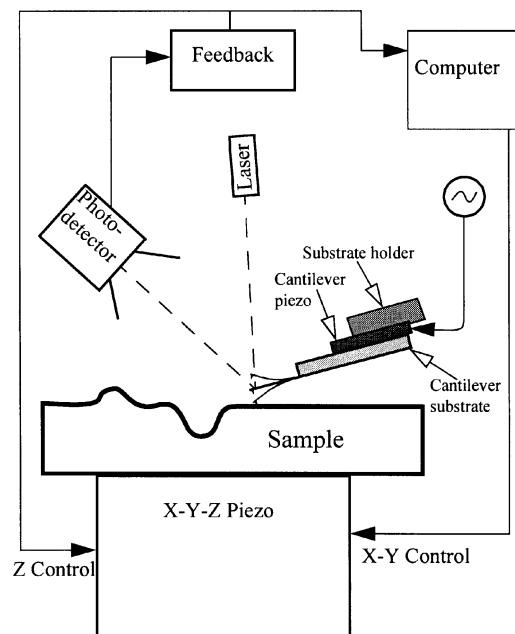


Fig. 1.6 Schematic of tapping mode used for surface roughness measurement

axis of the cantilever). Therefore a perpendicular scan may be more desirable. Generally, one picks a scanning angle that gives the same topographic data in both directions; this angle may be slightly different than that for the perpendicular scan.

For measurement of friction force being applied at the tip surface during sliding, left hand and right hand sets of quadrants of the photodetector are used. In the so-called friction mode, the sample is scanned back and forth in a direction orthogonal to the long axis of the cantilever beam. A friction force between the sample and the tip will produce a twisting of the cantilever. As a result, the laser beam will be reflected out of the plane defined by the incident beam and the beam reflected vertically from an untwisted cantilever. This produces an intensity difference of the laser beam received in the left hand and right hand sets of quadrants of the photodetector. The intensity difference between the two sets of detectors (FFM signal) is related directly to the degree of twisting and hence to the magnitude of the friction force. This method provides three-dimensional maps of friction force. One problem associated with this method is that any misalignment between the laser beam and the photodetector axis would introduce error in the measurement. However, by following the procedures developed by Ruan and Bhushan (1994a), in which the average FFM signal for the sample scanned in two opposite directions is subtracted from the friction profiles of each of the two scans, the misalignment effect is eliminated. By following the friction force calibration procedures developed by Ruan and Bhushan (1994a), voltages corresponding to friction forces can be converted to force unites. The coefficient of friction is obtained from the slope of friction force data measured as a function of normal loads typically ranging from 10 to 150 nN. This approach eliminates any contributions due to the adhesive forces (Bhushan et al. 1994). For calculation of the coefficient of friction based on a single point measurement, friction force should be divided by the sum of applied normal load and intrinsic adhesive force. Furthermore, it should be pointed out that for a single asperity contact,

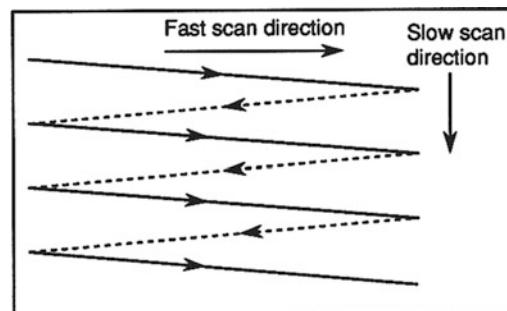


Fig. 1.7 Schematic of triangular pattern trajectory of the AFM tip as the sample is scanned in two dimensions. During imaging, data are recorded only during scans along the solid scan lines

the coefficient of friction is not independent of load.

The tip is scanned in such a way that its trajectory on the sample forms a triangular pattern, Fig. 1.7. Scanning speeds in the fast and slow scan directions depend on the scan area and scan frequency. Scan sizes ranging from less than $1 \text{ nm} \times 1 \text{ nm}$ to $125 \mu\text{m} \times 125 \mu\text{m}$ and scan rates from less than 0.5–122 Hz typically can be used. Higher scan rates are used for smaller scan lengths. For example, scan rates in the fast and slow scan directions for an area of $10 \mu\text{m} \times 10 \mu\text{m}$ scanned at 0.5 Hz are $10 \mu\text{m/s}$ and 20 nm/s , respectively.

1.4.2 Surface Force Apparatus (SFA)

SFAs are used to study both static and dynamic properties of the molecularly-thin liquid films sandwiched between two molecularly smooth surfaces. SFAs originally were developed by Tabor and Winterton (1969) and later by Israelachvili and Tabor (1972) to measure van der Waals forces between two mica surfaces as a function of separation in air or vacuum. Israelachvili and Adams (1978) developed a more advanced apparatus to measure normal forces between two surfaces immersed in a liquid so thin that their thickness approaches the dimensions of the liquid molecules themselves. A similar apparatus was also developed by Klein

(1980). SFAs, originally used in studies of adhesive and static interfacial forces, were first modified by Chan and Horn (1985) and later by Israelachvili et al. (1988) and Klein et al. (1991) to measure the dynamic shear (sliding) response of liquids confined between molecularly-smooth optically-transparent mica surfaces. Optically-transparent surfaces are required because the surface separation is measured using an optical interference technique. Van Alsten and Granick (1988) and Peachey et al. (1991) developed a new friction attachment which allow for the two surfaces to be sheared past each other at varying sliding speeds or oscillating frequencies, while simultaneously measuring both the friction force and normal force between them. Israelachvili (1989) and Luengo et al. (1997) also presented modified SFA designs for dynamic measurements including friction at oscillating frequencies. Because the mica surfaces are molecularly smooth, the actual area of contact is well defined and measurable, and asperity deformation does not complicate the analysis. During sliding experiments, the area of parallel surfaces is very large compared to the thickness of the sheared film and this provides an ideal condition for studying shear behavior because it permits one to study molecularly-thin liquid films whose thickness is well defined to the resolution of an angstrom. Molecularly thin liquid films cease to behave as a structural continuum with properties different from that of the bulk material (Van Alsten and Granick 1988, 1990a; Homola et al. 1989; Gee et al. 1990; Granick 1991).

Tonck et al. (1988) and Georges et al. (1993) developed a SFA used to measure the static and dynamic forces (in the normal direction) between a smooth fused borosilicate glass against a smooth and flat silicon wafer. They used a capacitance technique to measure surface separation, permitting the use of surfaces that are not optically transparent. Among others, metallic surfaces can be used at the interface. Georges et al. (1994) modified the original SFA so that a sphere can be moved towards and away from a plane and can be sheared at constant separation from the plane, for interfacial friction studies.

For a detailed review of various types of SFAs, see Israelachvili (1989, 1992), Horn (1990), and Homola (1993). SFAs based on their designs are commercially available from Sur-Force Corporation, Santa Barbara, California.

1.4.2.1 Israelachvili's and Granick's Design

The following review primarily is based on the papers by Israelachvili (1989) and Homola (1993). Israelachvili et al.'s design later followed by Granick et al. for oscillating shear studies, is used most commonly by researchers around the world.

Classical SFA Israelachvili and Adams (1978) developed the classical apparatus for measuring equilibrium or static intersurface forces in liquids and vapors. It consists of a small, air-tight stainless steel chamber in which two molecularly smooth curved mica surfaces can be translated towards or away from each other, see Fig. 1.8. The distance between the two surfaces also can be controlled independently to within ± 0.1 nm and the force sensitivity is about 10 nN. The technique utilizes two molecularly smooth mica sheets, each about 2 μm thick, coated with a semi reflecting 50–60 nm layer of pure silver. These are glued to rigid cylindrical silica disks of radius about 10 mm (silvered side down) mounted facing each other with their axes mutually at right angles (crossed cylinder position). This configuration is equivalent geometrically to a sphere contacting a flat surface. The adhesive glue which is used to affix the mica to the support is sufficiently compliant, so the mica will flatten under the action of adhesive forces or applied load to produce a contact zone in which the surfaces are locally parallel and planar. Outside of this contact zone, the separation between surfaces increases and the liquid, which is effectively in a bulk state, makes a negligible contribution to the overall response. The lower surface is supported on a cantilever spring that is used to push the two surfaces together with a known load. When the surfaces are forced into contact, they flatten elastically so that the contact zone is circular for duration of the static or

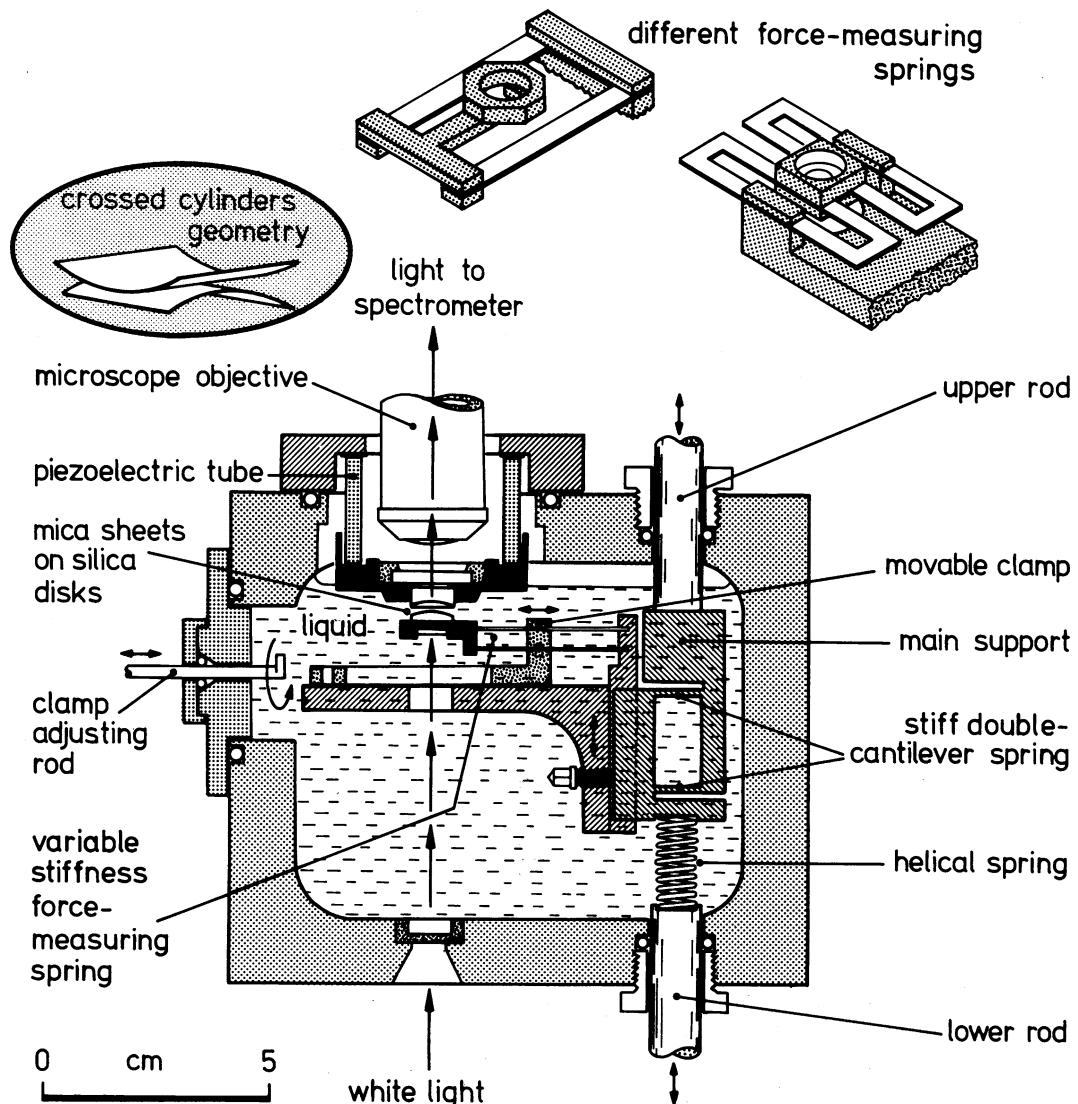
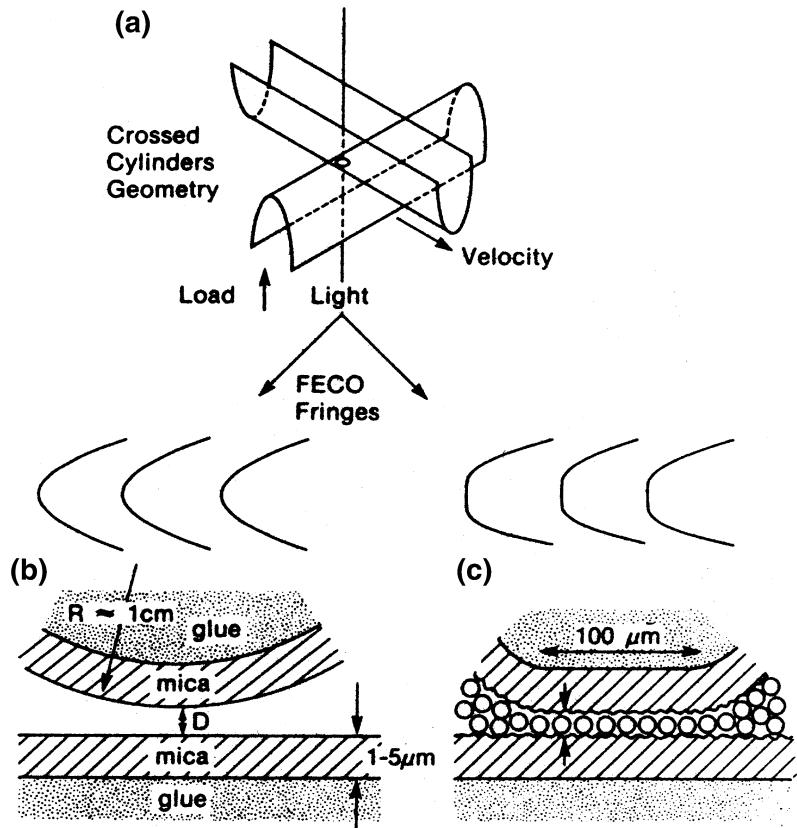


Fig. 1.8 Schematic of the surface force apparatus that employs the cross-cylinder geometry (Israelachvili and Adams 1978; Israelachvili 1989)

sliding interactions. The surface separation is measured using optical interference fringes of equal chromatic order (FECO) which enables the area of molecular contact and the surface separation to be measured to within 0.1 nm. For measurements, white light is passed vertically up through the two mica surfaces and the emerging beam is then focused onto the slit of a grating spectrometer. From the positions and shapes of the colored FECO fringes in the spectrogram, the

distance between the two surfaces and the exact shape of the two surfaces can be measured (see Fig. 1.9), as can the refractive index of the liquid (or material) between them. In particular, this allows for reasonably accurate determinations of the quantity of material deposited or adsorbed on the surfaces and the area of contact between two molecularly smooth surfaces. Any changes may be readily observed in both static and sliding conditions in real time (applicable to the design

Fig. 1.9 **a** Cross-cylinder configuration of mica sheet, showing formation of contact area. Schematic of the fringes of equal chromatic order (FECO) observed when two mica surfaces are **b** separated by distance D and **c** are flattened with a monolayer of liquid between them (Homola et al. 1990)



shown in Fig. 1.9) by monitoring the changing shapes of these fringes.

The distance between the two surfaces is controlled by use of a three-stage mechanism of increasing sensitivity. First, coarse control (upper rod) allows positioning of within about 1 μm . Then, the medium control (lower rod, which depresses the helical spring and that, in turn, bends the much stiffer double-cantilever spring by 1/1000 of this amount) allows positioning to about 1 nm. Finally, the piezoelectric crystal tube, which expands or contracts vertically by about 0.6 nm/V applied axially across the cylindrical wall, is used for final positioning to 0.1 nm.

The normal force is measured by expanding or contracting the piezoelectric crystal by a known amount and then measuring optically how much the two surfaces have actually moved. Any difference in the two values when multiplied by the stiffness of the force measuring spring gives

the force difference between the initial and final positions. In this way, both repulsive and attractive forces can be measured with a sensitivity of about 10 nN. The force measuring springs can be either single-cantilever or double-cantilever fixed-stiffness springs (as shown in Fig. 1.8), or the spring stiffness can be varied during an experiment (by up to a factor of 1000) by shifting the position of the dovetailed clamp using the adjusting rod. Other spring attachments, two of which are shown at the top of the figure, can replace the variable stiffness spring attachment (top right: nontilting nonshearing spring of fixed stiffness). Each of these springs are interchangeable and can be attached to the main support, allowing for greater versatility in measuring strong or weak and attractive or repulsive forces. Once the force F as a function of distance D is known for the two surfaces of radius R , the force between any other curved surfaces simply scales by R . Furthermore, the

adhesion energy (or surface or interfacial free energy) E per unit area between two flat surfaces is simply related to F by the so-called Derjaguin approximation (Israelachvili 1992) $E = F/2\pi R$. We note that the SFA is one of the few techniques available for directly measuring equilibrium force-laws (i.e., force versus distance at constant chemical potential of the surrounding solvent medium) (Israelachvili 1989). The SFA allows for both weak or strong and attractive or repulsive forces.

Most frequently, the molecularly smooth surface of mica is used in these measurements (Pashley 1981). However, silica (Horn et al. 1989) and sapphire (Horn and Israelachvili 1988) also have been used. It is also possible to deposit or coat each mica surface with metal films (Christenson 1988; Smith et al. 1988), carbon and metal oxides (Hirz et al. 1992), adsorbed polymer layers (Patel and Tirrell 1989), surfactant monolayers and bilayers (Christenson 1988; Israelachvili 1987, 1992; Israelachvili and McGuigan 1988). The range of liquids and vapors that can be used is almost endless.

Sliding Attachments for Tribological Studies

So far we have described a measurement technique that allows measurements of the normal forces between surfaces, that is, those occurring when two surfaces approach or separate from each other. However, in tribological situations, it is the transverse or shear forces that are of primary interest when two surfaces slide past each other. Essentially, there are two approaches used in studying the shear response of confined liquid films. In the first approach (constant velocity friction or steady-shear attachment), the friction is measured when one of the surfaces is moved at a constant speed over a distance of several hundreds of microns (Israelachvili et al. 1988; Gee et al. 1990; Homola 1989; Homola et al. 1990, 1991; Klein et al. 1991; Hirz et al. 1992). The second approach (oscillatory shear attachment) relies on the measurement of viscous dissipation and elasticity of confined liquids by using periodic sinusoidal oscillations over a range of amplitudes and frequencies (Van Alsten and Granick 1988, 1990a, b; Peachey et al. 1991; Hu et al. 1991).

For the constant velocity friction (steady-shear) experiments, the SFA was outfitted with a lateral sliding mechanism (Israelachvili et al. 1988; Israelachvili 1989; Gee et al. 1990; Homola 1989; Homola et al. 1990, 1991) allowing measurements of both normal and shearing forces (Fig. 1.10). The piezoelectric crystal tube mount supporting the upper silica disk of the basic apparatus shown in Fig. 1.8, is replaced. Lateral motion is initiated by a variable speed motor-driven micrometer screw that presses against the translation stage, which is connected via two horizontal double-cantilever strip springs to the rigid mounting plate. The translation stage also supports two vertical double-cantilever springs (Fig. 1.11) that at their lower end are connected to a steel plate supporting the upper silica disk. One of the vertical springs acts as a frictional force detector by having four resistance strain gauges attached to it, forming the four arms of a Wheatstone bridge and electrically connected to a chart recorder. Thus, by rotating the micrometer, the translation stage deflects, causing the upper surface to move horizontally and linearly at a steady rate. If the upper mica surface experiences a transverse frictional or viscous shearing force, this will cause the vertical springs to deflect, and this deflection can be measured by the strain gauges. The main support, force-measuring double-cantilever spring, movable clamp, white light, etc., are all parts of the original basic apparatus (Fig. 1.8), whose functions are to control the surface separation, vary the externally applied normal load, and measure the separation and normal force between the two surfaces, as already described. Note that during sliding, the distance between the surfaces, their true molecular contact area, their elastic deformation, and their lateral motion can all be monitored simultaneously by recording the moving FECO fringe pattern using a video camera (Gee et al. 1990).

The two surfaces can be sheared past each other at sliding speeds that can be varied continuously from 0.1 to 20 $\mu\text{m/s}$ while simultaneously measuring both the transverse (frictional) force and the normal (compressive or tensile) force between them. The lateral distances

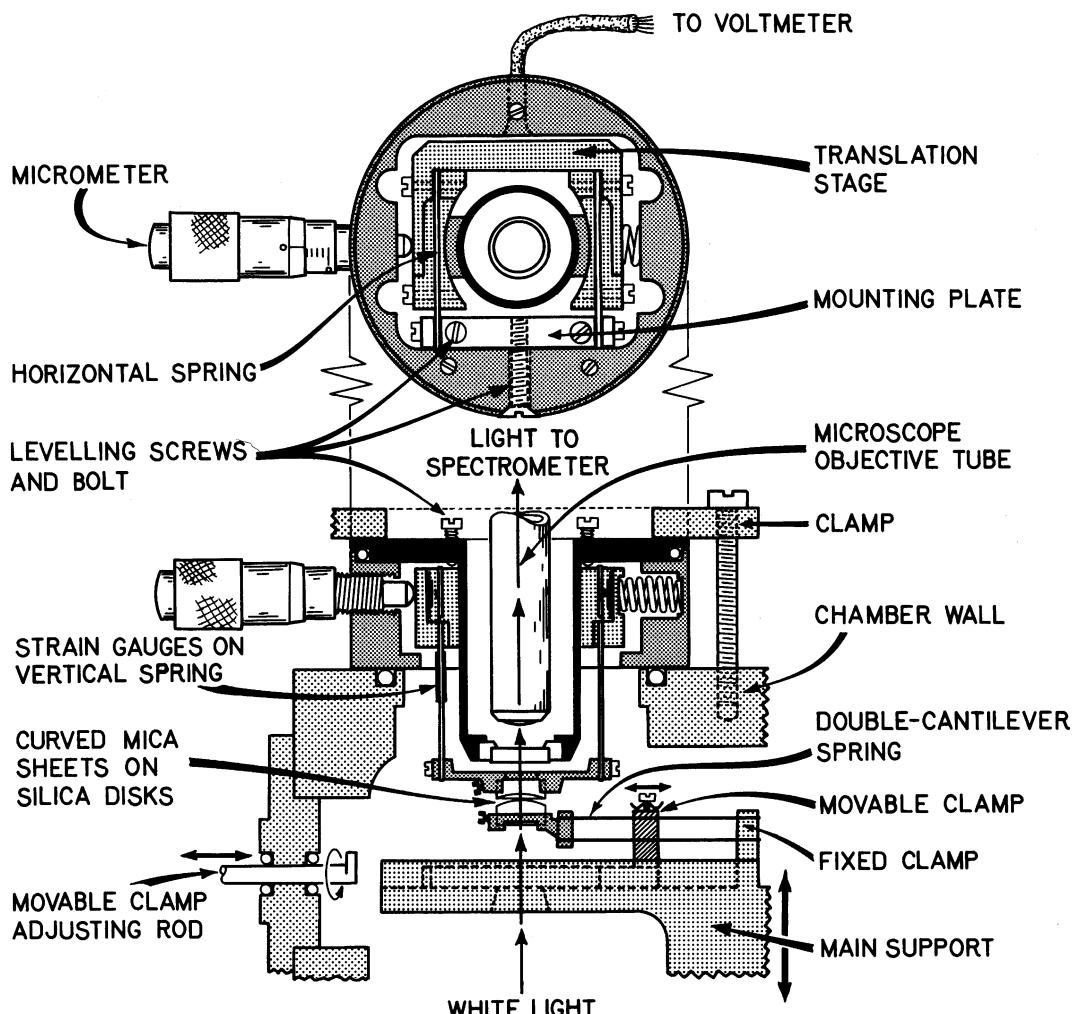


Fig. 1.10 Schematic of shear force apparatus. Lateral motion is initiated by a variable speed motor-driven micrometer screw that presses against the translation stage which is connected through two horizontal double-cantilever strip springs to the rigid mounting plate (Israelachvili et al. 1988; Israelachvili 1989)

traversed are on the order of a several hundreds of micrometers which correspond to several diameters of the contact zone.

With an oscillatory shear attachment, developed by Granick et al., viscous dissipation and elasticity and dynamic viscosity of confined liquids by applying periodic sinusoidal oscillations of one surface with respect to the other can be studied (Van Alsten and Granick 1988, 1990a, b; Peachey et al. 1991; Hu et al. 1991). This attachment allows for the two surfaces to be sheared past each other at varying sliding speeds

or oscillating frequencies, while simultaneously measuring both the transverse (friction or shear) force and the normal load between them. The externally applied load can be varied continuously, and both positive and negative loads can be applied. Finally the distance between the surfaces, their true molecular contact area, their elastic (or viscoelastic) deformation, and their lateral motion can all be monitored simultaneously by recording the moving interference fringe pattern using a video camera-recorder system.

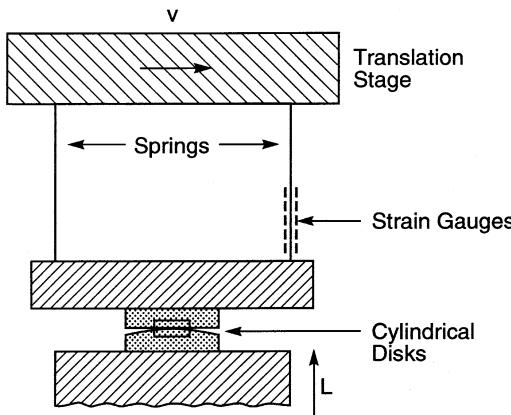


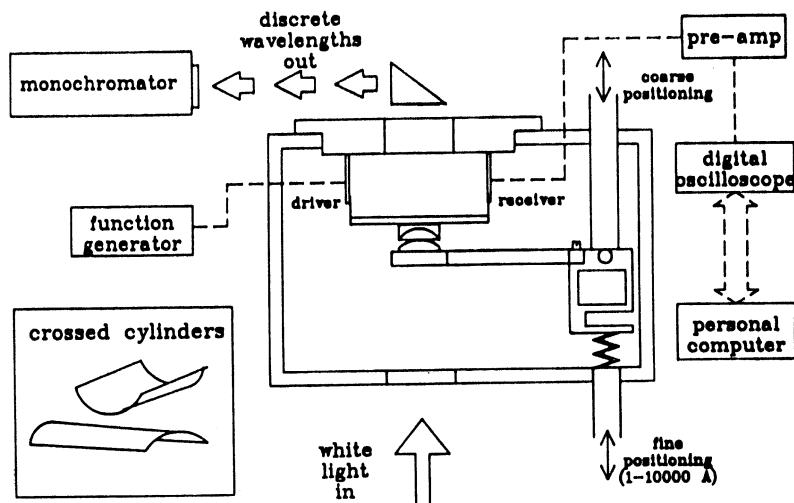
Fig. 1.11 Schematic of the sliding attachment. The translation stage also supports two vertical double-cantilever springs, which at their lower end are connected to a steel plate supporting the upper silica disk (Gee et al. 1990)

To produce shear while maintaining constant film thickness or constant separation of the surfaces, the top mica surface is suspended from the upper portion of the apparatus by two piezoelectric bimorphs. A schematic description of the surface force apparatus with the installed shearing device is shown in Fig. 1.12 (Van Alsten and Granick 1988, 1990a, b; Peachey et al. 1991; Hu et al. 1991). Israelachvili (1989) and Luengo et al. (1997) have also presented similar designs. The lower mica surface, as in the steady-shear sliding attachment, is stationary and sits at the tip of a double cantilever spring attached at the other

end to a stiff support. The externally applied load can be varied continuously by displacing the lower surface vertically. An AC voltage difference applied by a signal generator (driver) across one of the bimorphs tends to bend it in oscillatory fashion, while the frictional force resists that motion. Any resistance to sliding induces an output voltage across the other bimorphs (receiver), which can be easily measured by a digital oscilloscope. The sensitivity in measuring force is on the order of a few μN and the amplitudes of measured lateral displacement can range from a few nm to 10 μm .

The design is flexible and allows induction of time-varying stresses with different characteristic wave shapes simply by changing the wave form of the input electrical signal. For example, when measuring the apparent viscosity, a sine wave input is convenient to apply. Figure 1.13a shows an example of the raw data, obtained with a hexadecane film at a moderate pressure, when a sine wave was applied to one of the bimorphs (Van Alsten and Granick 1990b). By comparing the calibration curve with the response curve, which was attenuated in amplitude and lagged in phase, an apparent dynamic viscosity can be inferred. On the other hand, a triangular waveform is more suitable when studying the yield stress behavior of solid-like films, as in Fig. 1.13b. The triangular waveform, showing a linear increase and decrease of the applied force

Fig. 1.12 Schematic of the oscillatory shearing apparatus (Van Alsten and Granick 1988)



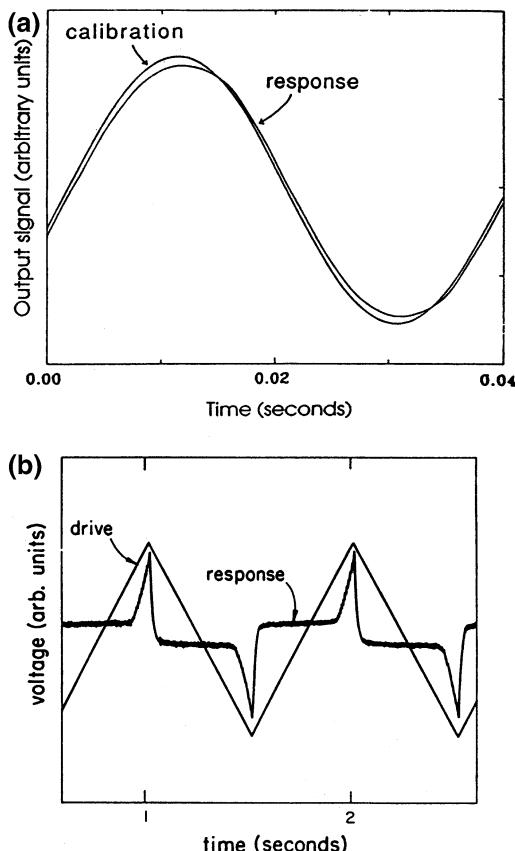


Fig. 1.13 **a** Two output signals induced by an applied sine wave (not shown) are displaced. The “calibration” waveform is obtained with the mica sheets completely separated. The response waveform is obtained with a thin liquid film between the sheets, which causes it to lag the calibration waveform, **b** the oscilloscope trace of the drive and response voltages used to determine critical shear stress. The drive waveform shows voltage proportional to induced stress on the sheared film and the response waveform shows voltage proportional to resulting velocity. Spikes in the response curve correspond to the stick-slip event (Van Alsten and Granick 1990b)

with time, is proportional to the driving force acting on the upper surface. The response waveform, which represents a resistance of the interface to shear, remains very small indicating that the surfaces are in a stationary contact with respect to each other until the applied stress reaches a yield point. At the yield point the slope of the response curve increases dramatically, indicating the onset of sliding.

Homola (1993) compared the two approaches—steady shear attachment and oscillating shear attachment. In experiments conducted by Israelachvili and his co-workers, the steady-shear attachment was employed to focus on the dynamic frictional behavior of the film after a sufficiently high shear stress was applied to exceed the yield stress and produce sliding at a constant velocity. In these measurements, the film was subjected to a constant shearing force for a time sufficiently long to allow them to reach a dynamic equilibrium, i.e., the molecules within the film had enough time to order and align with respect to the surface, both normally and tangentially. Under these conditions, dynamic friction was observed to be “quantized” according to the number of liquid layers between the solid surfaces and independent of the shear rate (Israelachvili et al. 1988). Clearly, in this approach, the molecular ordering is optimized by a steady shear, which imposes a preferred orientation on the molecules in the direction of shear.

The above mode of sliding is particularly important when the sheared film is made of a long chain lubricant molecules requiring a significantly long sliding time to order and align and even a longer time to relax (disorder) when sliding stops. This suggests that a steady-state friction is realized only when the duration of sliding exceeds the time required for an ensemble of the molecules to fully order in a specific shear field. It also suggests that static friction should depend critically on the sliding time and the extent of the shear induced ordering (Homola 1993).

In contrast, the oscillatory shear method, which utilizes periodic sinusoidal oscillations over a range of amplitudes and frequencies, addresses a response of the system to rapidly varying strain rates and directions of sliding. Under these conditions, the molecules, especially those exhibiting a solid-like behavior, cannot respond sufficiently fast to stress and are unable to order fully during the duration of a single pass, i.e., their dynamic and static behavior reflects an oscillatory shear-induced ordering that might or

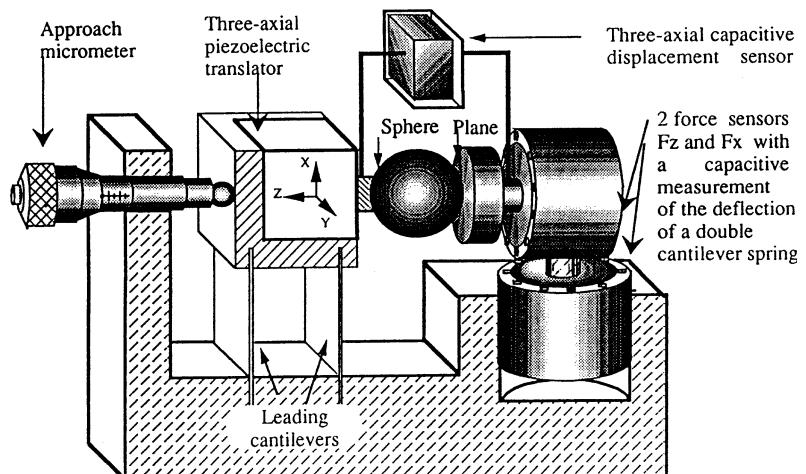
might not represent an equilibrium dynamic state. Thus, the response of the sheared film will depend critically on the conditions of shearing, i.e., the strain, the pressure, and the sliding conditions (amplitude and frequency of oscillations), which in turn will determine a degree of molecular ordering. This may explain the fact that the layer structure and “quantization” of the dynamic and static friction were not observed in these experiments, in contrast to results obtained when velocity was kept constant. Intuitively, this behavior is expected, considering that the shear-ordering tendency of the system frequently is disturbed by a shearing force of varying magnitude and direction. Nonetheless, the technique is capable of providing an invaluable insight into the shear behavior of molecularly thin films subjected to non-linear stresses, as is encountered frequently in practical applications. This is especially true under conditions of boundary lubrication where interacting surface asperities will be subjected to periodic stresses of varying magnitudes and frequencies (Homola 1993).

1.4.2.2 Georges et al.'s Design

The SFA developed by Tonck et al. (1988) and Georges et al. (1993) is used to measure static and dynamic forces in the normal direction between surfaces in close proximity. Georges et al. (1994) modified their original SFA to measure friction

forces. In this apparatus, in addition to having the sphere move normal to the plane, the sphere can be sheared at constant separation from the plane. The shear force apparatus is shown in Fig. 1.14. In this apparatus, a drop of liquid is introduced between a macroscopic spherical body and a plane. The sphere is moved towards and away from a plane using the expansion and the vibration of a piezoelectric crystal. Piezoelectric crystal is vibrated at low amplitude around an average separation for dynamic measurements to provide dynamic function of the interface. The plane specimen is supported by a double-cantilever spring. A capacitance sensor measures the elastic deformation of the cantilever and thus the force transmitted through the liquid to the plane. A second capacitance sensor is designed to measure the relative displacement between the supports of the two solids. The reference displacement signal is the sum of two signals: first, a ramp provides a constant normal speed from 50 to 0.01 nm/s, and second, the piezoelectric crystal is designed to provide a small sinusoidal motion in order to determine the dynamic behavior of sphere-plane interactions. A third capacitance sensor measures the electrical capacitance between the sphere and the plane. In all cases, the capacitance is determined by incorporating the signal of an oscillator in the inductive-capacitance (L-C) resonant-input stage of an oscillator to give a signal-dependent frequency in the range of 5–

Fig. 1.14 Schematic of shear force apparatus (Georges et al. 1994)



12 MHz. The resulting fluctuations in oscillation frequency are detected using a low-noise frequency discriminator. Simultaneous measurements of sphere-plane displacement, surface force, and the damping of the interface allows an analysis of all regimes of the interface (Georges et al. 1993). Loubet et al. (1993) used SFA in the crossed-cylinder geometry using two freshly-cleaved mica sheets similar to the manner used by Israelachvili and coworkers.

For shear experiments, three piezoelectric elements controlled by three capacitance sensors permit accurate motion control and force measurement along three orthogonal axes with displacement sensitivity of 10^{-3} nm and force sensitivity of 10^{-8} N. Adhesion and normal deformation experiments are conducted in the normal approach (z-axis). Friction experiments are conducted by introducing displacement in the X-direction at a constant normal force. In one of the experiment, Georges et al. (1994) used 2.95-mm dia. sphere made of cobalt-coated fused borosilicate glass and a silicon wafer for the plane.

1.4.3 Vibration Isolation

STM, AFM, and SFA should be isolated from sources of vibration in the acoustic and sub-acoustic frequencies, especially for atomic-scale measurements. Vibration isolation is provided generally by placing the instrument on a vibration isolation air table. For further isolation, the instrument should be placed on a pad of soft silicone. A cheaper alternative consists of a large mass of 100 or more N, suspended from elastic “bungee” cords. The mass should stretch the cords at least 0.3 m, but not so much that the cords reach their elastic limit. The instrument should be placed on the large mass. The system, including the microscope, should have a natural frequency of about 1 Hz or less both vertically and horizontally. Natural frequency can be tested by gently pushing on the mass and measure the rate at which its swings or bounces.

1.5 Magnetic Storage Devices and MEMS/NEMS and BioMEMS/NEMS

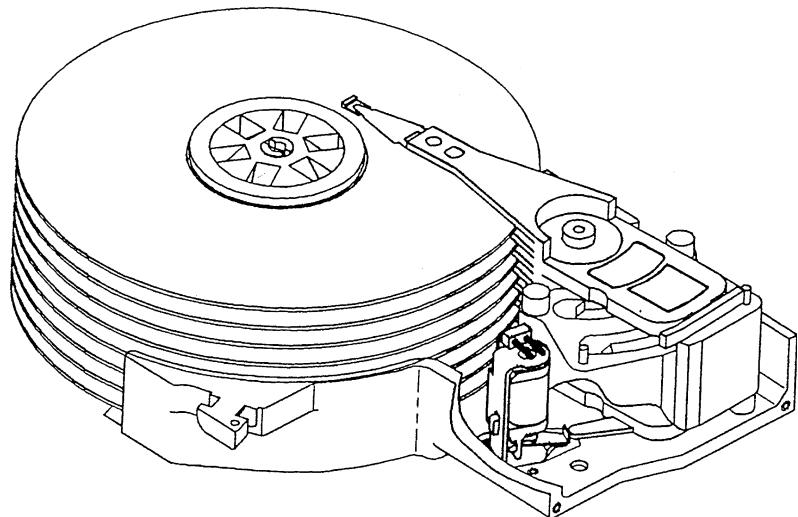
The interest in micro/nanotribology field grew from magnetic storage devices and later the applicability to the general field of micro/nanoelectromechanical systems (MEMS/NEMS) became clear. A few examples of magnetic storage devices and MEMS/NEMS and BioMEMS/NEMS are presented where micro/nanotribological and micro/nanomechanical tools and techniques are essential for interfacial studies.

1.5.1 Magnetic Storage Devices

Magnetic storage devices used for storage and retrieval are tape and rigid disk drives. These devices primarily are used for data storage applications. Tape drives are used for backup and mass storage given their highest volumetric density and lowest cost per unit storage. Rigid disk drives are used for online access and storage. Magnetic recording and playback involves the relative motion between a magnetic medium (tape or disk) against a read-write magnetic head. Heads are designed so that they develop a load-carrying hydrodynamic air film under steady operating conditions to minimize head-medium contact. However, physical contact between the medium and head occurs during starts and stops, referred to as contact-start-stops (CSS) (Bhushan 1996, 2000, 2001b; Castillo and Bhushan 2006). In modern magnetic storage devices, the flying heights (head-to-medium separation) are on the order of a couple of nm for rigid disk drives and on the order of 2–20 nm for tape drives. Roughnesses of the head and medium surfaces are on the order of 1–2 nm RMS. The need for ever-increasing recording densities requires that surfaces be as smooth as possible and the flying heights be as low as possible. High stiction (static friction) and wear are the limiting factors of this technology.

Figure 1.15 shows the schematic of a data processing rigid disk drive available with various

Fig. 1.15 Schematic of a data-processing magnetic rigid disk drive



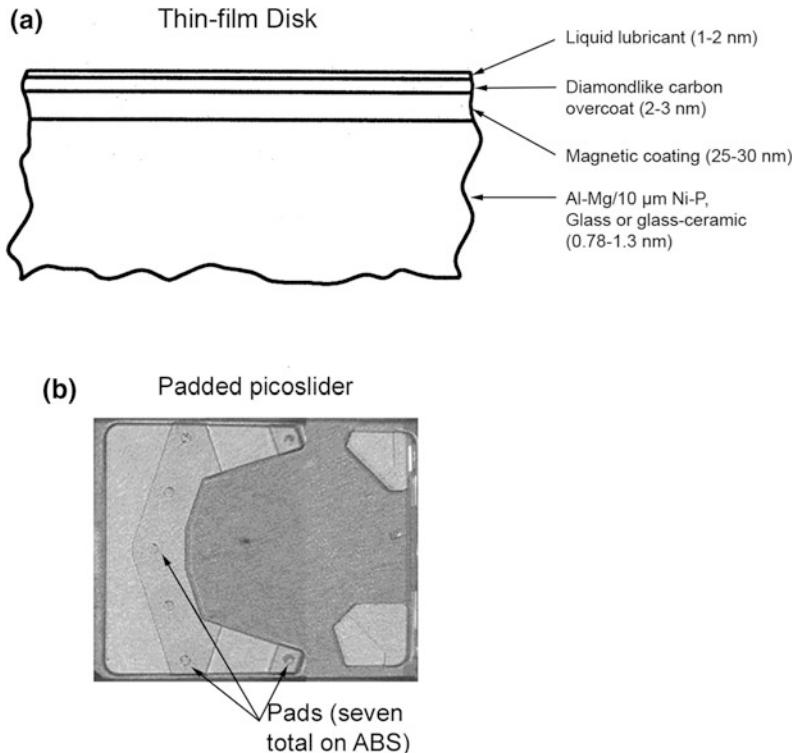
form factors (1.6-, 27.4-, 48-, 63.5-, 75-, and 95-mm). Nonremovable stacks of multiple disks mounted on ball bearings or hydrodynamic spindles, are rotated by an electric motor at constant angular speed ranging from about 5000 to in excess of 15,000 RPM, dependent upon the disk size. A head slider-suspension assembly (allowing one slider for each disk surface) is actuated by a stepper motor or a voice coil motor using a rotary actuator. Figure 1.16a shows the section view of a thin-film rigid disk. The substrate for rigid disks generally is a non-heat-treatable aluminum-magnesium alloy 5086, glass or glass ceramic. The protective overcoat commonly used for thin-film disks is sputtered diamondlike carbon (DLC), topically lubricated with perfluoropolyether-type of lubricants. Lubricants with polar-end groups generally are used for thin-film disks in order to provide partial chemical bonding to the overcoat surface. The disks used for CSS technology are laser textured in the landing zone. Figure 1.16b shows the schematic of a thin-film head picoslider with a step at the leading edge, and GMR read and inductive write. “Pico” refers to the small sizes of $1.25\text{ mm} \times 1\text{ mm}$. These sliders use $\text{Al}_2\text{O}_3\text{-TiC}$ (70–30 wt%) as the substrate material, with multilayered thin-film head structures coated with about a 3.5-nm thick DLC coating to prevent the thin-film structure from electrostatic discharge. The seven pads on the

padded slider are made of DLC, and are about 40 μm in diameter and 50 nm in height. A normal load of about 3 g is applied during use.

1.5.2 MEMS/NEMS and BioMEMS/NEMS

The advances in silicon photolithographic process technology led to the development of MEMS in the early 1980s (Bhushan 2017a). Later, lithographic and nonlithographic processes were developed to process nonsilicon (plastics or ceramics) materials. MEMS for mechanical applications include acceleration, pressure, flow, gas sensors, linear and rotary actuators, and other microstructures of microcomponents such as electric motors, gear trains, gas turbine engines, nozzles, fluid pumps, fluid valves, switches, grippers, and tweezers. MEMS for chemical applications include chemical sensors and various analytical instruments. Microoptoelectromechanical systems (or MOEMS) include micromirror arrays and fiber optic connectors. Radio frequency MEMS or RF-MEMS include inductors, capacitors, and antennas. High-aspect ratio MEMS (HARMEMS) have also been introduced. BioMEMS include biosensors, which include biofluidic chips (microfluidic chips) for chemical and biochemical analyses and a variety of biosensors in medical diagnostics, e.g., DNA,

Fig. 1.16 **a** Section views of a thin-film magnetic rigid disk, and **b** schematic of a padded picoslider used for CSS



RNA, proteins, cells, blood pressure and assays, and toxin identification, as well as implantable drug delivery devices. Killer applications include capacitive-type silicon accelerometers for automotive sensory applications, digital micromirror devices for projection displays, and a variety of biosensors. Any component requiring relative motions needs to be optimized for stiction and wear and interfacial adhesion is needed when depositing various films (including biological

molecular layers) on substrates (Bhushan 1997, 1998, 2001a, b, 2016, 2017a).

Figure 1.17 also shows two digital micromirror device (DMD) pixels used in digital light processing (DLP) technology for digital projection displays in portable and home theater projectors as well as table top and projection TVs (Hornbeck and Nelson 1988; Hornbeck 1999; Bhushan 2017a). The entire array (chip set) consists of a large number of rotatable aluminum

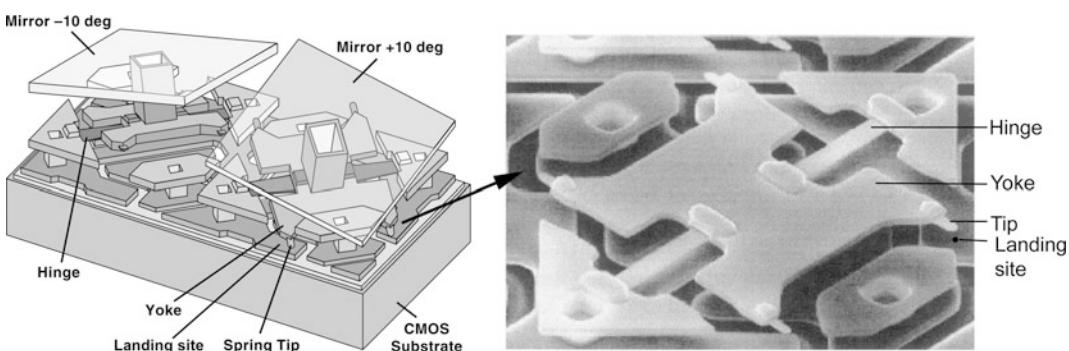


Fig. 1.17 Digital micromirror devices for projection displays (Hornbeck 1999)

micromirrors (digital light switches), which are fabricated on top of a CMOS static random access memory integrated circuit. The surface micromachined array consists of half of a million to more than two million of these independently controlled reflective, micromirrors (mirror size on the order of 14 μm square and 15 μm pitch), which flip backward and forward at a frequency of on the order of 5000 times a second. For the binary operation, micromirror/yoke structure mounted on torsional hinges is rotated $\pm 10^\circ$ (with respect to the horizontal plane) as a result of electrostatic attraction between the micromirror structure and the underlying memory cell, and is limited by a mechanical stop. Contact between cantilevered spring tips at the end of the yoke (four present on each yoke) with the underlying stationary landing sites is required for true digital (binary) operation. Stiction and wear during a contact between aluminum alloy spring tips and landing sites, hinge memory (metal creep at high operating temperatures), hinge fatigue, shock and vibration failure, and sensitivity to particles in the chip package and operating environment are some of the important issues affecting the reliable operation of a micromirror device (Liu and Bhushan 2004a, b, c). Perfluorodecanoic acid (PFDA) self-assembled monolayers are used on the tip and landing sites to reduce stiction and wear. The spring tip is used in order to use the spring stored energy to pop up the tip during pull-off. A lifetime estimate of over 100,000 operating hours with no degradation in image quality is the norm.

A biosensor based on a field-effect transistor (FET) is shown in Fig. 1.18a (Bhushan et al. 2005). FETs are sensitive to the electrical field produced by the charge at the surface of the gate insulator. In this sensor, the gate metal electrode of a metal-oxide semiconductor field effect transistor (MOSFET) is removed and replaced with a protein (receptor layer) whose cognate is the analyte (e.g., virus or bacteria) that is meant to be sensed. Various proteins may have 1–25 (positive or negative) charges per molecule. The binding of the receptor layer with the analyte produces a change in the effective charge, creating a change in the electrical field. This

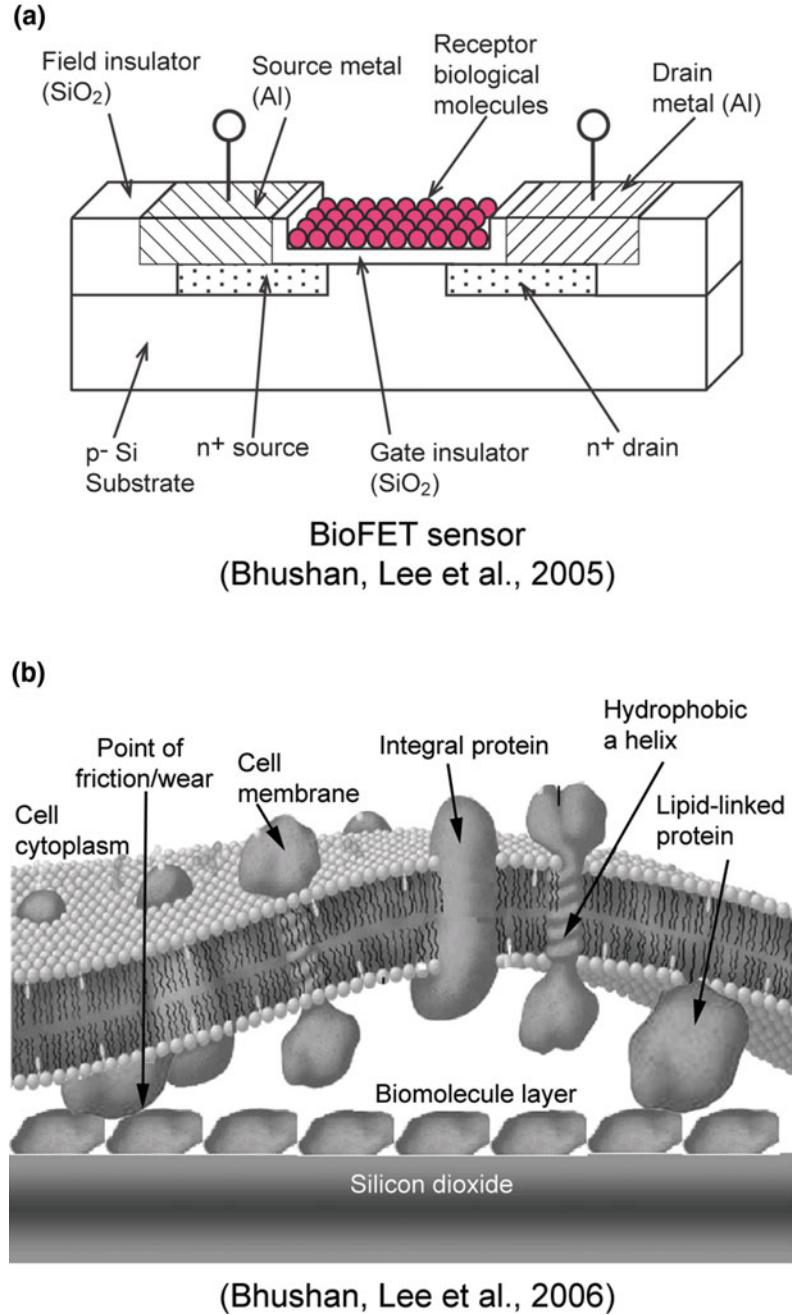
electrical field change may produce a measurable change in the current flow through the device. Adhesion between the protein layer and silica substrate affects the reliability of the biosensor. In the case of implanted biosensors, the biosensors come into contact with the exterior environment, such as tissues and fluids, and any relative motion of the sensor surface with respect to that exterior environment may result in surface damage. A schematic of the generation of friction and wear points when an implanted biosensor surface comes in contact with a living tissue is shown in Fig. 1.18b (Bhushan et al. 2005, 2006; Bhushan 2017a).

NEMS and BioNEMS are produced by nanomachining in a top-down or, more commonly, a bottom-up approach (from small to large) largely relying on nanochemistry (Bhushan 2017a). Examples of NEMS include nanocomponents, nanodevices, nanosystems, and nanomaterials such as microcantilever with integrated sharp nanotips for STM and AFM, quantum corral formed using STM by placing atoms one by one, AFM cantilever arrays (Millipede) for data storage, molecularly-thick films (e.g., in giant magnetoresistive or GMR heads and magnetic media), STM and AFM tips for nanolithography, dip-pen nanolithography for printing molecules, nanowires, carbon nanotubes, quantum wires, quantum boxes, quantum transistors, and molecular gears. Examples of BioNEMS include nanotube-based biosensors, biological (DNA) motors, and nanoparticles (e.g., nanomagnetic particles in magnetic media and targeted drug delivery) (Bhushan 2017a).

Figure 1.19 shows a single-walled carbon nanotube (SWNT) biosensor (Chen et al. 2004). The conductance of a carbon nanotube (CNT) device changes when proteins adsorb on the surface. The change in electrical resistance is a measure of protein adsorption. For high performance, adhesion should be strong between the adsorbent and the SWNT.

Figure 1.20 shows a conceptual model of a targeted intravascular drug delivery device using nanoparticles to search-and-destroy disease (tumor) cells (Martin and Grove 2001). The tumor cells have a 1–2 orders of magnitude higher

Fig. 1.18 **a** Schematic of a bioFET sensor (Bhushan et al. 2005), and **b** schematic showing the generation of friction and wear points due to interaction of implanted biomolecule layer on a biosensor with living tissue (Bhushan et al. 2006)



density of receptors than the normal cells and a lower pH. Some receptors are only expressed on tumor cells. Nanoparticles made of various materials including gold (Au), graphene, iron oxide, polymer, and silica have been studied in targeted drug delivery for cancer treatment. With lateral dimensions of 1 μm or less, the particles

are smaller than any blood cells. These particles can be injected into the blood stream and travel freely through the circulatory system. In order to direct these drug delivery nanoparticles to cancer sites, their external surfaces are modified chemically to carry molecules that have lock-and-key binding specificity with molecules that support a

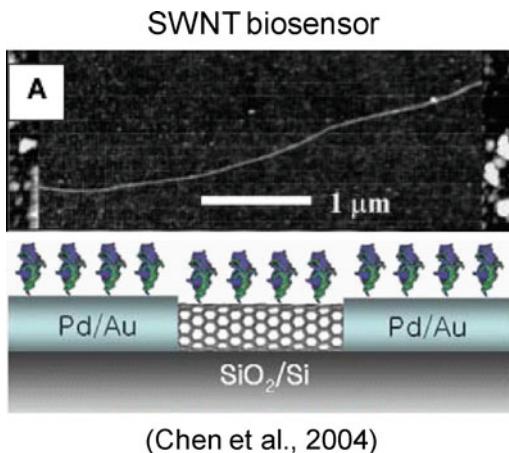


Fig. 1.19 SEM micrograph of SWNT biosensor; bottom schematic shows the adsorption of protein molecules to the SWNT (Chen et al. 2004)

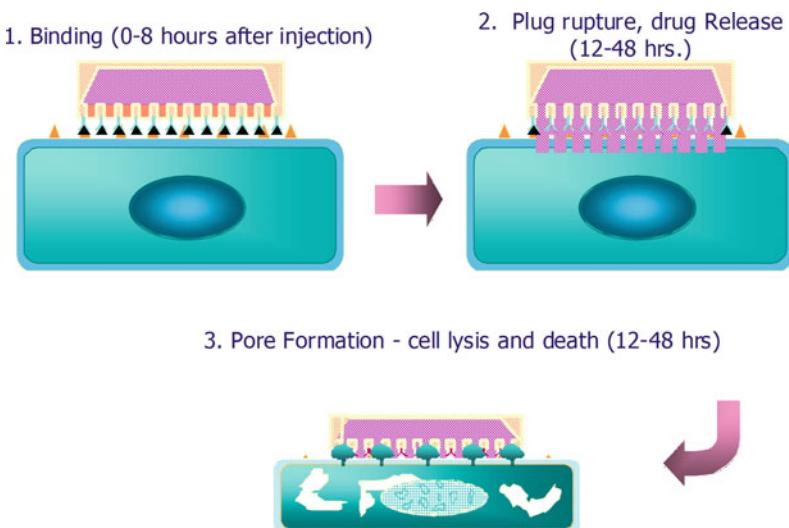
growing cancer mass. As the particles come into close proximity with the diseased cells, the ligands on the particle surfaces attach to the receptors on the cells. As soon as the particles dock on the cells, a compound is released that forms a pore on the membrane of the cells, which leads to cell death and, ultimately, death of the

cancer mass that was being nourished by the blood vessel. The adhesive interactions are regulated by specific (ligand-receptor binding) and non-specific (short-range van der Walls, electro static, steric) interactions (Decuzzi et al. 2005). Adhesion between nanoparticles and disease cells is required. Furthermore, the particles should travel close to the endothelium lining of vascular arteries to facilitate the interaction between the particles and diseased cells. Human capillaries can have radii as small as 4–5 μm.

1.6 Role of Nanotribology and Nanomechanics in Magnetic Storage Devices and MEMS/NEMS and BioMEMS/NEMS

Magnetic storage devices and MEMS/NEMS and BioMEMS/NEMS are the two examples where nanotribological and nanomechanical tools and techniques are essential for studies of nanoscale phenomena. Surface roughness and thicknesses of hard solid coatings and liquid lubricant

Fig. 1.20 Intravascular nanoparticles to search and destroy diseased blood cells. Life cycles include 1 target binding, 2 drug release, and 3 cell death (Martin and Grove 2001)



Intravascular nanoparticles for search and destroy diseased blood cells
(Martin and Grove, 2001)

coatings on the magnetic disk surface and flying height continue to decrease. Near-contact recording devices are also used. Interface studies of components with ultra-thin coatings can be performed ideally using nanotribological and nanomechanical tools and techniques.

In the case of MEMS/NEMS and BioMEMS/NEMS, adhesion, friction, and wear problems of moving components need to be addressed for high performance, long life, and reliability. Molecularly-thin films of solid and/or liquids are used for low adhesion, friction, and wear in many applications. In the case of BioMEMS/NEMS, adhesion between biological molecular layers and the substrate, and friction and wear of biological layers can be important. Again, interfacial phenomena in MEMS/NEMS and BioMEMS/NEMS can be ideally studied using nanotribological and nanomechanical tools and techniques.

1.7 Organization of the Book

The book integrates knowledge of nanotribology and nanomechanics. It starts with the definition of tribology, history of tribology and micro/nanotribology, its industrial significance, various measurement techniques employed, followed by various industrial applications. The remaining book is divided into five parts. The first part introduces scanning probe microscopy. The second part provides an overview of nanomechanics. The third part provides an overview of nanotribology. The fourth part provides an overview of molecularly-thick films for lubrication. And the last part focuses on nanotribology and nanomechanics studies conducted for various industrial applications.

References

- Alexander S, Hellemans L, Marti O, Schneir J, Elings V, Hansma PK (1989) An atomic-resolution atomic-force microscope implemented using an optical lever. *J Appl Phys* 65:164–167
- Amontons G (1699) De la resistance causee dans les machines. *Memoires de l'Academie Royale A*, pp 257–282
- Bhushan B (1996) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer, New York
- Bhushan B (1997) Micro/nanotribology and its applications, NATO ASI series E: applied sciences 330. Kluwer Academic Publishers, Dordrecht, Netherlands
- Bhushan B (1998) Tribology issues and opportunities in MEMS. Kluwer Academic Publishers, Dordrecht, Netherlands
- Bhushan B (1999) Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton, Florida
- Bhushan B (2000) Mechanics and reliability of flexible magnetic media, 2nd edn. Springer, New York
- Bhushan B (2001a) Fundamentals of tribology and bridging the gap between the macro- and micro/nanoscales. In: NATO science series II: mathematics, physics, and chemistry, vol 10. Kluwer Academic Publishers, Dordrecht, Netherlands
- Bhushan B (2001b) Macro- and microtribology of magnetic storage devices. In: Modern tribology handbook, vol 2—materials, coatings, and industrial applications. CRC Press, Boca Raton, Florida, pp 1413–1513
- Bhushan B (2005) Nanotribology and nanomechanics. *Wear* 259:1507–1531
- Bhushan B (2008a) Nanotribology, nanomechanics and nanomaterials characterization. *Phil Trans Roy Soc A* 366:1351–1381
- Bhushan B (2008b) Nanotribology and nanomechanics in nano/biotechnology. *Phil Trans Roy Soc A* 366:1499–1537
- Bhushan B (2010) Biophysics of human hair: structural, nanomechanical and nanotribological studies. Springer, Germany
- Bhushan B (2011) Nanotribology and nanomechanics I—measurement techniques and nanomechanics; II—nanotribology, biomimetics, and Industrial applications. Springer, Germany
- Bhushan B (2013a) Principles and applications of tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2016) Encyclopedia of nanotechnology, 2nd edn. Springer, Switzerland
- Bhushan B (2017a) Springer handbook of nanotechnology, 4th edn. Springer, Switzerland
- Bhushan B (2017b) Biophysics of skin and its treatments—structural, nanotribological, and nanomechanical studies. Springer, Switzerland
- Bhushan B, Ruan J (1994) Atomic-scale friction measurements using friction force microscopy: Part II—application to magnetic media. *ASME J Tribol* 116:389–396
- Bhushan B, Gupta BK (1991) Handbook of tribology: materials, coatings, and surface treatments. McGraw-Hill, New York; Reprinted with corrections, Krieger Publishing Co., Malabar, Florida (1997)
- Bhushan B, Li X (2003) Nanomechanical characterisation of solid surfaces and thin films, (Invited). *Inter Mater Rev* 48:125–164

- Bhushan B, Koinkar VN, Ruan J (1994) Microtribology of magnetic media. *Proc Inst Mech Eng Part J: J Eng Tribol* 208:17–29
- Bhushan B, Israelachvili JN, Landman U (1995) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996) Nanoindentation and picoindentation measurements using a capacitive transducer system in atomic force microscopy. *Philos Mag A* 74:1117–1128
- Bhushan B, Fuchs H, Hosaka S (2004) Applied scanning probe methods. Springer, Germany
- Bhushan B, Tokachichu DR, Keener MT, Lee SC (2005) Morphology and adhesion of biomolecules on silicon based surfaces. *Acta Biomater* 1:327–341
- Bhushan B, Tokachichu DR, Keener MT, Lee SC (2006) Nanoscale adhesion, friction, and wear studies of biomolecules on silicon based surfaces. *Acta Biomater* 2:39–49
- Binnig G, Rohrer H, Gerber Ch, Weibel E (1982) Surface studies by scanning tunnelling microscopy. *Phys Rev Lett* 49:57–61
- Binnig G, Quate CF, Gerber Ch (1986) Atomic force microscope. *Phys Rev Lett* 56:930–933
- Binnig G, Gerber Ch, Stoll E, Albrecht TR, Quate CF (1987) Atomic resolution with atomic force microscope. *Europhys Lett* 3:1281–1286
- Bowden FP, Tabor D (1950) The Friction and Lubrication of Solids, vol 1. Clarendon Press, Oxford
- Bowden FP, Tabor D (1964) The Friction and Lubrication of Solids, vol 2. Clarendon Press, Oxford
- Castillo J, Bhushan B (2006) Tribology of hard disk drives—magnetic data storage technology. In: Totten GE (ed) Handbook of lubrication and tribology, vol 1—application and maintenance, 2nd edn. CRC Press, Taylor & Francis Group, Boca Raton, Florida, pp 16-1–16-44
- Chan DYC, Horn RG (1985) The drainage of thin liquid films between solid surfaces. *J Chem Phys* 83:5311–5324
- Chen RJ, Choi HC, Bangsaruntip S, Yenilmez E, Tang X, Wang Q, Chang WL, Dai H (2004) An investigation of the mechanisms of electrode sensing of protein adsorption on carbon nanotube devices. *J Am Chem Soc* 126:1563–1568
- Christenson HK (1988) Adhesion between surfaces in unsaturated vapors—a reexamination of the influence of meniscus curvature and surface forces. *J Colloid Interface Sci* 121:170–178
- Coulomb CA (1785) Théorie des machines simples, en ayant regard au Frottement de leurs Parties et à la Roideur des Cordages. *Mem Math Phys X*, Paris 161–342
- Davidson CSC (1957) Bearings since the stone age. *Engineering* 183:2–5
- Decuzzi P, Lee S, Bhushan B, Ferrari M (2005) A theoretical model for the margination of particles with blood vessels. *Annals Biomed Eng* 33:179–190
- Dowson D (1998) History of tribology, 2nd edn. Institution Mechanical Engineers, London
- Gee ML, McGuigan PM, Israelachvili JN, Homola AM (1990) Liquid to solid-like transitions of molecularly thin films under shear. *J Chem Phys* 93:1895–1906
- Georges JM, Millot S, Loubet JL, Tonck A (1993) Drainage of thin liquid films between relatively smooth surfaces. *J Chem Phys* 98:7345–7360
- Georges JM, Tonck A, Mazoyer D (1994) Interfacial friction of wetted monolayers. *Wear* 175:59–62
- Granick S (1991) Motions and relaxations of confined liquids. *Science* 253:1374–1379
- Hirz SJ, Homola AM, Hadzioannou G, Frank SW (1992) Effect of substrate on shearing properties of ultrathin polymer films. *Langmuir* 8:328–333
- Holm R (1946) Electrical contacts. Springer, New York
- Homola AM (1989) Measurement of and relation between the adhesion and friction of two surfaces separated by thin liquid and polymer films. *ASME J Tribol* 111:675–682
- Homola AM (1993) Interfacial friction of molecularly thin liquid films. Miyoshi K, Chung YW (eds) Surface diagnostics in tribology. World Scientific Publishing, River Edge, New Jersey, pp 271–298
- Homola AM, Israelachvili JN, Gee ML, McGuigan PM (1989) Measurement of and relation between the adhesion and friction of two surfaces separated by thin liquid and polymer films. *ASME J Tribol* 111:675–682
- Homola AM, Israelachvili JN, McGuigan PM, Gee ML (1990) Fundamental experimental studies in tribology: the transition from interfacial friction of undamaged molecularly smooth surfaces. *Wear* 136:65–83
- Homola AM, Nguyen HV, Hadzioannou G (1991) Influence of monomer architecture on the shear properties of molecularly thin polymer melts. *J Chem Phys* 94:2346–2351
- Horn RG (1990) Surface forces and their action in ceramic materials. *Am Ceram Soc* 73:1117–1135
- Horn RG, Israelachvili JN (1988) Molecular organization and viscosity of a thin film of molten polymer between two surfaces as probed by force measurements. *Macromolecules* 21:2836–2841
- Horn RG, Smith DT, Haller W (1989) Surface forces and viscosity of water measured between silica sheets. *Chem Phys Lett* 162:404–408
- Hornbeck LJ (1999) A digital light processingTM update—status and future applications. *Proc SPIE* 3634:158–170
- Hornbeck LJ, Nelson WE (1988) Bistable deformable mirror devices. In: OSA technical digest series, vol 8, spatial light modulators and applications. OSA Washington, pp 107–110
- Hu WW, Carson GA, Granick S (1991) Relaxation time of confined liquids under shear. *Phys Rev Lett* 66:2758–2761
- Israelachvili JN (1987) Solvation forces and liquid structure—as probed by direct force measurements. *Acc Chem Res* 20:415–421
- Israelachvili JN (1989) Techniques for direct measurements of forces between surfaces in liquids at the atomic scale. *Chemtracts—Anal Phys Chem* 1:1–12

- Israelachvili JN (1992) Intermolecular and surface forces, 2nd edn. Academic Press, San Diego
- Israelachvili JN, Adams GE (1978) Measurement of friction between two mica surfaces in aqueous electrolyte solutions in the range 0–100 nm. *Chem Soc J Faraday Trans I* 74:975–1001
- Israelachvili JN, McGuiggan PM (1988) Forces between surface in liquids. *Science* 241:795–800
- Israelachvili JN, Tabor D (1972) The measurement of van der waals dispersion forces in the range of 1.5–130 nm. *Proc R Soc Lond A* 331:19–38
- Israelachvili JN, McGuiggan PM, Homola AM (1988) Dynamic properties of molecularly thin liquid films. *Science* 240:189–190
- Jost P (1966) Lubrication (tribology)—A report on the present position and industry's needs. Dept. of Education and Science, H.M. Stationery Office, London
- Jost P (1976) Economic impact of tribology. In: Proceedings mechanical failures prevention group, NBS Spec. Pub. 423, Gaithersburg, Maryland
- Klein J (1980) Forces between mica surfaces bearing layers of adsorbed polystyrene in cyclohexane. *Nature* 288:248–250
- Klein J, Perahia D, Warburg S (1991) Forces between polymer-bearing surfaces undergoing shear. *Nature* 352:143–145
- Layard AG (1853) Discoveries in the ruins of nineveh and babylon, I and II. John Murray, Albemarle Street, London
- Liu H, Bhushan B (2004a) Nanotribological characterization of digital micromirror devices using an atomic force microscope. *Ultramicroscopy* 100:391–412
- Liu H, Bhushan B (2004b) Investigation of nanotribological and nanomechanical properties of the digital micromirror device by atomic force microscope. *J Vac Sci Technol A* 22:1388–1396
- Liu H, Bhushan B (2004c) Bending and fatigue study on a nanoscale hinge by an atomic force microscope. *Nanotechnology* 15:1246–1251
- Loubet JL, Bauer M, Tonck A, Bec S, Gauthier-Manuel B (1993) Nanoindentation with a surface force apparatus. In: Nastasi M et al (ed) Mechanical properties and deformation behavior of materials having ultra-fine microstructures. Kluwer Academic Publishers, Dordrecht, pp 429–447
- Luengo G, Schmitt FJ, Hill R, Israelachvili JN (1997) Thin film bulk rheology and tribology of confined polymer melts: contrasts with bulk properties. *Macromolecules* 30:2482–2494
- Maharaj D, Bhushan B (2015) Friction, wear and mechanical behavior of nano-objects on the nanoscale. *Mater Sci Eng R* 95:1–43
- Martin FJ, Grove C (2001) Microfabricated drug delivery systems: concepts to improve clinical benefits. *Biomed Microdevices* 3:97–108
- Meyer G, Amer NM (1988) Novel optical approach to atomic force microscopy. *Appl Phys Lett* 53:1045–1047
- Parish WF (1935) Three thousand years of progress in the development of machinery and lubricants for the hand crafts. *Mill and Factory* 16 and 17
- Pashley RM (1981) Hydration forces between solid surfaces in aqueous electrolyte solutions. *J Colloid Interface Sci* 80:153–162
- Patel SS, Tirrell M (1989) Measurement of forces between surfaces in polymer fluids. *Annu Rev Phys Chem* 40:597–635
- Peachey J, Van Alsten J, Granick S (1991) Design of an apparatus to measure the shear response of ultrathin liquid films. *Rev Sci Instrum* 62:463–473
- Petroff NP (1883) Friction in Machines and the effects of the lubricant. *Engng. J.* (in Russian) St. Petersburg, pp 71–140, 228–279, 377–436, 535–564
- Reynolds OO (1886) On the theory of lubrication and its applications to Mr. Beauchamp tower's experiments. *Philos Trans R Soc (London)* 117:157–234
- Ruan J, Bhushan B (1994a) Atomic-scale friction measurements using friction force microscopy: Part I - general principles and new measurement techniques. *ASME J Tribol* 116:378–388
- Ruan J, Bhushan B (1994b) Atomic-scale and microscale friction of graphite and diamond using friction force microscopy. *J Appl Phys* 76:5022–5035
- Smith CP, Maeda M, Atanasoska L, White HS (1988) Ultrathin platinum films on mica and measurement of forces at the platinum/water interface. *J Phys Chem* 95:199–205
- Tabor D, Winterton RHS (1969) The direct measurement of normal and retarded van der waals forces. *Proc R Soc Lond A* 312:435–450
- Tonck A, Georges JM, Loubet JL (1988) Measurements of intermolecular forces and the rheology of dodecane between alumina surfaces. *J Colloid and Interf Sci* 126:1540–1563
- Tower B (1884) Report on friction experiments. *Proc Instn Mech Engrs* 632:29–35
- Van Alsten J, Granick S (1988) Molecular tribology of ultrathin liquid films. *Phys Rev Lett* 61:2570–2573
- Van Alsten J, Granick S (1990a) Shear rheology in a confined geometry—polysiloxane melts. *Macromolecules* 23:4856–4862
- Van Alsten J, Granick S (1990b) Tribology studied using atomically smooth surfaces. *Trib Trans* 33:436–446

Part I

Scanning Probe Microscopy

Scanning Probe Microscopy— Principle of Operation, Instrumentation, and Probes

Bharat Bhushan and Othmar Marti

Abstract

Since the introduction of the Scanning Tunneling Microscope (STM) in 1981 and Atomic Force Microscope (AFM) in 1985, many variations of probe-based microscopies, referred to as Scanning Probe Microscopes (SPM), have been developed. While the pure imaging capabilities of SPM techniques are dominated by the application of these methods at their early development stages, the physics of probe-sample interactions and the quantitative analyses of tribological, electronic, magnetic, biological, and chemical surfaces have now become of increasing interest. In this chapter, we introduce various STM and AFM designs, various operating modes, various probes (tips), and AFM instrumentation and analyses.

Keywords

Scanning probe microscopy • Atomic force microscopy • Scanning tunneling microscopy • Friction force microscopy • Electrical properties • Mechanical properties • Nanotribology • Nanomechanics

2.1 Introduction

The Scanning Tunneling Microscope (STM) developed by Dr. Gerd Binnig and his colleagues in 1981 at the IBM Zurich Research Laboratory, Rueschlikon, Switzerland, is the first instrument capable of directly obtaining three-dimensional (3D) images of solid surfaces with atomic resolution (Binnig et al. 1982). G. Binnig and H. Rohrer received a Nobel Prize in Physics in 1986 for their discovery. STMs can only be used to study surfaces that are electrically conductive to

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

O. Marti
Abteilung Experimentelle Physik, Universitaet Ulm,
Albert-Einstein-Allee 11, 89069 Ulm, Germany
e-mail: othmar.marti@uni-ulm.de

some degree. Based on their design of the STM, Binnig et al. developed in 1985 an Atomic Force Microscope (AFM) to measure ultrasmall forces (less than 1 μN) present between the AFM tip surface and the sample surface (Binnig et al. 1986; also see Binnig et al. 1987). AFMs can be used for measurement of all engineering surfaces that may be either electrically conductive or insulating. AFM has become a popular surface profiler for topographic and normal force measurements on the micro to nanoscale.

AFMs have been modified in order to measure both normal and lateral forces, and are called Lateral Force Microscope (LFM) or Friction Force Microscope (FFM) (Mate et al. 1987; Erlandsson et al. 1988; Marti et al. 1990; Meyer and Amer 1990b; Bhushan and Ruan 1994; Bhushan et al. 1995; Bhushan 1999a, 2011). FFM further have been modified to measure lateral forces in two orthogonal directions (Fujisawa et al. 1994a, b; Grafstrom et al. 1994; Overney et al. 1994; Warmack et al. 1994). A number of researchers have continued to improve the AFM/FFM designs and used them to measure adhesion and friction of solid and liquid surfaces on micro- and nanoscales (Burnham et al. 1990, 1991; Frisbie et al. 1994; Bhushan et al. 1995; Koinkar and Bhushan 1996; Scherer et al. 1997, 1999; Bhushan and Sundararajan 1998; Krotil et al. 1999; Bhushan and Dandavate 2000; Bhushan 1997, 1999a, 2001a, b, 2005, 2008, 2011, 2013a, b; Reinstaedtler et al. 2003). AFMs have been used for scratching, wear, and measurements of elastic/plastic mechanical properties (such as indentation hardness and modulus of elasticity) (Burnham and Colton 1989; Maivald et al. 1991; Bhushan et al. 1994, 1995, 1996; Bhushan and Koinkar 1994; DeVecchio and Bhushan 1997; Scherer et al. 1997; Bhushan and Sundararajan 1998; Bhushan 1997, 1999a, b, 2001a, b, c, 2005, 2008, 2011, 2013a, b; Amelio et al. 2001; Bhushan and Qi 2003; Kasai et al. 2004; Reinstaedtler et al. 2005).

AFMs have been used for manipulation of individual atoms of Xenon (Eigler and Schweizer 1990), molecules (Weisenhorn et al. 1990), silicon surfaces (Lyo and Avouris 1991), and

polymer surfaces (Leung and Goh 1992). STMs have been used for formation of nanofeatures by localized heating or by inducing chemical reactions under the STM tip (Abraham et al. 1986; Silver et al. 1987; Kobayashi et al. 1993) and nanomachining (Parkinson 1990). AFMs have been used for nanofabrication (Majumdar et al. 1992; Bhushan et al. 1994; Bhushan 1995, 1999a; Tsau et al. 1994) and nanomachining (Delawski and Parkinson 1992).

STMs and AFMs are used at extreme magnifications ranging from 10^3 to $10^9 \times$ in x, y, and z directions for imaging macro to atomic dimensions with high resolution information and for spectroscopy. These instruments can be used in any environment such as ambient air (Binnig et al. 1986; Bhushan and Blackman 1991), various gases (Burnham et al. 1990), liquid (Marti et al. 1987; Drake et al. 1989; Binggeli et al. 1993), vacuum (Binnig et al. 1982; Meyer and Amer 1988), low temperatures (lower than about 100 K) (Coombs and Pethica 1986; Kirk et al. 1988; Giessibl et al. 1991; Albrecht et al. 1992; Hug et al. 1993), and high temperatures (Basire and Ivanov 2000; Liu and Bhushan 2002). Imaging in liquid allows the study of live biological samples and also eliminates water capillary forces present in ambient air present at the tip-sample interface. Low temperature (liquid helium temperatures) imaging is useful for the study of biological and organic materials and the study of low-temperature phenomena such as superconductivity or charge-density waves. Low-temperature operation is also advantageous for high-sensitivity force mapping due to the reduction in thermal vibration. STMs and AFMs also have been used to image liquids such as liquid crystals and lubricant molecules on graphite surfaces (Foster and Frommer 1988; Smith et al. 1989, 1990; Andoh et al. 1992).

The pure imaging capabilities of variations of probe-based microscope techniques, referred to as Scanning Probe Microscopy (SPM), have dominated the application of these methods at their early development stages. However, the physics and chemistry of probe-sample interactions and the quantitative analyses of tribological, electronic, magnetic, biological, and

chemical surfaces now have become of increasing interest. Nanoscale science and technology are strongly driven by SPMs, which allow investigation and manipulation of surfaces down to the atomic scale. With growing understanding of the underlying interaction mechanisms, SPMs have found application in many fields outside basic research fields. In addition, various derivatives of all these methods have been developed for special applications, some of them targeting far beyond microscopy.

A family of instruments based on STMs and AFMs, SPMs have been developed for various applications of scientific and industrial interest. These include STM, AFM, FFM (or LFM), scanning electrostatic force microscopy (SEFM) (Martin et al. 1987; Stern et al. 1988), scanning force acoustic microscopy (SFAM) [or atomic force acoustic microscopy (AFAM)] (Yamanaka et al. 1994; Yamanaka and Tomita 1995; Rabe et al. 1996; Scherer et al. 1997, 1999; Amelio et al. 2001; Reinstaedtler et al. 2005), scanning magnetic microscopy (SMM) [or magnetic force microscopy (MFM)] (Martin and Wickramasinghe 1987; Rugar et al. 1990; Schoenberger and Alvarado 1990; Hartmann 1999; Avila and Bhushan 2010), scanning near field optical microscopy (SNOM) (Pohl et al. 1984; Betzig et al. 1991, 1992; Barbara et al. 1999), scanning thermal microscopy (SThM) (Williams and Wickramasinghe 1986, 1990; Majumdar 1999), scanning electrochemical microscopy (SECM) (Husser et al. 1989), scanning Kelvin Probe microscopy (SKPM) (Martin et al. 1988; Nonnenmacher et al. 1991; Weaver and Abraham

1991; DeVecchio and Bhushan 1998; Bhushan and Goldade 2000), scanning chemical potential microscopy (SCPM) (Williams and Wickramasinghe 1990), scanning ion conductance microscopy (SICM) (Hansma et al. 1989; Prater et al. 1991), and scanning capacitance microscopy (SCM) (Matey and Blanc 1985; Martin et al. 1988; Williams 1999; Lee et al. 2002). The subfamily of instruments that measure forces (e.g., AFM, FFM, SEFM, SFAM, and SMM) also are referred to as scanning force microscopes (SFM). Although these instruments offer atomic resolution and are ideal for basic research, they are most often used for cutting edge industrial applications that do not require atomic resolution.

Commercial production of SPMs started with STM in 1987 and AFM in 1989 by Digital Instruments Inc. (now Bruker Instruments. Inc.). For comparisons of SPMs with other microscopes, see Table 2.1 (Bruker Instruments, Inc.). Numbers of these instruments are equally divided into U.S., Japan, and Europe with industry/university and Government lab splits of about 50/50, 70/30, and 30/70, respectively. It is clear that research and industrial applications of SPMs are rapidly expanding. For various examples, see Bhushan et al. (2004, 2007, 2008), Bhushan and Fuchs (2006, 2007, 2009), Bhushan and Kawata (2007), Fuchs and Bhushan (2010), Bhushan (2010–2013, 2011).

This chapter presents an overview of STM and AFM with various operating modes and various probes (tips) used in these instruments. Then, details on AFM instrumentation and analyses are given.

Table 2.1 Comparison of various conventional microscopes with SPMs in 2015

	Optical	SEM/TEM	Confocal	SPM
Magnification	10^3	10^7	10^4	10^9
Instrument price (U.S. \$)	~\$20 k	~\$500 k	~\$60 k	~\$150 k
Technology age	200 years	50 years	30 years	30 years
Applications	Ubiquitous	Science & technology	New & unfolding	Science & technology
Approximate market 2015	\$2 B	\$1.5 B	\$400 M	\$600 M

2.2 Scanning Tunneling Microscope

The principle of electron tunneling was proposed by Giaever (1960). He proposed that if an electrical potential difference is applied to two metals separated by a thin insulating film, a current will flow across the barrier because of the ability of electrons to penetrate it. To be able to measure this tunneling current, the two metals must be spaced no more than 10 nm apart. Binnig et al. (1982) introduced vacuum tunneling combined with lateral scanning. The vacuum provides the ideal barrier for tunneling. The lateral scanning allows surface imaging with exquisite resolution—lateral less than 1 nm and vertical less than 0.1 nm—sufficient to define the position of single atoms. The very high vertical resolution is obtained because the tunneling current varies exponentially with the distance between the metal tip and the scanned surface acting as two electrodes. Typically, tunneling current decreases by a factor of 2 as the separation is increased by 0.2 nm. Very high lateral resolution depends upon the sharp tips. Binnig et al. (1982) overcame two key obstacles for damping external vibrations and for moving the tunneling probe in close proximity to the sample. Their instrument is called the scanning tunneling microscope (STM). Today's STMs can be used in the ambient environment for atomic-scale image of surfaces. Excellent reviews on this subject are presented by Hansma and Tersoff (1987), Sarid and Elings (1991), Durig et al. (1992); Frommer (1992), Guntherodt and Wiesendanger (1992), Wiesendanger and Guntherodt (1992), Bonnell (1993), Marti and Amrein (1993), Stroscio and Kaiser (1993), and Guntherodt et al. (1995).

The principle of STM is straightforward. A sharp metal tip (one electrode of the tunnel junction) is brought close enough (0.3–1 nm) to the surface to be investigated (second electrode) that, at a convenient operating voltage (10 mV–1 V), the tunneling current varies from 0.2 to 10 nA, which is measurable. The tip is scanned over a surface at a distance of 0.3–1 nm, while the tunneling current between it and the surface is sensed. The STM can be operated in either the constant current mode or the constant height

mode, Fig. 2.1. The left-hand column of Fig. 2.1 shows the basic constant current mode of operation. A feedback network changes the height of the tip z to keep the current constant. The displacement of the tip given by the voltage applied to the piezoelectric drives then yields a topographic map of the surface. Alternatively, in the constant height mode, a metal tip can be scanned across a surface at nearly constant height and constant voltage while the current is monitored, as shown in the right-hand column of Fig. 2.1. In this case, the feedback network responds only rapidly enough to keep the average current constant. The constant current mode is used generally for atomic-scale images. This mode is not practical for rough surfaces. A three-dimensional picture $[z(x, y)]$ of a surface consists of multiple scans $[z(x)]$ displayed laterally from each other in the y direction. It should be noted that if different atomic species are present in a sample, the different atomic species within a sample may produce different tunneling currents for a given bias voltage. Thus, the height data may not be a direct representation of the topography of the surface of the sample.

2.2.1 Binnig et al.'s Design

Figure 2.2 shows a schematic of one of Binnig and Rohrer's designs for operation in ultrahigh vacuum (Binnig et al. 1982; Binnig and Rohrer 1983). The metal tip was fixed to rectangular piezodrives P_x , P_y , and P_z made out of commercial piezoceramic material for scanning. The sample is mounted on either superconducting magnetic levitation or two-stage spring system to achieve a stability of the gap width of about 0.02 nm. The tunnel current J_T is a sensitive function of the gap width d that is $J_T \propto V_T \exp(-A\phi^{1/2}d)$, where V_T is the bias voltage, ϕ is the average barrier height (work function) and the constant $A = 1.025 \text{ eV}^{-1/2} \text{ \AA}^{-1}$. With a work function of a few eV, J_T changes by an order of magnitude for every angstrom change of d . If the current is kept constant to within, for example, 2%, then the gap d remains constant to within 1 pm. For operation in the constant current

Fig. 2.1 STM can be operated in either the constant-current or the constant-height mode. The images are of graphite in air (Hansma and Tersoff 1987)

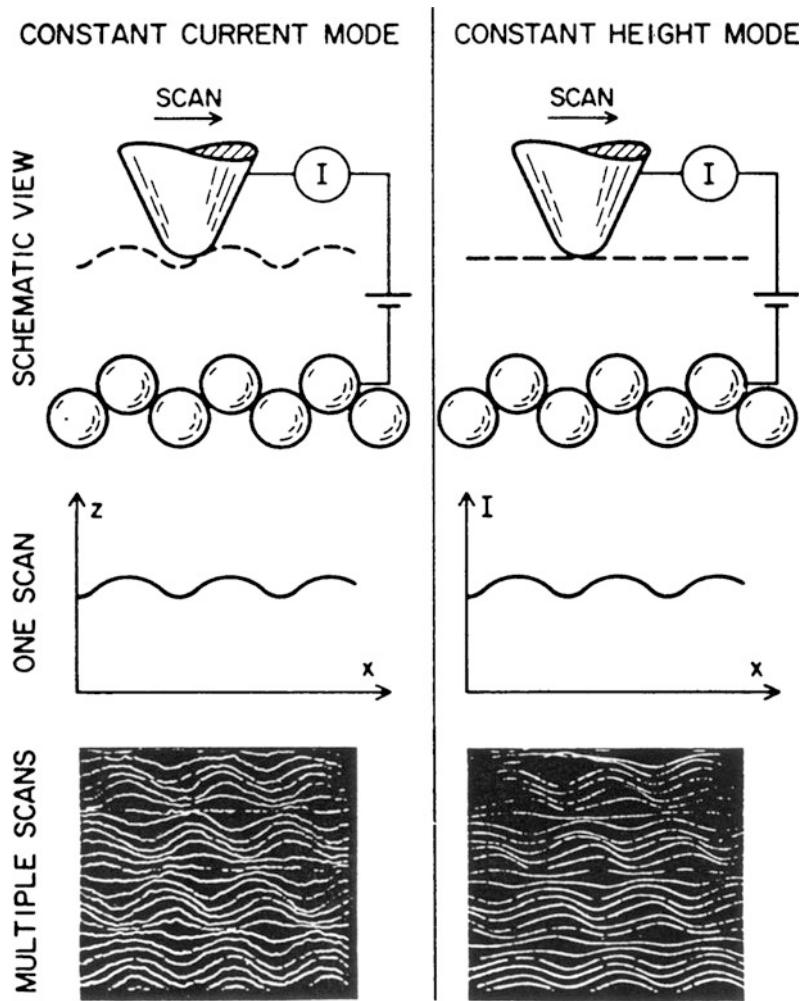
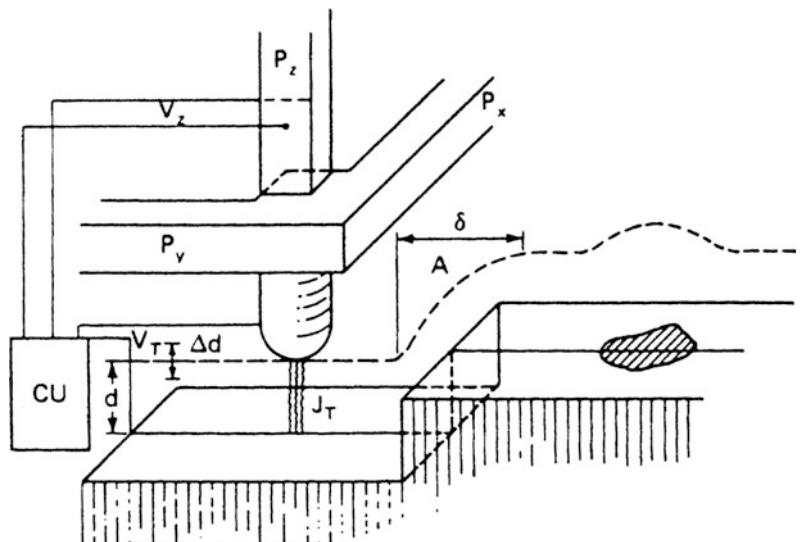


Fig. 2.2 Principle of operation of the STM made by Binnig and Rohrer (1983)



mode, the control unit (CU) applies a voltage V_z to the piezo P_z such that J_T remains constant when scanning the tip with P_y and P_x over the surface. At a constant work function ϕ , $V_z(V_x, V_y)$ yields the roughness of the surface $z(x, y)$ directly, as illustrated at a surface step at A. Smearing the step, δ (lateral resolution) is on the order of $(R)^{1/2}$, where R is the radius of the curvature of the tip. Thus, a lateral resolution of about 2 nm requires tip radii on the order of 10 nm. A 1-mm-diameter solid rod ground at one end at roughly 90° yields overall tip radii of only a few hundred nm, but with closest protrusion of rather sharp microtips on the relatively dull end yields a lateral resolution of about 2 nm. In situ sharpening of the tips by gently touching the surface brings the resolution down to the 1-nm range; by applying high fields (on the order of 10^8 V/cm). For example, applications of high field for half an hour, resolutions considerably below 1 nm could be reached. Most experiments were done with tungsten wires either ground or etched to a radius typically in the range of 0.1–10 μm . In some cases, in situ processing of the tips was done for further reduction of tip radii.

2.2.2 Commercial STMs

There are a number of commercial STMs available on the market. In 1987, Digital Instruments, Inc. (now Bruker Instruments) located in Santa Barbara, CA, introduced the first commercial STM, the Nanoscope I. In a later-model Nanoscope IV STM for operation in ambient air, the sample is held in position while a piezoelectric crystal in the form of a cylindrical tube (referred to as PZT tube scanner) scans the sharp metallic probe over the surface in a raster pattern while sensing and outputting the tunneling current to the control station, Fig. 2.3. The digital signal processor (DSP) calculates the desired separation of the tip from the sample by sensing the tunneling current flowing between the sample and the tip. The bias voltage applied between the sample and the tip encourages the tunneling current to flow. The DSP completes the digital feedback loop by outputting the desired voltage to

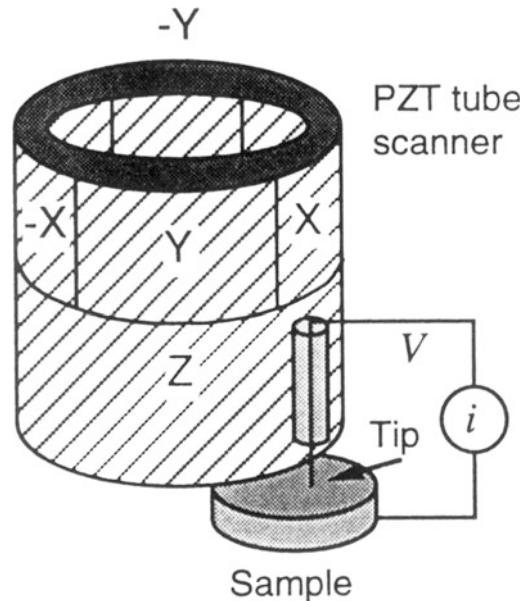


Fig. 2.3 Principle of operation of a commercial STM, a sharp tip attached to a piezoelectric tube scanner is scanned on a sample

the piezoelectric tube. The STM operates in both the “constant height” and “constant current” modes depending on a parameter selection in the control panel. In the constant current mode, the feedback gains are set high, the tunneling tip closely tracks the sample surface, and the variation in the tip height required to maintain constant tunneling current is measured by the change in the voltage applied to the piezo tube. In the constant height mode, the feedback gains are set low, the tip remains at a nearly constant height as it sweeps over the sample surface, and the tunneling current is imaged.

Physically, the Nanoscope STM consists of three main parts: the head, which houses the piezoelectric tube scanner for three dimensional motion of the tip and the preamplifier circuit (FET input amplifier) mounted on top of the head for the tunneling current; the base on which the sample is mounted; and the base support, which supports the base and head (Bhushan 1999a). The base accommodates samples up to 10 mm by 20 and 10 mm in thickness. Scan sizes available for the STM are 0.7 μm (for atomic resolution), 12, 75 and 125 μm^2 .

The scanning head controls the three dimensional motion of tip. The removable head consists of a piezo tube scanner, about 12.7 mm in diameter, mounted into an invar shell used to minimize vertical thermal drifts because of good thermal match between the piezo tube and the Invar. The piezo tube has separate electrodes for X, Y and Z which are driven by separate drive circuits. The electrode configuration (Fig. 2.3) provides x and y motions which are perpendicular to each other, minimizes horizontal and vertical coupling, and provides good sensitivity. The vertical motion of the tube is controlled by the Z electrode which is driven by the feedback loop. The x and y scanning motions are each controlled by two electrodes which are driven by voltages of same magnitudes, but opposite signs. These electrodes are called $-Y$, $-X$, $+Y$, and $+X$. Applying complimentary voltages allows a short, stiff tube to provide a good scan range without large voltages. The motion of the tip due to external vibrations is proportional to the square of the ratio of vibration frequency to the resonant frequency of the tube. Therefore, to minimize the tip vibrations, the resonant frequencies of the tube are high about 60 kHz in the vertical direction and about 40 kHz in the horizontal direction. The tip holder is a stainless steel tube with a 300 μm inner diameter for 250 μm diameter tips, mounted in ceramic in order to keep the mass on the end of the tube low. The tip is mounted either on the front edge of the tube (to keep mounting mass low and resonant frequency high) (Fig. 2.3) or the center of the tube for large range scanners, namely 75 and 125 μm (to preserve the symmetry of the scanning). This commercial STM accepts any tip with a 250 μm diameter shaft. The piezotube requires X-Y calibration which is carried out by imaging an appropriate calibration standard. Cleaved graphite is used for the small-scan length head while two dimensional grids (a gold plated ruling) can be used for longer range heads.

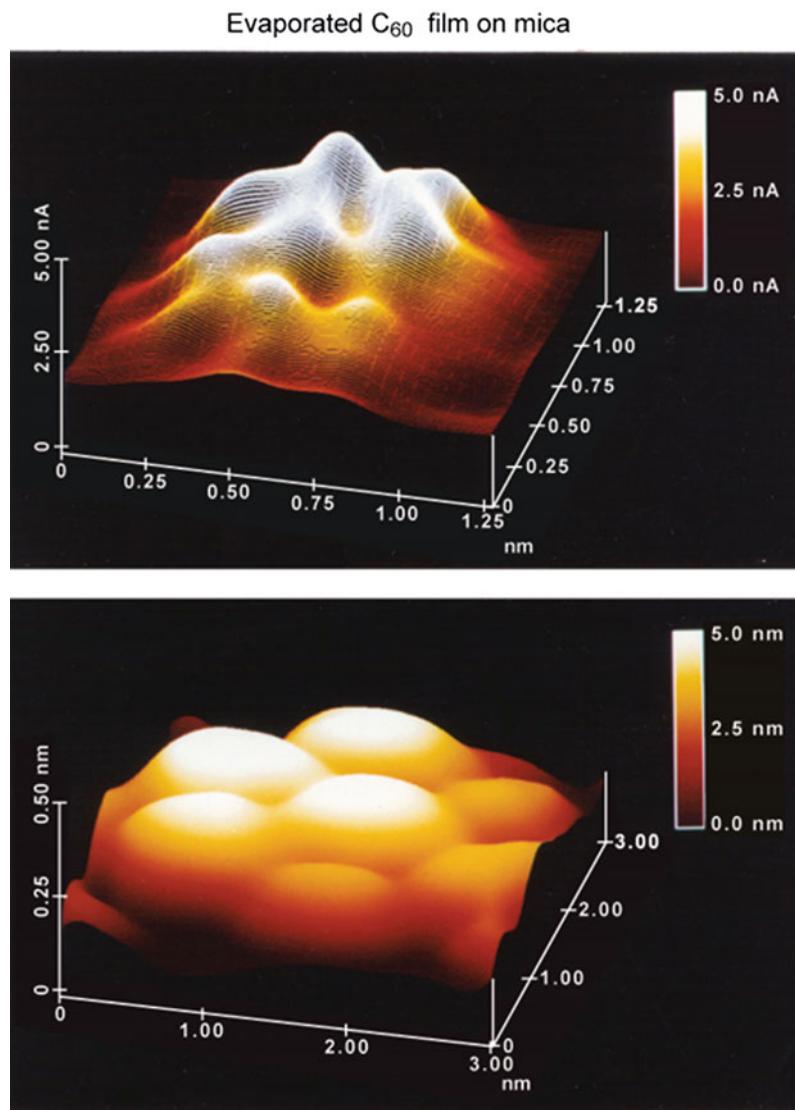
The Invar base holds the sample in position, supports the head, and provides coarse X-Y motion for the sample. A spring-steel sample clip with two thumb screws holds the sample in

place. An x-y translation stage built into the base allows the sample to be repositioned under the tip. Three precision screws arranged in a triangular pattern support the head and provide coarse and fine adjustment of the tip height. The base support consists of the base support ring and the motor housing. The stepper motor enclosed in the motor housing allows the tip to be engaged and withdrawn from the surface automatically.

Samples to be imaged with STM must be conductive enough to allow a few nanoamperes of current to flow from the bias voltage source to the area to be scanned. In many cases, nonconductive samples can be coated with a thin layer of a conductive material to facilitate imaging. The bias voltage and the tunneling current depend on the sample. Usually they are set at a standard value for engagement and fine-tuned to enhance the quality of the image. The scan size depends on the sample and the features of interest. A maximum scan rate of 122 Hz can be used. The maximum scan rate is usually related to the scan size. Scan rate above 10 Hz is used for small scans (typically 60 Hz for atomic-scale imaging with a 0.7 μm scanner). The scan rate should be lowered for large scans, especially if the sample surfaces are rough or contain large steps. Moving the tip quickly along the sample surface at high scan rates with large scan sizes will usually lead to a tip crash. Essentially, the scan rate should be inversely proportional to the scan size (typically 2–4 Hz for 1 μm , 0.5–1 Hz for 12 μm , and 0.2 Hz for 125 μm scan sizes). Scan rate in length/time, is equal to scan length divided by the scan rate in Hz. For example, for 10 $\mu\text{m} \times 10 \mu\text{m}$ scan size scanned at 0.5 Hz, the scan rate is 10 $\mu\text{m}/\text{s}$. Typically, 256 \times 256 data formats are most commonly used. The lateral resolution at larger scans is approximately equal to scan length divided by 256.

Figure 2.4 shows an example of STM images of an evaporated C₆₀ film on gold-coated freshly-cleaved mica taken at room temperature and ambient pressure (Bhushan et al. 1993). Images with atomic resolution at two scan sizes are obtained. Next we describe STM designs which are available for special applications.

Fig. 2.4 STM images of evaporated C₆₀ film on a gold-coated freshly-cleaved mica using a mechanically sheared Pt-Ir (80-20) tip in constant height mode (Bhushan et al. 1993)



2.2.2.1 Electrochemical STM

Electrochemical STM is used to perform and monitor the electrochemical reactions inside the STM. It includes a microscope base with an integral potentiostat, a short head with a 0.7 μm scan range and a differential preamp and the software required to operate the potentiostat and display the result of electrochemical reaction.

2.2.2.2 Standalone STM

Standalone STMs, which rest directly on the sample, are able to scan large samples. They are

similar to the standard STM, except the sample base has been eliminated.

2.2.3 STM Probe Construction

The STM probe should have a cantilever integrated with a sharp metal tip with a low aspect ratio (tip length/tip shank) to minimize flexural vibrations. Ideally, the tip should be atomically sharp, but in practice, most tip preparation methods produce a tip with a rather ragged profile and that

consists of several asperities with the one closest to the surface responsible for tunneling. STM cantilevers with sharp tips are typically fabricated from metal wires of tungsten (W), platinum-iridium (Pt-Ir), or gold (Au) and sharpened by grinding, cutting with a wire cutter or razor blade, field emission/evaporator, ion milling, fracture, or electrochemical polishing/etching (Nicolaides et al. 1988; Ibe et al. 1990). The two most commonly used tips are made from either a Pt-Ir (80/20) alloy or tungsten wire. Iridium is used to provide stiffness. The Pt-Ir tips are generally mechanically formed and are readily available. The tungsten tips are etched from tungsten wire with an electrochemical process, for example by using 1 molar KOH solution with a platinum electrode in an electrochemical cell at about 30 V. In general, Pt-Ir tips provide better atomic resolution than tungsten tips, probably due to the lower reactivity of Pt. But tungsten tips are more uniformly shaped and may perform better on samples with steeply sloped features. The tungsten wire diameter used for the cantilever is typically 250 μm with the radius of curvature ranging from 20 to 100 nm and a cone angle ranging from 10° to 60°, Fig. 2.5. The wire can be bent in an L shape, if required for use in the instrument. For calculations of normal spring constant and natural frequency of round cantilevers, see Sarid and Elings (1991).

For imaging of deep trenches, high-aspect-ratio controlled geometry (CG) Pt-Ir probes are used,

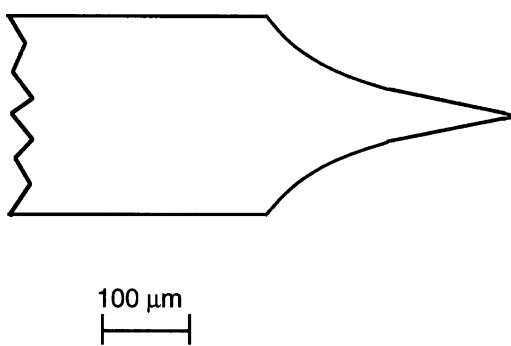


Fig. 2.5 Schematic of a typical tungsten cantilever with a sharp tip produced by electrochemical etching

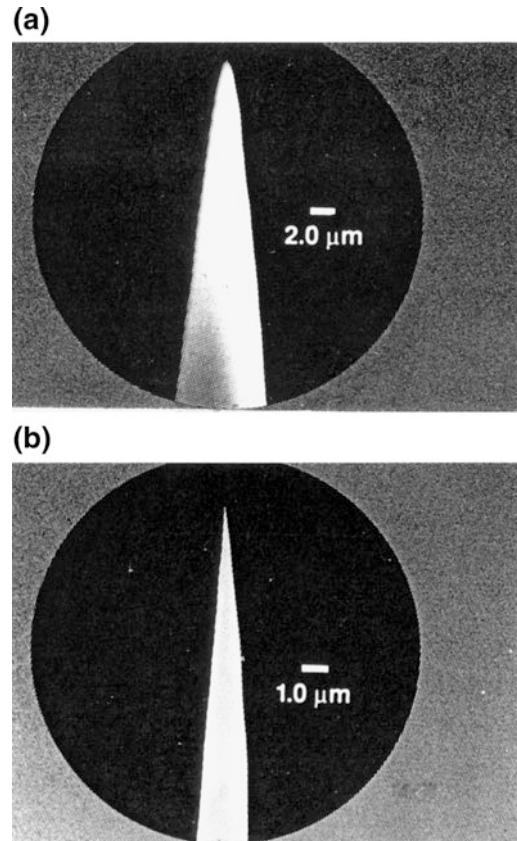


Fig. 2.6 Schematics of (a) CG Pt-Ir probe, and (b) CG Pt-Ir FIB milled probe

Fig. 2.6. These probes are electrochemically etched from Pt-Ir (80/20) wire and polished to a specific shape which is consistent from tip to tip. Probes have a full cone angle of approximately 15° and a tip radius of less than 50 nm. For imaging of very deep trenches ($>0.25 \mu\text{m}$) and nanofeatures, focused ion beam (FIB) milled CG probes with an extremely sharp tip (radius $<5 \text{ nm}$) are used. For electrochemistry, Pt/Ir probes are coated with a nonconducting film (not shown in the figure).

Pt alloy and W tips are very sharp and have high resolution, but are fragile and sometimes break when contacting a surface. Diamond tips have been used by Kaneko and Oguchi (1990). The diamond tip made conductive by boron ion implantation is found to be chip resistant.

2.3 Atomic Force Microscope

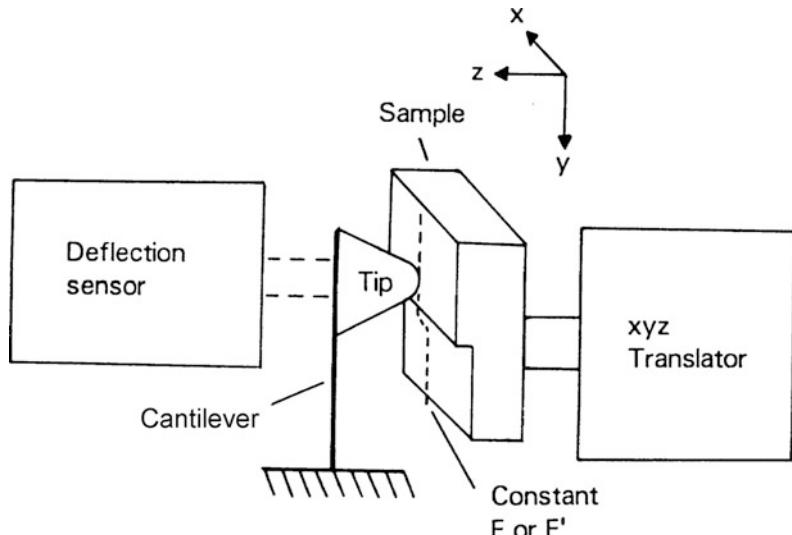
Like the STM, AFM relies on a scanning technique to produce very high resolution, 3-D images of sample surfaces. AFM measures ultrasmall forces (less than 1 nN) present between the AFM tip surface and a sample surface. These small forces are measured by measuring the motion of a very flexible cantilever beam having an ultrasmall mass. While STM requires that the surface to be measured be electrically conductive, AFM is capable of investigating surfaces of both conductors and insulators on an atomic scale if suitable techniques for measurement of cantilever motion are used. In the operation of high resolution AFM, the sample is generally scanned, rather than the tip as in STM, because AFM measures the relative displacement between the cantilever surface and reference surface, and any cantilever movement would add vibrations. For measurements of large samples, AFMs are available where the tip is scanned and the sample is stationary. As long as the AFM is operated in the so-called contact mode, little if any vibration is introduced.

The AFM combines the principles of the STM and the stylus profiler, Fig. 2.7. In an AFM, the force between the sample and tip is detected rather than the tunneling current to sense the proximity of the tip to the sample. The AFM can

be used either in static or dynamic mode. In the static mode, also referred to as repulsive mode or contact mode (Binnig et al. 1986), a sharp tip at the end of a cantilever is brought in contact with a sample surface. During initial contact, the atoms at the end of the tip experience a very weak repulsive force due to electronic orbital overlap with the atoms in the sample surface. The force acting on the tip causes a cantilever deflection which is measured by tunneling, capacitive, or optical detectors. The deflection can be measured to within 0.02 nm, so for typical cantilever spring constant of 10 N/m a force as low as 0.2 nN (corresponding normal pressure ~ 200 MPa for a Si_3N_4 tip with radius of about 50 nm against single-crystal silicon) can be detected. (To put these number in perspective, individual atoms and human hair are typically fraction of a nanometer and about 75 μm in diameter, respectively, and an eyelash and a drop of water have a mass of about 100 nN and 10 μN , respectively.)

In the dynamic mode of operation for the AFM, also referred to as attractive force imaging or noncontact imaging mode (noncontact-AFM or nc-AFM operation), the tip is brought in close proximity (within a few nm) to, and not in contact with the sample (Giessibl et al. 2011; Schirmeisen et al. 2011). The cantilever is deliberately vibrated either in amplitude

Fig. 2.7 Principle of operation of the AFM. Sample mounted on a piezoelectric scanner is scanned against a short tip and the cantilever deflection is measured, mostly, using a laser deflection technique. Force (contact mode) or force gradient (noncontact mode) is measured during scanning



modulation (AM) mode (Martin et al. 1987) or frequency modulation (FM) mode (Martin et al. 1987; Sarid and Elings 1991; Giessibl 1995; Anczykowski et al. 1996). Very weak van der Waals attractive forces are present at the tip-sample interface. In this technique, the normal pressure exerted at the interface is close to zero (desirable to avoid any surface deformation). In the two modes, surface topography is measured by laterally scanning the sample under the tip while simultaneously measuring the separation-dependent force or force gradient (derivative) between the tip and the surface, Fig. 2.7.

In the contact (static) mode, the interaction force between tip and sample is measured by measuring the cantilever deflection. In the non-contact (or dynamic) mode, the force gradient is obtained by vibrating the cantilever and measuring the shift of resonant frequency of the cantilever. To obtain topographic information, the interaction force is either recorded directly, or used as a control parameter for a feedback circuit that maintains the force or force derivative at a constant value. With an AFM operated in the contact mode, topographic images with a vertical resolution of less than 0.1 nm (as low as 0.01 nm) and a lateral resolution of about 0.2 nm have been obtained (Albrecht and Quate 1987; Binnig et al. 1987; Marti et al. 1987; Alexander et al. 1989; Meyer and Amer 1990a; Weisenhorn et al. 1991; Ruan and Bhushan 1994b). With a 0.01 nm displacement sensitivity, 10 nN–1 pN forces are measurable. These forces are comparable to the forces associated with chemical bonding, e.g., 0.1 μ N for an ionic bond and 10 pN for a hydrogen bond (Binnig et al. 1986). For further reading, see Rugar and Hansma (1990), Sarid (1991), Sarid and Elings (1991), Binnig (1992), Durig et al. (1992); Frommer (1992), Meyer (1992), Marti and Amrein (1993), Guntherodt et al. (1995) and Wickramasinghe (2000).

Lateral forces being applied at the tip during scanning in the contact mode affect roughness measurements (den Boef 1991). To minimize effects of friction and other lateral forces in the topography measurements in contact-mode AFMs and to measure topography of soft

surfaces, AFMs can be operated in the so-called or tapping mode or amplitude modulation mode (Maivald et al. 1991; Radmacher et al. 1992).

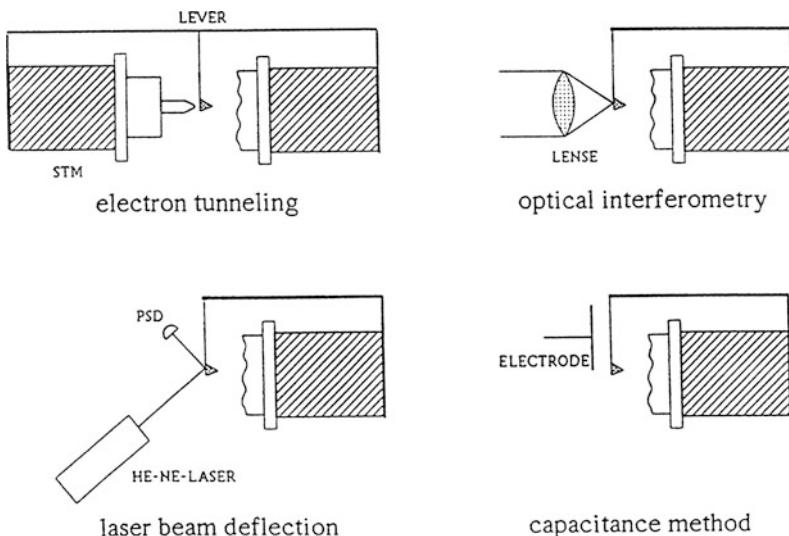
STM is ideal for atomic-scale imaging. To obtain atomic resolution with AFM, the spring constant of the cantilever should be weaker than the equivalent spring between atoms. For example, the vibration frequencies ω of atoms bound in a molecule or in a crystalline solid are typically 10^{13} Hz or higher. Combining this with the mass of the atoms m , on the order of 10^{-25} kg, gives interatomic spring constants k , given by $\omega^2 m$, on the order of 10 N/m (Rugar and Hansma 1990). (For comparison, the spring constant of a piece of household aluminum foil that is 4 mm long and 1 mm wide is about 1 N/m.) Therefore, a cantilever beam with a spring constant of about 1 N/m or lower is desirable. Tips have to be as sharp as possible. Tips with a radius ranging from 20 to 50 nm are commonly available.

Atomic resolution cannot be achieved with these tips at the normal load in the nN range. Atomic structures at these loads have been obtained from lattice imaging or by imaging of the crystal periodicity. Reported data show either perfectly ordered periodic atomic structures or defects on a larger lateral scale, but no well-defined, laterally-resolved, atomic-scale defects like those seen in images routinely obtained with STM. Interatomic forces with one or several atoms in contact are 20–40 or 50–100 pN, respectively. Thus, atomic resolution with AFM is only possible with a sharp tip on a flexible cantilever at a net repulsive force of 100 pN or lower (Ohnesorge and Binnig 1993). Upon increasing the force from 10 pN, Ohnesorge and Binnig (1993) observed that monoatomic step-lines were slowly wiped away and a perfectly ordered structure was left. This observation explains why mostly defect-free atomic resolution has been observed with AFM. Note that for atomic-resolution measurements, the cantilever should not be too soft to avoid jumps. Further note that measurements in the noncontact imaging mode may be desirable for imaging with atomic resolution.

The key component in AFM is the sensor for measuring the force on the tip due to its

interaction with the sample. A cantilever (with a sharp tip) with extremely low spring constants is required for high vertical and lateral resolutions at small forces (0.1 nN or lower), but at the same time, a high resonant frequency (about 10–100 kHz) is necessary in order to minimize the sensitivity to vibration noise from the building near 100 Hz. This requires a spring with extremely low vertical spring constant (typically 0.05–1 N/m) as well as low mass (on the order of 1 ng). Today, the most advanced AFM cantilevers are microfabricated from silicon or silicon nitride using photolithographic techniques. (For further details on cantilevers, see a later section). Typical lateral dimensions are on the order of 100 μm , with the thicknesses on the order of 1 μm . The force on the tip due to its interaction with the sample is sensed by detecting the deflection of the compliant lever with a known spring constant. This cantilever deflection (displacement smaller than 0.1 nm) has been measured by detecting tunneling current similar to that used in STM in the pioneering work of Binnig et al. (1986) and later used by Giessibl et al. (1991), by capacitance detection (Neubauer et al. 1990; Goddenhenrich et al. 1990), piezoresistive detection (Stahl et al. 1994; Kassing and Oesterschulze 1997), and by four optical techniques namely (1) by optical interferometry (Mate et al. 1987; Erlandsson et al. 1988; Mate 1992; Jarvis et al. 1993) and with the use of optical fibers (Rugar et al. 1989; Albrecht et al. 1992), (2) by optical polarization detection (Schoenenberger and Alvarado 1989, 1990), (3) by laser diode feedback (Sarid et al. 1988), and (4) by optical (laser) beam deflection (Meyer and Amer 1988, 1990a, b; Alexander et al. 1989; Marti et al. 1990). Schematics of the four more commonly used detection systems are shown in Fig. 2.8. The tunneling method originally used by Binnig et al. (1986) in the first version of AFM uses a second tip to monitor the deflection of the cantilever with its force sensing tip. Tunneling is rather sensitive to contaminants, and the interaction between the tunneling tip and the rear side of the cantilever can become comparable to the interaction between the tip and sample. Tunneling is rarely used and is mentioned earlier for historical purposes. Giessibl et al. (1991) have used it for a low temperature AFM/STM design. In contrast to tunneling, other deflection sensors are far away from the cantilever at distances of microns to tens of mm. The optical techniques are believed to be more sensitive, reliable, and easily implemented detection methods than others (Sarid and Elings 1991; Meyer 1992). Optical beam deflection method has the largest working distance, is insensitive to distance changes, and is capable of measuring angular changes (friction forces). Therefore, it is

Fig. 2.8 Schematics of the four detection systems for measurement of cantilever deflection. In each set up, the sample mounted on piezoelectric body is shown on the right, the cantilever in the middle, and the corresponding deflection sensor on the left (Meyer 1992)



the most commonly used in the commercial SPMs.

Almost all SPMs use piezo translators to scan the sample or, alternatively, to scan the tip. An electric field applied across a piezoelectric material causes a change in the crystal structure, with expansion in some directions and contraction in others. A net change in volume also occurs (Ashcroft and Mermin 1976). The first STM used a piezo tripod for scanning (Binnig et al. 1982). The piezo tripod is one way to generate three-dimensional movement of a tip attached to its center. However, the tripod needs to be fairly large (~ 50 mm) to get a suitable range. Its size and asymmetric shape makes it susceptible to thermal drift. Tube scanners are widely used in AFMs (Binnig and Smith 1986). These provide ample scanning range within a small size. Control electronics systems for AFMs can use either analog or digital feedback. Digital feedback circuits are better suited for ultra-low noise operation.

Images from AFMs need to be processed. An ideal AFM is a noise free device that images a sample with perfect tips of known shape and has perfect linear scanning piezo. In reality, scanning devices are affected by distortions for which corrections must be made. The distortions can be linear and nonlinear. Linear distortions mainly result from imperfections in the machining of the piezo translators, causing cross talk between the Z-piezo to the X- and Y-piezos, and vice versa. Nonlinear distortions mainly result because of presence of hysteresis loop in piezoelectric ceramics. These may also result if the scan frequency approaches the upper frequency limit of the X- and Y-drive amplifiers or the upper frequency limit of the feedback loop (z-component). In addition, electronic noise may be present in the system. The noise is removed by digital filtering in the real space (Park and Quate 1987) or in the spatial frequency domain (Fourier space) (Cooley and Tukey 1965).

Processed data consists of many tens of thousands of points per plane (or data set). The output of the first STM and AFM images were recorded on an X-Y chart recorder, with z-value plotted against the tip position in the fast scan

direction. Chart recorders have slow response so computers are used for display of the data. The data are displayed as wire mesh display or gray scale display (with at least 64 shades of gray).

2.3.1 Binnig et al.'s Design

In the first AFM design developed by Binnig et al. (1986), AFM images were obtained by measurement of the force on a sharp tip created by the proximity to the surface of the sample mounted on a 3-D piezoelectric scanner. Tunneling current between STM tip and the backside of the cantilever beam with attached tip was measured to obtain the normal force. This force was kept at a constant level with a feedback mechanism. The STM tip was also mounted on a piezoelectric element to maintain the tunneling current at a constant level.

2.3.2 Commercial AFM

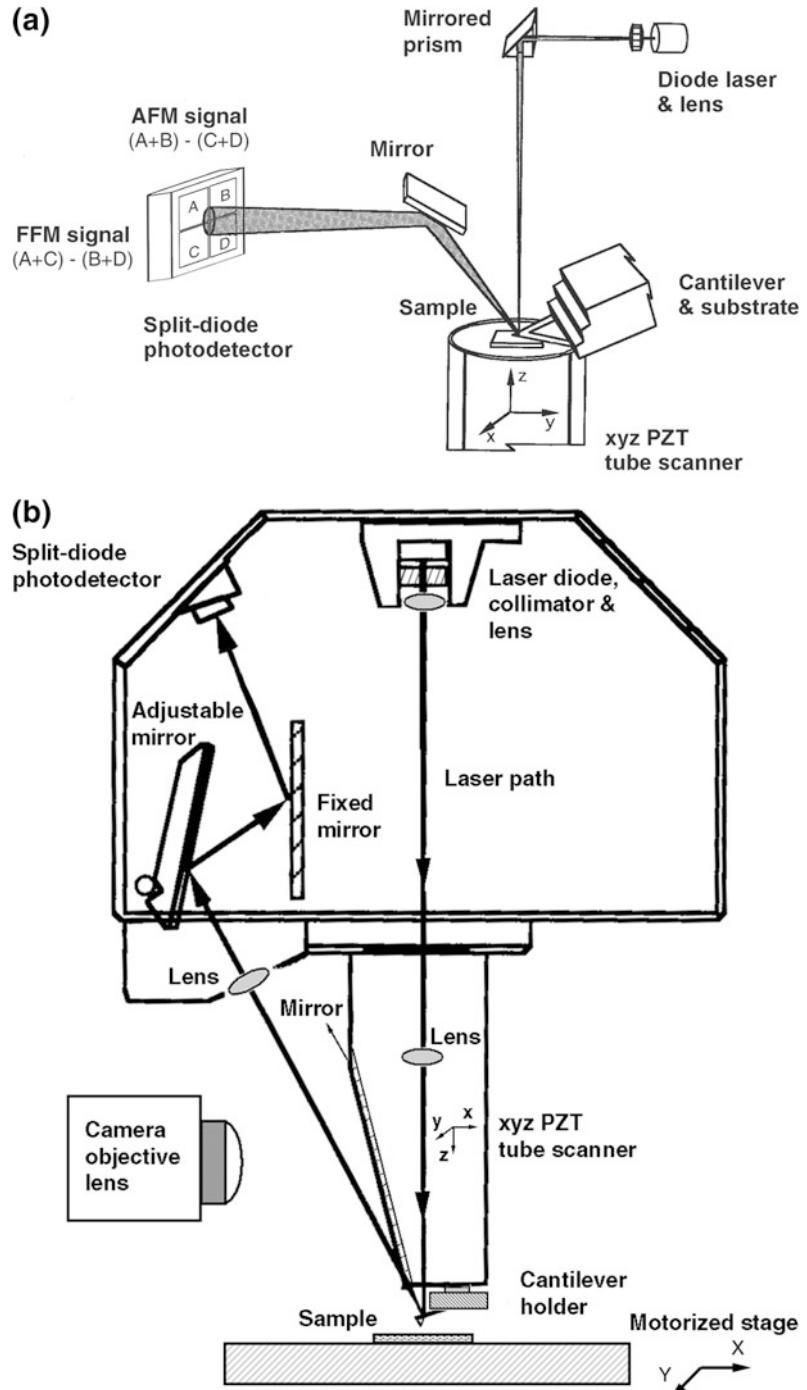
A review of early designs of AFMs is presented by Bhushan (1999a). There are a large number of commercial AFMs available on the market. Major manufacturers of AFMs for use in ambient environment are: Bruker Instruments, Inc. (formerly Digital Instruments Inc. and later Veeco Instruments Inc.), Santa Barbara, California; Agilent Technologies, Santa Clara, California, Park Systems Corp., Suwon, Korea; Asylum Research, Santa Clara, California; NT-MDT Spectrum Instruments, Moscow Russia; Nanosurf AF, Liestal, Switzerland; Seiko Instruments, Japan; and Olympus, Japan. A major manufacturer for AFM/STMs for use in UHV and low temperature (LT) environment is Scienta Omicron GMBH, Taunusstein, Germany.

We describe here two typical commercial AFMs—small sample and large sample AFMs—for operation in the contact mode, produced by Bruker Instruments, Santa Barbara, CA, with scanning lengths ranging from about 0.7 μm (for atomic resolution) to about 125 μm (Alexander et al. 1989; Bhushan and Ruan 1994; Ruan and Bhushan 1994a, b). The original design of these

AFMs comes from Meyer and Amer (1988). Basically, the AFM scans the sample in a raster pattern while outputting the cantilever deflection error signal to the control station. The cantilever

deflection (or the force) is measured using laser deflection technique, Fig. 2.9. The DSP in the workstation controls the z-position of the piezo based on the cantilever deflection error signal.

Fig. 2.9 Principles of operation of (a) a commercial small sample AFM/FFM, and (b) a large sample AFM/FFM



The AFM operates in both the “constant height” and “constant force” modes. The DSP always adjusts the height of the sample under the tip based on the cantilever deflection error signal, but if the feedback gains are low, the piezo remains at a nearly “constant height” and the cantilever deflection data is collected. With the high gains, the piezo height changes to keep the cantilever deflection nearly constant (therefore the force is constant) and the change in piezo height is collected by the system.

To describe further the operation principles of the small-sample commercial AFM shown in Fig. 2.9a, the sample, generally no larger than $10\text{ mm} \times 10\text{ mm}$, is mounted on a PZT tube scanner which consists of separate electrodes to scan precisely the sample in the x-y plane in a raster pattern, and to move the sample in the vertical (z) direction. A sharp tip at the free end of a flexible cantilever is brought in contact with the sample. Features on the sample surface cause the cantilever to deflect in the vertical and lateral directions as the sample moves under the tip. A laser beam from a diode laser (5 mW max peak output at 670 nm) is directed by a prism onto the back of a cantilever near its free end, tilted downward at about 10° with respect to the horizontal plane. The reflected beam from the vertex of the cantilever is directed through a mirror onto a quad photodetector (split photodetector with four quadrants) (commonly called position-sensitive detector or PSD, produced by Silicon Detector Corp., Camarillo, California). The differential signal from the top and bottom photodiodes provides the AFM signal, which is a sensitive measure of the cantilever vertical deflection. Topographic features of the sample cause the tip to deflect in the vertical direction as the sample is scanned under the tip. This tip deflection will change the direction of the reflected laser beam, changing the intensity difference between the top and bottom sets of photodetectors (AFM signal). In the AFM operating mode called the height mode, for topographic imaging or for any other operation in which the applied normal force is to be kept a constant, a feedback circuit is used to modulate the voltage applied to the PZT scanner to adjust

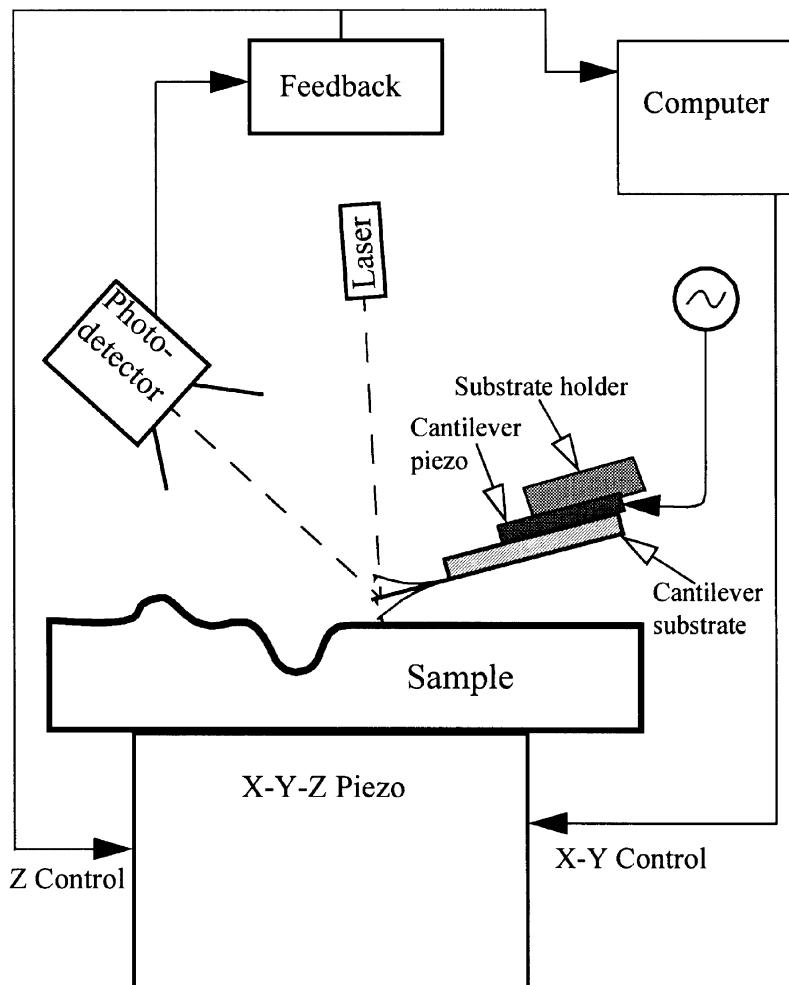
the height of the PZT. This ensures that the cantilever vertical deflection (given by the intensity difference between the top and bottom detector), and consequently normal load, will remain constant during scanning. The PZT height variation is thus a direct measure of the surface roughness of the sample.

In a large sample AFM, both force sensors using optical deflection method and scanning unit are mounted on the microscope head, Fig. 2.9b. Because of vibrations added by cantilever movement, lateral resolution of this design is somewhat poorer than the design in Fig. 2.9a in which the sample is scanned instead of cantilever beam. The advantage of the large sample AFM is that large samples can be measured readily.

Most AFMs can be used for topography measurements in the so-called tapping mode (intermittent contact mode), also referred to as dynamic force microscopy in the amplitude mode (AM) mentioned earlier. In the tapping mode, during scanning over the surface, the cantilever/tip assembly is sinusoidally vibrated by a piezo mounted above it, and the oscillating tip slightly taps the surface at the resonant frequency of the cantilever (70–400 kHz) with a constant (20–100 nm) oscillating amplitude introduced in the vertical direction with a feedback loop keeping the average normal force constant, Fig. 2.10. The oscillating amplitude is kept large enough so that the tip does not get stuck to the sample because of adhesive attractions. The tapping mode is used in topography measurements to minimize effects of friction and other lateral forces to measure topography of soft surfaces.

Topographic measurements are made at any scanning angle. At a first instance, scanning angle may not appear to be an important parameter. However, the friction force between the tip and the sample will affect the topographic measurements in a parallel scan (scanning along the long axis of the cantilever). Therefore a perpendicular scan may be more desirable. Generally, one picks a scanning angle that gives the same topographic data in both directions. This angle may be slightly different than that for the perpendicular scan.

Fig. 2.10 Schematic of tapping mode used for surface roughness measurements



For measurement of friction force being applied at the tip surface during sliding, left hand and right hand sets of quadrants of the photodetector are used. In the so-called friction mode, the sample is scanned back and forth in a direction orthogonal to the long axis of the cantilever beam. A friction force between the sample and the tip will produce a twisting of the cantilever. As a result, the laser beam will be reflected out of the plane defined by the incident beam and the beam reflected vertically from an untwisted cantilever. This produces an intensity difference of the laser beam received in the left hand and right hand sets of quadrants of the photodetector. The intensity difference between the two sets of detectors (FFM signal) is directly

related to the degree of twisting and hence to the magnitude of the friction force. This method provides three-dimensional maps of friction force. One problem associated with this method is that any misalignment between the laser beam and the photodetector axis introduces error in the measurement. However, by following the procedures developed by Ruan and Bhushan (1994a) in which the average FFM signal for the sample scanned in two opposite directions is subtracted from the friction profiles of each of the two scans, the misalignment effect is eliminated. By following the friction force calibration procedures developed by Ruan and Bhushan (1994a), voltages corresponding to friction forces can be converted to force units. The coefficient of

friction is obtained from the slope of friction force data measured as a function of normal loads typically ranging from 10 to 150 nN. This approach eliminates any contributions due to adhesive forces (Bhushan et al. 1994). For calculation of the coefficient of friction based on a single point measurement, friction force should be divided by the sum of applied normal load and intrinsic adhesive force. Furthermore, note that for a single asperity contact, the coefficient of friction is not independent of load. For further details, refer to a later section.

The tip is scanned in such a way that its trajectory on the sample forms a triangular pattern, Fig. 2.11. Scanning speeds in the fast and slow scan directions depend on the scan area and scan frequency. Scan sizes ranging from less than $1 \text{ nm} \times 1 \text{ nm}$ to $125 \mu\text{m} \times 125 \mu\text{m}$ and scan rates from less than 0.5–122 Hz typically can be used. Higher scan rates are used for smaller scan lengths. For example, scan rates in the fast and slow scan directions for an area of $10 \mu\text{m} \times 10 \mu\text{m}$ scanned at 0.5 Hz are $10 \mu\text{m/s}$ and 20 nm/s , respectively.

We now describe the construction of a small sample AFM in more detail. It consists of three main parts: the optical head that senses the cantilever deflection, a PZT tube scanner that controls the scanning motion of the sample mounted on its one end, and the base that supports the scanner and head and includes circuits for the deflection signal, Fig. 2.12a. The AFM connects directly to a control system. The optical head

consists of laser diode stage, photodiode stage preamp board, cantilever mount and its holding arm, and deflection beam reflecting mirror, Fig. 2.12b. The laser diode stage is a tilt stage used to adjust the position of the laser beam relative to the cantilever. It consists of the laser diode, collimator, focusing lens, baseplate, and the X and Y laser diode positioners. The positioners are used to place the laser spot on the end of the cantilever. The photodiode stage is an adjustable stage used to position the photodiode elements relative to the reflected laser beam. It consists of the split photodiode, the base plate, and the photodiode positioners. The deflection beam reflecting mirror is mounted on the upper left in the interior of the head which reflects the deflected beam toward the photodiode. The cantilever mount is a metal (for operation in air) or glass (for operation in water) block which holds the cantilever firmly at the proper angle, Fig. 2.12d. Next, the tube scanner consists of an Invar cylinder holding a single tube made of piezoelectric crystal which provides the necessary three-dimensional motion to the sample. Mounted on top of the tube is a magnetic cap on which the steel sample puck is placed. The tube is rigidly held at one end with the sample mounted on the other end of the tube. The scanner also contains three fine-pitched screws which form the mount for the optical head. The optical head rests on the tips of the screws which are used to adjust the position of the head relative to the sample. The scanner fits into the scanner support ring mounted on the base of the microscope, Fig. 2.12c. The stepper motor is controlled manually with the switch on the upper surface of the base and automatically by the computer during the tip engage and tip-withdraw processes.

The scan sizes available for these instruments are 0.7 , 12 , and $125 \mu\text{m}$. The scan rate must be decreased as the scan size is increased. A maximum scan rate of 122 Hz can be used. Scan rates of about 60 Hz should be used for small scan lengths ($0.7 \mu\text{m}$). Scan rates of 0.5 – 2.5 Hz should be used for large scans on samples with tall features. High scan rates help reduce drift, but they can only be used on flat samples with

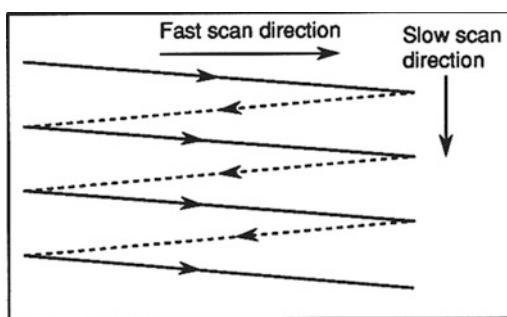
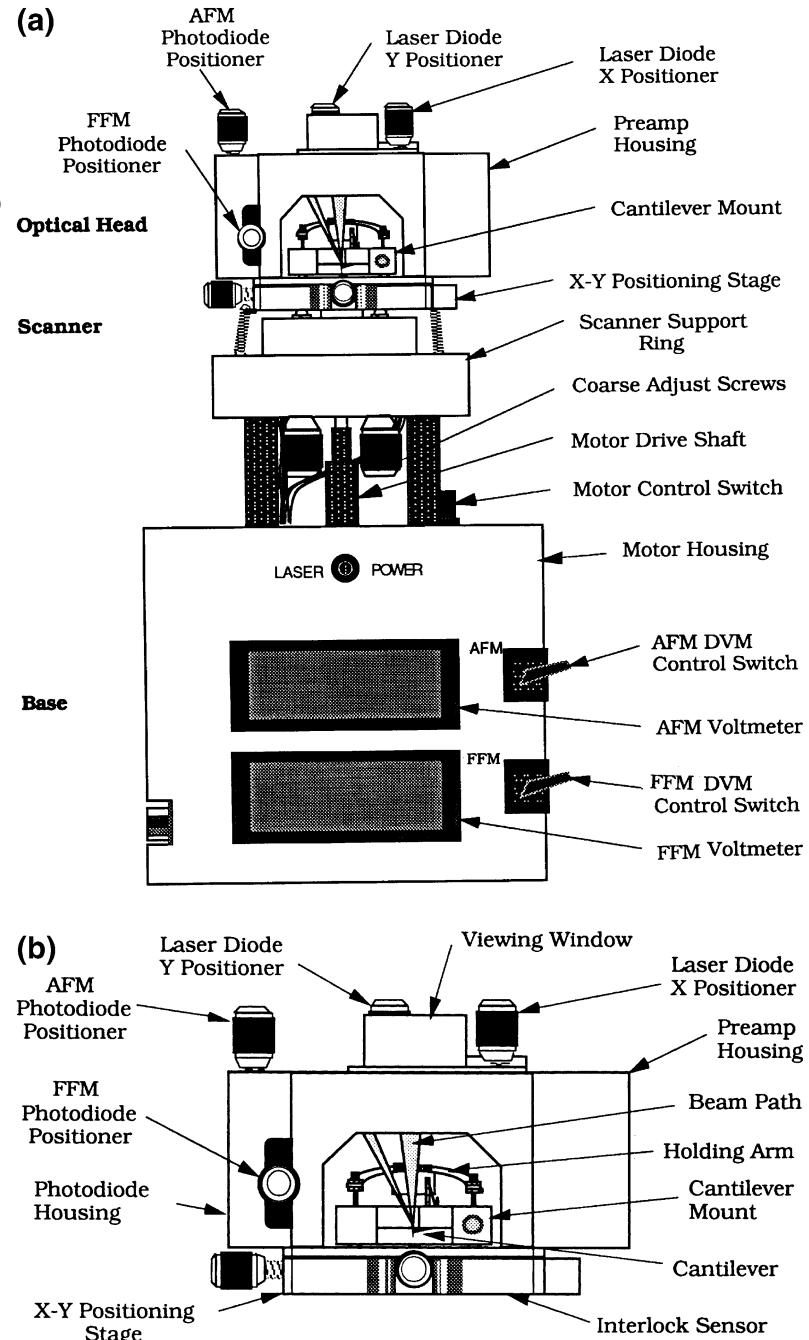


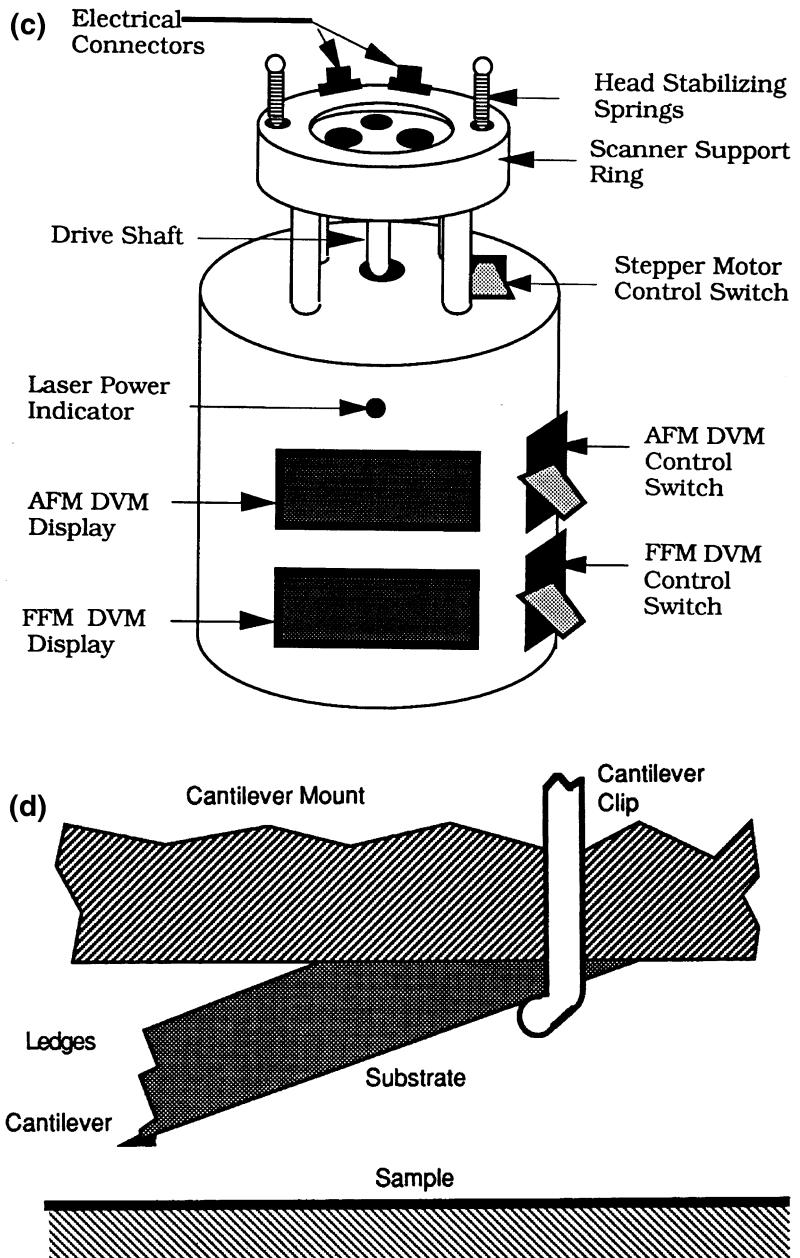
Fig. 2.11 Schematic of triangular pattern trajectory of the AFM tip as the sample is scanned in two dimensions. During imaging, data are recorded only during scans along the solid scan lines

Fig. 2.12 Schematics of a commercial Multimode AFM/FFM made by Bruker Instruments Inc. (a) front view, (b) optical head, (c) base, and (d) cantilever substrate mounted on cantilever mount (not to scale)



small scan sizes. Scan rate, or scanning speed in length/time in the fast scan direction, is equal to twice the scan length times the scan rate in Hz, and in the slow direction, it is equal to scan length times the scan rate in Hz divided by

number of data points in the transverse direction. For example, for $10 \mu\text{m} \times 10 \mu\text{m}$ scan size scanned at 0.5 Hz, the scan rates in the fast and slow scan directions are $10 \mu\text{m/s}$ and 20 nm/s , respectively. Normally 256×256 data points

Fig. 2.12 (continued)

are taken for each image. The lateral resolution at larger scans is approximately equal to scan length divided by 256. The piezo tube requires x-y calibration, which is carried out by imaging an appropriate calibration standard. Cleaved graphite is used for small scan heads while two-dimensional grids (a gold plating ruling) can be used for longer range heads.

Examples of AFM images of freshly-cleaved highly-oriented pyrolytic (HOP) graphite and mica surfaces are shown in Fig. 2.13 (Albrecht and Quate 1987; Marti et al. 1987; Ruan and Bhushan 1994b). Images with near atomic resolution are obtained.

Force calibration mode is used to study interaction between the cantilever and the sample

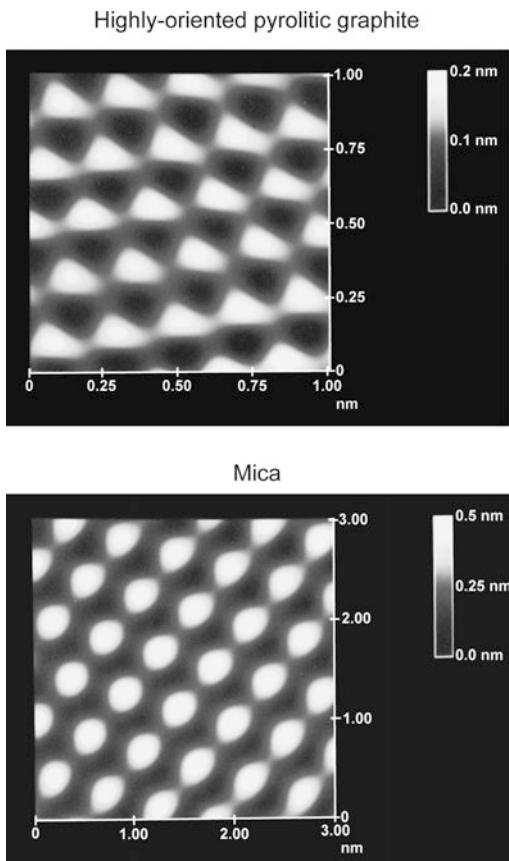


Fig. 2.13 Typical AFM images of freshly-cleaved highly-oriented pyrolytic graphite and mica surfaces taken using a square pyramidal Si_3N_4 tip

surface. In the force calibration mode, the X and Y voltages applied to the piezo tube are held at zero, and a sawtooth voltage is applied to the Z electrode of the piezo tube, Fig. 2.14a. The force measurement starts with the sample far away and the cantilever in its rest position. As a result of the applied voltage, the sample is moved up and down relative to the stationary cantilever tip. As the piezo moves the sample up and down, the cantilever deflection signal from the photodiode is monitored. The force-distance curve, a plot of the cantilever tip deflection signal as a function of the voltage applied to the piezo tube, is obtained. Figure 2.14b shows a typical force-distance curve showing the various features of the curve. The arrow heads reveal the direction of piezo travel. As the piezo extends, it

approaches the tip, which is at this point in free air and hence shows no deflection. This is indicated by the flat portion of the curve. As the tip approaches the sample within a few nanometers (point A), an attractive force exists between the atoms of the tip surface and the atoms of the sample surface. The tip is pulled towards the sample and contact occurs at point B on the graph. From this point on, the tip is in contact with the surface and as the piezo further extends, the tip gets further deflected. This is represented by the sloped portion of the curve. As the piezo retracts, the tip goes beyond the zero deflection (flat) line into the adhesive regime because of attractive forces (van der Waals forces and long range meniscus forces). At point C in the graph, the tip snaps free of the adhesive forces, and is in free air again. The horizontal distance between points B and C along the retrace line gives the distance moved by the tip in the adhesive regime. This distance multiplied by the stiffness of the cantilever gives the adhesive force. Incidentally, the horizontal shift between the loading and unloading curves results from the hysteresis in the PZT tube (Bhushan 1999a).

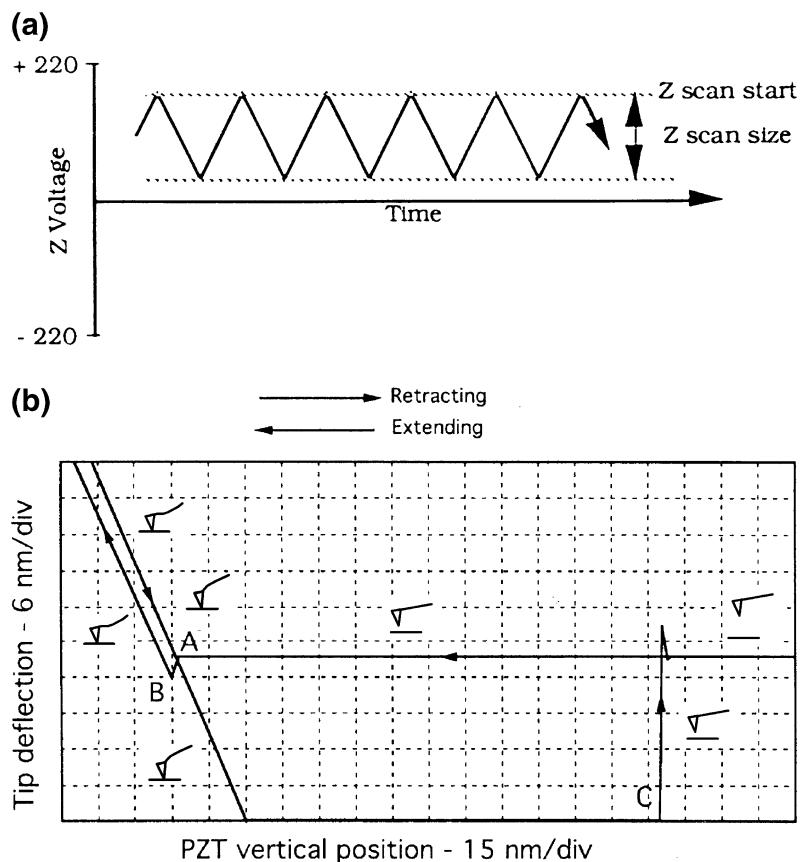
2.3.2.1 Multimode Capabilities

The multimode AFM can be used for topography measurements in the contact mode and AM or tapping mode as described earlier. It also can be used for measurements of lateral (friction) force, electric force gradients, and magnetic force gradients.

The multimode AFM using a grounded conducting tip can measure electric field gradients by oscillating the tip near its resonant frequency. When the lever encounters a force gradient from the electric field, the effective spring constant of the cantilever is altered, changing its resonant frequency. Depending on which side of the resonance curve is chosen, the oscillation amplitude of the cantilever increases or decreases due to the shift in the resonant frequency. By recording the amplitude of the cantilever, an image revealing the strength of the electric field gradient is obtained.

In a magnetic force microscope (MFM) used with a magnetically-coated tip, static cantilever

Fig. 2.14 (a) Force calibration Z waveform, and (b) a typical force-distance curve for a tip in contact with a sample. Contact occurs at point B; tip breaks free of adhesive forces at point C as the samples moves away from the tip



deflection is detected that occurs when a magnetic field exerts a force on the tip and the MFM images of magnetic materials can be produced. MFM sensitivity can be enhanced by oscillating the cantilever near its resonant frequency. When the tip encounters a magnetic force gradient, the effective spring constant, and hence the resonant frequency, is shifted. By driving the cantilever above or below the resonant frequency, the oscillation amplitude varies as the resonance shifts. An image of magnetic field gradients is obtained by recording the oscillation amplitude as the tip is scanned over the sample.

Topographic information is separated from the electric field gradients and magnetic field images by using a so-called lift mode. Measurements in lift mode are taken in two passes over each scan line. On the first pass, topographical information is recorded in the standard tapping mode where the oscillating cantilever lightly taps the surface.

On the second pass, the tip is lifted to a user-selected separation (typically 20–200 nm) between the tip and local surface topography. By using the stored topographical data instead of the standard feedback, the separation remains constant without sensing the surface. At this height, cantilever amplitudes are sensitive to electric field force gradients or relatively weak but long-range magnetic forces without being influenced by topographic features. Two-pass measurements are taken for every scan line, producing separate topographic and magnetic force images.

2.3.2.2 Electrochemical AFM

This option allows study of electrochemical reactions in the AFM. It includes a potentiostat, a fluid cell with a transparent cantilever holder and electrodes, and the software required to operate the potentiostat and display the results of the electrochemical reaction.

2.3.3 AFM Probe Construction

Various probes (cantilevers and tips) are used for AFM studies. The cantilever stylus used in the AFM should meet the following criteria: (1) low normal spring constant (stiffness), (2) a high resonant frequency, (3) a high quality factor of the cantilever Q , (4) high lateral spring constant (stiffness), (5) short cantilever length, (6) incorporation of components (such as mirror) for deflection sensing, and (7) a sharp protruding tip (Albrecht et al. 1990). In order to register a measurable deflection with small forces, the cantilever must flex with a relative low force (on the order of few nN) requiring vertical spring constants of 10^{-2} to 10^2 N/m for atomic resolution in the contact profiling mode. The data rate or imaging rate in the AFM is limited by the mechanical resonant frequency of the cantilever. To achieve a large imaging bandwidth, the AFM cantilever should have resonant frequency greater than about 10 kHz (preferably 30–100 kHz) in order to make the cantilever the least sensitive part of the system. Fast imaging rates are not just a matter of convenience, since the effects of thermal drifts are more pronounced with slow-scanning speeds. The combined requirements of a low spring constant and a high resonant frequency is met by reducing the mass of the cantilever. The quality factor Q ($=\omega_R/(c/m)$ where ω_R is the resonant frequency of the damped oscillator, c is the damping constant, and m is the mass of the oscillator) should have a high value for some applications. For example, resonance curve detection is a sensitive modulation technique for measuring small force gradients in noncontact imaging. Increasing the Q increases the sensitivity of the measurements. Mechanical Q values of 100–1000 are typical. In contact modes, Q is less important. A high lateral spring constant in the cantilever is desirable to reduce the effect of lateral forces in the AFM as frictional forces can cause appreciable lateral bending of the cantilever. Lateral bending results in error in the topography measurements. For friction measurements, cantilevers with less lateral rigidity are preferred. A sharp protruding tip must be formed at the end of the cantilever to

provide a well-defined interaction with the sample over a small area. The tip radius should be much smaller than the radii of corrugations in the sample in order for these to be measured accurately. The lateral spring constant depends critically on the tip length. Additionally, the tip should be centered at the free end.

In the past, cantilevers have been cut by hand from thin metal foils or formed from fine wires. Tips for these cantilevers were prepared by attaching diamond fragments to the ends of the cantilevers by hand, or, in the case of wire cantilevers, electrochemically etching the wire to a sharp point. Several cantilever geometries for wire cantilevers have been used. The simplest geometry is the L-shaped cantilever, usually made by bending a wire at a 90° angle. Other geometries include single-V and double-V geometries with a sharp tip attached at the apex of V, and double-X configuration with a sharp tip attached at the intersection (Marti et al. 1988; Burnham and Colton 1989). These cantilevers can be constructed with high vertical spring constants. For example, a double-cross cantilever with an effective spring constant of 250 N/m was used by Burnham and Colton (1989). The small size and low mass needed in the AFM make hand fabrication of the cantilever a difficult process with poor reproducibility. Conventional micro-fabrication techniques are ideal for constructing planar thin-film structures which have submicron lateral dimensions. Triangular (V-shaped) cantilevers have improved (higher) lateral spring constants in comparison to rectangular cantilevers. In terms of spring constants, triangular cantilevers are approximately equivalent to two rectangular cantilevers in parallel (Albrecht et al. 1990). Although the macroscopic radius of a photolithographically patterned corner is seldom much less than about 50 nm, microscopic asperities on the etched surface provide tips with near atomic dimensions.

Cantilevers and tips have been used from a whole range of materials. Most common are cantilevers/tips made of Si_3N_4 , Si, and stainless steel/diamond. Besides the geometry, the material parameters that determine the resonant frequency of the cantilever are Young's modulus

Table 2.2 Relevant properties of commonly used materials for cantilevers and diamond used in the some tips

Property	Diamond (for tips)	Si_3N_4	Si	Stainless steel	W	Ir
Young's Modulus (E) (GPa)	900–1050	310	130–188	210	350	530
Density (ρ_g) (kg/m^3)	3515	3180	2330	8050	19,310	—
Microhardness (GPa)	78.4–102	19.6	9–10	2–4	3.2	~3
Speed of sound ($\sqrt{E/\rho}$) (m/s)	17,000	9900	8200	5080	4250	5300

and density. Table 2.2 shows the relevant properties of commonly used materials and the speed of sound, indicative of the resonant frequency for a cantilever beam. Hardness is important to judge the durability of the tips, and is also listed in the table. Materials used for STM cantilevers are also included.

Silicon nitride cantilevers with integrated tips are less expensive than those made of other materials. They are very rugged and well suited to imaging in almost all environments. They are especially compatible to organic and biological materials. Microfabricated silicon nitride triangular beams with integrated square pyramidal tips made of plasma-enhanced chemical vapor deposition (PECVD) are the most commonly used (Albrecht et al. 1990). Figure 2.15a shows a schematic example of four cantilever beams, each with different sizes and spring constants, arranged in pairs, attached to a borosilicate glass substrate (Pyrex) and marketed by various manufacturers including Bruker Instruments, Inc. Each pair of triangular cantilevers measures 115 and 193 μm from the substrate to the apex, with base widths of 122 and 205 μm respectively. These cantilevers are commercially available with wide (top) and narrow (bottom) legs, all in the same thickness of 0.6 μm . Calculated spring constant and measured natural frequencies for each of the configurations are listed in Table 2.3. The most commonly used cantilever beam is the 115 μm -long, wide-legged cantilever (vertical spring constant = 0.58 N/m). Cantilevers with smaller spring constants should be used on softer samples. Pyramidal tips are highly symmetric with ends having a radius of about 20–50 nm. The tip side walls have a slope of 35°, and the length of the edges of the tip at the cantilever base is about 4 μm .

An alternative to silicon nitride cantilevers with integrated tips are microfabricated single-crystal silicon cantilevers with integrated tips. Si tips are sharper than Si_3N_4 tips because they are directly formed by the anisotropic etch in single-crystal Si rather than using an etch pit as a mask for deposited materials (Wolter et al. 1991). Etched single-crystal n-type silicon rectangular cantilevers with square pyramidal tips with a lower radius of less than 10 nm for contact and tapping mode (tapping mode etched silicon probe or TESP) AFMs are commercially available from various manufacturers including Bruker Instruments, Inc., Fig. 2.15b. Spring constants and resonant frequencies are also presented in the Fig. 2.15b.

Commercial triangular Si_3N_4 cantilevers have a typical width-thickness ratio of 10–30, which results in 100–1000 times stiffer spring constants in the lateral direction compared to the normal direction. Therefore these cantilevers are not well suited for torsion. For friction measurements, the torsional spring constant should be minimized in order to be sensitive to lateral forces. Rather long cantilevers with small thicknesses and large tip lengths are most suitable. Rectangular beams have lower torsional spring constants in comparison to the triangular (V-shaped) cantilevers. Table 2.4 lists the spring constants (with full length of the beam used) in three directions of typical rectangular beams. We note that lateral and torsional spring constants are about two orders of magnitude larger than the normal spring constants. A cantilever beam required for the tapping mode is quite stiff and may not be sensitive enough for friction measurements. Meyer et al. (1992) used a specially designed rectangular silicon cantilever with length = 200 m, width = 21 μm , thickness = 0.4 μm , tip

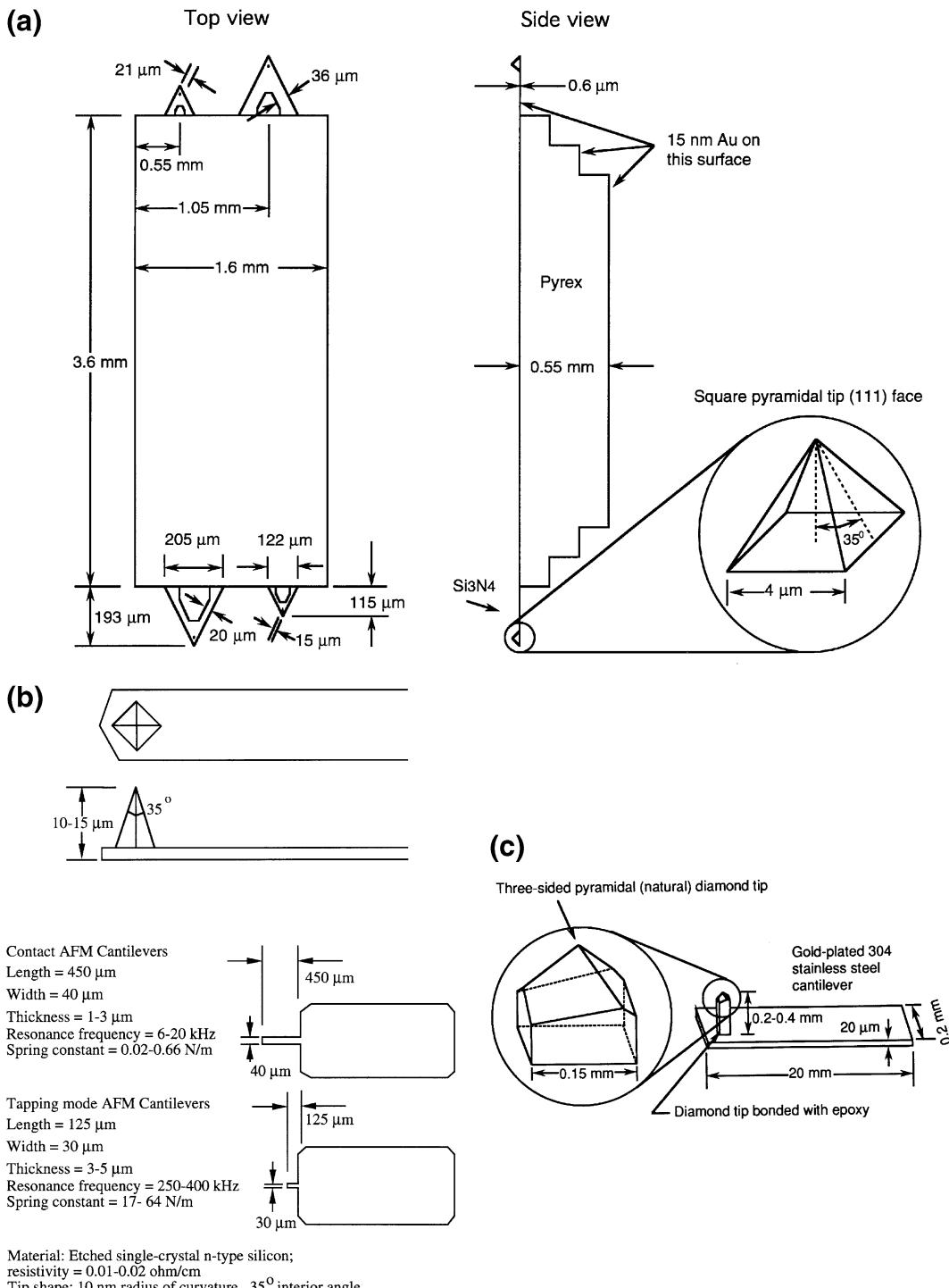


Fig. 2.15 Schematics of (a) triangular cantilever beam with square pyramidal tips made of PECVD Si₃N₄, (b) rectangular cantilever beams with square pyramidal tips made of etched single-crystal silicon, and (c) rectangular cantilever stainless steel beam with three-sided pyramidal natural diamond tip

Table 2.3 Measured vertical spring constants and natural frequencies of triangular (V-shaped) cantilevers made of PECVD Si₃N₄ (Data provided by Bruker Instruments, Inc.)

Cantilever dimension	Spring constant (k_z) (N/m)	Natural frequency (ω_0) (kHz)
115 μm long, narrow leg	0.38	40
115 μm long, wide leg	0.58	40
193 μm long, narrow leg	0.06	13–22
193 μm long, wide leg	0.12	13–22

Table 2.4 Vertical (k_z), lateral (k_y), and torsional (k_{yT}) spring constants of rectangular cantilevers made of Si (IBM) and PECVD Si₃N₄ (Bruker Instruments, Inc.)

Dimensions/stiffness	Si cantilever	Si ₃ N ₄ cantilever
Length (L) (μm)	100	100
Width (b) (μm)	10	20
Thickness (h) (μm)	1	0.6
Tip length (ℓ) (μm)	5	3
k_z (N/m)	0.4	0.15
k_y (N/m)	40	175
k_{yT} (N/m)	120	116
ω_0 (kHz)	~90	~65

Note $k_z = Ebh^3/4L^3$, $k_y = Eb^3h/4\ell^3$, $k_{yT} = Gbh^3/3L\ell^2$, and $\omega_0 = [k_z/(m_c + 0.24 bhL\rho)]^{1/2}$ where E is Young's modulus, G is the modulus of rigidity [=E/2(1 + v), where v is the Poisson's ratio], ρ is the mass density of the cantilever, and m_c is the concentrated mass of the tip (~4 ng) (Sarid and Elings 1991). For Si, E = 130 GPa, ρ_g = 2300 kg/m³, and v = 0.3. For Si₃N₄, E = 150 GPa, ρ_g = 3100 kg/m³, and v = 0.3

length = 12.5 μm , and shear modulus = 50 GPa, giving a normal spring constant of 0.007 N/m and torsional spring constant of 0.72 N/m which gives a lateral force sensitivity of 10 pN and an angle of resolution of 10^{-7} rad. With this particular geometry, sensitivity to lateral forces could be improved by about a factor of 100 compared with commercial V-shaped Si₃N₄ or rectangular Si or Si₃N₄ cantilevers used by Meyer and Amer (1990b) with torsional spring constant of ~100 N/m. Ruan and Bhushan (1994a) and Bhushan and Ruan (1994) used 115 μm -long, wide-legged V-shaped cantilevers made of Si₃N₄ for friction measurements.

For scratching, wear, and indentation studies, single-crystal natural diamond tips ground to the shape of a three-sided pyramid with an apex angle of either 60° or 80° whose point is sharpened to a radius of about 100 nm are commonly used (Bhushan et al. 1994; Bhushan 1999a)

(Fig. 2.15c). The tips are bonded with conductive epoxy to a gold-plated 304 stainless steel spring sheet (length = 20 mm, width = 0.2 mm, thickness = 20–60 μm) which acts as a cantilever. Free length of the spring is varied to change the beam stiffness. The normal spring constant of the beam ranges from about 5 to 600 N/m for a 20 μm thick beam. The tips are produced by various manufacturers including R-DEC Co., Tsukuba, Japan. SEM micrographs of silicon nitride, single-crystal silicon, and natural diamond tips are shown in Fig. 2.16.

For imaging within trenches by AFM, high aspect ratio tips are used. Examples of the two probes are shown in Fig. 2.17. The high-aspect ratio tip (Hart) probes are produced by starting with a conventional Si₃N₄ pyramidal probe. Through a combination of focused ion beam (FIB) and high-resolution scanning electron microscopy (SEM) techniques, a thin filament is

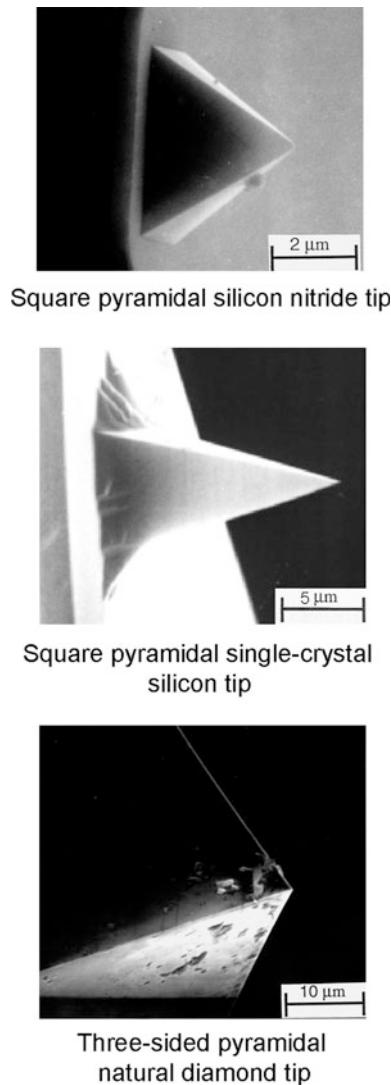


Fig. 2.16 SEM micrographs of a square-pyramidal PECVD Si_3N_4 tip, a square pyramidal etched single-crystal silicon tip, and a three-sided pyramidal natural diamond tip

grown at the apex of the pyramid. The probe filament is approximately 1 μm long and 0.1 μm in diameter. It tapers to an extremely sharp point (radius better than the resolution of most SEMs). The long thin shape and sharp radius make it ideal for imaging within “vias” of microstructures and trenches ($>0.25 \mu\text{m}$). Because of flexing of the probe, it is unsuitable for imaging structures at the atomic level since the flexing of the probe can create image artifacts. For

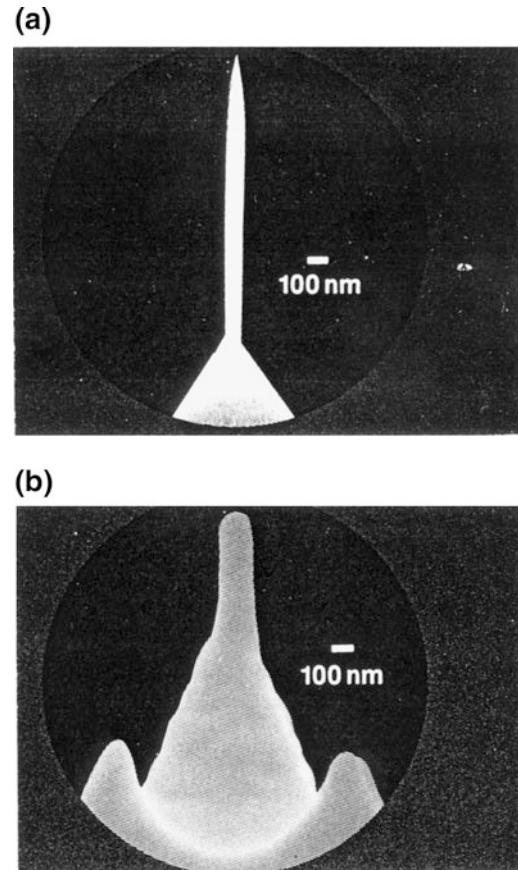


Fig. 2.17 Schematics of (a) HART Si_3N_4 probe, and (b) FIB milled Si_3N_4 probe

atomic-scale imaging, a FIB-milled probe is used which is relatively stiff yet allows for closely spaced topography. These probes start out as conventional Si_3N_4 pyramidal probes but the pyramid is FIB milled until a small cone shape is formed which has a high aspect ratio with 0.2–0.3 μm in length. The milled probes allow nanostructure resolution without sacrificing rigidity.

Carbon nanotube tips having small diameter and high aspect ratio are used for high resolution imaging of surfaces and of deep trenches, in the tapping mode or noncontact mode. Single-walled carbon nanotubes (SWNT) are microscopic graphitic cylinders that are 0.7–3 nm in diameter and up to many microns in length. Larger structures called multi-walled carbon nanotubes (MWNT) consist of nested, concentrically

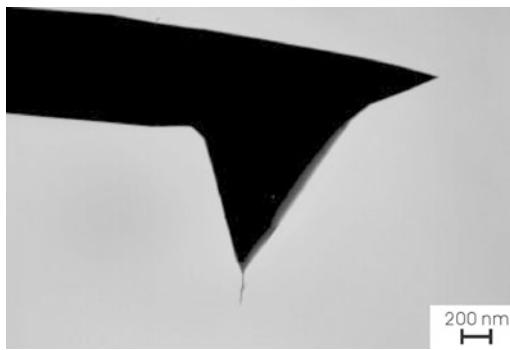


Fig. 2.18 SEM micrograph of a multi-walled carbon nanotube (MWNT) tip physically attached on the single-crystal silicon, square-pyramidal tip (Courtesy Piezomax Technologies, Inc.)

arranged SWNT and have diameters ranging from 3 to 50 nm. MWNT carbon nanotube AFM tips are produced by manual assembly (Dai et al. 1996), chemical vapor deposition (CVD) synthesis, and hybrid fabrication process (Hafner et al. 2001). Figure 2.18 shows a TEM micrograph of a carbon nanotube tip, ProbeMax™, commercially produced by mechanical assembly by Piezomax Technologies, Inc., Middleton, Wisconsin. For production of these tips, MWNT nanotubes are produced by carbon arc. They are physically attached on single-crystal silicon square-pyramidal tips in the SEM using a manipulator and the SEM stage to control the nanotubes and the tip independently. Once the nanotube is attached to the tip, it is usually too long to image with. It is shortened using an AFM applying voltage between the tip and the sample. Nanotube tips are also commercially produced by CVD synthesis by NanoDevices, Santa Barbara, California.

2.3.4 Friction Measurement Methods

Based on the work by Ruan and Bhushan (1994a), the two methods for friction measurements are now described in more detail. (Also see, Meyer and Amer 1990b; Palacio and Bhushan 2010.) A scanning angle is defined as the angle relative to the y-axis in Fig. 2.19a. This is also the long axis of the cantilever. A zero

degree scanning angle corresponds to the sample scanning in the y direction, and a 90° scanning angle corresponds to the sample scanning perpendicular to this axis in the xy plane (along the x axis). If the scanning direction is in both y and -y directions, we call this “parallel scan”. Similarly, a “perpendicular scan” means the scanning direction is in x and -x directions. The sample traveling direction for each of these two methods is illustrated in Fig. 2.19b.

In the method 1 (using “height” mode with parallel scans) in addition to topographic imaging, it is also possible to measure friction force when the scanning direction of the sample is parallel to y direction (parallel scan). If there were no friction force between the tip and the moving sample, the topographic feature would be the only factor which cause the cantilever to be deflected vertically. However, friction force does exist on all contact surfaces where one object is moving relative to another. The friction force between the sample and the tip will also cause a cantilever deflection. We assume that the normal force between the sample and the tip is W_0 when the sample is stationary (W_0 is typically in the range of 10 nN–200 nN), and the friction force between the sample and the tip is W_f as the sample scans against the tip. The direction of friction force (W_f) is reversed as the scanning direction of the sample is reversed from the positive (y) to the negative (-y) direction $\vec{W}_{f(y)} = -\vec{W}_{f(-y)}$.

When the vertical cantilever deflection is set at a constant level, it is the total force (normal force and friction force) applied to the cantilever that keeps the cantilever deflection at this level. Since the friction force is in opposite directions as the traveling direction of the sample is reversed, the normal force will have to be adjusted accordingly when the sample reverses its traveling direction, so that the total deflection of the cantilever will remain the same. We can calculate the difference of the normal force between the two traveling directions for a given friction force W_f . First, by means of a constant deflection, the total moment applied to the cantilever is constant. If we take the reference point

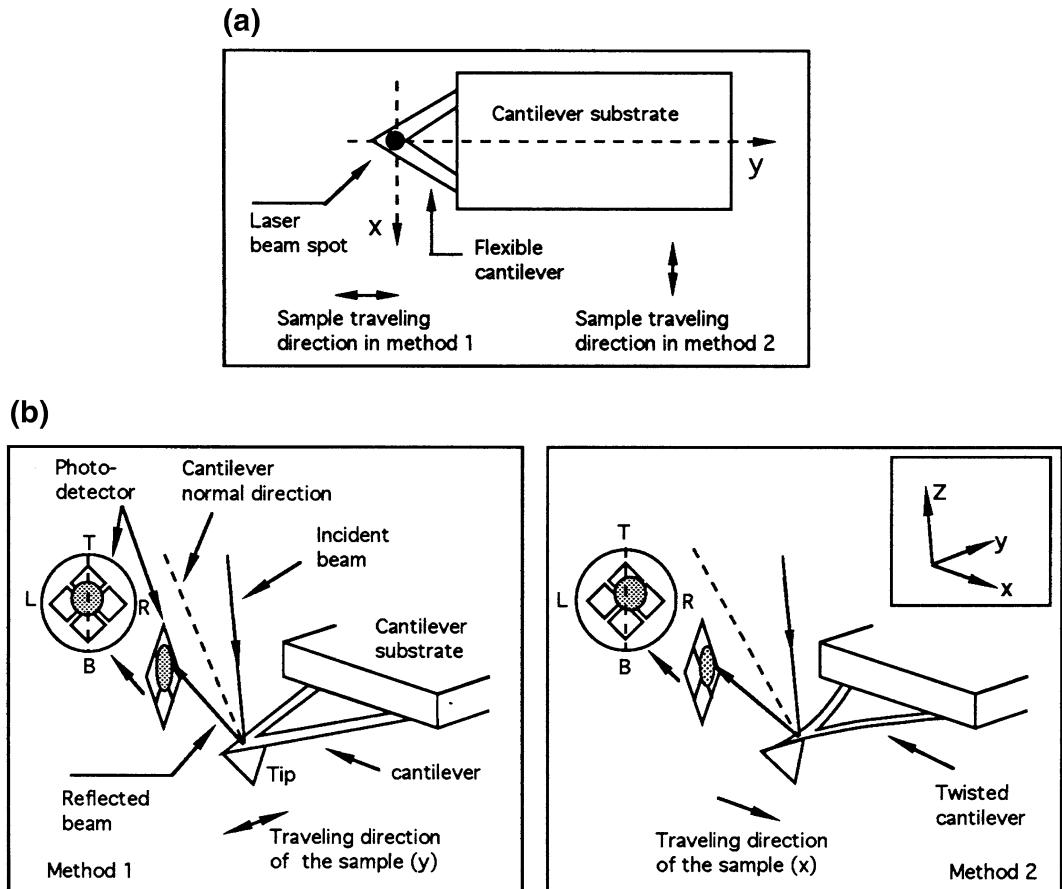


Fig. 2.19 (a) Schematic defining the x- and y-directions relative to the cantilever, and showing the sample traveling direction in two different measurement methods discussed in the text, (b) schematic of deformation of the tip and cantilever shown as a result of sliding in the x- and y-directions. A twist is introduced to the cantilever if the scanning is in the x-direction (**b, right**) (Ruan and Bhushan 1994a)

to be the point where the cantilever joins the cantilever holder (substrate), point P in Fig. 2.20, we have the following relationship:

$$(W_0 - \Delta W_1)L + W_f\ell = (W_0 + \Delta W_2)L - W_f\ell \quad (2.1)$$

or

$$(\Delta W_1 + \Delta W_2)L = 2W_f\ell \quad (2.2)$$

Thus

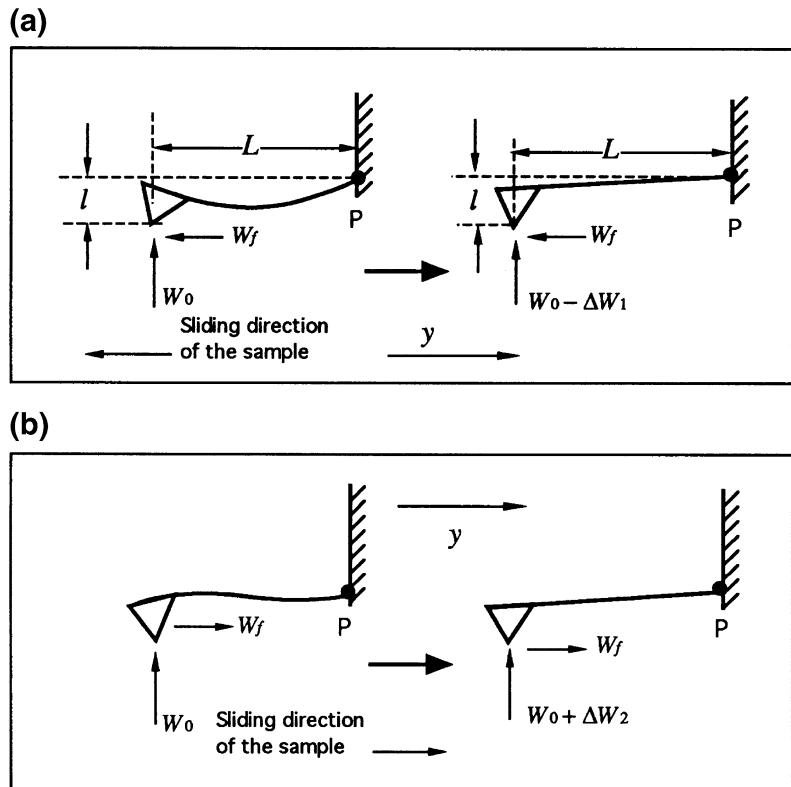
$$W_f = (\Delta W_1 + \Delta W_2)/2\ell \quad (2.3)$$

where ΔW_1 and ΔW_2 are the absolute value of the changes of normal force when the sample is traveling in $-y$ and y directions, respectively, as shown in Fig. 2.20; L is the length of the cantilever; and ℓ is the vertical distance between the end of the tip and point P. The coefficient of friction (μ) between the tip and the sample is then given as

$$\mu = \frac{W_f}{W_0} = \left[\frac{(\Delta W_1 + \Delta W_2)}{W_0} \right] \left(\frac{L}{2\ell} \right) \quad (2.4)$$

In all circumstances, there are adhesive and interatomic attractive forces between the

Fig. 2.20 **a** Schematic showing an additional bending of the cantilever—due to friction force when the sample is scanned in the y - or $-y$ -direction (left). **b** This effect will be canceled by adjusting the piezo height by a feedback circuit (right) (Ruan and Bhushan 1994a)



cantilever tip and the sample. The adhesive force can be due to water from the capillary condensation and other contaminants present at the surface which form meniscus bridges (Blackman et al. 1990; O’Shea et al. 1992; Bhushan 1999a) and the interatomic attractive force includes van

der Waals attraction (Burnham et al. 1991). If these forces (and indentation effect as well, which is usually small for rigid samples) can be neglected, the normal force W_0 is then equal to the initial cantilever deflection H_0 multiplied by the spring constant of the cantilever. $(\Delta W_1 + W_2)$ can be measured by multiplying the same spring constant by the height difference of the piezo tube between the two traveling directions (y and $-y$ directions) of the sample. This height difference is denoted as $(\Delta H_1 + \Delta H_2)$, shown schematically in Fig. 2.21. Thus, Eq. (2.4) can be rewritten as

$$\mu = \frac{W_f}{W_0} = \left[\frac{(\Delta H_1 + \Delta H_2)}{H_0} \right] \left(\frac{L}{2\ell} \right) \quad (2.5)$$

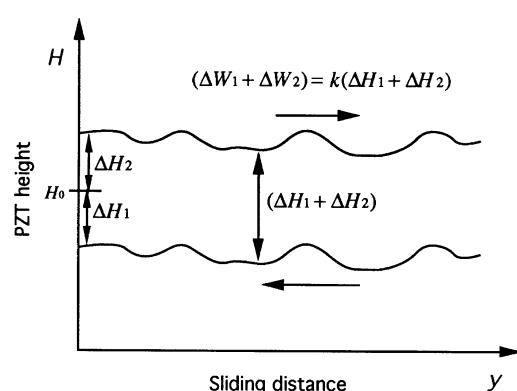


Fig. 2.21 Schematic illustration of the height difference of the piezoelectric tube scanner as the sample is scanned in y and $-y$ directions

Since the piezo tube vertical position is affected by the surface topographic profile of the sample in addition to the friction force being applied at the tip, this difference has to be taken point by point at the same location on the sample

surface as shown in Fig. 2.21. Subtraction of point by point measurements may introduce errors, particularly for rough samples. We will come back to this point later. In addition, precise measurement of L and ℓ (which should include the cantilever angle) are also required.

If the adhesive forces between the tip and the sample are large enough that it can not be neglected, one should include it in the calculation. However, there could be a large uncertainty in determining this force, thus an uncertainty in using Eq. (2.5). An alternative approach is to make the measurements at different normal loads and to use $\Delta(H_0)$ and $\Delta(\Delta H_1 + \Delta H_2)$ from the measurements in Eq. (2.5). Another comment on Eq. (2.5) is that, since only the ratio between $(\Delta H_1 + \Delta H_2)$ and H_0 comes into this equation, the piezo tube vertical position H_0 and its position difference $(\Delta H_1 + \Delta H_2)$ can be in the units of volts as long as the vertical traveling distance of the piezo tube and the voltage applied to it has a linear relationship. However, if there is a large nonlinearity between the piezo tube traveling distance and the applied voltage, this nonlinearity must be included in the calculation.

It should also be pointed out that, Eqs. (2.4) and (2.5) are derived under the assumption that the friction force W_f is the same for the two scanning directions of the sample. This is an approximation since the normal force is slightly different for the two scans and there may also be a directionality effect in friction. However, this difference is much smaller than W_0 itself. We can ignore the second order correction.

Method 2 (“aux” mode with perpendicular scan) to measure friction was suggested by Meyer and Amer (1990b). The sample is scanned perpendicular to the long axis of the cantilever beam (i.e., to scan along the x or $-x$ direction in Fig. 2.19a) and the output of the horizontal two quadrants of the photodiode-detector is measured. In this arrangement, as the sample moves under the tip, the friction force will cause the cantilever to twist. Therefore the light intensity between the left and right (L and R in Fig. 2.19b, right) detectors will be different. The differential signal between the left and right detectors is denoted as FFM signal $[(L - R)/(L + R)]$. This

signal can be related to the degree of twisting, hence to the magnitude of friction force. Again, because of a possible error in determining normal force due to the presence of an adhesive force at the tip-sample interface, the slope of the friction data (FFM signal vs. normal load) needs to be taken for an accurate value of coefficient of friction.

While friction force contributes to the FFM signal, friction force may not be the only contributing factor in commercial FFM instruments. One can notice this fact by simply engaging the cantilever tip with the sample. Before engaging, the left and right detectors can be balanced by adjusting the position of the detectors so that the intensity difference between these two detectors is zero (FFM signal is zero). Once the tip is engaged with the sample, this signal is no longer zero even if the sample is not moving in the xy plane with no friction force applied. This would be a detrimental effect. It has to be understood and eliminated from the data acquisition before any quantitative measurement of friction force becomes possible.

One of the fundamental reasons for this observation is the following. The detectors may not have been properly aligned with respect to the laser beam. To be precise, the vertical axis of the detector assembly (the line joining T-B in Fig. 2.22) is not in the plane defined by the incident laser beam and the beam reflected from an untwisted cantilever (we call this plane “beam plane”). When the cantilever vertical deflection changes due to a change of applied normal force (without having the sample scanned in the xy plane), the laser beam will be reflected up and down and form a projected trajectory on the detector. (Note that this trajectory is in the defined beam plane). If this trajectory is not coincident with the vertical axis of the detector, the laser beam will not evenly bisect the left and right quadrants of the detectors, even under the condition of no torsional motion of the cantilever, see Fig. 2.22. Thus when the laser beam is reflected up and down due a change of the normal force, the intensity difference between the left and right detectors will also change. In other words, the FFM signal will change as the normal

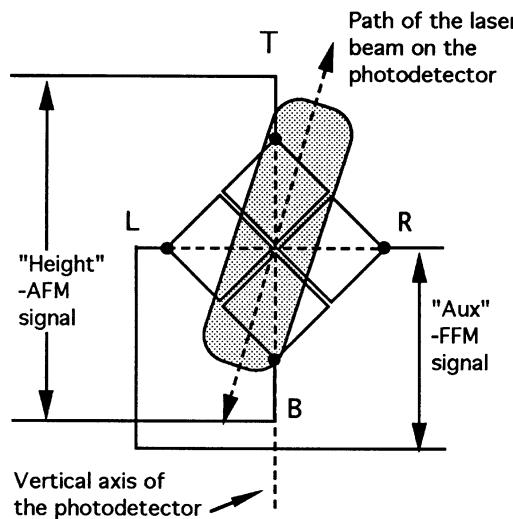


Fig. 2.22 The trajectory of the laser beam on the photodetectors in as the cantilever is vertically deflected (with no torsional motion) for a misaligned photodetector with respect to the laser beam. For a change of normal force (vertical deflection of the cantilever), the laser beam is projected at a different position on the detector. Due to a misalignment, the projected trajectory of the laser beam on the detector is not parallel with the detector vertical axis (the line joint T-B) (Ruan and Bhushan 1994a)

force applied to the tip is changed, even if the tip is not experiencing any friction force. This (FFM) signal is unrelated to friction force or to the actual twisting of the cantilever. We will call this part of FFM signal “ FFM_F ,” and the part which is truly related to friction force “ FFM_T ”.

The FFM_F signal can be eliminated. One way of doing this is as follows. First the sample is scanned in both x and $-x$ directions and the FFM signal for scans in each direction is recorded. Since friction force reverses its directions when the scanning direction is reversed from x to $-x$ direction, the FFM_T signal will have opposite signs as the scanning direction of the sample is reversed ($FFM_T(x) = -FFM_T(-x)$). Hence, the FFM_T signal will be canceled out if we take the sum of the FFM signals for the two scans. The average value of the two scans will be related to FFM_F due to the misalignment,

$$FFM(x) + FFM(-x) = 2FFM_F \quad (2.6)$$

This value can therefore be subtracted from the original FFM signals of each of these two scans to obtain the true FFM signal (FFM_T). Or, alternately, by taking the difference of the two FFM signals, one directly gets the FFM_T value

$$\begin{aligned} FFM(x) - FFM(-x) &= FFM_T(x) - FFM_T(-x) \\ &= 2FFM_T(x) \end{aligned} \quad (2.7)$$

Ruan and Bhushan (1994a) have shown that error signal (FFM_F) can be very large compared to friction signal FFM_T , thus correction is required.

Now we compare the two methods. The method of using “height” mode and parallel scan (method 1) is very simple to use. Technically, this method can provide 3D friction profiles and the corresponding topographic profiles. However, there are some problems with this method. Under most circumstances, the piezo scanner displays a hysteresis when the traveling direction of the sample is reversed. Therefore the measured surface topographic profiles will be shifted relative to each other along the y -axis for the two opposite (y and $-y$) scans. This would make it difficult to measure the local height difference of the piezo tube for the two scans. However, the average height difference between the two scans and hence the average friction can still be measured. The measurement of average friction can serve as an internal means of friction force calibration. Method 2 is a more desirable approach. The subtraction of FFM_F signal from FFM for the two scans does not introduce error to local friction force data. An ideal approach in using this method would be to add the average value of the two profiles in order to get the error component (FFM_F) and then subtract this component from either profiles to get true friction profiles in either directions. By making measurements at various loads, we can get the average value of the

coefficient of friction which then can be used to convert the friction profile to the coefficient of friction profile. Thus any directionality and local variations in friction can be easily measured. In this method, since topography data are not affected by friction, accurate topography data can be measured simultaneously with friction data and better localized relationship between the two can be established.

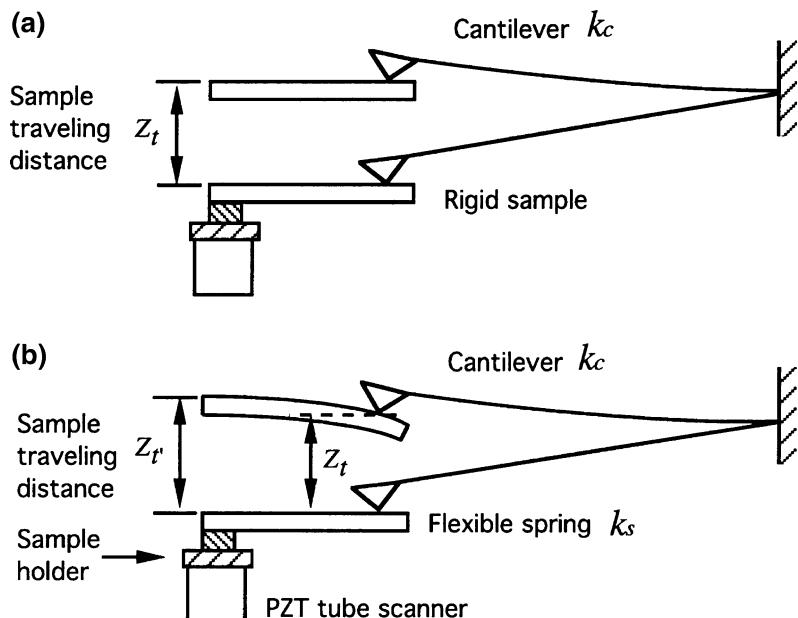
2.3.5 Normal Force and Friction Force Calibrations of Cantilever Beams

Based on Ruan and Bhushan (1994a), we now discuss normal force and friction force calibrations. In order to calculate the absolute value of normal and friction forces in Newtons using the measured AFM and FFM_T voltage signals, it is necessary to first have an accurate value of the spring constant of the cantilever (k_c). The spring constant can be calculated using the geometry and the physical properties of the cantilever material (Albrecht et al. 1990; Meyer and Amer 1990b; Sarid and Elings 1991). However, the properties of the PECVD Si₃N₄ (used in

fabricating cantilevers) could be different from those of bulk material. For example, by using an ultrasonic measurement, we found the Young's modulus of the cantilever beam to be about 238 ± 18 GPa which is less than that of bulk Si₃N₄ (310 GPa). Furthermore the thickness of the beam is nonuniform and difficult to measure precisely. Since the stiffness of a beam goes as the cube of thickness, minor errors in precise measurements of thickness can introduce substantial stiffness errors. Thus one should experimentally measure the spring constant of the cantilever. Cleveland et al. (1993) measured the normal spring constant by measuring resonant frequencies of the beams.

For normal spring constant measurement, Ruan and Bhushan (1994a) used a stainless steel spring sheet of known stiffness (width = 1.35 mm, thickness = 15 μm , free hanging length = 5.2 mm). One end of the spring was attached to the sample holder and the other end was made to contact with the cantilever tip during the measurement, see Fig. 2.23. They measured the piezo traveling distance for a given cantilever deflection. For a rigid sample (such as diamond), the piezo traveling distance Z_t (measured from the point where the tip touches the sample)

Fig. 2.23 Illustration showing the deflection of cantilever as it is pushed by (a) a rigid sample or by (b) a flexible spring sheet (Ruan and Bhushan 1994a)



should equal the cantilever deflection. To keep the cantilever deflection at the same level using a flexible spring sheet, the new piezo traveling distance $Z_{t'}$ would be different from Z_t . The difference between $Z_{t'}$ and Z_t corresponds to the deflection of the spring sheet. If the spring constant of the spring sheet is k_s , the spring constant of the cantilever k_c can be calculated by

$$(Z_{t'} - Z_t)k_s = Z_t k_c$$

or

$$k_c = \frac{k_s(Z_{t'} - Z_t)}{Z_t} \quad (2.8)$$

The spring constant of the spring sheet (k_s) used in this study is calculated to be 1.54 N/m. For a wide-legged cantilever used in our study (length = 115 μm, base width = 122 μm, leg width = 21 μm, and thickness = 0.6 μm), k_c was measured to be 0.40 N/m instead of 0.58 N/m reported by its manufacturer—Digital Instruments Inc. To relate photodiode detector output to the cantilever deflection in nm, they used the same rigid sample to push against the AFM tip. Since for a rigid sample the cantilever vertical deflection equals the sample traveling distance measured from the point where the tip touches the sample, the photodiode output as the tip is pushed by the sample can be converted directly to cantilever deflection. For these measurements, they found the conversion factor to be 20 nm/V.

The normal force applied to the tip can be calculated by multiplying the cantilever vertical deflection by the cantilever spring constant for samples which have very small adhesive force with the tip. If the adhesive force between the sample and the tip is large, it should be included in the normal force calculation. This is particularly important in atomic-scale force measurement because in this region, the typical normal force that is measured is in the range of a few hundreds of nN to a few mN. The adhesive force could be comparable to the applied force.

The conversion of friction signal (from FFM_T to friction force) is not as straightforward. For example, one can calculate the degree of twisting

for a given friction force using the geometry and the physical properties of the cantilever (Meyer and Amer 1988; O’Shea et al. 1992). One would need the information on the detectors such as the quantum efficiency of the detector, the laser power, the instrument’s gain, etc. in order to be able convert the signal into the degree of twisting. Generally speaking, this procedure cannot be accomplished without having some detailed information about the instrument. This information is not usually provided by the manufacturers. Even if this information is readily available, error may still occur in using this approach because there will always be variations as a result of the instrumental set up. For example, it has been noticed that the measured FFM_T signal could be different for the same sample when different AFM microscopes of the same kind are used. The essence is that, one cannot calibrate the instrument experimentally using this calculation. O’Shea et al. (1992) did perform a calibration procedure in which the torsional signal was measured as the sample is displaced a known distance laterally while ensuring that the tip does not slide over the surface. However, it is difficult to verify if the tip sliding does not occur.

Apparently, a new method of calibration is required. There is a more direct and simpler way of doing this. The first method described (method 1) to measure friction can directly provide an absolute value of coefficient of friction. It can therefore be used just as an internal means of calibration for the data obtained using method 2. Or for a polished sample which introduces least error in friction measurement using method 1, method 1 can be used to obtain calibration for friction force for method 2. Then this calibration can be used for measurement on all samples using method 2. In method 1, the length of the cantilever required can be measured using an optical microscope; the length of the tip can be measured using a scanning electron microscope. The relative angle between the cantilever and the horizontal sample surface can be measured directly. Thus the coefficient of friction can be measured with few unknown parameters. The friction force can then be calculated by

multiplying the coefficient of friction by the normal load. The FFM_T signal obtained using method 2 can then be converted into friction force. For their instrument, they found the conversion to be 8.6 nN/V.

2.4 AFM Instrumentation and Analyses

The performance of AFMs and the quality of AFM images greatly depend on the instrument available and the probes (cantilever and tips) in use. This section describes the mechanics of cantilevers, instrumentation and analysis of force detection systems for cantilever deflections, and scanning and control systems.

2.4.1 The Mechanics of Cantilevers

2.4.1.1 Stiffness and Resonances of Lumped Mass Systems

Any one of the building blocks of an AFM, be it the body of the microscope itself or the force measuring cantilevers, are mechanical resonators. These resonances can be excited either by the surrounding or by the rapid movement of the tip or the sample. To avoid problems due to building or air induced oscillations, it is of paramount importance to optimize the design of the AFMs for high resonant frequencies. This usually means to decrease the size of the microscope (Pohl 1986). By using cube-like or sphere-like structures for the microscope, one can considerably increase the lowest eigen-frequency. The fundamental natural frequency, ω_0 , of any spring is given by

$$\omega_0 = \frac{1}{2\pi} \sqrt{\frac{k}{m_{eff}}} \quad (2.9)$$

where k is the spring constant (stiffness) in the normal direction and m_{eff} is the effective mass. The spring constant k of a cantilever beam with uniform cross section (Fig. 2.24) is given by (Thomson and Dahleh 1998)

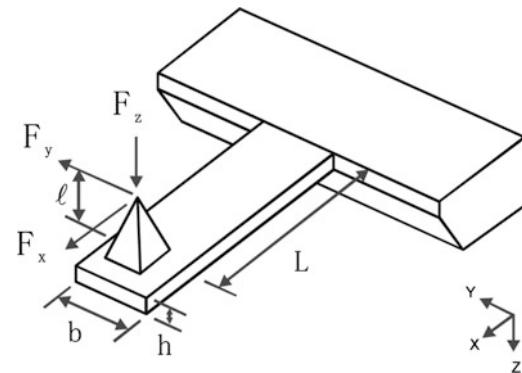


Fig. 2.24 A typical AFM cantilever with length L , width b , and height h . The height of the tip is ℓ . The material is characterized by Young's modulus E , the shear modulus G and a mass density ρ . Normal (F_z), axial (F_x), and lateral (F_y) forces exist at the end of the tip

$$k = \frac{3EI}{L^3} \quad (2.10)$$

where E is the Young's modulus of the material, L is the length of the beam and I is the moment of inertia of the cross section. For a rectangular cross section with a width b (perpendicular to the deflection) and a height h one obtains an expression for I

$$I = \frac{bh^3}{12} \quad (2.11)$$

Combining Eqs. (2.9), (2.10) and (2.11) we get an expression for ω_0

$$\omega_0 = \sqrt{\frac{Ebh^3}{4L^3 m_{eff}}} \quad (2.12)$$

The effective mass can be calculated using Raleigh's method. The general formula using Raleigh's method for the kinetic energy T of a bar is

$$T = \frac{1}{2} \int_0^L m \left(\frac{\partial z(x)}{\partial t} \right)^2 dx \quad (2.13)$$

For the case of a uniform beam with a constant cross section and length L one obtains for the deflection $z(x) = z_{max}[1 - (3x/2L) +$

$(x^3/2L^3)$. Inserting z_{\max} into Eq. (2.13) and solving the integral gives

$$\begin{aligned} T &= \frac{1}{2} \int_0^L \frac{m}{L} \left[\frac{\partial z_{\max}(x)}{\partial t} \left(1 - \frac{3x}{2L} \right) + \left(\frac{x^3}{L^3} \right) \right]^2 dx \\ &= \frac{1}{2} m_{\text{eff}} (z_{\max} t)^2 \end{aligned}$$

which gives

$$m_{\text{eff}} = \frac{9}{20} m \quad (2.14)$$

Substituting the Eq. (2.14) into Eq. (2.12) and noting that $m = \rho Lbh$, where ρ is the mass density, one obtains the following expression

$$\omega_0 = \left(\frac{\sqrt{5}}{3} \sqrt{\frac{E}{\rho}} \right) \frac{h}{L^2} \quad (2.15)$$

It is evident from Eq. (2.15), that one way to increase the natural frequency is to choose a material with a high ratio E/ρ ; see Table 2.2 for typical values of $\sqrt{E/\rho}$ of various commonly used materials. Another way to increase the lowest eigen-frequency is also evident in Eq. (2.15). By optimizing the ratio h/L^2 , one can increase the resonant frequency. However it does not help to make the length of the structure smaller than the width or height. Their roles will just be interchanged. Hence the optimum structure is a cube. This leads to the design rule, that long, thin structures like sheet metal should be avoided. For a given resonant frequency, the quality factor Q should be as low as possible. This means that an inelastic medium such as rubber should be in contact with the structure to convert kinetic energy into heat.

2.4.1.2 Stiffness and Resonances of Cantilevers

Cantilevers are mechanical devices specially shaped to measure tiny forces. The analysis given in the previous section is applicable. However, to

understand better the intricacies of force detection systems we will discuss the example of a cantilever beam with uniform cross section, Fig. 2.24. The bending of a beam due to a normal load on the beam is governed by the Euler equation (Thomson and Dahleh 1998)

$$M = EI(x) \frac{d^2 z}{dx^2} \quad (2.16)$$

where M is the bending moment acting on the beam cross section. $I(x)$ the moment of inertia of the cross section with respect to the neutral axis defined by

$$I(x) = \int_z \int_y z^2 dy dz \quad (2.17)$$

For a normal force F_z acting at the tip,

$$M(x) = (L - x)F_z \quad (2.18)$$

since the moment must vanish at the endpoint of the cantilever. Integrating Eq. (2.16) for a normal force F_z acting at the tip and observing that EI is a constant for beams with a uniform cross section, one gets

$$z(x) = \frac{L^3}{6EI} \left(\frac{x}{L} \right)^2 \left(3 - \frac{x}{L} \right) F_z \quad (2.19)$$

The slope of the beam is

$$z'(x) = \frac{Lx}{2EI} \left(2 - \frac{x}{L} \right) F_z \quad (2.20)$$

From Eqs. (2.19) and (2.20), at the end of the cantilever, i.e. for $x = L$, for a rectangular beam, and by using an expression for I in Eq. (2.11), one gets,

$$z(L) = \frac{4}{Eb} \left(\frac{L}{h} \right)^3 F_z \quad (2.21)$$

$$z'(L) = \frac{3}{2} \left(\frac{z}{L} \right) \quad (2.22)$$

Now the stiffness in the normal (z) direction, k_z , is

$$k_z = \frac{F_z}{z(L)} = \frac{Eb}{4} \left(\frac{h}{L} \right)^3 \quad (2.23)$$

and a change in angular orientation of the end of cantilever beam is

$$\Delta\alpha = \frac{3}{2} \frac{z}{L} = \frac{6}{Ebh} \left(\frac{L}{h} \right)^2 F_z \quad (2.24)$$

Now we ask what will, to first order, happen if we apply a lateral force F_y to the end of the tip (Fig. 2.24). The cantilever will bend sideways and it will twist. The stiffness in the lateral (y) direction, k_y , can be calculated with Eq. (2.23) by exchanging b and h

$$k_y = \frac{Eh}{4} \left(\frac{b}{L} \right)^3 \quad (2.25)$$

Therefore the bending stiffness in lateral direction is larger than the stiffness for bending in the normal direction by $(b/h)^2$. The twisting or torsion on the other side is more complicated to handle. For a wide, thin cantilever ($b \gg h$) we obtain torsional stiffness along y axis, k_{yT}

$$k_{yT} = \frac{Gb h^3}{3L\ell^2} \quad (2.26)$$

where G is the modulus of rigidity [=E/2(1 + v), where v is the Poisson's ratio]. The ratio of the torsional stiffness to the lateral bending stiffness is

$$\frac{k_{yT}}{k_y} = \frac{1}{2} \left(\frac{\ell b}{h L} \right)^2 \quad (2.27)$$

where we assume $v = 0.333$. We see that thin, wide cantilevers with long tips favor torsion while cantilevers with square cross sections and short tips favor bending. Finally we calculate the ratio between the torsional stiffness and the normal bending stiffness,

$$\frac{k_{yT}}{k_z} = 2 \left(\frac{L}{\ell} \right)^2 \quad (2.28)$$

Equations (2.26) to (2.28) hold in the case where the cantilever tip is exactly in the middle axis of the cantilever. Triangular cantilevers and cantilevers with tips not on the middle axis can be dealt with by finite element methods.

The third possible deflection mode is the one from the force on the end of the tip along the cantilever axis, F_x (Fig. 2.24). The bending moment at the free end of the cantilever is equal to the $F_x\ell$. This leads to the following modification of Eq. (2.18) for the case of forces F_z and F_x

$$M(x) = (L - x)F_z + F_x\ell \quad (2.29)$$

Integration of Eq. (2.16) now leads to

$$z(x) = \frac{1}{2EI} \left[Lx^2 \left(1 - \frac{x}{3L} \right) F_z + \ell x^2 F_x \right] \quad (2.30)$$

and

$$z'(x) = \frac{1}{EI} \left[\frac{Lx}{2} \left(2 - \frac{x}{L} \right) F_z + \ell x F_x \right] \quad (2.31)$$

Evaluating Eqs. (2.30) and (2.31) at the end of the cantilever, we get the deflection and the tilt

$$z(L) = \frac{L^2}{EI} \left(\frac{L}{3} F_z - \frac{\ell}{2} F_x \right) \quad (2.32)$$

$$z'(L) = \frac{L}{EI} \left(\frac{L}{2} F_z + \ell F_x \right)$$

From these equations, one gets

$$F_z = \frac{12EI}{L^3} \left(z(L) - \frac{Lz'(L)}{2} \right) \quad (2.33)$$

$$F_x = \frac{2EI}{\ell L^2} (2Lz'(L) - 3z(L))$$

A second class of interesting properties of cantilevers is their resonance behavior. For cantilever beams one can calculate the resonant frequencies (Colchero 1993; Thomson and Dahleh 1998)

$$\omega_n^{free} = \frac{\lambda_n^2}{2\sqrt{3}L^2} \sqrt{\frac{E}{\rho}} \quad (2.34)$$

with $\lambda_0 = (0.596864...) \pi$, $\lambda_1 = (1.494175...) \pi$, $\lambda_n \rightarrow (n + 1/2)\pi$. The subscript n represents the order of the frequency, e.g., fundamental, second mode, and the nth mode.

A similar equation as Eq. (2.34) holds for cantilevers in rigid contact with the surface. Since there is an additional restriction on the movement of the cantilever, namely the location of its end point, the resonant frequency increases. Only the λ_n 's terms change to (Colchero 1993)

$$\begin{aligned} \lambda'_0 &= (1.2498763 ...) \pi, \lambda'_1 \\ &= (2.2499997 ...) \pi, \lambda'_n \rightarrow (n + 1/4)\pi \end{aligned} \quad (2.35)$$

The ratio of the fundamental resonant frequency in contact to the fundamental resonant frequency not in contact is 4.3851.

For the torsional mode we can calculate the resonant frequencies as

$$\omega_0^{tors} = 2\pi \frac{h}{Lb} \sqrt{\frac{G}{\rho}} \quad (2.36)$$

For cantilevers in rigid contact with the surface, we obtain the expression for the fundamental resonant frequency (Colchero 1993)

$$\omega_0^{tors,contact} = \frac{\omega_0^{tors}}{\sqrt{1 + 3(2L/b)^2}} \quad (2.37)$$

The amplitude of the thermally induced vibration can be calculated from the resonant frequency using

$$\Delta z_{therm} = \sqrt{\frac{k_b T}{k}} \quad (2.38)$$

where k_b is Boltzmann's constant and T is the absolute temperature. Since AFM cantilevers are resonant structures, sometimes with rather high Q , the thermal noise is not evenly distributed as

Eq. (2.38) suggests. The spectral noise density below the peak of the response curve is (Colchero 1993)

$$z_0 = \sqrt{\frac{4k_b T}{k\omega_0 Q}} \quad \left\{ \text{in m}/\sqrt{\text{Hz}} \right\} \quad (2.39)$$

where Q is the quality factor of the cantilever, described earlier.

2.4.2 Instrumentation and Analyses of Detection Systems for Cantilever Deflections

A summary of selected detection systems was provided in Fig. 2.8. Here we discuss in detail pros and cons of various systems.

2.4.2.1 Optical Interferometer Detection Systems

Interferometers

Soon after the first papers on the AFM (Binnig et al. 1986), which used a tunneling sensor, an instrument based on an interferometer was published (McClelland et al. 1987). The sensitivity of the interferometer depends on the wavelength of the light employed in the apparatus. Figure 2.25 shows the principle of such an interferometric design. The light incident from the left is focused by a lens on the cantilever. The reflected light is collimated by the same lens and interferes with the light reflected at the flat. To separate the reflected light from the incident light a $\lambda/4$ plate converts the linear polarized incident light to circular polarization. The reflected light is made linear polarized again by the $\lambda/4$ -plate, but with a polarization orthogonal to that of the incident light. The polarizing beam splitter then deflects the reflected light to the photo diode.

Homodyne Interferometer

To improve the signal to noise ratio of the interferometer the cantilever is driven by a piezo near its resonant frequency. The amplitude Δz of the cantilever as a function of driving frequency Ω is

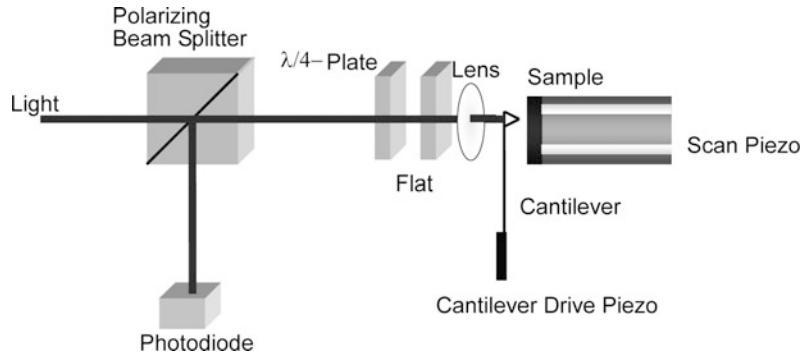


Fig. 2.25 Principle of an interferometric AFM. The light of the laser light source is polarized by the polarizing beam splitter and focused on the back of the cantilever. The light passes twice through a quarter wave plate and is hence orthogonally polarized to the incident light. The second arm of the interferometer is formed by the flat. The interference pattern is modulated by the oscillating cantilever

$$\Delta z(\Omega) = \Delta z_0 \frac{\Omega_0^2}{\sqrt{(\Omega^2 - \Omega_0^2)^2 + \frac{\Omega^2 \Omega_0^2}{Q^2}}} \quad (2.40)$$

where Δz_0 is the constant drive amplitude and Ω_0 the resonant frequency of the cantilever. The resonant frequency of the cantilever is given by the effective potential

$$\Omega_0 = \sqrt{\left(k + \frac{\partial^2 U}{\partial z^2}\right) \frac{1}{m_{eff}}} \quad (2.41)$$

where U is the interaction potential between the tip and the sample. Equation (2.41) shows that an attractive potential decreases Ω_0 . The change in Ω_0 in turn results in a change of the Δz (see Eq. (2.40)). The movement of the cantilever changes the path difference in the interferometer. The light reflected from the cantilever with the amplitude $A_{\ell,0}$ and the reference light with the amplitude $A_{r,0}$ interfere on the detector. The detected intensity $I(t) = \{A_{\ell}(t) + A_r(t)\}^2$ consists of two constant terms and a fluctuating term

$$2A_{\ell}(t)A_r(t) = A_{\ell,0}A_{r,0} \sin\left[\omega t + \frac{4\pi\delta}{\lambda}\right] + \frac{4\pi\Delta z}{\lambda} \sin(\Omega t) \quad (2.42)$$

Here ω is the frequency of the light, λ is the wavelength of the light, δ is the path difference in the interferometer, and Δz is the instantaneous amplitude of the cantilever, given according to Eqs. (2.40) and (2.41) as a function of Ω , k , and U . The time average of Eq. (2.42) then becomes

$$\begin{aligned} \langle 2A_{\ell}(t)A_r(t) \rangle_T &\propto \cos\left[\frac{4\pi\delta}{\lambda} + \frac{4\pi\Delta z}{\lambda} \sin(\Omega t)\right] \\ &\approx \cos\left(\frac{4\pi\delta}{\lambda}\right) - \sin\left[\frac{4\pi\Delta z}{\lambda} \sin(\Omega t)\right] \\ &\approx \cos\left(\frac{4\pi\delta}{\lambda}\right) - \frac{4\pi\Delta z}{\lambda} \sin(\Omega t) \end{aligned} \quad (2.43)$$

Here all small quantities have been omitted and functions with small arguments have been linearized. The amplitude of Δz can be recovered with a lock-in technique. However, Eq. (2.43) shows that the measured amplitude is also a function of the path difference δ in the interferometer. Hence this path difference δ must be very stable. The best sensitivity is obtained when $\sin(4\delta/\lambda) \approx 0$.

Heterodyne Interferometer

This influence is not present in the heterodyne detection scheme shown in Fig. 2.26. Light incident from the left with a frequency ω is split in a reference path (upper path in Fig. 2.26) and a

measurement path. Light in the measurement path is shifted in frequency to $\omega_1 = \omega + \Delta\omega$ and focused on the cantilever. The cantilever oscillates at the frequency Ω , as in the homodyne detection scheme. The reflected light $A_\ell(t)$ is collimated by the same lens and interferes on the photo diode with the reference light $A_r(t)$. The fluctuating term of the intensity is given by

$$2A_\ell(t)A_r(t) = A_{\ell,0}A_{r,0} \sin \left[(\omega + \Delta\omega)t + \frac{4\pi\delta}{\lambda} \right] \sin(\Omega t) \quad (2.44)$$

where the variables are defined as in Eq. (2.42). Setting the path difference $\sin(4\pi\delta/\lambda) \approx 0$ and taking the time average, omitting small quantities and linearizing functions with small arguments we get

$$\begin{aligned} \langle 2A_\ell(t)A_r(t) \rangle_T &\propto \cos \left[\Delta\omega t + \frac{4\pi\delta}{\lambda} + \frac{4\pi\Delta z}{\lambda} \sin(\Omega t) \right] \\ &= \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \cos \left[\frac{4\pi\Delta z}{\lambda} \sin(\Omega t) \right] \\ &\quad - \sin \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \sin \left[\frac{4\pi\Delta z}{\lambda} \sin(\Omega t) \right] \\ &\approx \cos \left(\frac{4\pi\delta}{\lambda} \right) - \sin \left[\frac{4\pi\Delta z}{\lambda} \sin(\Omega t) \right] \\ &\approx \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \left[1 - \frac{8\pi^2\Delta z^2}{\lambda^2} \sin(\Omega t) \right] \\ &\quad - \frac{4\pi\Delta z}{\lambda} \sin \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \sin(\Omega t) \\ &= \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) - \frac{8\pi^2\Delta z^2}{\lambda^2} \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \sin(\Omega t) \\ &\quad - \frac{4\pi\Delta z}{\lambda} \sin \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \sin(\Omega t) \\ &= \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) - \frac{4\pi^2\Delta z^2}{\lambda^2} \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \\ &\quad + \frac{4\pi^2\Delta z^2}{\lambda^2} \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \cos(2\Omega t) \\ &\quad - \frac{4\pi\Delta z}{\lambda} \sin \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \sin(\Omega t) \\ &= \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \left(1 - \frac{4\pi^2\Delta z^2}{\lambda^2} \right) \\ &\quad + \frac{2\pi^2\Delta z^2}{\lambda^2} \left\{ \cos \left[(\Delta\omega + 2\Omega)t + \frac{4\pi\delta}{\lambda} \right] \right. \\ &\quad \left. + \cos \left[(\Delta\omega - 2\Omega)t + \frac{4\pi\delta}{\lambda} \right] \right\} \\ &\quad + \frac{2\pi\Delta z}{\lambda} \left\{ \cos \left[(\Delta\omega + \Omega)t + \frac{4\pi\delta}{\lambda} \right] + \cos \left[(\Delta\omega - \Omega)t + \frac{4\pi\delta}{\lambda} \right] \right\} \end{aligned} \quad (2.45)$$

Multiplying electronically the components oscillating at $\Delta\omega$ and $\Delta\omega + \Omega$ and rejecting any product except the one oscillating at Ω we obtain

$$\begin{aligned} A &= \frac{2\Delta z}{\lambda} \left(1 - \frac{4\pi^2\Delta z^2}{\lambda^2} \right) \cos \left[(\Delta\omega + 2\Omega)t + \frac{4\pi\delta}{\lambda} \right] \cos \left(\Delta\omega t + \frac{4\pi\delta}{\lambda} \right) \\ &= \frac{\Delta z}{\lambda} \left(1 - \frac{4\pi^2\Delta z^2}{\lambda^2} \right) \left\{ \cos \left[(2\Delta\omega + \Omega)t + \frac{8\pi\delta}{\lambda} \right] + \cos(\Omega t) \right\} \\ &\approx \frac{\pi\Delta z}{\lambda} \cos(\Omega t) \end{aligned} \quad (2.46)$$

Unlike in the homodyne detection scheme the recovered signal is independent from the path difference δ of the interferometer. Furthermore a lock-in amplifier with the reference set $\sin(\Delta\omega t)$ can measure the path difference δ independent of the cantilever oscillation. If necessary, a feedback circuit can keep $\delta = 0$.

Fiber-optical Interferometer

The fiber-optical interferometer (Rugar et al. 1989) is one of the simplest interferometers to build and use. Its principle is sketched in Fig. 2.27. The light of a laser is fed into an optical fiber. Laser diodes with integrated fiber pigtailed are convenient light sources. The light is split in a fiber-optic beam splitter into two fibers. One fiber is terminated by index matching oil to avoid any reflections back into the fiber. The end of the other fiber is brought close to the cantilever in the AFM. The emerging light is partially reflected back into the fiber by the cantilever. Most of the light, however, is lost. This is not a big problem since only 4% of the light is reflected at the end of the fiber, at the glass-air interface. The two reflected light waves interfere with each other. The product is guided back into the fiber coupler and again split into two parts. One half is analyzed by the photodiode. The other half is fed back into the laser. Communications grade laser diodes are sufficiently resistant against feedback to be operated in this environment. They have, however, a bad coherence length, which in this case does not matter, since the optical path difference is in any case no larger than 5 μm . Again the end of the fiber has to be positioned on a piezo drive to set the distance between the fiber and the cantilever to $\lambda(n + 1/4)$.

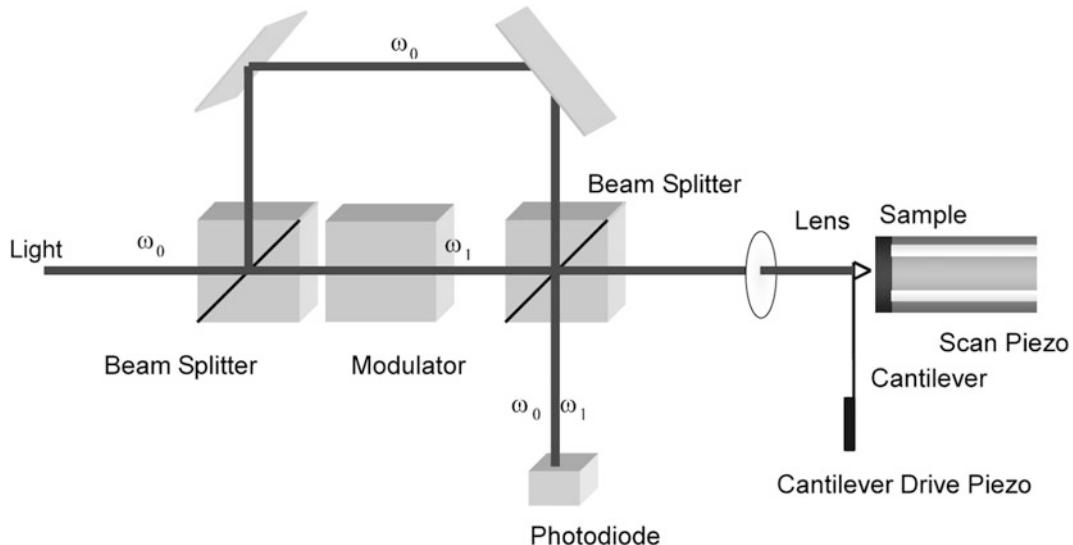


Fig. 2.26 Principle of a heterodyne interferometric AFM. Light with the frequency ω_0 is split into a reference path (upper path) and a measurement path. The light in the measurement path is frequency shifted to ω_1 by an acousto-optical modulator (or an electro-optical modulator). The light reflected from the oscillating cantilever interferes with the reference beam on the detector

Nomarski-Interferometer

Another solution to minimize the optical path difference is to use the Nomarski interferometer (Schoenenberger and Alvarado 1989). Figure 2.28 shows a schematic of the microscope. The light of a laser is focused on the cantilever by lens. A birefringent crystal (for instance calcite) between the cantilever and the lens with its optical axis 45° off the polarization

direction of the light splits the light beam into two paths, offset by a distance given by the length of the crystal. Birefringent crystals have varying indexes of refraction. In calcite, one crystal axis has a lower index than the other two. This means, that certain light rays will propagate at a different speed through the crystal than the others. By choosing a correct polarization, one can select the ordinary ray, the extraordinary ray

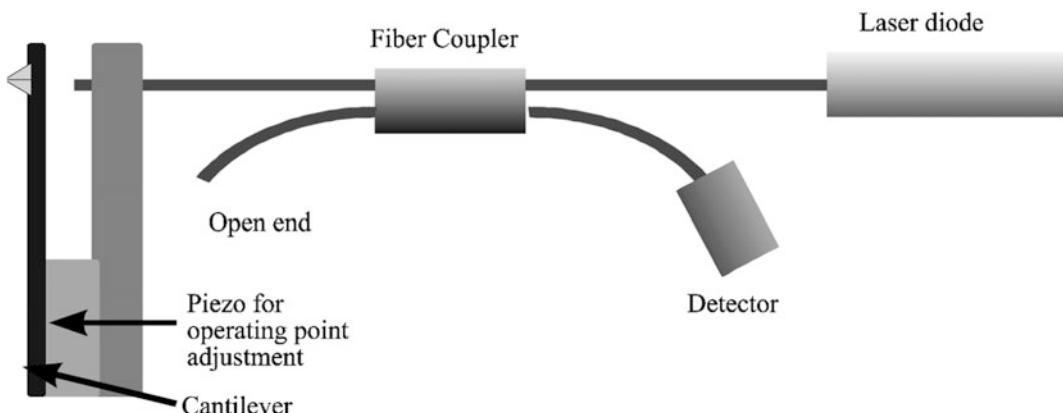


Fig. 2.27 A typical setup for a fiber optic interferometer readout

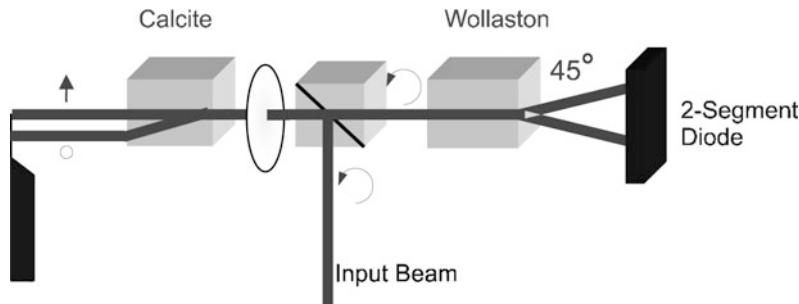


Fig. 2.28 Principle of Nomarski AFM. The circular polarized input beam is deflected to the *left* by a non-polarizing beam splitter. The light is focused onto a cantilever. The calcite crystal between the lens and the cantilever splits the circular polarized light into two spatially separated beams with orthogonal polarizations. The two light beams reflected from the lever are superimposed by the calcite crystal and collected by the lens. The resulting beam is again circular polarized. A Wollaston prism produces two interfering beams with $\pi/2$ phase shift between them. The minimal path difference accounts for the excellent stability of this microscope

or one can get any distribution of the intensity amongst those two rays. A detailed description of birefringence can be found in textbooks (e.g., Shen 1984). A calcite crystal deflects the extraordinary ray at an angle of 6° within the crystal. By choosing a suitable length of the calcite crystal, any separation can be set.

The focus of one light ray is positioned near the free end of the cantilever while the other is placed close to the clamped end. Both arms of the interferometer pass through the same space, except for the distance between the calcite crystal and the lever. The closer the calcite crystal is placed to the lever, the less influence disturbances like air currents have.

Sensitivity

Sarid (1991) has given values for the sensitivity of the different interferometric detection systems. Table 2.5 presents a summary of his results.

2.4.2.2 Optical Lever

The most common cantilever deflection detection system is the optical lever (Meyer and Amer 1988; Alexander et al. 1989). This method, depicted in Fig. 2.29, employs the same technique as light beam deflection galvanometers. A fairly well collimated light beam is reflected off a mirror and projected to a receiving target. Any change in the angular position of the mirror will change the position, where the light ray hits

Table 2.5 Noise in Interferometers

	Homodyne interferometer, fiber optic interferometer	Heterodyne interferometer	Nomarski interferometer
Laser noise $\langle \delta i^2 \rangle_L$	$\frac{1}{4} \eta^2 F^2 P_i^2 RIN$	$\eta^2 (P_R^2 + P_S^2) RIN$	$\frac{1}{16} \eta^2 P^2 \delta\theta$
Thermal noise $\langle \delta i^2 \rangle_T$	$\frac{16\pi^2}{\lambda^2} \eta^2 F^2 P_i^2 \frac{4k_B T B Q}{\omega_0 k}$	$\frac{4\pi^2}{\lambda^2} \eta^2 P_d^2 \frac{4k_B T B Q}{\omega_0 k}$	$\frac{\pi^2}{\lambda^2} \eta^2 P^2 \frac{4k_B T B Q}{\omega_0 k}$
Shot Noise $\langle \delta i^2 \rangle_S$	$4e\eta P_d B$	$2e\eta(P_R + P_S)B$	$\frac{1}{2} e\eta P B$

F is the finesse of the cavity in the homodyne interferometer, P_i the incident power, P_d is the power on the detector, η is the sensitivity of the photodetector and RIN is the relative intensity noise of the laser. P_R and P_S are the power in the reference and sample beam in the heterodyne interferometer. P is the power in the Nomarski interferometer, $\delta\theta$ is the phase difference between the reference and the probe beam in the Nomarski interferometer. B is the bandwidth, e is the electron charge, λ is the wavelength of the laser, k the cantilever stiffness, ω_0 is the resonant frequency of the cantilever, Q is the quality factor of the cantilever, T is the temperature, and δi is the variation of current i

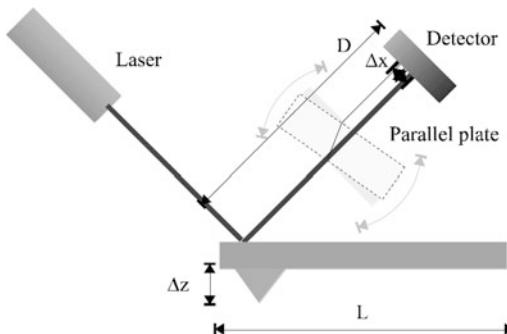


Fig. 2.29 The setup of optical lever detection microscope

the target. Galvanometers use optical path lengths of several meters and scales projected to the target wall as a read-out help.

For the AFM using the optical lever method a photodiode segmented into two (or four) closely spaced devices detects the orientation of the end of the cantilever. Initially, the light ray is set to hit the photodiodes in the middle of the two sub-diodes. Any deflection of the cantilever will cause an imbalance of the number of photons reaching the two halves. Hence the electrical currents in the photodiodes will be unbalanced too. The difference signal is further amplified and is the input signal to the feedback loop. Unlike the interferometric AFMs, where often a modulation technique is necessary to get a sufficient signal to noise ratio, most AFMs employing the optical lever method are operated in a static mode. AFMs based on the optical lever method are universally used. It is the simplest method to construct an optical readout and it can be confined in volumes smaller than 5 cm on the side.

The optical lever detection system is a simple yet elegant way to detect normal and lateral force signals simultaneously (Meyer and Amer 1988, 1990b; Alexander et al. 1989; Marti et al. 1990). It has the additional advantage that it is the fact that it is a remote detection system.

Implementations

Light from a laser diode or from a super luminescent diode is focused on the end of the cantilever. The reflected light is directed onto a quadrant diode that measures the direction of the

light beam. A Gaussian light beam far from its waist is characterized by an opening angle β . The deflection of the light beam by the cantilever surface tilted by an angle α is 2α . The intensity on the detector then shifts to the side by the product of 2α and the separation between the detector and the cantilever. The readout electronics calculates the difference of the photocurrents. The photocurrents in turn are proportional to the intensity incident on the diode.

The output signal is hence proportional to the change in intensity on the segments

$$I_{sig} \propto 4 \frac{\alpha}{\beta} I_{tot} \quad (2.47)$$

For the sake of simplicity, we assume that the light beam is of uniform intensity with its cross section increasing proportional to the distance between the cantilever and the quadrant detector. The movement of the center of the light beam is then given by

$$\Delta x_{Det} = \Delta z \frac{D}{L} \quad (2.48)$$

The photocurrent generated in a photodiode is proportional to the number of incoming photons hitting it. If the light beam contains a total number of N_0 photons then the change in difference current becomes

$$\Delta(I_R - I_L) = \Delta I = const \Delta z D N_0 \quad (2.49)$$

Combining Eqs. (2.48) and (2.49), one obtains that the difference current ΔI is independent of the separation of the quadrant detector and the cantilever. This relation is true, if the light spot is smaller than the quadrant detector. If it is greater, the difference current ΔI becomes smaller with increasing distance. In reality the light beam has a Gaussian intensity profile. For small movements Δx (compared to the diameter of the light spot at the quadrant detector), Eq. (2.49) still holds. Larger movements Δx , however, will introduce a nonlinear response. If the AFM is operated in a constant force mode, only small movements Δx of the light spot will

occur. The feedback loop will cancel out all other movements.

The scanning of a sample with an AFM can twist the microfabricated cantilevers because of lateral forces (Mate et al. 1987; Marti et al. 1990; Meyer and Amer 1990b) and affect the images (den Boef 1991). When the tip is subjected to lateral forces, it will twist the cantilever and the light beam reflected from the end of the cantilever will be deflected perpendicular to the ordinary deflection direction. For many investigations this influence of lateral forces is unwanted. The design of the triangular cantilevers stems from the desire, to minimize the torsion effects. However, lateral forces open up a new dimension in force measurements. They allow, for instance, a distinction of two materials because of the different friction coefficient, or the determination of adhesion energies. To measure lateral forces the original optical lever AFM has to be modified. The only modification compared with Fig. 2.29 is the use of a quadrant detector photodiode instead of a two-segment photodiode and the necessary readout electronics, see Fig. 2.9a. The electronics calculates the following signals:

$$\begin{aligned} U_{\text{Normal Force}} &= \alpha [(I_{\text{Upper Left}} + I_{\text{Upper Right}}) \\ &\quad - (I_{\text{Lower Left}} + I_{\text{Lower Right}})] \\ U_{\text{Lateral Force}} &= \beta [(I_{\text{Upper Left}} + I_{\text{Lower Left}}) \\ &\quad - (I_{\text{Upper Right}} + I_{\text{Lower Right}})] \end{aligned} \quad (2.50)$$

The calculation of the lateral force as a function of the deflection angle does not have a simple solution for cross-sections other than circles. An approximate formula for the angle of twist for rectangular beams is (Baumeister and Marks 1967)

$$\theta = \frac{M_t L}{\beta G b^3 h} \quad (2.51)$$

where $M_t = F_y \ell$ is the external twisting moment due to lateral force, F_y , and β a constant determined by the value of h/b . For the equation to hold, h has to be larger than b .

Inserting the values for a typical microfabricated cantilever with integrated tips

$$\begin{aligned} b &= 6 \times 10^{-7} \text{ m} \\ h &= 10^{-5} \text{ m} \\ L &= 10^{-4} \text{ m} \\ \ell &= 3.3 \times 10^{-6} \text{ m} \\ G &= 5 \times 10^{10} \text{ Pa} \\ \beta &= 0.333 \end{aligned} \quad (2.52)$$

into Eq. (2.51) we obtain the relation

$$F_y = 1.1 \times 10^{-4} N \times \theta \quad (2.53)$$

Typical lateral forces are of order 10^{-10} N.

Sensitivity

The sensitivity of this setup has been calculated in various papers (Colchero et al. 1991; Sarid 1991; Colchero 1993). Assuming a Gaussian beam the resulting output signal as a function of the deflection angle is dispersion like. Equation (2.47) shows that the sensitivity can be increased by increasing the intensity of the light beam I_{tot} or by decreasing the divergence of the laser beam. The upper bound of the intensity of the light I_{tot} is given by saturation effects on the photodiode. If we decrease the divergence of a laser beam we automatically increase the beam waist. If the beam waist becomes larger than the width of the cantilever we start to get diffraction. Diffraction sets a lower bound on the divergence angle. Hence one can calculate the optimal beam waist w_{opt} and the optimal divergence angle β (Colchero et al. 1991; Colchero 1993)

$$\begin{aligned} w_{\text{opt}} &\approx 0.36 b \\ \theta_{\text{opt}} &\approx 0.89 \frac{\lambda}{b} \end{aligned} \quad (2.54)$$

The optimal sensitivity of the optical lever then becomes

$$e[\text{mW}/\text{rad}] = 1.8 \frac{b}{\lambda} I_{\text{tot}}[\text{mW}] \quad (2.55)$$

The angular sensitivity optical lever can be measured by introducing a parallel plate into the beam. A tilt of the parallel plate results in a displacement of the beam, mimicking an angular deflection.

Additional noise source can be considered. Of little importance is the quantum mechanical uncertainty of the position (Colchero et al. 1991; Colchero 1993), which is for typical cantilevers at room temperature

$$\Delta z = \sqrt{\frac{\hbar}{2m\omega_0}} = 0.05 \text{ fm} \quad (2.56)$$

where \hbar is the Planck constant ($=6.626 \times 10^{-34}$ Js). At very low temperatures and for high frequency cantilevers this could become the dominant noise source. A second noise source is the shot noise of the light. The shot noise is related to the particle number. We can calculate the number of photons incident on the detector

$$n = \frac{I\tau}{\hbar\omega} = \frac{I\lambda}{2\pi Bhc} = 1.8 \times 10^9 \frac{I[\text{W}]}{B[\text{Hz}]} \quad (2.57)$$

where I is the intensity of the light, τ the measurement time, $B = 1/\tau$ the bandwidth, and c the speed of light. The shot noise is proportional to the square root of the number of particles. Equating the shot noise signal with the signal resulting for the deflection of the cantilever one obtains

$$\Delta z_{shot} = 68 \frac{L}{w} \sqrt{\frac{B[\text{kHz}]}{I[\text{mW}]}} [\text{fm}] \quad (2.58)$$

where w is the diameter of the focal spot. Typical AFM setups have a shot noise of 2 pm. The thermal noise can be calculated from the equipartition principle. The amplitude at the resonant frequency is

$$\Delta z_{therm} = 129 \sqrt{\frac{B}{k[\text{N/m}]\omega_0 Q}} [\text{pm}] \quad (2.59)$$

A typical value is 16 pm. Upon touching the surface, the cantilever increases its resonant frequency by a factor of 4.39. This results in a new thermal noise amplitude of 3.2 pm for the cantilever in contact with the sample.

2.4.2.3 Piezoresistive Detection

Implementations

An alternative detection system which is not as widely spread as the optical detection schemes are piezoresistive cantilevers (Ashcroft and Mermin 1976; Stahl et al. 1994; Kassing and Oesterschulze 1997). These cantilevers are based on the fact that the resistivity of certain materials, in particular of Si, changes with the applied stress. Figure 2.30 shows a typical implementation of a piezo-resistive cantilever. Four resistances are integrated on the chip, forming a Wheatstone bridge. Two of the resistors are in unstrained parts of the cantilever, the other two are measuring the bending at the point of the maximal deflection. For instance when an AC voltage is applied between terminals a and c one can measure the detuning of the bridge between terminals b and d. With such a connection the output signal varies only due to bending, but not due to changing of the ambient temperature and thus the coefficient of the piezoresistance.

Sensitivity

The resistance change is (Kassing and Oester-schulze 1997)

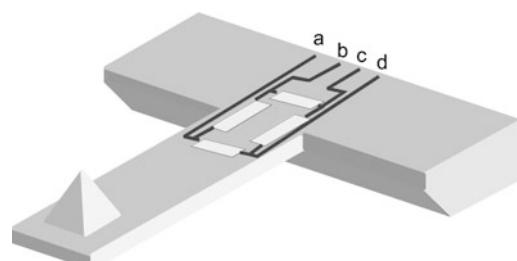


Fig. 2.30 A typical setup for a piezoresistive readout

$$\frac{\Delta R}{R_0} = \Pi \delta \quad (2.60)$$

where Π is the tensor element of the piezo-resistive coefficients, δ the mechanical stress tensor element and R_0 the equilibrium resistance. For a single resistor they separate the mechanical stress and the tensor element in longitudinal and transversal components

$$\frac{\Delta R}{R_0} = \Pi_t \delta_t + \Pi_\ell \delta_\ell \quad (2.61)$$

The maximum value of the stress components are $\Pi_t = -64.0 \times 10^{-11} \text{ m}^2/\text{N}$ and $\Pi_\ell = -71.4 \times 10^{-11} \text{ m}^2/\text{N}$ for a resistor oriented along the (110) direction in silicon (Kassing and Oesterschulze 1997). In the resistor arrangement of Fig. 2.30, two of the resistors are subject to the longitudinal piezo-resistive effect and two of them are subject to the transverse piezo-resistive effect. The sensitivity of that setup is about four times that of a single resistor, with the advantage that temperature effects cancel to first order. The resistance change is then calculated as

$$\frac{\Delta R}{R_0} = \Pi \frac{3Eh}{2L^2} \Delta z = \Pi \frac{6L}{bh^2} F_z \quad (2.62)$$

where $\Pi = 67.7 \times 10^{-11} \text{ m}^2/\text{N}$ is the averaged piezo-resistive coefficient. Plugging in typical values for the dimensions (Fig. 2.24) ($L = 100 \mu\text{m}$, $b = 10 \mu\text{m}$, $h = 1 \mu\text{m}$) one obtains

$$\frac{\Delta R}{R_0} = \frac{4 \times 10^{-5}}{nN} F_z \quad (2.63)$$

The sensitivity can be tailored by optimizing the dimensions of the cantilever.

2.4.2.4 Capacitance Detection

The capacitance of an arrangement of conductors depends on the geometry. Generally speaking, the capacitance increases for decreasing separations. Two parallel plates form a simple capacitor (see Fig. 2.31, upper left), with the capacitance

$$C = \frac{\varepsilon \varepsilon_0 A}{x} \quad (2.64)$$

where A is the area of the plates, assumed equal, and x is the separation. Alternatively, one can consider a sphere versus an infinite plane (see Fig. 2.31, lower left). Here the capacitance is (Sarid 1991)

$$C = 4\pi \varepsilon_0 R \sum_{n=2}^{\infty} \frac{\sin h(\alpha)}{\sin h(n\alpha)} \quad (2.65)$$

where R is the radius of the sphere, and α is defined by

$$\alpha = \ln \left[1 + \frac{z}{R} + \sqrt{\frac{z^2}{R^2} + 2\frac{z}{R}} \right] \quad (2.66)$$

One has to keep in mind that capacitance of a parallel plate capacitor is a nonlinear function of the separation. Using a voltage divider one can circumvent this problem. Figure 2.32a shows a low pass filter. The output voltage is given by

$$U_{out} = U_{\approx} \frac{\frac{1}{j\omega C}}{R + \frac{1}{j\omega C}} = U_{\approx} \frac{1}{j\omega CR + 1} \cong \frac{U_{\approx}}{j\omega CR} \quad (2.67)$$

Here, U_{\approx} is the driving voltage, C is given by Eq. (2.64), ω is the excitation frequency and j is

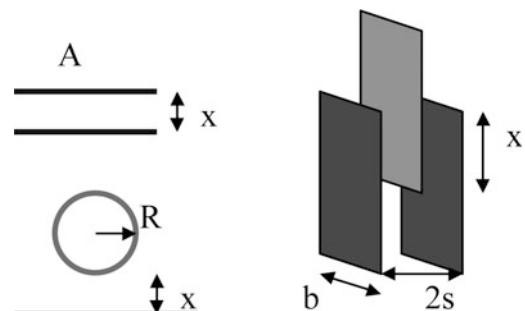


Fig. 2.31 Three possible arrangements of a capacitive readout. The *upper left* shows the cross section through a parallel plate capacitor. The *lower left* shows the geometry sphere versus plane. The *right side* shows the more complicated, but linear capacitive readout

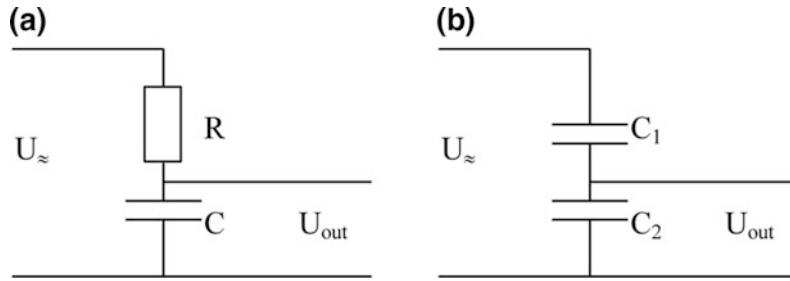


Fig. 2.32 Measuring the capacitance. The *left side*, (a) shows a low pass filter, the *right side*, (b) shows a capacitive divider. C (left) or C_2 (right) are the capacitances under test

the imaginary unit. The approximate relation in the end is true when $\omega CR \gg 1$. This is equivalent to the statement that C is fed by a current source, since R must be large in this setup. Plugging Eq. (2.64) into Eq. (2.67) and neglecting the phase information one obtains

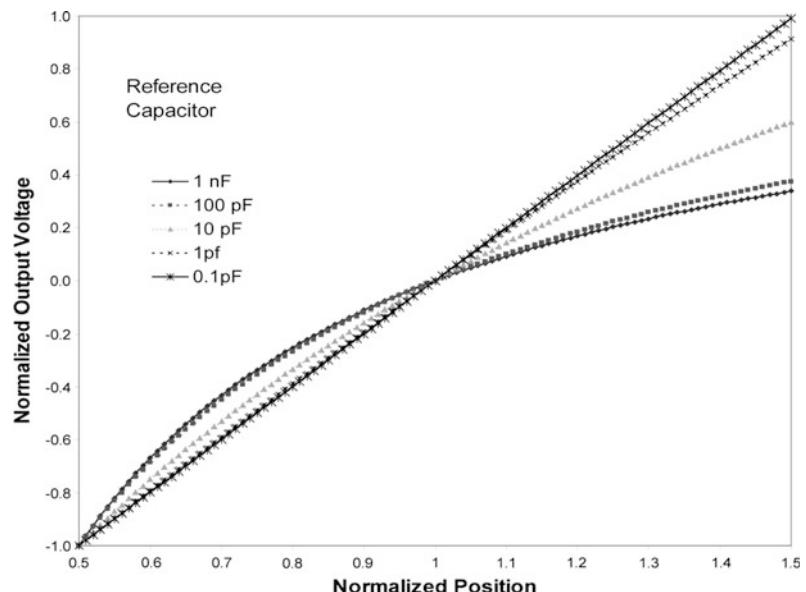
$$U_{out} = \frac{U_{\approx}x}{\omega R \epsilon \epsilon_0 A} \quad (2.68)$$

which is linear in the displacement x .

Figure 2.32b shows a capacitive divider. Again the output voltage U_{out} is given by

$$U_{out} = U_{\approx} \frac{C_1}{C_2 + C_1} = U_{\approx} \frac{C_1}{\frac{\epsilon \epsilon_0 A}{x} + C_1} \quad (2.69)$$

Fig. 2.33 Linearity of the capacitance readout as a function of the reference capacitor



If there is a stray capacitance C_s then Eq. (2.69) is modified as

$$U_{out} = U_{\approx} \frac{C_1}{\frac{\epsilon \epsilon_0 A}{x} + C_s + C_1} \quad (2.70)$$

Provided $C_s + C_1 \ll C_2$ one has a system which is linear in x . The driving voltage U_{\approx} has to be large (more than 100 V) to have the output voltage in the range of 1 V. The linearity of the readout depends on the capacitance C_1 (Fig. 2.33).

Another idea is to keep the distance constant and to change the relative overlap of the plates (see Fig. 2.31, right side). The capacitance of the

moving center plate versus the stationary outer plates becomes

$$C = C_s + 2 \frac{\varepsilon \varepsilon_0 b x}{S} \quad (2.71)$$

where the variables are defined in Fig. 2.31. The stray capacitance comprises all effects, including the capacitance of the fringe fields. When length x is comparable to the width b of the plates one can safely assume that the stray capacitance is constant, independent of x . The main disadvantage of this setup is that it is not as easily incorporated in a microfabricated device as the others.

Sensitivity

The capacitance itself is not a measure of the sensitivity, but its derivative is indicative of the signals one can expect. Using the situation described in Fig. 2.31, upper left, and in Eq. (2.64) one obtains for the parallel plate capacitor

$$\frac{dC}{dx} = -\frac{\varepsilon \varepsilon_0 A}{x^2} \quad (2.72)$$

Assuming a plate area A of $20 \mu\text{m}$ by $40 \mu\text{m}$ and a separation of $1 \mu\text{m}$ one obtains a capacitance of 31 fF (neglecting stray capacitance and the capacitance of the connection leads) and a dC/dx of $3.1 \times 10^{-8} \text{ F/m} = 31 \text{ fF}/\mu\text{m}$. Hence it is of paramount importance to maximize the area between the two contacts and to minimize the distance x . The latter however is far from being trivial. One has to go to the limits of microfabrication to achieve a decent sensitivity.

If the capacitance is measured by the circuit shown in Fig. 2.32 one obtains for the sensitivity

$$\frac{dU_{out}}{U_{\approx}} = \frac{dx}{\omega R \varepsilon \varepsilon_0 A} \quad (2.73)$$

Using the same value for A as above, setting the reference frequency to 100 kHz , and selecting $R = 1 \text{ G}\Omega$, we get the relative change of the output voltage U_{out} to

$$\frac{dU_{out}}{U_{\approx}} = \frac{22.5 \times 10^{-6}}{\overset{\circ}{\text{A}}} \times dx \quad (2.74)$$

A driving voltage of 45 V then translates to a sensitivity of $1 \text{ mV}/\text{\AA}$. A problem in this setup is the stray capacitances. They are in parallel to the original capacitance and decrease the sensitivity considerably.

Alternatively one could build an oscillator with this capacitance and measure the frequency. RC-oscillators typically have an oscillation frequency of

$$f_{res} \propto \frac{1}{RC} = \frac{x}{R \varepsilon \varepsilon_0 A} \quad (2.75)$$

Again the resistance R must be of the order of $1 \text{ G}\Omega$, when stray capacitances C_s are neglected. However C_s is of the order of 1 pF . Therefore one gets $R = 10 \text{ M}\Omega$. Using these values the sensitivity becomes

$$df_{res} = \frac{C dx}{R(C + C_s)^2 x} \approx \frac{0.1 \text{ Hz}}{\overset{\circ}{\text{A}}} dx \quad (2.76)$$

The bad thing is that the stray capacitances have made the signal nonlinear again. The linearized setup in Fig. 2.31 has a sensitivity of

$$\frac{dC}{dx} = 2 \frac{\varepsilon \varepsilon_0 b}{s} \quad (2.77)$$

Substituting typical values, $b = 10 \mu\text{m}$, $s = 1 \mu\text{m}$ one gets $dC/dx = 1.8 \times 10^{-10} \text{ F/m}$. It is noteworthy that the sensitivity remains constant for scaled devices.

Implementations

The readout of the capacitance can be done in different ways (Neubauer et al. 1990; Godden-henrich et al. 1990). All include an alternating current or voltage with frequencies in the 100 kHz to the 100 MHz range. One possibility is to build a tuned circuit with the capacitance of the cantilever determining the frequency. The resonance frequency of a high quality Q tuned circuit is

$$\omega_0 = (LC)^{-\frac{1}{2}} \quad (2.78)$$

where L is the inductance of the circuit. The capacitance C includes not only the sensor capacitance but also the capacitance of the leads. The precision of a frequency measurement is mainly determined by the ratio of L and C

$$Q = \left(\frac{L}{C}\right)^{\frac{1}{2}} \frac{1}{R} \quad (2.79)$$

Here R symbolizes the losses in the circuit. The higher the quality the more precise the frequency measurement. For instance a frequency of 100 MHz and a capacitance of 1 pF gives an inductance of 250 μ H. The quality becomes then 2.5×10^8 . This value is an upper limit, since losses are usually too high.

Using a value of $dC/dx = 31 \text{ fF}/\mu\text{m}$ one gets $\Delta C/\text{\AA} = 3.1 \text{ aF}/\text{\AA}$. With a capacitance of 1 pF one gets

$$\frac{\Delta\omega}{\omega} = \frac{1}{2} \frac{\Delta C}{C} \quad (2.80)$$

$$\Delta\omega = 100 \text{ MHz} \times \frac{1}{2} \frac{3.1 \text{ aF}}{1 \text{ pF}} = 155 \text{ Hz}$$

This is the frequency shift for 1 \AA deflection. The calculation shows, that this is a measurable quantity. The quality also indicates that there is no physical reason why this scheme should not work.

2.4.3 Combinations for 3D-Force Measurements

Three dimensional force measurements are essential if one wants to know all the details of the interaction between the tip and the cantilever. The straightforward attempt to measure three forces is complicated, since force sensors such as interferometers or capacitive sensors need a minimal detection volume, which often is too large. The second problem is that the force-sensing tip has to be held by some means. This implies that one of the three Cartesian axes is stiffer than the others.

However by the combination of different sensors one can achieve this goal. Straight cantilevers are employed for these measurements, because they can be handled analytically. The key observation is, that the optical lever method does not determine the position of the end of the cantilever. It measures the orientation. In the previous sections, one has always made use of the fact, that for a force along one of the orthogonal symmetry directions at the end of the cantilever (normal force, lateral force, force along the cantilever beam axis) there is a one to one correspondence of the tilt angle and the deflection. The problem is, that the force along the cantilever beam axis and the normal force create a deflection in the same direction. Hence what is called the normal force component is actually a mixture of two forces. The deflection of the cantilever is the third quantity, which is not considered in most of the AFMs. A fiber optic interferometer in parallel to the optical lever measures the deflection. Three measured quantities then allow the separation of the three orthonormal force directions, as is evident from Eqs. (2.27) and (2.33) (Fujisawa et al. 1994a, b; Grafstrom et al. 1994; Overney et al. 1994; Warmack et al. 1994).

Alternatively one can put the fast scanning direction along the axis of the cantilever. Forward and backward scans then exert opposite forces F_x . If the piezo movement is linearized, both force components in AFM based on the optical lever detection can be determined. In this case, the normal force is simply the average of the forces in the forward and backward direction. The force from the front, F_x , is the difference of the forces measured in forward and backward direction.

2.4.4 Scanning and Control Systems

Almost all SPMs use piezo translators to scan the tip or the sample. Even the first STM (Binnig et al. 1982; Binnig and Rohrer 1983) and some of the predecessor instruments (Young et al. 1971, 1972) used them. Other materials or setups

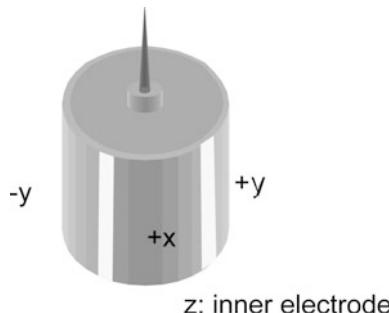


Fig. 2.34 Schematic drawing of a piezoelectric tube scanner. The piezo ceramic is molded into a tube form. The outer electrode is separated into four segments and connected to the scanning voltages. The z-voltage is applied to the inner electrode

for nano-positioning have been proposed, but were not successful (Gerber and Marti 1985; Garcia Cantù and Huerta Garnica 1990).

2.4.4.1 Piezo Tubes

A popular solution is tube scanners (Fig. 2.34). They are now widely used in SPMs due to their simplicity and their small size (Binnig and Smith 1986; Chen 1992a). The outer electrode is segmented in four equal sectors of 90°. Opposite sectors are driven by signals of the same magnitude, but opposite sign. This gives, through bending, a two-dimensional movement on, approximately, a sphere. The inner electrode is normally driven by the z signal. It is possible, however, to use only the outer electrodes for scanning and for the z-movement. The main drawback of applying the z-signal to the outer electrodes is, that the applied voltage is the sum of both the x- or y-movement and the z-movement. Hence a larger scan size effectively reduces the available range for the z-control.

2.4.4.2 Piezo Effect

An electric field applied across a piezoelectric material causes a change in the crystal structure, with expansion in some directions and contraction in others. Also, a net volume change occurs (Ashcroft and Mermin 1976). Many SPMs use the transverse piezo electric effect, where the

applied electric field \vec{E} is perpendicular to the expansion/contraction direction.

$$\Delta L = L(\vec{E} \cdot \vec{n})d_{31} = L \frac{V}{t} d_{31} \quad (2.81)$$

where d_{31} is the transverse piezoelectric constant, V is the applied voltage, t is the thickness of the piezo slab or the distance between the electrodes where the voltage is applied, L is the free length of the piezo slab, and \vec{n} is the direction of polarization. Piezo translators based on the transverse piezoelectric effect have a wide range of sensitivities, limited mainly by mechanical stability and breakdown voltage.

2.4.4.3 Scan Range

The calculation of the scanning range of a piezotube is difficult (Carr 1988; Chen 1992a, b). The bending of the tube depends on the electric fields and the nonuniform strain induced. A finite element calculation where the piezo tube was divided into 218 identical elements was used (Carr 1988) to calculate the deflection. On each node the mechanical stress, stiffness, strain and piezoelectric stress was calculated when a voltage was applied on one electrode. The results were found to be linear on the first iteration and higher-order corrections were very small even for large electrode voltages. It was found that to first order the x- and z-movement of the tube could be reasonably well approximated by assuming that the piezo tube is a segment of a torus. Using this model one obtains

$$dx = (V_+ - V_-)|d_{31}| \frac{L^2}{2td} \quad (2.82)$$

$$dz = (V_+ + V_- - 2V_z)|d_{31}| \frac{L}{2t} \quad (2.83)$$

where $|d_{31}|$ is the coefficient of the transversal piezoelectric effect, L is tube's free length, t is tube's wall thickness, d is tube's diameter, V_+ is voltage on positive outer electrode while V_- is voltage of the opposite quadrant negative electrode, and V_z is voltage of inner electrode. The

cantilever or sample mounted on the piezotube has an additional lateral movement because the point of measurement is not in the end plane of the piezotube. The additional lateral displacement of the end of the tip is $\ell \sin\varphi \approx \ell\varphi$, where ℓ is the tip length and φ is the deflection angle of the end surface. Assuming that the sample or cantilever are always perpendicular to the end of the walls of the tube and calculating with the torus model one gets for the angle

$$\varphi = \frac{L}{R} = \frac{2dx}{L} \quad (2.84)$$

where R is the radius of curvature the piezo tube. Using the result of Eq. (2.84) one obtains for the additional x-movement

$$dx_{add} = \ell\varphi = \frac{2dx\ell}{L} = (V_+ - V_-)|d_{31}| \frac{\ell L}{td} \quad (2.85)$$

and for the additional z-movement due to the x-movement

$$dz_{add} = \ell - \ell \cos \varphi = \frac{\ell\varphi^2}{2} = \frac{2\ell(dx)^2}{L^2} \\ = (V_+ - V_-)^2 |d_{31}|^2 \frac{\ell L^2}{2t^2 d^2} \quad (2.86)$$

Carr (1988) assumed for his finite element calculations that the top of the tube was completely free to move and, as a consequence, the top surface was distorted, leading to a deflection angle about half that of the geometrical model. Depending on the attachment of the sample or the cantilever this distortion may be smaller, leading to a deflection angle in-between that of the geometrical model and the one of the finite element calculation.

2.4.4.4 Nonlinearities and Creep

Piezo materials with a high conversion ratio, i.e. a large d_{31} or small electrode separations, with large scanning ranges are hampered by substantial hysteresis resulting in a deviation from

linearity by more than 10%. The sensitivity of the piezo ceramic material (mechanical displacement divided by driving voltage) decreases with reduced scanning range, whereas the hysteresis is reduced. A careful selection of the material for the piezo scanners, the design of the scanners, and of the operating conditions is necessary to get optimum performance.

2.4.4.5 Linearization Strategies

Passive Linearization: Calculation

The analysis of images affected by piezo nonlinearities (Libioulle et al. 1991; Stoll 1991; Durselen et al. 1995; Fu 1995) shows that the dominant term is

$$x = AV + BV^2 \quad (2.87)$$

where x is the excursion of the piezo, V the applied voltage and A and B two coefficients describing the sensitivity of the material. Equation (2.87) holds for scanning from $V = 0$ to large V . For the reverse direction the equation becomes

$$x = \tilde{A}V - \tilde{B}(V - V_{max})^2 \quad (2.88)$$

where \tilde{A} and \tilde{B} are the coefficients for the back scan, and V_{max} is the applied voltage at the turning point. Both equations demonstrate the true x-travel is small at the beginning of the scan and becomes larger towards the end. Therefore images are stretched at the beginning and compressed at the end.

Similar equation hold for the slow scan direction. The coefficients, however, are different. The combined action causes a greatly distorted image. This distortion can be calculated. The data acquisition systems record the signal as a function of V . However the data is measured as a function of x . Therefore we have to distribute the x -values evenly across the image this can be done by inverting an approximation of Eq. (2.87). First we write

$$x = AV \left(1 - \frac{B}{A} V \right) \quad (2.89)$$

For $B \ll A$ we can approximate

$$V = \frac{x}{A} \quad (2.90)$$

We now substitute Eq. (2.90) into the non-linear term of Eq. (2.89). This gives

$$\begin{aligned} x &= AV \left(1 + \frac{Bx}{A^2} \right) \\ V &= \frac{x}{A} \frac{1}{1 + Bx/A^2} \approx \frac{x}{A} \left(1 - \frac{Bx}{A^2} \right) \end{aligned} \quad (2.91)$$

Hence an equation of the type

$$x_{true} = x(\alpha - \beta x/x_{max}) \quad \text{with } 1 = \alpha - \beta \quad (2.92)$$

takes out the distortion of an image. α and β are dependent on the scan range, the scan speed and on the scan history and have to be determined with exactly the same settings as for the measurement. x_{max} is the maximal scanning range. The condition for α and β guarantees that the image is transformed onto itself.

Similar equations as the empirical one shown above Eq. (2.92) can be derived by analyzing the movements of domain walls in piezo ceramics.

Passive Linearization: Measuring the Position

An alternative strategy is to measure the position of the piezo translators. Several possibilities exist.

- (1) The interferometers described above can be used to measure the elongation of the piezo elongation. Especially the fiber optic interferometer is easy to implement. The coherence length of the laser only limits the measurement range. However the signal is of periodic nature. Hence a direct use of the signal in a feedback circuit for the position is not possible. However as a measurement tool and, especially, as a calibration tool the

interferometer is without competition. The wavelength of the light, for instance in a HeNe laser is so well defined that the precision of the other components determines the error of the calibration or measurement.

- (2) The movement of the light spot on the quadrant detector can be used to measure the position of a piezo (Barrett and Quate 1991). The output current changes by $0.5 \frac{\text{A}}{\text{cm}} \times \frac{P[\text{W}]}{R[\text{cm}]}$. Typical values ($P = 1 \text{ mW}$, $R = 0.001 \text{ cm}$) give 0.5 A/cm . The noise limit is typically $0.15 \text{ nm} \times \sqrt{\frac{\Delta f[\text{Hz}]}{H[\text{W/cm}^2]}}$. Again this means that the laser beam above would have a 0.1 nm noise limitation for a bandwidth of 21 Hz. The advantage of this method is that, in principle, one can linearize two axes with only one detector.
- (3) A knife-edge blocking part of a light beam incident on a photodiode can be used to measure the position of the piezo. This technique, commonly used in optical shear force detection (Betzig et al. 1992; Toledo-Crow et al. 1992), has a sensitivity of better than 0.1 nm .
- (4) The capacitive detection (Griffith et al. 1990; Holman et al. 1996) of the cantilever deflection can be applied to the measurement of the piezo elongation. Equations (2.64) to (2.79) apply to the problem. This technique is used in some commercial instruments. The difficulties lie in the avoidance of fringe effects at the borders of the two plates. While conceptually simple, one needs the latest technology in surface preparation to get a decent linearity. The electronic circuits used for the readout are often proprietary.
- (5) Linear Variable Differential Transformers (LVDT) are a convenient means to measure positions down to 1 nm. They can be used together with a solid state joint setup, as often used for large scan range stages. Unlike the capacitive detection there are few difficulties to implementation. The sensors and the detection circuits LVDTs are available commercially.

- (6) A popular measurement technique is the use of strain gauges. They are especially sensitive when mounted on a solid state joint where the curvature is maximal. The resolution depends mainly on the induced curvature. A precision of 1 nm is attainable. The signals are low—a Wheatstone bridge is needed for the readout.

Active Linearization

Active linearization is done with feedback systems. Sensors need to be monotonic. Hence all the systems described above, with the exception of the interferometers are suitable. The most common solutions include the strain gauge approach, the capacitance measurement or the LVDT, which are all electronic solutions. Optical detection systems have the disadvantage that the intensity enters into the calibration.

2.4.4.6 Alternative Scanning Systems

The first STMs were based on piezo tripods (Binnig et al. 1982). The piezo tripod (Fig. 2.35) is an intuitive way to generate the three dimensional movement of a tip attached to its center. However, to get a suitable stability and scanning range, the tripod needs to be fairly large (about 50 mm). Some instruments use piezo stacks instead of monolithic piezoactuators. They are arranged in the tripod arrangement. Piezo stacks are thin layers of piezoactive materials glued together to form a device with up to 200 μm of actuation range. Preloading with a suitable metal casing reduces the nonlinearity.

If one tries to construct a homebuilt scanning system, the use of linearized scanning tables is

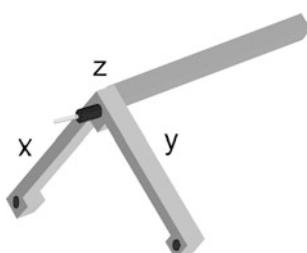


Fig. 2.35 An alternative type of piezo scanners: the tripod

recommended. They are built around solid state joints and actuated by piezo stacks. The joints guarantee that the movement is parallel with little deviation from the predefined scanning plane. Due to the construction it is easy to add measurement devices such as capacitive sensors, LVDTs or strain gauges which are essential for a closed loop linearization. Two-dimensional tables can be bought from several manufacturers. They have a linearity of better than 0.1% and a noise level of 10^{-4} – 10^{-5} of the maximal scanning range.

2.4.4.7 Control Systems

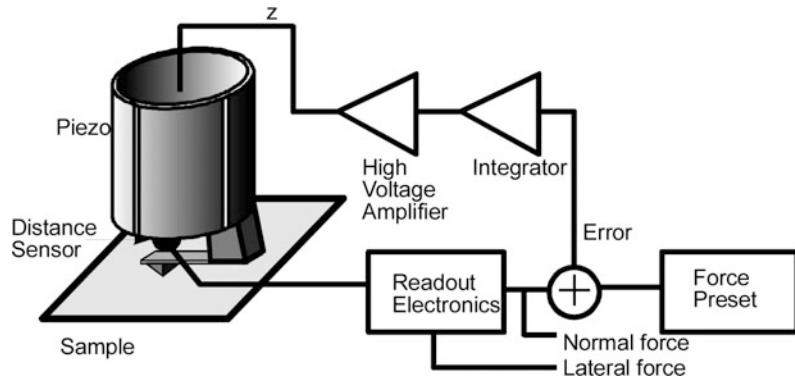
Basics

The electronics and software play an important role in the optimal performance of an SPM. Control electronics and software are supplied with commercial SPMs. Control electronic systems can use either analog or digital feedback. While digital feedback offers greater flexibility and the ease of configuration, analog feedback circuits might be better suited for ultralow noise operation. We will describe here the basic setups for AFM.

Figure 2.36 shows a block schematic of a typical AFM feedback loop. The signal from the force transducer is fed into the feedback loop consisting mainly of a subtraction stage to get an error signal and an integrator. The gain of the integrator (high gain corresponds to short integration times) is set as high as possible without generating more than 1% overshoot. High gain minimizes the error margin of the current and forces the tip to follow the contours of constant density of states as good as possible. This operating mode is known as Constant Force Mode. A high voltage amplifier amplifies the outputs of the integrator. As AFMs using piezotubes usually require ± 150 V at the output, the output of the integrator needs to be amplified by a high voltage amplifier.

In order to scan the sample, additional voltages at high tension are required to drive the piezo. For example, with a tube scanner, four scanning voltages are required, namely $+V_x$, $-V_x$, $+V_y$ and $-V_y$. The x- and y-scanning voltages are generated in a scan generator (analog or computer controlled). Both voltages are

Fig. 2.36 Block schematic of the feedback control loop of an AFM



input to the two respective power amplifiers. Two inverting amplifiers generate the input voltages for the other two power amplifiers. The topography of the sample surface is determined by recording the input-voltage to the high voltage amplifier for the z-channel as a function of x and y (Constant Force Mode).

Another operating mode is the Variable Force Mode. The gain in the feedback loop is lowered and the scanning speed increased such that the force on the cantilever is not any more constant. Here the force is recorded as a function of x and y.

Force Spectroscopy

Four modes of spectroscopic imaging are in common use with force microscopes: measuring lateral forces, $\partial F/\partial z$, $\partial F/\partial x$ spatially resolved, and measuring force versus distance curves. Lateral forces can be measured by detecting the deflection of a cantilever in a direction orthogonal to the normal direction. The optical lever deflection method most easily does this. Lateral force measurements give indications of adhesion forces between the tip and the sample.

$\partial F/\partial z$ measurements probe the local elasticity of the sample surface. In many cases the measured quantity originates from a volume of a few cubic nanometers. The $\partial F/\partial z$ or local stiffness signal is proportional to Young's modulus, as far as one can define this quantity. Local stiffness is measured by vibrating the cantilever by a small amount in z-direction. The expected signal for very stiff samples is zero: for very soft samples one gets, independent of the stiffness, also a

constant signal. This signal is again zero for the optical lever deflection and equal to the driving amplitude for interferometric measurements. The best sensitivity is obtained when the compliance of the cantilever matches the stiffness of the sample.

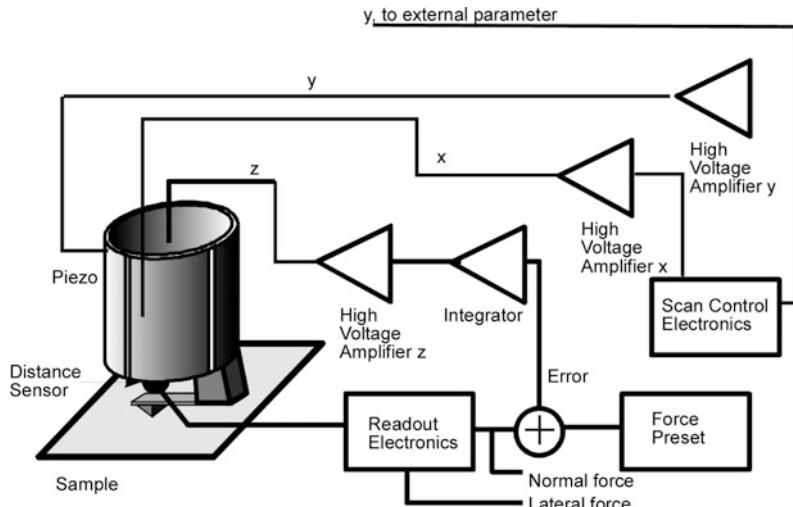
A third spectroscopic quantity is the lateral stiffness. It is measured by applying a small modulation in the x-direction on the cantilever. The signal is again optimal when the lateral compliance of the cantilever matches the lateral stiffness of the sample. The lateral stiffness is, in turn, related to the shear modulus of the sample.

Detailed information on the interaction of the tip and the sample can be gained by measuring force versus distance curves. It is necessary to have cantilevers with high enough compliance to avoid instabilities due to the attractive forces on the sample.

Using the Control Electronics as a Two-Dimensional Measurement Tool

Usually the control electronics of an AFM is used to control the x- and y-piezo signals while several data acquisition channels record the position dependent signals. The control electronics can be used in another way: it can be viewed as a two-dimensional function generator. What is normally the x- and y-signal can be used to control two independent variables of an experiment. The control logic of the AFM then ensures that the available parameter space is systematically probed at equally spaced points.

Fig. 2.37 Wiring of an AFM to measure friction force curves along a line



An example is friction force curves measured along a line across a step on graphite.

Figure 2.37 shows the connections. The z-piezo is connected as usual, like the x-piezo. However the y-output is used to command the desired input parameter. The offset of the y-channel determines the position of the tip on the sample surface, together with the x-channel.

2.4.4.8 Some Imaging Processing Methods

The visualization and interpretation of images from AFMs is intimately connected to the processing of these images. An ideal AFM is a noise-free device that images a sample with perfect tips of known shape and has perfect linear scanning piezos. In reality, AFMs are not that ideal. The scanning device in AFMs is affected by distortions. The distortions are both linear and nonlinear. Linear distortions mainly result from imperfections in the machining of the piezo-translators causing crosstalk from the Z-piezo to the X- and Y-piezos, and vice versa. Among the linear distortions, there are two kinds which are very important. First, scanning piezos invariably have different sensitivities along the different scan axes due to the variation of the piezo material and uneven sizes of the electrode areas. Second, the same reasons might cause the scanning axes not to be orthogonal. Furthermore, the plane in which the piezoscaner moves for

constant height z is hardly ever coincident with the sample plane. Hence, a linear ramp is added to the sample data. This ramp is especially bothersome when the height z is displayed as an intensity map.

The nonlinear distortions are harder to deal with. They can affect AFM data for a variety of reasons. First, piezoelectric ceramics do have a hysteresis loop, much like ferromagnetic materials. The deviations of piezoceramic materials from linearity increase with increasing amplitude of the driving voltage. The mechanical position for one voltage depends on the previously applied voltages to the piezo. Hence, to get the best position accuracy, one should always approach a point on the sample from the same direction. Another type of nonlinear distortion of the images occurs when the scan frequency approaches the upper frequency limit of the X- and Y-drive amplifiers or the upper frequency limit of the feedback loop (z-component). This distortion, due to feedback loop, can only be minimized by reducing the scan frequency. On the other hand, there is a simple way to reduce distortions due to the X- and Y-piezo drive amplifiers. To keep the system as simple as possible, one normally uses a triangular waveform for driving the scanning piezos. However, triangular waves contain frequency components as multiples of the scan frequency. If the cutoff frequency of the X- and Y-drive electronics or of

the feedback loop is too close to the scanning frequency (two or three times the scanning frequency), the triangular drive voltage is rounded off at the turning points. This rounding error causes, first, a distortion of the scan linearity and, second, through phase lags, the projection of part of the backward scan onto the forward scan. This type of distortion can be minimized by carefully selecting the scanning frequency and by using driving voltages for the X- and Y-piezos with waveforms like trapezoidal waves, which are closer to a sine wave. The values measured for X, Y, or Z piezos are affected by noise. The origin of this noise can be either electronic, some disturbances, or a property of the sample surface due to adsorbates. In addition to this incoherent noise, interference with main and other equipment nearby might be present. Depending on the type of noise, one can filter it in the real space or in Fourier space. The most important part of image processing is to visualize the measured data. Typical AFM data sets can consist of many thousands to over a million points per plane. There may be more than one image plane present. The AFM data represents a topography in various data spaces.

Most commercial data acquisition systems use implicitly some kind of data processing. Since the original data is commonly subject to slopes on the surface, most programs use some kind of slope correction. The least disturbing way is to substrate a plane $z(x, y) = Ax + By + C$ from the data. The coefficients are determined by fitting $z(x, y)$ to the data. Another operation is to subtract a second order function such as $z(x, y) = Ax^2 + By^2 + Cxy + Dx + Ey + F$. Again, the parameters are determined with a fit. This function is appropriate for almost plane data, where the nonlinearity of the piezos caused such a distortion.

In the image processing software from Digital Instruments, up to three operations are performed on the raw data. First, a zero-order flatten is applied. The flatten operation is used to eliminate image bow in the slow scan direction (caused by physical bow in the instrument itself), slope in the slow scan direction, bands in the image (caused by differences in the scan height from

one scan line to the next). The flatten operation takes each scan line and subtracts the average value of the height along each scan line from each point in that scan line. This brings each scan line to the same height. Next, a first-order planefit is applied in the fast scan direction. The planefit operation is used to eliminate bow and slope in the fast scan direction. The planefit operation calculated a best-fit plane for the image and subtracts it from the image. This plane has a constant non-zero slope in the fast scan direction. In some cases, higher-order polynomial “plane” may be required. Depending upon the quality of the raw data, the flatten operation and/or the planefit operation may not be required at all.

2.5 Closure

Since introduction of the STM in 1981 and AFM in 1985, many variations of probe-based microscopies, referred to as SPMs, have been developed. While the pure imaging capabilities of SPM techniques are dominated by the application of these methods at their early development stages, the physics of probe-sample interactions and the quantitative analyses of tribological, electronic, magnetic, biological, and chemical surfaces have now become of increasing interest. Nanoscale science and technology are strongly driven by SPMs which allow investigation and manipulation of surfaces down to the atomic scale. With growing understanding of the underlying interaction mechanisms, SPMs have found application in many fields outside basic research fields. In addition, various derivatives of all these methods have been developed for special applications, some of them targeting far beyond microscopy.

References

- Abraham DW, Mamin HJ, Ganz E, Clark J (1986) Surface modification with the scanning tunneling microscope. IBM J Res Dev 30:492–499
- Albrecht TR, Quate CF (1987) Atomic resolution imaging of a nonconductor by atomic force microscopy. J Appl Phys 62:2599–2602

- Albrecht TR, Akamine S, Carver TE, Quate CF (1990) Microfabrication of cantilever styli for the atomic force microscope. *J Vac Sci Technol A* 8:3386–3396
- Albrecht TR, Grutter P, Rugar D, Smith DPE (1992) Low temperature force microscope with all-fiber interferometer. *Ultramicroscopy* 42–44:1638–1646
- Alexander S, Hellemans L, Marti O, Schneir J, Elings V, Hansma PK (1989) An atomic-resolution atomic-force microscope implemented using an optical lever. *J Appl Phys* 65:164–167
- Amelio S, Goldade AV, Rabe U, Scherer V, Bhushan B, Arnold W (2001) Measurements of mechanical properties of ultra-thin diamond-like carbon coatings using atomic force acoustic microscopy. *Thin Solid Films* 392:75–84
- Anczykowski B, Krueger D, Babcock KL, Fuchs H (1996) Basic properties of dynamic force spectroscopy with the scanning force microscope in experiment and simulation. *Ultramicroscopy* 66:251–259
- Andoh Y, Oguchi S, Kaneko R, Miyamoto T (1992) Evaluation of very thin lubricant films. *J Phys D Appl Phys* 25:A71–A75
- Ashcroft NW, Mermin ND (1976) Solid state physics. Holt Reinhart and Winston, New York
- Avila A, Bhushan B (2010) “Electrical measurement techniques in atomic force microscopy”, (Invited). *Crit Rev Solid State Mater Sci* 35:38–51
- Barbara PF, Adams DM, O’Connor DB (1999) Characterization of organic thin film materials with near-field scanning optical microscopy (NSOM). *Annu Rev Mater Sci* 29:433–469
- Barrett RC, Quate CF (1991) Optical scan-correction system applied to atomic force microscopy. *Rev Sci Instrum* 62:1393
- Basire C, Ivanov DA (2000) Evolution of the lamellar structure during crystallization of a semicrystalline-amorphous polymer blend: time-resolved hot-stage SPM study. *Phys Rev Lett* 85:5587–5590
- Baumeister T, Marks SL (1967) Standard handbook for mechanical engineers, 7th edn. McGraw-Hill, New York
- Betzig E, Troutman JK, Harris TD, Weiner JS, Kostelak RL (1991) Breaking the diffraction barrier—optical microscopy on a nanometric scale. *Science* 251:1468–1470
- Betzig E, Finn PL, Weiner JS (1992) Combined shear force and near-field scanning optical microscopy. *Appl Phys Lett* 60:2484
- Bhushan B (1995) Micro/nanotribology and its applications to magnetic storage devices and MEMS. *Tribol Int* 28:85–96
- Bhushan B (1997) Micro/nanotribology and its applications, vol E330. Kluwer Academic Pub, Dordrecht, Netherlands
- Bhushan B (1999a) Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton, Florida
- Bhushan B (1999b) Wear and mechanical characterisation on micro-to picoscales using AFM. *Int Mat Rev* 44:105–117
- Bhushan B (2001a) Modern tribology handbook, Vol. 1: principles of tribology. CRC Press, Boca Raton, Florida
- Bhushan B (2001b) Fundamentals of tribology and bridging the gap between the macro- and micro/nanoscales. NATO Science Series II, vol 10. Kluwer Academic Pub., Dordrecht, Netherlands
- Bhushan B (2001c) Nano- to microscale wear and mechanical characterization studies using scanning probe microscopy. *Wear* 251:1105–1123
- Bhushan B (2005) Nanotribology and nanomechanics. *Wear* 259:1507–1531
- Bhushan B (2008) Nanotribology, nanomechanics and nanomaterials characterization. *Phil Trans R Soc A* 366:1351–1381
- Bhushan B (2010–2013) Scanning probe microscopy in nanoscience and nanotechnology, vols 1–3. Springer, Heidelberg, Germany
- Bhushan B (2011) Nanotribology and Nanomechanics I & II, 3rd edn. Springer, Heidelberg, Germany
- Bhushan B (2013a) Principles and applications of tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B, Blackman GS (1991) Atomic force microscopy of magnetic rigid disks and sliders and its applications to tribology. *ASME J Tribol* 113:452–458
- Bhushan B, Dandavate C (2000) Thin-film friction and adhesion studies using atomic force microscopy. *J Appl Phys* 87:1201–1210
- Bhushan B, Fuchs H (2006) Applied scanning probe methods II—scanning probe microscopy techniques, III—characterization, and IV—industrial applications. Springer, Heidelberg, Germany
- Bhushan B, Fuchs H (2007) Applied scanning probe methods VII—biomimetics and industrial applications. Springer, Heidelberg, Germany
- Bhushan B, Fuchs H (2009) Applied scanning probe methods XI—scanning probe microscopy techniques; XII—characterization; and XIII—biomimetics and industrial applications. Springer, Heidelberg, Germany
- Bhushan B, Goldade AV (2000) Measurements and analysis of surface potential change during wear of single-crystal silicon (100) at ultralow loads using Kelvin probe microscopy. *Appl Surf Sci* 157:373–381
- Bhushan B, Kawata S (2007) Applied scanning probe methods VI—characterization. Springer, Heidelberg, Germany
- Bhushan B, Koinkar VN (1994) Nanoindentation hardness measurements using atomic force microscopy. *Appl Phys Lett* 75:5741–5746
- Bhushan B, Qi J (2003) Phase contrast imaging of nanocomposites and molecularly-thick lubricant films in magnetic media. *Nanotechnology* 14:886–895
- Bhushan B, Ruan J (1994) Atomic-scale friction measurements using friction force microscopy: Part II—application to magnetic media. *ASME J Tribol* 116:389–396

- Bhushan B, Sundararajan S (1998) Micro/nanoscale friction and wear mechanisms of thin films using atomic force and friction force microscopy. *Acta Mater* 46:3793–3804
- Bhushan B, Ruan J, Gupta BK (1993) A scanning tunnelling microscopy study of fullerene films. *J Phys D Appl Phys* 26:1319–1322
- Bhushan B, Koinkar VN, Ruan J (1994) Microtribology of magnetic media. *Proc Inst Mech Eng Part J: J Eng Tribol* 208:17–29
- Bhushan B, Israelachvili JN, Landman U (1995) Nanotribology: friction, wear, and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996) Nano/Picoindentation measurements using capacitive transducer in atomic force microscopy. *Philos Mag A* 74:1117–1128
- Bhushan B, Fuchs H, Hosaka S (2004) Applied scanning probe methods. Springer, Heidelberg, Germany
- Bhushan B, Fuchs H, Kawata S (2007) Applied scanning probe methods V—scanning probe microscopy techniques. Springer, Heidelberg, Germany
- Bhushan B, Fuchs H, Tomitori M (2008) Applied scanning probe methods VIII—scanning probe microscopy techniques; IX—characterization; and X—biomimetics and industrial applications. Springer, Heidelberg, Germany
- Binggeli M, Christoph R, Hintermann HE, Colchero J, Marti O (1993) Friction force measurements on potential controlled graphite in an electrolytic environment. *Nanotechnology* 4:59–63
- Binnig G (1992) Force microscopy. *Ultramicroscopy* 42–44:7–15
- Binnig G, Rohrer H (1983) Scanning tunnelling microscopy. *Surf Sci* 126:236–244
- Binnig G, Smith DPE (1986) Single-tube three-dimensional scanner for scanning tunneling microscopy. *Rev Sci Instrum* 57:1688
- Binnig G, Rohrer H, Gerber C, Weibel E (1982) Surface studies by scanning tunneling microscopy. *Phys Rev Lett* 49:57–61
- Binnig G, Quate CF, Gerber Ch (1986) Atomic force microscope. *Phys Rev Lett* 56:930–933
- Binnig G, Gerber Ch, Stoll E, Albrecht TR, Quate CF (1987) Atomic resolution with atomic force microscope. *Europhys Lett* 3:1281–1286
- Blackman GS, Mate CM, Philpott MR (1990) Interaction forces of a sharp tungsten tip with molecular films on silicon surface. *Phys Rev Lett* 65:2270–2273
- Bonnell DA (ed) (1993) Scanning tunneling microscopy and spectroscopy—theory, techniques, and applications. VCH, New York
- Burham NA, Colton RJ, Pollock HM (1991) Interpretation issues in force microscopy. *J Vac Sci Technol A* 9:2548–2556
- Burnham NA, Colton RJ (1989) Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *J Vac Sci Technol A* 7:2906–2913
- Burnham NA, Domiguez DD, Mowery RL, Colton RJ (1990) Probing the surface forces of monolayer films with an atomic force microscope. *Phys Rev Lett* 64:1931–1934
- Carr RG (1988) Finite element analysis of PZT tube scanner motion for scanning tunneling microscopy. *J Microsc* 152:379–385
- Chen CJ (1992a) In situ testing and calibration of tube piezoelectric scanners. *Ultramicroscopy* 42–44:1653–1658
- Chen CJ (1992b) Electromechanical deflections of piezoelectric tubes with quartered electrodes. *App Phys Lett* 60:132
- Cleveland JP, Manne S, Bocek D, Hansma PK (1993) A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Rev Sci Instrum* 64:403–405
- Colchero J (1993) Reibungskraftmikroskopie, Ph.D. Dissertation, Physics Faculty, University of Konstanz, Konstanz, Germany
- Colchero J, Marti O, Bielefeldt H, Mlynek J (1991) Scanning force and friction microscopy. *Phys Stat Sol* 131:73–75
- Cooley JW, Tukey JW (1965) An algorithm for machine calculation of complex fourier series. *Math Comput* 19:297
- Coombs JH, Pethica JB (1986) Properties of vacuum tunneling currents: anomalous barrier heights. *IBM J Res Dev* 30:455–459
- Dai HJ, Hafner JH, Rinzler AG, Colbert DT, Smalley RE (1996) Nanotubes as nanoprobe in scanning probe microscopy. *Nature* 384:147–150
- Delawski E, Parkinson BA (1992) Layer-by-layer etching of two-dimensional metal chalcogenides with the atomic force microscope. *J Am Chem Soc* 114:1661–1667
- den Boef AJ (1991) The influence of lateral forces in scanning force microscopy. *Rev Sci Instrum* 62:88–92
- DeVecchio D, Bhushan B (1997) Localized surface elasticity measurements using an atomic force microscope. *Rev Sci Instrum* 68:4498–4505
- DeVecchio D, Bhushan B (1998) Use of a nanoscale kelvin probe for detecting wear precursors. *Rev Sci Instrum* 69:3618–3624
- Drake B, Prater CB, Weisenhorn AL, Gould SAC, Albrecht TR, Quate CF, Cannell DS, Hansma HG, Hansma PK (1989) Imaging crystals, polymers and processes in water with the atomic force microscope. *Science* 243:1586–1589
- Durig U, Zuger O, Stalder A (1992) Interaction force detection in scanning probe microscopy: methods and applications. *J Appl Phys* 72:1778–1797
- Durselen R, Grunewald U, Preuss W (1995) Calibration and applications of a high precision piezo scanner for nanometrology. *Scanning* 17:91–96
- Eigler DM, Schweizer EK (1990) Positioning single atoms with a scanning tunnelling microscope. *Nature* 344:524–528

- Erlandsson R, McClelland GM, Mate CM, Chiang S (1988) Atomic force microscopy using optical interferometry. *J Vac Sci Technol A* 6:266–270
- Foster J, Frommer J (1988) Imaging of liquid crystal using a tunneling microscope. *Nature* 333: 542–547
- Frisbie CD, Rozsnyai LF, Noy A, Wrighton MS, Lieber CM (1994) Functional group imaging by chemical force microscopy. *Science* 265:2071–2074
- Frommer J (1992) Scanning tunneling microscopy and atomic force microscopy in organic chemistry. *Angew Chem Int Ed Engl* 31:1298–1328
- Fu J (1995) In situ testing and calibrating of Z-Piezo of an atomic force microscope. *Rev Sci Instrum* 66:3785–3788
- Fuchs H, Bhushan B (2010) Biosystems—investigated by scanning probe microscopy. Springer, Heidelberg, Germany
- Fujisawa S, Ohta M, Konishi T, Sugawara Y, Morita S (1994a) Difference between the forces measured by an optical lever deflection and by an optical interferometer in an atomic force microscope. *Rev Sci Instrum* 65:644–647
- Fujisawa S, Kishi E, Sugawara Y, Morita S (1994b) Fluctuation in 2-dimensional stick-slip phenomenon observed with 2-dimensional frictional force microscope. *Jpn J Appl Phys (Part 1)* 33:3752–3755
- Garcia Cantù R, Huerta Garnica MA (1990) Long-scan imaging by STM. *J Vac Sci Technol A* 8:354
- Gerber C, Marti O (1985) Magnetostrictive positioner. *IBM Tech Discl Bull* 27:6373
- Giaever I (1960) Energy gap in superconductors measured by electron tunneling. *Phys Rev Lett* 5:147–148
- Giessibl FJ (1995) Atomic resolution of the silicon (111)-(7×7) surface by atomic force microscopy. *Science* 267:68–71
- Giessibl FJ, Gerber Ch, Binnig G (1991) A low-temperature atomic force/scanning tunneling microscope for ultrahigh vacuum. *J Vac Sci Technol B* 9:984–988
- Giessibl FJ, Sugawara Y, Morita S, Hosoi H, Sueoka K, Mukasa K, Sasahara A, Onishi H (2011) Noncontact atomic force microscopy and related topics. In: Bhushan B (ed) Nanotribology and nanomechanics I, 3rd edn. Springer, Heidelberg, Germany, pp 195–237
- Goldenhenrich T, Lemke H, Hartmann U, Heiden C (1990) Force microscope with capacitive displacement detection. *J Vac Sci Technol A* 8:383–387
- Grafstrom S, Ackermann J, Hagen T, Neumann R, Probst O (1994) Analysis of lateral force effects on the topography in scanning force microscopy. *J Vac Sci Technol B* 12:1559–1564
- Griffith JE, Miller GL, Green CA (1990) A scanning tunneling microscope with a capacitance-based position monitor. *J Vac Sci Technol B* 8:2023–2027
- Guntherodt HJ, Wiesendanger R (eds) (1992) Scanning tunneling microscopy I: general principles and applications to clean and adsorbate-covered surfaces. Springer, Berlin
- Guntherodt HJ, Anselmetti D, Meyer E (eds) (1995) Forces in scanning probe methods, vol E286. Kluwer Academic Publishers, Dordrecht, Netherlands
- Hafner JH, Cheung CL, Woolley AT, Lieber CM (2001) Structural and functional imaging with carbon nanotube AFM probes. *Prog Biophys Mol Biol* 77:73–110
- Hansma PK, Tersoff J (1987) Scanning tunneling microscopy. *J Appl Phys* 61:R1–R23
- Hansma PK, Drake B, Marti O, Gould SAC, Prater CB (1989) The scanning ion-conductance microscope. *Science* 243:641–643
- Hartmann U (1999) Magnetic force microscopy. *Annu Rev Mater Sci* 29:53–87
- Holman AE, Laman CD, Scholte PMLO, Heerens WC, Tuinstra F (1996) A calibrated scanning tunneling microscope equipped with capacitive sensors. *Rev Sci Instrum* 67:2274–2280
- Hug HJ, Moser A, Jung Th, Fritz O, Wadas A, Parashikov I, Guntherodt HJ (1993) Low temperature magnetic force microscopy. *Rev Sci Instrum* 64:2920–2925
- Husser OE, Craston DH, Bard AJ (1989) Scanning electrochemical microscopy—high resolution deposition and etching of materials. *J Electrochem Soc* 136:3222–3229
- Ibe JP, Bey PP, Brandon SL, Brizzolara RA, Burnham NA, DiLella DP, Lee KP, Marrian CRK, Colton RJ (1990) On the electrochemical etching of tips for scanning tunneling microscopy. *J Vac Sci Technol A* 8:3570–3575
- Jarvis SP, Oral A, Weihs TP, Pethica JB (1993) A novel force microscope and point contact probe. *Rev Sci Instrum* 64:3515–3520
- Kaneko R, Oguchi S (1990) Ion-implanted diamond tip for a scanning tunneling microscope. *Jap J Appl Phys* 28:1854–1855
- Kasai T, Bhushan B, Huang L, Su C (2004) Topography and phase imaging using the torsional resonance mode. *Nanotechnology* 15:731–742
- Kassing R, Oesterschulze E (1997) Sensors for scanning probe microscopy. In: Bhushan B (ed) Micro/nanotribology and its applications. Kluwer Academic Publishers, Dordrecht, Netherlands, pp 35–54
- Kirk MD, Albrecht T, Quate CF (1988) Low-temperature atomic force microscopy. *Rev Sci Instrum* 59:833–835
- Kobayashi A, Grey F, Williams RS, Ano M (1993) Formation of nanometer-scale grooves in silicon with a scanning tunneling microscope. *Science* 259:1724–1726
- Koinkar VN, Bhushan B (1996) Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J Vac Sci Technol A* 14:2378–2391
- Krotil U, Stifter T, Waschipky H, Weishaupt K, Hild S, Marti O (1999) Pulse force mode: a new method for the investigation of surface properties. *Surf Interface Anal* 27:336–340
- Lee DT, Pelz JP, Bhushan B (2002) Instrumentation for direct, low frequency scanning capacitance

- microscopy, and analysis of position dependent stray capacitance. *Rev Sci Instrum* 73:3523–3533
- Leung OM, Goh MC (1992) Orientation ordering of polymers by atomic force microscope tip-surface interactions. *Science* 225:64–66
- Libioulle N, Ronda A, Taborelli M, Gilles JM (1991) Deformations and nonlinearity in scanning tunneling microscope images. *J Vac Sci Technol B* 9:655–658
- Liu H, Bhushan B (2002) Investigation of nanotribological properties of self-assembled monolayers with alkyl and biphenyl spacer chains. *Ultramicroscopy* 91:185–202
- Lyo IW, Avouris Ph (1991) Field-induced nanometer-to-atomic-scale manipulation of silicon surfaces with the STM. *Science* 253:173–176
- Maivald P, Butt HJ, Gould SAC, Prater CB, Drake B, Gurley JA, Elings VB, Hansma PK (1991) Using force modulation to image surface elasticities with the atomic force microscope. *Nanotechnology* 2:103–106
- Majumdar A (1999) Scanning thermal microscopy. *Annu Rev Mater Sci* 29:505–585
- Majumdar A, Oden PI, Carrejo JP, Nagahara LA, Graham JJ, Alexander J (1992) Nanometer-scale lithography using the atomic force microscope. *Appl Phys Lett* 61:2293–2295
- Marti O, Amrein M (eds) (1993) STM and SFM in biology. Academic Press, San Diego
- Marti O, Drake B, Hansma PK (1987) Atomic force microscopy of liquid-covered surfaces: atomic resolution images. *Appl Phys Lett* 51:484–486
- Marti O, Gould S, Hansma PK (1988) Control electronics for atomic force microscopy. *Rev Sci Instrum* 59:836–839
- Marti O, Colchero J, Mlynek J (1990) Combined scanning force and friction microscopy of mica. *Nanotechnology* 1:141–144
- Martin Y, Wickramasinghe HK (1987) Magnetic imaging by force microscopy with 1000 Å resolution. *Appl Phys Lett* 50:1455–1457
- Martin Y, Williams CC, Wickramasinghe HK (1987) Atomic force microscope-force mapping and profiling on a sub 100-Å scale. *J Appl Phys* 61:4723–4729
- Martin Y, Abraham DW, Wickramasinghe HK (1988) High-resolution capacitance measurement and potentiometry by force microscopy. *Appl Phys Lett* 52:1103–1105
- Mate CM (1992) Atomic-force-microscope study of polymer lubricants on silicon surfaces. *Phys Rev Lett* 68:3323–3326
- Mate CM, McClelland GM, Erlandsson R, Chiang S (1987) Atomic-scale friction of a tungsten tip on a graphite surface. *Phys Rev Lett* 59:1942–1945
- Matey J, Blanc J (1985) Scanning capacitance microscopy. *J Appl Phys* 57:1437–1444
- McClelland GM, Erlandsson R, Chiang S (1987) Atomic force microscopy: general principles and a new implementation. In: Thompson DO, Chimenti DE (eds) Review of progress in quantitative nondestructive evaluation, vol 6B, Plenum, New York, pp 1307–1314
- Meyer E (1992) Atomic force microscopy. *Surf Sci* 41:3–49
- Meyer G, Amer NM (1988) Novel optical approach to atomic force microscopy. *Appl Phys Lett* 53:1045–1047
- Meyer G, Amer NM (1990a) Optical-beam-deflection atomic force microscopy: the NaCl (001) surface. *Appl Phys Lett* 56:2100–2101
- Meyer G, Amer NM (1990b) Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. *Appl Phys Lett* 57:2089–2091
- Meyer E, Overney R, Luthi R, Brodbeck D et al (1992) Friction force microscopy of mixed Langmuir-Blodgett films. *Thin Solid Films* 220:132–137
- Neubauer G, Coben SR, McClelland GM, Horne D, Mate CM (1990) Force microscopy with a bidirectional capacitance sensor. *Rev Sci Instrum* 61:2296–2308
- Nicolaides RL, Yong WE, Packard WF, Zhou HA et al (1988) Scanning tunneling microscope tip structures. *J Vac Sci Technol A* 6:445–447
- Nonnenmacher M, O’Boyle MP, Wickramasinghe HK (1991) Kelvin probe force microscopy. *Appl Phys Lett* 58:2921–2923
- O’Shea SJ, Welland ME, Rayment T (1992) Atomic force microscope study of boundary layer lubrication. *Appl Phys Lett* 61:2240–2242
- Ohnesorge F, Binnig G (1993) True atomic resolution by atomic force microscopy through repulsive and attractive forces. *Science* 260:1451–1456
- Overney RM, Takano H, Fujihira M, Paulus W, Ringsdorf H (1994) Anisotropy in friction and molecular stick-slip motion. *Phys Rev Lett* 72:3546–3549
- Palacio M, Bhushan B (2010) Normal and lateral force calibration techniques for AFM cantilevers. *Crit Rev Solid State* 35:73–104
- Park SI, Quate CF (1987) Digital filtering of STM images. *J Appl Phys* 62:312
- Parkinson B (1990) Layer-by-layer nanometer scale etching of two-dimensional substrates using the scanning tunneling microscopy. *J Am Chem Soc* 112:7498–7502
- Pohl DW (1986) Some design criteria in STM. *IBM J Res Dev* 30:417
- Pohl DW, Denk W, Lanz M (1984) Optical stethoscopy—image recording with resolution $\lambda/20$. *Appl Phys Lett* 44:651–653
- Prater CB, Hansma PK, Tortonese M, Quate CF (1991) Improved scanning ion-conductance microscope using microfabricated probes. *Rev Sci Instrum* 62:2634–2638
- Rabe U, Janser K, Arnold W (1996) Vibrations of free and surface-coupled atomic force microscope: theory and experiment. *Rev Sci Instrum* 67:3281–3293
- Radmacher M, Tillman RW, Fritz M, Gaub HE (1992) From molecules to cells: imaging soft samples with the atomic force microscope. *Science* 257:1900–1905
- Reinstaedter M, Rabe U, Scherer V, Hartmann U, Goldade A, Bhushan B, Arnold W (2003) On the

- nanoscale measurement of friction using atomic force microscope cantilever torsional resonances. *Appl Phys Lett* 82:2604–2606
- Reinstaedtler M, Kasai T, Rabe U, Bhushan B, Arnold W (2005) Imaging and measurement of elasticity and friction using the TR mode. *J Phys D Appl Phys* 38: R269–R282
- Ruan J, Bhushan B (1994a) Atomic-scale friction measurements using friction force microscopy: part I—general principles and new measurement techniques. *ASME J Tribol* 116:378–388
- Ruan J, Bhushan B (1994b) Atomic-scale and microscale friction of graphite and diamond using friction force microscopy. *J Appl Phys* 76:5022–5035
- Rugar D, Hansma PK (1990) Atomic force microscopy. *Phys Today* 43:23–30
- Rugar D, Mamin HJ, Guethner P (1989) Improved fiber-optical interferometer for atomic force microscopy. *Appl Phys Lett* 55:2588–2590
- Rugar D, Mamin HJ, Guethner P, Lambert SE, Stern JE, McFadyen I, Yogi T (1990) Magnetic force microscopy—general principles and application to longitudinal recording media. *J Appl Phys* 63:1169–1183
- Sarid D (1991) Scanning force microscopy. Oxford University Press, New York
- Sarid D, Elings V (1991) Review of scanning force microscopy. *J Vac Sci Technol B* 9:431–437
- Sarid D, Iams D, Weissenberger V, Bell LS (1988) Compact scanning-force microscope using laser diode. *Opt Lett* 13:1057–1059
- Scherer V, Bhushan B, Rabe U, Arnold W (1997) Local elasticity and lubrication measurements using atomic force and friction force microscopy at ultrasonic frequencies. *IEEE Trans Mag* 33:4077–4079
- Scherer V, Arnold W, Bhushan B (1999) Lateral force microscopy using acoustic friction force microscopy. *Surf Interface Anal* 27:578–587
- Schirmeisen A, Anczykowski B, Hoelscher H, Fuchs F (2011) Dynamic modes of atomic force microscopy. In: Bhushan B (ed) Nanotribology and nanomechanics I, 3rd edn. Springer, Heidelberg, Germany, pp 307–353
- Schoenenberger C, Alvarado SF (1989) A differential interferometer for force microscopy. *Rev Sci Instrum* 60:3131–3135
- Schoenenberger C, Alvarado SF (1990) Understanding magnetic force microscopy. *Z Phys B* 80:373–383
- Shen YR (1984) The principles of nonlinear optics. Wiley, New York
- Silver RM, Ehrichs EE, de Lozanne AL (1987) Direct writing of submicron metallic features with a scanning tunnelling microscope. *Appl Phys Lett* 51:247–249
- Smith D, Horber H, Gerber C, Binnig G (1989) Smectic liquid crystal monolayers on graphite observed by scanning tunneling microscopy. *Science* 245:43–45
- Smith D, Horber J, Binnig G, Nejoh H (1990) Structure, registry and imaging mechanism of alkyl-cyanobiphenyl molecules by tunnelling microscopy. *Nature* 344:641–644
- Stahl U, Yuan CW, Delozanne AL, Tortonese M (1994) Atomic force microscope using piezoresistive cantilevers and combined with a scanning electron microscope. *Appl Phys Lett* 65:2878–2880
- Stern JE, Terris BD, Mamin HJ, Rugar D (1988) Deposition and imaging of localized charge on insulator surfaces using a force microscope. *Appl Phys Lett* 53:2717–2719
- Stoll EP (1991) Restoration of STM images distorted by time-dependent piezo driver aftereffects. *Ultramicroscopy* 42–44:1585–1589
- Stroscio JA, Kaiser WJ (eds) (1993) Scanning tunneling microscopy. Academic Press, Boston
- Thomson WT, Dahleh MD (1998) Theory of vibration with applications, 5th Edn. Prentice Hall, Upper Saddle River, New Jersey
- Toledo-Crow R, Yang PC, Chen Y, Vaez-Iravani M (1992) Near-field differential scanning optical microscope with atomic force regulation. *Appl Phys Lett* 60:2957–2959
- Tsau L, Wang D, Wang KL (1994) Nanometer scale patterning of silicon (100) surface by an atomic force microscope operating in air. *Appl Phys Lett* 64:2133–2135
- Warmack RJ, Zheng XY, Thundat T, Allison DP (1994) Friction effects in the deflection of atomic force microscope cantilevers. *Rev Sci Instrum* 65:394–399
- Weaver JMR, Abraham DW (1991) High resolution atomic force microscopy potentiometry. *J Vac Sci Technol B* 9:1559–1561
- Weisenhorn AL, MacDougall JE, Gould JAC, Cox SD, Wise WS, Massie J, Maivald P, Elings VB, Stucky GD, Hansma PK (1990) Imaging and manipulating of molecules on a zeolite surface with an atomic force microscope. *Science* 247:1330–1333
- Weisenhorn AL, Egger M, Ohnesorge F, Gould SAC, Heyn SP, Hansma HG, Sinsheimer RL, Gaub HE, Hansma PK (1991) Molecular resolution images of Langmuir-Blodgett films and DNA by atomic force microscopy. *Langmuir* 7:8–12
- Wickramasinghe HK (2000) Progress in scanning probe microscopy. *Acta Mater* 48:347–358
- Wiesendanger R, Guntherodt HJ (eds) (1992) Scanning tunneling microscopy, II: further applications and related scanning techniques. Springer, Berlin
- Williams CC (1999) Two-dimensional dopant profiling by scanning capacitance microscopy. *Annu Rev Mater Sci* 29:471–504
- Williams CC, Wickramasinghe HK (1986) Scanning thermal profiler. *Appl Phys Lett* 49:1587–1589
- Williams CC, Wickramasinghe HK (1990) Microscopy of chemical-potential variations on an atomic scale. *Nature* 344:317–319
- Wolter O, Bayer T, Greschner J (1991) Micromachined silicon sensors for scanning force microscopy. *J Vac Sci Technol B* 9:1353–1357

- Yamanaka K, Tomita E (1995) Lateral force modulation atomic force microscope for selective imaging of friction forces. *Jpn J Appl Phys* 34:2879–2882
- Yamanaka K, Ogisco H, Kolosov O (1994) Ultrasonic force microscopy for nanometer resolution subsurface imaging. *Appl Phys Lett* 64:178–180
- Young R, Ward J, Scire F (1971) Observation of metal-vacuum-metal tunneling, field emission, and the transition region. *Phys Rev Lett* 27:922
- Young R, Ward J, Scire F (1972) The Topographiner: an instrument for measuring surface microtopography. *Rev Sci Instrum* 43:999

Calibration of Normal and Lateral Forces in Cantilevers Used in Atomic Force Microscopy

Manuel L.B. Palacio and Bharat Bhushan

Abstract

Atomic force microscopy (AFM) is an indispensable technique for nanoscale topographic imaging, as well as quantification of normal and lateral forces exerted on the AFM tip while interacting with the surface of materials. In order to measure these forces, an accurate determination of the normal and lateral forces exerted on the AFM cantilever is necessary. In this chapter, we present a critical review of various techniques for measuring cantilever stiffness in the normal and lateral/torsional directions in order to calibrate the normal and lateral forces exerted on AFM cantilevers. The key concepts of each technique are presented, along with a discussion of their advantages and disadvantages.

Keywords

Atomic force microscopy (AFM) · Scanning probe microscopy (SPM) · Tip-cantilever assembly · Normal force · Lateral force · Cantilever calibration · Nanotribology

3.1 Introduction

Scanning probe microscopy techniques, developed in the 1980s, are used to obtain high resolution three-dimensional (3-D) images, as well as physical and chemical characterization. These techniques rely on the use of a probe tip, which scans a specimen in order to generate an image. The most widely-used SPM is the atomic force microscope (AFM), first introduced in 1985. In AFM, the force between the tip and sample is used to produce an image (Bhushan 2011, 2017).

M.L.B. Palacio · B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

M.L.B. Palacio
e-mail: manuel.palacio@gmail.com

The tip can be in contact with the sample during imaging with AFM. Therefore, the AFM can be used for a variety of measurements. Over the years, the use of AFMs has expanded beyond surface profiling to other capabilities such as the measurement of adhesion, friction, wear, elastic/plastic mechanical properties, electrical, magnetic and thermal properties, as well as in situ nanofabrication and nanomanipulation.

Multiple AFM designs exist and examples of commercial small-sample and large-sample AFMs are shown in Fig. 3.1a, b, respectively. In the small-sample AFM design, the sample, generally no larger than $10\text{ mm} \times 10\text{ mm}$, is mounted onto a lead zirconate titanate (PZT) tube scanner which consists of separate electrodes to scan the sample precisely in the x–y plane in a raster pattern and to move the sample in the vertical (z) direction. In the large-sample AFM design, the sample is stationary while the tip is scanned. In both designs, a sharp tip at the free end of a flexible cantilever is brought into contact with the sample. Features on the sample surface cause the cantilever to deflect vertically and laterally during scanning, which is what is detected by the AFM. A laser beam from a diode laser (5 mW max peak output at 670 nm wavelength) is directed by a prism onto the back of a cantilever near its free end, tilted downward at about 12° with respect to the horizontal plane. The reflected beam from the vertex of the cantilever is directed through a mirror onto a split photodetector with four quadrants (commonly called position-sensitive detector or PSD). The differential signal from the top and bottom photodiodes provides the AFM signal which is a sensitive measure of the cantilever vertical deflection. For surface imaging, the tip is scanned either along or transverse to the longitudinal axis of the cantilever. Topographical features of the sample cause the tip to deflect in the vertical direction as the sample is scanned. This tip deflection will change the direction of the reflected laser beam, changing the intensity difference between the top and bottom sets of photodetectors (AFM signal).

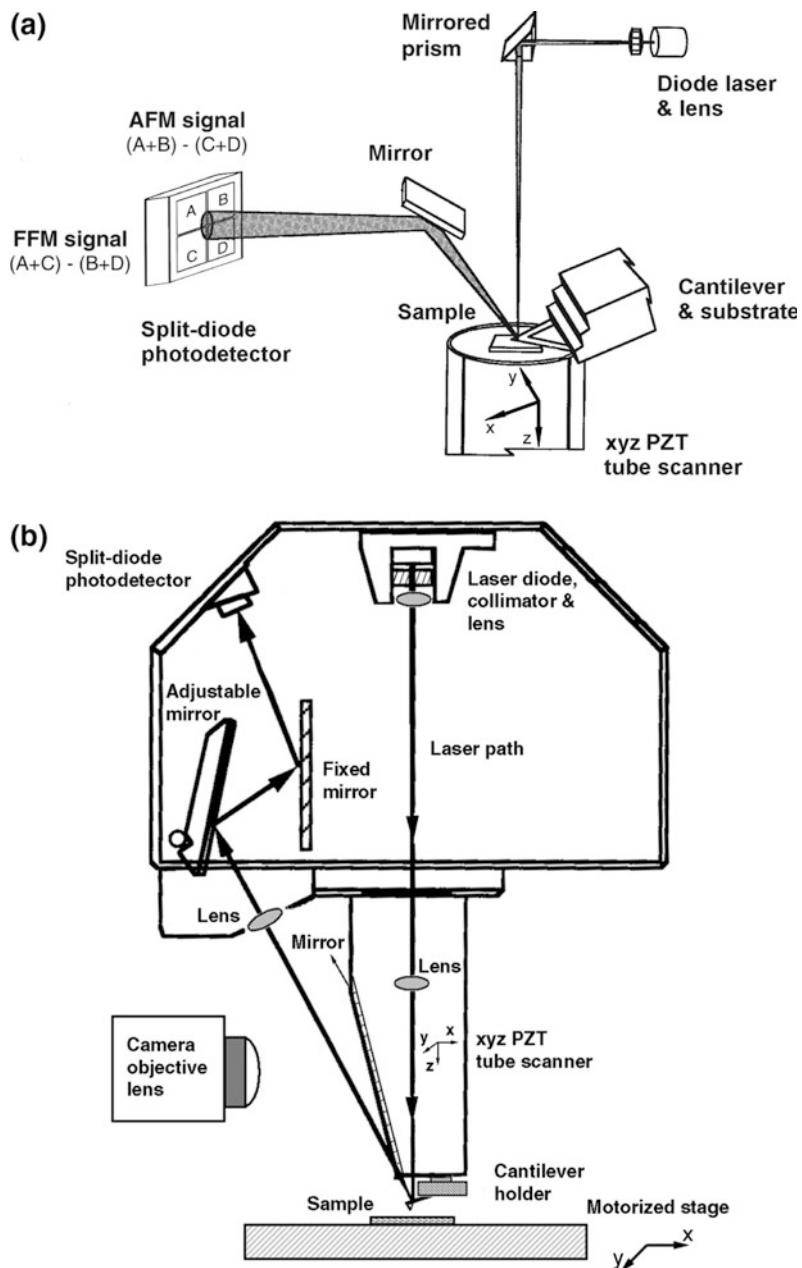
In the AFM operating mode called the height mode, for topographical imaging or for any other

operation in which the applied normal force is to be kept constant, a feedback circuit is used to modulate the voltage applied to the PZT scanner to adjust the height of the PZT, so that the cantilever vertical deflection (given by the intensity difference between the top and bottom detector) will remain constant during scanning. The PZT height variation is thus a direct measure of the surface roughness of the sample. Imaging is carried out either in contact mode or tapping mode. In contact mode, the tip is in contact at all times as it slides on the sample surface with the applied normal force kept constant. In tapping mode, the tip is oscillating as it slides on the sample surface, and is therefore not in contact at all times. The amplitude of oscillation is kept constant during the scan.

The friction force being applied at the tip during sliding can be measured using the quadrants on the left and right sides of the photodetector. In the so-called friction mode, the sample is scanned back and forth in a direction transverse to the long axis of the cantilever beam. Friction force between the sample and the tip will produce a twisting of the cantilever. As a result, the laser beam will be reflected out of the plane defined by the incident beam and the vertically reflected beam from an untwisted cantilever. This produces an intensity difference of the laser beam received in the left hand and right hand sets of quadrants of the photodetector. The intensity difference between the two sets of detectors (friction force microscope or FFM signal) is directly related to the degree of twisting and hence to the magnitude of the friction force. This method provides maps of the friction force (Meyer and Amer 1990; Bhushan 2011, 2017).

Illustrations of the two common cantilever configurations, triangular (or V-shaped) and rectangular are shown in Fig. 3.2. These cantilevers are made from a wide range of materials, the most common being Si_3N_4 (triangular) and silicon (rectangular). Silicon nitride cantilevers are less expensive than those made of silicon. They are very rugged and well suited to imaging in various conditions (Bhushan 2011, 2017). Microfabricated silicon nitride triangular beams with integrated square pyramidal tips

Fig. 3.1 Principle of operation of commercial
a Small-sample AFM, and
b large-sample AFM



made by plasma-enhanced chemical vapor deposition (PECVD) are used most commonly (Albrecht et al. 1990). These cantilevers are typically coated with a thin gold film to increase the laser signal reflected to the photodetector. AFM cantilevers are commercially available from Bruker (Camarillo, CA), Nanosensors GmbH (Aidlingen, Germany), and NT-MDT

(Moscow, Russia), among others. Typical specifications of a silicon nitride cantilever are 115–196 μm length, 17–41 μm width, and 0.6 μm thickness (NP series, Bruker). The pyramidal tips are highly symmetric with the end having a radius of about 20–50 nm. The tip side walls have a slope of 35°, and the height of the tip is about 3 μm .

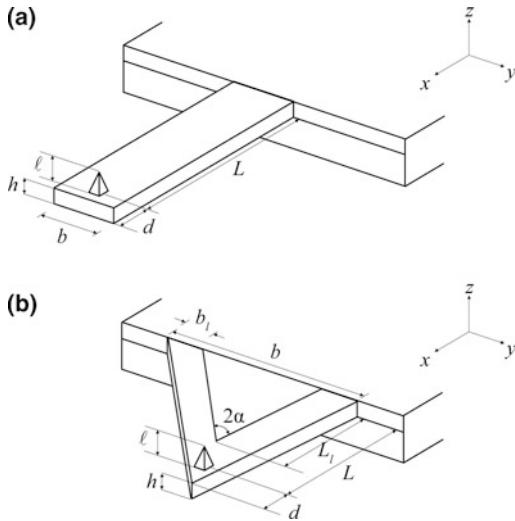


Fig. 3.2 Typical AFM cantilevers with **a** rectangular, and **b** triangular geometries. The cantilevers have length L , width b , and height (or thickness) h . The height of the tip is ℓ . The cantilever material is characterized by Young's modulus E , the shear modulus G and a mass density ρ

Microfabricated single-crystal silicon cantilevers with integrated tips are also used. Si tips are sharper than Si_3N_4 tips because they are directly formed by the anisotropic etch in single-crystal Si rather than using an etch pit as a mask for deposited materials (Wolter et al. 1991). Etched single-crystal n-type silicon rectangular cantilevers with square pyramidal tips typically have an end radius of less than 10 nm for contact and tapping mode. These cantilevers may be coated with a thin aluminum reflective film to increase the laser signal. Typical specifications for a rectangular cantilever for tapping mode applications are 125 μm length, 30 μm width, and 4 μm thickness. The tip side walls have a slope of 22.5°, and the tip height ranges from 10 to 15 μm (TESP series, Bruker). Since Si_3N_4 cantilevers are produced by thin film deposition, thinner cantilevers can be produced compared to those made with Si. Si_3N_4 is used for cantilevers with a normal stiffness of up to about 1 N/m and Si is used for higher stiffness applications.

An accurate determination of normal, bending, and torsion forces is necessary in order to measure the interaction forces between the

surface and the AFM tip. This requires determining the stiffness of the cantilever used, as well as the PSD calibration in order to convert the measured signal (in volts) into a force (in Newtons) (Bhushan 2011, 2017). There are a number of techniques reported in the literature to obtain the PSD calibration.

With regards to the estimation of cantilever stiffness in general, it will be shown in the next section that the stiffnesses are directly proportional to the cube of the thickness of the cantilever. This presents a source of uncertainty for the stiffness determined by using fundamental beam theory. Microfabricated cantilevers often have non-uniformity in their dimensions, including their thickness. Average measurement errors for the cantilever length, width, and thickness are about 1, 4, and 5%, respectively (Ohler 2010). The non-uniformity of the reflective coating used (Au or Al) also adds to the thickness uncertainty. Scanning electron microscopy (SEM) should be used instead of optical microscopy in order to measure these dimensions accurately. However, this could be time-consuming and difficult to implement in a routine manner (Sader et al. 1995).

There is also some variation coming from the mechanical properties of the cantilever material used in the calculations. For example, by using an ultrasonic measurement, Ruan and Bhushan (1994a) found the Young's modulus of the cantilever beam to be about 238 ± 18 GPa, which is less than that of bulk Si_3N_4 , 310 GPa (Bhushan 2011, 2017). Ohler (2010) estimates the error from the modulus measurement at 5% for Si and 20% for Si_3N_4 . The error for Si is smaller because the cantilever is made from the bulk material, while Si_3N_4 cantilevers are fabricated through a deposition process which causes its material properties to be different from the bulk.

Another common method for determining cantilever stiffness is through finite element analysis (FEA) (Neumeister and Ducker 1994). This approach is especially useful for non-rectangular configurations, e.g., triangular cantilevers. FEA suffers from inaccuracy of the dimensions of the cantilever, as well as the mechanical properties needed. Since there are

uncertainties in stiffness determination inherent from the beam theory-based approach and FEA methods, there is a need to experimentally determine the stiffness of AFM cantilevers in order to accurately measure normal and lateral forces.

The objective of this chapter is to review various techniques for calibrating the normal and lateral forces of AFM cantilevers. The key concepts will be discussed, along with advantages and disadvantages for each technique. A number of articles are available in the literature comparing various specific techniques (e.g., Clifford and Seah 2005; Cook et al. 2006; Pettersson et al. 2007; Palacio and Bhushan 2010). Palacio and Bhushan (2010) provided a comprehensive review of the theory, features, limitations, and experimental uncertainties, which is needed to guide the experimentalist in the proper implementation of various normal and lateral force calibration techniques. In this chapter, Sect. 3.2 discusses the analytical methods used to determine the normal and lateral stiffness. Section 3.3 reviews the various experimental techniques proposed for evaluating the normal stiffness, and Sect. 3.4 reviews the different techniques for calibrating the lateral force and/or lateral stiffness of AFM cantilevers. This chapter is based in part on a review article by Palacio and Bhushan (2010) on this subject.

3.2 Analytical Approaches for Determining the Cantilever Normal and Lateral Stiffness

The analytical methods used to determine the normal and lateral stiffness of the cantilever is briefly reviewed in this section.

The normal stiffness can be calculated using the geometrical and physical properties of the cantilever material (Albrecht et al. 1990; Meyer and Amer 1990; Sarid and Elings 1991). By definition, the normal stiffness k_z (also commonly referred to as the spring constant or spring stiffness) is given by $F_z = k_z \Delta z$, where F_z is the normal force, and Δz is the deflection of the

cantilever beam. From fundamental beam theory, the expression for the stiffness of a rectangular cantilever beam with a uniform cross section, with the load applied on its end and experiencing small deflections, is given by Timoshenko and Goodier (1970) and Thomson and Dahleh (1998)

$$k_z = \frac{3EI}{L^3} \quad (3.1)$$

where E is the Young's modulus of the material, L is the length of the beam, and I is the area moment of inertia of the cross section. Equation 3.1 assumes that the cantilever is homogeneous, isotropic, and exhibits linear elastic behavior. For a rectangular cross section with a width b and a height (or thickness) h , one obtains an expression for I (Timoshenko and Goodier 1970; Thomson and Dahleh 1998)

$$I = \frac{bh^3}{12} \quad (3.2)$$

$$k_z = \frac{Eb}{4} \left(\frac{h}{L} \right)^3 \quad (3.3)$$

Equation 3.3 is the intrinsic stiffness of the cantilever. For the purpose of calculating the stiffness of an AFM cantilever, a correction has to be made which accounts for the cantilever's tilt (on the order of 12°) relative to the horizontal axis. The "effective" stiffness is given by Hutter (2005) and Ohler (2010)

$$k_{z,eff} = \frac{k_z}{\cos^2 \gamma} \quad (3.4)$$

where, γ is the tilt angle of the cantilever.

If a lateral force F_y is applied to the end of the cantilever beam, the cantilever will bend sideways. The bending stiffness in the lateral direction, k_{yB} , is given by $F_y = k_{yB} \Delta y$, where Δy is the deflection in the y -direction. The quantity k_{yB} can be calculated with Eq. 3.3 by exchanging b and h (Timoshenko and Goodier 1970; Thomson and Dahleh 1998)

$$k_{yB} = \frac{Eh}{4} \left(\frac{b}{L} \right)^3 \quad (3.5)$$

Therefore, the bending stiffness in the lateral direction is larger than the stiffness for bending in the normal direction by $(b/h)^2$.

When the lateral force F_y is applied at the end of the tip, as in the case of the AFM cantilever, lateral deflection is accompanied by twisting or torsion along the cantilever axis. Instead of k_{yb} , we now define k_{yT} from $F_y = k_{yT}\Delta y$. For a wide, thin cantilever ($b \gg h$) experiencing a rotation ϕ , the torsional stiffness of the beam, k_ϕ , is given as $M = k_\phi\phi$, where M is the torque or torsion moment. From beam theory, the torsional stiffness is defined as follows (Timoshenko and Goodier 1970; Young and Budynas 2002)

$$k_\phi = \frac{Gb^3}{3L} \quad (3.6)$$

where G is the modulus of rigidity or shear modulus [= $E/2(1 + v)$, where v is the Poisson's ratio]. The relationship between E and G is valid for homogeneous, isotropic, and linear elastic materials. The lateral stiffness of the cantilever-tip assembly with the lateral load applied at the end of the tip (torque) is defined as (Timoshenko and Goodier 1970; Bhushan 2011, 2017)

$$k_{yT} = \frac{k_\phi}{\ell^2} \quad (3.7)$$

where ℓ is the length of the tip mounted at the end of the cantilever. For completeness, we note that the ratio k_{yT}/k_z is independent of the cantilever thickness and width, and it is much larger than one.

The normal stiffness of triangular cantilevers has been approximated by assuming that the cantilever is composed of two rectangular cantilevers in parallel. This "parallel beam approximation" (PBA) was first proposed by Albrecht et al. (1990), and modifications were suggested by Butt et al. (1993) and Sader (1995). In the latter, Sader (1995) argued that inappropriate width and length of the two rectangular cantilevers were used by Albrecht et al. (1990) and Butt et al. (1993). Based on finite element

analysis, the formula for the normal stiffness for the triangular cantilever is of the form (Sader 1995)

$$k_z = \frac{Eh^3b_l}{2(L+d)^3} \left(1 + \frac{4b_l^3}{b^3} (3 \cos \alpha - 2) \right)^{-1} \quad (3.8)$$

where α is one-half the included angle between the legs of the cantilever.

An approach considered to be more accurate than the PBA was proposed by Neumeister and Ducker (1994), who modeled the triangular cantilever by subdividing it into two parts, namely, a triangular plate corresponding to the front part and two prismatic beams corresponding to the legs of the cantilever. They derived expressions for the normal, lateral, and torsional stiffnesses. This approach was further modified by Clifford and Seah (2005).

The normal stiffness of commercial triangular Si_3N_4 cantilevers is typically from 0.01 to 0.6 N/m (Anonymous 2016). These cantilevers have a typical width to thickness ratio of 10–30. The width to thickness ratio and the triangular geometry results in 100–1000 times greater stiffness values in the lateral direction compared to the normal direction. Using the Young's modulus value of Si(111) which is 181 GPa, the rectangular Si cantilever described earlier (TESP series, Bruker) has stiffness values of typically 40, 2500, 2230 N/m and 0.36×10^{-6} N m in the normal (k_z), lateral bending (k_{yB}), lateral bending due to torque (k_{yT}), and torsional (k_ϕ) directions, respectively. A cantilever beam required for tapping mode is quite stiff and may not be sensitive enough and is therefore not well suited for measuring torsion. For friction measurements, the torsional stiffness should be minimized in order to be sensitive to the lateral forces. Long cantilevers with small thickness and large tip length are most suitable. An example of the dimensions for a rectangular silicon cantilever for high lateral force sensitivity is 200 μm length, 21 μm width, 0.4 μm thickness, and 12.5 μm tip length, which gives stiffness values of 0.007, 21, 1 N/m and

1.6×10^{-10} N m in the normal (k_z), lateral bending (k_{yB}), lateral bending due to torque (k_{yT}), and torsional (k_ϕ) directions, respectively. This cantilever has a lateral force sensitivity of 10 pN, assuming an angular resolution of 10^{-7} rad. With this particular geometry, sensitivity to lateral forces is improved by about a factor of 100 or more compared with more commonly used triangular Si_3N_4 or rectangular Si or Si_3N_4 cantilevers (Bhushan 2011, 2017).

Triangular cantilevers have been assumed to have a high lateral bending stiffness (relative to rectangular cantilevers), which minimizes lateral deflection of the cantilever during imaging (Albrecht and Quate 1987; Albrecht et al. 1990). This assumption has been questioned by Sader and co-workers, who conducted theoretical analysis and experiments to test this assumption (Sader 2003; Sader and Sader 2003). Their theoretical approach involved the determination of the lateral resistance R , defined as follows

$$R_x = \frac{k_x}{k_z} \quad (3.9)$$

$$R_y = \frac{k_y}{k_z} \quad (3.10)$$

where k_x is the lateral stiffness in the direction parallel to the longitudinal axis of the cantilever. The lateral resistances of the rectangular and triangular cantilevers are comparable only when the latter has narrow legs (i.e., b_l/b of the triangular cantilever shown in Fig. 3.2 is close to 0.1). However, typical triangular cantilevers have $b_l/b > 0.1$. In general, it was found that the rectangular cantilever is stiffer (less prone to bending) than the comparable triangular cantilever by as much as 7 and 4 times in the x- and y-directions, respectively (Sader 2003). Experimentally, this was confirmed on model macroscopic cantilevers where a rod was attached on the ends. By applying a torque and measuring the rotation angle, it was found that rectangular cantilevers are less prone to rotation, as predicted by theory (Sader and Sader 2003). These results imply that the assumption stated above is incorrect as rectangular cantilevers are actually less

susceptible to lateral forces (Sader 2003; Sader and Sader 2003). It appears that the use of triangular cantilevers is historical with no obvious advantage.

As indicated in the Introduction, errors in the measured cantilever dimensions and the mechanical properties of the cantilever material will lead to uncertainties in calculating the stiffness by using the analytical approaches presented in this section, such that both the normal and lateral cantilever stiffness have to be determined experimentally. Sections 3.3 and 3.4 describe the various techniques for calibrating the normal stiffness and the lateral force/stiffness, respectively.

3.3 Normal Stiffness Calibration Techniques

Experimental methods for measuring the normal stiffness of cantilevers are described in this section. The methods are mainly divided into static and dynamic techniques. Static methods rely on the cantilever deflection, either through a reference cantilever or reference mass. Dynamic methods depend on the oscillation of the cantilever. For the dynamic methods, a common parameter is the frequency f , of the cantilever, defined as (Thomson and Dahleh 1998)

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{k_z}{m_{eff}}} \quad (3.11)$$

where ω is the angular frequency and m_{eff} is the effective mass, which varies depending on the cantilever used. For instance, for a rectangular cantilever with $L/b > 5$, its effective mass is approximated as $m^* \approx 0.2427$ m, where m is the mass of the cantilever (with $m = \rho_c b h L$ and ρ_c is the density of the cantilever). Another important parameter is the quality factor, Q (also known as the Q-factor). The Q-factor is a metric which describes energy loss during an oscillation. A high Q-factor value indicates low energy dissipation per cycle, and is therefore desirable for cantilevers.

3.3.1 Static Methods

3.3.1.1 Reference Cantilever Method

One way of determining the cantilever stiffness is to use another larger cantilever with a known stiffness, i.e., a reference cantilever. Ruan and Bhushan (1994a) used a stainless steel spring sheet of known stiffness (width = 1.35 mm, thickness = 15 μm , free hanging length = 5.2 mm). One end of the spring was attached to the sample holder, and the other end was made to contact the cantilever tip during the measurement, see Fig. 3.3. They measured the piezo traveling distance for a given cantilever deflection. For a rigid sample (such as diamond, Fig. 3.3a), the piezo traveling distance Z_t (measured from the point where the tip touches the sample) should equal the cantilever deflection. Z_t should be determined beforehand in order to differentiate between the deflection of the cantilever and the spring sheet, as shown in Fig. 3.3b. To keep the cantilever deflection at the same level using a flexible spring sheet, the new piezo traveling distance Z'_t would be different from Z_t . The difference between Z'_t and Z_t corresponds to the deflection of the spring sheet. If the stiffness of the reference is k_s , the stiffness of the cantilever k_z can be calculated by Ruan and Bhushan (1994a)

$$(Z'_t - Z_t)k_s = Z_t k_z$$

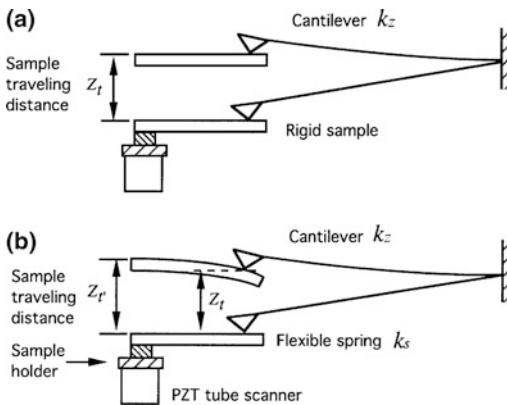


Fig. 3.3 Illustration showing the deflection of cantilever as it is pushed by **a** a rigid sample, or by **b** a flexible spring sheet (Ruan and Bhushan 1994a)

or

$$k_z = k_s(Z'_t - Z_t)/Z_t \quad (3.12)$$

A source of uncertainty in this method is the offset from positioning the cantilever relative to the end of the reference. Since the reference gets stiffer as load is applied farther from the end, the cantilever should contact the reference as close to its end as possible. If there is an offset in the positioning (such as that shown schematically in Fig. 3.2), k_z can be corrected, e.g., for a rectangular beam by using (Sader et al. 1995)

$$k_z = k_{z,off} \left(\frac{L+d}{L} \right)^3 \quad (3.13)$$

where $k_{z,off}$ is the normal stiffness measured with the offset, $L + d$ is the total length of the reference cantilever, and d is the distance away from the end of the reference that the load is applied.

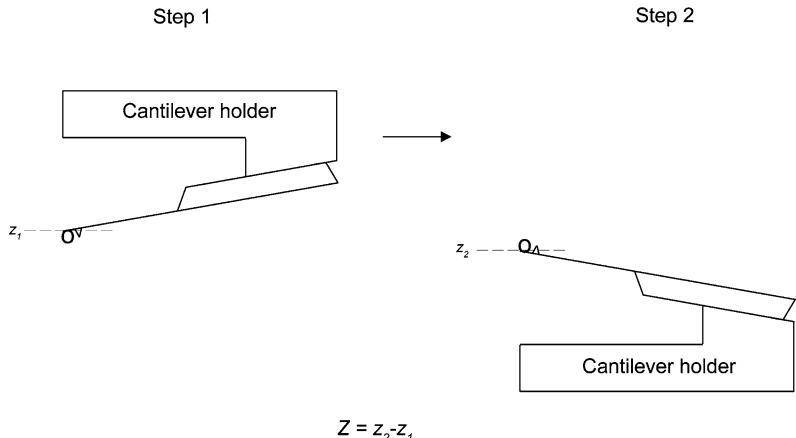
The advantage of the reference cantilever method is that it is experimentally simple. The deflection voltage signals can be used without the need to convert into meters as the calibration constant will cancel out in Eq. 3.12. However, this technique requires accurate positioning of the cantilever of interest relative to the reference cantilever, as discussed above. In addition, since the sharp tip is in contact with a single asperity, extreme roughness on the reference surface may cause variations on the measured piezo travel distance.

Commercial reference cantilevers referred to as a “force calibration chip” contain three rectangular cantilevers on single crystal silicon, and are available from Park Scientific Instruments (Tortonese and Kirk 1997). The cantilevers have lengths of 97, 197, and 397 μm , and have equal thickness of 2 μm and equal width of 29 μm . For these cantilevers, the stiffness ranged from 0.16 to 10 N/m, as determined using the resonance method (to be discussed in Sect. 3.2.2).

3.3.1.2 Inverted Loaded Cantilever Method

In this method, particles are attached to the end of the cantilever. As shown in Fig. 3.4, the

Fig. 3.4 Schematic of the inverted cantilever method



deflection of the cantilever is first measured after the addition of the particle and then remeasured after the cantilever has been inverted. The difference in the deflection, Z (in volts), is twice the deflection due to gravity. The stiffness is expressed as (Senden and Ducker 1994)

$$k_z = \frac{2m_s g}{z} = \frac{8\pi r^3 \rho_s g}{3z} \quad (3.14)$$

where the mass of the spherical particle is given in terms of r and ρ_s , its radius and density, and g is the gravitational constant. In Eq. 3.14, $z = \Omega Z$, where z and Z correspond to the deflection difference in meters and volts, respectively, and Ω is the deflection calibration which relates the cantilever deflection to the detector voltage reading. Senden and Ducker (1994) used tungsten spheres that are 10–50 μm in diameter.

The attachment of a particle and the inversion procedure both pose a risk of damage to the cantilever being analyzed. In addition, uncertainties in the radius and density of the particle will lead to errors in the measured k_z value. SEM would be a more preferable technique to optical microscopy in obtaining accurate particle dimensions. Optical microscopy is known to give an average error of about 5–10% in determining the dimensions of the particle (Ohler 2010).

3.3.1.3 Pendulum Method

A schematic illustrating this technique is shown in Fig. 3.5. As the name implies, the stiffness is

determined by pushing the tip against the pendulum and measuring the deflection of the pendulum and the cantilever, z_p and z_c , respectively. The force exerted by the cantilever and the component of the gravitational force of the pendulum in the direction of motion are the same at equilibrium (Butt et al. 1993)

$$F = k_z z_c = \frac{m_p g z_p}{l_p} \quad (3.15)$$

where m_p is the mass of the pendulum, and l_p is the length of the pendulum. The stiffness is given by

$$k_z = \frac{m_p g z_p}{l_p z_c} \quad (3.16)$$

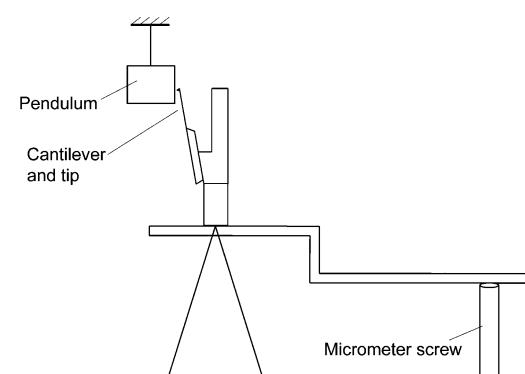


Fig. 3.5 Schematic of the pendulum method

The main disadvantage of this technique is the need for the pendulum set-up. The experiment itself presents a risk of damaging the cantilever. Calibration of the pendulum deflection is necessary, and this could be a source of measurement errors.

3.3.1.4 Miscellaneous Methods

The following are examples of methods where additional instrumentation needs to be implemented to determine the normal stiffness. The Electrostatic Force Balance (EFB), its schematic shown in Fig. 3.6a, was developed at the United States National Institute of Standards and Technology (Pratt et al. 2005; Shaw et al. 2007). Electrostatic force acting along the vertical axis is generated when voltages are applied to the coaxial cylinders. The displacement is then monitored with an interferometer. Another example is the so-called Nano Force Calibrator (NFC), its schematic shown in Fig. 3.6b, was

developed at the Korea Research Institute of Standards and Science (Kim et al. 2006, 2007). In this setup, the cantilever is placed in contact with a precision balance, and controlled displacement is applied by a stage capable of moving in the nanometer range. It should be mentioned that other comparable force balances, as well as stiffness artifacts (standard cantilevers) have been developed in the National Physical Laboratory in the United Kingdom (Cumpson and Hedley 2003; Leach et al. 2006) and in the Physikalisch-Technische Bundesanstalt (PTB) in Germany (Behrens et al. 2003; Nesterov 2007). The main advantage of these techniques is that the measured force is traceable to the Systeme International d'Unites (SI units). The disadvantage is that the necessity for additional instrumentation, and especially in the case of the NFC, environmental factors such as acoustic noise, thermal fluctuations, and air flow fluctuations will cause measurement inaccuracies.

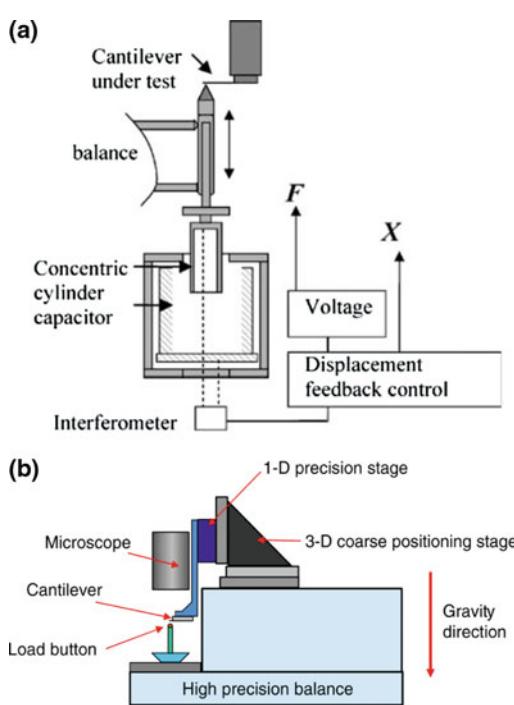


Fig. 3.6 Schematics of the **a** electrostatic force balance (adapted from Shaw et al. 2007), and **b** nano force calibrator (Kim et al. 2006)

3.3.2 Dynamic Methods

3.3.2.1 Added Mass Method

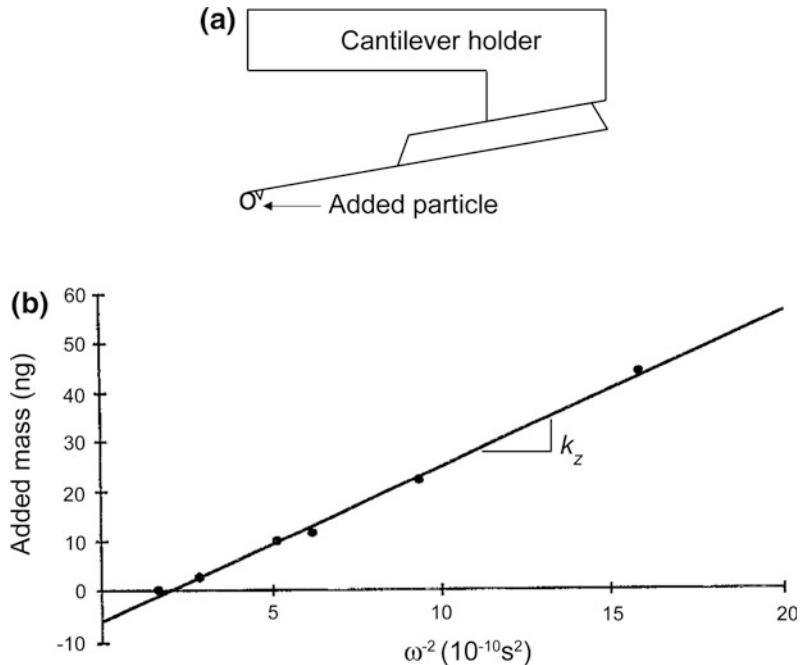
The added mass method was proposed by Cleveland et al. (1993) and is also referred to as the “frequency scaling” technique. In this method, the resonance frequencies of cantilevers are measured before and after the addition of small masses (such as tungsten particles with diameter 7–16 μm and mass 2.8–44 ng) at the tip of the cantilever. A schematic is shown in Fig. 3.7a.

The normal stiffness of the cantilever can be obtained from its effective mass and angular resonance frequency. By using Eqs. 3.3 and 3.11, and with $m_{\text{eff}} = m^*$, the angular resonance frequency can be expressed as

$$\omega_0 \approx \frac{h}{2L^2} \left(\frac{E}{0.2427\rho_c} \right)^{1/2} \approx \frac{h}{L^2} \left(\frac{E}{\rho_c} \right)^{1/2} \quad (3.17)$$

Equation 3.17 and the approximation for the effective mass given above both require the dimensions and material properties of the cantilever, which as described earlier are susceptible

Fig. 3.7 **a** Schematic of the side view of a cantilever with an added mass, and **b** plot of added mass versus ω^{-2} for a single cantilever, where the slope gives the stiffness (adapted from Cleveland et al. 1993)



to measurement errors. In addition, the resonance frequency defined in Eq. 3.17 is for vacuum, not in air (A correction that accounts for the surrounding medium is discussed in Sect. 3.3.2.2). Therefore, there is a need to determine the resonance frequency and effective mass directly from the measurements in order to obtain k_z more accurately.

In the added mass method, a spherical particle with mass m_s is added to the cantilever, such that the effective mass is $m_{\text{eff}} = m^* + m_s$. The angular resonance frequency ω changes correspondingly as a result of particle addition ($\omega \neq \omega_0$). The stiffness expression in Eq. 3.11 can be modified to account for the added mass as follows (Cleveland et al. 1993)

$$m_s = \frac{k_z}{\omega^2} - m^* \quad (3.18)$$

The resonance frequency is determined during the tuning procedure, where the cantilever is vibrated at a certain frequency range, usually between 10 and 400 kHz, and the frequency which produces maximum amplitude response is determined. By adding a number of different

masses to the cantilever and measuring the resonance frequency at each time, a plot of m_s versus ω^{-2} yields a straight line with a slope of k_z and a y-intercept of m^* , as shown in Fig. 3.7b.

In the method described in the previous paragraph, multiple masses are used, and k_z is obtained from the line fit. One can also obtain k_z using a single mass. In this case, k_z can be solved for by expressing Eq. 3.11 with and without the added mass, such that

$$k_z = \frac{m_s}{(1/\omega^2 - 1/\omega_0^2)} \quad (3.19)$$

The data obtained with the added mass method was found to follow a scaling law where the stiffness is proportional to the cube of the unloaded resonance frequency

$$k_z = b(\omega_0 L \sqrt{\rho_c})^3 / 4\sqrt{E} \quad (3.20)$$

where the length L , width w , elastic modulus E , and density ρ_c are known. The stiffness can also be evaluated with Eq. 3.20; however, the use of cantilever dimensions introduces error.

Therefore, the result is not as precise as that obtained by the addition of various end masses (Eq. 3.18).

A correction could be made to account for the distance that the added particle is offset from the end of the cantilever by d (Fig. 3.2). As an example, for a rectangular beam, m_s can be corrected by Sader et al. (1995)

$$m_s = m_{s,off} \left(\frac{L}{L+d} \right)^3 \quad (3.21)$$

where $m_{s,off}$ is the effective mass of the particle measured with the offset, which can then be used in either Eq. 3.18 or 3.19 in place of m_s .

This method is time-intensive and has the risk of damaging the cantilever during the addition of the particles. In practice, the difficulty of positioning the particle on the cantilever is also an issue (Sader et al. 1995; Ohler 2010). The measured stiffness has an uncertainty of approximately 15–30%. The variation occurs mainly from errors in measuring the diameter of the added particle. The particles may not be a perfect sphere, so the volume should be calculated as an ellipsoid and careful measurements must be taken with the SEM. Another source of error is the mass of the cantilever, as its dimensions and density are used in the calculations.

3.3.2.2 Resonance Method

Sader et al. (1995) suggested calculating k_z from the measurement of cantilever resonance frequency in vacuum (in the absence of a fluid medium) ω_{vac} , using Eq. 3.11 with $m_{eff} = m^*$.

As stated earlier, the resonance frequency can be obtained during the cantilever tuning procedure. This approach poses some limitations. Measurement of cantilever thickness, density and mass are sources of error. Moreover, this approach assumes a cantilever frequency in vacuum and does not account for the surrounding medium (air or liquid). Sader (1998) proposed an approach which accounts for the fluid medium and eliminates the need for measuring the density of the cantilever. The resonance frequency for a rectangular cantilever in vacuum is related to the resonance frequency in fluid ω_f by

$$\omega_{vac} = \omega_f \left(1 + \frac{\pi \rho_f b}{4 \rho_c h} \Gamma_r(\omega_f) \right)^{1/2} \quad (3.22)$$

The cantilever density is given by the following

$$\rho_c h = \frac{\pi \rho_f}{4} [Q\Gamma_i(\omega_f) - \Gamma_r(\omega_f)] \quad (3.23)$$

In Eqs. 3.22–3.23, ρ_f is the density of the fluid medium, and Γ_r and Γ_i are the real and imaginary components of the hydrodynamic function, Γ , which, in turn, depends on the Reynolds number $Re = \rho_f \omega b^2 / 4\eta$. The hydrodynamic function is obtained using an analytical expression applicable to rectangular beams. The quantity η is the viscosity of the fluid medium, which is independent of cantilever thickness and density. Substituting Eqs. 3.22 and 3.23 into 3.11, the stiffness is given by Sader et al. (1999)

$$k_z = 0.1906 \rho_f b^2 L Q \omega_f^2 \Gamma_i(\omega_f) \quad (3.24)$$

This technique is not prone to the experimental uncertainties inherent in determining thickness and density of the cantilever, but it still requires knowledge of its width and length. In addition, the model requires knowledge of the Reynolds number of the fluid medium, which is a disadvantage as it could change a lot depending on the elevation and must therefore be determined accurately. The hydrodynamic function calculation is not straightforward and can only be solved for by approximate analytical solutions such as the one by Sader (1998). However, for relatively stiff cantilevers (where the expected $k_z > 1$ N/m), this method is highly desirable. Since measurement of the resonance frequency is highly accurate (average error $\sim 0.1\%$), the uncertainty of the resulting k_z is low, $\sim 4\%$ (Ohler 2010).

It must be noted that the resonance method discussed here is based on the geometry of a rectangular cantilever beam. However, the analytical solution to the hydrodynamic function needs to be extended to the geometry of triangular cantilevers. The normal stiffness obtained from the resonance method on a rectangular

cantilever can be used to calibrate a triangular cantilever as long as both rectangular and triangular cantilevers are present on a single wafer. The assumption is that the rigidity Eh^3 of the rectangular and triangular cantilevers is the same since they are on the same wafer (Sader et al. 1995). For the rectangular cantilever, Eh^3 is solved for by using Eq. 3.3, with b and L being known and k_z obtained from the resonance method. Afterwards, this value for Eh^3 is used on the triangular cantilever by using the appropriate expression for the normal stiffness, provided that the length, width, and included angle of the triangular cantilever are known. Wafers containing rectangular and triangular cantilevers are not widely available. Most commercial cantilevers are fabricated such that the wafer only contains either rectangular or triangular cantilevers, not both.

More recently, Sader et al. (2005) proposed a general expression that enables one to determine the stiffness for any geometry of the cantilever. The stiffness is derived based on its relationship to the energy dissipation of the cantilever during oscillation. Through dimensional analysis, the stiffness is related to the density of the fluid, the cantilever length, the Q factor, resonance frequency, and the Reynolds number of the fluid as follows

$$k_z = \rho_f L^3 \Omega(\text{Re}) Q \omega_f^2 \quad (3.25)$$

The dimensionless function $\Omega(\text{Re})$ is defined depending on the geometry. For a rectangular cantilever, an analytical solution is known such that $\Omega(\text{Re})$ can be expressed in terms of the hydrodynamic function $\Gamma(\omega_f)$, which leads to the expression in Eq. 3.24. For a triangular cantilever, an analytical solution is not available and $\Omega(\text{Re})$ has to be determined by performing experiments in a fluid of interest to measure the resonance frequency and Q factor as a function of fluid pressure. Then, an expression for the normal stiffness is obtained from a curvefit.

The resonance method can also be implemented using a Laser Doppler vibrometer (LDV) instead of an AFM (Ohler 2007). This provides independent validation of the results

obtained from AFM. However, the disadvantage of LDV is that this requires additional instrumentation aside from the AFM, which might not be easily accessible.

3.3.2.3 Thermal Noise Method

A harmonic oscillator in equilibrium with its surroundings will fluctuate in response to thermal noise (Hutter and Bechhoefer 1993). As shown in Fig. 3.8a, the AFM tip-sample system is modeled as a spherical tip held at a distance from the sample surface by a spring (representing the cantilever). The spring is assumed to behave like a simple harmonic oscillator. The normal stiffness of the AFM cantilever can be related to its thermal energy during its vibration by the equipartition theorem, leading to the following relationship

$$\frac{1}{2} k_z \langle z^2 \rangle = \frac{1}{2} k_B T \quad (3.26)$$

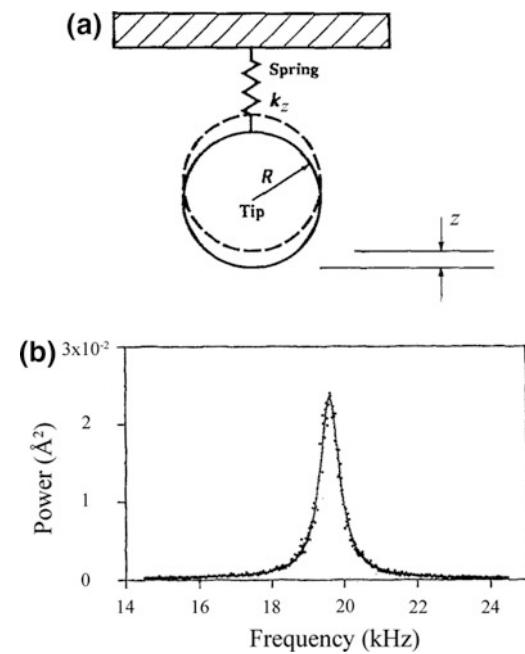


Fig. 3.8 a Illustration of the principle behind the thermal noise method, and b power spectral density plot of the cantilever deflection fluctuations (adapted from Hutter and Bechhoefer 1993)

where k_B is the Boltzmann constant, T is the temperature, and $\langle z^2 \rangle$ is the mean square deflection of the cantilever, which fluctuates due to thermal noise.

An example of the power spectrum of cantilever deflection fluctuations is shown in Fig. 3.8b. From experimental data, p , the area of the power spectrum of the thermal fluctuations is equal to $\langle z^2 \rangle$, so the stiffness can be expressed as (Hutter and Bechhoefer 1993)

$$k_z = k_B T / p \quad (3.27)$$

The thermal noise method (also referred to as the thermal tune method) for determining the normal stiffness is widely regarded as being less prone to experimental uncertainties, as it eliminates the need for the dimensions and the mechanical properties of the cantilever. However, a disadvantage of this method is that cantilevers do not behave perfectly like simple harmonic oscillators, such that Eq. 3.27 is only an approximation. Butt and Jaschke (1995) proposed a correction which accounts for the bending shapes for each vibration mode. Moreover, they accounted for the additional error arising from the fact that the inclination of the cantilever ($dz(L)/dx$) is measured instead of true displacement. They proposed the following equation (Butt and Jaschke 1995)

$$k_z = 0.817 k_B T / p \quad (3.28)$$

Hutter (2005) recognized that another correction is necessary since the cantilever is mounted at an angle γ relative to the horizontal axis, leading to this equation

$$k_z = 0.817 k_B T \cos^2 \gamma / p \quad (3.29)$$

The thermal noise technique is accurate and relatively simple to perform. However, for some AFM systems, implementing this technique requires additional instrumentation, such as a spectrometer or lock-in amplifier to collect the thermal noise data. One limitation is that this method is most suitable for calibrating soft cantilevers (where the expected $k_z < 1$ N/m) where

the thermal noise is higher than the noise from the deflection measurement (Ohler 2010).

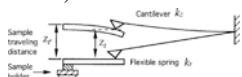
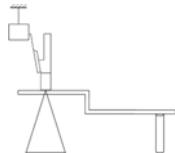
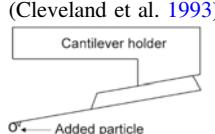
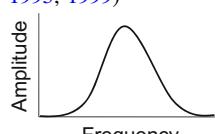
The thermal noise technique can also be implemented using a Laser Doppler vibrometer (LDV) instead of an AFM (Ohler 2007). The advantage and disadvantage of LDV as a complementary technique has been discussed above (Sect. 3.3.2.2).

3.3.3 Discussion

The main highlights of the normal stiffness calibration methods discussed here are shown in Table 3.1. Three static measurement techniques were reviewed, namely, the reference, inverted cantilever, and pendulum methods. The reference cantilever method is simple and straightforward. Its only limitation is that it requires accurate positioning of the cantilever relative to the reference during the experiment in order to obtain accurate results. In the inverted loaded cantilever experiment, the addition of a sphere presents the risk of damaging the cantilever, rendering it unsuitable for further use. In addition, the necessity for calibrating the observed deflection signal corresponding to cantilever motion further complicates the experiment. The pendulum method is based on a concept similar to the inverted cantilever technique in the sense that it is also a gravity-based experiment. In this case, the mass of the pendulum is required, and this quantity can be accurately determined. However, calibrating the pendulum deflection is needed, which could be a source of measurement uncertainty. The Electrostatic Force Balance and the Nano Force Calibrator are examples of experimental normal stiffness calibration techniques where the force application is traceable to SI units. However, these setups need to be built and additional instrumentation is necessary.

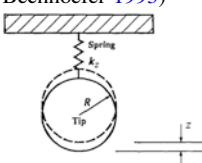
The dynamic experiments discussed above are the added mass, resonance, and thermal noise methods. Similar to the static inverted cantilever method, the added mass method involves adding a particle to the cantilever, which is time-consuming and has the risk of damaging the cantilever if the particle is added improperly. In

Table 3.1 Summary of techniques for calibrating the normal stiffness

Technique	Principle/key equation	Pros	Cons
<i>Static methods</i>			
Reference cantilever (Ruan and Bhushan 1994a)	A calibrated reference is used to calibrate probes $k_z = k_s(Z'_t - Z_t)/Z_t$ 	Simple; easy to implement	Requires accurate positioning of cantilever relative to reference
Inverted loaded cantilever (Senden and Ducker 1994)	A particle is attached to cantilever and the deflection of the inverted cantilever is measured $K_z = \frac{8\pi r^3 \rho_s g}{3z}$ 	–	Requires attachment of a particle to cantilever; risk of cantilever damage
Pendulum (Butt et al. 1993)	Tip is pushed against a pendulum, displacements of tip and pendulum are measured $k_z = \frac{m_p g z_p}{l_p \zeta_c}$ 	–	Requires pendulum set-up; risk of cantilever damage; calibration of pendulum needed
Electrostatic force balance (Pratt et al. 2005)	Electrostatic force is applied to the cantilever and displacement is measured by an interferometer	–	Additional instrumentation is required
Nano force calibrator (Kim et al. 2006)	Precision stage controls the displacement while a balance measures the force	–	Additional instrumentation is required
<i>Dynamic methods</i>			
Added mass (Cleveland et al. 1993)	A number of particles are individually attached to the end of the cantilever and the resonance frequency is measured $m_s = \frac{k_z}{\omega^2} - m^*$ 	–	Requires attachment of particles to cantilever; risk of cantilever damage; uncertainty in dimensions and errors in placement of added particle can lead to significant error
Resonance (Sader et al. 1995, 1999)	The resonance frequency, quality factor and dimensions are measured For a rectangular beam, $k_z = 0.1906 \rho_f b^2 L Q \omega^2 \Gamma_i(\omega)$ 	Relatively simple; desirable for stiff cantilevers	Method relies on cantilever dimensions; requires Reynolds number of the fluid medium and calculation of the hydrodynamic function

(continued)

Table 3.1 (continued)

Technique	Principle/key equation	Pros	Cons
Thermal noise (Hutter and Bechhoefer 1993) 	Thermal fluctuation of the cantilever is measured $k_z = k_B T/l_p$	Accurate; relatively simple; desirable for soft cantilevers	Potential inaccuracy when applied to stiff cantilevers

addition, uncertainties in the dimensions of the cantilever and errors in placement of the particle lead to significant error. The resonance method has been shown to be accurate for stiffer cantilevers, i.e. normal stiffness higher than 1 N/m. It requires measurement of the cantilever width, which gives the result some uncertainty. Also, since the model is derived for a rectangular cantilever beam, there is a need to use the extended theoretical treatment so that it can be applied to triangular cantilevers as well.

The thermal noise method does not require any parameters related to the cantilever beam dimension, which makes it accurate. The precision of reproducibility of the obtained stiffness using the thermal noise method has been reported by Hutter and Bechhoefer (1993) and Matei et al. (2006) to be as high as 5%. However, this method is mostly limited for calibrating soft cantilevers where the thermal noise is higher than the noise from the deflection measurement. Depending on the AFM system, the technique may require additional instrumentation such as a spectrometer or lock-in amplifier (which are commonly available instruments) in order to obtain the thermal noise spectra.

The tilt of the AFM cantilever affects the normal stiffness, and a correction could be necessary depending on the calibration method used. The added mass and resonance methods yield the intrinsic stiffness, i.e., the value that one would obtain assuming no tilt, so results from these methods should be modified by using Eq. 3.4 (introduced in Sect. 3.2). This correction is not necessary for the reference cantilever method since the “effective” stiffness is obtained in the measurement (Ohler 2010).

Selecting a technique for calibrating the normal stiffness depends on both the available instrumentation and the user’s experience in data analysis. For high accuracy measurements, the resonance and thermal noise methods are the most preferable for stiff and soft cantilevers, respectively. A simpler technique such as the reference cantilever method is highly recommended for calibrating the normal stiffness as well.

3.4 Lateral Force and Stiffness Calibration Techniques

Static and dynamic methods for calibrating the lateral force and stiffness are reviewed in this section. Static methods require either bending or torsion of the cantilever, while dynamic methods involve the determination of the torsional vibration characteristics of the cantilever. We describe various methods for measuring lateral forces by using a direct method for measuring the coefficient of friction, where the lateral force is calculated from (Sects. 3.4.1.1 to 3.4.1.2). We also describe methods for the determination of either the torsional or torsional and lateral stiffness (Sects. 3.4.1.3, 3.4.1.4, 3.4.2.1 and 3.4.2.2). A discussion on calculating the lateral force from the torsional stiffness is given in Sect. 3.4.2.3.

3.4.1 Static Methods

3.4.1.1 Axial Sliding Method

Based on the work by Ruan and Bhushan (1994a), the axial friction measurement method

is described. A scanning angle is defined as the angle relative to the x-axis in Fig. 3.9a. This is also the long axis of the cantilever. A 0° scanning angle corresponds to the sample scanning in the x direction, and a 90° scanning angle corresponds to the sample scanning perpendicular to this axis in the xy plane (in y axis). If the scanning direction is in both x and $-x$ directions, this is called a “parallel scan.” Similarly, a “perpendicular scan” means scanning is done in the y and $-y$ directions. The sample traveling direction for each of these two scanning directions is illustrated in Fig. 3.9b. Parallel scanning is discussed as “method 1,” where the coefficient of

friction is obtained. Perpendicular scanning is described below as “method 2,” where lateral forces are determined based on the measured signal due to sliding in the lateral direction. In order to convert the measured lateral signal to a force, a conversion factor is calculated based on the coefficient of friction determined using method 1. This is then used to obtain a friction force in three dimensions.

In method 1, aside from topographic imaging, it is also possible to measure friction force. If no friction force existed between the tip and the moving sample, the topographic feature would be the only factor which causes the cantilever to

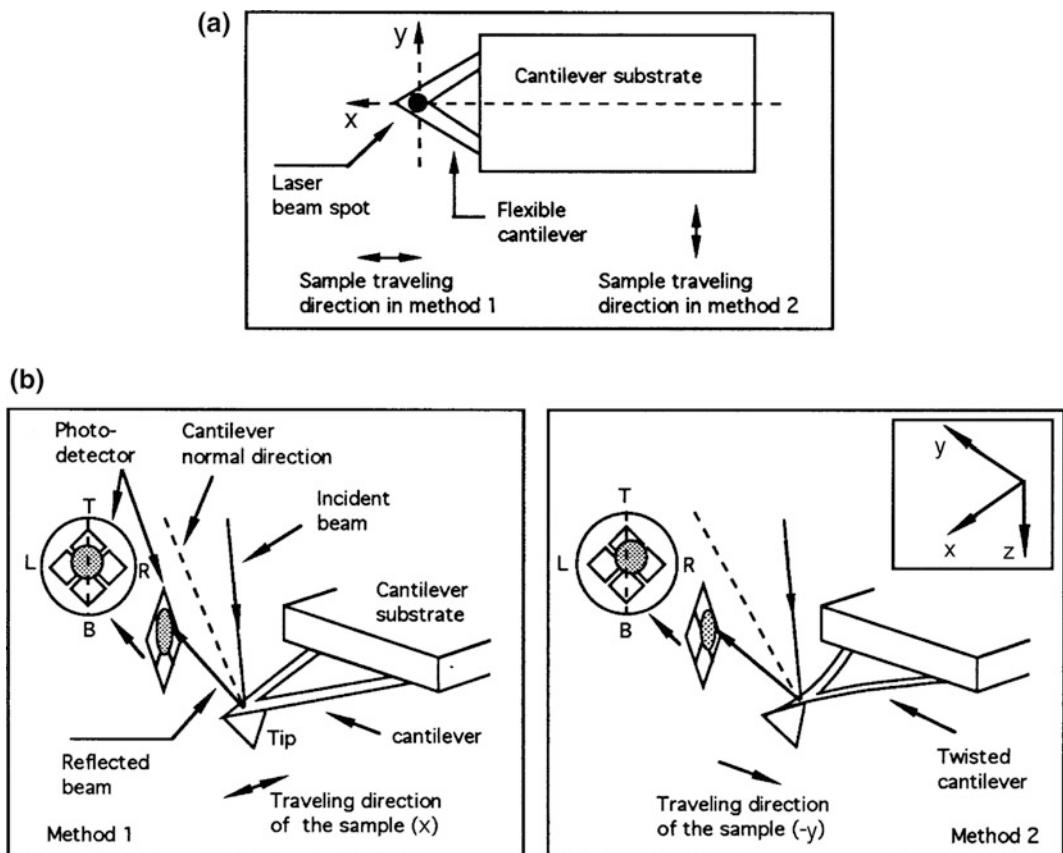


Fig. 3.9 a Schematic defining the x- and y-directions relative to the cantilever, and showing the sample traveling direction in two different measurement methods discussed in the text, and b schematic of deformation of the tip and cantilever shown as a result of sliding in the x- and y-directions. A twist is introduced to the cantilever if the scanning is in the y-direction [(b), right] (Ruan and Bhushan 1994a)

be deflected vertically. However, friction force does exist on all contact surfaces where one object is moving relative to another. The friction force between the sample and the tip will also cause cantilever deflection. We assume that the normal force between the sample and the tip is W when the sample is stationary (W is typically in the range of 1–200 nN), and the friction force (lateral force) between the sample and the tip is F as the sample scans against the tip (Fig. 3.10). The direction of friction force is reversed as the scanning direction of the sample is reversed from positive (x) to negative ($-x$) directions, i.e., $\vec{F}_x = -\vec{F}_{-x}$.

When the vertical cantilever deflection is set at a constant level, it is both normal and friction forces applied to the cantilever that keeps the cantilever deflection at this level. Since the friction force is in opposite directions as the traveling direction of the sample is reversed, the normal force is adjusted accordingly when the sample reverses its traveling direction, so that the cantilever deflection remains the same. In order to maintain constant deflection, the bending of the cantilever (left side of Fig. 3.10) is canceled by adjusting the piezotube height by a feedback circuit (right side of Fig. 3.10). The observed cantilever deflection is the contribution of two bending moments. The first contribution is from the friction force, and the second is from the normal force. Based on the basic definition for the angular deflection $\psi = \int \frac{M}{EI} dx$ (Timoshenko and Goodier 1970; Young and Budynas 2002), the angular deflections due to the friction force and the normal load, ψ_F and ψ_W , respectively, will be calculated. It will be further assumed that the total angular deflection due to the friction and normal force is constant in either sliding direction when the cantilever deflection is kept constant. Friction force may have a directionality effect, and forces in the forward and reverse directions may be slightly different.

We first consider a simple case where the cantilever is not tilted, as illustrated in Fig. 3.10a. The angular deflection contributed by the moment due to the friction force (ψ_F) is given by

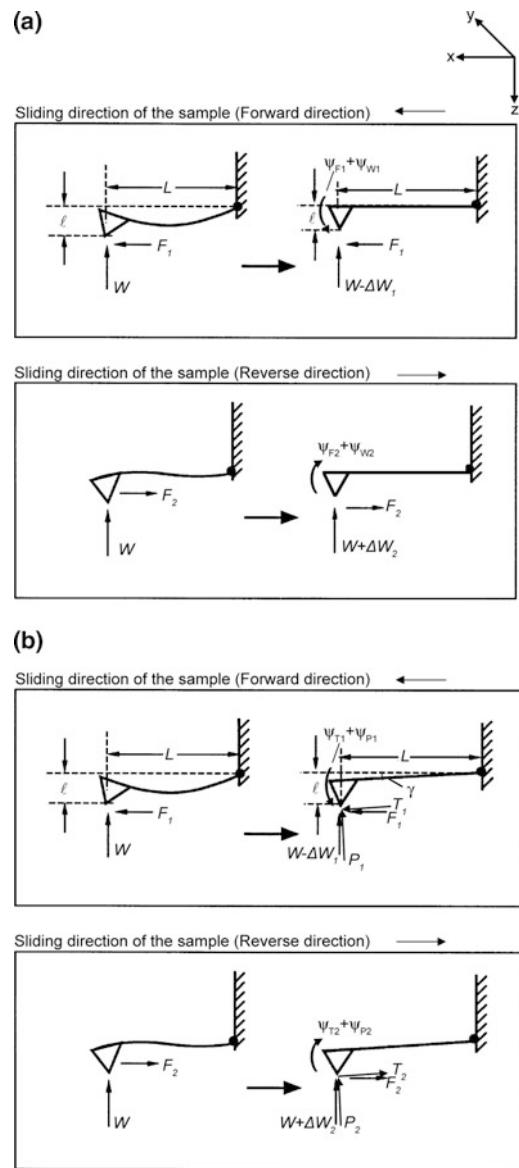


Fig. 3.10 Schematic showing an additional bending of the cantilever due to friction force when the sample is scanned in the x or $-x$ direction (left), and this effect will be canceled by adjusting the piezo height by a feedback circuit (right) for a horizontal tip, and **b** tip tilted at an angle γ

$$\begin{aligned}\psi_{F1} &= \frac{12L}{Ebh^3} F_1 \ell \\ \psi_{F2} &= -\frac{12L}{Ebh^3} F_2 \ell\end{aligned}\quad (3.30)$$

in the forward (subscript “1”) and reverse (subscript “2”) sliding directions, respectively. The angular deflection contributed by the moment due to the normal force (ψ_W) is

$$\begin{aligned}\psi_{W1} &= \frac{6L}{Ebh^3}(W - \Delta W_1)L \\ \psi_{W2} &= \frac{6L}{Ebh^3}(W + \Delta W_2)L\end{aligned}\quad (3.31)$$

The total angular deflection, which is the sum of the friction and normal force contributions, is given by

$$\begin{aligned}\psi_{F1} + \psi_{W1} &= \frac{12L}{Ebh^3}F_1\ell + \frac{6L}{Ebh^3}(W - \Delta W_1)L \\ \psi_{F2} + \psi_{W2} &= -\frac{12L}{Ebh^3}F_2\ell + \frac{6L}{Ebh^3}(W + \Delta W_2)L\end{aligned}\quad (3.32)$$

As stated earlier, $\psi_{F1} + \psi_{W1} = \psi_{F2} + \psi_{W2}$. Simplifying the resulting equation, and rearranging to get an expression for the average friction force $(F_1 + F_2)/2$, leads to

$$\frac{F_1 + F_2}{2} = (\Delta W_1 + \Delta W_2) \frac{L}{4\ell} \quad (3.33)$$

The coefficient of friction can be calculated using Eq. 3.33 as

$$\mu = \frac{F_1 + F_2}{2W} = \left[\frac{(\Delta W_1 + \Delta W_2)}{W} \right] \frac{L}{4\ell} \quad (3.34)$$

In all circumstances, there are adhesive and interatomic attractive forces between the cantilever tip and the sample. The adhesive force can be due to water from capillary condensation and other contaminants present at the surface which form meniscus bridges and the interatomic attractive force includes van der Waals attraction (Bhushan 2011, 2017). There is an indentation effect as well, which is usually small for rigid samples. If these forces can be neglected, the normal force W is then equal to the initial cantilever deflection H_0 multiplied by the stiffness of the cantilever. $(\Delta W_1 + \Delta W_2)$ can be measured by multiplying the same stiffness by the height difference of the piezo tube between the two

traveling directions (forward and reverse) of the sample. This height difference is denoted as $(\Delta H_1 + \Delta H_2)$, shown schematically in Fig. 3.11 a. Thus, Eq. 3.34 can be rewritten as

$$\mu = \frac{F_1 + F_2}{2W} = \left[\frac{(\Delta H_1 + \Delta H_2)}{H_0} \right] \left(\frac{L}{4\ell} \right) \quad (3.35)$$

Since the piezo tube vertical position is affected by the surface topographic profile of the sample in addition to the friction force being applied at the tip, this difference has to be taken point by point at the same location on the sample surface as shown in Fig. 3.11a. Subtraction of point by point measurements may introduce errors, particularly for rough samples. In addition, precise measurements of L and ℓ (which should include the cantilever angle) are also required. Since only the ratio between $(\Delta H_1 + \Delta H_2)$ and H_0 comes into Eq. 3.35, the piezo tube vertical position H_0 and its position difference $(\Delta H_1 + \Delta H_2)$ can be in the units of volts as long as the vertical traveling distance of the piezo tube and the voltage applied to it has a linear relationship. However, if there is a large nonlinearity between the piezo tube traveling distance and the applied voltage, this nonlinearity must be included in the calculation (Ruan and Bhushan 1994a).

If the adhesive forces between the tip and the sample are large enough that it cannot be neglected, one should include it in the calculation. However, there could be a large uncertainty in determining this force, and thus an uncertainty in using Eq. 3.35. An alternative approach is to measure the height difference of the PZT $(\Delta H_1 + \Delta H_2)$ at different normal loads (H_0) and to use the slope of $(\Delta H_1 + \Delta H_2)$ versus the slope of (H_0) from the measurements in Eq. 3.35 to calculate μ . Figure 3.11b shows the data from three sets of measurements at various loads on a Pt sample, where it is seen that a linear fit is obtained. The coefficient of friction for this sample was found to be 0.054 (Ruan and Bhushan 1994a).

Now we consider the cantilever tilt angle γ relative to the horizontal axis, as illustrated in Fig. 3.10b. In this case, the friction force (T) and normal load (P) components are resolved in terms

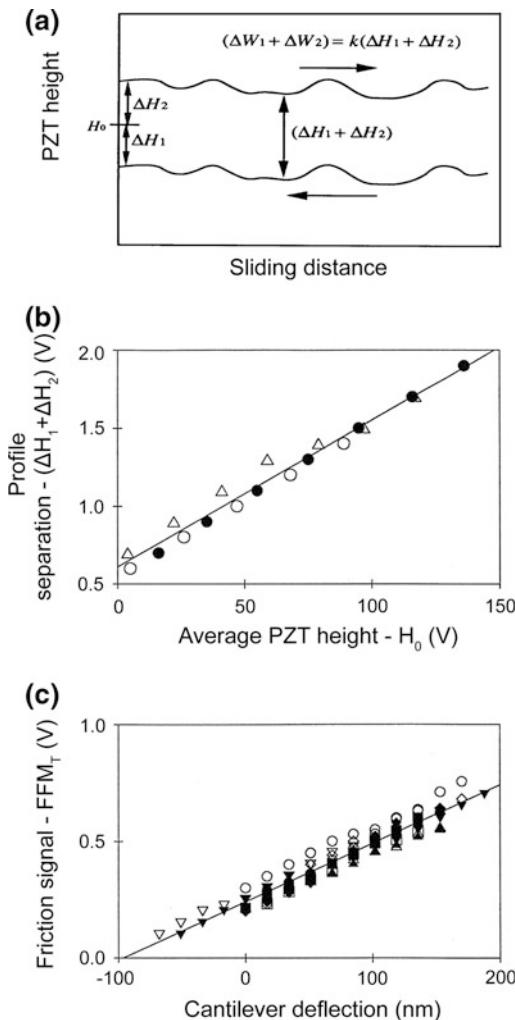


Fig. 3.11 **a** Schematic of the height difference of the piezoelectric tube scanner as the sample is scanned in y and $-y$ directions, **b** the vertical height difference as a function of the PZT center position between the two sliding directions on a Pt sample (method 1). The three symbols represent three sets of repeated measurements. The slope of the linear fit is proportional to the coefficient of friction between the Si_3N_4 tip and Pt. **c** Friction signal as a function of cantilever vertical deflection for Pt (method 2). Different symbols represent eleven sets of repeated measurements. The slope of the linear fit is proportional to the coefficient of friction between the Si_3N_4 tip and Pt (Ruan and Bhushan 1994a)

of the measured horizontal force (F) and normal load (W) along the x and z axes (Wang et al. 2007).

The resolved horizontal and normal force components along x and z axes in the forward sliding direction are given by

$$\begin{aligned} T_1 &= F_1 \cos \gamma - (W - \Delta W_1) \sin \gamma \\ P_1 &= F_1 \sin \gamma + (W - \Delta W_1) \cos \gamma \end{aligned} \quad (3.36)$$

and in the reverse direction by

$$\begin{aligned} T_2 &= F_2 \cos \gamma + (W + \Delta W_2) \sin \gamma \\ P_2 &= -F_2 \sin \gamma + (W + \Delta W_2) \cos \gamma \end{aligned} \quad (3.37)$$

Using Eq. 3.30, the angular deflection contributed by the force along the longitudinal axis of the cantilever is given by

$$\begin{aligned} \psi_{T1} &= \frac{12L}{Ebh^3} (F_1 \cos \gamma - (W - \Delta W_1) \sin \gamma) \ell \\ \psi_{T2} &= -\frac{12L}{Ebh^3} (F_2 \cos \gamma + (W + \Delta W_2) \sin \gamma) \ell \end{aligned} \quad (3.38)$$

Based on Eq. 3.31, the angular deflection contributed by the normal load perpendicular to the longitudinal axis of the cantilever is given by

$$\begin{aligned} \psi_{P1} &= \frac{6L}{Ebh^3} (F_1 \sin \gamma + (W - \Delta W_1) \cos \gamma) L \\ \psi_{P2} &= \frac{6L}{Ebh^3} (-F_2 \sin \gamma + (W + \Delta W_2) \cos \gamma) L \end{aligned} \quad (3.39)$$

The total angular deflection, obtained by adding the friction force and normal load contributions is given by

$$\begin{aligned} \psi_{T1} + \psi_{P1} &= \frac{12L\ell}{Ebh^3} (\mu(W - \Delta W_1) \cos \gamma - (W - \Delta W_1) \sin \gamma) \\ &\quad + \frac{6L^2}{Ebh^3} (\mu(W - \Delta W_1) \sin \gamma + (W - \Delta W_1) \cos \gamma) \end{aligned} \quad (3.40)$$

$$\begin{aligned} \psi_{T2} + \psi_{P2} &= -\frac{12L\ell}{Ebh^3} (\mu(W + \Delta W_2) \cos \gamma + (W + \Delta W_2) \sin \gamma) \\ &\quad + \frac{6L^2}{Ebh^3} (-\mu(W + \Delta W_2) \sin \gamma + (W + \Delta W_2) \cos \gamma) \end{aligned}$$

where μ is defined as $F_1/(W - \Delta W_1) = F_2/(W + \Delta W_2)$. Since it is assumed that $\psi_{T1} + \psi_{P1} = \psi_{T2} + \psi_{P2}$, by simplifying the resulting equation, the expression for μ is obtained as follows

$$\begin{aligned}\mu &= \frac{\Delta W_1 + \Delta W_2}{2W - \Delta W_1 + \Delta W_2} \frac{\left(-\frac{12L\ell}{Ebh^3} \sin \gamma + \frac{6L^2}{Ebh^3} \cos \gamma\right)}{\left(\frac{12L\ell}{Ebh^3} \cos \gamma + \frac{6L^2}{Ebh^3} \sin \gamma\right)} \\ &= \frac{\Delta W_1 + \Delta W_2}{2W - \Delta W_1 + \Delta W_2} \frac{(-2\ell \sin \gamma + L \cos \gamma)}{(2\ell \cos \gamma + L \sin \gamma)}\end{aligned}\quad (3.41)$$

This can be simplified by expressing $(P_1 + P_2)$ and $(P_1 - P_2)$ in terms of the height difference of the piezotube.

$$\begin{aligned}\mu &= \frac{\Delta H_1 + \Delta H_2}{2H_o - \Delta H_1 + \Delta H_2} \frac{(-2\ell \sin \gamma + L \cos \gamma)}{(2\ell \cos \gamma + L \sin \gamma)} \\ &\approx \frac{\Delta H_1 + \Delta H_2}{2H_o} \frac{(-2\ell \sin \gamma + L \cos \gamma)}{(2\ell \cos \gamma + L \sin \gamma)}\end{aligned}\quad (3.42)$$

For $\gamma = 0$, Eq. 3.41 reduces to Eq. 3.34, and Eq. 3.42 reduces to Eq. 3.35.

Method 2 (“aux” mode with perpendicular scan) is used to obtain 3D maps of friction. The sample is scanned perpendicular to the long axis of the cantilever beam (i.e., to scan along the y or $-y$ direction in Fig. 3.9a), and the output of the two horizontal quadrants of the photodiode-detector is measured. In this arrangement, as the sample moves under the tip, the friction force will cause the cantilever to twist. Therefore, the light intensity between the left and right (L and R in Fig. 3.9b, right) detectors will be different. The differential signal between the left and right detectors is denoted as the friction force microscopy or FFM signal $[(L - R)/(L + R)]$. This signal can be related to the degree of twisting, hence to the magnitude of friction force. Again, because of a possible error in determining normal force due to the presence of an adhesive force at the tip-sample interface, the slope of the friction data (FFM signal vs. normal load) needs to be taken for an accurate value of coefficient of friction.

While friction force contributes to the FFM signal, it may not be the only contributing factor in commercial FFM instruments. One can notice this by engaging the cantilever tip with the

sample. Before engaging, the left and right detectors can be balanced by adjusting the position of the detectors so that the intensity difference between these two detectors is zero (FFM signal is zero). Once the tip is engaged with the sample, this signal is no longer zero even if the sample is not moving in the xy plane with no friction force applied. This would be a detrimental effect. It has to be understood and eliminated from the data acquisition before any quantitative measurement of friction force becomes possible.

When the laser beam is reflected up and down due to a change of the normal force, the intensity difference between the left and right detectors will also change (Ruan and Bhushan 1994a). In other words, the FFM signal will change as the normal force applied to the tip is changed, even if the tip is not experiencing any friction force. This FFM signal is unrelated to friction force or to the actual twisting of the cantilever, but rather to cross-talk effects, which can either be optical or mechanical in nature. We will call this part of the FFM signal “ FFM_F ,” and the part which is truly related to friction force “ FFM_T .” The FFM_F signal can be eliminated. One way of doing this is as follows. First the sample is scanned in both y and $-y$ directions, and the FFM signal for scans in each direction is recorded. Since friction force reverses its direction when the scanning direction is reversed from y to $-y$ direction, the FFM_T signal will have opposite signs as the scanning direction of the sample is reversed ($FFM_T(y) = -FFM_T(-y)$). Hence the FFM_T signal will be canceled out if we take the sum of the FFM signals for the two scans. The average value of the two scans will be related to FFM_F due to the misalignment,

$$FFM(y) + FFM(-y) = 2FFM_F \quad (3.43)$$

This value can therefore be subtracted from the original FFM signals of each of these two scans to obtain the true FFM signal (FFM_T). Alternately, by taking the difference of the two FFM signals, one directly gets the FFM_T value

$$\begin{aligned} \text{FFM}(y) - \text{FFM}(-y) &= \text{FFM}_T(y) - \text{FFM}_T(-y) \\ &= 2\text{FFM}_T(y) \end{aligned} \quad (3.44)$$

Ruan and Bhushan (1994a) have shown that an error signal (FFM_F) can be very large compared to the friction signal FFM_T , thus correction is required. An implementation of this method is shown in Fig. 3.11c, where the true friction signal as a function of cantilever vertical deflection is shown for Pt. The coefficient of friction obtained using method 1 is then used to convert the friction force into nN.

One can measure L and ℓ for a full implementation of the axial sliding method (Eq. 3.35). Instead of doing this, we use the following procedure. It has been observed that the coefficient of friction for a Si_3N_4 or Si tip on single crystal silicon in ambient air is approximately 0.05. Since the coefficient of friction is an interface property, it is expected to be the same for Si_3N_4 or Si tips with varying dimensions (Tambe 2005). Hence, for a given Si_3N_4 or Si tip, measurements are first made on a single crystal silicon sample in a perpendicular scan, and the conversion factor for the lateral voltage signal is obtained based on a coefficient of friction of 0.05. This conversion factor is then used to analyze friction data on the test sample. The normal load is calculated from the normal stiffness and the photodetector sensitivity (in V/nm). Since both the normal and lateral forces are known, the coefficient of friction can be calculated.

The advantages of this technique are simplicity and ease of implementation. However, the need for measuring the cantilever length and tip height provides a source of measurement errors for this method.

3.4.1.2 Wedge Method

A schematic for the wedge method is shown in Fig. 3.12a. In the original work by Ogletree et al. (1996), the tip is scanned across a calibration sample containing triangular features with two well-defined slopes. This is based on the knowledge that when the tip is scanned across the sample surface, the measured friction forces are

generated by both material effects as well as topography-induced effects (Meyer and Amer 1990; Ruan and Bhushan 1994b, c; Koinkar and Bhushan 1997; Sundararajan and Bhushan 2000).

The approach for obtaining the coefficient of friction is as follows (Ogletree et al. 1996; Varenberg et al. 2003). At any given load, the friction, F , and normal, W , forces depend on the direction of motion, either uphill (denoted by subscript “u”) or downhill (denoted by subscript “d”) and can be resolved in terms of the applied load, P , and horizontal force, T , components. The total normal force includes the external applied force plus the intrinsic adhesive force, A . For uphill motion

$$\begin{aligned} W_u &= P \cos \theta + T_u \sin \theta + A_u \\ F_u &= -P \sin \theta + T_u \cos \theta \end{aligned} \quad (3.45)$$

For downhill motion

$$\begin{aligned} W_d &= P \cos \theta - T_d \sin \theta + A_d \\ F_d &= P \sin \theta + T_d \cos \theta \end{aligned} \quad (3.46)$$

It is assumed that the friction force is linearly related to the total normal force (Bhushan 1999, 2013), such that

$$\mu = \frac{F_u}{W_u} = \frac{F_d}{W_d} \quad (3.47)$$

Substituting Eqs. 3.45–3.46 into 3.47 and assuming that $A_u = A_d = A$ leads to

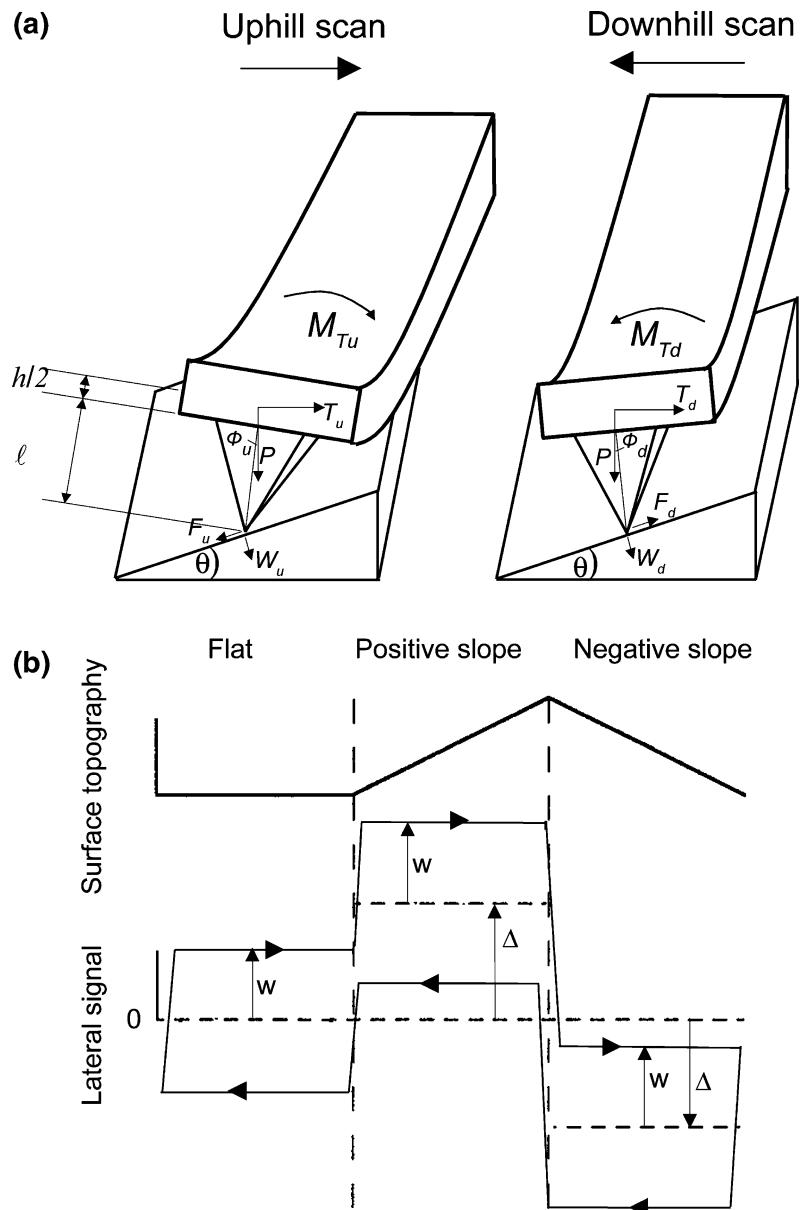
$$\begin{aligned} \mu &= \frac{-P \sin \theta + T_u \cos \theta}{P \cos \theta + T_u \sin \theta + A} \\ &= \frac{P \cos \theta - T_d \sin \theta}{P \cos \theta - T_d \sin \theta + A} \end{aligned} \quad (3.48)$$

In order to solve for μ , T_u and T_d will be related to experimentally measurable parameters.

Equations 3.45–3.46 can be rearranged into

$$\begin{aligned} T_u &= \frac{P \sin \theta + \mu P \cos \theta + \mu A}{\cos \theta - \mu \sin \theta} \\ T_d &= \frac{-P \sin \theta + \mu P \cos \theta + \mu A}{\cos \theta + \mu \sin \theta} \end{aligned} \quad (3.49)$$

Fig. 3.12 Schematics of (a) a cantilever torsion while sliding up and down on an inclined surface, the basis of the wedge method, and (b) friction loops for flat, positively sloped and negatively sloped surfaces



Furthermore, T_u and T_d can be related to the torsion moment M by

$$\begin{aligned} M_{Tu} &= T_u(l + h/2) \\ M_{Td} &= T_d(l + h/2) \end{aligned} \quad (3.50)$$

where ℓ and h pertain to the tip length and cantilever thickness, respectively (Fig. 3.2). Next,

the friction loops for the flat, positively sloped and negatively sloped surfaces are considered. Focusing on the two sloped regions shown in Fig. 3.12b, the half-width of the friction loop, w , and the friction loop offset, Δ , are measured experimentally. These two quantities are related to the torsion moment as follows

$$w = \frac{M_{Tu} - M_{Td}}{2} \quad (3.51)$$

$$\Delta = \frac{M_{Tu} + M_{Td}}{2} \quad (3.52)$$

Substituting Eqs. 3.49 and 3.50 into 3.51 and 3.52 yields the following

$$w = (\ell + h/2) \frac{\mu P + \mu A \cos \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta} \quad (3.53)$$

$$\Delta = (\ell + h/2) \frac{P \sin \theta \cos \theta + \mu^2 P \sin \theta \cos \theta + \mu^2 A \sin \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta} \quad (3.54)$$

The lateral forces vary with the applied load such that it is necessary to define the slopes $w' = dw/dP$ and $\Delta' = d\Delta/dP$, obtained by taking the first derivative of Eqs. 3.53 and 3.54

$$w' = \alpha_c w'_o = \frac{dw}{dP} = (\ell + h/2) \frac{\mu}{\cos^2 \theta - \mu^2 \sin^2 \theta} \quad (3.55)$$

$$\Delta' = \alpha_c \Delta'_o = \frac{d\Delta}{dP} = (\ell + h/2) \frac{(1 + \mu^2) \sin \theta \cos \theta}{\cos^2 \theta - \mu^2 \sin^2 \theta} \quad (3.56)$$

where w_o and Δ_o are the experimentally-determined half-width of the friction loop and the friction loop offset, respectively (both in volts), and α_c is the lateral force calibration factor (in N/V). In the limiting case of no friction, $w' \rightarrow 0$ and $\Delta' \rightarrow \tan \theta$, as expected. It should be noted that by taking the derivatives (Eqs. 3.55 and 3.56), the adhesive force is eliminated from the equations needed for determining μ . Finally, the expression for the coefficient of friction is obtained by dividing Eq. 3.56 by 3.55

$$\mu + \frac{1}{\mu} = \frac{2\Delta'}{w' \sin 2\theta} \quad (3.57)$$

After the coefficient of friction is determined, it can be used in either Eq. 3.50 or 3.51 to calculate α_c , and to obtain w and Δ in their proper

units. Two values of μ are obtained from the quadratic equation in Eq. 3.57, and both roots are considered to be equally good solutions. However, one of the roots could give a value for μ_s that is large enough such that the denominator in either Eq. 3.55 or 3.56 becomes negative (i.e., negative α_c). As this is artificial, then this spurious value is disregarded, leaving only one acceptable μ .

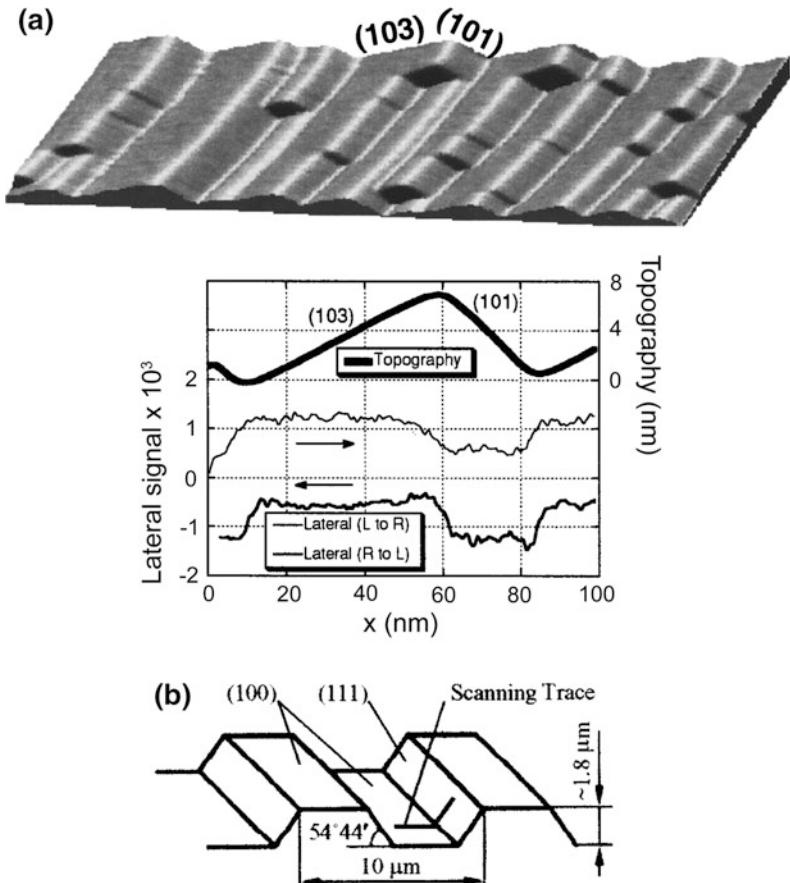
An image of the SrTiO₃ calibration sample with an inclination of 54°44' used by Ogletree et al. (1996), along with an example of the data they obtained using the wedge method are presented in Fig. 3.13a. A limitation of the method is that it is only suitable for sharp tips because large, blunt tips (such as colloidal probes) will give unreliable data while sliding on the SrTiO₃ surface, which is relatively steep. In addition, crosstalk between deflection and torsion signals, signal drift, and laser or cantilever misalignment, cause an uncertainty in identifying the zero point in the torsion signal, affecting the value of the friction loop offset (Varenberg et al. 2003).

Varenberg et al. (2003) applied the wedge method to a silicon calibration grating instead of SrTiO₃, shown in Fig. 3.13b. The use of the Si grating enables the measurement of friction on a flat surface, aside from the sloped surfaces (with slope of 54°44') measured in the original method. This allows measurements on tips with large radii. Taking the measurement on a flat surface eliminates the uncertainty of determining the zero point in the torsion signal, which is a limitation of the original wedge method. In the following equations, the distinction between the sloped and flat surface is made by the subscripts *s* and *f*. The coefficient of friction on the sloped surface μ_s is obtained by taking the ratio of Eqs. 3.53 and 3.54 such that

$$\frac{\Delta_s}{w_s} = \frac{P \sin \theta \cos \theta + \mu_s^2 P \sin \theta \cos \theta + \mu_s^2 A \sin \theta}{\mu_s P + \mu_s A \cos \theta}$$

or

Fig. 3.13 **a** AFM image of the SrTiO₃ calibration surface used in the original wedge method (*top*), and experimental data for the wedge method (*bottom*). Lateral deflection signal for each direction and topography measured on the (101) and (103) facets of the SrTiO₃(305) calibration sample (Ogletree et al. 1996), and **b** schematic of the silicon calibration grating used by Varenberg et al. (2003)



$$\begin{aligned} & \sin \theta(P \cos \theta + A)\mu_s^2 \\ & - \frac{\Delta_s}{w_s}(P + A \cos \theta)\mu_s + P \sin \theta \cos \theta \\ & = 0 \end{aligned} \quad (3.58)$$

In Eq. 3.58, θ , P and A are known. It should be recalled that the voltage outputs $\Delta_{o,s}$ and $w_{o,s}$ are being measured instead of Δ_s and w_s . Since the ratio Δ_s/w_s is taken in Eq. 3.58, α_c cancels out and the measured voltage outputs are sufficient. Also, $\Delta_{o,s} = \Delta_{o,s,measured} - \Delta_{o,f}$ because the friction loop offset voltage of the flat surface $\Delta_{o,f}$ is nonzero and must therefore be subtracted from the measured value on the sloped surface.

After determining a value for μ_s , the calibration constant α_c is then obtained from either Eq. 3.53 or 3.54 so that the lateral force can be calculated from the voltage data. The quadratic equation (Eq. 3.58) yields two values for μ_s . The real

solution is one that would not result in a negative value for the quantities in either Eq. 3.53 or 3.54.

The advantage of using Eq. 3.58 over the ratio of the derivatives (Eq. 3.57) is that by using the silicon grating, one set of load and friction loop is sufficient in calculating the coefficient of friction, whereas for SrTiO₃, multiple load settings are necessary.

The coefficient of friction on the flat surface μ_f can be determined by substituting $\theta = 0^\circ$ into Eq. 3.53 such that

$$\begin{aligned} \mu_f &= \frac{1}{(\ell + h/2)} \frac{w_f}{(P + A)} = \frac{\beta_c}{(\ell + h/2)} \frac{w_o}{(P + A)} \\ &= \frac{\beta_c w_o}{(P + A)} \end{aligned} \quad (3.59)$$

where β_c is a calibration factor in (N m/V) given as $\alpha_c = \frac{\beta_c}{(\ell + h/2)}$.

The two values for the coefficient of friction, μ_s and μ_f , corresponding to the sloped and flat surfaces, respectively, may not necessarily be equal, but has been shown by Varenberg et al. (2003) to be close to each other. The advantages of this using the Si grating over the wedge method with SrTiO₃ are: Si is a commercially-available calibration grating, the method can be performed at any single applied load, and that all types of cantilevers (sharp and colloidal tips with a radius of curvature up to 2 μm) can be calibrated. The use of silicon grating limits the applicability of the wedge experiment to tips with small cone angles. For tips with larger cone angles and radii of curvature, the data becomes unreliable due to the high slope of the silicon grating surface. To address this, Tocha et al. (2006) proposed another modification to the calibration sample used. They used a silicon surface milled by a focused ion beam (FIB) such that notches with slopes of 20°, 25°, 30° and 35° (relative to the wafer surface) were present. The advantage of scanning on less steep slopes is that larger tips, such as colloidal probes with radius of curvature greater than 2 μm, can be calibrated, in addition to the sharp integrated tips.

The wedge method is not as convenient as the axial sliding technique (Sect. 3.4.1.1). For a flat portion of the calibration grating, it depends not only on the signal output of the AFM, but also requires the cantilever dimensions (Eq. 3.59). The wedge method requires a calibration standard (either SrTiO₃ or Si), and in addition, the method presented above is computationally not straightforward. It has also been shown that the pull-off force value used in the calculations is a major source of measurement errors in the implementation of the wedge method (Ling et al. 2007). Expressions for the friction loop parameters were derived to take into account adhesive forces. Other sources of error related to the detection of cantilever deflection have been identified, which affects the measured μ . This includes a non-zero lateral deflection of the cantilever in the absence of applied torque, variation in cantilever deflection (which in turn affects lateral deflection) due to feedback response limitations, and the

susceptibility of the lateral signal to optical interference effects. The latter two are relevant to other techniques as well.

3.4.1.3 Lever Method

The lever method for torsional stiffness determination is shown schematically in Fig. 3.14, where a lever assembly consisting of a glass fiber and a silica sphere is attached to a tipless cantilever (Feiler et al. 2000). It should be noted that a similar technique was proposed by Bogdanovic et al. (2000), but is only suitable for rectangular cantilevers. A hammerhead cantilever configuration has been proposed by Reitsma (2007) which allows in situ calibration and friction measurements. Placement of a colloidal probe off the center of the cantilever (Quintanilla and Goddard 2008) is another adaptation of the lever method that has been proposed.

In the lever method, the force-distance curve is obtained using the cantilever of interest prior to lever attachment in order to get the vertical detector sensitivity of the cantilever, c_z (in m/V) which is the slope of the constant compliance region (Feiler et al. 2000)

$$c_z = \frac{z^0}{\Delta V_z} \quad (3.60)$$

where z^0 is the cantilever deflection without the lever, and ΔV_z is the change in the vertical voltage. Next, the cantilever deflection after attachment of a lever with length l is considered. If a cantilever with the attached lever is pressed against the sample, the calibration factor for the

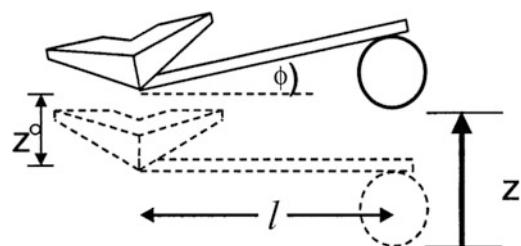


Fig. 3.14 Schematic of the lever method (adapted from Feiler et al. 2000)

vertical deflection with an attached lever, $c_{z,L}$ (in m/V) is given by

$$c_{z,L} = \frac{z}{\Delta V_z} \quad (3.61)$$

where z is the resulting deflection. The calibration factor ε , which converts the measured lateral voltage ΔV_y into an applied torque M_T , is defined as

$$M_T = \varepsilon \Delta V_y \quad (3.62)$$

The torque is from the applied force F acting on the lever, i.e., $M_T = Fl$. This applied force causes a change in both the vertical and lateral voltage signals. In the vertical direction, $F = k_z z^0$. Substituting this and Eq. 3.60 into the basic definition for torque yields

$$M_T = k_z c_z \Delta V_z l \quad (3.63)$$

By combining Eqs. 3.62 and 3.63, an expression for ε is obtained

$$\varepsilon = c_z k_z l \frac{\Delta V_z}{\Delta V_y} \quad (3.64)$$

By applying a known torque to the cantilever while measuring the change in the angular deflection of the cantilever ϕ , the torsional stiffness can be directly calculated from $M_T = k_\phi \phi$. The angular deflection is obtained from the difference between the vertical movement of the lever and that of the cantilever itself, divided by the lever length such that

$$\phi = \frac{(z - z^0)}{l} = \frac{(z - c_z \Delta V_z)}{l} = \frac{z(1 - c_z/c_{z,L})}{l} \quad (3.65)$$

By substituting this to Eq. 3.63, the expression for k_ϕ is obtained as

$$k_\phi = \frac{k_z l^2}{(c_{z,L}/c_z - 1)} \quad (3.66)$$

In this method, the friction force acting on the tip with height ℓ (with the lever no longer present)

can be determined independently from the torsional stiffness as long as the calibration factor ε (Eq. 3.64) is known. In this case, $M_T = F \ell$. By relating this to Eq. 3.64, the following is obtained

$$F = \frac{\varepsilon \Delta V_y}{\ell} \quad (3.67)$$

The main disadvantage of this technique is that it is only practically applicable to tipless cantilevers due to the necessity of attaching a lever to the cantilever of interest in the calibration procedure, which in itself is difficult and time consuming in practice. It can be used on cantilevers with integrated tips as long as the tip is not affected when the glass lever is attached to the cantilever of interest.

3.4.1.4 Particle Interaction Apparatus

The so-called particle interaction apparatus relies on the presence of a reference cantilever oriented perpendicular to the cantilever of interest, which in turn, contains an attached spherical particle (colloidal probe). As shown in Fig. 3.15a, this part is denoted as the calibration experiment (Ecke et al. 2001). The second part is the friction experiment where the cantilever with the colloidal probe slides on a flat surface. This is shown in Fig. 3.15b, along with the definition of the relevant parameters for this technique.

The lateral sensitivity c_y (in m/V) is defined as the ratio of the total lateral displacement on the lateral voltage ΔV_y . In this method, two lateral sensitivities are determined. First, the calibration experiment is performed where the reference cantilever is pushed against the cantilever of interest, and the lateral sensitivity of the reference c_y^{ref} is obtained as (Ecke et al. 2001)

$$c_y^{ref} = \frac{y_o + y_{ref}}{\Delta V_y} \quad (3.68)$$

where the displacement comes from both the tip y_o and the reference y_{ref} . Afterwards, the friction experiment is performed on a rigid flat surface instead of the reference cantilever in order to obtain the lateral sensitivity of the cantilever on the rigid surface c_y^0

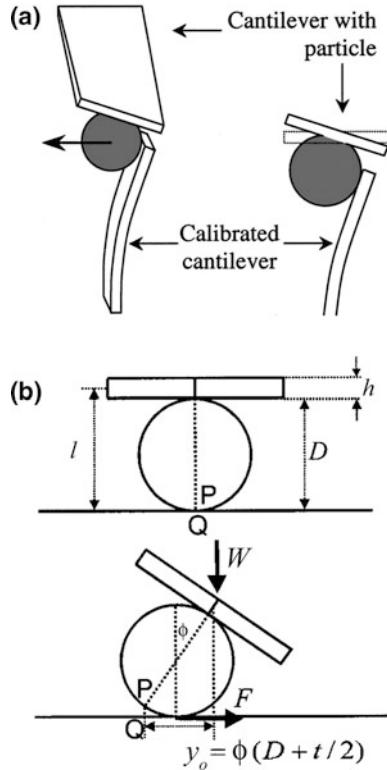


Fig. 3.15 **a** Schematic of the particle and reference cantilever in the particle interaction apparatus, and **b** schematic of the front view of a particle attached to a cantilever in contact with the surface in the particle interaction apparatus. The equilibrium position is shown on *top*. At the bottom illustration, normal force has been applied, leading to rotation of the particle by angle θ (adapted from Ecke et al. 2001)

$$c_y^0 = \frac{y_o}{\Delta V_y} \quad (3.69)$$

where the displacement is solely from the tip. By substituting Eq. 3.69 into Eq. 3.68, the sensitivity of the reference cantilever is expressed as

$$c_y^{ref} = c_y^0 + \frac{y_{ref}}{\Delta V_y} \quad (3.70)$$

During the calibration experiment the forces acting on the cantilever of interest and the reference are equal such that

$$k'_T y' = k_z^{ref} y_{ref} \quad (3.71)$$

By substituting Eq. 3.71 into 3.70, the lateral stiffness of the cantilever, k'_T , is related to the normal stiffness of the reference cantilever, k_z^{ref} , by

$$k'_T = \frac{c_y^{ref} - c_y^0}{c_y^0} k_z^{ref} \quad (3.72)$$

The torque applied to the cantilever by a lateral force F is given by $M_T = Fl$, where l is the length of the lever arm. During calibration with the reference, the length of the lever arm is given by

$$l' = \frac{1}{2}(D + h) \quad (3.73)$$

In the friction experiment on the rigid flat surface, l is given by

$$l = \left(D + \frac{h}{2} \right) \quad (3.74)$$

where $h/2$ takes into consideration that the twisting axis is located in the middle of the thickness of the cantilever.

Since the torque is also related to the twist of the cantilever ϕ (Fig. 3.15b) by $M_T = k_\phi \phi$, the lateral stiffness can be expressed as

$$k_{yT} = \frac{k}{l^2} \quad (3.75)$$

Equation 3.75 can be used to relate the lateral force constant from the calibration k'_T and that from the friction experiment, k_{yT} , as follows

$$\frac{k_{yT}}{k'_T} = \left(\frac{l'}{l} \right)^2 \quad (3.76)$$

The expressions for the length of the lever arm can then be substituted into Eq. 3.76 such that

$$k_{yT} = k'_T \left(\frac{D + h}{2D + h} \right)^2 \quad (3.77)$$

By substituting Eq. 3.77 into Eq. 3.72 the lateral stiffness is obtained

$$k_{yt} = \frac{c_y^{ref} - c_y^0}{c_y^0} k_z^{ref} \left(\frac{D+h}{2D+h} \right)^2 \quad (3.78)$$

The lateral force during the friction experiment is obtained once k_{yt} is known by using

$$F = k_{yt} c_y \Delta V_L \quad (3.79)$$

where

$$c_y = c_y^0 \left(\frac{D+h}{2D+h} \right) \quad (3.80)$$

As presented above, this particle interaction approach only gives a value for the lateral stiffness, and the torsional stiffness is determined indirectly by using the relationship $k_{yt} = k_\phi/l^2$. A disadvantage of this method is that it is limited in applicability; it is only suited for colloidal probes. In addition, the normal stiffness of the reference cantilever is required. Another limitation is that this method requires the measurement of the cantilever thickness and the diameter of the attached particle, which requires SEM for highest accuracy.

3.4.1.5 Miscellaneous Methods

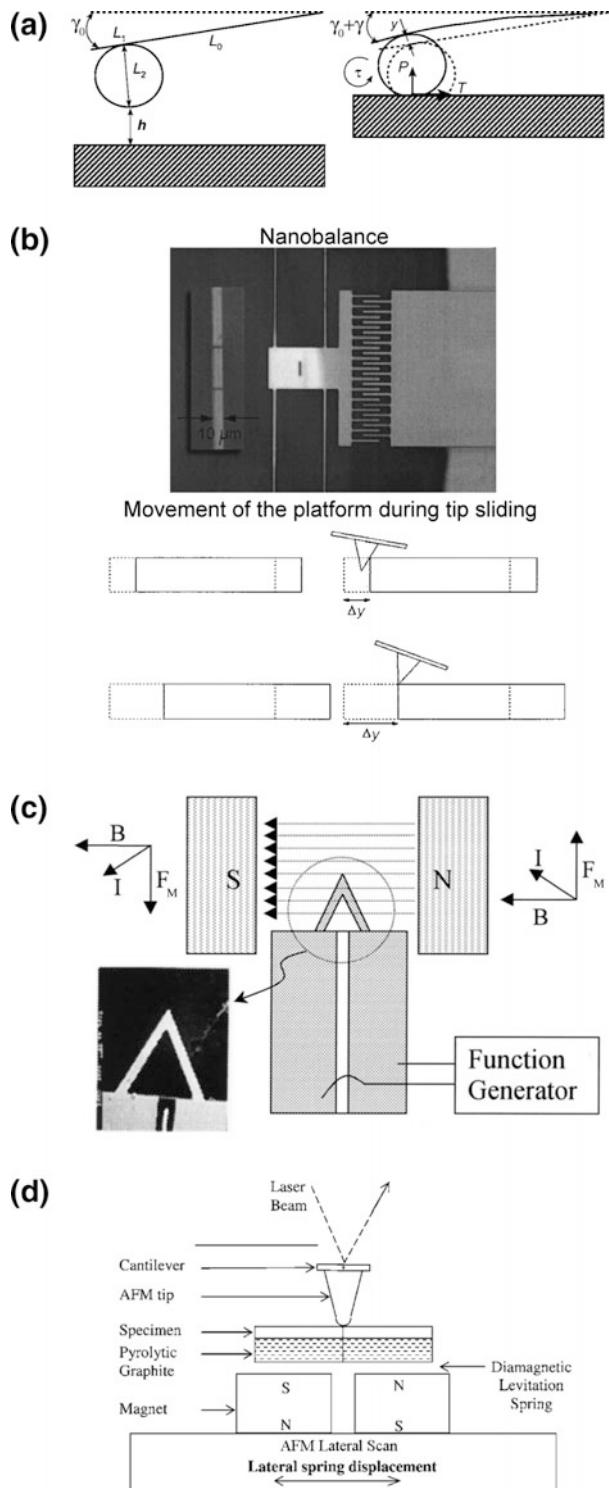
The following are examples of methods where either limitations in the technique or the need for additional instrumentation for measuring the lateral force or the torsional stiffness may not facilitate their routine implementation.

Stiernstedt et al. (2005) proposed a method for measuring the coefficient of friction while sliding in the axial direction. This method is based on the observed hysteresis of the slope of the photodetector signal versus piezo distance during approach and withdrawal of the tip on the surface (Attard et al. 1999; Stiernstedt et al. 2005, 2006; Attard 2007). The hysteresis is caused by the tilt of the cantilever (approx. 10°–15°) relative to the horizontal axis, which causes the tip to slide on the substrate during the acquisition of the curve for the normal load (“force curve”). As shown in the left side of Fig. 3.16a, L_0 is the length of the flexible part of the cantilever, and L_1 is the length of the adhesive holding the probe. A tip (or

colloidal probe) with length L_2 is attached at the point $L_1 + L_0$ from the base of the cantilever. γ_0 is the angle of the cantilever relative to the horizontal axis. Before contact, the vertical separation between tip and substrate is h . During contact, a normal force W and torque M_T applied to the end of the cantilever causes a deflection at the end of the beam, y , and an additional deflection, γ , such that the actual angle of the tip is $\gamma_0 + \gamma$ (right side of Fig. 3.16a). This compliance hysteresis analysis accounts for the presence of friction between the tip and substrate during contact, and removes the effect of friction from the measured forces. One main disadvantage of this method is that it is not a direct means for converting the observed deflection signal into the lateral force. The authors used the lever method (discussed in Sect. 3.4.1.3) to calibrate their cantilever. On the other hand, this method provides a direct way of determining the coefficient of friction, provided that the angle of tip tilt relative to the horizontal axis is accurately known. This method is also unsuitable for materials exhibiting viscoelasticity, where the observed hysteresis is from relaxation of the stresses in the material and not from friction. In addition, the method outlined above is computationally not straightforward.

The lateral electrical nanobalance (LEN), shown in Fig. 3.16b, was developed at the National Physical Laboratory in the United Kingdom. It is a comb actuator-based MEMS device fabricated by silicon-on-insulator micro-machining (Cumpson et al. 2005). The nanobalance consists of a gold-coated silicon platform suspended on cantilever beams. The platform contains a 3 μm wide slit, which the AFM tip will enter as it scans the surface. The displacement of the platform as the tip enters the slit is a measure of lateral displacement. Prior to a cantilever calibration, the nanobalance displacement is calibrated in order to obtain its static and dynamic displacement properties, and to separate actual mechanical displacement from parasitic capacitances. In the cantilever calibration experiment, the tip is scanned in and out of the nanobalance platform in this device. The lateral

Fig. 3.16 Schematics of a the tip-surface contact showing horizontal sliding of the tip in the compliance hysteresis method (adapted from Stiernstedt et al. 2005), b lateral electrical nanobalance and movement of the platform (Δy) during the tip sliding (Cumpson et al. 2005), c Lorentz force induction technique (Jeon et al. 2004), and d diamagnetic levitation spring system (Li et al. 2006)



force is the product of the lateral spring constant and the lateral displacement of the platform as during the scan. The torsional stiffness is obtained by taking the ratio of the stiffness of the device and the slope of the recorded lateral force signal when scanning on the platform. The authors reported that the precision of the measured lateral force is around $\pm 7\%$. However, a limitation of this technique is that the deflection signal is susceptible to nonlinear behavior and crosstalk between lateral and normal force signals.

Jeon et al. (2004) measured the torsional stiffness without the application of a normal load by inducing the Lorentz force when the cantilever of interest was positioned between two strong permanent magnets (Fig. 3.16c). The Lorentz force (F_M) was developed by applying an electric current to the cantilever in a magnetic field. The torsional stiffness of the cantilever of interest can be obtained by using the relationship between the magnetic torque and the torque induced by the twisting of the cantilever

$$k_\phi = \frac{1}{\phi} NISB \sin \theta \quad (3.81)$$

where N is the number of current loops (1), I is the current and S is the area of the cantilever tip, B is the magnetic field, and θ is the angle between the applied magnetic field and the normal to the cantilever plane. Unfortunately, the technique is not straightforward to implement, as it requires knowledge of the cantilever geometry, as well as the need for a magnetic field source, and the requirement that the cantilever should be metal-coated in order to form a circuit.

Another instrumentation-based lateral force calibration technique was proposed by Li et al. (2006). As shown in Fig. 3.16d, the calibration of the lateral force constant was performed by placing the cantilever on a set-up where the specimen is mounted on a pyrolytic graphite sheet and four magnets. Since the graphite sheet levitates in a magnetic field, it acts as a spring system. When the tip is in contact with the surface, the magnets and the AFM base are scanned

laterally, with the normal load held constant. Then, the photodetector voltage is recorded relative to the lateral force displacement. The stiffness of the levitation system is evaluated from the natural vibration frequencies of the system. In the cantilever calibration experiment, the lateral photodetector output is monitored as a function of the lateral spring displacement, which is assumed to be predominantly from the magnetic spring (tip displacement assumed to be negligible). This allows the measurement of the lateral stiffness with a reported accuracy on the order of 0.1%. The major disadvantage of this technique is the need for additional instrumentation for the diamagnetic levitation system.

3.4.2 Dynamic Methods

3.4.2.1 Torsional Added Mass Method

The added mass method for determining stiffness was extended to enable the calibration of the torsional stiffness (Green et al. 2004). They used an approach similar to that used in the determination of the normal stiffness. In this case, the radial resonance frequency of torsional vibration, ω_T , changes upon addition of a mass

$$\omega_T^2 = \frac{k_\phi}{J_s + J_e} \quad (3.82)$$

where k_ϕ is the torsional stiffness, J_e is the effective mass moment of inertia, and J_s is the mass moment of inertia of the added mass m_s , where $J_s = \frac{7}{5}m_s r^2$. The mass of the sphere can be expressed in terms of the sphere's radius r and density ρ_s such that

$$J_s = \frac{7}{5}m_s r^2 = \frac{28}{15}\pi\rho_s r^5 \quad (3.83)$$

Substitution of Eq. 3.83 into 3.82 gives an expression for the torsional stiffness (Green et al. 2004)

$$\frac{28\pi\rho_s r^5}{15} = \frac{k_\phi}{\omega_T^2} - J_e \quad (3.84)$$

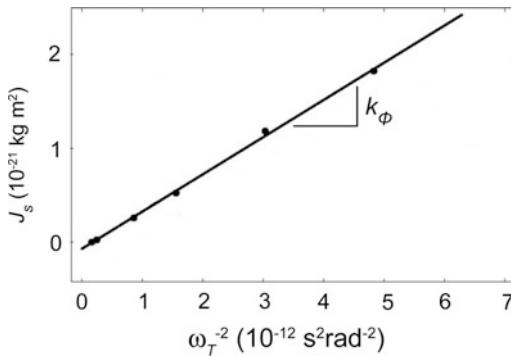


Fig. 3.17 Mass moment of inertia as a function of the square of the angular frequency of torsional vibration, where the slope is the torsional stiffness (adapted from Green et al. 2004)

A plot of J_s versus ω_T^{-2} will give a straight line with a slope equal to k_ϕ . An example of the application of the torsional added mass method is shown in Fig. 3.17.

Similar to the added mass method for normal stiffness determination, this technique can potentially damage the cantilever, and errors in determining the radius and density of the added mass directly affect the measured torsional stiffness. In addition, placement of the added mass away from the axis or away from the free end of the cantilever can change the k_ϕ by as much as 25% (Green et al. 2004).

3.4.2.2 Torsional Resonance Method

In this technique by Green et al. (2004), the resonance method for normal stiffness determination (Sader et al. 1999, 2005) was extended for determining the torsional stiffness. On the assumption that $L \gg b \gg h$, k_ϕ for a rectangular beam is given by Love (1959)

$$k_\phi = \frac{1}{3\pi} \rho_c b^3 h L \omega_{T,vac}^2 \quad (3.85)$$

where $\omega_{T,vac}$ is the resonance frequency of torsional vibration in vacuum. The quantity $\omega_{T,vac}$ can be related to its corresponding value in a fluid medium (e.g., air) ω_T by Green et al. (2004)

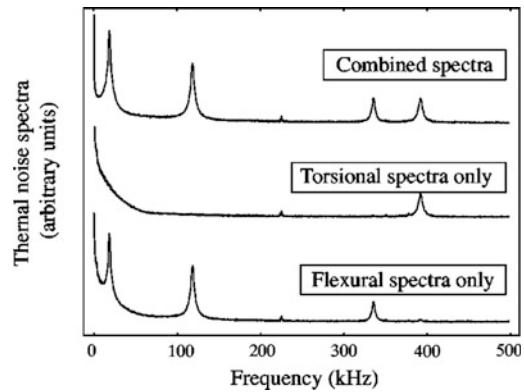


Fig. 3.18 Thermal noise spectra due to flexural vibration, torsional vibration and a combination of the two (Green et al. 2004)

$$\omega_{T,vac} = \omega_T \left(1 + \frac{3\pi\rho_f b}{2\rho h} \Gamma_r^T(\omega_T) \right)^{1/2} \quad (3.86)$$

where Γ_r^T is the real part of the hydrodynamic function. The density of the cantilever is given by the following

$$\rho_c h = \frac{3\pi\rho_f b}{2} [Q_T \Gamma_i^T(\omega_T) - \Gamma_r^T(\omega_T)] \quad (3.87)$$

where Q_T is the quality factor of the resonance peak of torsional vibration in a fluid medium. By substituting Eqs. 3.86 and 3.87 into 3.85, the torsional stiffness is given by Green et al. (2004)

$$k_\phi = 0.1592 \rho_f b^4 L Q_T \omega_T^2 \Gamma_i^T(\omega_T) \quad (3.88)$$

Figure 3.18 shows the thermal noise spectra due to flexural vibration, torsional vibration and a combination of the two (Green et al. 2004). All three are shown to emphasize that data is usually a combination of the flexural and torsional vibration, and that the source of each peak should be analyzed with care. The peaks due to flexural and torsional vibration can be distinguished by picking the appropriate detector signals. Data from the middle plot should be used to obtain the torsional stiffness using this resonance technique.

An advantage of this technique is that it is relatively simple. However, the model relies on

cantilever dimensions, which are sources of measurement uncertainty. In addition, it relies on knowledge of the Reynolds number of the fluid medium and the determination of the hydrodynamic function. It must be noted that the resonance method discussed here is based on the geometry of a rectangular cantilever beam. A solution to the hydrodynamic function is needed to extend this technique to triangular cantilevers. Alternatively, a wafer with both rectangular and triangular cantilevers present can be used, as described in Sect. 3.3.2.2

3.4.2.3 Calculation of Lateral Force from Torsional Stiffness

In the dynamic methods discussed in Sects. 3.4.2.1 and 3.4.2.2, the torsional stiffness k_ϕ (in N m) is obtained. Calculation of the lateral force from k_ϕ is not straightforward in this case and an additional step is necessary. Photodetector calibration is necessary in order to obtain the torsional sensitivity (in V/rad). Pettersson et al. (2007) describes a procedure for the torsional calibration of the photodetector. Alternatively, if k_ϕ is converted into k_{yT} , the lateral sensitivity can be determined (Cannara et al. 2006).

An example of a procedure which can be implemented in general is the “optical geometry” method proposed by Liu et al. (1996). They derived a relationship for the lateral force as the product of the half-width of the friction loop multiplied by a constant involving the stiffness parameter. This parameter is the product of k_ϕ , k_z , the sensitivity of the input channel of the FFM and the tip length, divided by the product of the normal sensitivity of the force curve slope, the coefficient relating the sensitivities of the FFM and AFM photodiode pairs and a constant which relates the normal bending angle and the bending force acting on the cantilever tip. The normal bending is calculated with the knowledge of the cantilever’s material properties and dimensions. These parameters are necessary in order to relate cantilever bending and torsion. The half-width of the friction loop and the sensitivity of the FFM input channel are related to the change in the photodiode signal resulting from torsion. This, in

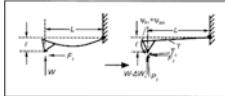
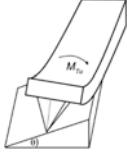
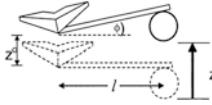
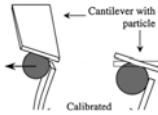
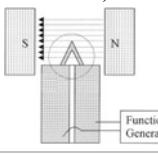
turn, is related to the change in photodiode signal from bending, the coefficient relating the sensitivities of the FFM and AFM and the cantilever’s normal bending angle. As discussed earlier, the material properties and cantilever dimensions are sources of measurement uncertainties such that this method for calculating the lateral force from the torsional stiffness may not be accurate.

3.4.3 Discussion

Table 3.2 lists the highlights for the lateral force and stiffness calibration methods discussed above. Quantitative measurement and calibration of the lateral force is not as straightforward as the corresponding methods for the normal force. This is because lateral motion involves bending or torsion of the cantilever during sliding in the parallel or perpendicular direction, relative to the long axis of the cantilever, respectively. Aside from classifying the methods discussed above as either static or dynamic, they can also be differentiated based on the measurable parameters that each model provides. The axial bending, wedge, and compliance hysteresis techniques all result in a value for the coefficient of friction. The lateral force can be easily calculated afterwards. Meanwhile, the lever, particle interaction, torsional added mass, and torsional resonance methods all give a value of the stiffness (torsional, or torsional and lateral). The stiffness can be used to calculate the lateral force, although it is not straightforward in the case of the torsional added mass and torsional resonance methods.

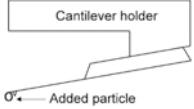
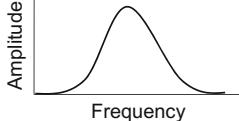
The distinction can also be made based on classifying them as “one-step” and “two-step” methods. Some methods, such as axial sliding and wedge, provide a direct means (one-step) for obtaining the lateral force. Meanwhile, other calibration techniques, such as the lever particle interaction, torsional added mass, and torsional resonance methods, consist of two steps. The first step is the determination of the torsional or lateral stiffness. The second step is the determination of the photodetector sensitivity (V/rad or V/m).

Table 3.2 Summary of techniques for calibrating the lateral force and torsional stiffness

Technique	Principle/key equation	Pros	Cons
<i>Static methods</i>			
Axial sliding ^a (Ruan and Bhushan 1994a)	Height variation of piezo tube is monitored during sliding in the long axis of cantilever $\mu = \frac{F}{W} = \left[\frac{(\Delta H_1 + \Delta H_2)}{H_0} \right] \left(\frac{L}{4\ell} \right)$ 	Simple; easy to implement	Requires measurement of cantilever length and tip height
Wedge method ^a (Ogletree et al. 1996)	Cantilever is scanned across calibration standard with two well-defined slopes $\mu + \frac{1}{\mu} = \frac{2a'_o}{w'_o \sin 2\theta}$ 	Does not require cantilever dimensions	Requires calibration standard; pull-off force value may cause measurement errors; computation method is not straightforward
Lever method (Feiler et al. 2000)	Lever with glass fiber and sphere is attached to cantilever and angular deflection of cantilever is measured $k_\phi = \frac{k_l l^2}{(c_{z,L}/c_z - 1)}$ 	-	Only practical for tipless cantilevers; requires attachment of lever to cantilever of interest, which is difficult and time consuming in practice
Particle interaction apparatus (Ecke et al. 2001)	Reference cantilever is oriented perpendicular to cantilever with spherical tips (colloidal probe) and the twist of the cantilever is measured which is used to calculate k_y and k_ϕ $k_{yT} = \frac{k_\phi}{l^2} = \frac{c_y^{ref} - c_y^0}{c_y^0} k_z^{ref} \left(\frac{D+h}{2D+h} \right)^2$ 	-	Need special setup; only applicable to colloidal probes
Compliance hysteresis method ^a (Stiernstedt et al. 2005)	Tilt of cantilever and the hysteresis of the compliance lines during approach and withdrawal are measured	-	Requires accurate tip tilt angle value; computation method is not straightforward
Lateral electrical nanobalance (Cumpson et al. 2005)	Cantilever is calibrated by scanning tip in and out of the nanobalance platform	-	Requires fabricated nanobalance; signal is susceptible to nonlinearity and crosstalk
Lorentz force induction (Jeon et al. 2004)	Lorentz force is induced when an electric current is applied to cantilever in a magnetic field and the torsional resonance frequency is measured $k_\phi = \frac{1}{\phi} NISB \sin \theta$ 	-	Requires magnetic field source, metal coating on cantilever, and knowledge of cantilever geometry
Diamagnetic levitation spring system (Li et al. 2006)	Cantilever is placed in a set-up where specimen is on a pyrolytic graphite sheet and four magnets and the levitating graphite sheet is related to the lateral tip displacement	-	Additional instrumentation is required

(continued)

Table 3.2 (continued)

Technique	Principle/key equation	Pros	Cons
<i>Dynamic methods</i>			
Torsional added mass method (Green et al. 2004) 	A number of particles are individually attached to the end of the cantilever and the resonance frequency is measured $\frac{28\pi\rho_s s^5}{15} = \frac{k_\phi}{\omega_T^2} - J_e$	-	Requires attachment of particles to cantilever; risk of cantilever damage; uncertainty in dimensions and errors in placement of added particle lead to significant error; conversion of torsional stiffness to lateral force is not straightforward.
Torsional resonance method (Green et al. 2004) 	The resonance frequency, quality factor and dimensions are measured For a rectangular beam, $k_\phi = 0.1592\rho_f b^4 L Q_T \omega_T^2 \Gamma_i(\omega_T)$	Relatively simple for torsional stiffness determination	Model relies on cantilever dimensions; requires Reynolds number of the fluid medium and calculation of the hydrodynamic function; conversion of torsional stiffness to lateral force is not straightforward

^aThese techniques allow direct measurement of the lateral force

In evaluating the various techniques for measuring the lateral force of the cantilever, both the ease of implementation as well as the limitations of the respective models should be taken into consideration. The axial bending method is simple to implement, but the requirement of measuring the length of the cantilever and the tip introduces uncertainty into the resulting value of the coefficient of friction.

The wedge method depends solely on the signals obtained during the friction experiment (on the sloped surfaces) and not on any cantilever dimension or property, which makes this technique desirable. Another benefit of this technique is that it can be used to ascertain the state of tip geometry (Cain et al. 2001). However, it should be noted that the procedures described for performing the wedge method calibration have their own caveats. The use of a relatively steep calibration standard limits the applicability of the wedge method to sharp tips and colloidal probes with small radii of curvature. As pointed out earlier, inaccuracies in the pull-off force could affect the coefficient of friction obtained. Also, this method is not mathematically straightforward since careful consideration has to be made in selecting

between the two values obtained in the quadratic equation solving for the coefficient of friction.

The lever and particle interaction methods are based on measuring the angular deflection of the cantilever and can measure both the torsional stiffness and lateral force. However, both the lever and particle interaction methods appear to be complicated to implement experimentally. The lever method is only practical for tipless cantilevers, while there is a risk of damaging the cantilever or dislodging the colloidal probe in the particle interaction method. The particle interaction method is also limited because it is only applicable to colloidal probes.

The compliance hysteresis method is performed in the longitudinal direction and yields a value of the coefficient of friction similar to the axial sliding and wedge methods. However, the model requires the tilt of the cantilever and its length, which may cause artifacts if measured incorrectly. Also, the computation method is not very straightforward.

The lateral electrical nanobalance (LEN), Lorentz force induction, and diamagnetic levitation spring methods are examples of novel calibration techniques that require the use of

additional instrumentation in order to induce cantilever twisting and measure the resulting lateral displacement. It should be pointed out that there are other instrument-based methods that have been proposed in the literature. However, they were not included in the review, because either the measurement principle is similar to the ones discussed here, or the method is largely dependent on cantilever properties (such as mechanical properties of the cantilever and its dimensions). Two of the calibration procedures discussed in Sect. 3.4.1.5 are noteworthy due to their reported accuracy. Cumpson et al. (2005) reported that in the LEN method, the lateral force uncertainty is as low as $\pm 7\%$, while the diamagnetic levitation method was reported to measure stiffness values with accuracy on the order of 0.1%. The Lorentz force induction method is of interest because it does not require the application of normal load, such that deflection is nominally zero, such that a pure torsional response is obtained. However, these potential advantages are outweighed by the requirement to acquire (or construct) and calibrate the measurement set-ups that will then be used to calibrate the cantilever.

The dynamic techniques (added mass and resonance) for determining the torsional stiffness have advantages and disadvantages similar to the normal stiffness methods that they were originally derived from. For the added mass technique, the main concern is the possibility of cantilever damage from the added sphere. In addition, the use of an inaccurate value for the sphere, as well as errors in placement of the sphere (e.g. off-axis), will lead to significant errors. For the resonance technique, an advantage is that the procedure enables the simultaneous determination of both the torsional and normal stiffness. However, the need for the dimensions of the cantilever (length and width) is a disadvantage as it adds some uncertainty in the measurement. Another disadvantage is that the lateral force cannot be determined easily from the torsional stiffness using these dynamic methods. An additional procedure to determine the sensitivity of the photodetector is necessary in order to complete the quantification of the lateral force.

Among all the techniques presented, the axial sliding, wedge and torsional resonance methods are preferred for calibrating lateral forces. These techniques are not limited to cantilever geometry and can be used on both integrated tips and colloidal probes. The first one (axial sliding) is recommended due to its relative simplicity, and the latter two, though not as convenient to implement, also have minimal susceptibility to cantilever damage. However, it should be pointed out that the wedge and torsional resonance methods have some susceptibility to measurement uncertainties, as described earlier.

It should be noted that one of the sources of error not discussed so far is the crosstalk (either optical or mechanical in nature) between the signals corresponding to lateral and normal cantilever deflection, as this is difficult to eliminate in general. Optical crosstalk takes place because of the rotational misalignment of the PSD. This causes an error in the topography and friction measurements, and is particularly pronounced for nanostructures with significant local variations in frictional properties (Piner and Ruoff 2002). Various approaches exist in the literature to analyze optical crosstalk (Bhushan 2011, 2017; Cannara et al. 2006). Mechanical crosstalk is caused by the positional offset of the tip relative to the symmetry axis of the cantilever. The mechanical crosstalk artifact can be corrected for by using a set of relationships that account for the offset caused by the non-orthogonal alignment of the PSD in combination with the AFM tip being off-centered relative to the cantilever (Asay and Kim 2006).

3.5 Outlook

Over the years, the use of the AFM has expanded beyond surface profiling to capabilities such as the nanomanipulation and the measurement of various physical properties. There is a need to calibrate the normal and lateral forces of AFM cantilevers in order to use AFM data properly in studies of tribological and mechanical properties, and as presented in this chapter, various techniques are available.

For calibrating the normal forces, the reference cantilever, resonance, and thermal noise methods are desirable. The use of a reference cantilever is recommended due to the simplicity of this method. The resonance and thermal noise methods are both relatively simple to implement, and the latter technique is regarded to provide accurate values of the normal stiffness as it is the least susceptible to measurement uncertainties. For lateral forces, the axial sliding, wedge, and torsional resonance methods are desirable. The axial sliding technique is a simple technique. The wedge and torsional resonance methods are noteworthy as they possess minimal susceptibility to cantilever damage. However, there are other sources of measurement uncertainties involved with these techniques, most notably the effect of the pull-off force on the wedge method and the need for cantilever dimensions in the torsional resonance method.

There is a need to further understand the effect of optical and mechanical crosstalk, with the latter having been less studied so far. It would be of great benefit to experimentalists to have simple tools to deconvolute signal artifacts, thus ensuring accuracy in the measurement and analysis of normal and lateral forces. Lastly, it should be pointed out that most of the available calibration techniques are more applicable to the rectangular cantilever due to the simplicity of its geometry. Since, in practice, triangular cantilevers are as widely-used as the rectangular cantilever, more geometry-independent experimental techniques should be developed.

Appendix—Nomenclature

Roman Symbols

<i>A</i>	Adhesive force
<i>a</i>	Amplitude
<i>b</i>	Cantilever width
<i>b_l</i>	Width of the leg in a triangular cantilever
<i>c</i>	Photodetector sensitivity
<i>d</i>	Distance of the tip to the edge of the cantilever
<i>E</i>	Young's modulus

<i>F</i>	Friction force (lateral force)
<i>f</i>	Frequency of the cantilever
<i>G</i>	Shear modulus
<i>H</i>	Piezo tube height in axial sliding method
<i>h</i>	Cantilever thickness
<i>I</i>	Area moment of inertia
<i>J</i>	Mass moment of inertia
<i>k_B</i>	Boltzmann's constant
<i>k_x</i>	Cantilever stiffness in the direction parallel to the longitudinal axis
<i>k_{yB}</i>	Cantilever stiffness in the direction perpendicular to the longitudinal axis due to bending
<i>k_{yT}</i>	Cantilever stiffness in the direction perpendicular to the longitudinal axis due to applied torque
<i>k_z</i>	Cantilever stiffness in the normal direction
<i>k_ϕ</i>	Cantilever torsional stiffness
<i>L</i>	Cantilever length
<i>l</i>	Lever length
<i>ℓ</i>	Tip length
<i>M_T</i>	Torsion moment
<i>m</i>	Mass of the cantilever
<i>m[*]</i>	Effective mass of the cantilever
<i>m_s</i>	Mass of added particle
<i>P</i>	Normal load component in axial sliding method, or applied load in the wedge method
<i>p</i>	Area of the power spectrum in the thermal noise method
<i>Q</i>	Quality factor
<i>r</i>	Radius of added particle
<i>T</i>	Temperature (in thermal tune method), or friction/horizontal force component (in axial sliding wedge and compliance hysteresis methods)
<i>W</i>	Normal load
<i>w</i>	Half width of the friction loop in the wedge method
<i>z</i>	Cantilever deflection.

Greek Symbols

α	One-half the included angle between the legs of a triangular cantilever
α_c	Lateral force calibration factors in the wedge method
β_c	

γ	Cantilever tilt relative to horizontal axis	Bogdanovic G, Meurk A, Rutland MW (2000) Tip friction—torsional spring constant determination. <i>Coll Surf B</i> 19:397–405
Δ	Friction loop offset in the wedge method	Butt H-J, Jaschke M (1995) Calculation of thermal noise in atomic force microscopy. <i>Nanotechnology</i> 6:1–7
$\delta_{I,II}$	Deflection of the cantilever in the parts I and II	Butt H-J, Siedle P, Seifert K, Fendler K, Seeger T, Bamberg E, Weisenhorn AL, Goldie K, Engel A (1993) Scan speed limit in atomic force microscopy. <i>J Microscopy</i> 169:75–84
ε	Calibration factor in the lever method	Cain RG, Reitsma MG, Biggs S, Page NW (2001) Quantitative comparison of three calibration techniques for the lateral force microscope. <i>Rev Sci Instrum</i> 72:3304–3312
η	Viscosity of fluid medium	Cannara RJ, Eglin M, Carpick RW (2006) Lateral force calibration in atomic force microscopy: a new lateral force calibration method and general guidelines for optimization. <i>Rev Sci Instrum</i> 77:053701
μ	Coefficient of friction	Cleveland JP, Manne S, Bocek D, Hansma PK (1993) A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. <i>Rev Sci Instrum</i> 64:403–405
ν	Poisson's ratio	Clifford CA, Seah MP (2005) The determination of atomic force microscope cantilever spring constants via dimensional methods for nanomechanical analysis. <i>Nanotechnology</i> 16:1666–1680
ρ	Density	Cook SM, Lang KM, Chynoweth KM, Wigton M, Simmonds RW, Schaffer TE (2006) Practical implementation of dynamic methods for measuring atomic force microscope cantilever spring constants. <i>Nanotechnology</i> 17:2135–2145
θ	Inclination of calibration standard in the wedge method	Cumpson PJ, Hedley J (2003) Accurate analytical measurements in the atomic force microscope: a micro-fabricated spring constant standard potentially traceable to the SI. <i>Nanotechnology</i> 14:1279–1288
θ_{II}	Rotation of the legs of a triangular cantilever in the longitudinal direction	Cumpson PJ, Hedley J, Clifford CA (2005) Microelectromechanical device for lateral force calibration in the atomic force microscope: lateral electrical nanobalance. <i>J Vac Sci Technol B</i> 23:1992–1997
ϕ	Cantilever rotation from applied torque	Ecke S, Raiteri R, Bonaccorso E, Reiner C, Deisereth HJ, Butt HJ (2001) Measuring normal and friction forces acting on individual fine particles. <i>Rev Sci Instrum</i> 72:4164–4170
ω	Angular frequency of the cantilever	Feiler A, Attard P, Larson I (2000) Calibration of the torsional spring constant and the lateral photodiode response of frictional force microscopes. <i>Rev Sci Instrum</i> 71:2746–2750
Γ	Hydrodynamic function in the resonance method	Green CP, Lioe H, Cleveland JP, Proksch R, Mulvaney P, Sader JE (2004) Normal and torsional spring constants of atomic force microscope cantilevers. <i>Rev Sci Instrum</i> 75:1988–1996
ς	Damping ratio	Hutter J (2005) Comment on tilt of atomic force microscope cantilevers: effect on spring constant and adhesion measurements. <i>Langmuir</i> 21:2630–2632

References

- Albrecht TR, Quate CF (1987) Atomic resolution imaging of a nonconductor by atomic force microscopy. *J Appl Phys* 62:2599–2602
- Albrecht TR, Akamine S, Carver TE, Quate CF (1990) Microfabrication of cantilever styli for the atomic force microscope. *J Vac Sci Technol A* 8:3386–3396
- Anonymous (2016) AFM probes, tips, and cantilevers. Bruker AFM Probes, Camarillo, CA. <http://www.brukerafmprobes.com>
- Asay DB, Kim SH (2006) Direct force balance method for atomic force microscopy lateral force calibration. *Rev Sci Instrum* 77:043903
- Attard P (2007) Measurement and interpretation of elastic and viscoelastic properties with the atomic force microscope. *J Phys Condens Matter* 19:473201
- Attard P, Carambassis A, Rutland MW (1999) Dynamic surface force measurement. 2. Friction and the atomic force microscope. *Langmuir* 15:553–563
- Behrens I, Doering L, Peiner E (2003) Piezoresistive cantilever as portable micro force calibration standard. *J Micromech Microeng* 13:S171–S177
- Bhushan B (1999) Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton
- Bhushan B (2011) Nanotribology and nanomechanics I and II, 3rd edn. Springer, Heidelberg
- Bhushan B (2013) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2017) Springer handbook of nanotechnology, 4th edn. Springer International, Switzerland
- Bogdanovic G, Meurk A, Rutland MW (2000) Tip friction—torsional spring constant determination. *Coll Surf B* 19:397–405
- Butt H-J, Jaschke M (1995) Calculation of thermal noise in atomic force microscopy. *Nanotechnology* 6:1–7
- Butt H-J, Siedle P, Seifert K, Fendler K, Seeger T, Bamberg E, Weisenhorn AL, Goldie K, Engel A (1993) Scan speed limit in atomic force microscopy. *J Microscopy* 169:75–84
- Cain RG, Reitsma MG, Biggs S, Page NW (2001) Quantitative comparison of three calibration techniques for the lateral force microscope. *Rev Sci Instrum* 72:3304–3312
- Cannara RJ, Eglin M, Carpick RW (2006) Lateral force calibration in atomic force microscopy: a new lateral force calibration method and general guidelines for optimization. *Rev Sci Instrum* 77:053701
- Cleveland JP, Manne S, Bocek D, Hansma PK (1993) A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Rev Sci Instrum* 64:403–405
- Clifford CA, Seah MP (2005) The determination of atomic force microscope cantilever spring constants via dimensional methods for nanomechanical analysis. *Nanotechnology* 16:1666–1680
- Cook SM, Lang KM, Chynoweth KM, Wigton M, Simmonds RW, Schaffer TE (2006) Practical implementation of dynamic methods for measuring atomic force microscope cantilever spring constants. *Nanotechnology* 17:2135–2145
- Cumpson PJ, Hedley J (2003) Accurate analytical measurements in the atomic force microscope: a micro-fabricated spring constant standard potentially traceable to the SI. *Nanotechnology* 14:1279–1288
- Cumpson PJ, Hedley J, Clifford CA (2005) Microelectromechanical device for lateral force calibration in the atomic force microscope: lateral electrical nanobalance. *J Vac Sci Technol B* 23:1992–1997
- Ecke S, Raiteri R, Bonaccorso E, Reiner C, Deisereth HJ, Butt HJ (2001) Measuring normal and friction forces acting on individual fine particles. *Rev Sci Instrum* 72:4164–4170
- Feiler A, Attard P, Larson I (2000) Calibration of the torsional spring constant and the lateral photodiode response of frictional force microscopes. *Rev Sci Instrum* 71:2746–2750
- Green CP, Lioe H, Cleveland JP, Proksch R, Mulvaney P, Sader JE (2004) Normal and torsional spring constants of atomic force microscope cantilevers. *Rev Sci Instrum* 75:1988–1996
- Hutter J (2005) Comment on tilt of atomic force microscope cantilevers: effect on spring constant and adhesion measurements. *Langmuir* 21:2630–2632
- Hutter JL, Bechhoefer J (1993) Calibration of atomic-force microscope tips. *Rev Sci Instrum* 64:1868–1873
- Jeon S, Braiman Y, Thundat T (2004) Torsional spring constant obtained for an atomic force microscope cantilever. *Appl Phys Lett* 84:1795–1797

- Kim MS, Choi JJ, Park YK, Kim JH (2006) Atomic force microscope cantilever calibration device for quantified force metrology at micro- or nano-scale regime: the nano force calibrator (NFC). *Metrologia* 43:389–395
- Kim MS, Choi JJ, Kim JH, Park YK (2007) Si-traceable determination of spring constants of various atomic force microscope cantilevers with a small uncertainty of 1%. *Meas Sci Technol* 18:3351–3358
- Koinkar VN, Bhushan B (1997) Effect of scan size and surface roughness on microscale friction measurements. *J Appl Phys* 81:2472–2479
- Leach R, Chetwynd D, Blunt L, Haycocks J, Harris P, Jackson K, Oldfield S, Reilly S (2006) Recent advances in traceable nanoscale dimension and force metrology in the UK. *Meas Sci Technol* 17:467–476
- Li Q, Kim KS, Rydberg A (2006) Lateral force calibration of an atomic force microscope with a diamagnetic levitation spring system. *Rev Sci Instrum* 77:065105
- Ling X, Butt H-J, Kappl M (2007) Quantitative measurement of friction between single microspheres by friction force microscopy. *Langmuir* 23:8392–8399
- Liu E, Blanpain B, Celis JP (1996) Calibration procedures for frictional measurements with a lateral force microscope. *Wear* 192:141–150
- Love AEH (1959) A treatise on the mathematical theory of elasticity. Pergamon, London
- Matei GA, Thoreson EJ, Pratt JR, Newell DB, Burnham NA (2006) Precision and accuracy of thermal calibration of atomic force microscopy cantilevers. *Rev Sci Instrum* 77:083703
- Meyer G, Amer NM (1990) Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. *Appl Phys Lett* 57:2089–2091
- Nesterov V (2007) Facility and methods for the measurement of micro and nano forces in the range below 10^{-5} N with a resolution of 10^{-12} N (development concept). *Meas Sci Technol* 18:360–366
- Neumeister JM, Ducker WA (1994) Lateral, normal and longitudinal spring constants of atomic force microscopy cantilevers. *Rev Sci Instrum* 65:2527–2531
- Ogletree DF, Carpick RW, Salmeron M (1996) Calibration of frictional forces in atomic force microscopy. *Rev Sci Instrum* 67:3298–3306
- Ohler B (2007) Cantilever spring constant calibration using laser Doppler vibrometry. *Rev Sci Instrum* 78:063701
- Ohler B (2010) Practical advice on the determination of cantilever spring constants. <http://nanoscaleworld.bruker-axs.com/nanoscaleworld/media/p/143.aspx>
- Palacio MLB, Bhushan B (2010) Normal and lateral force calibration techniques for AFM cantilevers. *Crit Rev Solid State Mater Sci* 35:73–104, 36:261
- Pettersson T, Nordgren N, Rutland MW, Feiler A (2007) Comparison of different methods to calibrate torsional spring constant and photodetector for atomic force microscopy friction measurements in air and liquid. *Rev Sci Instrum* 78:093702
- Piner R, Ruoff RS (2002) Cross talk between friction and height signals in atomic force microscopy. *Rev Sci Instrum* 73:3392–3394
- Pratt JR, Kramar JA, Newell DB, Smith DT (2005) Review of SI traceable force metrology for instrumented indentation and atomic force microscopy. *Meas Sci Technol* 16:2129–2137
- Quintanilla MAS, Goddard DT (2008) A calibration method for lateral forces for use with colloidal probe force microscopy cantilevers. *Rev Sci Instrum* 79:023701
- Reitsma MG (2007) Lateral force calibration using a modified atomic force microscope cantilever. *Rev Sci Instrum* 78:106102
- Ruan J, Bhushan B (1994a) “Atomic-scale friction measurements using friction force microscopy: part i-general principles and new measurement techniques. *ASME J Tribol* 116:378–388
- Ruan J, Bhushan B (1994b) Atomic-scale and microscale friction of graphite and diamond using friction force microscopy. *J Appl Phys* 76:5022–5035
- Ruan J, Bhushan B (1994c) Frictional behavior of highly oriented pyrolytic graphite. *J Appl Phys* 76:8117–8120
- Sader JE (1995) Parallel beam approximation for V-shaped atomic force microscope cantilevers. *Rev Sci Instrum* 75:4583–4586
- Sader JE (1998) Frequency response of cantilever beams immersed in various fluids with applications to the atomic force microscope. *J Appl Phys* 84:64–76
- Sader JE (2003) Susceptibility of atomic force microscopy cantilevers to lateral forces. *Rev Sci Instrum* 74:2438–2443
- Sader JE, Sader RC (2003) Suitability of atomic force microscope cantilevers to lateral forces: experimental verification. *Appl Phys Lett* 83:3195–3197
- Sader JE, Larson I, Mulvaney P, White LR (1995) Method for the calibration of atomic force microscope cantilevers. *Rev Sci Instrum* 66:3789–3798
- Sader JE, Chon JWM, Mulvaney P (1999) Calibration of rectangular atomic force microscopy cantilevers. *Rev Sci Instrum* 70:3967–3969
- Sader JE, Pacifico J, Green CP, Mulvaney P (2005) General scaling law for stiffness measurement of small bodies with applications to the atomic force microscope. *J Appl Phys* 97:124903
- Sarid D, Elings V (1991) Review of scanning force microscopy. *J Vac Sci Technol B* 9:431–437
- Senden TJ, Ducker WA (1994) Experimental determination of spring constants in atomic force microscopy. *Langmuir* 10:1003–1004
- Shaw GA, Kramar J, Pratt J (2007) SI-traceable spring constant calibration of microfabricated cantilevers for small force measurement. *Exp Mech* 47:143–151
- Stiernstedt J, Rutland MW, Attard P (2005) A novel technique for the in situ calibration and measurement of friction with the atomic force microscope. *Rev Sci Instrum* 76:083710

- Stiernstedt J, Rutland MW, Attard P (2006) Erratum: a novel technique for the in situ calibration and measurement of friction with the atomic force microscope. *Rev Sci Instrum* 77:019901
- Sundararajan S, Bhushan B (2000) Topography-induced contributions to friction forces measured using an atomic force/friction force microscope. *J Appl Phys* 88:4825–4831
- Tambe NS (2005) Nanotribological investigations of materials, coatings and lubricants for nanotechnology applications at high sliding velocities. Ph.D. dissertation, The Ohio State University; available from <http://www.ohiolink.edu/etd/send-pdf.cgi?osu1109949835>
- Thomson WT, Dahleh MD (1998) Theory of vibration with applications, 5th edn. Prentice Hall, Upper Saddle River
- Timoshenko SP, Goodier JN (1970) Theory of elasticity, 3rd edn. McGraw-Hill, New York
- Tocha E, Schonherr H, Vancso GJ (2006) Quantitative nanotribology by AFM: a novel universal calibration platform. *Langmuir* 22:2340–2350
- Tortonese M, Kirk M (1997) Characterization of application specific probes for SPMs. *Proc SPIE* 3009:53–60
- Varenberg M, Etsion I, Halperin G (2003) An improved wedge calibration method for lateral force in atomic force microscopy. *Rev Sci Instrum* 74:3362–3367
- Wang YL, Zhao XZ, Zhou FQ (2007) Improved parallel scan method for nanofriction force measurement with atomic force microscopy. *Rev Sci Instrum* 78: 036107
- Wolter O, Bayer T, Greschner J (1991) Micromachined silicon sensors for scanning force microscopy. *J Vac Sci Technol B* 9:1353–1357
- Young WC, Budynas RG (2002) Roark's formulas for stress and strain, 7th edn. McGraw-Hill, New York

Biomedical Sensing with the Atomic Force Microscope

4

Constanze Lamprecht, Jürgen Strasser, Melanie Köhler,
Sandra Posch, Yoo Jin Oh, Rong Zhu,
Lilia A. Chtcheglova, Andreas Ebner
and Peter Hinterdorfer

Abstract

In this chapter we highlight the use and advantages of the atomic force microscope (AFM) in life science. Our aim is to present the wealth of experimental possibilities provided by this powerful toolbox with special regard to biomedical sensing applications. Originally invented in the 1980s to visualize solid surfaces on the nanometer scale, today AFM imaging is routinely used to nondestructively map the surface-ultrastructure of soft biological samples under physiological conditions with unprecedented lateral resolution. Owing to its force detection sensitivity that ranges from nano-Newton down to a few pico-Newton the AFM has become an established technique for exploring kinetic and structural details of inter- and intramolecular interactions and biomolecular recognition processes. The combination of such single molecule force measurements with topographical imaging has led to the development of recognition imaging, which allows for identification and mapping of specific components in complex biological samples with high spatial accuracy. In the following, the basic principles of biologically relevant AFM imaging modes as well as the methods of single molecule force spectroscopy (SMFS) and simultaneous topography and recognition imaging (TREC) will be introduced and discussed. Selected experiments

C. Lamprecht · M. Köhler · S. Posch · Y.J. Oh ·
R. Zhu · L.A. Chtcheglova · A. Ebner ·
P. Hinterdorfer (✉)
Institute of Biophysics, Johannes Kepler University
Linz, 4020 Linz, Austria
e-mail: peter.hinterdorfer@jku.at

C. Lamprecht
e-mail: constanze.lamprecht@jku.at

M. Köhler
e-mail: melanie.kohler@uclouvain.be

S. Posch
e-mail: sandra.posch@jku.at

Y.J. Oh
e-mail: yoo_jin.oh@jku.at

R. Zhu
e-mail: rong.zhu@jku.at

A. Ebner
e-mail: andreas.ebner@jku.at

J. Strasser · L.A. Chtcheglova
Center for Advanced Bioanalysis GmbH, 4020 Linz,
Austria
e-mail: juergen.strasser@cbl.at

L.A. Chtcheglova
e-mail: lilia.chtcheglova@jku.at; lilia.chtcheglova@cbl.at

will be presented in more detail to illustrate the combined application of these techniques in the elucidation of questions in molecular biology, pharmaceutical science and the medical field.

Keywords

Atomic force microscopy • Single molecule force spectroscopy • Recognition imaging • Biomedical sensing

4.1 Introduction

Invented in 1986 as a further development of the scanning tunneling microscope (Binnig et al. 1986), the atomic force microscope (AFM) has successfully complemented scanning and transmission electron microscopy as an instrument to visualize cells, biological membranes and molecules from the micrometer scale down to the sub-nanometer range. Surface textures and shapes of biological specimens are visualized as three-dimensional maps with an exceptionally good vertical resolution by scanning the surface with a sharp tip attached to a sensitive cantilever. Where electron microscopy operates under vacuum and requires extensive sample preparation, including steps such as freeze drying, staining or metal coating, AFM measurements can be carried out in aqueous and physiological environments without laborious sample preparation, staining or labeling. Structures and physiological functions of biological samples like membrane proteins, native membranes (Muller et al. 2002; Horber and Miles 2003; Scheuring et al. 2005) and live cells (Jeremic et al. 2003; Zaman et al. 2005; Franz and Muller 2005) are preserved over an extended period of time making it possible to follow structural changes and other dynamic processes at the nanoscale (Rico et al. 2013; Ando et al. 2014; Preiner et al. 2014; Preiner et al. 2015). Next to topographical imaging, the AFM-tip can be used to probe physical properties of biomolecules, cells and tissues (Hinterdorfer et al. 1996; Withers and Aston 2006; Kasas and Dietler 2008; Kurland et al. 2012), and to

measure inter- and intramolecular forces (Lee et al. 1994; Florin et al. 1994; Hinterdorfer et al. 1996) on the single molecule level with outstanding signal-to-noise ratio and pico-Newton resolution (Hugel and Seitz 2001; Alessandrini and Facci 2005; Carvalho and Santos 2012). The use of biomolecule functionalized AFM-tips for observation and quantification of single ligand-receptor interactions, their affinity, rate constants and conformation in comparison with results from ensemble-average techniques is also of particular interest in life sciences and biomedicine (Carvalho and Santos 2012; Kienberger et al. 2006).

In recognition of the rapidly growing literature on AFM life science applications (e.g. Casuso et al. 2011; Dorobantu and Gray 2010; Mueller and Dufrene 2011; Pillet et al. 2014) we will highlight the use and advantages of AFM for biomedical sensing ranging from the elucidation of basic biological processes to pharmaceutical research (Lamprecht et al. 2014) and diagnostics. After a brief introduction to the most relevant modes for topographical imaging of biological samples, a special emphasis will be placed on the principles and theory of single molecule force spectroscopy (SMFS), biochemical functionalization of the AFM-tip, and simultaneous topographical and recognition imaging (TREC) (Raab et al. 1999; Stroh et al. 2004) as the combination of topographical imaging with biosensing AFM-tips. Representative studies on isolated membrane preparations, mammalian cells and bacteria will be presented to emphasize the biomedical sensing capabilities of the AFM toolbox.

4.2 Topographical Imaging of Biological Samples

AFM imaging offers the possibility to nondestructively map the topography of soft biological samples under physiological conditions and yields three-dimensional images of the surface ultra-structure with a high lateral resolution of a few nm. In general, a cantilever carrying a sharp tip is scanned laterally over a specimen. Due to the interaction with features on the surface the cantilever is bent. In most commercial setups a laser beam is focused on the back of the cantilever and reflected onto a four-segment photodiode to monitor the bending and separate the movements of the laser beam in the horizontal and vertical direction. The up- and down deflection leads to a voltage difference which can be directly attributed to changes in the z direction. Lateral movement of the cantilever leads to a shift of the laser beam between the left and the right section of the photodiode. Tip scanning can be performed in numerous different modes, the most basic and biologically relevant ones will be discussed in this section. In addition the issue of imaging conditions and artifacts, as well as some general sample preparation procedures will be addressed.

4.2.1 Basic Imaging Modes for Bio-Application

AFM **contact mode** is the most basic imaging mode in bio-applications. The tip is in permanent contact with the sample surface (Fig. 4.1a). Either the force (constant force mode) acting on the sample or the height (constant height mode) between substrate and cantilever is held constant. The latter mode is applied rather to flat samples and hard surfaces, as high shear forces might damage soft biological samples. The topography image results from the deflection of the cantilever as the tip scans over the surface. In constant force mode, the feedback loop keeps the deflection of the cantilever, i.e. the force between the tip and the sample constant. The topography image is then created from the traveling distance of the z

—piezo element in the scanner. The finite reaction rate of the feedback loop decreases the fastest possible scanning velocity as opposed to constant height scanning, but due to the control of forces, less sample damage can be expected. Using very soft cantilevers with a spring constant in the range of 0.1–0.01 N/m constant force imaging can be well suited even for delicate biological applications. An example of contact mode imaging in buffer solution of a nuclear membrane preparation spanned across a PDMS grid (Rangl et al. 2009) is presented in Fig. 4.2a. Images at higher magnification nicely resolve the embedded individual doughnut-shaped nuclear pore complexes (NPCs) with an outer diameter of ~ 120 nm and inner diameter of ~ 80 nm. Imaging of soft tissues and usually cells requires the restriction to blunt AFM tips (opening semi-angles 18° or 35°), very low loading forces (<1 nN) and reduced scan speed (<1 line/s) to prevent penetration of the cell membrane by the tip.

To overcome the contact mode speed issue and to eliminate frictional forces, intermittent contact or **tapping mode** was developed (Hansma et al. 1994). A piezoelectric actuator excites the cantilever to oscillate and hence the tip just briefly taps the sample surface, which minimizes the lateral forces on the sample (Fig. 4.1b). The feedback loop keeps the oscillating amplitude constant by adjusting the distance and thus the acting force between tip and sample. The topography image is generated from the scanner's z—motion. A related dynamic imaging mode is **MAC mode** imaging, where a magnetically coated cantilever is excited to vibrate in a magnetic field. In liquid, the amplitude of the cantilevers is, in contrast to tapping mode, well defined by the driven damped harmonic oscillator model and can be easily determined. MAC mode is often the method of choice for the almost unlimited biological applications. Figure 4.2b shows some examples for Mac mode imaging DNA strands (i Mahut et al. 2012), bacteria (ii, iii Oh et al. 2012) and living blood platelets (iv Posch et al. 2013). Although dynamic imaging modes induce far less shear force compared to contact mode, the viscoelastic

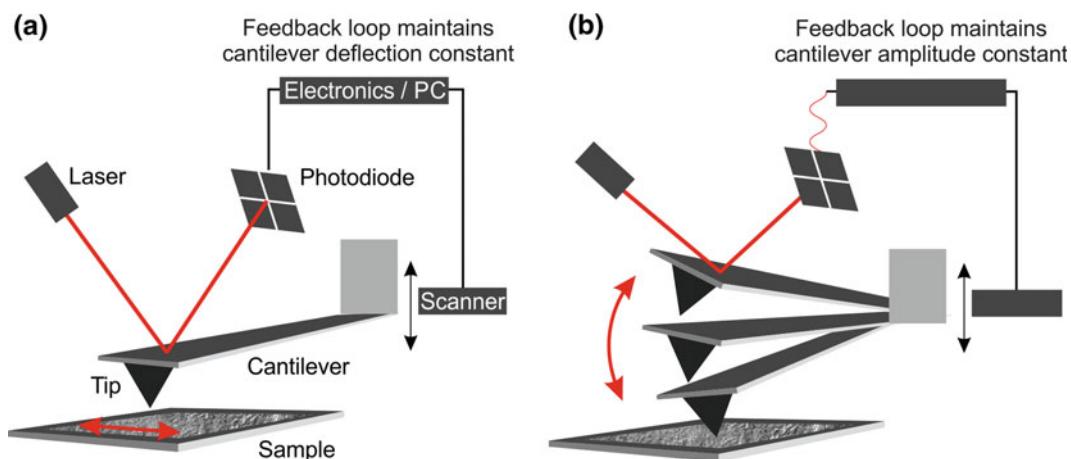


Fig. 4.1 Schematic illustration of the working principle of AFM. A sample is probed by an ultra-sharp tip mounted on a cantilever that scans laterally over the sample surface. Deflections of the cantilever are reported by a laser beam that is reflected from the back of the cantilever onto a photodiode. The deflection signal is fed into a feedback loop that controls the piezo (scanner) movement by which the tip is lifted from or lowered to the surface in order to keep the force of the tip–sample interaction constant. A topographic image of the sample is then reconstructed from the piezo-movement. **a** In contact mode the deflection is kept at a constant value and **b** in tapping and MAC mode the amplitude of the oscillation cantilever is maintained constant.

response and dynamic nature of soft samples such as live cells essentially reduce the resolution to approximately 50–100 nm.

A more recent imaging technique is higher harmonics imaging (Gracia 2010; Garcia and Herruzo 2012), which achieves a higher lateral resolution compared to conventional AFM imaging. In liquid, the amplitude of the second harmonic was found to be an order of magnitude higher than in air and thus more sensitive to local variations on the sample surface. Higher harmonic contributions in the movement of an oscillating cantilever are generated by nonlinear tip–sample interactions, yielding additional information on structure and physical properties such as sample stiffness. The second harmonic amplitude image—as well as the topography image can be recorded simultaneously during scanning in tapping mode and under physiological conditions. With this technique, second harmonic amplitude images of bacterial surface layers and a layer of human rhinovirus serotype 2 (HRV2, Fig. 4.2c) showed detailed substructures with a resolution of about 0.5 nm, which were not visible in topographical images (Preiner et al. 2007). Later approaches used multi-harmonic

AFM to map local material properties of intact and disrupted virions *in vitro* at high resolution (Cartagena et al. 2013) and live cells (Dulebo et al. 2009).

4.2.2 Sample Preparation and Imaging Conditions

A crucial requirement for successful and reliable biological AFM experiments is proper sample preparation. Although there are many standardized nondestructive approaches for sample immobilization, they are usually only applicable for a specific system and need to be adjusted to the situation. Substrates like mica, glass and silicon oxide turned out to be the ideal support. The atomically flat mica is suitable for immobilizing e.g. DNA, single proteins or supported lipid films. Glass cover slips are rougher than mica surfaces and are thus mainly used for imaging cells and other large structures. A good alternative for mica can be silicon oxide wafers, which are, more expensive but have a much smoother surface than glass (El Kirat et al. 2005).

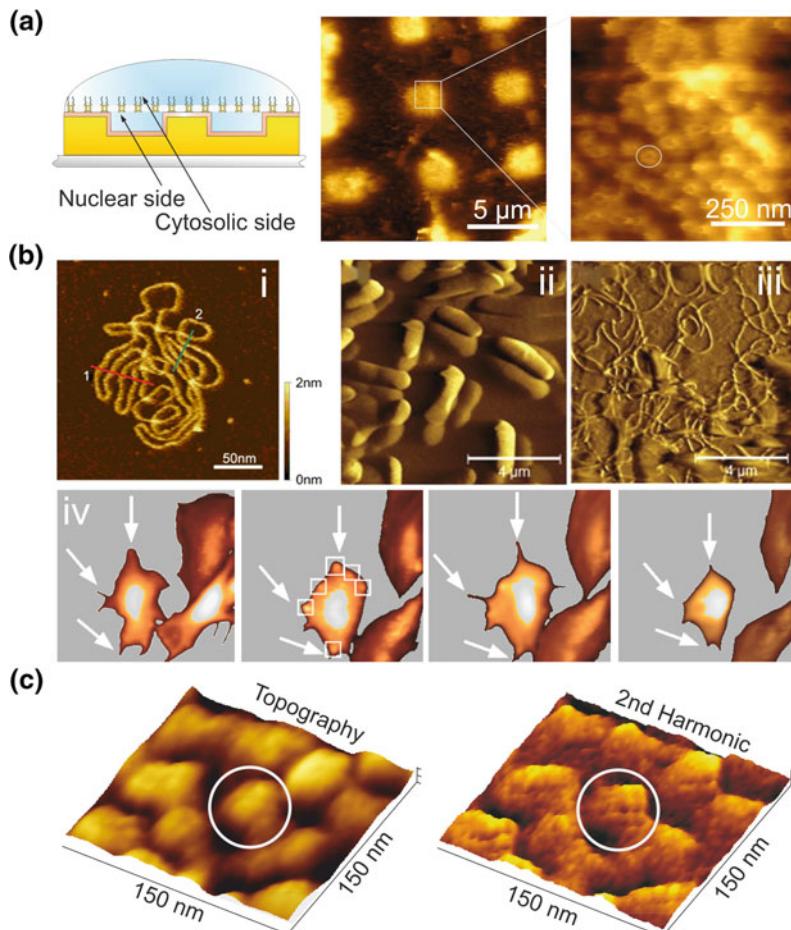


Fig. 4.2 Basic topographical imaging modes for bio-applications. **a** Contact mode imaging of nuclear envelope (NE) patches on micro-structured PDMS (left, yellow) (Rangl et al. 2009). The thickness of the PDMS grid is about ~1 mm with plateaus and wells of 3 μm in width and 0.7 μm in height, which form micro-chambers into which the nuclear baskets of the nuclear pore can extend. An overview scan image of the membrane preparation performed in contact mode shows the elevated plateaus and lower regions where the membrane spans the micro-wells. A zoom-in on a plateau region reveals the doughnut shaped nuclear pore complexes (white circle). **b** Tapping and MAC mode imaging examples of *i* a relaxed plasmid DNA cluster on 3-aminopropyl-modified mica (Mahut et al. 2012), *ii*, *iii* amplitude images of wild type *E. coli* bacteria on a gelatin-coated mica surface (*ii*) and flagella-like nanowire structures of CsgA (-) mutant bacteria after removal of the cells with an AFM tip (Oh et al. 2012), and *iv* visualization of shape changes of living blood platelets during activation where a series of filopodia extrusion and blebbing steps can be observed (Posch et al. 2013) before the cells shrink and detach (image size: 15 μm , color range: 0–800 nm). **c** Higher harmonics imaging of a dense layer of human rhinovirus serotype 2 (HRV2) on mica in buffer. The simultaneously recorded second harmonic amplitude image reveals substructures of the viral capsides (circle) that remain hidden in the topographical image (scan size: 350 \times 350 nm^2 ; color range: 0–13 nm, 0–0.35 V).

For the immobilization of biological samples either covalent or non-covalent (adsorption) protocols can be used. Nonspecific adsorption relies on non-covalent bonds such as hydrophobic interactions, stacking, hydrogen bonds, and

electrostatic attraction. Biomolecules can be adsorbed on mica or covalently bound to a chemically modified surface (<http://www.jku.at/biophysics/content/e257042>). Supported lipid bilayers are often prepared by the fusion of lipid

vesicles on solid surfaces or by the Langmuir–Blodgett technique (Ulman 2013). Animal cells have the ability to adhere to solid supports. Coating the surface with adhesive proteins may enhance cell immobilization. Additional chemical fixation with e.g. glutaraldehyde can prevent cell damage or detachment. Microbial cells, such as bacteria, yeast and fungi do not have the ability to spread over a surface. Thus, they have to be bound covalently to the substrate (Ebner et al. 2008a; Oh et al. 2016), physically adsorbed (Doktycz et al. 2003) or mechanically immobilized in e.g. a porous membrane (Dufrene 2002). In general, it is also highly recommended to investigate the effect of pH and ionic strength of the sample buffer on the image quality (Dufrene 2002). A more detailed overview of sample preparation protocols and imaging conditions may be found elsewhere (e.g. El Kirat et al. 2005).

4.2.3 Imaging Artifacts

AFM setups are commonly very well shielded against external vibrations (acoustical, mechanical) or electronic malfunctions and AFM imaging of biological specimen is quite straight forward. Nevertheless, imaging artifacts may occur rather frequently and lead to misinterpretation of data, if not recognized. To achieve a high resolution image, all scan parameters must be controlled conscientiously. Setting feedback parameters (i.e. proportional and integral gain) too high during imaging is a common mistake (Fig. 4.3a). Forces acting between the AFM probe and the sample during imaging have to be carefully monitored. Changes of pH and increasing salt concentrations due to evaporation during long term imaging need to be considered. Temperature drifts especially in heated sample chambers influence imaging stability and the effective force might increase. Large forces can dramatically reduce the image resolution, and apart from causing molecular damage, may lead to tip contaminations, to name a few complications.

Apart from the imaging conditions the validity of AFM images is significantly influenced by the cantilever tip as well as by the scanner (Fig. 4.3b,

c). If the tip curvature radius, which is typically between 5 and 10 nm, is comparable or bigger than the structures on the sample surface, the lateral size of the structures appears overestimated. In this case, the FWHM (full width at half maximum) value of the sample features can be determined by analyzing the cross-section of AFM topography images. The following relation connects the FWHM value with the actual diameter d of the investigated structures, by implying the tip radius R : $FWHM = 2\sqrt{d(\frac{R+d}{4})}$. In general, to achieve high-resolution AFM images of biological samples, it is important to use a sharp tip (Fig. 4.3b, iii) and a soft cantilever in order to nondestructively map the sample features on the surface (Raposo et al. 2007).

Also the piezo electric scanner may be directly a source of errors. Especially when imaging large areas, i.e. when the scanner has to expand far, deviations from the linear expansion behavior can occur (Fig. 4.3c). To avoid biased images of the sample topography, the scanner should be perpendicular to the surface. Hysteresis in the response of the piezo element ('piezo creep') may cause edge overshoots and together with thermal scanner drift may cause image distortion. These effects are often observed when zooming-in.

4.3 Single Molecule Force Spectroscopy (SMFS)

Measurements of interaction forces traditionally rely on ensemble techniques such as shear flow detachment (Bongrand et al. 1988) and the surface force apparatus (Israelachvili 2011). Single molecule techniques, on the other hand, can reveal molecular processes, pathways and rare events that would be masked in ensemble measurements by analyzing individual events in large heterogeneous populations. Though they typically require more effort and time to reliably generate high-quality data, direct measurements can provide invaluable insight into complex systems. In AFM based force spectroscopy the cantilever is used as molecular sensor to quantify interaction forces between the tip of the

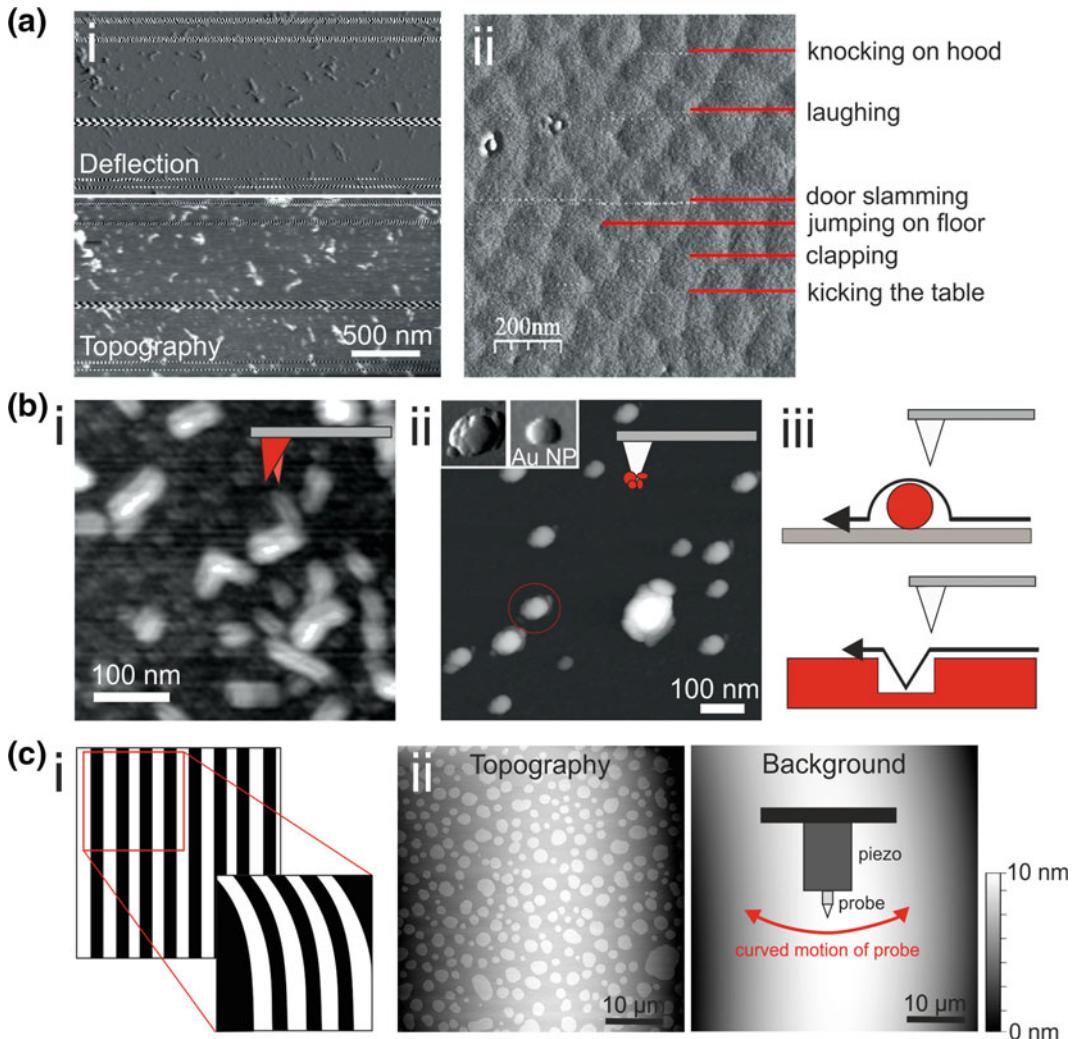


Fig. 4.3 Common imaging artifacts. **a** *i* The feedback parameters (proportional and integral gain) are set too high leading to z-oscillation of the piezo electric scanner that show as ‘zebra-stripe’-like artifacts. Scanner oscillations appear stronger in the deflection and become visible in the topography, when they are too pronounced. *ii* Perturbations in the acoustical and mechanical frequency range (e.g. vibration) are picked up by the AFM image, if setups are not well enough shielded (by e.g. active vibration isolation table, acoustic enclosure). **b** AFM-tip associated artifacts commonly appear in the form of *i* doubled structure in the topography image, due to a double tip, or *ii* in shape of a repetitive pattern with a preferential orientation, which is caused by tip contaminations. *iii* The tip geometry radius determines the lateral resolution in AFM imaging and the tip geometry causes a broadening effect and occasionally incorrect recording of the sample features. **c** Artifacts caused by the piezo electric scanner are foremost *i* a so-called scanner creep after zoom-in due to adaption of the piezo crystal to the new set voltage, resulting in stretching and compression of features at the beginning of the new image, and *ii* the nonlinearity of a tube scanner in z-direction, that accounts for a nonlinear (mostly parabolic) background, which is commonly removed by post-processing of images (flattening). The topographical image shows an overview scan of a phase separated supported lipid bilayer (DOPC:SM:Chol 2:2:1). The liquid ordered phase (*round patches*) is about 1 nm higher than the surrounding liquid disordered phase.

cantilever and the substrate. The deflection of the AFM cantilever Δx relates to the force F acting on the cantilever through Hooke’s law

$F = k \cdot \Delta x$, with the spring constant k of the cantilever. The force resolution of the AFM, $\Delta F = (k_B T \cdot k)^{1/2}$, is limited by the thermal

noise of the cantilever which, in turn, is determined by its spring constant. With commercially available (bio-)cantilevers interaction forces between single molecules in the range of only a few pico-Newton (pN) (Kienberger et al. 2006) and adhesion forces between a single cell and a substrate in the order of a few nano-Newton (nN) (Friedrichs et al. 2013) are routinely resolved. For single molecule force spectroscopy (SMFS) the small size of AFM tips between 2 and 50 nm provides unprecedented high lateral resolution.

Over the past 20 years the technique has been refined constantly and become a versatile tool for characterizing molecular interactions (Lee et al. 1994; Florin et al. 1994; Hinterdorfer et al. 1996; Ott et al. 2016). During this time various single molecular pairs, such as avidin/streptavidin and biotin, antibodies and their antigens, proteins and specialized aptamers or DNA and DNA-binding proteins, have been studied with great success (Bizzarri and Cannistraro 2010; Ritzefeld et al. 2013). Besides the detection of intermolecular forces, AFM also shows great potential in measuring intramolecular forces acting within molecules such as mechanical proteins, membrane proteins, DNA or polysaccharides (Ott et al. 2016; Ritzefeld et al. 2013; Marszalek and Dufrêne 2012; Engel and Gaub 2008; Whited and Park 2014). In these experiments, the molecule is clamped between the tip and the surface and its viscoelastic properties as well as unfolding and refolding pathways are studied by force-distance cycles (Marszalek and Dufrêne 2012).

SMFS is not only limited to the reporting of interaction forces either, and can readily be used to obtain interaction rate constants of molecular association and dissociation as well as length scales and relative heights of energy barriers traversing the free energy surface. To this end, the rate at which a pulling force is applied to the bond is varied, revealing a dependence of this “loading rate” and kinetic and thermodynamic parameters. This technique of dynamic force spectroscopy is explained in greater in this section.

4.3.1 Basic Principle of SMFS

Basic single molecule force spectroscopy (SMFS) requires immobilization of one of the binding partners to the apex of the AFM tip and the other to a hard surface such as mica or glass. The functionalized tip is then repeatedly brought into contact with and retracted from the substrate, while recording the cantilever deflection, i.e. the force acting on the cantilever as illustrated in Fig. 4.4a. During most of the approach phase, when the distance of the tip to the surface is sufficiently large (1–3), the cantilever deflection remains zero. Upon contact (4) the cantilever bends upwards due to a repulsive force that increases linearly (4–5) as the tip is pushed further into the surface. After a pre-defined contact force (5) has been reached, the cantilever movement is reversed and the deflection relaxes (5–6). If binding had occurred, the cantilever will bend downwards as it is further retracted from the surface (6–7). In case the binding partners were tethered through flexible cross-linkers (Kienberger et al. 2000; Hinterdorfer et al. 2000), the shape of the attractive force-distance profile is nonlinear. It then depends on the elastic properties of the cantilever (commonly modeled as a harmonic spring), the cross-linker, which exhibits parabolic-like characteristics, reflecting an increase of the spring constant during extension and the bond itself. When the ramping force reaches a critical value the binding partners dissociate and the cantilever jumps back into the resting position (7). The force value at dissociation is termed unbinding force and is quantified by the vertical distance from the critical bond force to the baseline (7–2). Similarly, the unbinding length is given by the horizontal distance from the contact point to the extension at rupture (6–7, horizontal component). In absence of interaction between tip bound ligand and surface receptors the retract curve will closely follow the approach curve. A typical SMFS experiment at a given pulling velocity comprises 1000 so called force-distance cycles (FDC) resulting in a distribution of unbinding forces, which can be either plotted as a histogram or as a continuous probability density function

(PDF) (Fig. 4.4b). The maximum of the distribution is termed most probable unbinding force F^* . The specificity of the observed (un-)binding event is usually demonstrated by blocking the interaction either through addition of free ligands

to mask receptor sites on the surface, or by saturating the binding partner on the AFM-tip, drastically reducing the binding probability (i.e. the number of unbinding events observed per 1000 cycles) (Fig. 4.4c).

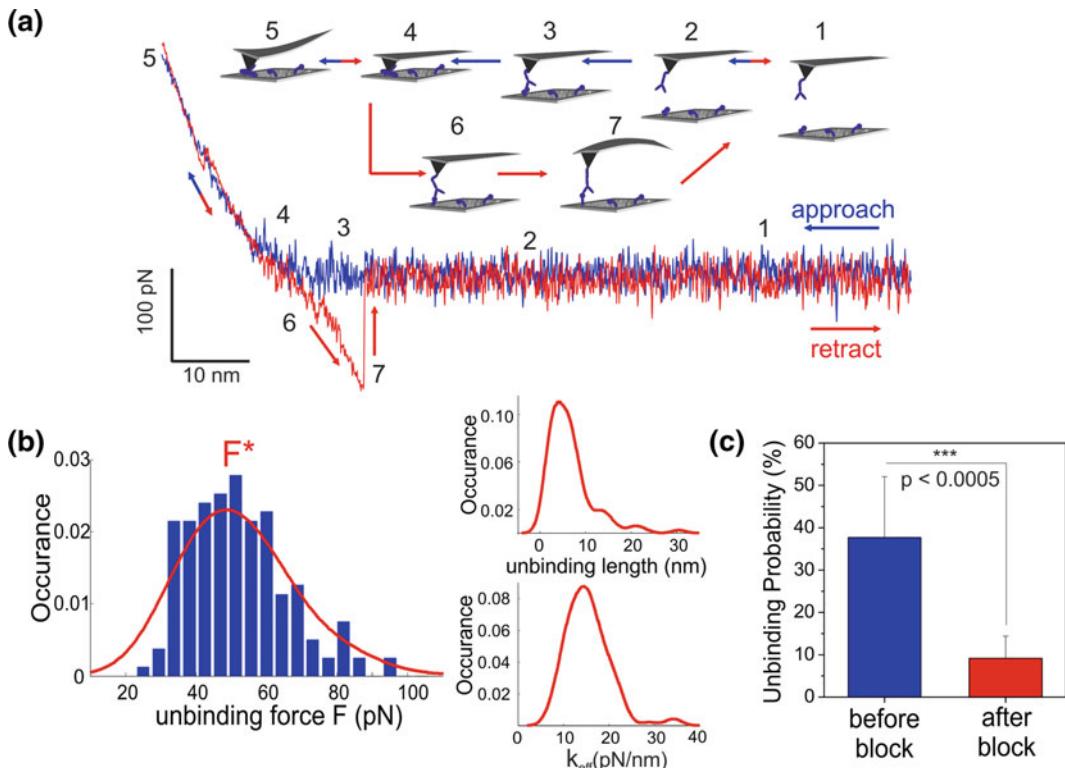


Fig. 4.4 Single molecule force spectroscopy (SMFS). **a** A force-distance cycle (FDC) (e.g. sweep-rate of 200 nm/s) with a ligand functionalized AFM-tip and a dense layer of receptor molecules immobilized on mica. Binding occurs towards the end of the approach phase (1–5) when the tip comes into contact (4) and is pressed into the surface. Under further approach the tip bends upwards exerting a linear increasing force on the sample (5). During tip retraction the bending relaxes again. When a ligand-receptor bond has formed, the cantilever progressively bends downwards (6) visible as a characteristic parabolic slope in the retract curve reflecting the extension of the distensible cross-linker ligand-receptor connection. The force increases until rupture (unbinding) of the ligand-receptor bond occurs (7). After rupture the retract curve closely follows the approach curve. **b** Statistical analysis of a great number of unbinding events at a given sweep-rate and indentation force yields distributions of unbinding force F , unbinding length l and effective spring constant k_{eff} that can be displayed a histogram (blue) or probability density functions (PDF) (red). In a PDF the data are weighed by their uncertainty and presented in a continuous curve that offers a measure of reliability not provided by simple binning. This eases determination of the most probable values for each parameter, e.g. F^* as the most probable unbinding force. **c** The specificity of the ligand-receptor interaction needs to be verified by a block-experiment where typically one of the binding partners is saturated by free agonist molecules. This will result in a significant decrease of the number of detected unbinding events, i.e. the binding probability. The statistical significance can be rated e.g. by an unpaired two-sample t-test.

4.3.2 Theory of Single Molecule Binding and Force Spectroscopy

Molecular recognition is mediated by a multitude of non-covalent interactions, the energy of which is only slightly higher than thermal energy. Due to the power law dependence of these interactions on distance, the attractive forces between non-covalently interacting molecules are extremely short-ranged. A close geometrical and chemical fit within the binding interface is therefore a prerequisite for efficient association. The weak bonds that govern molecular cohesion are believed to be formed in a spatially and temporarily correlated fashion. Protein binding often involves structural rearrangements, either localized or global, that commonly bear functional significance by modulating the activity of the interaction partners. Signaling pathways, enzyme activity, and the activation and inactivation of genes all depend on conformational changes induced by ligand binding.

The strength of binding is usually given by the binding energy ΔG , which amounts to the free energy difference between the bound and the free state, and which can readily be determined by ensemble measurements. ΔG determines the ratio of bound complexes $[RL]$ to the product of free reactants $[R][L]$ at equilibrium and is related to the equilibrium dissociation constant K_D through $\Delta G = -RT \ln(K_D)$, where R is the gas constant and T the temperature in Kelvin. K_D itself is related to the empirical association (k_{on}) and dissociation (k_{off}) rate constants through $K_D = k_{off}/k_{on}$. In order to get an estimate for the interaction forces F, from binding energies ΔG , the depth of the binding pocket may be used as a characteristic length scale l. Using typical values of $\Delta G = 20 \text{ k}_\text{B}T$ and $l = 0.5 \text{ nm}$, an order of magnitude estimate of $F (= \Delta G/l) \sim 170 \text{ pN}$ is obtained for the binding strength of a single molecular pair. Classical mechanics describes bond strength as the gradient in energy along the direction of separation. Unbinding therefore occurs when the applied force exceeds the steepest gradient in energy. This purely

mechanical description of molecular bonds, however, does not provide insights into the microscopic determinants of bond formation and rupture.

Bonds will resist detachment when a pulling force is applied. Notably, the unbinding force may approach and even exceed the adiabatic limit given by the steepest energy gradient of the interaction potential if rupture occurs in less time than needed for diffusive relaxation (10^{-10} – 10^{-9} s for biomolecules in viscous aqueous medium) and friction effects become dominant (Grubmüller et al. 1996). Therefore, unbinding forces do not resemble unitary values and the dynamics of the experiment critically affects the measured bond strengths. At the time scale of AFM experiments (milliseconds to seconds), thermal impulses govern the unbinding process. In the thermal activation model, the lifetime of a molecular complex in solution is described by a Boltzmann ansatz, $\tau(0) = \tau_{osc} \cdot \exp(E_b/k_\text{B}T)$ (Bell 1978), where τ_{osc} is the inverse of the natural oscillation frequency and E_b is the height of the energy barrier for dissociation. This gives a simple Arrhenius dependency of dissociation rate on barrier height.

Force acting on a complex deforms the interaction free energy landscape and lowers barriers for dissociation (Fig. 4.5a). As a result, bond lifetime is shortened. The lifetime $\tau(F)$ of a bond loaded with a constant force F is given by: $\tau(F) = \tau_{osc} \cdot \exp(E_b - x_\beta F)k_\text{B}T$ (Bell 1978), where x_β marks the thermally averaged projection of the energy barrier along the direction of the force.

4.3.2.1 The Bell-Evans Model

A detailed analysis of the relation between bond strength and lifetime was performed by Evans and Ritchie (1997), using Kramers' theory for over-damped kinetics. For a sharp barrier, the lifetime $\tau(F)$ of a bond subjected to a constant force F relates to its characteristic lifetime, $\tau(0)$, according to: $\tau(F) = \tau(0)\exp(-x_\beta F/k_\text{B}T)$ (Hinterdorfer et al. 1996). However, in most pulling experiments the applied force is not constant but varies with time, and the measured

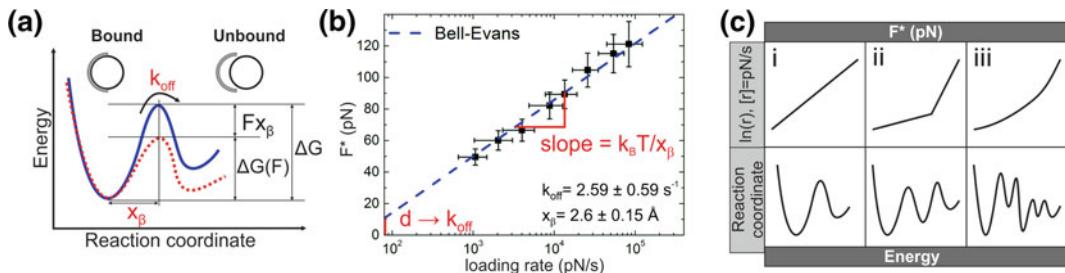


Fig. 4.5 Theory and interpretation of single molecule binding and force spectroscopy. **a** In a simple model of molecular dissociation over a single sharp energy barrier, a constant force F lowers the barrier's height proportionally. This gives rise to a characteristic length scale x_β that is interpreted as the distance of the energy barrier from the energy minimum along the projection of the force. The barrier is crossed with the characteristic unbinding rate k_{off} . After (Bizzarri and Cannistraro 2010). **b** In the Bell-Evans model the most probable unbinding force F^* scales with the natural logarithm of the loading rate. The intersect with the force axis yields the characteristic unbinding rate k_{off} and the slope reveals the characteristic length scale x_β . **c** In the classical Bell-Evans interpretation linear segments in the logarithmic plot of the loading rate dependence reflect multiple independent barriers along the energy profile of a biological bond (*ii*). For non-linear curved spectra very complex energy landscapes may result from this interpretation. After (Bizzarri and Cannistraro 2010).

bond strength depends on the loading rate (Grubmüller et al. 1996; Evans and Ritchie 1997; Izrailev et al. 1997). In accordance with this, experimentally measured unbinding forces do not assume unitary values but rather vary with both pulling velocity (Baumgartner et al. 2000; Fritz et al. 1998) and cantilever spring constant (Lee et al. 1994). Nevertheless, contributions arising from thermal activation manifest mostly near the point of detachment. Therefore, the change of force with time, i.e. the loading rate r ($=dF/dt$), can be derived from the product of the pulling velocity and the effective spring constant at the end of the force curve, just before unbinding occurs.

Dynamic force spectroscopy (DFS) exploits this dependence of bond strength on the loading rate to obtain detailed insights into intra- and intermolecular interactions. The dependence of the rupture force on the loading rate (force spectrum), in the thermally activated regime was first derived by Evans and Ritchie (1997) and described further by Strunz et al. (2000). Rupture itself is a stochastic process and the likelihood of bond survival is expressed in the master equation as a time-dependent probability $N(t)$ to be in the bound state under a steady ramp of force, namely $dN(t)/dt = -k_{\text{off}}(rt)N(t)$ (Strunz et al. 2000). This results in a distribution of unbinding forces

$P(F)$ parameterized by the loading rate (Evans and Ritchie 1997; Strunz et al. 2000; Evans and Ritchie 1999). The most probable force for unbinding F^* , given by the maximum of the distribution (Fig. 4.4b), relates to the loading rate through $F^* = F_\beta \ln(r) k_{\text{off}}^{-1}/F_\beta$, where the force scale F_β is set by the ratio of thermal energy to x_β (Fig. 4.5a) (Evans and Ritchie 1997; Strunz et al. 2000). Thus, the unbinding force scales linearly with the logarithm of the loading rate. This phenomenological approach to relating data from pulling experiments to the innate kinetic and thermodynamic parameters of the probed interaction is commonly referred to as Bell-Evans model.

For a single barrier, the Bell-Evans model would give rise to a simple, linear force spectrum F^* versus $\ln(r)$ (Fig. 4.5c, i). In cases where the escape path traverses several barriers, the curve will follow a sequence of linear regimes, each marking a particular barrier (Evans and Ritchie 1997; Evans and Ritchie 1999; Merkel et al. 1999) (Fig. 4.5c, ii). Transition from one regime to the other is associated with an abrupt change of slope determined by the characteristic barrier length scale and signifying that a crossover between barriers has occurred. The slopes of the force vs. loading rate curves contain information about the length scale x_β of prominent energy

barriers along the force-driven dissociation pathway, which may be related to the depth of the binding pocket of the interaction (Fig. 4.5a, b) (Baumgartner et al. 2000).

4.3.2.2 Limitations of the Bell-Evans Model and New Approaches to DFS Data Analysis

The predicted linear dependence of the unbinding force on the logarithm of the loading rate r holds well when the barriers are stationary with force, as found by a large number of unbinding and unfolding experiments. However, if the position of the transition state is expected to vary along the reaction coordinate with the force, as for example when the curvature at the top of the barrier is small, this simple dependence gives way to more complex forms. Schlierf and Rief (2006) used a Kramers' diffusion model to calculate the probability force distributions when the barrier cannot be assumed stationary. A failure to fit force distributions at high loading rates using a Bell-Evans model was also reported in by Neuert et al. (2006) for the interaction of digoxigenin and anti-digoxigenin. Poor matches were observed in the crossover region between the two linear regimes of the force spectrum as well.

Additionally, fitting each regime in the force spectrum with a separate function according to the Bell-Evans model can lead to very complex underlying energy landscapes even for simple experimental systems. For curved force spectra (Fig. 4.5c, iii) this interpretation may result in an arbitrary number of barriers, each characterized by at least two fit parameters, as well as values for x_β several orders of magnitude below 1 Å (Bizzarri and Cannistraro 2010; Friddle et al. 2012). Lastly, the phenomenological model described above holds for linear energy surfaces with $\Delta G \rightarrow \infty$, a simplification which has been shown to be applicable to a large number of biomolecular interactions. For some systems, however, different energy barrier and energy well geometries might better correspond to experimental data (Dudko et al. 2006).

Alternative models which predict inherently nonlinear force spectra have emerged in recent years, extending the framework laid down by Evans and Ritchie. The newer models described below include additional fit parameters, most notably the height of the free energy barrier, which does away with the necessity of estimating ΔG using the Arrhenius equation. Dudko et al. (2006) modified the Bell-Evans model by accounting for the effects of the applied force on the position of the transition state. In their treatment they find that a dependence of the mean rupture force on the loading rate r in the form $(\ln r)^\nu$, with the exponential constant ν allowing for either cusp-like ($\nu = 1/2$) or linear-cubic ($\nu = 2/3$) free-energy surfaces, corresponds more closely to experimental force spectra than a linear approach ($\nu = 1$, corresponds to the Bell-Evans model). They also find the distribution of unbinding forces to be asymmetrical, as indeed observed many times. In a subsequent publication (Dudko et al. 2008) introduced a method based on transforming rupture force distributions from DFS into force-dependent lifetimes akin to those obtained from constant-force experiments, and show how the molecular linker used in the experiment affects the loading rate.

The model put forward by Friddle et al. supports the notion of reversible rebinding at low loading rates when using soft springs (Friddle et al. 2012), a process which is not taken into account by the other approaches to DFS presented here. The Friddle model introduces two distinct regimes during molecular unbinding: An equilibrium regime at low loading rates which supports cycles of reversible unbinding and rebinding, and a kinetic regime at higher loading rates where unbinding is irreversible. In addition to k_{off} and x_β , an additional fit parameter is introduced, the equilibrium force F_{eq} . This is the force where dissociation and association rates are in balance. F_{eq} is the lowest force at which a bond breaks for a given pulling apparatus geometry and pulling speed. This also means that at forces below F_{eq} , no complete dissociation will occur, which marks an important difference between a free ligand-receptor pair, where

spontaneous unbinding is possible, and a system where the interaction partners are tethered to force transducers. An interesting application of this model was presented by Alsteens et al. (2015), who used it to analyze DFS data of the protease-activated receptor-1 (PAR1) in native membranes and its ligands. Whereas the majority of DFS models operate on most-probable rupture forces obtained from pooling many hundreds of measurements, Alsteens and colleagues present a method to apply the work of Friddle et al. (2012) to data clouds of individual unbinding events. In addition to these new models for the analysis of DFS data, several modifications of the Bell-Evans model itself were proposed over the years (Bizzarri and Cannistraro 2010; Sekatskii et al. 2013; Walton et al. 2008). These reflect new findings about parameters influencing experiments, such as the stiffness of the force transducer (Walton et al. 2008) or the range of loading rates in which this phenomenological approach is applicable (Sekatskii et al. 2013).

4.3.3 Data Evaluation and Determination of Interaction Rate Constants

A comprehensive guide to modern analysis of DFS is given by Bizzarri and Cannistraro (2010) and Noy and Friddle (2013), whereas Hane et al. (2014) have compared the three dominant models in the field in great detail. The choice of model relies first and foremost on the obtained force spectrum: Non-linear alternatives to the Bell-Evans model should only be used if the forces measured experimentally actually do not scale linearly with $\ln(r)$, since fitting a linear dependence with a non-linear function, while possible, may result in fit parameters with little to no meaning.

While the dissociation rate constant k_{off} is a fit parameter in DFS models (Fig. 4.5), quantification of the association rate constant k_{on} has to be determined via the interaction time that is required to achieve half of the maximum possible binding probability ($t_{1/2}$). This can be obtained

from experiments where the encounter time between receptor and ligand is varied over a broad range (Baumgartner et al. 2000). Given that the concentration of ligand molecules on the tip available for interaction with the surface-bound receptors c_{eff} is known, the association rate constant can be derived from $k_{on} = t_{0.5}^{-1} c_{eff}^{-1}$. Determination of the effective ligand concentration requires knowledge of the effective volume V_{eff} explored by the tip-tethered ligand which, in turn, depends on the tether length. Therefore, only order of magnitude estimates of k_{on} can be gained from such measurements (Baumgartner et al. 2000).

Note that the application of force acts to select the dissociation path. Since the kinetics of reactions are pathway dependent, such a selection implies that kinetic parameters extracted from force-probe experiments may differ from those obtained from assays conducted in the absence of external force. At extremely fast complexation and decomplexation kinetics the forces can be independent of the loading rate, indicating that the experiments were carried out under thermodynamic equilibrium (Auletta et al. 2004). These concerns notwithstanding, SMFS has been demonstrated to be a versatile tool enabling the estimation of association and dissociation rate constants, as well as the measurement of structural parameters of the binding pocket (Bizzarri and Cannistraro 2010; Whited and Park 2014).

4.3.4 AFM-Tip Chemistry

For SMFS experiments the usually inert AFM tip has to be converted into a (mono-) molecular sensor by attaching one or a few ligand molecules to the apex of the tip. Based on the tip surface material, i.e. silicon nitride or gold, different approaches are possible. While the use of gold coated tips allows functionalization based on self-assembled monolayers (SAMs) of thiolated linker molecules, the common silicon or silicon nitride tips have to be chemically modified in the first step (e.g. by amino-functionalization protocols). More simple strategies like adsorption or chemisorptions may provide for efficient coupling

under suitable circumstances. A more detailed description can be found in Ebner et al. (2008a). Physisorption, meaning the nonspecific adsorption of molecules, which relies on non-covalent bonds including hydrophobic interactions, stacking, hydrogen bonds, and electrostatic attraction, offers an easy but effective coupling strategy. However, AFM tip functionalization requires very firm coupling. Though this might be realized by physisorption (e.g. with the help of cross-linking Fritz et al. 1998; Chen and Moy 2000), only covalent attachment protocols will yield reliable and reproducible coupling results that render the tip non-adhesive and introduce a distensible linker between ligand and AFM tip. Figure 4.6 illustrates common strategies for functionalization of commercial silicon/silicon nitride and cold coated AFM tips.

4.3.4.1 Tip Functionalization of Silicon or Silicon Nitride Tips

The most common approach to make inert silicon and silicon nitride tips chemically addressable is the amino-functionalization (Fig. 4.6a). The gas phase protocol using aminopropyl-triethoxysilane (APTES) as well as the liquid phase method employing ethanolamine-hydrochloride in DMSO are frequently used. Both yield sufficient reactive sites ($\sim 1000\text{--}2000$ amino-groups/ μm^2) while avoiding the formation of sponge like multilayers, which are not suitable for force spectroscopy experiments (Ebner et al. 2007a). A good understanding of the chemical process helps to perform successful amino-functionalization: Under ambient conditions the surface of silicon and silicon nitride tips has a silicon-oxide layer with high numbers of silanol ($-\text{Si}-\text{OH}$) groups. APTES, as the most commonly used amino-silanization reagent, reacts with these silanol groups. The reaction itself has to be performed in an argon flooded desiccator to control the amount of water. The AFM cantilever chips are thoroughly cleaned with chloroform and might be treated by a mild plasma before placing them into the desiccator. In a six liter argon filled desiccator 30 μl APTES and 10 μl triethylamine as catalyst are allowed to react for 120 min followed by careful purging with argon

gas. A two days “curing” phase at room temperature ensures sufficiently strong binding of the generated APTES layer on the cantilever. The amino-functionalization protocol based on ethanol-hydrochloride is a simpler alternative to gas phase APTES coating, but results in less well defined chemical coating. In a nutshell, ethanolamine-hydrochloride is dissolved in DMSO, molecular sieves are added and the cleaned tips are incubated overnight followed by a careful cleaning procedure (Riener et al. 2003).

In the next step a heterobifunctional cross-linker is bound to the amino residues on the tip. Polyethylenglycol (PEG) consisting of 10–30 ethyleneglycol units turned out to be best suited, due to the properties of this polymer. PEG itself is an inert, water-soluble, and distensible polymer with the advantage to be protein repulsive. In addition, the risk of denaturation of the bio-ligands is markedly lowered compared to coupling with a short PEG-linker. In very special cases the linker already contains the ligand molecule itself, e.g. PEG-biotin (Riener et al. 2003) or fluorescein (Ros et al. 1998). These molecules can be bound as part of the linker, since their structure is stable under typical coupling conditions like chloroform as solvent and the application of triethylamine as catalyst.

Most other ligands—usually proteins—require aqueous coupling conditions. Proteins are often coupled to the PEG cross-linker via their lysine residues. NHS-PEG-aldehyde (Ebner et al. 2007b) or NHS-PEG-Acetal (Wildling et al. 2005) linkers are both well-suited, though the former is slightly prone to loop formation. Ligands with a free thiol at the end (e.g. cysteine mutants) may either be tethered by NHS-PEG-PDP linkers (Kamruzzahan et al. 2006) with the advantage that the bond can be cleaved, or by a NHS-PEG-maleimide linker. A very general strategy is realized by the use of the NHS-PEG-benzophenone linker (Ebner et al. 2008a). After binding the linker via its NHS end to the amino-functionalized tips, benzophenone can be activated by UV-irradiation (366 nm) and allows stable covalent bond formation with aliphatic carbons. Admittedly, the above mentioned linkers do not allow for site directed binding.

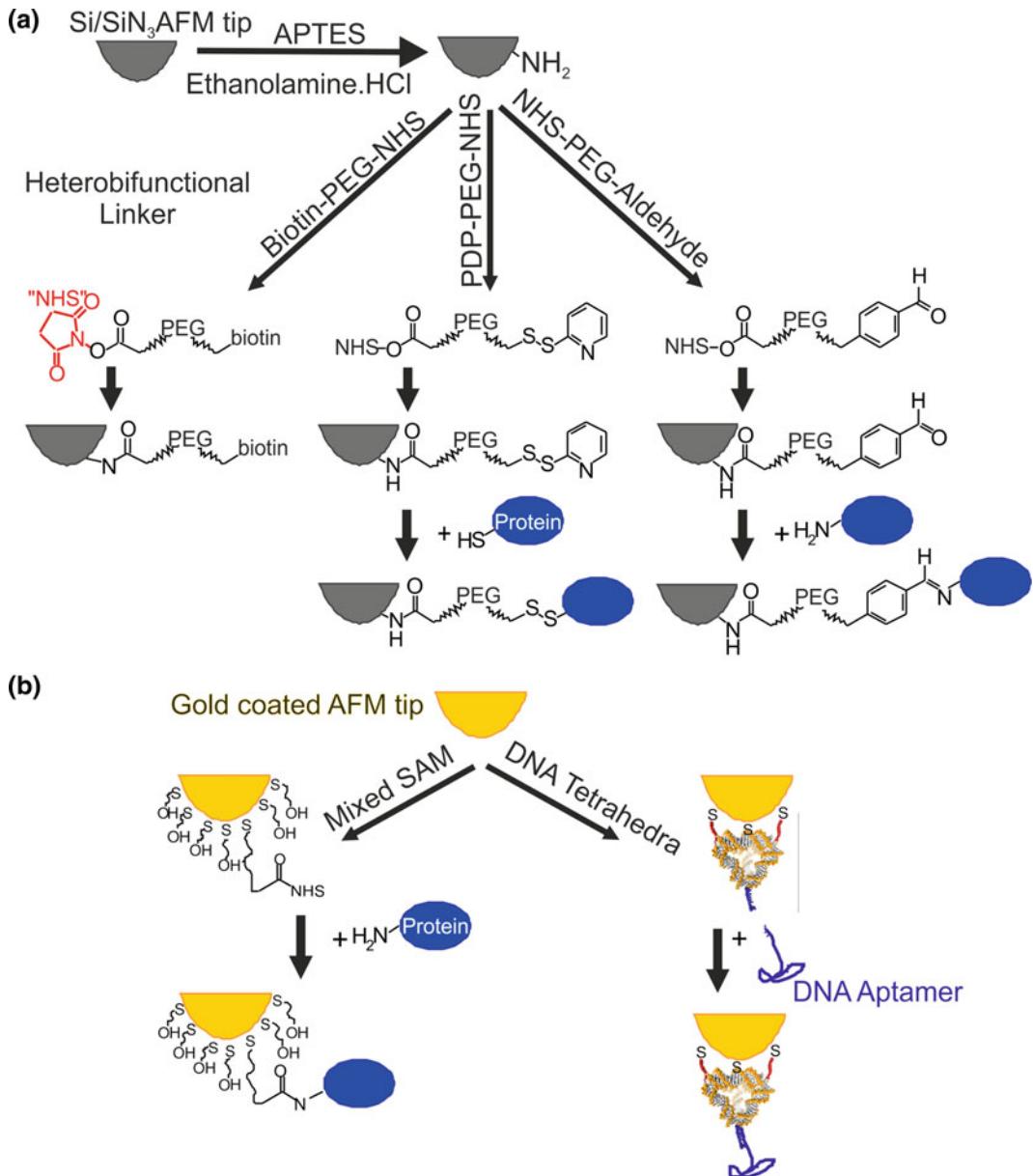


Fig. 4.6 Functionalization strategies for the covalent coupling of biological ligands to AFM tips. **a** A common approach for silicon and silicon nitride tips requires amino-functionalization of the tip surface for attachment of a heterobifunctional cross-linker that will tether the ligand. **b** Functionalization of gold coated tips exploits thiol-chemistry. Building blocks such as mixed self-assembled monolayers or DNA tetrahedra are used to link the ligand to the tip.

Ligands usually offer more than one coupling group. For instance, when the amino acid sequence of a protein contains more than one lysine residue. Synthetic proteins, however, are

often tagged by hexa- or deca-histidine for purification (His_6 and His_{12} , respectively), which can be exploited for site directed coupling by the use of nitrile-tri-acetic acid (NTA) functionalized

Table 4.1 Overview of heterobifunctional PEG linkers for silicon/silicon nitride tips

Typical target molecule	Linker	Reactive site	
Biotin	NHS-PEG-biotin	–	Ebner et al. (2007a, 2008a), Riener et al. (2003)
Protein, R-NH ₂	NHS-PEG-aldehyde	Lysine (R-NH ₂)	Ebner et al. (2007b, 2008a)
Protein, R-NH ₂	NHS-PEG-aldehyde	Lysine (R-NH ₂)	Ebner et al. (2008a), Wildling et al. (2005)
Protein, R-SH	NHS-PEG-PDP	Cysteine, R-SH	Ebner et al. (2008a), Kamruzzahan et al. (2006)
Protein, R-SH	NHS-PEG-maleimide	Cysteine, R-SH	Ebner et al. (2008a), Lin et al. (2006)
Aliphatic carbons (e.g. DNA)	NHS-PEG-benzophenone	Aliphatic carbon	Ebner et al. (2008a)
His ₆ -tagged protein	Multi-step procedure	His ₆ -tag	Ebner et al. (2008a), Tang et al. (2009)

linkers. Here the coupling protocol requires somewhat more effort since the synthesis of an NHS-PEG-NTA linker is not possible. A detailed description can be found in Tang et al. (2009). Table 4.1 gives an overview of the described linkers, their reactive sites and the target group in the bio-ligand. For in detail AFM-tip functionalization protocols the reader may refer to Ebner et al. (2008a) and Wildling et al. (2011). For completeness, it might be added that apart from these most common strategies, a number of other coupling protocols are available. These are covered for instance by Barattin and Voyer (2011).

4.3.4.2 Functionalization of Gold Coated Tips

Gold coated AFM-tips are readily commercially available, but have the drawback that the gold coating requires pre-coating with an adhesive layer (e.g. chromium) that increases the tip radius by 10 nm or more. Still, gold is an inert noble metal that can be cleaned at harsh conditions without any troublesome effects like surface oxidation. Functionalization can be achieved with extremely high efficiency utilizing the thiophilic properties of gold (Fig. 4.6b). To ensure sufficiently strong binding of thiolated cross-linkers, either a SAM of bifunctional alky-mercaptanes, or linkers containing a higher number of sulfur residues should be employed. In general, successful SAM formation on the AFM tip is achieved by a mixed SAM, where

one component (e.g. a ω -mercaptoalkanol) is used as a non-adhesive part and the other component affords covalent coupling of the ligand of interest. In contrast, the use of thiolated building blocks such as thiolated DNA tetrahedra offers a one- or two-step functionalization procedure for the generation of reactive coatings with highly defined spacing (Leitner et al. 2011).

4.4 Simultaneous Topography and Recognition Imaging (TREC)

In addition to measuring interaction forces, mapping the specific recognition sites on biological samples with high spatial accuracy is of great interest in many fields of life science. In contrast to the well-established but rather slow force volume mode (FV) (Ludwig et al. 1997; Willemsen et al. 1998; Grandbois et al. 2000; Lehenkari et al. 2000; Almqvist et al. 2004; Kim et al. 2011, 2015), the simultaneous topography and recognition (TREC) imaging technique has developed into a versatile tool for quick local mapping of binding sites in molecular biology and bionanotechnology (Stroh et al. 2004a, b). Although a lot of progress has been made in terms of performance and speed (Adamcik et al. 2011; Rico et al. 2011; Alsteens et al. 2012; Heu et al. 2012), TREC is still considerably faster than many force-volume techniques. For

comparison a complete FV image at 30 nm resolution requires typically 20 min, and increasing the resolution dramatically lowers the recording speed. In contrast, TREC images with at 2 nm resolution are recorded in less than 4.5 min.

4.4.1 Principle of TREC

TREC is based on a dynamic imaging mode in fluid, where the image is generated by oscillating the tip close to its resonance frequency (see Sect. 2.1). On top of that the topographical imaging of receptor molecules is combined with molecular recognition by tethering a ligand to the AFM tip using a distensible PEG cross-linker (see Sect. 3.4.1). The tip-bound ligand is capable of binding to cognate receptors on the sample surface. With this configuration and a modified feedback loop, two independent images are recorded simultaneously—a topography image of the sample surface and a lateral map of the corresponding receptor recognition sites. The separation into topographical and recognition signals is achieved by splitting the cantilever oscillation amplitude into downward and upward parts with respect to the cantilever's zero position using a special electronic circuit (PicoTREC, Keysight, formerly Agilent, Santa Rosa, CA, USA) (Fig. 4.7a). Provided that the cantilever's Q-factor is ~ 1 when in contact with the sample in liquid the maxima (or peaks) of oscillation are independently influenced by the tip sample interaction, i.e. topographic features measurably effect only the downward amplitude, whereas the physical connection of the tip to the surface during ligand-receptor binding causes reduction of the upward amplitude of the tip oscillation. Consequently, the envelope of the lower amplitudes generates the topographic image (Fig. 4.7 b), and the envelope of the upper amplitudes yields the map of the recognition sites (Fig. 4.7 c). As the tip contacts the surface only intermittently, this technique provides very gentle tip-surface interactions and the high oscillation frequency provides for more than a hundred times faster mapping than conventional force volume techniques.

4.4.2 Parameter Optimization in TREC

In 2004, the first proof of concept on biological samples was realized by visualizing the highly specific lysozyme—HyHEL5 antibody binding in single-molecule interaction maps while simultaneously carrying out high-resolution topographic imaging (Stroh et al. 2004). Shortly afterwards Ebner et al. (2005) employed the well-characterized, high-affinity avidin-biotin interaction to determine key factors for stable, robust and reliable recognition imaging results. In 2009, Preiner et al. (2009) investigated the physical properties of the tip-sample ensemble in much detail and established general handling procedures and a protocol for new TREC users to conduct meaningful TREC imaging experiments.

Unbiased values for topographical heights require operation of the feedback loop with only the lower half of the cantilever amplitude, which responds to the topographical features. Use of the full amplitude would include the influence of the recognitions signals, causing faulty height information (Fig. 4.8a). In order to reduce feedback artifacts (termed as topographical crosstalk) and unspecific adhesion in the recognition image the driving frequency should be adjusted to a value smaller than the cantilever resonance frequency at the surface. The most important factor, however, is the magnitude of the oscillation amplitude itself. A minimum-amplitude is required for the linker to exert a measurable force on the cantilever to cause an effect on the top peaks of the oscillations for the generation of the recognition signal (Fig. 4.8b). On the other hand, the receptor-ligand bond is broken at each oscillation cycle when using too large amplitudes, and recognition events pass undetected during scanning. The range of appropriate amplitudes for recognition imaging is determined by the effective linker length (including the position and the length of the molecule attached to the linker), and therefore rather sharply localized. Variation of the amplitude while keeping the ratio of amplitude set-point and free amplitude constant can thus be used as a quick specificity proof without perturbation of the receptor-ligand system as opposed

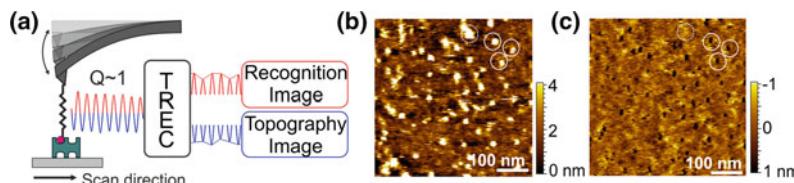


Fig. 4.7 Simultaneous topography and recognition imaging (TREC). **a** The cantilever oscillation is split into *lower* and *upper* parts allowing for simultaneous acquisition of topography and recognition image provided that the cantilever's Q-factor is ~ 1 . **b** Topography image of electrostatically adsorbed avidin molecules on a freshly cleaved mica imaged with a biotin-tethered tip. **c** Simultaneously recorded recognition image. Topography and recognition show very good correlation with avidin molecules causing corresponding recognition spots (*solid circles*). Some topographical features lack recognition (*dashed circle*) as they do not specifically interact with the tip bound biotin. Scan size 500 nm. After (Stroh et al. 2004; Ebner et al. 2005).

to the widely applied block of either functionalized AFM-tip or surface receptors by incubation with free the analog.

4.4.3 Application of TREC

TREC imaging has become a robust, accurate and reliable technique within the past decade. It has been applied to identify and localize specific receptor binding sites on a variety isolated systems and more complex heterogeneous bio-surfaces such as membranes, cells and whole tissues. On nanoarrays of self-assembled bacterial s-layer proteins Tang et al. (2008) determined the precision of TREC in the detection of recognition sites. Strep-tagII, an eight amino acid and artificial ligand for *Strep-Tactin* had been fused to a truncated form of the S-layer protein SbpA of *Lysinibacillus sphaericus* CCM 2177. Wild type SbpA s-layer protein mixed with rSbpA-Strep-tagII at a ratio of 7:1 was brought to self-assemble into a square (p4) lattice. TREC imaging with a *Strep-Tactin* functionalized AFM-tip yielded high resolution maps of the Strep-tagII binding sites with a surprisingly high positional accuracy of 1.5 nm (Fig. 4.9a).

In one of the first medical studies involving TREC the density of cystic fibrosis transmembrane conductor regulator (CFTR) on the membrane of red blood cells was investigated (Ebner et al. 2008b). CFTR is involved in the transport of fluids across epithelia. Mutations in the gene encoding CFTR can lead to dysfunctions and

cause cystic fibrosis (CF), which is characterized by a depletion of CFTR in the membrane of cells. A specific role of CFTR in erythrocytes is unknown at the current stage and no major hematological disorders have been reported for CF patients. Nevertheless, TREC confirmed a significant depletion of CFTR in red blood cells. About 800 CFTR proteins were localized in the erythrocyte plasma membranes of healthy donors while in CF patients the number was decreased to about 250. The advantages of TREC over methods relying on for instance quantum-dot-labeled antibodies (Dufrene 2002) are the prevention of possible steric hindrance during binding of labeled antibodies to the target molecule and a direct visualization of molecules (Fig. 4.9b).

Another medical application for recognition imaging was carried out by Rankl et al. (2010), where the distribution of corneodesmosin on the surface of stratum corneum (SC) was mapped at the nanoscale. Corneodesmosin, a protein known to be present in the SC, plays an important role in its physical integrity. With TREC, it was possible to characterize the morphological changes and the distribution of corneodesmosin on the surface of SC. The results are applied in the design of cosmetic compounds that will allow better interaction with the SC to the benefit and improvement of skin appearance (Fig. 4.9c).

Just recently, TREC imaging was applied to visualize uncoupling protein 1 (UCP1) reconstituted into lipid bilayers and to analyze the ATP-protein interaction at a single molecule level (Koehler et al. 2017). UCP1 is a mitochondrial

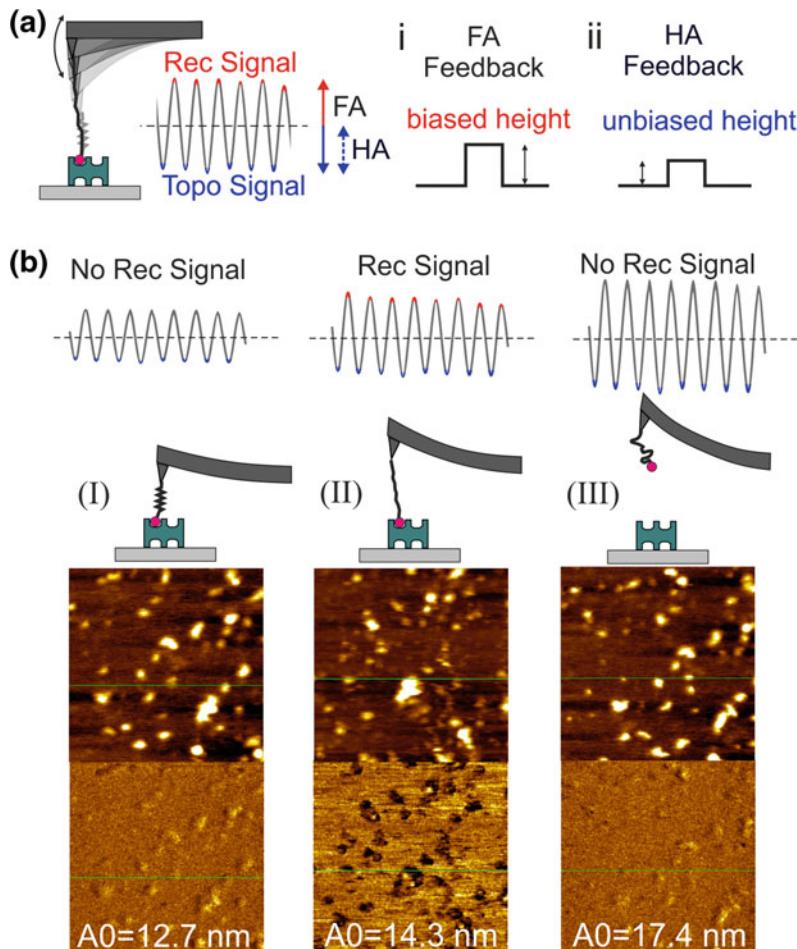


Fig. 4.8 Parameter optimization in TREC. **a** Comparison of half-amplitude feedback (HA) versus full amplitude feedback (FA). When the tip-tethered ligand has bound to surface-attached receptor, the topography is sensed at the bottom of the oscillation (blue) and stretching of the linker molecule reduces the movement in the upper parts of the oscillation (red). Consequently the FA parameter is biased by the additional amplitude reduction due to molecular recognition. In response the piezo-actuator will exaggerate vertical adjustment, which shows as an increased topographical height of the detected receptor molecule. Only operation of TREC using the lower HA feedback parameter will yield correct height information. **b** The oscillation amplitude is the most critical parameter in recognition imaging. If the oscillation amplitude is too small (*I*) the linker does not exert an efficient force on the cantilever upward swing and no recognition signal is generated. In the second regime (*II*), the amplitude has been increased, causing an efficient stretching of the linker without rupture of the ligand-receptor complex in the upward swing. This results in a pronounced recognition signal. In the third regime (*III*), the amplitude is too high; causing the rupture of the bond in every top peak of the oscillation and no stable recognition signal is generated.

membrane protein and involved in non-shivering thermogenesis under cold-acclimating conditions. In the last years, this protein has become increasingly important as a target protein for the therapy of obesity as well as inflammatory, neurodegenerative and ischemic diseases. Based on this system, a new mode for mapping and studying interaction sites at low lateral density

was devised by combining recognition imaging and force spectroscopy. This method makes it possible to select a single receptor molecule reconstituted in a supported lipid membrane and then quantify receptor-ligand unbinding forces. Apart from these few selected examples a comprehensive overview of TREC studies can be found in Table 4.2.

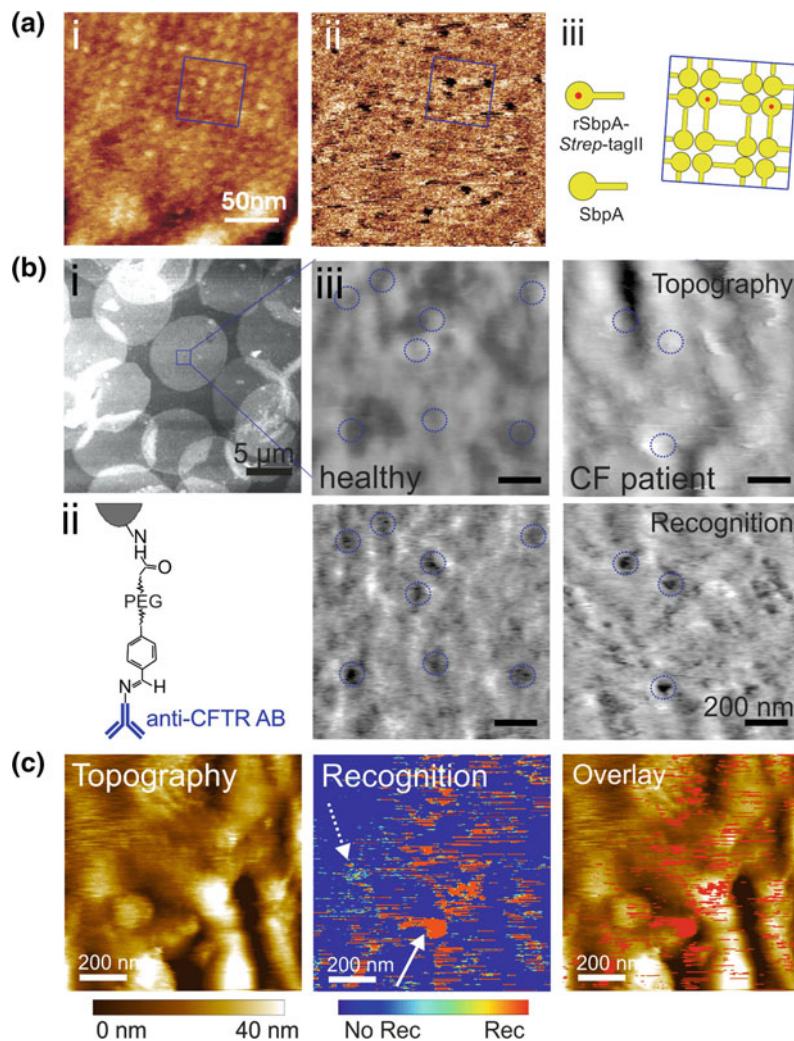


Fig. 4.9 Bio-sensing TREC applications. **a** Topography (*i*) and recognition (*ii*) images of the S-layer proteins rSbpA-Strep-tagII and wild-type SbpA (in a molar ratio of 1:7) co-crystallized on a silicon chip as imaged with a Strep-tagII binding Strep-Tactin functionalized AFM-tip. The schematic drawing illustrates the regular arrangement of the co-crystallized proteins and the ratio of Strep-tagII which is perfectly reflected in the recognition image at a lateral resolution of 1.5 nm. **b** Medical sensing application of TREC for identification of cystic fibrosis (CF) patients. *i* Inside-out oriented erythrocyte membrane patches were scanned with an anti-CFTR antibody functionalized tip for quantification of CFTR receptors (*ii*). *iii* Topography (*top*) and recognition (*bottom*) images of healthy and CF donors reveal fewer recognition events between tip-bound anti-CFTR and CFTR in the erythrocyte membrane as a clear indication of the disease. **c** Simultaneous topography and recognition images of the stratum corneum of a skin sample using a tip coupled with an antibody against corneodesmosin, which is implicated in the process of skin peeling. Recognition spots of a few (*dashed arrow*) up to 100 nm (*white arrow*) were found. The overlay of both images shows that most of the recognition spots are located on top of pronounced topographical features.

Table 4.2 Overview of TREC studies in the years 2006–2016 (no claim to completeness)

Sample	Study	Author, year, reference
Endothelial cells	Identification and visualization of vascular endothelial (VE)-cadherin that is known to play a crucial role in calcium-dependent, hemophilic cell-to-cell adhesion	Chtcheglova et al. (2007, 2010)
DNA aptamers	In vitro selection and evaluation of DNA aptamers selected to bind histone H4 tails	Lin et al. (2007)
Red blood cell membranes	Density of cystic fibrosis transmembrane conductor regulator (CFTR) on the membrane of red blood cells was investigated	Ebner et al. (2008b)
Cardiac cells	Inhibition of hERG K ⁺ channel is major cause of long QT syndromes inducing fatal cardiac arrhythmias mediated by ergtoxin. TREC was used to visualize ergtoxin receptors on the voltage sensing domain of hERG K ⁺ channels	Chtcheglova et al. (2008)
Bacterial surface layer (S-layer)	Visualizing binding sites on bacterial S-layer nanoarrays containing affinity-tags (Strep-tagII) to with a positional accuracy of 1.5 nm. The nanoarrays were then used as a template for constructing highly ordered molecular binding blocks	Tang et al. (2008)
Chromatin	To provide insights into epigenetic processes, Wang et al. used TREC to perform single-epitope recognition imaging of native chromatin	Wang et al. (2008)
Stratum corneum	Mapping the distribution of corneodesmosin on the surface of stratum corneum (SC) at the nanoscale for designing new cosmetic compounds that will better interact with the SC surface for a specific skin benefit	Rankl et al. (2010)
Macrophages	Precise nano-mapping revealed compact cluster organization of Fc γ Rs that would favor the rapid recognition of huge pathogens and thus their effective internalization	Ahmad et al. (2011)
Human eye lens	Recognition imaging of whole, untreated human tissue samples to map the distribution of abnormal protein aggregates on the anterior lens capsule, which is a well-known cause of irreversible blindness	Creasey et al. (2010, 2011)
Protein-nucleic acid interaction	TREC can also be utilized to study protein—nucleic acid interactions which play important roles in the physiological behavior of living cells and mapped molecular details of the ricin-aptamer interaction	Wang et al. (2012)
Plant cells	Mapping the structural changes of the cell wall of natural and pretreated plant cells provides better understanding of surface structural changes which leads to a potential improvement of biomass pretreatment for biofuel production	Zhang et al. (2013)
Mitochondrial Membrane	Visualization of uncoupling protein 1 (UCP1) reconstituted into lipid bilayers and analysis of the ATP-protein interaction at a single molecule level. Based on this study, TREC was further developed to quantify also binding forces on interaction sites at low lateral density	Zhu et al. (2013), Koehler et al. (2017)
Monocytes	CD1d molecule, a monomorphic major histocompatibility complex class I-like molecule, presents different types of glycolipids to invariant natural killer T (iNKT) cells that play an important role in immunity to infection and tumors, as well as in regulating autoimmunity. TREC analysis was conducted to nanomap single CD1d molecules on THP1 cells that were loaded with different glycolipids	Duman et al. (2013)

(continued)

Table 4.2 (continued)

Sample	Study	Author, year, reference
Bladder cancer cells	Human gonadotropin-releasing hormone receptor (GnRH-R; or type I GnRH-R) that expressed in tumor cells, has gained a great attention as a specific target for cancer therapy. Given the clinical utility, the improved nano-organization GnRH-R is an important step in the development of more effective and possibly new therapeutic strategies	Zhang et al. (2014)
Melanoma cells	Hsp70 is known to be present at the extracellular leaflet of the plasma membrane (surface Hsp70, sHsp70) of most tumor cells but not that of the corresponding benign cells. TREC analysis was conducted to elucidate its pro-tumor role	Nimmervoll et al. (2015)

4.5 AFM Biomedical Sensing—Examples

In the previous sections AFM techniques have been introduced that present a multifunctional toolkit for an ever increasing diversity of biological and life sciences applications. The AFM has evolved into an enabling instrument to observe structural details of molecular assemblies, isolated membrane preparations and cell surfaces, and to monitor cellular responses to external stimuli such as antibodies or drugs. In the following studies topographical imaging, SMFS, TREC and their combination were employed to answer questions in basic molecular biology, pharmaceutical science and the medical field that were inaccessible by other techniques, so far. Thus, the selected studies illustrate and highlight the biomedical sensing capabilities of the AFM.

4.5.1 Isolated Membranes

4.5.1.1 SMFS Investigation of Blood Coagulation—Von Willebrand Factor

The von Willebrand factor (VWF) is a glycoprotein that plays a central role in hemostasis, the first stage of wound healing. It triggers platelet adhesion in areas of vascular damage by binding to exposed sub-endothelial collagen and thus causes wound closure. In an initial step the VWF

binds to collagen with specific binding sites that are located in domains A1 and A3 of the glycoprotein. Collagen III is believed to interact with the A3-domain, and collagen VI with the A1-domain. The forces and the dynamics involved in these interactions were investigated by SMFS with the VWF tethered to the AFM tip and substrates coated with dense layer of collagen (Posch 2015). Different mutations in the A3-domain of the VWF A1–A2–A3 construct were investigated. These mutations (H1786R, Q1734H and S1731T) correspond to the type 2M of the current von Willebrand disease (VWD) classification system, and are characterized by a defect in the collagen binding domain but exhibit otherwise normal multimeric distribution. Surprisingly, patients carrying these mutations show no or mild symptoms, or do have alternative reason for impeded coagulation. These mutations were diagnosed purely by accident and thus, their clinical significance was not clarified up to now. It could be shown that defective interaction between one VWF domain and one type of collagen can be compensated by alternative binding events. This observation might explain the inconspicuous bleeding tendency found in patients carrying these VWF A3 domain mutations.

In a follow-up study the interplay of the VWF domains A1 and A2 during the onset of blood coagulation were investigated more closely (Aponte-Santamaría et al. 2015). The VWF domain A1 is responsible for mediating platelet adhesion under flow through the platelet

glycoprotein Ib α (GPIb α). The adjacent VWF domain A2 unfolds under shear upon which it exposes a cleaving site for the protease ADAMTS13 to prevent thrombosis. A combination of molecular dynamics (MD) simulations, AFM (Fig. 4.10a), and microfluidic experiments revealed that the A1/platelet interaction is blocked by mutual A1/A2 binding under low shear. In order to elucidate the exact shielding mechanism, SMFS was used to mimic typical elongation forces imposed by the shear flow of blood (Posch et al. 2016). The

interaction strength between the A1 and A2 domain (Fig. 4.10b, c) was probed together with the unfolding behavior of A2 (Fig. 4.10d) as an indicator of the shielding capacity. As a result only a small extent of A2, if at all, unfolded before dissociation from A1. These findings suggest that the platelet binding site in domain A1 is made accessible by the slip-bond dissociation of domain A1 from A2 and that domain A2 remains largely folded during this process, which keeps it protected against cleavage and degradation (Fig. 4.10e).

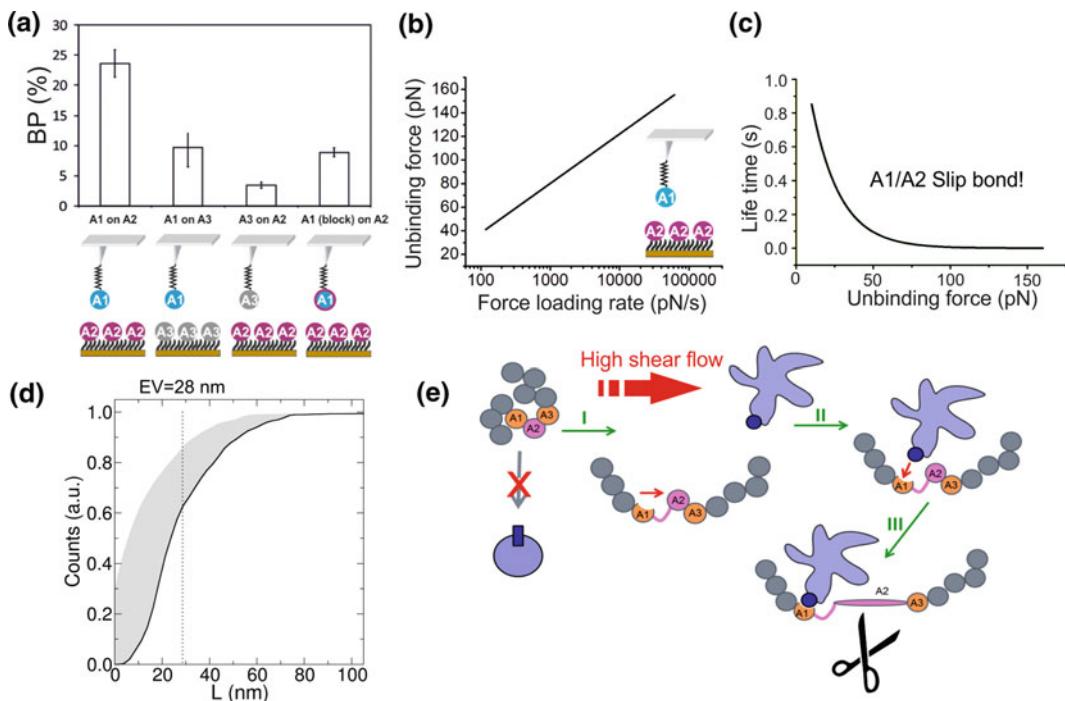


Fig. 4.10 (a–d) SMFS investigations of the functions of the von Willebrand factor (VWF) in blood coagulation. **a** Number of specific binding events between VWF A1 and A2 domains, where A1 was coupled to the AFM tip and A2 was immobilized on the surface. In control experiments either the A2 domains on the surface or the A1 domain at the AFM tip were replaced by the unspecific VWF A3 domain or the tip-bound A1 was saturated by soluble A2. **b** Loading rate dependence of the VWF A1/A2 interaction. **c** The life time *versus* unbinding force-plot clearly indicates slip bond behavior of the A1/A2 interaction. **d** Cumulative distribution of the elongation L of the A1–A2 complex prior to dissociation (black line) and the corrected value after subtraction of the size of A1 and the linkers (gray area). The dotted line indicates the expectation value (EV) for L. The measured EV (*28 nm) was substantially lower than the extension of a fully stretched unfolded A2 domain (*80 nm). **e** Suggested model for the VWF A1/A2 interaction. Higher shear stress due to e.g. an injury activates VWF to interact with the platelet GPIb α receptor. The GPIb α binding site in domain A1 is made accessible by dissociation of domain A1 from A2 under high shear flow conditions (I, II). Domain A2 remains largely folded during this process protecting it from cleavage and degradation (II). SMFS data imply that dissociation predominantly occurs before exposure of A2 to cleavage. Thus, the A2 domain might prevent platelet binding in a force-dependent manner to ensure that VWF is entirely activated before cleavage.

4.5.1.2 TREC Studies to Elucidate ATP-Binding to UCP1 in the Mitochondrial Membrane in Non-shivering Thermogenesis

Uncoupling protein 1 (UCP1) is a transmembrane protein in the mitochondrial membrane of brown adipose tissue (BAT). The protein dissipates the energy of the proton gradient over the inner mitochondrial membrane as heat. This process is also known as non-shivering thermogenesis under cold-acclimating conditions (Klingenberg and Echternach 2001) and is essential to newborns to counter heat loss, but also prevents energy storage as fat. Thus, UCP1 has been proposed as a pharmaceutical target for the treatment of obesity in adults (Nedergaard et al. 2005; Lidell and Enerbäck 2010). Long-chain fatty acids activate UCP1 proton transport, whereas purine nucleotides have an inhibitory effect (Nicholls et al. 1978; Lin and Klingenberg 1980; Nicholls 1976; Urbánková et al. 2003; Beck et al. 2007). Using TREC the mechanism of nucleotide-mediated binding and inhibition of UCP1 inside the mitochondrial membrane was elucidated on a structural level (Zhu et al. 2013).

Figure 4.11a shows the experimental setup for TREC measurements of UCP1 reconstituted into lipid bilayers on a mica surface and a cantilever tip functionalized either with ATP or anti-UCP1 antibody specific to amino acid residues 145–159 of UCP1 using acetal-linker chemistry. Only about half of the protein molecules that were visible in the topographic image were also recognized by anti-UCP1 AB (Fig. 4.11b). This result shows that the orientation of UCP1 protein in the planar lipid bilayer is random, as expected. The specificity of this interaction was proved by injection of a peptide that competed with UCP1-anti-UCP1 AB interaction. Contrarily, when the cantilever was functionalized with ATP instead of the UCP1-AB, co-localization of proteins in the topography and signals in the recognition image was close to 100%. As the orientation of UCP1 in the membrane was determined to be random, this result was

unexpected and implies that the nucleotide binding sites are accessible irrespective of the orientation of the protein in the reconstituted membrane. Whereas the size of protein spots in the topographic image was uniform, the size of the recognition spots appeared different. Two groups with mean diameters of 12.65 ± 1.03 and 17.68 ± 0.86 nm were determined (Fig. 4.11c, ii). Given that the ATP binding site remains accessible no matter which side of the protein is immersed in the layer or facing upwards, the existence of two different sizes of recognition spots can be explained by differences in submersion depths of the ATP molecule on the tip and the linker to estimate the localization of the binding site, ATP functionalized tips with different lengths of linkers were used (Fig. 4.11d, iii). From these data the depth of the ATP binding site in UCP1 was calculated to be 1.27 ± 0.10 nm from the membrane surface.

Taken the NMR structure of UCP2 (57% homology to UCP1) (Berardi et al. 2011; Pebay-Peyroula et al. 2003) and the proposed three-step purine nucleotide (PN) binding model for UCP1 (Modrianský et al. 1997) into account we hypothesized that only binding of ATP to the cytosolic side (c-side) would lead to a conformational change in the protein, which then causes inhibition. PN binding from the matrix facing side (m-side) would not induce a conformational change and thus not trigger UCP1 inhibition (Fig. 4.11e). This suggestion might plausibly explain one of the unresolved bio-energetical questions: how can the protein be in the active state despite the presence of mM concentrations of ATP in the cytosol? In the absence of the crystallographic structure of UCP1, TREC data yield valuable insights into the nucleotide binding mechanism in the process of non-shivering thermogenesis. High-resolution topographical AFM images provided for the first time visualization of UCPs reconstituted into lipid bilayer membranes. Simultaneously measured recognition images revealed the existence of a single ATP binding site that can be accessed by nucleotides from both the c-side and the m-side.

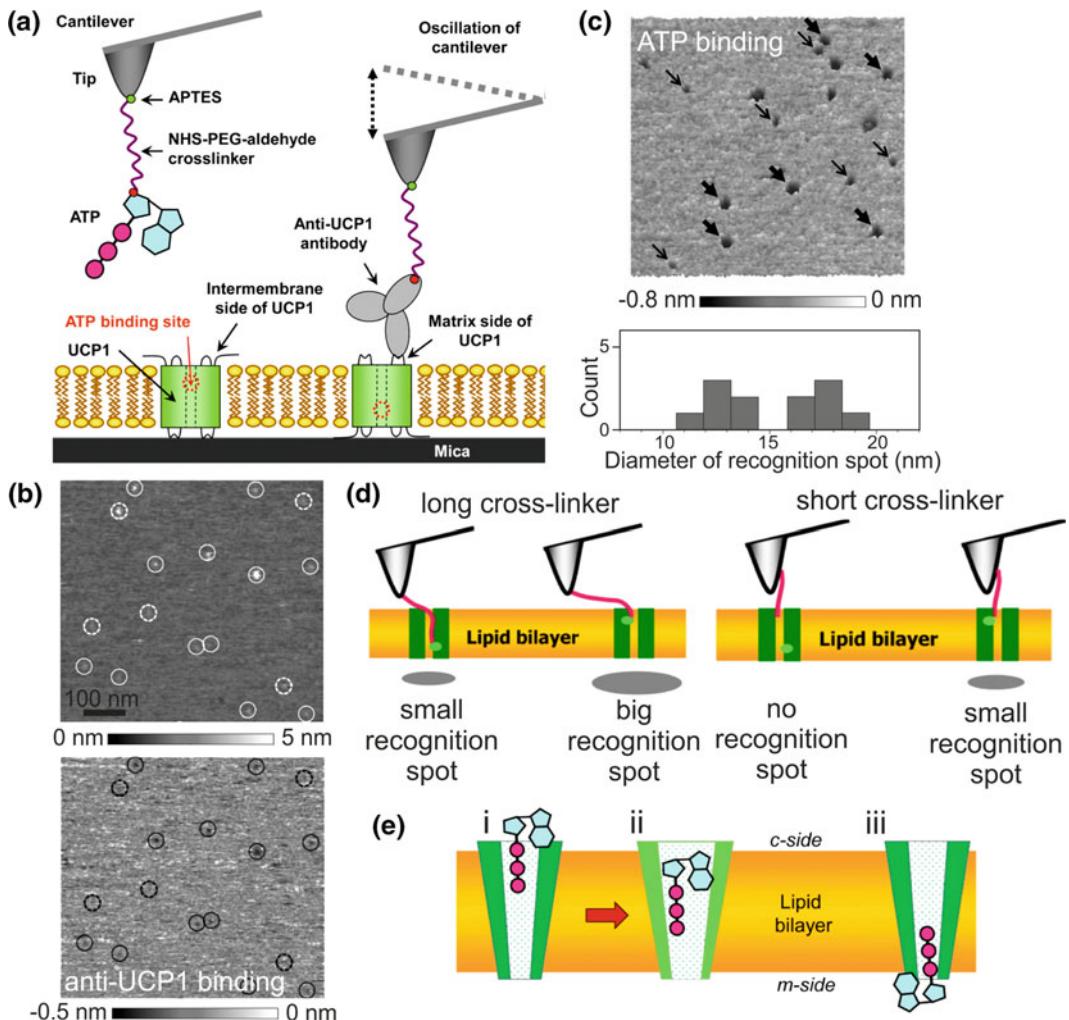


Fig. 4.11 Investigation of ATP-binding to UCP1 using TREC. **a** Illustration of the TREC experiment with the uncoupling protein 1 (UCP1) reconstituted into a supported lipid bilayer on a mica surface and a UCP1 specific antibody or ATP coupled to the AFM-tip. **b** High-resolution topographical (top) and UCP1 antibody-recognition (bottom) images. Correlation shows that some UCP1 molecule are recognized (circles of solid line), whereas a few remain unrecognized (dashed). **c** High-resolution ATP-recognition image of UCP1 in the bilayer reveals large (thick arrows) and small (thin arrows) spots, whereas the molecules size in the topography (not shown) were homogeneous in size. Evaluation of the diameter of the recognition spots shows a bi-modal statistical distribution. **d** Schematic illustration of the expected dependence of recognition spot size on the orientation of UCP1 in the bilayer and on the length of the cross-linker that tethers ATP to the AFM-tip. **e** Proposed mechanism of ATP binding: initial binding from the c-side (*i*) followed by conformational change of UCP1 (*ii*) and protein inhibition. *iii* Protein binding from the m-side without conformational change.

4.5.2 Mammalian Cells

4.5.2.1 Drug Binding to the Serotonin Transporter (SERT)

The serotonin transporter (SERT) is a member of the neurotransmitter sodium symporter

(NSS) family (Kristensen et al. 2011) and retrieves extracellular serotonin back into the presynaptic neuron to terminate neurotransmission. Some neurological conditions, e.g. depression, anxiety disorder, attention-deficit hyperactivity disorder (ADHD), autism,

Parkinson's disease, and addiction are related to NSS dysregulation. SERT is a target both for therapeutic medicines such as citalopram (CIT) to treat depression and for illicit drugs like cocaine which induces severe health problems. X-ray analysis of SERT homologues, i.e. bacterial LeuT_{Aa} (Yamashita et al. 2005) and dDAT (the *Drosophila melanogaster* dopamine

transporter) (Penmatsa et al. 2013; Wang et al. 2015) indicate a primary binding site (S1-site) for substrates in the membrane-spanning region. Other studies provided hint for a second S2-site, located within the extracellular vestibule (Plenge and Mellerup 1997; Plenge et al. 2012). To elucidate the number and the mechanism of ligand binding site(s) in SERT, we used SMFS to

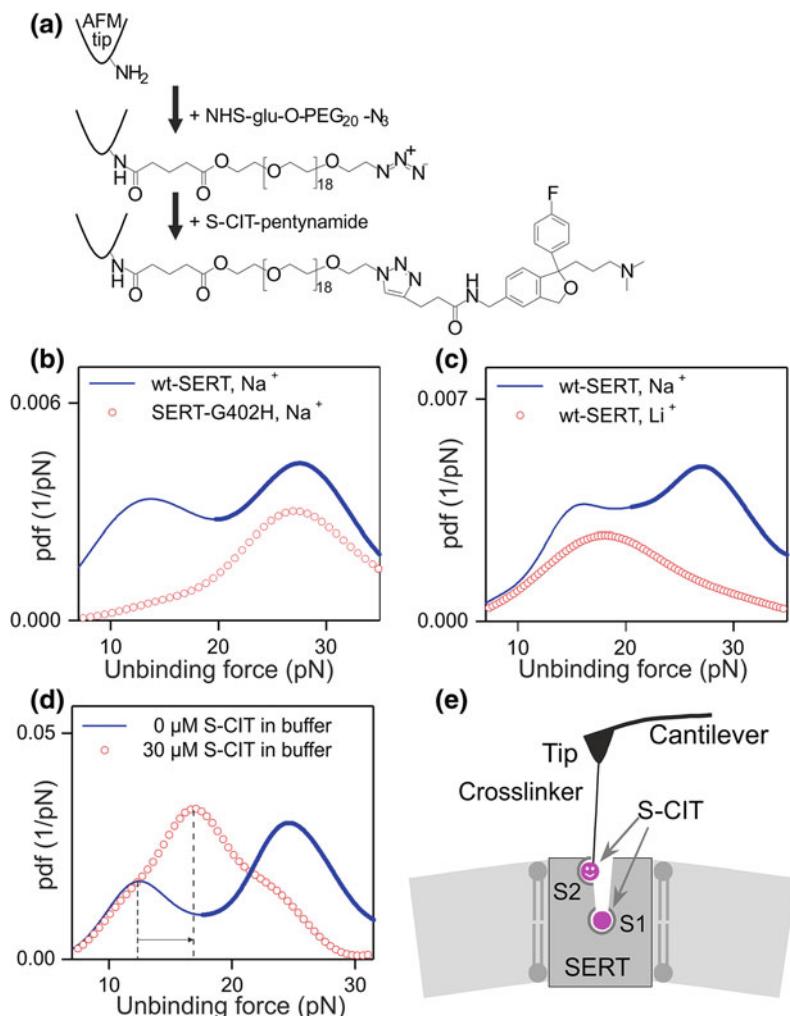


Fig. 4.12 Drug binding to the serotonin transporter (SERT). **a** Conjugation of S-CIT to the AFM cantilever tip. After the cantilever was amino-functionalized with gas phase APTES, it was pegylated with NHS-glu-O-PEG20-N3. The alkyne-modified S-CIT analogue was coupled to the azide-terminated PEG via co-catalyst-accelerated copper(I)-catalyzed azide-alkyne cycloaddition. **b** Comparison of force distribution measured on wild type (wt) SERT (solid line) and mutant SERT-G402H (dotted line). **c** Comparison of force distribution measured in Na^+ buffer (solid line) and Li^+ buffer (dotted line). **d** Comparison of force distribution measured on wt-SERT in Na^+ buffer without (solid line) or with 30 μM S-CIT (dotted line). After the S-CIT in solution bound at S1 site (e), the binding strength between S-CIT and S2 site was enhanced, shown as the shift of the peak position in (d), indicating a positive allosteric.

measure the interaction forces between SERT and S-enantiomers of CIT (S-CIT) (Zhu et al. 2015).

An alkyne-modified 5-aminomethyl analogue (Kumar et al. 2014) of the S-CIT was covalently conjugated onto the AFM cantilever tip (Fig. 4.12a) using click chemistry. SMFS was conducted on living CHO-K1 cells expressing human SERT (Wildling et al. 2012) in HEPES buffer at room temperature. Force curves with unbinding events were observed during the repeated measurements, and the experimental probability density function (PDF) of forces was thus generated (Fig. 4.12b, solid curve). The two distinct peaks in the force PDF reflect the S-CIT rupture from binding sites with two different strengths of interaction. To identify the location of the binding sites for the two populations, force measurements were performed on SERT-G402H which contains a point mutation in the outer vestibule (Plenge et al. 2012). The force PDF from SERT-G402H (Fig. 4.12b, dotted curve) showed only a single peak, which corresponded to the second peak in the force PDF from wild type (wt) SERT. This indicates that the first peak in the force PDF from wt SERT originates from the vestibular S2 site. In addition, we verified that Na^+ ions were important for accessibility to the SERT binding sites by recording forces of S-CIT binding to wt SERT in Li^+ buffer. The PDF showed a single peak in the force PDF (Fig. 4.12c, dotted curve) that coincided with the first peak in the force PDF observed in Na^+ buffer (Fig. 4.12c, solid curve). The absence of the second peak is consistent with the notion that the central S1 site is Na^+ -dependent. The persistence of the first peak implies that access to the vestibule S2 site does not require Na^+ .

To examine if there is allosteric effect (Monod et al. 1963) between the S1 and S2 sites, we performed competition experiments with 30 μM S-CIT in solution so that the higher affinity S1 site was occupied whereas the lower affinity S2 site was still accessible to S-CIT on the tip. Consequently, only a single peak remained in the force PDF (Fig. 4.12d, dotted curve), which was attributed to the S2 site. This is due to the long life span (~ 20 min) of S-CIT at the S1 site

(Plenge and Mellerup 1997). Meanwhile, the AFM-tip tethered S-CIT, which has a local concentration of ~ 3.7 mM (corresponding to one S-CIT molecule within a hemisphere of radius ~ 6 nm, defined by the length of PEG linker), is likely to effectively compete with the S-CIT in solution for binding to the S2 site. The remaining peak shifted to the right (arrows in Fig. 4.12d), indicating that the binding of S-CIT at the central S1 site exerts a positive allosteric effect on the S2 site (Fig. 4.12e). Such allosterically enhanced binding strength of S-CIT at the S2 site should provide the possibility for the X-ray crystallography to capture the S2 site. Three months after the demonstration of the two S-CIT binding sites in SERT by our SMFS study (Zhu et al. 2015), the X-ray structure of SERT bound with two S-CIT molecules was published (Coleman et al. 2016) which verified our results.

Given the importance of allosteric regulation in biology and pharmacology, the SMFS based nanopharmacological approach paved a new avenue to explore transient binding sites in clinically relevant membrane transporters and opened the door to quantitatively address the modulation of interaction forces between ligands and allosterically coupled binding sites.

4.5.2.2 Fluorescence-Guided SMFS to Confirm Increased GLUT4 Translocation

The combination of AFM instrumentation with optical microscopy (Pfister et al. 2005) has greatly eased biomedical sensing applications of the AFM on cells. It provides the possibility to correlate independently or simultaneously acquired topographical and optical/fluorescence images (Duman et al. 2010; El-Kirat-Chatel and Duffrene 2012) and to guide the AFM tip in SMFS experiments (Madl et al. 2006). In a recent application optically-guided AFM was employed to reveal the translocation of glucose transporter 4 (GLUT4) in cells (Lanzendorfer et al. 2014). GLUT4 is the key molecule for uptake of glucose into adipocytes and muscle cells and thus reduces blood glucose levels to prevent hyperglycemia. In a basal state, GLUT4 is trapped in intracellular compartments. Insulin induces the translocation

of GLUT4 into the plasma-membrane and thus triggers glucose uptake. Resistance to insulin, an important parameter of type-2 diabetes, strongly reduces this translocation process (Govers 2014).

There is a great demand for analytical techniques capable of quantifying the GLUT4 translocation efficiency in order to search for compounds that are able to enhance the GLUT4 translocation process in the absence of insulin (insulin mimetics substances). Most approaches such as western blot analysis of membrane fractions, ELISA guided binding assays, flow cytometric analysis (Bruzzone et al. 2012; Kadan

et al. 2013; Kristiansen and Richter 2002) and immuno-electron techniques (Carvalho et al. 2004) suffer from being indirect, time-consuming, non-quantitative and/or more expensive. An important improvement was the insertion of different epitope tags (mainly myc and HA) in the first extracellular loop of the GLUT4 protein in combination with a genetically encoded fluorophore molecules (mainly GFP) at the c-terminus of the transporter (Jiang et al. 2002; Baus et al. 2008) as depicted in Fig. 4.13a. Here, dual-color total internal reflection fluorescence (TIRF) microscopy has been

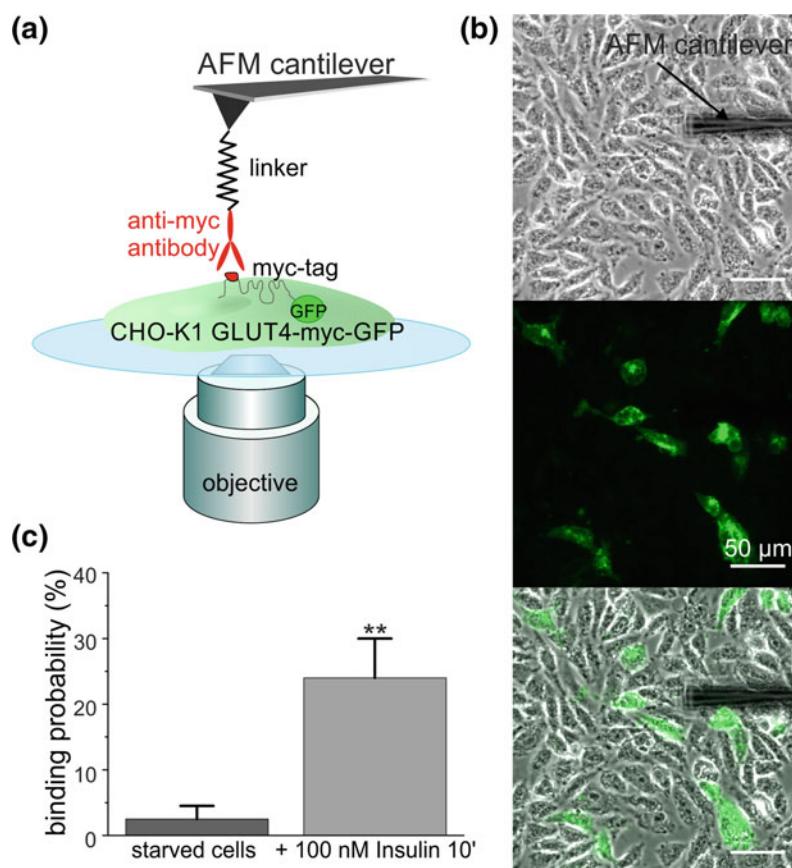


Fig. 4.13 Fluorescence-guided AFM force detection of GLUT4 molecules in the membrane of CHO-K1 hInsR/GLUT4-myc-GFP cells. **a** Schematic representation of an integrated platform combining an AFM with an inverted light microscope to visualize and recognize the insulin dependent GLUT4 glucose transporter by using a genetically encoded fluorophore (GFP) at the c-terminus and a myc-tag in the first extracellular loop. An anti-myc antibody functionalized AFM tip was used to exclusively identify membrane inserted GLUT4 molecules. **b** By overlapping highly contrasted optical (phase contrast) and fluorescence images it was possible to visualize the AFM cantilever and thus to guide it on a cell of interest. **c** Force measurements results (binding probability) of starved and insulin-treated (cells were incubated with 100 nM of insulin for 10 min) GLUT4-myc-GFP hInsR cells. **P < 0.01, significant increase with respect to the starved cells.

used to track GLUT4 translocation in the CHO-K1-hIR-GLUT4-myc-GFP cell line. Nevertheless, a critical point of such TIRF experiments is to proof that the increase of fluorescence in the evanescent field in response to insulin or insulin mimetic compounds is solely caused by the increase of GLUT4 molecules in the plasma membrane, and not by the increase of recruited vesicles in the TIRF zone. The use of fluorescently labeled antibodies that target the myc-tag in the first extracellular loop of GLUT4-myc-GFP is a possibility to further reduce uncertainty. However, before incubation with the reporter molecule, cells are commonly fixed with paraformaldehyde, which may slightly permeabilize the cells, potentially leading to staining of intracellular GLUT4 molecules, as well. In order to clarify this uncertainty, fluorescence-guided SMFS experiments were performed with the anti-myc antibody covalently coupled to the AFM tip. In this configuration the antibody can only interact with GLUT4-myc-GFP fraction on the cell membrane.

For SMFS experiments the anti-myc antibody was covalently bound to the AFM tip (nominal spring constant 0.02 N/m, MSCT, Bruker, CA, USA) via a bifunctional NHS-PEG-acetal linker (Fig. 4.13a). Highly contrasted optical (phase contrast) and epi-fluorescence images were acquired with a 20 \times objective to enable positioning of the AFM cantilever above CHO-K1 cells that stably express the GLUT4-myc-GFP construct (Fig. 4.13b) and force-distance curves (FDCs) were collected. For each condition, i.e. either deprived of serum (starved) or insulin stimulated, at least three different cells expressing GLUT4 receptors were tested using in sum three different AFM tips. Each AFM/epi-fluorescence experiment was performed in duplicate using independent cell preparations. For each data set of FDCs (>500 FDCs per set) the binding probability (BP) was determined as the number of force curves that show specific unbinding events of anti-myc antibody and GLUT4-myc-GFP molecules on the cell membrane divided by total number of collected force curves. A significant difference

between starved (BP \sim 2.4%) and insulin stimulated cells (BP \sim 24.6%) was clearly recognized (Fig. 4.13c).

This fluorescence-guided SMFS study elegantly and decisively confirmed the improved accessibility of the myc-tag in the first extracellular loop of GLUT4 only after insulin stimulation. In a follow-up study additional AFM based techniques such as TREC will help to uncover the nano-organization of GLUT4 proteins on cell membranes, their oligomeric states, and as well the cortical cytoskeleton rearrangement. A detailed comparison of the effect of insulin and insulin mimetic phytochemicals will help to identify new compounds that increase GLUT4 translocation, oligomerization, domain formation, and cytoskeletal coupling for treatment of type-2 diabetes.

4.5.3 Bacteria

For decades, the methods of choice for studying physiology and behavior of prokaryotic organisms were growth in batch cultures and on agar plates (Persat et al. 2015). In the last few years it has been recognized that apart from chemical cues also mechanical interactions through cell-cell and cell-surface contact play a significant role in microbial biology on surfaces. Thus, characterization of surface structures, binding mechanisms and identification of main molecular players have moved into spotlight for understanding fundamental processes such as bacterial adhesion, surface recognition, biofilm development, and the incipient stage of infection.

Advances in AFM techniques provide us with powerful tools to explore the physical and mechanical properties of bacteria in unprecedented detail. Apart from analyzing nanostructures on the cell surface by topographical imaging, indentation measurements yield physical properties of single bacteria such as cell wall stiffness, Young's modulus, internal turgor pressure, and the stretching modulus as a function of environmental conditions like medium ionic strength. Such measurements have revealed that the nano-mechanical properties of bacteria

can impose mechanics to drive the adaptive behavior (Oh et al. 2012). Additionally, force indentation and force spectroscopy approaches using AFM to study these effects have become available only recently by measuring processes occurring at the interface between micro-organisms and cells (Dufrêne 2015).

4.5.3.1 Bacterial Surface Layer (S-Layer) of *Tannerella Forsythia*

Crystalline bacterial cell surface layers (S-layer) are the most common surface structure in bacteria. S-layers may hold great merit in nanotechnology due to their inherent ability to self-assemble into an ordered protein lattice on solid surfaces (Tang et al. 2008). However, from a clinical point of view the outermost cell surface

is a virulence factor, as it mediates adhesion to and invasion of host cells. Thus, it is an important task to characterize the biochemical nature and structure of S-layers in pathogenic bacteria under live cell conditions. *Tannerella forsythia* (*T. forsythia*) is a gram negative oral pathogen belonging to the so-called “red-complex” bacteria together with *Porphyromonas gingivalis* and *Treponema denticola*, and has been implicated in onset and development of periodontal diseases in human (Chalabi et al. 2010). Clinical evidence accumulates that *T. forsythia* also impacts systemic health, but the underlying mechanisms of pathogenicity are only slowly beginning to unravel (Sekot et al. 2011).

MAC mode imaging was employed to determine the S-layer ultra-structure on intact *T. forsythia* cells at the nanometer scale (Fig. 4.14)

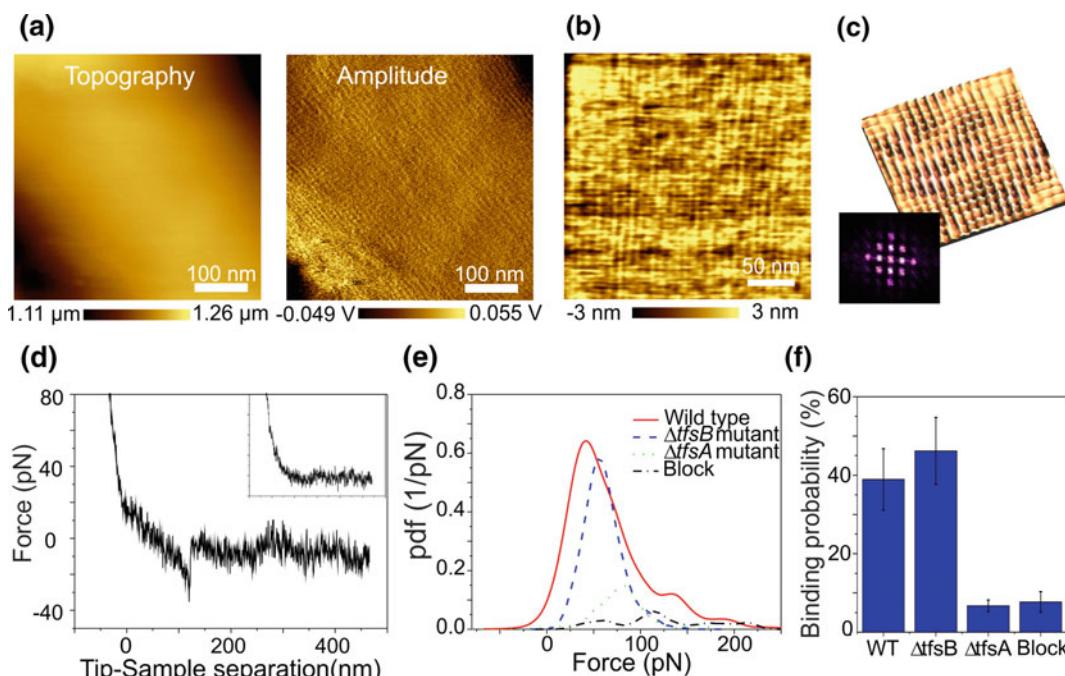


Fig. 4.14 Bacterial surface layer (S-layer) of *Tannerella forsythia*. **a–d** MAC-mode imaging of *T. forsythia* wild-type bacteria: **a** topography and amplitude image, **b** magnified topographic image, and **c** reconverted FFT 3D image of topography; *inset*, power spectrum of lattice. **d–f** SMFS to identify the specific interactions between anti-*TfsA* antibody and *TfsA* glycoprotein (*TfsA-GP*) in the S-layer lattice. **d** Typical unbinding event in a force-distance cycle (*inset* blocking of the specific interaction by injecting of recombinant *TfsA* protein into the bath solution), **e** distribution of unbinding forces between anti-*TfsA* antibody and the bacterial cell surface (wild type, *TfsA*-deficient mutant ($\Delta tfsA$), and *TfsB*-deficient mutant ($\Delta tfsB$)) at a retraction velocity of 1000 nm/s, and **f** comparison of binding probabilities of the anti-*TfsA* antibody-conjugated AFM tip to wild type, $\Delta tfsB$, and $\Delta tfsA$ cells, as well as the blocking experiment. For discussion see text.

(Sekot et al. 2012; Oh et al. 2013). For noninvasive imaging, 1 ml of *T. forsythia* culture was collected, washed and immobilized by mechanical trapping on 0.8 μm polycarbonate membranes. The wild type (wt) cells are $3.5 \pm 1 \mu\text{m}$ in length and $0.6 \pm 0.1 \mu\text{m}$ in width, resulting in a large curvature that makes proper scaling of the height image difficult. Hence, the amplitude image provides significantly better local contrast and the fine structure in the surface topology is more clearly resolved (Fig. 4.14b). The *T. forsythia* wt bacteria S-layer possesses a distinct, periodic square (p4) lattice, as clearly in the high-resolution reconverted 3D fast Fourier transform (FFT) image obtained from an enlarged topography (Fig. 4.14c). Peaks are due to surface periodicities, which are represented in the power spectra as the bright peaks near to the origin (inset in Fig. 4.14c). The S-layer lattice periodicity is $8.74 \pm 0.35 \text{ nm}$ with an angle of $90 \pm 4^\circ$, which is comparable to earlier observations made by electron microscopy on freeze-fractured, freeze dried, and ultrathin sectioned preparations (Sekot et al. 2012).

The AFM data clearly indicate that the outer membrane of intact live *T. forsythia* wt bacteria has a continuous, symmetrical 2 dimensional lattice structure built by two distinct S-layer protein species (*TfsA* and *TfsB*). In order to further explore the arrangement of *TfsA* and *TfsB* SMFS was used to identify the specific interactions between an anti-*TfsA* antibody and the *TfsA* glycoprotein (*TfsA-GP*) within the *TfsA/TfsB* S-layer lattice. The anti-*TfsA* antibody was

tethered to the AFM tip via a standard NHS-PEG-aldehyde linker utilizing a lysine residue as reactive site in the antibody. Force-distance curves were measured on intact *T. forsythia* wt cells indicating specific binding of the antibody to *TfsA-GP* on the bacterial S-layer (Fig. 4.14d). Each rupture event was used to calculate the distribution of specific unbinding forces, fitted with a Gaussian function, and added up to yield the empirical density function (PDF) as shown in Fig. 4.14e. Most probable unbinding forces between the anti-*TfsA* and S-layer *TfsA-GP* at a retraction velocity of 1000 nm/s were 41 pN with wt cells and 56 pN with *TfsB*-GP-deficient cells, where only the *TfsA-GP* is exposed on the cell surface. Differences of unbinding forces between wt and *TfsB*-deficient cells may occur due to the orientation and flexibility of the binding site of the *TfsA-GP* on the bacterial cell surface. The *TfsA-GP* on wt bacteria is embedded in the S-layer lattice structure, whereas the *TfsA-GP* does not appear to be that strictly fixed on the surface of the *TfsB*-deficient mutant where the space-filling counterpart (*TfsB-GP*) is missing. Binding was also highly specific, as verified from the binding probability (Fig. 4.14f).

4.5.3.2 AFM Indentation Experiments to Measure Physical Properties of *Escherichia Coli*

Bacteria produce various types of filamentous structures such as flagella, pili, and curli that

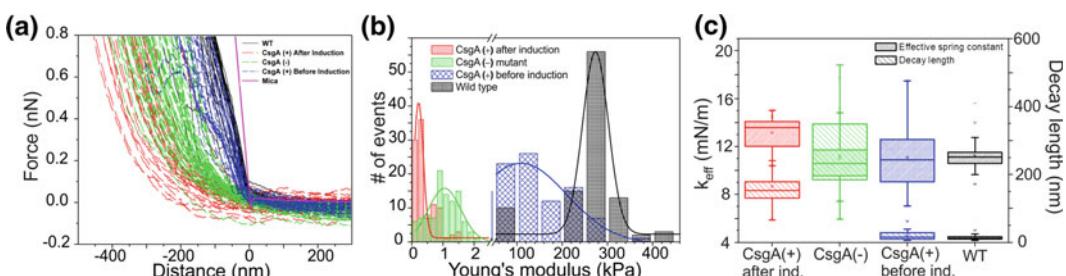


Fig. 4.15 AFM indentation measurements to determine the physical properties of *T. forsythia* bacteria depending on their capacity to produce curli. **a** Approach indentation curves measured on a mica surface, wild type, CsgA knock-out mutant (CsgA $^{-}$), and CsgA overexpressed mutant (CsgA $^{+}$) before curli expression, and after curli expression (by IPTG induction) on the bacterial surface. **b** Distribution of Young's modulus for the different mutant strains. For discussion see text.

promote bacterial adhesion onto the surface. In AFM indentation experiments the effects of curli on the physical properties of the bacterial membrane in *Escherichia coli* (*E. coli*) and the two mutant strains curli fibre protein CsgA deficient mutant (CsgA $^{-}$) and CsgA overexpressed mutant (CsgA $^{+}$) were studied (Oh et al. 2012). Figure 4.15a shows typical approach curves measured on the bacterial cell surfaces that reflect well the different stiffness values of the three strains in relation to the production of curli structures. The measured curves were fitted using the Hertz model for a conical tip to extract the Young's modulus values (Touhami et al. 2003; Chen et al. 2009; Kuznetsova et al. 2007) (Fig. 4.15b).

The effective spring constant of the cell wall and the decay length of the steric force sensed upon approaching the bacterial surface are a measure for the cell's turgor pressure (Chen et al. 2009; Arnoldi et al. 2000) and the thickness of the protruding proteins on the bacterial surface, respectively. The late linear parts of the approaching force-indentation curves were used to calculate the bacterial spring constant (effective spring constants), which were fitted using the models of Alexander and de Gennes, as modified for AFM analysis to quantify the steric interactions (e.g. decay lengths) (Alexander 1977; Comesano and Logan 2000; de Gennes 1987; Strauss et al. 2009) (Fig. 4.15c). The data show that the effective spring constant, i.e. the turgor pressure of the bacterial cells, was not affected by bacterial surface structures such as flagella, and curli fibers. Since there was no change in bacterial turgor pressure of the mutants compared to that of wt *E. coli*, this parameter most likely reflects the mechanical properties of the cell wall. Consequently, the observed decrease in Young's modulus has been induced by changes in bacterial surface properties. The decay length of the bacterial surface was clearly affected by CsgA curli production. These differences evidently indicate the presence of the CsgA protein on CsgA $^{+}$ overexpressed mutant after curli expression and the presence of flagella-like structure on CsgA $^{-}$ curli deficient mutant,

respectively, which both act as a polymer brush on the bacterial cell surface.

Overall the curli over expressed mutant after IPTG induction produced protruding CsgA protein structures that affected cell wall mechanical properties. This over-production of curli fiber decreased Young's modulus but increased the decay length on the bacterial surface, while maintaining the inner turgor pressure. This quantitative information obtained by AFM suggests that many critical issues of how mechanical parameters influence cellular behavior can be properly addressed using this method.

4.6 Perspectives and Concluding Remarks

A few years ago live cell AFM experiments still required a high-level of expertise, technologically advanced AFM equipment, and consequently were restricted to a few laboratories. Today, several manufacturers offer highly developed bio-AFM setups that are mounted on inverted optical microscopes or provided with upright optical configurations. They are equipped with environmental chambers and sample stages that offer great flexibility and control of experimental conditions. Data processing and analysis also for SMFS is being simplified more and more, as increasingly adequate software is integrated directly with the AFM instrumentation and specialized software has become available from different companies. All these advancements greatly facilitate operation of an AFM by scientist from various life science disciplines. Nevertheless, a basic understanding of the techniques and critical assessment of recorded data remain crucial to AFM experiments. Proper sample preparation and control of environmental conditions, surface chemistry, and the quality of the AFM probe, are equally important and which may alter over the course of an experiment (e.g. due to contamination) (El Kirat et al. 2005). Live-cell experiments in buffer or medium, for instance, have become a routine task and temperature control is available to most setups. However, such experiments

continue to require a great deal of patience and newcomers to AFM will generally need to practice before obtaining good data.

SMFS experiments are straightforward, but careful sample and AFM tip preparation together with critical data evaluation are of high importance (Dufrene et al. 2011; Noy 2011). While standardized tip-functionalization protocols become more and more available (Ebner et al. 2008c), validation of successful tip functionalization, i.e. that the ligand is attached and oriented correctly for binding usually happens during the actual experiment. As the interaction force is calculated from the bending of the cantilever, the spring constant has to be calibrated properly for each cantilever and may vary tremendously from the reported manufacturer value (te Riet et al. 2011). Further considerations in SMFS include the separation of specific (wanted) interactions from nonspecific (unwanted) interactions. Their distinction can be facilitated through the use of defined PEG-cross-linkers as described further above, but may remain challenging even under well-controlled conditions. Unequivocal demonstration of binding specificity by blocking experiments remains absolutely necessary. Finally, data interpretation and automation of force-spectroscopy analyses continue to be time consuming tasks that (partially) require in-house programmed algorithms. However, many AFM manufacturers have started to implement analysis algorithms into their data processing software and continue to develop automated routines that will contribute to simplify SMFS data evaluation.

Apart from its capabilities AFM remains primarily a surface technique. Consequently, in the past ligand receptor interactions on rigid surfaces and on cell membranes were acquired ‘blindly’, meaning that force curves were collected at individual, randomly selected sites, with no corresponding image of receptor distribution. The combination of AFM instrumentation with optical microscopy has eliminated this shortcoming and improved imaging and force spectroscopy applications on cells, with the possibility to correlate independently or simultaneously acquired topographic and fluorescence images. The continued

development of combined AFM and optical techniques like CLSM (confocal laser scanning microscopy) (Poole et al. 2004; Kassies et al. 2005; Doak et al. 2008), and high-resolution correlative microscopy methods such as TIRF (total internal reflection fluorescence) (Mathur et al. 2000; Trache and Lim 2009) or STED (stimulated emission depletion) and STORM (stochastic optical reconstruction microscopy) (Chacko et al. 2013) will further improve live cell applications and allow sophisticated experimental settings. In addition, remarkable advances have been made in developing AFM instruments with increased data acquisition rates, giving access to unprecedented time scales (millisecond resolution) (Casuso et al. 2011). Time resolution is a crucial factor that currently limits investigation of rapid and highly dynamic biological processes with conventional AFM instrumentation. Acquisition of ‘high’ resolution images takes several minutes, where dynamic processes in biology usually occur at much shorter time-scales. The ongoing improvements in the field of high-speed AFM open up fascinating new perspectives to explore cellular dynamics and monitor molecule interactions with receptors (Shibata et al. 2010; Ando 2012; Rico et al. 2013; Ando et al. 2014; Preiner et al. 2014; Karner et al. 2016). In the future, high-speed AFM force distance based imaging modes, will become more and more important in this field. These techniques will allow direct observations of individual molecules in action at high spatiotemporal resolution and provide simultaneous nanoscale mapping of their properties.

To summarize this chapter, we have provided a snapshot of AFM techniques with a special focus on force spectroscopy and recognition imaging for biomedical sensing applications. Our declared aim was to introduce in particular researchers of the life sciences, who may be new to these techniques, to the wealth of experimental possibilities provided by the AFM toolbox. The highlighted examples in the last section illustrate the considerable potential of the force-measuring tool SMFS and the recognition imaging technique TREC to decipher receptor and cell responses to external stimuli and to explore the

interaction of cell receptors with individual (drug) molecules. Studies on artificial and native membranes, tissue samples and live cells hold great promise to increase our understanding of medically relevant molecular processes in disease patterns or pathogen invasion and may help in the development of new therapeutic approaches and pharmaceuticals.

Acknowledgements Authors and projects presented in this chapter received in part funding from the European Union's Framework Programme for Research and Innovation Horizon 2020 under Marie Skłodowska-Curie grant No. 656842 (to CL), the Austrian Research Promotion Agency (FFG) under MNT-ERA-net project No. 823980 (to AE), project No. 850681 (to LAC), the Austrian Science Foundation (FWF) under project(s) SFB-F35 (to PH), P25357 (to PH), I 1767-B11 (to PH), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under FOR1543 (to PH), a fellowship from the Austrian Programme for Advanced Research and Technology (APART) of the Austrian Academy of Science and a Marie Curie International Incoming Fellowship of the 7th European Community Framework Program (both to YJO).

References

- Adamcik J, Berquand A, Mezzenga R (2011) Single-step direct measurement of amyloid fibrils stiffness by peak force quantitative nanomechanical atomic force microscopy. *Appl Phys Lett* 98(19):3701
- Ahmad SF et al (2011) Nanosensing of Fc gamma receptors on macrophages. *Anal Bioanal Chem* 399 (7):2359–2367
- Arnoldi M et al (2000) Bacterial turgor pressure can be measured by atomic force microscopy. *Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics* 62(1 Pt B):1034–1044
- Alessandrini A, Facci P (2005) AFM: a versatile tool in biophysics. *Meas Sci Technol* 16(6):R65–R92
- Alexander S (1977) Adsorption of chain molecules with a polar head a scaling description. *J Phys France* 38 (8):983–987
- Almqvist N et al (2004) Elasticity and adhesion force mapping reveals real-time clustering of growth factor receptors and associated changes in local cellular rheological properties. *Biophys J* 86(3):1753–1762
- Alsteens D et al (2012) High-resolution imaging of chemical and biological sites on living cells using peak force tapping atomic force microscopy. *Langmuir* 28(49):16738–16744
- Alsteens D et al (2015) Imaging G protein-coupled receptors while quantifying their ligand-binding free-energy landscape. *Nat Methods* 12(9):845–851
- Ando T (2012) High-speed atomic force microscopy coming of age. *Nanotechnology* 23(6):062001
- Ando T, Uchihashi T, Scheuring S (2014) Filming biomolecular processes by high-speed atomic force microscopy. *Chem Rev*
- Aponte-Santamaría C et al (2015) Force-sensitive autoinhibition of the von Willebrand factor is mediated by interdomain interactions. *Biophys J* 108(9):2312–2321
- Auletta T et al (2004) Beta-cyclodextrin host-guest complexes probed under thermodynamic equilibrium: thermodynamics and AFM force spectroscopy. *J Am Chem Soc* 126(5):1577–1584
- Barattin R, Voyer N (2011) Chemical modifications of atomic force microscopy tips. In: Braga PC, Ricci D (ed) *Atomic force microscopy in biomedical research*. Humana Press, pp 457–483
- Baumgartner W et al (2000) Cadherin interaction probed by atomic force microscopy. *Proc Natl Acad Sci USA* 97(8):4005–4010
- Baus D et al (2008) Identification of a novel AS160 splice variant that regulates GLUT4 translocation and glucose-uptake in rat muscle cells. *Cell Signal* 20 (12):2237–2246
- Beck V et al (2007) Polyunsaturated fatty acids activate human uncoupling proteins 1 and 2 in planar lipid bilayers. *FASEB J* 21(4):1137–1144
- Bell GI (1978) Models for the specific adhesion of cells to cells. *Science* 200(4342):618–627
- Berardi MJ et al (2011) Mitochondrial uncoupling protein 2 structure determined by NMR molecular fragment searching. *Nature* 476(7358):109–113
- Binnig G, Quate CF, Gerber C (1986) Atomic force microscope. *Phys Rev Lett* 56(9):930–933
- Bizzarri AR, Cannistraro S (2010) The application of atomic force spectroscopy to the study of biological complexes undergoing a biorecognition process. *Chem Soc Rev* 39(2):734–749
- Bongrand P et al (1988) Use of hydrodynamic flows to study cell adhesion. *Phys Basis of Cell–Cell Adhes* 125–156
- Bruzzone S et al (2012) The plant hormone abscisic acid increases in human plasma after hyperglycemia and stimulates glucose consumption by adipocytes and myoblasts. *Faseb J* 26(3):1251–1260
- Camesano TA, Logan BE (2000) Probing bacterial electrosteric interactions using atomic force microscopy. *Environ Sci Technol* 34(16):3354–3362
- Cartagena A et al (2013) Mapping in vitro local material properties of intact and disrupted virions at high resolution using multi-harmonic atomic force microscopy. *Nanoscale* 5(11):4729–4736
- Carvalho FA, Santos NC (2012) Atomic force microscopy-based force spectroscopy in biological and biomedical applications. *IUBMB Life* 64 (6):465–472
- Carvalho E et al (2004) GLUT4 overexpression or deficiency in adipocytes of transgenic mice alters the composition of GLUT4 vesicles and the subcellular

- localization of GLUT4 and insulin-responsive aminopeptidase. *J Biol Chem* 279(20):21598–21605
- Casuso I, Rico F, Scheuring S (2011) Biological AFM: where we come from—where we are—where we may go. *J Mol Recognit* 24(3):406–413
- Chacko JV, Zanacchi FC, Diaspro A (2013) Probing cytoskeletal structures by coupling optical superresolution and AFM techniques for a correlative approach. *Cytoskeleton* 70(11):729–740
- Chalabi M et al (2010) Periodontopathic bacteria and herpesviruses in chronic periodontitis. *Mol Oral Microbiol* 25(3):236–240
- Chen A, Moy VT (2000) Cross-linking of cell surface receptors enhances cooperativity of molecular adhesion. *Biophys J* 78(6):2814–2820
- Chen Y-Y et al (2009) Surface rigidity change of *Escherichia coli* after filamentous bacteriophage infection. *Langmuir* 25(8):4607–4614
- Chtcheglova LA et al (2007) Nano-scale dynamic recognition imaging on vascular endothelial cells. *Biophys J* 93(2):L11–L13
- Chtcheglova LA et al (2008) Localization of the ergtoxin-1 receptors on the voltage sensing domain of hERG K⁺ channel by AFM recognition imaging. *Pflügers Archiv-Eur J Physiol* 456(1):247–254
- Chtcheglova LA et al (2010) AFM functional imaging on vascular endothelial cells. *J Mol Recognit* 23 (6):589–596
- Coleman JA, Green EM, Gouaux E (2016) X-ray structures and mechanism of the human serotonin transporter. *Nature* 532(7599):334–339
- Creasey R et al (2010) Detecting protein aggregates on untreated human tissue samples by atomic force microscopy recognition imaging. *Biophys J* 99 (5):1660–1667
- Creasey R et al (2011) Atomic force microscopy-based antibody recognition imaging of proteins in the pathological deposits in Pseudoexfoliation syndrome. *Ultramicroscopy* 111(8):1055–1061
- de Gennes PG (1987) Polymers at an interface; a simplified view. *Adv Colloid Interface Sci* 27 (3):189–209
- Doak S et al (2008) High-resolution imaging using a novel atomic force microscope and confocal laser scanning microscope hybrid instrument: essential sample preparation aspects. *Histochem Cell Biol* 130 (5):909–916
- Doktycz MJ et al (2003) AFM imaging of bacteria in liquid media immobilized on gelatin coated mica surfaces. *Ultramic* 97(1–4):209–216
- Dorobantu LS, Gray MR (2010) Application of atomic force microscopy in bacterial research. *Scanning* 32 (2):74–96
- Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96(10):108101
- Dudko OK, Hummer G, Szabo A (2008) Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci USA* 105(41):15755–15760
- Dufrene YF (2002) Atomic force microscopy, a powerful tool in microbiology. *J Bacteriol* 184(19):5205–5213
- Dufrêne YF (2015) Sticky microbes: forces in microbial cell adhesion. *Trends Microbiol* 23(6):376–382
- Dufrene YF et al (2011) Five challenges to bringing single-molecule force spectroscopy into living cells. *Nat Methods* 8(2):123–127
- Dulebo A et al (2009) Second harmonic atomic force microscopy imaging of live and fixed mammalian cells. *Ultramicroscopy* 109(8):1056–1060
- Duman M et al (2010) Improved localization of cellular membrane receptors using combined fluorescence microscopy and simultaneous topography and recognition imaging. *Nanotechnology* 21(11):115504
- Duman M et al (2013) Nanomapping of CD1d-glycolipid complexes on THP1 cells by using simultaneous topography and recognition imaging. *J Mol Recognit* 26(9):408–414
- Ebner A et al (2005) Localization of single avidin–biotin interactions using simultaneous topography and molecular recognition imaging. *ChemPhysChem* 6 (5):897–900
- Ebner A, Hinterdorfer P, Gruber HJ (2007a) Comparison of different aminofunctionalization strategies for attachment of single antibodies to AFM cantilevers. *Ultramicroscopy* 107(10–11):922–927
- Ebner A et al (2007b) A new, simple method for linking of antibodies to atomic force microscopy tips. *Bioconjug Chem* 18(4):1176–1184
- Ebner A et al (2008a) Functionalization of probe tips and supports for single-molecule recognition force Microscopy. *Top Curr Chem* 285:29–76
- Ebner A et al (2008b) Determination of CFTR densities in erythrocyte plasma membranes using recognition imaging. *Nanotechnology* 19(38):384017
- Ebner A et al (2008c) Functionalization of probe tips and supports for single-molecule recognition force Microscopy. In: Samori P (ed) *Stm and Afm studies on*. Springer-Verlag Berlin, Berlin, pp 29–76
- El Kirat K et al (2005) Sample preparation procedures for biological atomic force microscopy. *J Microsc* 218 (3):199–207
- El-Kirat-Chatel S, Dufrene YF (2012) Nanoscale imaging of the Candida—macrophage Interaction using correlated fluorescence-atomic force microscopy. *ACS Nano* 6(12):10792–10799
- Engel A, Gaub HE (2008) Structure and mechanics of membrane proteins. *Annu Rev Biochem* 77:127–148
- Evans E, Ritchie K (1997) Dynamic strength of molecular adhesion bonds. *Biophys J* 72(4):1541–1555
- Evans E, Ritchie K (1999) Strength of a weak bond connecting flexible polymer chains. *Biophys J* 76 (5):2439–2447
- Florin EL, Moy VT, Gaub HE (1994) Adhesion forces between individual ligand-receptor pairs. *Science* 264 (5157):415–417

- Franz CM, Muller DJ (2005) Analyzing focal adhesion structure by atomic force microscopy. *J Cell Sci* 118 (22):5315–5323
- Friddle RW, Noy A, De Yoreo JJ (2012) Interpreting the widespread nonlinear force spectra of intermolecular bonds. *Proc Natl Acad Sci* 109(34):13573–13578
- Friedrichs J et al (2013) A practical guide to quantify cell adhesion using single-cell force spectroscopy. *Methods* 60(2):169–178
- Fritz J et al (1998) Force-mediated kinetics of single P-selectin/ligand complexes observed by atomic force microscopy. *Proc Natl Acad Sci USA* 95(21):12283–12288
- Garcia R, Herruzo ET (2012) The emergence of multi-frequency force microscopy. *Nat Nano* 7(4):217–226
- Govers R (2014) Molecular mechanisms of GLUT4 regulation in adipocytes. *Diabetes & Metabolism* 40 (6):400–410
- Gracia R (ed) (2010) Amplitude modulation atomic force microscopy. Wiley-VCH Verlag GmbH & Co. KGaA, 6 Oct 2010
- Grandbois M et al (2000) Affinity imaging of red blood cells using an atomic force microscope. *J Histochem Cytochem* 48(5):719–724
- Grubmüller H, Heymann B, Tavan P (1996) Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271 (5251):997–999
- Hane F, Attwood S, Leonenko Z (2014) Comparison of three competing dynamic force spectroscopy models to study binding forces of amyloid- β (1–42). *Soft Matter* 10(12):1924–1930
- Hansma P et al (1994) Tapping mode atomic force microscopy in liquids. *Appl Phys Lett* 64(13):1738–1740
- Heu C et al (2012) Glyphosate-induced stiffening of HaCaT keratinocytes, a peak force tapping study on living cells. *J Struct Biol* 178(1):1–7
- Hinterdorfer P et al (1996) Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proc Natl Acad Sci USA* 93(8):3477–3481
- Hinterdorfer P et al (2000) Poly (ethylene glycol): an ideal spacer for molecular recognition force microscopy/spectroscopy. *Single Mol* 1(2):99–103
- Horber JKH, Miles MJ (2003) Scanning probe evolution in biology. *Science* 302(5647):1002–1005
- Hugel T, Seitz M (2001) The study of molecular interactions by AFM force spectroscopy. *Macromol Rapid Commun* 22(13):989–1016
- Israelachvili JN (2011) Intermolecular and surface forces, 3rd edn. Academic Press, London & New York
- Izrailev S et al (1997) Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys J* 72 (4):1568–1581
- Jeremic A et al (2003) Reconstituted fusion pore. *Biophys J* 85(3):2035–2043
- Jiang ZY et al (2002) A phosphatidylinositol 3-kinase-independent insulin signaling pathway to N-WASP/Arp2/3/F-actin required for GLUT4 glucose transporter recycling. *J Biol Chem* 277(1):509–515
- Kadan S et al (2013) In vitro evaluations of cytotoxicity of eight antidiabetic medicinal plants and their effect on GLUT4 translocation. *Evid-Based Complement Altern Med*
- Kamruzzahan ASM et al (2006) Antibody linking to atomic force microscope tips via disulfide bond formation. *Bioconjug Chem* 17(6):1473–1481
- Karner A et al (2016) Tuning membrane protein mobility by confinement into nanodomains. *Nat Nanotechnol* doi:10.1038/nnano.2016.236
- Kasas S, Dietler G (2008) Probing nanomechanical properties from biomolecules to living cells. *Pflügers Archiv Eur J Physiol* 456(1):13–27
- Kassies R et al (2005) Combined AFM and confocal fluorescence microscope for applications in bio-nanotechnology. *J Microsc* 217(Pt 1):109–116
- Kienberger F et al (2000) Static and dynamical properties of single poly (ethylene glycol) molecules investigated by force spectroscopy. *Single Mol* 1(2):123–128
- Kienberger F et al (2006) Molecular recognition imaging and force spectroscopy of single biomolecules. *Acc Chem Res* 39(1):29–36
- Kim IH et al (2011) Nanoscale mapping and affinity constant measurement of signal-transducing proteins by atomic force microscopy. *Anal Chem* 83(5):1500–1503
- Kim DH et al (2015) Cytosolic targeting factor AKR2A captures chloroplast outer membrane-localized client proteins at the ribosome during translation. *Nat Commun* 6:6843
- Klingenberg M, Echternach KS (2001) Uncoupling proteins: the issues from a biochemist point of view. *Biochim Biophys Acta* 1504(1):128–143
- Koehler M et al (2017) Combined recognition imaging and force spectroscopy: a new mode for mapping and studying interaction sites at low lateral density. *Sci Adv Mater* 9(1):128–134
- Kristensen AS et al (2011) SLC6 neurotransmitter transporters: structure, function, and regulation. *Pharmacol Rev* 63(3):585–640
- Kristiansen S, Richter EA (2002) GLUT4-containing vesicles are released from membranes by phospholipase D cleavage of a GPI anchor. *Am J Physiol-Endocrinol Metabolism* 283(2):E374–E382
- Kumar V et al (2014) Novel and high affinity fluorescent ligands for the serotonin transporter based on (S)-citalopram. *ACS Med Chem Lett* 5(6):696–699
- Kurland NE, Drira Z, Yadavalli VK (2012) Measurement of nanomechanical properties of biomolecules using atomic force microscopy. *Micron* 43(2–3):116–128
- Kuznetsova TG et al (2007) Atomic force microscopy probing of cell elasticity. *Micron* 38(8):824–833
- Lamprecht C, Hinterdorfer P, Ebner A (2014) Applications of biosensing atomic force microscopy in monitoring drug and nanoparticle delivery. *Expert Opin Drug Deliv* 11(8):1237–1253
- Lanzendorfer P et al (2014) Identification of novel insulin mimetic drugs by quantitative total internal reflection fluorescence (TIRF) microscopy. *Br J Pharmacol* 171 (23):5237–5251

- Lee GU, Kidwell DA, Colton RJ (1994) Sensing discrete streptavidin-biotin interactions with atomic force microscopy. *Langmuir* 10(2):354–357
- Lehenkari P et al (2000) Adapting atomic force microscopy for cell biology. *Ultramicroscopy* 82(1):289–295
- Leitner M et al (2011) Single-molecule AFM characterization of individual chemically tagged DNA tetrahedra. *ACS Nano* 5(9):7048–7054
- Lidell ME, Enerbäck S (2010) Brown adipose tissue—a new role in humans? *Nat Rev Endocrinol* 6(6):319–325
- Lin C, Klingenberg M (1980) Isolation of the uncoupling protein from brown adipose tissue mitochondria. *FEBS Lett* 113(2):299–303
- Lin L et al (2006) Recognition imaging with a DNA aptamer. *Biophys J* 90(11):4236–4238
- Lin L et al (2007) In vitro selection of histone H4 aptamers for recognition imaging microscopy. *J Am Chem Soc* 129(47):14568–14569
- Ludwig M, Dettmann W, Gaub H (1997) Atomic force microscope imaging contrast based on molecular recognition. *Biophys J* 72(1):445
- Madl J et al (2006) A combined optical and atomic force microscope for live cell investigations. *Ultramicroscopy* 106(8–9):645–651
- Mahut M et al (2012) Time-resolved chloroquine-induced relaxation of supercoiled plasmid DNA. *Anal Bioanal Chem* 402(1):373–380
- Marszalek PE, Dufrêne YF (2012) Stretching single polysaccharides and proteins using atomic force microscopy. *Chem Soc Rev* 41(9):3523–3534
- Mathur AB, Truskey GA, Reichert WM (2000) Atomic force and total internal reflection fluorescence microscopy for the study of force transmission in endothelial cells. *Biophys J* 78(4):1725–1735
- Merkel R et al (1999) Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature* 397(6714):50–53
- Modrianský M et al (1997) Identification by site-directed mutagenesis of three arginines in uncoupling protein that are essential for nucleotide binding and inhibition. *J Biol Chem* 272(40):24759–24762
- Monod J, Changeux J-P, Jacob F (1963) Allosteric proteins and cellular control systems. *J Mol Biol* 6 (4):306–329
- Mueller DJ, Dufrene YF (2011) Atomic force microscopy: a nanoscopic window on the cell surface. *Trends Cell Biol* 21(8):461–469
- Muller DJ et al (2002) Observing structure, function and assembly of single proteins by AFM. *Prog Biophys Mol Biol* 79(1–3):1–43
- Nedergaard J, Ricquier D, Kozak LP (2005) Uncoupling proteins: current status and therapeutic prospects. *EMBO Rep* 6(10):917–921
- Neuert G et al (2006) Dynamic force spectroscopy of the digoxigenin-antibody complex. *FEBS Lett* 580 (2):505–509
- Nicholls DG (1976) Hamster Brown-adipose-tissue mitochondria. *Eur J Biochem* 62(2):223–228
- Nicholls DG, Bernson VS, Heaton GM (1978) The identification of the component in the inner membrane of brown adipose tissue mitochondria responsible for regulating energy dissipation, in *Effectors of Thermoogenesis*. Springer, pp 89–93
- Nimmervoll B et al (2015) Cell surface localised Hsp70 is a cancer specific regulator of clathrin-independent endocytosis. *FEBS Lett* 589(19):2747–2753
- Noy A (2011) Force spectroscopy 101: how to design, perform, and analyze an AFM-based single molecule force spectroscopy experiment. *Curr Opin Chem Biol* 15(5):710–718
- Noy A, Friddle RW (2013) Practical single molecule force spectroscopy: how to determine fundamental thermodynamic parameters of intermolecular bonds with an atomic force microscope. *Methods* 60(2):142–150
- Oh YJ et al (2012) Characterization of Curli A production on living bacterial surfaces by scanning probe microscopy. *Biophys J* 103(8):1666–1671
- Oh YJ et al (2013) Characterizing the S-layer structure and anti-S-layer antibody recognition on intact *Tannerella forsythia* cells by scanning probe microscopy and small angle X-ray scattering. *J Mol Recognit: JMR* 26(11):542–549
- Oh et al (2016) Curli mediate bacterial adhesion to fibronectin via tensile multiple bonds. *Sci Rep* 6:33909
- Ott W et al (2016) Single-molecule force spectroscopy on polyproteins and receptor–ligand complexes: the current toolbox. *J Struct Biol*
- Pebay-Peyroula E et al (2003) Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* 426(6962):39–44
- Penmatsa A, Wang KH, Gouaux E (2013) X-ray structure of dopamine transporter elucidates antidepressant mechanism. *Nature* 503(7474):85–90
- Persat A et al (2015) The mechanical world of bacteria. *Cell* 161(5):988–997
- Pfister G et al (2005) Detection of HSP60 on the membrane surface of stressed human endothelial cells by atomic force and confocal microscopy. *J Cell Sci* 118(Pt 8):1587–1594
- Pillet F et al (2014) Atomic force microscopy and pharmacology: from microbiology to cancerology. *Biochim Biophys Acta* 1840(3):1028–1050
- Plenge P, Mellerup ET (1997) An affinity-modulating site on neuronal monoamine transport proteins. *Pharmacol Toxicol* 80(4):197–201
- Plenge P et al (2012) Steric hindrance mutagenesis in the conserved extracellular vestibule impedes allosteric binding of antidepressants to the serotonin transporter. *J Biol Chem* 287(47):39316–39326
- Poole K et al (2004) The effect of raft lipid depletion on microvilli formation in MDCK cells, visualized by atomic force microscopy. *FEBS Lett* 565(1–3):53–58

- Posch S (2015) Collagen surface functionalization—a new strategy for molecular recognition force spectroscopy. *Imaging Microsc* 1
- Posch S et al (2013) Activation induced morphological changes and integrin $\alpha IIb\beta 3$ activity of living platelets. *Methods* 60(2):179–185
- Posch S et al (2016) Mutual A domain interactions in the force sensing protein von Willebrand Factor. *J Struct Biol*, under review
- Preiner J et al (2007) Higher harmonic atomic force microscopy: imaging of biological membranes in liquid. *Phys Rev Lett* 99(4):046102
- Preiner J et al (2009) Simultaneous topography and recognition imaging: physical aspects and optimal imaging conditions. *Nanotechnology* 20(21):215103
- Preiner J et al (2014) IgGs are made for walking on bacterial and viral surfaces. *Nat Commun* 5:4394
- Preiner J et al (2015) High-speed AFM images of thermal motion provide stiffness map of interfacial membrane protein moieties. *Nano Lett* 15(1):759–763
- Raab A et al (1999) Antibody recognition imaging by force microscopy. *Nat Biotech* 17(9):902–905
- Rangl M et al (2009) Stable, non-destructive immobilization of native nuclear membranes to micro-structured PDMS for single-molecule force spectroscopy. *Chem-PhysChem* 10(9–10):1553–1558
- Rankl C et al (2010) Detection of corneodesmosin on the surface of stratum corneum using atomic force microscopy. *Exp Dermatol* 19(11):1014–1019
- Raposo M, Ferreira Q, Ribeiro P (2007) A guide for atomic force microscopy analysis of soft-condensed matter. *Mod Res Educ Top Microsc* 1:758–769
- Rico F, Su C, Scheuring S (2011) Mechanical mapping of single membrane proteins at submolecular resolution. *Nano Lett* 11(9):3983–3986
- Rico F et al (2013) High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science* 342(6159):741–743
- Riener CK et al (2003) Simple test system for single molecule recognition force microscopy. *Anal Chim Acta* 479(1):59–75
- Ritzefeld M et al (2013) Analysis of DNA interactions using single-molecule force spectroscopy. *Amino Acids* 44(6):1457–1475
- Ros R et al (1998) Antigen binding forces of individually addressed single-chain Fv antibody molecules. *Proc Natl Acad Sci USA* 95(13):7402–7405
- Scheuring S, Levy D, Rigaud JL (2005) Watching the components of photosynthetic bacterial membranes and their in situ organisation by atomic force microscopy. *Biochim Biophys Acta* 1712(2):109–127
- Schlierf M, Rief M (2006) Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape. *Biophys J* 90(4):L33–L35
- Sekatskii S, Benedetti F, Dietler G (2013) Dependence of the most probable and average bond rupture force on the force loading rate: first order correction to the Bell-Evans model. *J Appl Phys* 114(3):034701
- Sekot G et al (2011) Potential of the *Tannerella forsythia* S-layer to delay the immune response. *J Dent Res* 90 (1):109–114
- Sekot G et al (2012) Analysis of the cell surface layer ultrastructure of the oral pathogen *Tannerella forsythia*. *Arch Microbiol* 194(6):525–539
- Shibata M et al (2010) High-speed atomic force microscopy shows dynamic molecular processes in photoactivated bacteriorhodopsin. *Nat Nano* 5(3):208–212
- Strauss J, Burnham NA, Camesano TA (2009) Atomic force microscopy study of the role of LPS O-antigen on adhesion of *E. coli*. *J Mol Recognit* 22(5):347–355
- Stroh C et al (2004a) Single-molecule recognition imaging microscopy. *Proc Natl Acad Sci USA* 101 (34):12503–12507
- Stroh CM et al (2004b) Simultaneous topography and recognition imaging using force microscopy. *Biophys J* 87(3):1981–1990
- Strunz T et al (2000) Model energy landscapes and the force-induced dissociation of ligand-receptor bonds. *Biophys J* 79(3):1206–1212
- Tang J et al (2008) Recognition imaging and highly ordered molecular templating of bacterial S-layer nanoarrays containing affinity-tags. *Nano Lett* 8 (12):4312–4319
- Tang J et al (2009) Detection of metal binding sites on functional S-layer nanoarrays using single molecule force spectroscopy. *J Struct Biol* 168(1):217–222
- te Riet J et al (2011) Interlaboratory round robin on cantilever calibration for AFM force spectroscopy. *Ultramicroscopy* 111(12):1659–1669
- Touhami A, Nysten B, Dufrêne YF (2003) Nanoscale mapping of the elasticity of microbial cells by atomic force microscopy. *Langmuir* 19(11):4539–4543
- Trache A, Lim SM (2009) Integrated microscopy for real-time imaging of mechanotransduction studies in live cells. *J Biomed Opt* 14(3):034024
- Ulman A (2013) An introduction to ultrathin organic films: from Langmuir–Blodgett to self-assembly. Academic press
- Urbánková E et al (2003) Transport kinetics of uncoupling proteins analysis of UCP1 reconstituted in planar lipid bilayers. *J Biol Chem* 278(35):32497–32500
- Walton EB, Lee S, Van Vliet KJ (2008) Extending Bell's model: how force transducer stiffness alters measured unbinding forces and kinetics of molecular complexes. *Biophys J* 94(7):2621–2630
- Wang H et al (2008) Single-epitope recognition imaging of native chromatin. *Epigenetics Chromatin* 1(1):1
- Wang B et al (2012) High-resolution single-molecule recognition imaging of the molecular details of ricin-aptamer interaction. *J Phys Chem B* 116(17):5316–5322
- Wang KH, Penmatsa A, Gouaux E (2015) Neurotransmitter and psychostimulant recognition by the dopamine transporter. *Nature* 521(7552):322–327
- Whited AM, Park PSH (2014) Atomic force microscopy: a multifaceted tool to study membrane proteins and

- their interactions with ligands. *Biochim Biophys Acta (BBA)-Biomemb* 1838(1):56–68
- Wildling L et al (2005) Linking of underivatized antibody molecules to AFM tips via long, distensible poly (ethylene glycol) chains. *Biophys J* 88(1):154A–155A
- Wildling L et al (2011) Linking of sensor molecules with amino groups to amino-functionalized AFM tips. *Bioconjug Chem* 22(6):1239–1248
- Wildling L et al (2012) Probing binding pocket of serotonin transporter by single molecular force spectroscopy on living cells. *J Biol Chem* 287(1):105–113
- Willemsen OH et al (1998) Simultaneous height and adhesion imaging of antibody-antigen interactions by atomic force microscopy. *Biophys J* 75(5):2220–2228
- Withers JR, Aston DE (2006) Nanomechanical measurements with AFM in the elastic limit. *Adv Colloid Interface Sci* 120(1–3):57–67
- Yamashita A et al (2005) Crystal structure of a bacterial homologue of Na⁺/Cl⁻-dependent neurotransmitter transporters. *Nature* 437(7056):215–223
- Zaman MS et al (2005) Imaging and analysis of *Bacillus anthracis* spore germination. *Microsc Res Tech* 66(6):307–311
- Zhang M et al (2013) Mapping out the structural changes of natural and pretreated plant cell wall surfaces by atomic force microscopy single molecular recognition imaging. *Biotechnol Biofuels* 6(1):1
- Zhang J et al (2014) Nanoscale organization of human GnRH-R on human bladder cancer cells. *Anal Chem* 86(5):2458–2464
- Zhu R et al (2013) Mapping the nucleotide binding site of uncoupling protein 1 using atomic force microscopy. *J Am Chem Soc* 135(9):3640–3646
- Zhu R et al (2015) Nanopharmacological force sensing to reveal allosteric coupling in transporter binding sites. *Angew Chem Int Ed*

Part II
Nanomechanics

Depth-sensing Nanoindentation Techniques for Nanomechanical Characterization

5

Bharat Bhushan

Abstract

To measure nanomechanical properties of surface layers of bulk materials and thin films, depth-sensing nanoindentation measurement techniques are used commonly. The nanoindentation apparatus continuously monitors the load and the position of the indenter relative to the surface of the specimen (depth of an indent or displacement) during the indentation process. Indentation experiments can be performed at a penetration depth of as low as about 5 nm. This chapter presents an overview of various nanoindentation techniques, various measurement options, and data analysis. Data on elastic-plastic deformation behavior, hardness, elastic modulus, scratch resistance, film-substrate adhesion, residual stresses, time-dependent creep and relaxation properties, fracture toughness, and fatigue are presented.

Keywords

Nanoindentation · Depth-sensing nanoindentation · Picoindentation · Continuous stiffness · Thin films · Hardness · Elastic modulus · Scratch resistance · Adhesion · Stiffness · Creep · Toughness · Fatigue

5.1 Introduction

Mechanical properties of solid surfaces and thin films are important in various applications including those requiring tribological performance (Bhushan 1996, 1999a, 2001, 2011, 2013a, b, 2017). Among the mechanical properties of interest are elastic-plastic deformation behavior, hardness, elastic modulus, scratch resistance, film-substrate adhesion, residual

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: bhushan.2@osu.edu

stresses, time-dependent creep and relaxation properties, fracture toughness, and fatigue. Indentation measurements can assess structural heterogeneities on and underneath the surface such as the presence of buried layers, grain boundaries, graded material, and multilayered structures. These properties can be obtained using commercial and specialized indentation testers.

Physical contacts between sliding interfaces in magnetic storage devices and micro/nanoelectromechanical systems (MEMS/NEMS) occur at very low loads. Thus, friction and wear of sliding surfaces is controlled primarily by the physical and chemical properties of a few surface atomic layers. Ultrathin films, as thin as 1–2 nm, and nanocomposites are used in many applications including magnetic storage devices and MEMS/NEMS (Bhushan 1996, 2017). Measurements of surface layers and ultrathin films and nanocomposites require specialized instrumentation. Depth-sensing nanoindenter and atomic force microscope (AFM) are used commonly for measurements on the nanoscale (Bhushan et al. 1996a; Bhushan 1999a; Bhushan and Li 2003).

Mechanical properties of materials on the nanoscale are size-dependent (Bhushan et al. 1996a, 2014; Bhushan and Li 2003). Nanomechanical characterizations have been motivated partly by the development of nanocomposites and the application of nanometer-thick films for miniaturization of engineering and electronic components, and partly by newly available methods of probing mechanical properties in small volumes. Compared to their bulk material counterparts, nano-objects exhibit enhanced mechanical, electrical, magnetic, chemical, friction, and wear-reducing properties (Bhushan et al. 2014; Maharaj and Bhushan 2015). Nano-objects include nanoparticles, nanorods, and nanowires. Nanomechanical characterization of nano-objects is needed.

A depth-sensing nanoindenter is an important tool for probing the mechanical properties of small volumes of material. Indentation load-displacement data contain a wealth of information. From the load-displacement data, many mechanical properties such as hardness

and elastic modulus can be determined without imaging the indentation impression. The nanoindenter also has been used to measure the fracture toughness and fatigue properties of ultrathin films, which cannot be measured by conventional indentation tests. Nanoscratch, nanowear, and adhesion tests can be performed at ramping loads using a tangential force sensor. AFMs are ideal for imaging of nanometer-scale indents, providing useful information about nanoindentation deformation (Bhushan 1999a, 2011, 2017). When an indentation system is used in conjunction with an AFM, *in situ* imaging can be obtained.

In this chapter, we review two commercial depth-sensing nanoindentation test apparatuses, followed by various measurement options, data analysis, and their use for determination of various mechanical properties of interest (Bhushan et al. 1996a; Bhushan 1999a, 2011; Bhushan and Li 2003). Emphasis here is on measurements of surface layers of bulk materials and ultrathin films at shallow indentation depths as small as 5–20 nm.

5.2 Nanoindentation Measurement Techniques and Data Analysis

Hardness measures resistance to local deformation. The most commonly used hardness measurements are: scratch hardness and static indentation hardness (Tabor 1951). Scratch hardness is the oldest form of hardness measurement. It depends on the ability of one material to scratch another or to be scratched by another solid. Solid and thin-film surfaces are scratched by a sharp stylus made of hard material, typically diamond. As a measure of scratch hardness and/or interfacial adhesion, either the loads required to scratch or fracture the surface or delaminate the film are used or the normal/tangential load-scratch size relationships are used (Heavens 1950; Tabor 1951, 1970; Benjamin and Weaver 1960; Campbell 1970; Ahn et al. 1978; Bhushan 1987; Wu 1991; Bhushan et al. 1995, 1996a; Bhushan and Gupta 1995; Hainsworth et al. 1996; Blau et al. 1997; Bhushan 1999a; Bhushan and Li 2003).

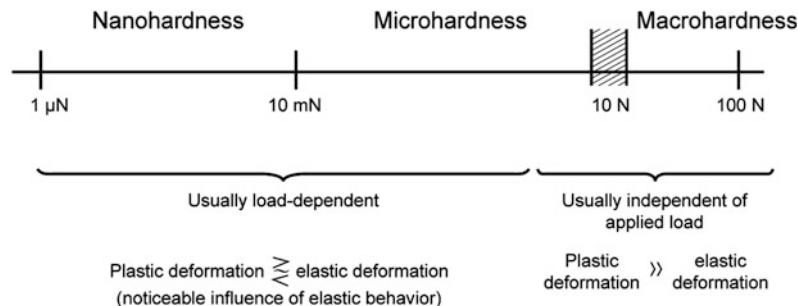
The methods most widely used in determining the hardness of materials are the (quasi) static indentation methods. Indentation hardness is essentially a measure of the plastic deformation properties of materials and only to a secondary extent a measure of their elastic properties. There is a large hydrostatic component of stress around the indentation, and since this plays no part in plastic flow, the indentation pressure is appreciably higher than the uniaxial flow stress of the materials. For many materials, the indentation pressure is about three times as large, but if the material shows appreciable elasticity, the yielding of the elastic hinterland imposes less constraint on plastic flow and the factor of proportionality may be considered less than 3. Indentation hardness depends on the time of loading and on the temperature and other operating environmental conditions. In the indentation tests, a spherical, conical, or pyramidal indenter is forced into the surface of the material, which forms a permanent (plastic) indentation in the surface of the material to be examined. The hardness number (GPa or kg/mm²), equivalent to the average pressure under the indenter, is calculated as the applied normal load divided by either the curved (surface) area (Brinell, Rockwell, and Vickers hardness numbers) or the projected area (Knoop and Berkovich hardness numbers) of the contact between the indenter and the material being tested under load (Lysaght 1949; Berkovich 1951; Tabor 1951, 1970; Mott 1957; O'Neill 1967; Westbrook and Conrad 1973; Blau and Lawn 1986; Bhushan and Gupta 1997; VanLandingham et al. 2001; Fisher-Cripps 2002).

In a conventional indentation hardness test, the contact area is determined by measuring the indentation size using a microscope after the sample is unloaded. For metals, there is a small change in the size of the indentation on unloading so that the conventional hardness test is essentially a test of hardness under load, although it is subject to some error due to varying elastic contraction of the indentation (Stilwell and Tabor 1961). The first depth-sensing indentation hardness test was developed in the early 1980s. In this test, the contact area is determined by measuring the indentation depth during the

loading/unloading cycle (Pethica et al. 1983; Nastasi et al. 1993; Bhushan et al. 1996a; Bhushan and Gupta 1997; Bhushan 1999a; Bhushan and Li 2003). Hardness data can be obtained without imaging the indentations with high reproducibility. This is useful particularly for small indents required for hardness measurements at very shallow depths. Depth measurements have, however, a major weakness arising from “piling-up” and “sinking-in” of material around the indentation. The measured indentation depth needs to be corrected for the depression (or the hump) of the sample around the indentation before it can be used for calculation of the hardness (Doerner and Nix 1986; Doerner et al. 1986; Nix 1989; Oliver and Pharr 1992). Elastic modulus is the slope of the stress-strain curve in the elastic regime. It can be obtained from the slope of the unloading curve (Oliver and Pharr 1992).

In addition to measurements of hardness and elastic modulus, static indentation tests have been used for measurements of a wide variety of material properties such as elastic-plastic deformation behavior (Pethica et al. 1983; Doerner and Nix 1986; Fabes et al. 1992; Oliver and Pharr 1992; Bhushan et al. 1996a; Bhushan 1999a; Bhushan and Li 2003), flow stress (Tabor 1951), scratch resistance and film-substrate adhesion (Heavens 1950; Tabor 1951; Benjamin and Weaver 1960; Campbell 1970; Ahn et al. 1978; Bhushan 1987; Wu 1991; Bhushan et al. 1995, 1996a; Bhushan and Gupta 1995; Gupta and Bhushan 1995a, b; Blau et al. 1997), residual stresses (Swain et al. 1977; Marshall and Lawn 1979; LaFontaine et al. 1991), creep (Westbrook 1957; Mulhearn and Tabor 1960; Atkins et al. 1966; Walker 1973; Chu and Li 1977; Li et al. 1991; Raman and Berriche 1992; Bhushan et al. 1996a), stress relaxation (Chu and Li 1980; Mayo and Nix 1988; LaFontaine et al. 1990; Wu 1991; Nastasi et al. 1993; Bhushan et al. 1996a), fracture toughness and brittleness (Lawn et al. 1980; Chantikul et al. 1981; Lawn 1993; Bhushan et al. 1996a, b; Li et al. 1997; Hainsworth et al. 1998; Li and Bhushan 1998a), and fatigue (Li and Chu 1979; Wu et al. 1991; Li and Bhushan 2002b, c).

Fig. 5.1 Extended load range of static indentation hardness testing



The extended load range of static indentation testing is shown schematically in Fig. 5.1. We note that only the lower micro- and nanohardness load range can be employed successfully for measurements of extremely thin (submicron-thick) films. The intrinsic hardness of surface layers or thin films becomes meaningful only if the influence of the substrate material can be eliminated. The minimum load for most commercial microindentation testers available is about 10 mN. Loads on the order of 1 μ N or less to 1 mN are desirable if the indentation depths are required to be a few nanometers. In this case, the indentation size sometimes reaches the resolution limit of a light microscope, and it is impossible to find such a small imprint if the measurement is made with a microscope after the indentation load has been removed. Hence, either the indentation apparatuses are placed *in situ* a scanning electron microscope (SEM) or *in situ* indentation depth measurements are made. In addition, the latter measurements would offer the advantages of observing the penetration process itself. In viscoelastic/viscoplastic materials, since indentation size changes with time, *in situ* measurements of the indentation size are particularly useful. In addition, they can provide more complete creep and relaxation data for the materials.

In this section, we will review two commercial depth-sensing nanoindentation test apparatuses, followed by data analysis and various measurement options.

5.2.1 Nanoindentation Apparatuses

Earlier research by Alekhin et al. (1972), Ternovskii et al. (1973), and Bulychev et al. (1975, 1979) led to the development of depth-sensing apparatuses (Pethica et al. 1983). In the depth-sensing indentation apparatus, the load-indentation depth is monitored continuously during the loading and unloading processes (Newey et al. 1982; Pethica et al. 1983; Wierenga and Franken 1984; Bhushan et al. 1985, 1988; Hannula et al. 1986; Wierenga et al. 1986; Tsukamoto et al. 1987; Williams et al. 1988; Yanagisawa and Motamura 1987; Wu et al. 1988; Loubet et al. 1993; Bhushan et al. 1996a; Randall et al. 1996). For detailed reviews, see Bhushan (1999a) and Bhushan and Li (2003).

In commercially available nanoindenter, an indenter is pushed into the sample until bulk plastic deformation occurs, and then unloaded. Indentation load is applied by various methods including using an electromagnet or electrostatic actuation. The indentation displacement is measured *in situ* using a capacitive displacement sensor in almost all cases. The apparatus continuously monitors the load and the position of the indenter relative to the surface of the specimen (depth of an indent) during the indentation process. The area of the indent is then calculated from a knowledge of the geometry of the tip of the diamond indenter.

5.2.1.1 General Description and Principle of Operation

Commercially available nanoindenters are made by various companies. The major manufacturers include Keysight Technologies, Santa Rosa, CA, U.S.A. (first major manufacturer established in the mid-1980s as Nano Instruments, Inc., Knoxville, TN, later acquired by MTS System Corp., followed by Agilent Technologies), Micro Materials, Wrexham, U.K., and CSM Instruments Inc., Neuchâtel, Switzerland. These can be used with a minimum indentation depth of about 20 nm. Hysitron Inc., Minneapolis, MN, launched in the late 1990s, used an AFM platform and electrostatic actuation for loading/unloading. The minimum depth achieved was about 10 nm, and it remains a major research tool (Bhushan et al. 1996a). One of the benefits of this apparatus is that the nanoindenter head can be used to replace the AFM head in a commercial AFM without requiring new scanning and imaging capabilities.

Next we briefly describe the Keysight and Hysitron indenters.

Keysight Nanoindenter

Initial development of this apparatus has been described by Pethica et al. (1983) and Oliver et al. (1986). Mechanical properties measurements can be made at a minimum penetration depth of about 20 nm (or a plastic depth of about 15 nm) (Oliver et al. 1986). Specifications and commonly-used operating parameters for a commercial nanoindenter are given in Table 5.1 (Anonymous 2014).

The nanoindenter consists of three major components: the indenter head, an X-Y-Z motorized precision table for positioning and transporting the sample, and an AFM and/or optical microscope, Fig. 5.2a. The load on the indenter is generated using a voice-coil in a permanent magnet assembly attached to the top of the indenter (loading) column. The generated load is simply the vector product of the current through the coil and the magnetic field strength of the permanent magnet. This type of load

application allows for very fast, closed-loop feedback control over the displacement as it completely separates the load application system and the displacement measuring system. Various interchangeable indenter heads are available with load ranges typically from 0 to 500 mN, with a high-load option up to 10 N. The displacement of the indenter is measured using a three-plate capacitive displacement sensor. All three plates are circular disks approximately 1.5 mm thick. The two outer plates have a diameter of approximately 50 mm, and the inner, moving plate is half that size. The indenter column is attached to the moving plate. Two outer plates are maintained at equal and opposite drive voltages (2 V at 12.5 kHz). The output voltage of the center pick-up plate is related to the position of that plate in the capacitive gap uniquely. The pick-up voltage is measured using an analog to digital converter (ADC) running at a typical rate of 4 kHz. The displacement system is calibrated using a laser-interferometer. This plate-and-indenter assembly is supported by two leaf springs cut to have very low stiffness. The motion is damped by airflow around the central plate of the capacitor, which is attached to the loading column. At the bottom of the indenter rod, an indenter, typically a three-sided pyramidal diamond tip (Berkovich indenter, to be discussed later) is attached.

The indenter head assembly is attached to the "U" beam rigidly, below which the X-Y-Z table rides, Fig. 5.2a. An AFM and/or optical microscope is also attached to the beam for in situ imaging capabilities. The position of an indent on a specimen is selected using either an optical microscope or the AFM. The specimens are held on an X-Y-Z table whose position relative to the microscope or the indenter is controlled. The spatial resolution of the position of the table in the X-Y plane is on the order of 100 nm. The three components just described are enclosed in a cabinet to ensure the thermal stability of the samples. The entire apparatus is placed on a vibration-isolation table.

A continuous stiffness measurement (CSM) option (to be described later) can be used to make continuous measurements of the stiffness

Table 5.1 Typical specifications and commonly-used operating parameters in Keysight Nano Indenter G200 (Anonymous 2014)

<i>Standard XP indentation head</i>	
Displacement resolution	<0.01 nm
Total indenter travel	1.5 mm
Maximum indentation depth	>500 μ m
Load application	Coil/magnet assembly
Displacement measurement	Capacitance gauge
Loading capability	500 mN
Maximum load (standard)	30 mN
Maximum load with DCM II option	10 N
Maximum load with High Load option	50 nN
Load resolution	<1.0 μ N
Contact Force	$\sim 5 \times 10^6$ N/m
Load frame stiffness	
Indentation placement	100 mm \times 100 mm
Useable surface area	Automated remote with mouse
Position control	1 μ m
Position accuracy	
Microscope	25 \times
Video screen	10 \times and 40 \times
Objective	
<i>DCM II indentation head option</i>	
Displacement resolution	<1 nm
Range of indenter travel	70 μ m
Loading column mass	<150 mg
Load application	Coil/magnet assembly
Displacement measurement	Capacitance gauge
Typical leaf spring stiffness	\sim 100 N/m
Typical damping coefficient	0.02 Ns/m
Typical resonant frequency	120 Hz
Lateral stiffness	80.000 N/m
Loading capability	30 mN (13 g)
Maximum load	3 nN (0.3 μ g)
Load resolution	
<i>Express test option</i>	
Time per indentation	Standard <5.0 s
<i>LFM option</i>	
Maximum lateral force	>250 mN
Lateral resolution	<2 μ N
Maximum scratch distance	>100 mm
Scratch speed	100 nm/s up to 2 mm/s
<i>High load option</i>	
Maximum force	10 N
Load resolution	\leq 1 mN

(continued)

Table 5.1 (continued)

Maximum indentation depth	$\geq 500 \mu\text{m}$
Displacement resolution	0.01 nm
Frame stiffness	$\geq 5 \times 10^6 \text{ N/m}$
<i>NanoVision option</i>	
X-Y scan range	$100 \mu\text{m} \times 100 \mu\text{m}$
Z scan range	Indentation head dependent
Positioning accuracy	$\leq 20 \text{ nm}$
Resonant frequency	>120 Hz

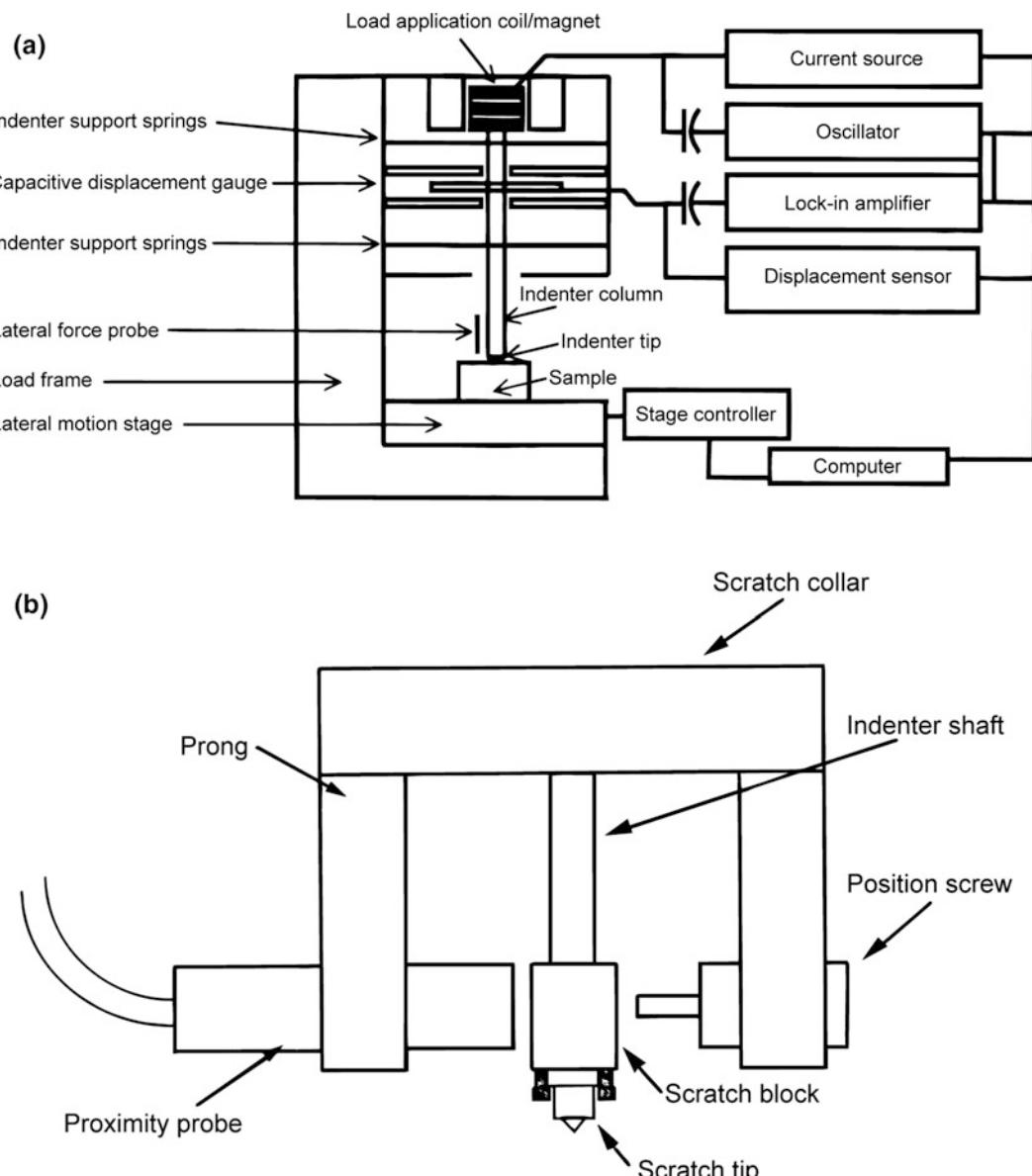


Fig. 5.2 Schematics of an Keysight nanoindenter design **a** showing the major components: the indenter head, a X-Y-Z motorized precision table, and tangential force option, and **b** detailed schematic of the tangential force option hardware (not to scale) (adapted from Keysight brochures)

of a sample, which allows the elastic modulus to be calculated as a function of time (or indentation depth) (Pethica and Oliver 1989; Li and Bhushan 2002a). Because of the relatively small time constant of the measurements, the device is particularly useful in studies of time-dependent properties of materials.

Hysitron Nanoindenter

For nanometer-scale indentation depths, in situ capabilities of imaging of the indents at extremely high magnifications is desirable. AFMs are ideal for in situ imaging of surfaces before and after indentation. They have been used for indentation studies with nanometer-scale depths (Bhushan and Koinkar 1994; Bhushan et al. 1995). However, load-displacement curves cannot be obtained with high accuracy for the following reasons. First, to obtain the indentation depth, the sample displacement needs to be subtracted from the tip displacement, requiring subtraction of two large numbers. Second, piezoelectric transducers (PZTs) generally are used for sample motion, which exhibit nonlinearity, hysteresis, and creep. Nonlinearity can be corrected; however, hysteresis and creep can result in displacements on the order of the indentation depths of interest. Finally, a large, lateral deflection of the cantilever beam may be required to apply high loads, which changes the indentation location during loading.

An indentation system with a three-plate transducer with electrostatic actuation for loading/unloading applications and capacitive sensor for displacement measurements was developed by Hysitron Inc. The system can apply loads ranging from 1 μN or less to 10 mN directly, and make load-displacement measurements with subnanometer indentation depth resolution (Bhushan et al. 1996a). When used in conjunction with an AFM, in situ imaging also can be obtained. A schematic of the indentation system using an AFM (such as from Bruker Instruments) as a platform is shown in Fig. 5.3a (Bhushan et al. 1996a). The indentation system consists of a three-plate transducer with electrostatic actuation hardware used for direct application of normal load, and a capacitive sensor

used for measurements of vertical displacement (Fig. 5.3b). The AFM head is replaced with this transducer head, and the sample is mounted on the PZT scanner. The transducer has a three-plate (Be–Cu) capacitive structure, which provides high sensitivity, a large dynamic range, and a linear output signal with respect to load or displacement. The tip is connected to the center plate, which is spring mounted to the housing. Four springs are mounted on the top and another four are mounted at the bottom with a total stiffness of about 100 Nm^{-1} . A titanium shank brazed to the diamond indenter tip is glued to a glass-reinforced plastic (Ryton) holder, which, in turn, is screwed into the center plate (Fig. 5.3c).

Vertical displacement of the tip (indentation depth) is measured by measuring the displacement of the center electrode relative to the two outer electrodes using the capacitance technique. During measurements, the sample remains stationary. The load is applied by an electrostatic force generated between the center (pick-up) electrode and drive plate (upper or lower) when a voltage is applied to the drive plate. The applied load is proportional to the square of the voltage. The load resolution is 100 nN or better, and the displacement resolution is 0.1 nm. A load range of 1 μN –10 mN typically is used. Loading rates can be varied, changing the load–unload period typically from 180 to 950 s. The AFM functions as a platform for the indenter system and also provides in situ imaging before and after indentation with a lateral resolution of about 10 nm and a vertical resolution of about 0.2 nm. Load-displacement data during loading and unloading are obtained as contact depth as low as 5–10 nm. A continuous stiffness option is also available. The instrument also can be used for scratching, wear, and adhesion studies (Palacio and Bhushan 2010; Kumar and Bhushan 2015; Cho and Bhushan 2016).

This instrument also can be installed inside an SEM or TEM to allow in situ imaging of the indentation process.

5.2.1.2 The Indenters

The indenter should have a high elastic modulus, virtually no plastic deformation, low friction, smooth surface, and a well-defined geometry that

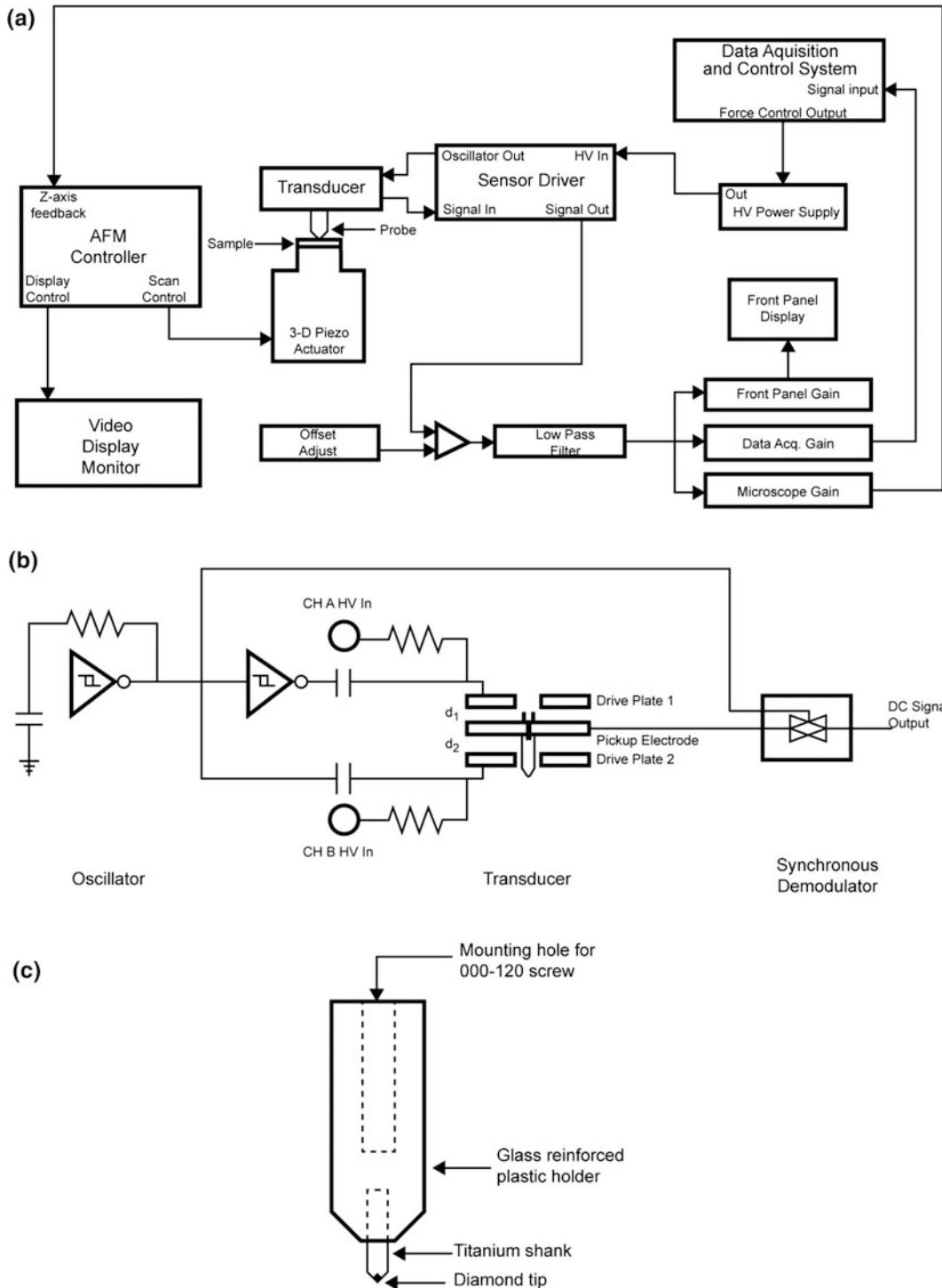


Fig. 5.3 Schematic of the Hysitron nanoindenter system **a** showing the major components, **b** the three-plate transducer with electrostatic actuation hardware and capacitance sensor, and **c** tip-holder mount assembly (adapted from Bhushan et al. 1996a)

is capable of making a well-defined indentation impression. The first four requirements are satisfied by choosing natural diamond material for the tip. A well-defined perfect tip shape is difficult to achieve. Berkovich is a three-sided pyramid and provides a sharply pointed tip compared with the Vickers or Knoop indenters, which are four-sided pyramids and have a slight offset (0.5–1 μm) (Tabor 1970; Bhushan 1999a; Bhushan and Li 2003). Because any three nonparallel planes intersect at a single point, it is relatively easy to grind a sharp tip on an indenter if Berkovich geometry is used. However, an indenter with a sharp tip suffers from a finite, but an exceptionally difficult-to-measure, tip bluntness. In addition, pointed indenters produce a virtually-constant plastic strain impression, and there is the additional problem of assessing the elastic modulus from the continuously-varying unloading slope. Spherical indentation overcomes many of the problems associated with pointed indenters. With a spherical indenter, one is able to follow the transition from elastic to plastic behavior and thereby define the yield stress (Bell et al. 1992). However, a sharper tip is desirable, especially for extremely thin films requiring shallow indentation. Therefore, a Berkovich indenter is used most commonly for measurements of nanomechanical properties. Experimental procedures have been developed to correct for the tip shape, to be described later.

The Berkovich indenter, directly brazed to a 304 stainless steel holder, is a three-sided (triangular-based) pyramidal diamond, with a nominal angle of 65.3° between the (side) face and the normal to the base at apex, an angle of 76.9° between edge and normal, and with a radius of the tip less than 0.1 μm (Fig. 5.4) (Berkovich 1951). The typical indenter is shaped to be used for a minimum indentation (penetration) depth of 10–20 nm. The indentations appear as equilateral triangles (Fig. 5.4c) and the height of triangular indent ℓ is related to the depth h as

$$\frac{h}{\ell} = \left(\frac{1}{2}\right) \cot 76.9 = \frac{1}{8.59} \quad (5.1a)$$

The relationship $h(\ell)$ is dependent on the shape of the indenter. The height of the triangular indent, ℓ , is related to the length of one side of the triangle, a , as

$$\ell = 0.866 a \quad (5.1b)$$

and

$$\frac{h}{a} = \frac{1}{7.407} \quad (5.1c)$$

The projected contact area (A) for the assumed geometry is given as

$$A = 0.433a^2 = 23.76h^2 \quad (5.2)$$

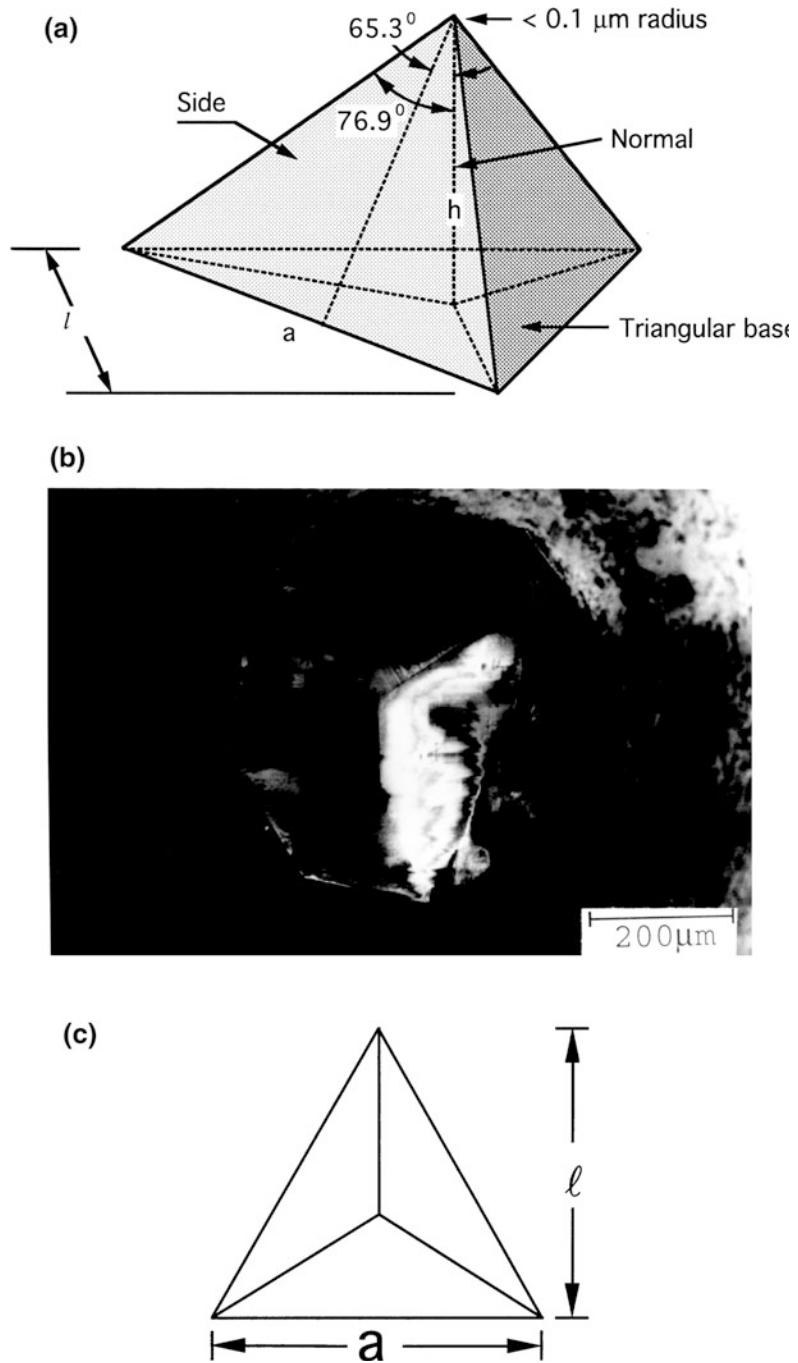
Another three-sided pyramidal indenter, the cube corner indenter, can displace more than three times the volume of the Berkovich indenter at the same load, thereby producing much higher stresses and strains in the vicinity of the contact and reducing the cracking threshold. This makes this indenter ideal for the estimation of indentation fracture toughness at relatively small scales indentation (Li et al. 1997). The spherical indenter initiates elastic contact and then causes elastic-plastic contact at higher loads. This indenter is well-suited for examinations of yielding and work hardening. However, it is very difficult to obtain a precise sphere with a diameter of less than 100 μm made of diamond. This fact limits its application in nanoindentation testing.

The exact shape of the indenter tip needs to be measured for determination of hardness and elastic modulus. Since the indenter is quite blunt, direct imaging of indentations of small size in the SEM is difficult. Determination of tip area function will be discussed later.

5.2.1.3 Indentation Procedure

The hardness and elastic modulus are measured in compliance with ISO 14577. An indentation test involves moving the indenter to the surface of the material and measuring the forces and displacements associated during indentation. The surface is located for each indentation by

Fig. 5.4 **a** Schematic and **b** photograph of the shape of a Berkovich indenter, and **c** indent impression



lowering the indenter at a constant loading rate against the suspending springs and detecting a change in velocity on contact with the surface. In the testing mode, the load is incremented in order to maintain a constant loading rate or constant

displacement rate. The load and indentation depths are measured during indentation both in the loading and unloading cycles. The force contribution of the suspending springs and the displacements associated with the measured

compliance of the instrument are removed (Bhushan 1999a; Bhushan and Li 2003).

Typically, a Berkovich tip with a radius of about 50 nm is used. Multiple loading and unloading steps are performed to examine the reversibility of the deformation, ensuring that the unloading data used for analysis purposes are mostly elastic. A typical indentation experiment consists of a combination of several segments, e.g., approach, load, hold, and unload. Either constant loading or constant displacement experiments can be performed (Oliver and Pharr 1992; Bhushan 1999a). A typical constant loading indentation experiment consists of eight steps: approaching the surface at 10 nm/s; loading to peak load at a constant loading rate (10% of peak load/s); unloading 90% of peak load at a constant unloading rate (10% of peak load/s); reloading to peak load; holding the indenter at peak load for 10 s; unloading 90% of peak load, holding the indenter after 90% unloading; finally unloading completely. The first hold step is included to incorporate the corrections due to thermal drift. Figure 5.5 shows a representative load-displacement curve of an indentation made at 15 mN peak indentation load for Si(100) during a loading/unloading sequence in a constant loading experiment (Bhushan and Li 2003).

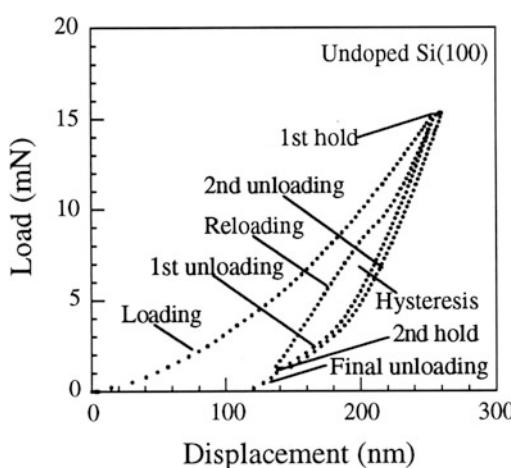


Fig. 5.5 Representative load-displacement curve during a loading/unloading sequence in a constant loading experiment (adapted from Bhushan and Li 2003)

5.2.1.4 Acoustic Emission Measurements During Indentation

Acoustic emission (AE) measurement is a very sensitive technique to monitor cracking of the surfaces and subsurfaces (Bhushan 1999a; Bhushan and Li 2003). The nucleation and growth of cracks result in a sudden release of energy within a solid, then some of the energy is dissipated in the form of elastic waves. These waves are generated by sudden changes in stress and displacement that accompany the deformation. If the release of energy is sufficiently large and rapid, then elastic waves in the ultrasonic frequency regime (AE) will be generated, and these can be detected using PZTs via expansion and compression of the PZT crystals (Scruby 1987; Bhushan 1996).

An AE sensor is used to detect cracking during indentation tests in the nanoindenter. It can be mounted underneath the sample. The energy dissipated during crack growth can be estimated by the rise time of the AE signal. The transducer converts the AE signal into voltage and is used for continuous display of the AE signal. Any correlation between the AE signal and local changes in the load-displacement curves can be observed (Bhushan 1999a; Bhushan and Li 2003).

5.2.1.5 Nanoscratch and Lateral Force Measurements

There are several commercially available micro- and nanoscratch testers, such as the Taber shear/scratch tester (Taber Industries, North Tonawanda, New York) for thick films; Revetest Xpress scratch tester (CSM Instruments, Neuchatel, Switzerland) for thin films (Perry 1981, 1983; Sekler et al. 1988), and nanoindenter with scratch and lateral force measurement (LFM) options for ultrathin films (Wu 1991; Bhushan et al. 1995; Bhushan and Gupta 1995; Gupta and Bhushan 1995a, b; Bhushan 1999a; Bhushan and Li 2003). AFMs with a sharp diamond tip also are used for nanoscratch, nano-wear, and lateral force measurements (Bhushan 1999b; Sundararajan and Bhushan 2001).

Here, the nanoscratch and LFM option in a commercial Keysight nanoindenter is described.

Scratches of various lengths at programmable loads can be made (Bhushan 1999a; Bhushan and Li 2003). Lateral (friction) forces and/or AE signal can also be measured simultaneously. The lateral force option includes a set of proximity (capacitance) probes for measurement of lateral displacement or force in the two lateral directions along X and Y, and a special “scratch collar,” which mounts around the indenter shaft with hardness indenter, Fig. 5.2b. A scratch block is mounted on the end of the indenter shaft, in line with the proximity probes and the positioning screws. The scratch tip is attached to the scratch block with two Allen head screws. The scratch tip can be a Berkovich indenter or a conventional conical diamond tip with a tip radius of about 1–5 μm and an included angle of 60°–90° (typically 1 μm of tip radius with 60° of included angle; Bhushan et al. 1995; Palacio and Bhushan 2010; Kumar and Bhushan 2015; Cho and Bhushan 2016). A larger included angle of 90° may be desirable for a more durable tip. The tip radius should not be very small as it will blunt readily.

During scratching, a load is applied up to a specified indentation load or up to a specified indentation depth, and the lateral motion of the sample is measured. During scratching, load and indentation depth are monitored. Scratches can be made either with a constant load or at ramp-up load. Measurement of lateral force allows the calculations of the coefficient of friction during scratching. The resolution of the proximity probe provides resolution of lateral force on the order of 2 μN ; therefore, a minimum load of about 20 μN can be measured. Consequently, a minimum normal load of about 0.2 mN typically is used for a sample with a coefficient of friction of about 0.1. Microscopy of the scratch produced at ramp-up load allows measurement of the critical load required to break the film (if any), scratch width, and general observations of scratch morphology. Typically, draw acceleration ($\mu\text{m}/\text{s}^2$) and draw velocity ($\mu\text{m}/\text{s}$) are 10 $\mu\text{m}/\text{s}^2$ and 5 $\mu\text{m}/\text{s}$, respectively. The data not only provide measures of scratch resistance or wear resistance, but also provide insight into ductile/brittle fracture modes (Bhushan and Gupta 1995).

5.2.2 Analysis of Indentation Data

Load, W , as a function of displacement (or indentation depth or penetration depth), h , is monitored continuously and recorded during the loading and unloading process, and is referred to as an indentation curve or load-displacement curve (Bhushan 1999a; Bhushan and Li 2003). Figure 5.6 shows stress-strain curves, typical indentation curves, the deformed surfaces after tip removal, and residual impressions of indentation for ideal elastic, rigid-perfectly plastic, elastic-perfectly plastic, and real elastic-plastic solids (Bhushan 1999a). For an elastic solid, the sample deforms elastically according to elastic modulus, and the deformation is recovered during unloading. As a result, there is no impression of the indentation after unloading. For a rigid-perfectly plastic solid, no deformation occurs until yield stress is reached, when plastic flow takes place. There is no recovery during unloading and the impression remains unchanged. In the case of elastic-plastic solid, it deforms elastically according to elastic modulus and then it deforms plastically. The elastic deformation is recovered during unloading. In the case of an elastic-perfectly plastic solid, there is no work hardening.

Most engineering surfaces follow real elastic-plastic deformation behavior with work hardening (Johnson 1985; Bhushan 2013a, b). Figure 5.7a schematically shows the deformation pattern of a real elastic-plastic sample during and after indentation (Oliver and Pharr 1992; Hainsworth et al. 1996; Oliver 2001). The contact depth (h_c) is defined as the depth of indenter in contact with the sample under load. The depth measured during the indentation (h) includes the depression of the sample around the indentation, in addition to the contact depth. The depression of the sample around the indentation ($h_s = h - h_c$) is caused by elastic displacements and must be subtracted from the data to obtain the actual depth of indentation or actual hardness. At peak load, the load and displacement are W_{max} and h_{max} , respectively, and the radius of the contact circle is a . Upon unloading, the elastic displacements in the contact region are recovered

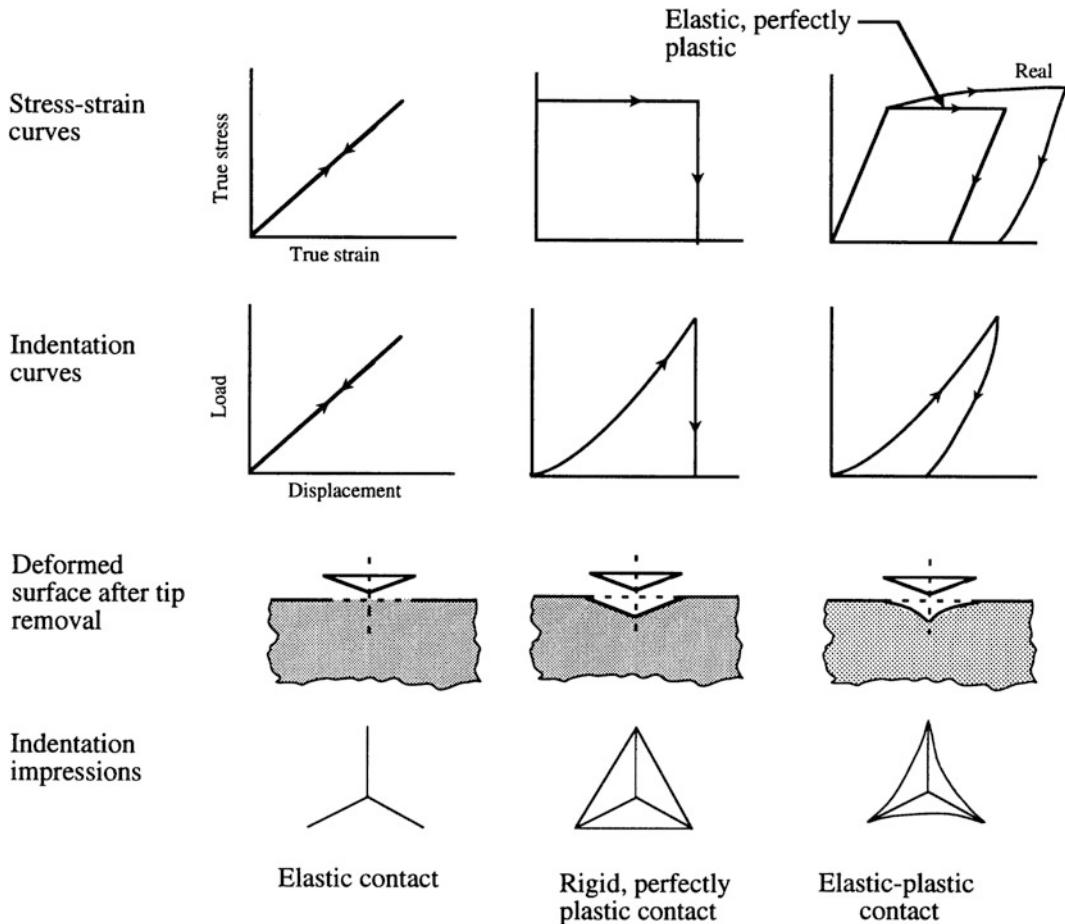


Fig. 5.6 Schematics of stress-strain curves, typical indentation curves, deformed surfaces after tip removal, and residual impressions of indentation, for ideal elastic, rigid-perfectly plastic, elastic-perfectly plastic (ideal) and real elastic-plastic solids

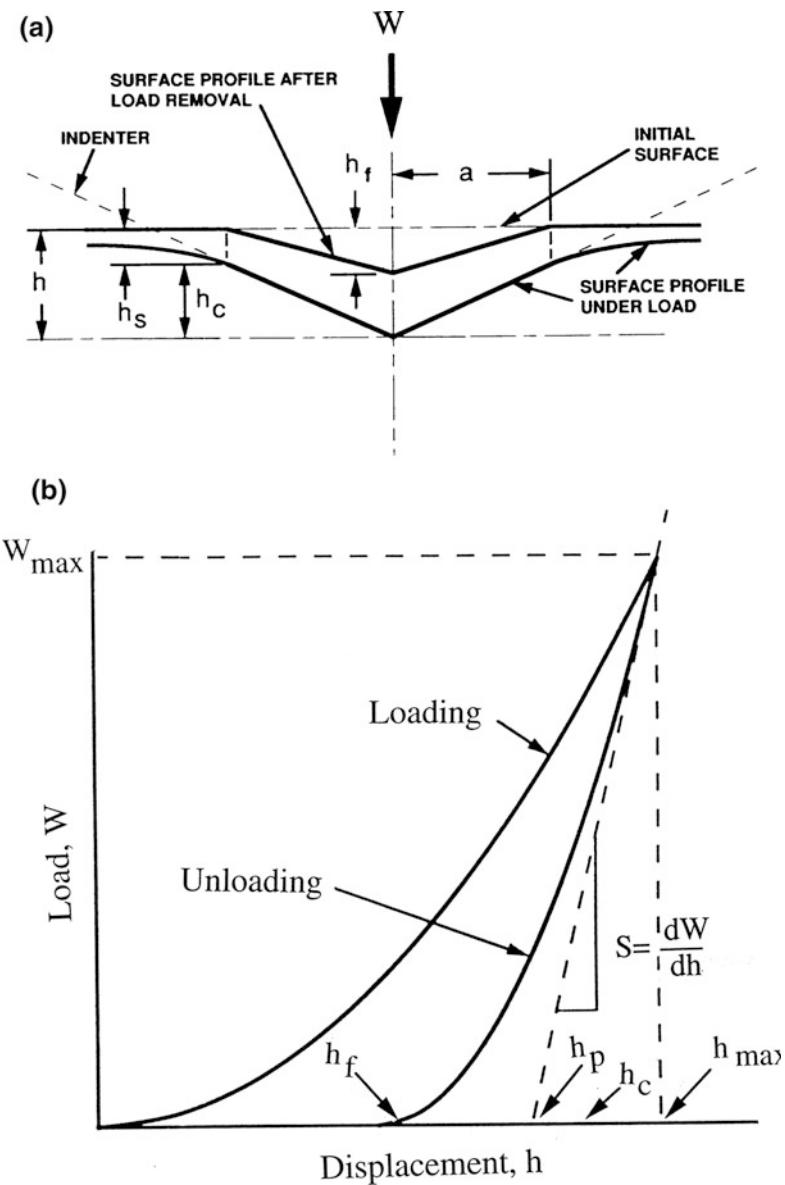
and when the indenter is fully withdrawn, the final depth of the residual hardness impression is h_f . A schematic of a load-displacement curve is shown in Fig. 5.7b.

In order to predict the deflection of the surface at the contact perimeter for a conical indenter and a paraboloid of revolution, Oliver and Pharr (1992) developed an expression, based on the work of Sneddon (1965), for h_c at the maximum load (required for hardness calculation) from h_{max} ,

$$h_c = h_{max} - \varepsilon W_{max}/S_{max} \quad (5.3a)$$

where $\varepsilon = 0.72$ for the conical indenter, $\varepsilon = 0.75$ for the paraboloid of revolution, and $\varepsilon = 1$ for the flat punch; and S_{max} is the stiffness (=1/compliance) equal to the slope of unloading curve (dW/dh) at the maximum load. (Hay et al. (1999) have proposed corrections to Sneddon's equations.) Oliver and Pharr (1992) assumed that behavior of the Berkovich indenter is similar to that of the conical indenter, since cross-sectional areas of both types of indenters varies as the square of the contact depth and their geometries are singular at the tip. Therefore, for a Berkovich indenter, $\varepsilon \sim 0.72$. Thus, h_c is slightly larger than plastic indentation depth (h_p), which is given by

Fig. 5.7 **a** Schematic representation of the indenting process illustrating the depression of the sample around the indentation and the decrease in indentation depth upon unloading (adapted from Oliver and Pharr 1992), and **b** schematic of load-displacement curve



$$h_p = h_{max} - W_{max}/S_{max} \quad (5.3b)$$

and Nix 1986; Bhushan 1999a; Swadener et al. 2002)

Based on the finite element analysis of the indentation process, Laursen and Simo (1992) showed that h_c cannot be assumed equal to h_p for indenters that do not have a flat-punch geometry.

Projected contact-area-to-depth relationship for a Vickers indenter with ideal pyramidal geometry (ideally sharp tip) is given as (Doerner

$$A = 24.5h_c^2 \quad (5.4a)$$

Since the area to depth relationship is equivalent for both typical Berkovitch and Vickers pyramids, Eq. 5.4a holds for the Berkovitch indenter as well. Although a slightly different

expression for $A(h)$ is presented in Eq. 5.2 for the assumed Berkovich indenter geometry, this relationship is used most commonly in the analysis of the indentation hardness data.

As shown in Fig. 5.8a, the indenter tip generally is rounded so that ideal geometry is not maintained near the tip. To study the effect of tip radius on the elastic-plastic deformation (load vs. displacement curve), Shih et al. (1991) modeled the blunt-tip geometry by a spherical tip of various radii. They derived a geometric relationship (assuming no elastic recovery) between the projected contact area of the indenter to the actual contact depth. Figure 5.8b shows the measured contact area vs. indentation depth data by Pethica et al. (1983) for nickel and by Doerner and Nix (1986) for annealed α -brass. From this figure, it seems that a tip radius of 1 μm fits the data best. If there is elastic recovery, the experimental data are smaller than what they should be, and then the tip radius would be even larger than 1 μm . Shih et al. (1991) used the finite element method to simulate an indentation test. They showed that load-indentation depth data obtained using nanoindenter for nickel by Pethica et al. (1983) can be fitted with a simulated profile for a tip radius of about 1 μm , Fig. 5.8c.

Figure 5.8 shows that the actual indentation depth, h_c , produces a larger contact area than would be expected for an indenter with an ideal shape. For the real indenter used in the measurements, the nominal shape is characterized by an area function $F(h_c)$, which relates projected contact area of the indenter to the contact depth (Eq. 5.4a),

$$A^{1/2} = F(h_c) \quad (5.4b)$$

The functional form must be established experimentally prior to the analysis (to be described later).

5.2.2.1 Hardness

Berkovich hardness HB (or H_B) is defined as the load divided by the projected contact area of the indentation. It is the mean pressure that a material will support under load. From the indentation curve, we can obtain hardness at the maximum load as

$$HB = W_{max}/A \quad (5.5)$$

where W_{max} is the maximum indentation load, and A is the projected contact area at the peak load. The contact area at the peak load is determined by the geometry of the indenter and the corresponding contact depth h_c using Eqs. 5.3a and 5.4b. A plot of hardness as a function of indentation depth for polished single-crystal Si (111), with and without tip-shape calibration, is shown in Fig. 5.9. We note that, for this example, tip-shape calibration is necessary. For these measurements, the hardness is independent of corrected depth.

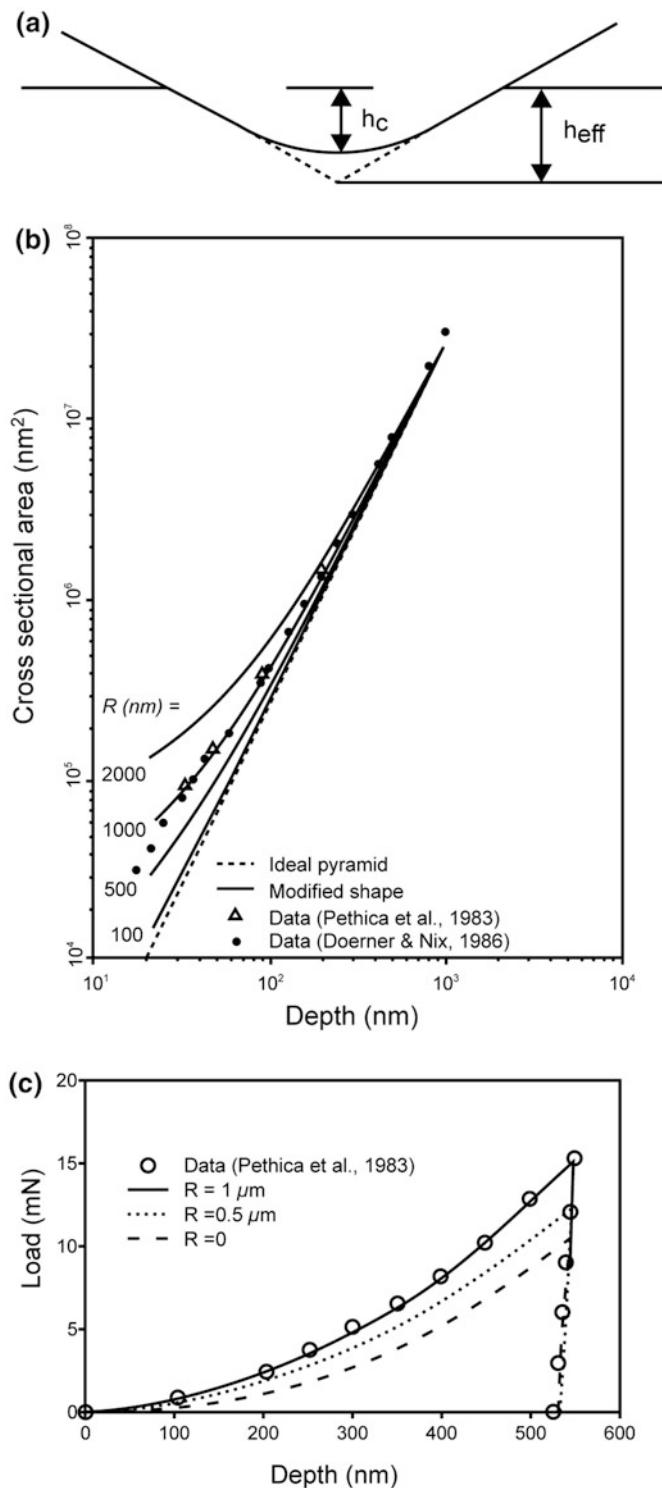
Note that hardness measured using this definition may be different from that obtained from the more conventional definition in which the area is determined by direct measurement of the size of the residual hardness impression. The reason for the difference is that, in some materials, a small portion of the contact area under load is not plastically deformed, and, as a result, the contact area measured by observation of the residual hardness impression may be less than that at peak load. However, for most materials, measurements using two techniques give similar results.

5.2.2.2 Modulus of Elasticity

The initial unloading is an elastic event, though during loading a sample undergoes elastic-plastic deformation. Therefore, elastic modulus of the specimen can be inferred from the initial slope of the unloading curve (dW/dh) called stiffness (1/compliance) (at the maximum load) (Fig. 5.7 b). It should be noted that the contact stiffness is measured only at the maximum load, and no restrictions are placed on the unloading data being linear during any portion of the unloading (Bhushan 1999a; Bhushan and Li 2003).

An approximate elastic solution is obtained by analyzing a flat punch whose area in contact with the specimen is equal to the projected area of the actual punch, assuming that the area in contact remains constant during initial unloading. Based on the analysis of indentation of an elastic half space by a flat cylindrical punch by Sneddon

Fig. 5.8 **a** Schematic of an indenter tip with a nonideal shape. The contact depth and the effective depth are also shown, **b** predicted projected contact area as a function of indentation depth curves for various tip radii and measured data for nickel (adapted from Pethica et al. 1983), and for annealed α -brass (adapted from Doerner and Nix 1986), and **c** predicted load as a function of indentation depth curves for various tip radii and measured data for nickel (adapted from Shih et al. 1991)



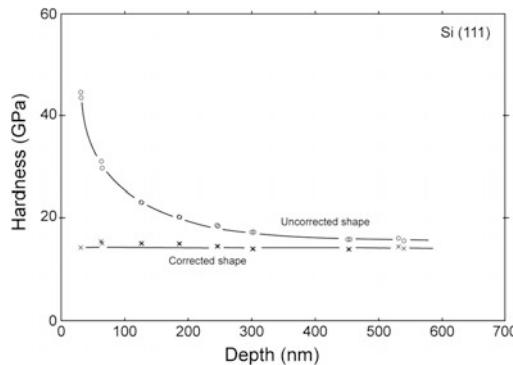


Fig. 5.9 Hardness as a function of indentation depth for polished single-crystal Si(111) calculated from the area function with and without tip shape calibration (adapted from Doerner and Nix 1986)

(1965), Loubet et al. (1984) calculated the elastic deformation of an isotropic elastic material with a flat-ended cylindrical punch. They obtained an approximate relationship for compliance (dh/dW) for the Vickers (square) indenter. King (1987) solved the problem of flat-ended cylindrical, quadrilateral (Vickers and Knoop), and triangular (Berkovitch) punches indenting an elastic half-space. He found that the compliance for the indenter is approximately *independent of the shape* (with a variation of at most 3%) if the projected area is fixed. Pharr et al. (1992) also verified that compliance of a paraboloid of revolution of a smooth function is the same as that of a spherical or a flat-ended cylindrical punch. The relationship for the compliance C (inverse of stiffness S) for an (Vickers, Knoop, and Berkovich) indenter is given as

$$C = \frac{1}{S} = \frac{dh}{dW} \sim \frac{1}{2E_r} \left(\frac{\pi}{A} \right)^{1/2} \quad (5.6)$$

where

$$\frac{1}{E_r} = \frac{1 - v_s^2}{E_s} + \frac{1 - v_i^2}{E_i}$$

dW/dh is the slope of the unloading curve at the maximum load (Fig. 5.7b), E_r , E_s , and E_i are the reduced modulus and elastic moduli of the specimen and the indenter, and v_s and v_i are the Poisson's ratios of the specimen and indenter.

C (or S) is the experimentally measured compliance (or stiffness) at the maximum load during unloading, and A is the projected contact area at the maximum load.

The contact depth h_c is related to the projected area of the indentation A for a real indenter by Eq. 5.4b. With a plot of the measured compliance (dh/dW) the reciprocal of the corrected indentation depth obtained from various indentation curves (one data point at maximum load for each curve) should yield a straight line with slope proportional to $1/E_r$ (Fig. 5.10) (Doerner and Nix 1986). E_s can then be calculated, provided Poisson's ratio with great precision is known to obtain a good value of the modulus. For a diamond indenter, $E_i = 1140$ GPa and $v_i = 0.07$ are taken (Bhushan 1999a; Bhushan and Li 2003). In addition, the y-intercept of the compliance vs. the reciprocal indentation depth plot should give any additional compliance that is independent of the contact area. The compliance of the loading column is generally removed from the load-displacement curve, whose measurement techniques will be described later.

To measure initial unloading stiffness (S), Doerner and Nix (1986) fitted a straight line to the

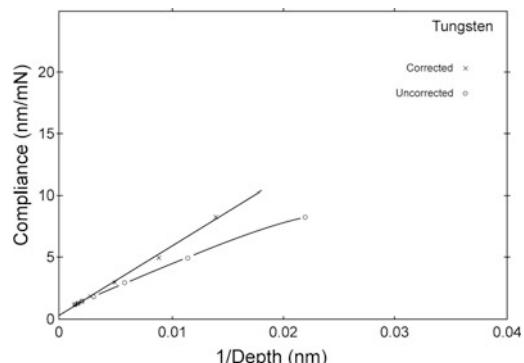


Fig. 5.10 Compliance as a function of the inverse of indentation depth for tungsten with and without tip shape calibration. A constant modulus with $1/\text{depth}$ would be indicated by the straight line. The slope of the corrected curve is 480 GPa, which compares reasonably well to the known modulus of tungsten (420 GPa). The small y-intercept of about 0.3 nm/mN is attributed to load frame compliance, not removed (adapted from Doerner and Nix 1986)

upper about one third portion of the unloading curve. The problem with this is that for nonlinear loading data, the measured stiffness depends on how much of the data is used in the fit. Oliver and Pharr (1992) proposed a new procedure. They found that the entire unloading data are well described by a simple power law relation

$$W = B(h - h_f)^m \quad (5.7)$$

where the constants B and m are determined by a least-square fit. The initial unloading slope is then found analytically, differentiating this expression and evaluating the derivative at the maximum load and maximum depth. As we have pointed out earlier, unloading data used for the calculations should be obtained after several loading/unloading cycles and with peak hold periods.

This analysis is based on an elastic solution, which only accounts for sink-in (the indented material around the indenter below its original surface). However, in the more realistic case of elastic-plastic contact, sink-in or pile-up (the indented material around the indenter above its original surface) can occur depending on the specific mechanical properties of the material. For pile-up situations, the just-described Oliver-Pharr method would underestimate the true contact area by as much as 50%. This in turn leads to overestimation of the hardness and elastic modulus. Based on some modeling, pile-up is significant only when $h_f/h_{max} > 0.7$ and the material does not appreciably work harden. Note that h_f/h_{max} equal to zero corresponds to fully elastic deformation and a value of 1 corresponds to rigid-plastic behavior. Compressive residual stresses result in pile-up, whereas tensile stresses result in sink-in. Although some correction procedures have been proposed (Pharr 1998; Tsui and Pharr 1999), the real contact area measurement requires imaging of indentation impressions.

5.2.2.3 Determination of Load Frame Compliance and Indenter Area Function

As stated earlier, measured displacements are the sum of the indentation depths in the specimen and the displacements of suspending springs and

the displacements associated with the measuring instruments, referred to as load frame compliance. Therefore, to determine accurately the specimen depth, load frame compliance must be known (Bhushan 1999a; Bhushan and Li 2003). This is especially important for large indentations made with high modulus for which the load frame displacement can be a significant fraction of the total displacement. The exact shape of the diamond indenter tip needs to be measured because hardness and elastic modulus depend on the contact areas derived from measured depths. The tip gets blunt (Fig. 5.8a) and its shape significantly affects the prediction of mechanical properties (Figs. 5.9 and 5.10).

Oliver and Pharr (1992) proposed a method for determining area functions. Their method is based only on one assumption: that elastic modulus is independent of indentation depth. They also proposed a method to determine load frame compliance. We first describe the methods for determining of load frame compliance followed by the method for area function.

They modeled the load frame and the specimen as two springs in series; thus,

$$C = C_s + C_f \quad (5.8)$$

where C , C_s , and C_f are the total measured compliance, specimen compliance and load frame compliance, respectively. From Eqs. 5.6 and 5.8, we get

$$C = C_f + \frac{1}{2E_r} \left(\frac{\pi}{A} \right)^{1/2} \quad (5.9)$$

From Eq. 5.9, we note that if the modulus of elasticity is constant, a plot of C as a function of $A^{-1/2}$ is linear and the vertical intercept gives C_f . It is obvious that the most accurate values of C_f are obtained when the specimen compliance is small; i.e., for large indentations.

To determine the area function and the load frame compliance, relatively large indentations in aluminum are made because of its low hardness (Bhushan 1999a; Bhushan and Li 2003). In addition, for the larger aluminum indentations (typically 700–4000 nm deep), the area function

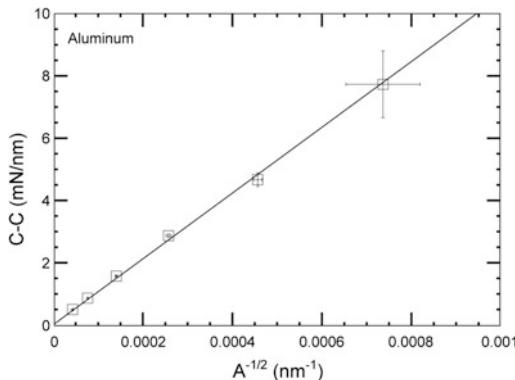


Fig. 5.11 Plot of $(C - C_f)$ as a function $A^{-1/2}$ for aluminum. The error bars are two standard deviations in length (adapted from Oliver and Pharr 1992)

for a perfect Berkovich indenter (Eq. 5.4a) can be used to provide a first estimate of the contact. Values of C_f and E_r are thus obtained by plotting C as a function of $A^{-1/2}$ for the large indentations, Fig. 5.11.

Contact areas for indentations are made using the measured C_f value at shallow depths on the aluminum with measured E_r and/or on a harder fused silica surface with published values of E_r , by rewriting Eq. 5.9 as

$$A = \frac{\pi}{4} \frac{1}{E_r^2} \frac{1}{(C - C_f)^2} \quad (5.10)$$

from which an initial guess at the area function is made by fitting A as a function h_c data to an eighth-order polynomial

$$A = 24.5h_c^2 + C_1h_c + C_2h_c^{1/2} + C_3h_c^{1/4} + \dots + C_8h_c^{1/128} \quad (5.11)$$

where C_1 through C_8 are constants. The first term describes the perfect shape of the indenter; the others describe deviations from the Berkovich geometry due to blunting of the tip. A convenient fitting routine is that contained in the Kaleida-graph software. A weighted procedure can be used to assure that data points with small and large magnitudes are of equal importance. An iterative approach can be used to refine the values of C_f and E_r further.

For calculations of load frame compliance and indenter area function, a series of indents are made in two standard materials—aluminum and fused quartz. It is assumed that both these materials are elastically isotropic, their moduli are well known, and their moduli are independent of indentation depth (Oliver and Pharr 1992). The first step is to determine the load frame compliance precisely by indenting a well-annealed, high-purity aluminum, typically at indentation depths from 700 to 4000 nm. As an example, the following load time history can be used. (1) Approach and contact surface, (2) load to peak load, (3) unload to 90% of peak load and hold for 100 s, (4) reload to peak load and hold for 10 s, and (5) unload completely. The lower hold is used to establish thermal drift and the upper hold to minimize time-dependent plastic effects. The final unloading data are used to determine the unloading compliances using the power law fitting procedure described earlier. The load frame compliance is determined from the aluminum data by plotting the measured compliance as a function of area calculated, assuming the ideal Berkovich indenter. Calculated E_r is checked with known elastic constants for aluminum, $E = 70.4$ GPa and $v = 0.347$.

The problem with using aluminum to extend the area function to small depths is that because of its low hardness, small indentations in aluminum require very small loads, and a limit is set by the force resolution of the indentation system (Bhushan 1999a; Bhushan and Li 2003). This problem can be avoided by making the small indentations in fused quartz, a much harder, isotropic material available in optically-finished plate form. Typically, measurements are made at six or so peak loads at depths ranging from about 15 to 700 nm. Above 700 nm of depth, indenter can be assumed to have a perfect shape. The contact areas and contact depths are then determined using Eq. 5.11 and h_c in conjunction with the reduced modulus computed from the elastic constant for fused quartz, $E = 72$ GPa and $v = 0.170$. The machine compliance is known from the aluminum analysis. The area function is only good for the depth range used in the

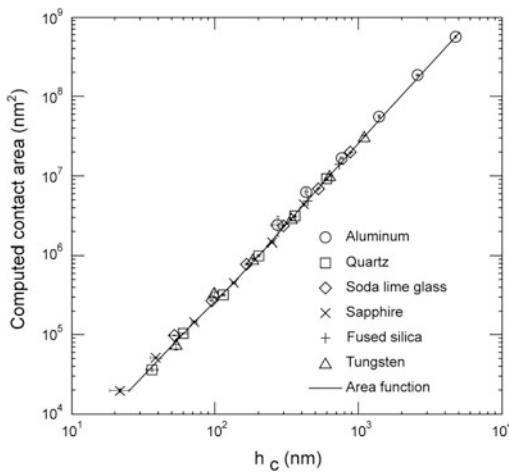


Fig. 5.12 The computed contact areas as a function of contact depths for six materials. The error bars are two standard deviations in length (adapted from Oliver and Pharr 1992)

calculations. Typical data of contact areas as a function of contact depths for six materials is shown in Fig. 5.12.

In situ scanning electron microscope (SEM) and transmission electron microscope (TEM) indentation tests are expected to give an in-depth understanding of elastic/plastic deformation behavior. Some nanoindenter instruments can be operated inside an SEM or TEM.

5.2.2.4 Hardness/Modulus² Parameter

Calculations of hardness and modulus described so far require the calculations of the indent's projected contact area from the indentation depth, which are based on the assumption that the test surface be smooth to dimensions much smaller than the projected area (Bhushan 1999a; Bhushan and Li 2003). Therefore, data obtained from rough samples show considerable scatter. Joslin and Oliver (1990) developed an alternative method for data analysis without requiring the calculations of the projected contact area of the indent. This method provides measurement of a parameter hardness/modulus², which provides a measure of the resistance of the material to plastic penetration.

Joslin and Oliver (1990) showed that for several types of rigid punches (cone, flat punch,

parabola of revolution, and sphere) as long as there is a single contact between the indenter and the specimen,

$$H/E_r^2 = (4/\pi)(W/S^2) \quad (5.12)$$

where S is the stiffness obtained from the unloading curve. E_r is related to E_s by a factor of $1 - v_s^2$ for materials with moduli significantly less than diamond (Eq. 5.6). The H/E_s^2 parameter represents a materials resistance to plastic penetration. We clearly see that calculation of projected contact area and knowledge of area function are not required. However, this method does not give the hardness and modulus values separately.

5.2.2.5 Continuous Stiffness Measurement (CSM)

The continuous stiffness measurement technique allows measurement of sample stiffness during indentation without the need for discrete unloading cycles, and with a time constant that is at least three orders of magnitude smaller than the time constant of the more conventional method of determining stiffness from the slope of an unloading curve (Pethica and Oliver 1989; Li and Bhushan 2002a). Furthermore, the measurements can be made at exceedingly small penetration depths. Thus, this method is ideal for determining the stiffness and, hence, the elastic modulus and hardness of films a few tens of nanometers thick. Furthermore, its small time constant makes it especially useful in measuring the properties of some polymeric materials.

Measurement of continuous stiffness is accomplished by the superposition of a very small AC current of a known relatively high frequency (typically 69.3 Hz) on the loading coil of the indenter, Fig. 5.13. This current, which is much smaller than the DC current that determines the nominal load on the indenter, causes the indenter to vibrate with a frequency related to the stiffness of the sample and to the indenter contact area. A comparison of the phase and amplitude of the indenter vibrations (determined with a lock-in amplifier) with the phase and amplitude of the imposed AC signal allows the stiffness to be calculated either in terms of amplitude or phase.

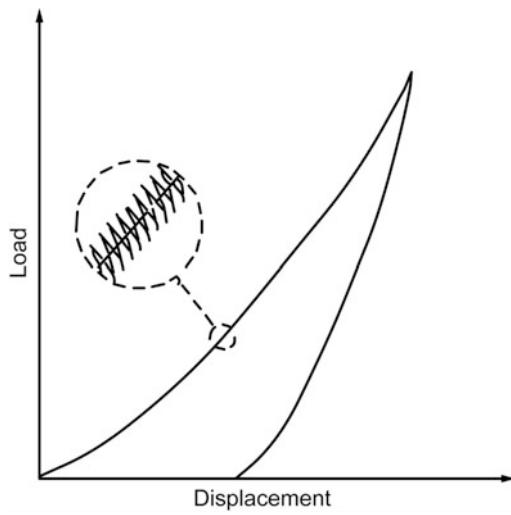


Fig. 5.13 Schematic of loading cycle in the continuous stiffness option

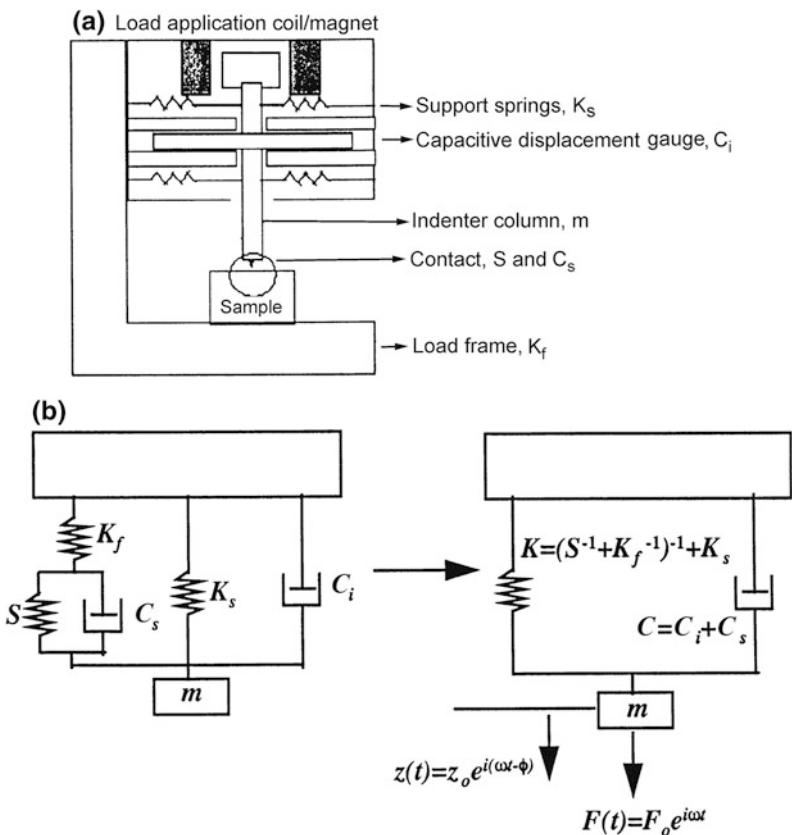
For calculation of stiffness, the dynamic response of the indentation system has to be determined (Li and Bhushan 2002a; Bhushan

Fig. 5.14 a Schematic of a nanoindenter with each component represented in the simple harmonic oscillator models, and b the dynamic model of the indentation system (adapted from Bhushan and Li 2003)

and Li 2003). The relevant components are the mass m of the indenter, the spring constant K_f of the leaf springs that support the indenter (load frame stiffness), the stiffness of the indenter frame K_s , and the damping constant C_i due to the air in the gaps of the capacitor plate displacement-sensing system. These combine with the sample stiffness, S , and damping constant of the contact zone, C_s , as shown schematically in Fig. 5.14 to produce the overall response. If the imposed driving force is $F(t) = F_0 \exp(i\omega t)$ and the displacement response of the indenter is $z(t) = z_0 \exp(i\omega t - \phi)$, the ratio of amplitudes of the imposed force and the displacement response is given by (Pethica and Oliver 1989)

$$\left| \frac{F_0}{z_0} \right| = \left[(K - m\omega^2)^2 + \omega^2 C^2 \right]^{1/2} \quad (5.13)$$

and the phase angle, ϕ , between the driving force and the response is



$$\tan \phi = \omega C / (K - m\omega^2) \quad (5.14)$$

Equations 5.13 and 5.14 may be solved simultaneously for K and C . The stiffness and damping of the contact are given by

$$S = \left[\frac{1}{\frac{F_0}{z_0} \cos \phi - (K_s - m\omega^2)} - \frac{1}{K_f} \right]^{-1} \quad (5.15a)$$

and

$$C_s \omega = \frac{F_0}{z_0} \sin \phi - C_i \omega \quad (5.15b)$$

With the exception of S and C_s terms, all the terms in Eqs. 5.15a, 5.15b can be measured independently. The parameters m , K_s , and C_i are determined by analyzing the dynamic response of the system when the indenter is hanging free, which is done in the factory. Details on K_f have been presented earlier. In a CSM experiment, the excitation frequency (ω) is set. The AC input to the force coil is generated with a standard AC signal generator, and any frequency between about 10 and 150 Hz may be selected. The displacement amplitude (z_0) and phase angle (ϕ) (using a lock-in amplifier) are measured. Using Eqs. 5.15a, 5.15b, S and $C_s \omega$ are calculated.

5.2.2.6 Bending Experiments by Beam Deflection Measurement

Bending and bending fatigue properties of suspended cantilevered and doubled anchored (fixed) beams can be measured using a nanoindenter or an AFM (Sundararajan et al. 2002; Li and Bhushan 2003; Li et al. 2003; Wei et al. 2005; Palacio et al. 2007a, b). The advantage of a nanoindenter is that loads up to about 400 mN higher than those in AFM (up to about 100 μ N) can be used for some structures requiring high loads. Figure 5.15 shows a schematic of the beam bending in the normal and lateral direction using a nanoindenter (Palacio et al. 2007a, b). To avoid a sharp tip pushing into the beam specimen, a blunt tip is used.

For bending experiments, load as a function of displacement is measured, and the slope of the

curve is obtained. For a cantilever beam with one end clamped, the elastic modulus E is given as (Young et al. 2012),

$$E = \frac{\ell^3}{3I} m \quad (5.16a)$$

where ℓ is the beam length, I is the area moment of inertia for the beam cross-section, and m is the slope of the linear region of the force displacement curve. For a fixed elastic beam loaded at the center of the span in the normal direction, as shown in Fig. 5.15 (middle), the elastic modulus is expressed as (Young et al. 2012)

$$E = \frac{\ell^3}{192I} m \quad (5.16b)$$

for a beam with a rectangular cross-section of width w and thickness t , $I = wt^3/12$.

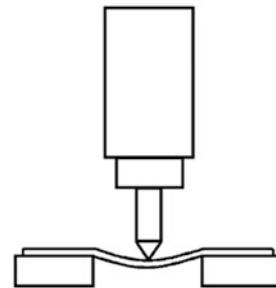
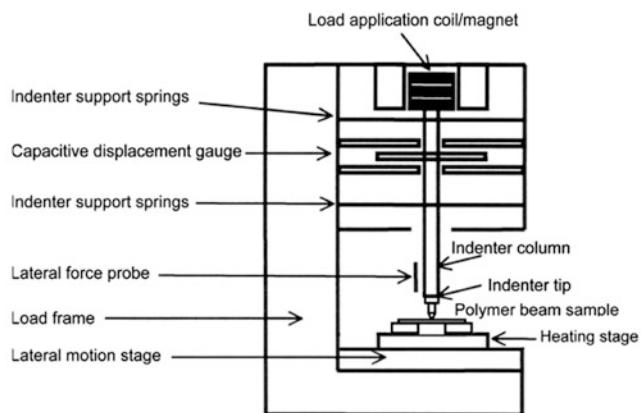
The application of simple beam theory to the load deflection data of beams also enable one to determine the yield strength of the microbeam material. For a homogeneous cantilever beam under load at the end of the beam, the strain in the beam varies linearly through the thickness such that the maximum strain at a given length occurs at the top and at the bottom of the beam. The top of the beam is in tension and the bottom of the beam is in compression. In addition, the maximum stress in the beam is located at the fixed end of the beam where the applied moment is greatest. When this maximum stress reaches the yield strength of the material, the beam begins to deform plastically. The onset of such deformation can be recognized in the plot of load versus deflection by a deviation from linearity. The load that marks this deviation is defined as the yield load W_y , and the yield strength for a rectangular beam is given by (Young et al. 2012),

$$\sigma_y = 6W_y \ell / (bt^2) \quad (5.17)$$

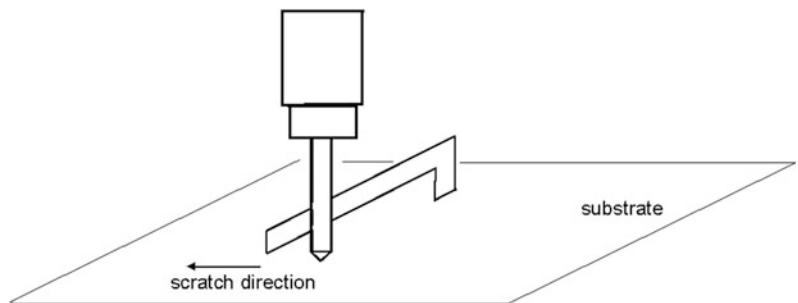
Following yielding, the beam continues to bend as more load is applied, and some of the strain in the beam is plastic. Thus, it is unrecoverable during unloading, and the beam eventually fails. The shape of the load versus

Fig. 5.15 Schematic of cantilever beam bending test in the normal and lateral direction using a nanoindenter (adapted from Palacio et al. 2007b)

Schematic of normal beam bending test



Schematic of lateral beam bending test



deflection curves after yielding and prior to unloading depends on the elastic and plastic properties of the material.

5.2.2.7 Determination of Hardness and Modulus of Elasticity of Thin Films from the Composite Response of Film and Substrate

It is widely accepted that to measure true hardness of films, the indentation depth should not exceed 10% of the film thickness (Tabor 1951) to 30% of the film thickness (Anonymous 1979). Otherwise, measured hardness is affected by the substrate properties. A number of expressions have been derived that relate thin-film hardness to substrate hardness, composite hardness (measured on the coated substrate), and film thickness. These analyses allow the calculation of the thin-film hardness from the measured composite hardness data (Buckle 1973; Doerner and Nix 1986; Sargent 1986; Burnett and Rickerby 1987b; King 1987; Bhattacharya and Nix 1988a, b; Bull and Rickerby 1990; Vinci and Vlassak 1996; Korsunsky et al. 1998; McGurk and Page 1999; Tsui and Pharr 1999). Here we discuss two models based on the volume law of mixtures (volume fraction model) (Sargent 1986) and numerical analysis (King 1987; Bhattacharya and Nix 1988a, b; Bhushan and Venkatesan 2005; Bhushan 2013a, b).

The hardness of a film/substrate composite generally is believed to be a weighted average of the volume of plastically-deformed material in the film (V_f) and that in the substrate (V_s) (Sargent 1986),

$$H = H_f \frac{V_f}{V} + H_s \frac{V_s}{V} \quad (5.18)$$

where $V = V_f + V_s$. The deformed volumes of film and substrate can be calculated using expanding spherical cavity model (Johnson 1985).

Bhattacharya and Nix (1988a, b) modeled the indentation process using the finite-element method to study the elastic-plastic response of materials. Bhattacharya and Nix (1988b)

calculated elastic and plastic deformation associated with submicron indentation by a conical indenter of thin films on substrates, using the finite-element method. The effects of the elastic and plastic properties of both the film and substrate on the hardness of the film/substrate composite were studied by determining the average pressure under the indenter as a function of the indentation depth. They developed empirical equations for film/substrate combinations for which the substrate is either harder or softer than the film. For the case of a soft film on a harder substrate, the effect of substrate on film hardness can be described as

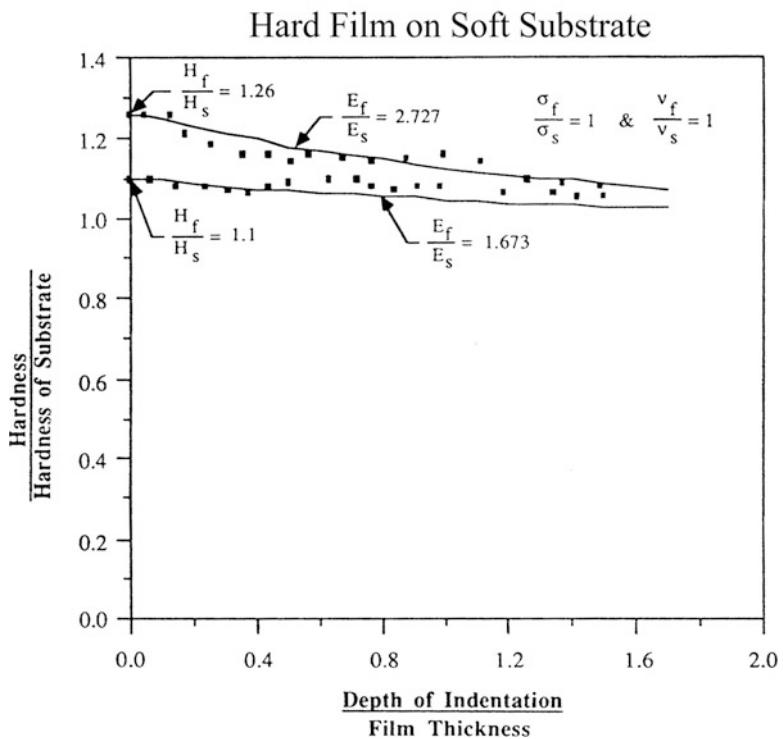
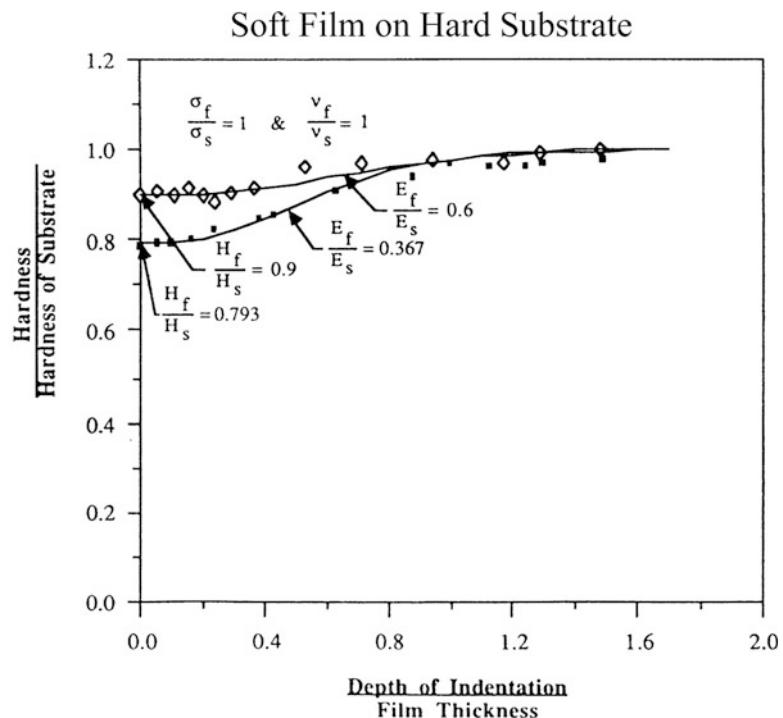
$$\frac{H}{H_s} = 1 + \left(\frac{H_f}{H_s} - 1 \right) \exp \left[- \frac{(\sigma_f/\sigma_s)}{(E_f/E_s)} (h_c/t_f)^2 \right] \quad (5.19a)$$

where E_f and E_s are the elastic moduli, σ_f and σ_s are the yield strengths, and H_f and H_s are the hardnesses of the film and substrate, respectively. H is the hardness of the composite, h_c is the contact indentation depth, and t_f is the film thickness. Similarly, for the case of a hard film on a softer substrate, the hardness can be expressed as

$$\frac{H}{H_s} = 1 + \left(\frac{H_f}{H_s} - 1 \right) \exp \left[- \frac{(H_f/H_s)}{(\sigma_f/\sigma_s)(E_f/E_s)^{1/2}} (h_c/t_f) \right] \quad (5.19b)$$

Composite hardness results were found to depend only very weakly on Poisson's ratio (ν). Figure 5.16 presents plots showing the effect of relative elastic moduli of the film and the substrate on the composite hardness as a function of the depth of indentation for a soft film on a hard substrate and a hard film on a soft substrate. For cases in which the film and substrate have different elastic moduli, hardness is observed to be independent of the substrate for indentation depths less than about 30% of the film thickness, after which the hardness slowly increases/decreases because of the presence of the substrate (also see Bhushan and Venkatesan 2005). Thus, the "30% rule" is preferred. They also reported that for cases

Fig. 5.16 Effect of relative elastic moduli of the film and the substrate (E_f/E_s) on the composite hardness as a function of the depth of indentation for a soft film on a hard substrate and a hard film on a soft substrate. Yield strengths (σ) and Poisson's ratio (ν) are same for both substrates and films (adapted from Bhattacharya and Nix 1988b)



in which the film and substrate have different yield strengths, it is observed that the variation of hardness with depth of indentation in these cases is qualitatively similar to cases in which the film and substrate have different elastic moduli.

Doerner and Nix (1986) modeled the influence of the substrate on the elastic measurement of very thin film empirically in an indentation test using the following expression for the compliance:

$$C = \frac{dh}{dW} = \frac{1}{2} \left(\frac{\pi}{A} \right)^{1/2} \left\{ \frac{1 - v_f^2}{E_f} \left[1 - \exp \left(\frac{-\alpha t_f}{\sqrt{A}} \right) \right] + \frac{1 - v_s^2}{E_s} \exp \left(\frac{-\alpha t_f}{\sqrt{A}} \right) + \frac{1 - v_i^2}{E_i} \right\} + b$$

where the subscripts *f*, *s*, and *i* refer to the film, substrate, and indenter, respectively. The term \sqrt{A} is equal to $(24.5)^{1/2} h_c$ for the Vickers or Berkovich indenter. The film thickness is t_f , and *b* is the y-intercept for the compliance versus 1/depth plot, obtained for the bulk substrate, which can be neglected in most cases. The weighing factors $[1 - \exp(-\alpha t_f / \sqrt{A})]$ and $\exp(-\alpha t_f / \sqrt{A})$ have been added to account for the changing contributions of the substrate and film to the compliance. The factor α can be determined empirically.

King's analysis (1987) verified that Eq. 5.20 is an excellent functional form for describing the influence of the substrate and theoretically determined the values of α for various indenter shapes. The value of α was found to depend on the indenter shape and size and film thickness, and was found to be independent of E_i/E_s . The values of α as a function of \sqrt{A}/t_f for Berkovich (triangular) indenters are shown in Fig. 5.17. The values of α were found to be similar for square and triangular indenters. Bhattacharya and Nix (1988b) analyzed the deformations of a layered medium in contact with a conical indenter using the finite-element method. Their analysis also verified the relationship given in Eq. 5.20.

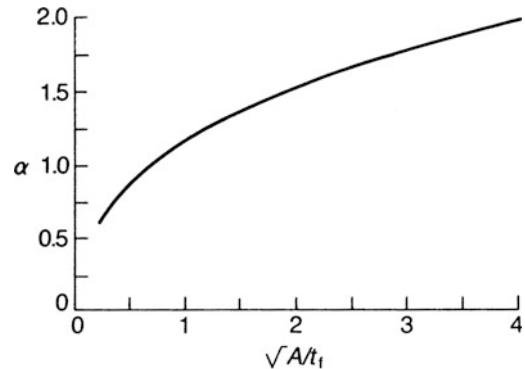


Fig. 5.17 Parameter α as a function of normalized indenter size for Berkovich indenter indenting a layered solid surface (adapted from King 1987)

5.3 Measured Mechanical Properties of Engineering Materials

We present representative mechanical properties data obtained on various materials and coatings (Bhushan 1999a; Bhushan and Li 2003).

5.3.1 Load–Displacement Curves

A variety of mechanical phenomena, such as transition from elastic to plastic deformation, creep deformation, formation of subsurface cracks, and crystallographic phase transition, can be studied by the load-displacement curves obtained at different loading conditions (Pethica et al. 1983; Doerner and Nix 1986; LaFontaine et al. 1990, 1991; Pharr et al. 1990; Page et al. 1992; Oliver and Pharr 1992; Pharr 1992; Whitehead and Page 1992; Gupta et al. 1993, 1994; Gupta and Bhushan 1994, 1995a, b; Bhushan et al. 1995, 1996a, b; Bhushan and Gupta 1995; Bhushan 1999a; Bhushan and Li 2003; Palacio and Bhushan 2010, 2013; Kumar and Bhushan 2015; Cho and Bhushan 2016).

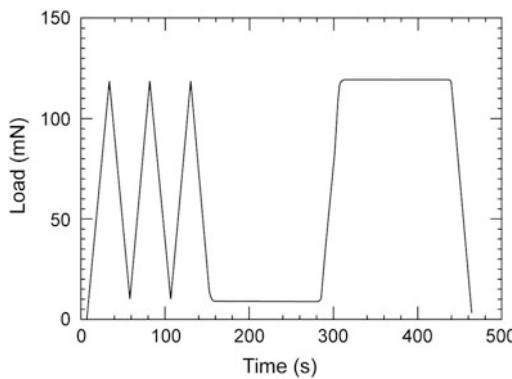


Fig. 5.18 A typical load-time sequence for mechanical property measurements (adapted from Oliver and Pharr 1992)

Mechanical properties measurements often are made using the following loading sequence: three loading-unloading cycles, hold for 100 s at 10% of the peak load, reload, hold for 100 s, and unload (Fig. 5.18) (Oliver and Pharr 1992). Load-displacement curves for electropolished single-crystal tungsten, fused silica, and single-crystal silicon (110) are shown in Figs. 5.19, 5.20 and 5.21, respectively. Tungsten data are typical of materials in which the hardness is relatively small compared to the modulus, as is observed in most metals; most of the indenter displacement in these metals is accommodated plastically and only a small portion is recovered on unloading. Fused silica and silicon are harder which show larger elastic recovery during unloading, the largest being that for fused silica.

The unloading/reloading behaviors of the various materials are different. For tungsten, Fig. 5.19 shows that the peak load displacements shift to higher values in successive loading/unloading cycles (Oliver and Pharr 1992). In addition, the relatively large displacement just prior to final unloading is due to creep during the 100 s hold period at peak load. Indentation at a very low load of 0.5 mN caused only elastic displacements, Fig. 5.19b. At two higher peak loads, the indentation is elastic and plastic, as shown in Fig. 5.19a, c, respectively. When a threshold load of about 1 mN is reached, a sudden jump in displacement corresponding to

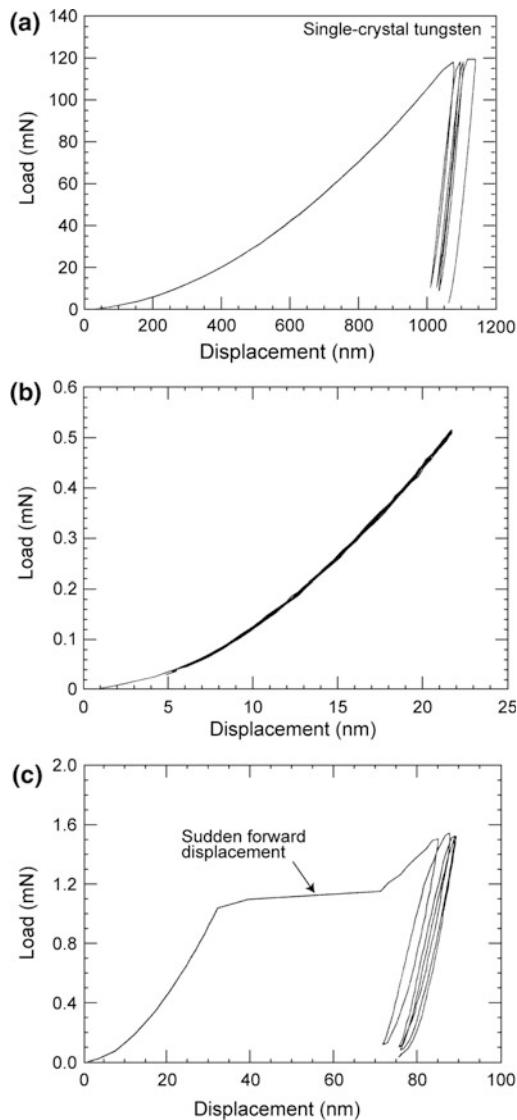


Fig. 5.19 Load-displacement curves for an electro-polished single-crystal tungsten **a** at a peak load of 120 mN, **b** at a peak load of 0.5 mN (elastic contact), and **c** at a peak load of 1.5 mN showing the yield point (adapted from Oliver and Pharr 1992)

the onset of plasticity is observed, and a permanent hardness impression is formed. At mid-load (Fig. 5.19c), a distinct hysteresis loop is observed, as might be expected if there were a small amount of reverse plasticity upon loading. However, the looping degenerates with cycling after three or four cycles, the load-displacement behavior is largely elastic. The

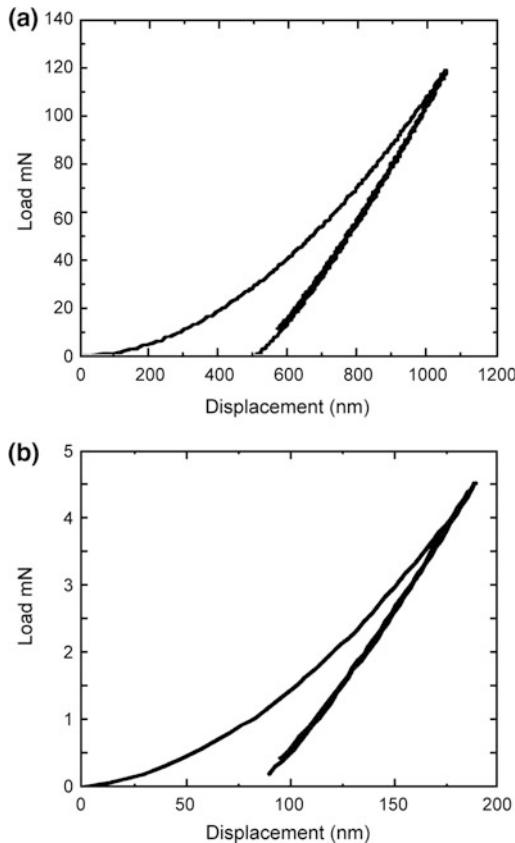


Fig. 5.20 Load-displacement curves for fused silica **a** at a peak load of 120 mN, and **b** at a peak load of 4.5 mN (adapted from Pharr 1992)

unloading/reloading curves for fused silica are nearly the same, Fig. 5.20 (Pharr 1992). The near-perfect reversibility suggests that at peak loads of 120 and 4.5 mN, deformation after initial unloading is almost entirely elastic.

The loading/unloading behavior of silicon shown in Fig. 5.21, is in sharp contrast to other materials (Figs. 5.19 and 5.20) (Pharr et al. 1989, 1990; Page et al. 1992; Pharr 1992). The data presented in Fig. 5.21 were taken over two cycles of loading and unloading. At high peak loads, the initial unloading curve for silicon is not at all smooth, but exhibits a discontinuity, Fig. 5.21a. At lower peak loads, the behavior changes, and below some critical value the discontinuity is no longer observed, Fig. 5.21b. However, at this load the load-displacement

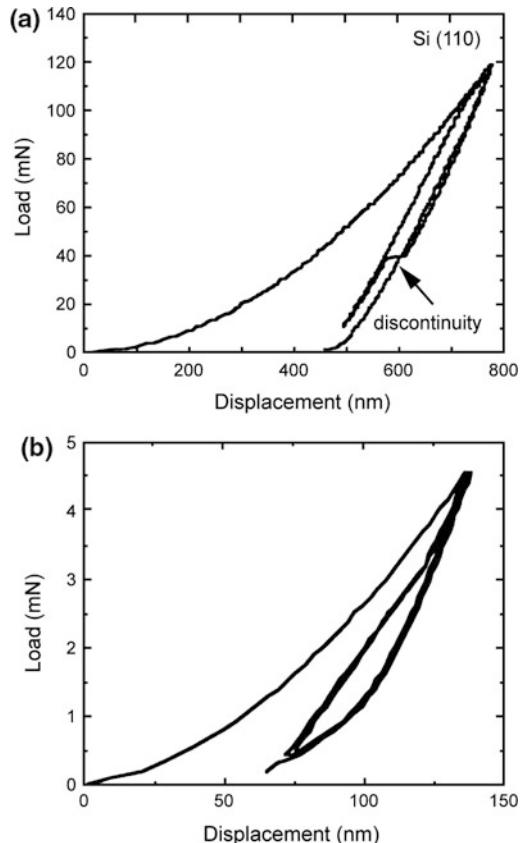


Fig. 5.21 Load-displacement curves for single-crystal Si (110) **a** at a peak load of 120 mN, and **b** at a peak load of 4.5 mN (adapted from Pharr 1992)

behavior shows another anomalous feature—a large hysteresis, which shows no sign of degeneration through several cycles of deformation. The fact that the curves are highly hysteretic implies that deformation is not entirely elastic. The discontinuity at high loads and the nondegenerative hysteresis at low loads are quite unique to silicon and are observed in each of the (100), (110), and (111) orientations. The load below which the discontinuity disappears and the hysteresis becomes apparent is generally in the range 5–20 mN. Pharr (1992) concluded that larger hysteresis observed in the unloading curve at low loads is due to a pressure-induced phase transformation from its normal diamond cubic form to a β -tin metal phase. At some point in the transformation, an amorphous phase is formed,

whose evidence is reported by Callahan and Morris (1992). The discontinuity in displacement observed during unloading at peak loads of greater than about 15 mN are believed to be due to formation of lateral cracks which form at the base of the median crack which results in the surface of the specimen being thrust upward, schematically shown in Fig. 5.22 (Pharr 1992). Lateral cracking is aided by the phase transformations.

AE signal collected during the indentation process allows understanding of the deformation process. The load-displacement curve for Si(100) with a 120 mN load is shown in Fig. 5.23a, and the first AE signal that was recorded for this test is shown in Fig. 5.23b (Weihs et al. 1992). The AE signals correlated with small jumps in tip displacement. After testing, radial cracks were visible at the corners of indentation. AE events such as the one shown in Fig. 5.23b were

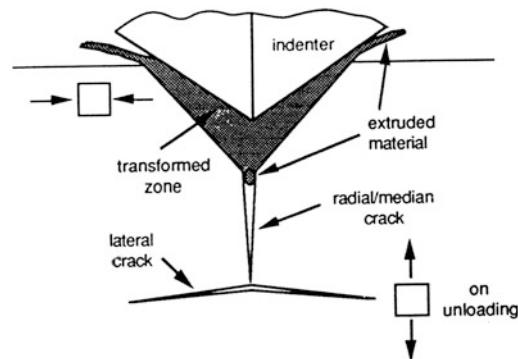


Fig. 5.22 A schematic illustration of the coupling of the phase transformation and cracking during indentation of single-crystal silicon (adapted from Pharr 1992)

recorded at applied load as low as 48 mN on loading. In the final stages of unloading, small AE signals that had an inverted shape compared to Fig. 5.23b were detected occasionally. Goken and Kempf (2001) reported that the

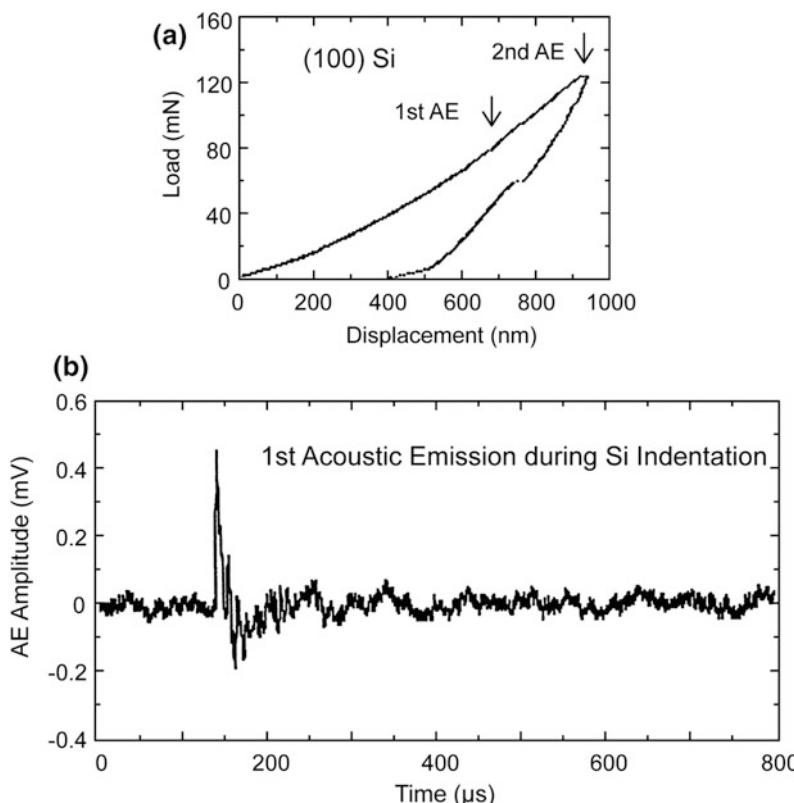


Fig. 5.23 **a** Load-displacement curves for single-crystal Si(100), and **b** AE amplitude as a function of time during loading (adapted from Weihs et al. 1992)

load-displacement curves obtained by nanoindentations on most metallic and intermetallic materials show discontinuities or pop-ins during the initial part of loading. These pop-ins mark a sharp transition from pure elastic loading to a plastic deformation of the specimen surface, thus correspond to an initial yield point. On smooth surfaces pop-ins are observed frequently, but not on surfaces with a high roughness. Step edges on the surface are believed to act as dislocation sources for the initial yield events. It should be noted that the pop-in load determines a minimum load necessary to generate plastic indentation.

Gupta et al. (1993, 1994) and Gupta and Bhushan (1994) reported that hysteresis in cyclic indentation and discontinuity kinks in the unloading curve are considerably reduced by ion implantation of compound forming species O⁺ and N⁺ into single-crystal silicon. They further reported that amorphous silicon films did not

exhibit either hysteresis in the cyclic indentation or a discontinuity kink in the indentation loads ranging from 1 to 90 mN. This suggests that for a perfect amorphous structure, hysteresis does not occur because of the absence of crystallographic pressure-induced phase transition during cyclic loading and unloading. In addition, the disordered structure does not allow the nucleation and propagation of the lateral cracks beneath the indentation.

The load-displacement curve for indented Ni films on a glass substrate is shown in Fig. 5.24a (Weihs et al. 1992). The Ni films debonded from their substrates at forces ranging between 130 and 250 mN. The debonding events were marked by the indenter tip jumping downward as “chunks” of the Ni film buckled away from the underneath. The indenter tip jumped a distance equal to the film thickness as it initially debonded. Figure 5.24b also shows the corresponding

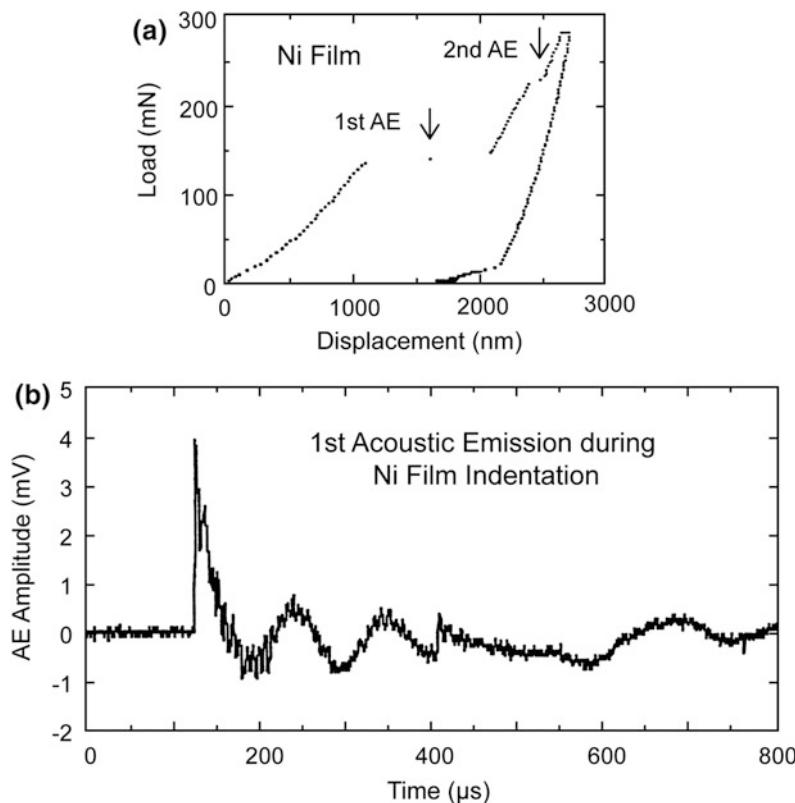


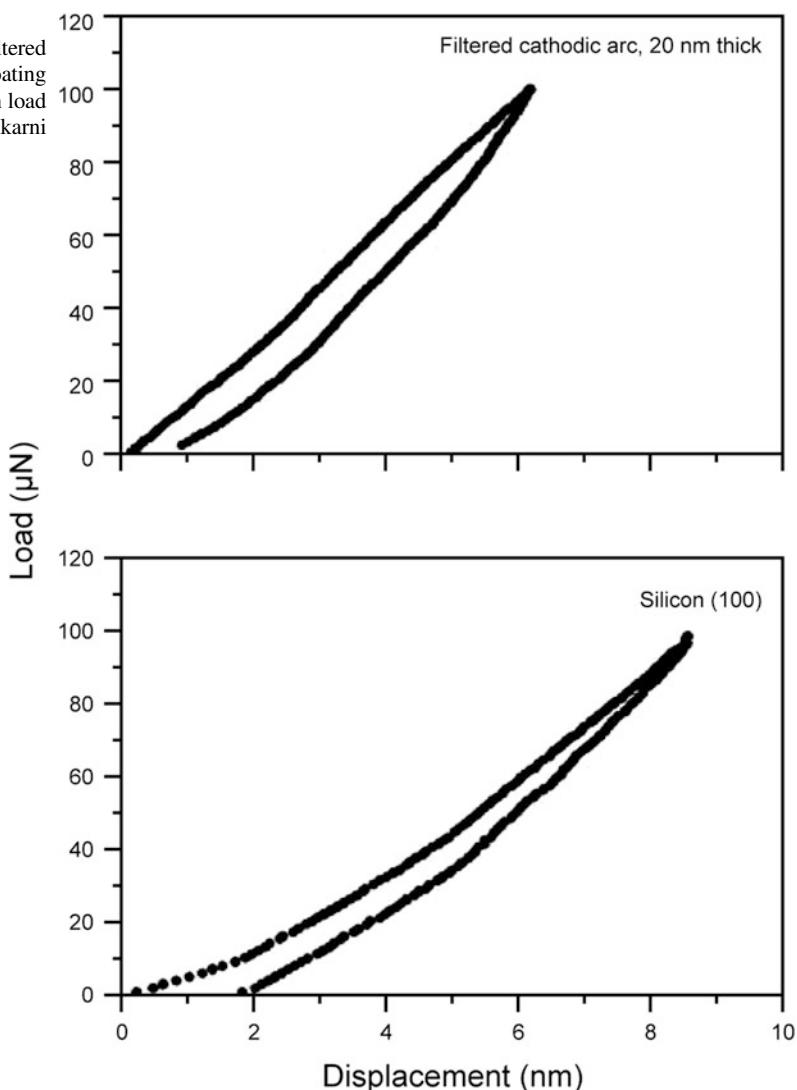
Fig. 5.24 **a** Load-displacement curves for Ni film on glass substrate, and **b** AE amplitude as a function of time during loading (adapted from Weihs et al. 1992)

AE trace with a rise time of 1.8 μ s. In this particular test, debonding continued at higher forces and a second AE event was recorded. After each test, optical microscopy confirmed the delamination of the film from underneath the indenter.

Li and Bhushan (1999c) measured mechanical properties of 100-nm thick DLC coatings deposited on a silicon substrate at a peak load of 0.2 mN using a Keysight-type nanoindenter by four different deposition techniques: filtered cathodic arc (FCA); ion beam (IB); electron cyclotron resonance-chemical vapor deposition

(ECR-CVD); and sputtered (SP). Kulkarni and Bhushan (1997) measured mechanical properties of a 20-nm thick DLC coating deposited by FCA at a peak load of 100 μ N using a Hysitron nanoindenter, Fig. 5.25. The data for Si(100) also is shown for comparison. The indentation depths at the peak load ranged from 6 to 8 nm, smaller than the coating thickness. Indentation depth of the coating at the peak load was slightly lower than that of the Si substrate, exhibiting less plastic deformation, which suggests higher hardness.

Fig. 5.25 Load-displacement curves for a 20 nm thick filtered cathodic arc deposited DLC coating and Si(100) at peak indentation load of 100 μ N (adapted from Kulkarni and Bhushan 1997)



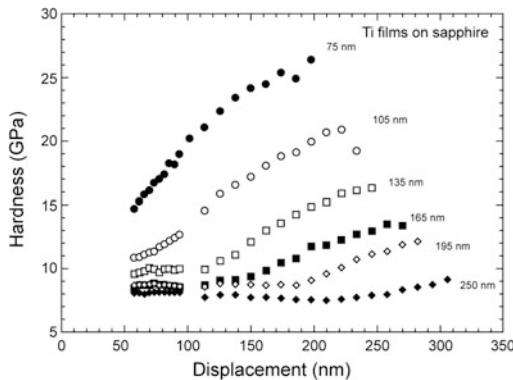
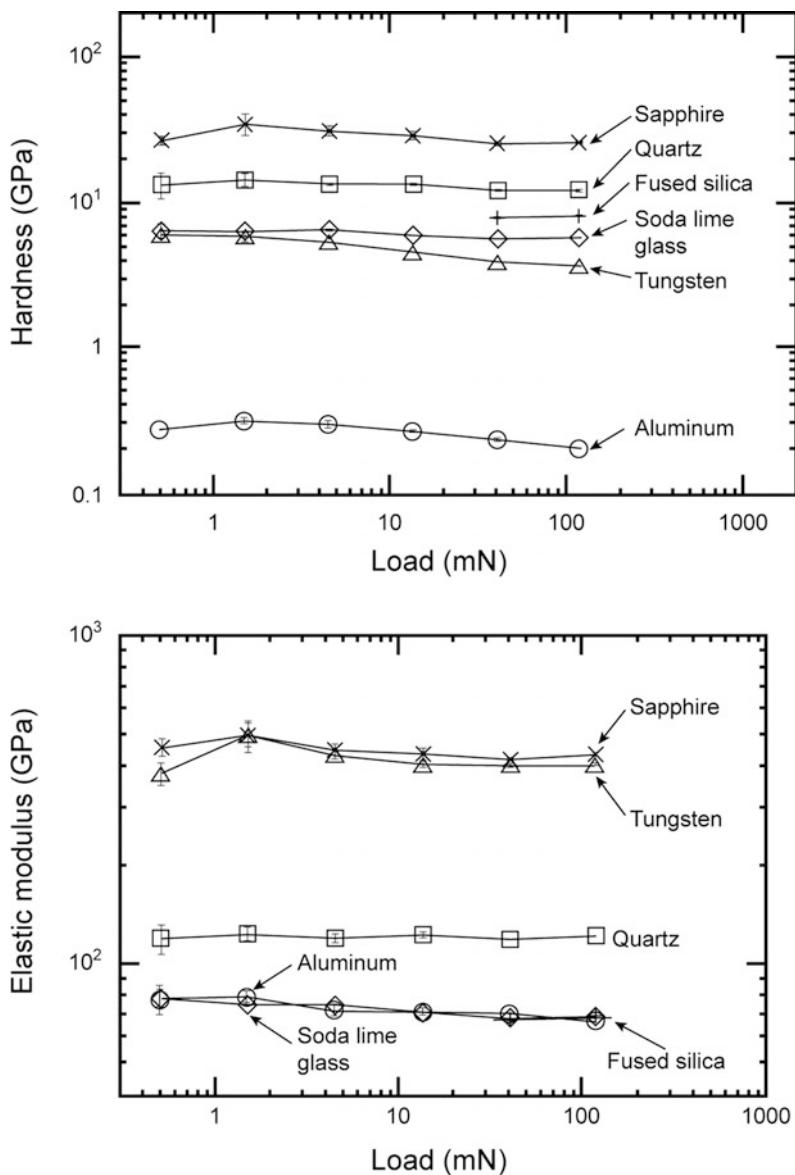


Fig. 5.26 Indentation hardness as a function of displacement for Ti films on sapphire substrates. Numbers next to each set of data correspond to the coating thickness (adapted from Fabes et al. 1992)

Fig. 5.27 Hardness and elastic modulus as a function of load for six bulk materials—mechanically polished single-crystal aluminum, electro-polished single-crystal tungsten, soda-lime glass, fused silica, (001) single-crystal sapphire, and (001) single-crystal quartz (adapted from Oliver and Pharr 1992)

5.3.2 Hardness and Elastic Modulus Measurements

It has been reported earlier that true hardness of the films can be obtained if the indentation depth does not exceed about 30% of the film thickness. At higher indentation depths, the composite hardness changes with the indentation depth. Measured hardness values of soft, Ti films on a hard, sapphire substrate are presented in Fig. 5.26. We note that hardness increases with a decrease in the film thickness or increase in the indentation depth, as expected. The film hardness is the steady-state hardness (left portion), independent of the indentation depth.



Hardness and elastic moduli for six bulk materials are shown in Fig. 5.27. The data show that there is a very small indentation size effect in several materials on the hardness values. The moduli remains more or less constant over the entire range of load. The hardness and modulus values at the two highest loads are comparable with the literature values (Oliver and Pharr 1992). Measured values of various magnetic disk substrates and single-crystal crystal, as a reference, are shown in Fig. 5.28. A summary of measured nanoscale values and published micro/macroscale data are shown in Table 5.2. Scale effects are observed. Fracture toughness data will be discussed later.

5.3.3 Scale Effects on Nanomechanical Properties

In order to study the effect of contact depth, including shallow depths on the order of 5–10 nm, an AFM-based, Hysitron-type indenter is used frequently (Bhushan et al. 1996a). Figure 5.29a shows the load-displacement data for Si(100) and GaAs(100) at displacements as low as 2 nm (Bhushan et al. 1996a). Note that the loading and unloading curves are not smooth, but exhibit sharp discontinuities, particularly at high loads, shown by arrows. Discontinuities in the loading parts of the curve result from slip. Pharr et al. (1990) and Pharr (1992) reported the evidence of discontinuities in the unloading curves, and hypothesized that a sharp discontinuity is due to the formation of a lateral crack, which forms at the base of median crack, causing the surface of the specimen to be thrust upwards.

Figure 5.29b shows the load-displacement curves during three loading and unloading cycles for single-crystal Si(100) (Bhushan et al. 1996a). The unloading and reloading curves show a large “hysteresis,” which shows no sign of degeneration through three cycles of deformation, and the peak load displacements shift to higher values in successive loading–unloading cycles. Pharr et al. (1989, 1990), and Pharr

(1992) also have observed hysteresis behavior in silicon at similar loads using a nanoindenter. The fact that the curves are highly hysteretic implies that deformation is not entirely elastic. It is believed that large hysteresis is due to a pressure-induced phase transformation from its normal diamond cubic form to a β -tin metal phase (Bhushan 1999a).

The nanohardness values for single-crystal Si (100) as a function of residual depth are plotted in Fig. 5.30 (Bhushan et al. 1996a). It is clear that the hardness of silicon increases with a decrease in the load and corresponding indentation depth. A nanoindentation hardness value was about 11.5 GPa at a peak indentation depth of 750 nm, which is slightly higher than the bulk hardness value of silicon of 9–10 GPa (Anonymous 1988). The data shows that the hardness exhibits size effect. At smaller volumes, there is a lower probability of encountering material defects. Furthermore, at small volumes, there is an increase in the stress necessary to operate dislocation sources (Gane and Cox 1970; Sargent 1986). According to the strain gradient plasticity theory advanced by Fleck et al. (1994), large strain gradients inherent in small indentations lead to accumulation of geometrically necessary dislocations that cause enhanced hardening. In addition, an increase in hardness at lower indentation depths may result from contributions of the surface films. These are some of the plausible explanations for the increase in hardness at smaller volumes.

Kulkarni and Bhushan (1996a, b) measured the effect of contact depth on mechanical properties for single-crystal Al(100). Figure 5.31a shows the load-displacement curves at six different peak indentation loads of 15, 25, 50, 100, 300, and 500 μ N. For the case of 15 μ N peak load, the residual depth is about 3 nm. The peak depth at 500 μ N load is almost 200 nm. The indenter displacement results in plastic deformation with very little elastic recovery (Bhushan et al. 1996a; Kulkarni and Bhushan 1996a, b). The arrow marks in the loading portion of the curve indicate discontinuities (forward jump) in displacement at higher loads. As indicated earlier,

Fig. 5.28 Hardness and elastic modulus as a function of displacement (load ranging from 0.1 to 2.5 mN) for mechanically-polished 10- μm thick electroless Ni-P film on Al-Mg alloy 5086, chemically-strengthened alkali-aluminosilicate glass, chain silicate glass-ceramic (polycrystalline) (canasite by corning), and single-crystal Si (111) (adapted from Bhushan and Gupta 1995)

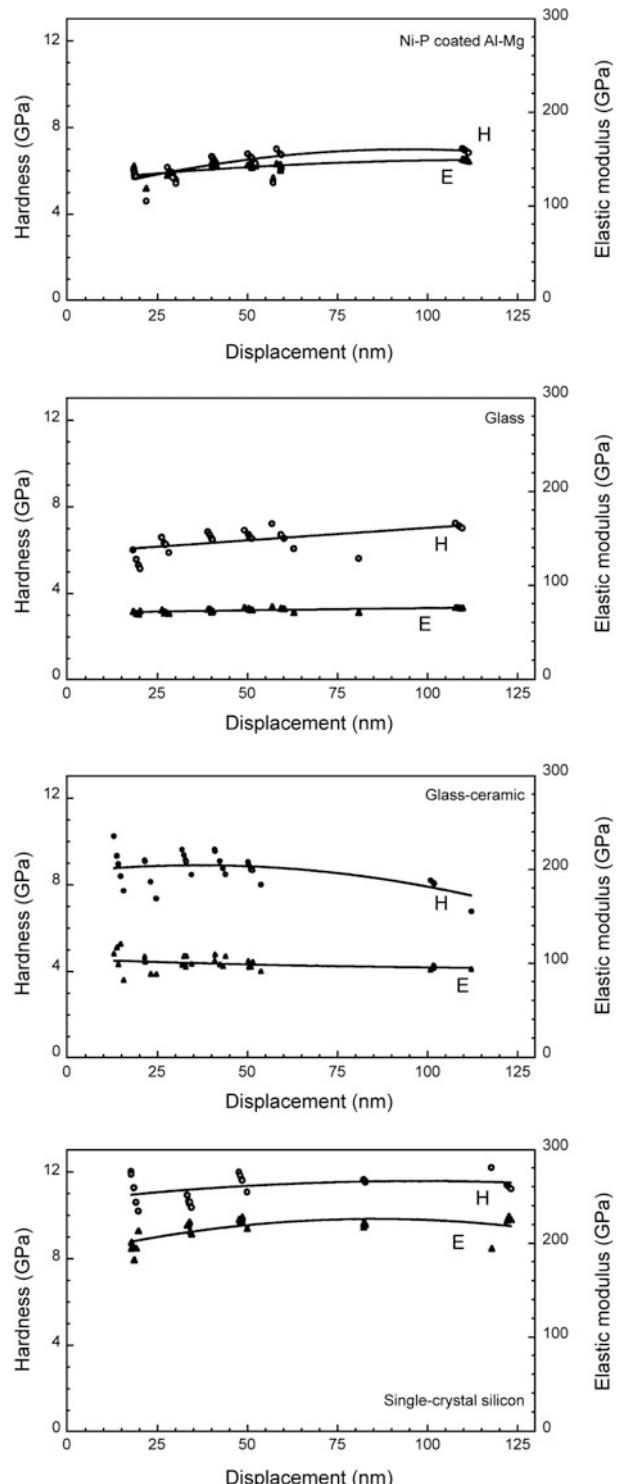


Table 5.2 RMS roughness values measured by an atomic force microscope, hardness and elastic modulus values measured by using a nanoindenter (at an indentation depth of about 20 nm), and fracture toughness measured by using a Vickers microindenter (Bhushan and Gupta 1995)

Material	RMS roughness (nm)	Hardness ^a (GPa)	Elastic modulus ^b (GPa)	Crack length, c (μm)	Fracture toughness ^c (MPa m ^{1/2})
Ni–P/Al–Mg	3.6	6.0 (5.5)	130 (200)	No cracks	—
Chemically-strengthened glass	1.1	6.0 (5.8)	85 (73)	Significant cracking	(0.9)
Glass-ceramic	6.1	8.5 (5.5)	100 (83)	59.4	0.65 (4.0)
Single-crystal Si(111)	0.95	11.0 (9–10)	200 (180)		

^aValues in parentheses are the reported values measured by conventional Vickers indentation method

^bValues in parentheses are the reported values measured by conventional tensile pull test method

^cValues in parentheses are the reported values measured by the chevron-notched short-bar method

discontinuities probably result from slip of the tip (arising because of stick-slip friction effects) and/or because of lateral cracks developed at the corners of indentation (Weihs et al. 1992).

Figure 5.31b shows the load displacement curves during three loading and unloading cycles. A hysteresis was observed after cyclic loading and unloading at peak indentation load of 100 μN. The peak load displacement shifts to higher values in successive cycles with no sign of degeneration, indicating that the deformation is not entirely elastic, even after three cycles (Oliver and Pharr 1992; Bhushan et al. 1996a). The hardness and elastic modulus as a function of load (or indentation depth) are shown in Fig. 5.32. The elastic modulus remains almost constant for all values of loads owing to the constant modulus assumption. Hardness decreases with increasing load. These results are consistent with earlier data on Si(100). As compared to single-crystal Si(100), aluminum has a higher dislocation density. Thus, the probability of the tip encountering defects is high. In addition, surface films may play a role in the higher hardness value observed at lower loads.

5.3.4 Time-Dependent Viscoelastic/Plastic Properties

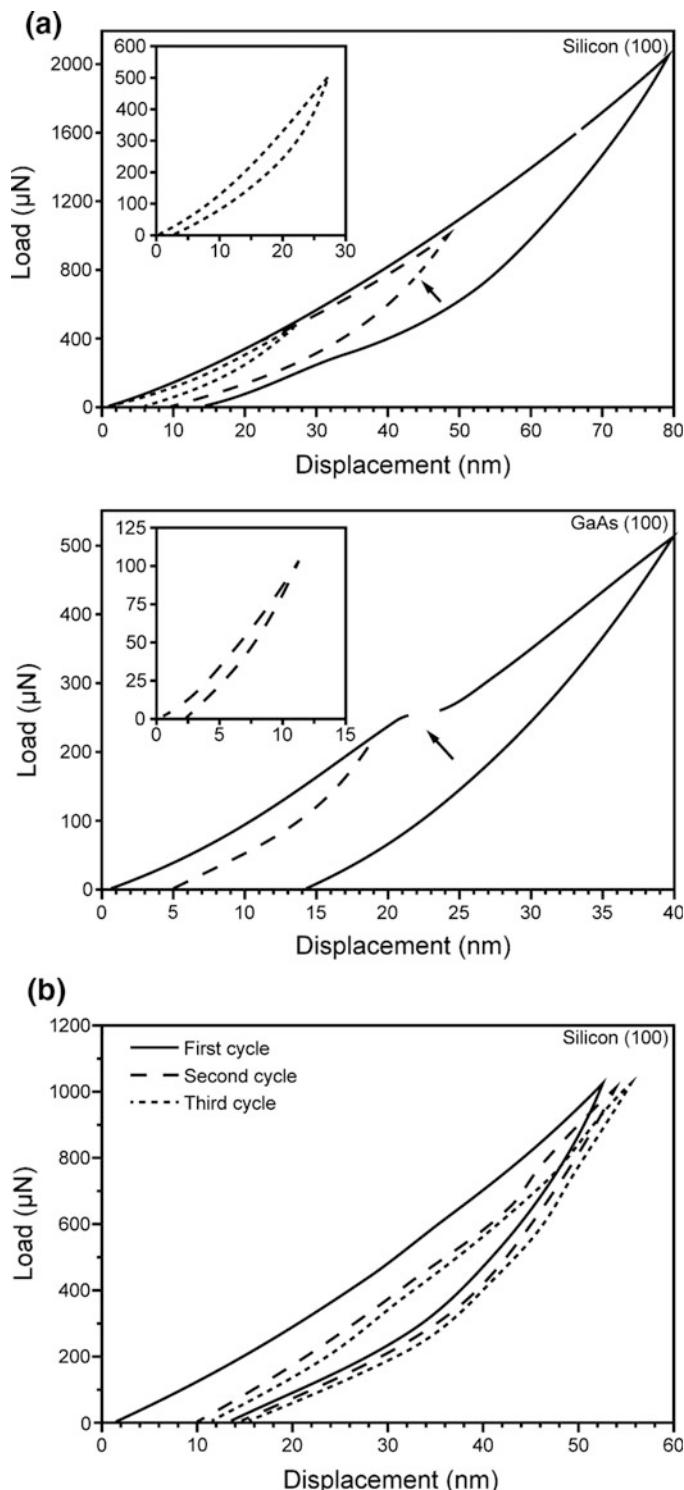
Viscoelastic solids such as polymers creep at temperatures as low as room temperature. Most

materials, including ceramics and even diamond, creep some at temperatures well below half their melting points. Indentation creep, load relaxation, and strain-rate sensitivity tests are used for measurement of the time-dependent flow of materials. These three tests provide the measure of viscoelastic/viscoplastic properties, and any of them can be used (Bhushan 1999a; Bhushan and Li 2003).

In the indentation creep test, the hardness indenter maintains its load over a period of time under well-controlled conditions, and changes in indentation size are monitored (Wesbrook, 1957; Mulhearn and Tabor 1960; Walker 1973; Bhushan et al. 1996a; Kulkarni and Bhushan 1996b, 1997; Li et al. 2001). The analysis of creep is more complex than that of creep data obtained using a conventional technique because of the shape of the tip. Indentation stress acting on the sample decreases with time as the contact area increases.

In the indentation load relaxation test, the indenter first is pushed into the sample at a fixed displacement rate until a predetermined load or displacement is achieved and the position of the indenter then is fixed (Bhushan 1999a; Bhushan and Li 2003). The material below the indenter is elastically supported and will continue to deform in a non-elastic manner, thereby tending to push the indenter farther into the sample. Load relaxation is achieved by conversion of elastic strain in the sample into inelastic strain in the sample. During the test, the load and position of the

Fig. 5.29 Load-displacement curves **a** at various peak loads for Si(100) and GaAs (100), and **b** during repeated loading-unloading cycles for Si(100). Sharp discontinuities in **(a)** are shown by arrows (adapted from Bhushan et al. 1996a)



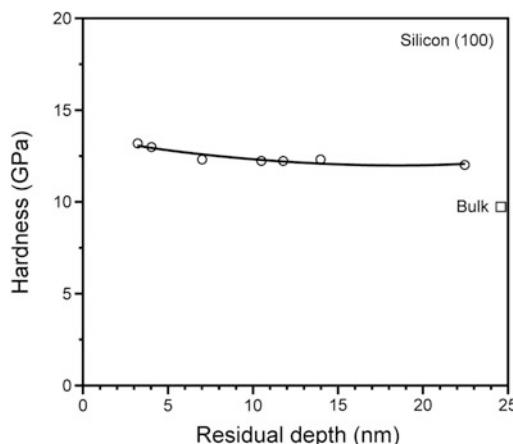


Fig. 5.30 Indentation hardness as a function of residual indentation depth for Si(100). Bulk hardness value is shown for reference (adapted from Bhushan et al. 1996a)

indenter and the specimen are monitored continuously. Typically, the indenter motion is held constant and the changes in the load are monitored as a function of time.

To measure strain rate sensitivity, individual indentations are performed at a prescribed loading rate that is varied from one indentation to another (Mayo and Nix 1988; Bhushan et al. 1996a; Kulkarni and Bhushan 1996b). The values of the indentation pressure and strain rates from tests performed at different loading rates are compared at a common indentation depth.

Li et al. (1991) conducted creep experiments using a nanoindenter on silicon at several hundred μN load. Their data exhibited significant creep only at high temperatures (greater than or equal to 0.25 times the melting point of silicon). The mechanism of dislocation glide plasticity was believed to dominate the indentation creep process. Using a Hysitron nanoindenter, indentation creep and strain rate sensitivity experiments were conducted on single-crystal silicon, single-crystal aluminum, and DLC coatings at low loads on the order of 1 μN or less (Bhushan et al. 1996a; Kulkarni and Bhushan 1996b, 1997). Figure 5.33a shows the load-displacement curves for single crystal Si(100) at various peak loads held at 180 s (Bhushan et al. 1996a). Note that significant creep occurs at room temperature.

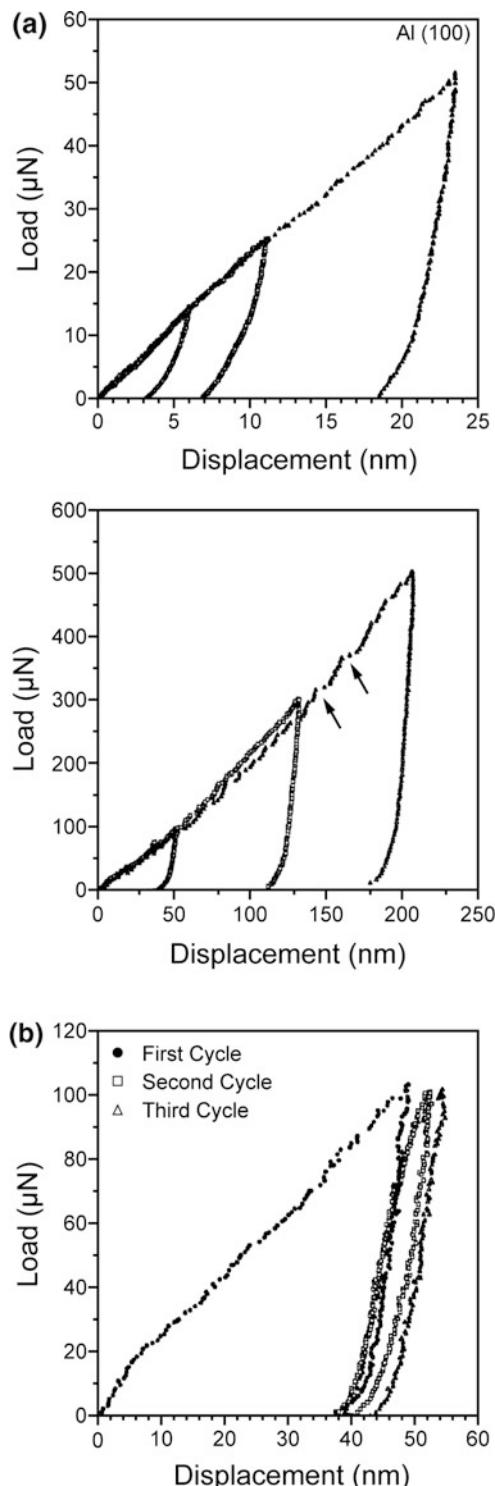
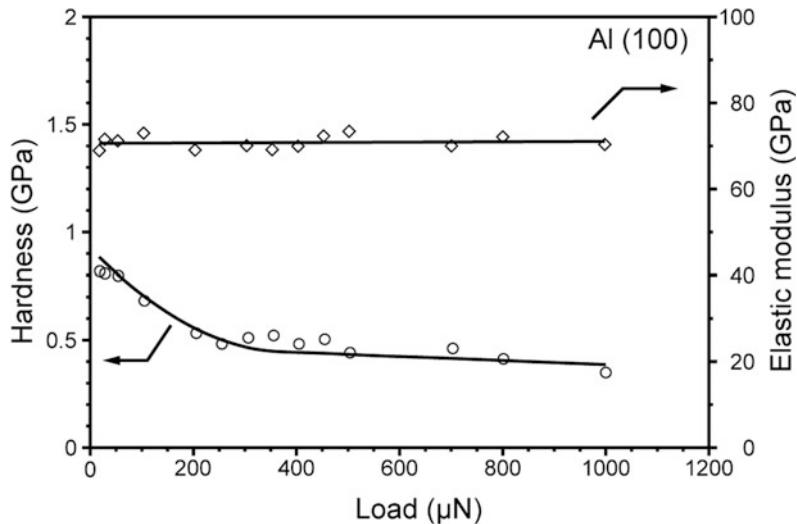


Fig. 5.31 Load-displacement curves for single-crystal Al(100) **a** at peak loads of 15, 25, 50, 100, 300, and 500 μN , and **b** during three successive loading/unloading cycles at a peak load of 100 μN . Arrows in (a) indicate the forward jump in the displacement (adapted from Kulkarni and Bhushan 1996b)

Fig. 5.32 Hardness and elastic modulus of elasticity as a function of load for single-crystal Al(100) (adapted from Kulkarni and Bhushan 1996b)



To demonstrate the creep effects further, the load-displacement curve for the lowest peak load of 500 μN held at 30 s also is shown as an inset, along with the data at no hold (0 s) for comparison. Again, creep is observed even at a hold period of 30 s. In this study, since indentation tests were performed at very low loads in ambient conditions, it is believed that creep of hydrated layers were being measured, rather than that of the bulk silicon. Hydrated layers are expected to exhibit viscoelastic properties. To study strain-rate sensitivity of silicon, experiments were conducted at two different rates of loading (Fig. 5.33b). Note that a change in the loading rate of a factor of about five results in a change in the load-displacement data. Creep effects and strain-rate sensitivity to single-crystal Al(111) and DLC coatings also were reported by Kulkarni and Bhushan (1996b, 1997).

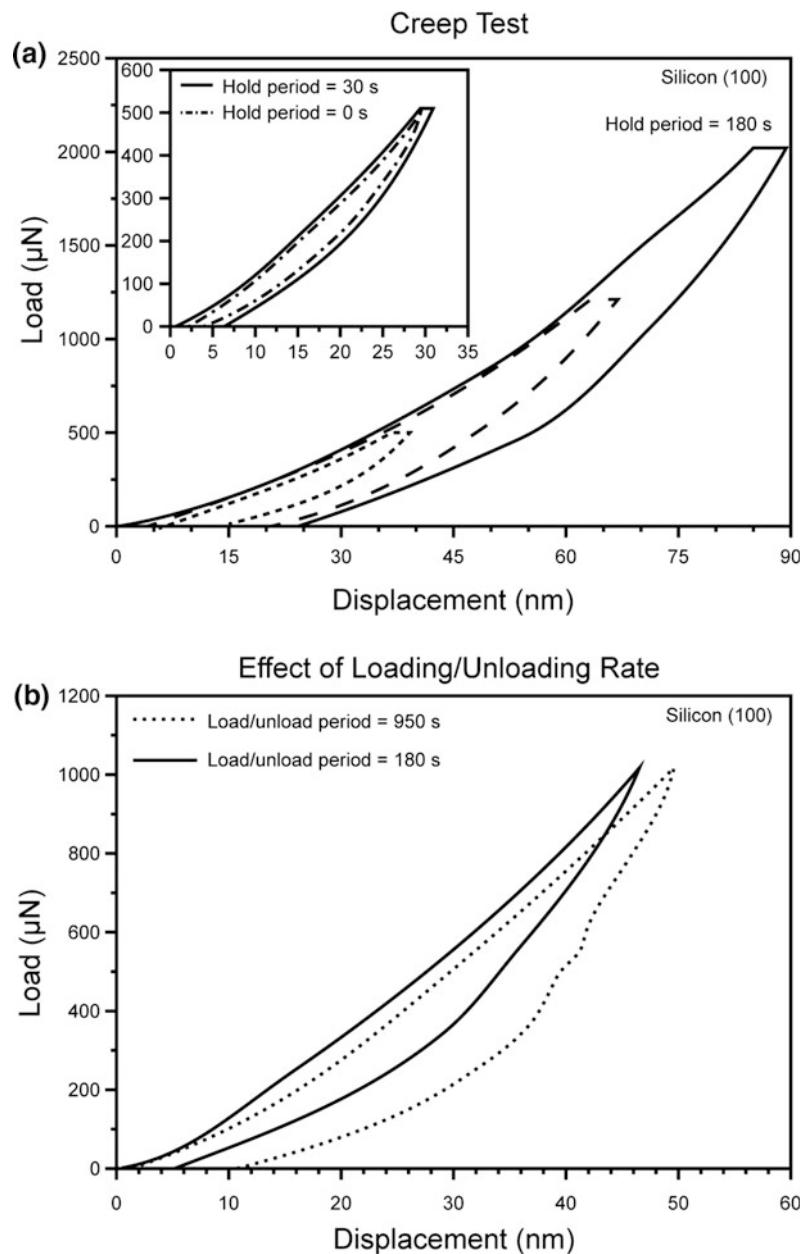
5.3.5 Continuous Stiffness Measurement

Continuous stiffness measurements (CSM) are used for continuous measurement of sample stiffness or compliance (related to elastic modulus) during indentation (Oliver and Pharr 1992; Bhushan 1999a; Li and Bhushan 2002a; Bhushan and Li 2003). A representative continuous stiffness data for single-crystal tungsten for

the indentation data in Fig. 5.19b are presented in Fig. 5.34a for fully elastic contact. Comparisons of these data show that the measured contact stiffness and, thus, the contact area does increase and decrease in the way that would be expected based on the loading history. The continuous stiffness data for indentation of single-crystal tungsten at 1.5 mN load (plastic contact, Fig. 5.19c) is shown in Fig. 5.34b for a fully plastic contact. It is seen that for each of the four unloadings, the contact stiffness changes immediately and continuously as the specimen is unloaded. Thus, the contact area, which varies in the same way as the contact stiffness, is not constant during the unloading of the plastic hardness impression, even during the initial stages of unloading.

The CSM method can be used to study graded materials and multilayered structures (Bhushan 1999a; Li and Bhushan 2002a; Bhushan and Li 2003). For a uniform material, from Eqs. 5.4a and 5.6, contact stiffness, S , is linearly proportional to contact depth, h_c . For a non-uniform material, a linear relationship between S and h_c does not exist. Therefore, the CSM technique can be used to study the mechanical properties of graded materials (such as magnetic tapes) and multilayered structures (such as magnetic rigid disks) by monitoring the change in contact stiffness, elastic modulus, and hardness as a function of contact depth (Li and Bhushan 2000a, b,

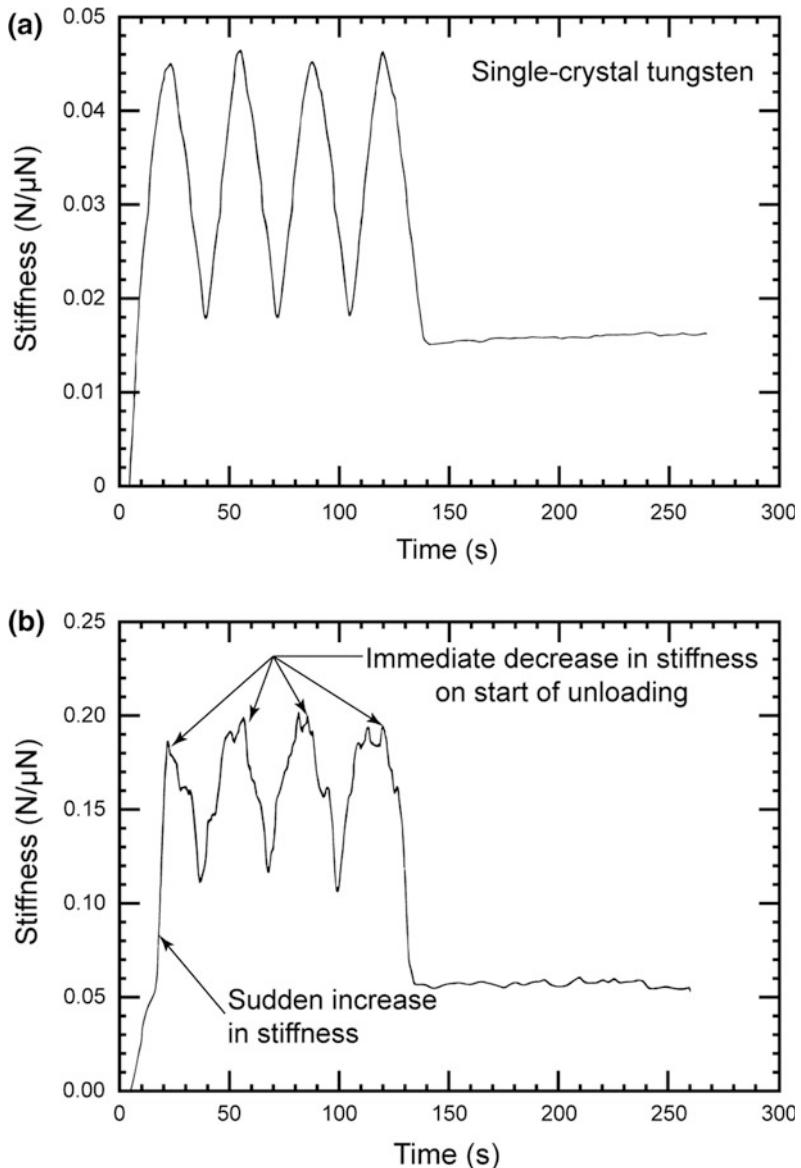
Fig. 5.33 **a** Creep behavior and **b** loading/unloading rate behavior of single-crystal Si (100) (adapted from Bhushan et al. 1996a)



2001a, b, 2002a; Li et al. 2001). The schematics of contact stiffness as a function of contact depth for indentations made on uniform and graded materials are shown in Fig. 5.35. The contact stiffness, elastic modulus, and hardness as a function of contact depth for a magnetic rigid disk and a magnetic tape with multilayered structures are shown in Fig. 5.36a, b, respectively. For the magnetic disk, from the variations

in contact stiffness, elastic modulus and hardness, one can distinguish one layer from another easily (Fig. 5.36a). In the case of the magnetic tape, observed continuous decrease in the rate of change of the contact stiffness suggests that the tape coating has graded properties (Fig. 5.36b). The elastic modulus values obtained at a shallow indentation depth were the same as those measured from tensile tests for the tape.

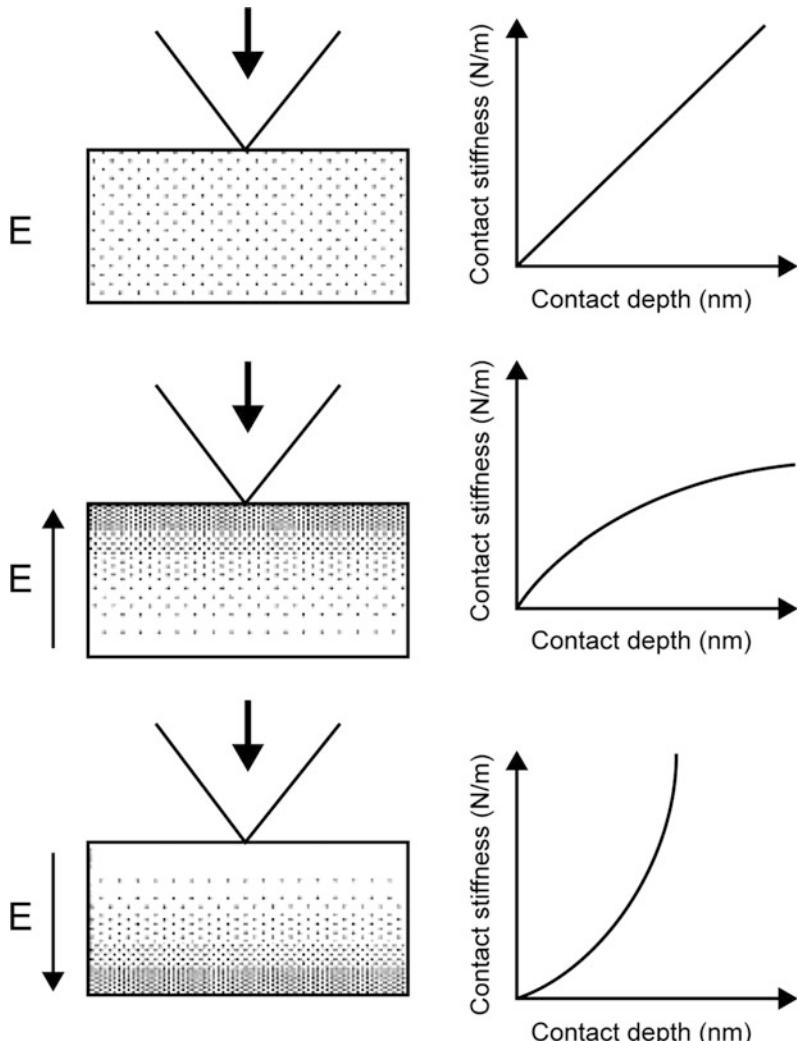
Fig. 5.34 Contact stiffness versus time **a** for a fully elastic contact and **b** for a fully plastic contact on an electropolished single-crystal tungsten measured with the continuous stiffness technique (adapted from Oliver and Pharr 1992)



The CSM technique also has been used to perform nanoscale indentation creep testing (Li and Bhushan 2000b). In an indentation creep test, a constant load is applied to the indenter and the change in indentation depth (displacement) is monitored as a function of time. Compared to conventional tensile creep tests, the CSM indentation creep experiments are particularly useful as they simulate creep resulting from asperity contacts in a sliding contact. The CSM technique has been used to study the creep

behavior of bulk materials (Syed Asif and Pethica 1997), graded materials, and multilayered solids (Li and Bhushan 200b, 2001a, b, 2002a; Li et al. 2001). Figure 5.37 shows the CSM indentation creep results for a magnetic tape. The tape exhibits an increase in indentation displacement and a decrease in mean stress with time, indicating that stress relaxation occurred during the hold segment. The contact stiffness remains nearly constant during the 600 s hold segment,

Fig. 5.35 Schematics of contact stiffness as a function of contact depth for indentations made on uniform and graded materials (adapted from Li and Bhushan 2002a)



indicating that the contact between the tip and the tape does not change significantly.

Nanoscale damage caused by fatigue is of critical importance to the reliability of ultra-thin protective overcoats and micro/nanostructures. The cyclic loading used in the CSM makes the technique useful for evaluation of nanofatigue, which will be described later.

5.3.6 Mechanical Properties of Nano-objects

Nano-objects can be described as a single material or composite having at least one dimension

that is between 1 and 100 nm (Bhushan et al. 2014; Maharaj and Bhushan 2015). They come in a variety of discrete geometries that include spheres, tubes, rods, wires, and pillars. Compared to their bulk material counterparts, many nano-objects exhibit enhanced mechanical, electrical, magnetic, chemical, and friction, and wear-reducing properties. This makes them attractive for use in many applications on the macro- to nanoscale.

Nano-objects are used in various applications that include drug delivery for cancer treatment, oil detection, contaminant removal, catalysis, and tribology on the macro- to nanoscale. As an example, nano-objects made of various materials

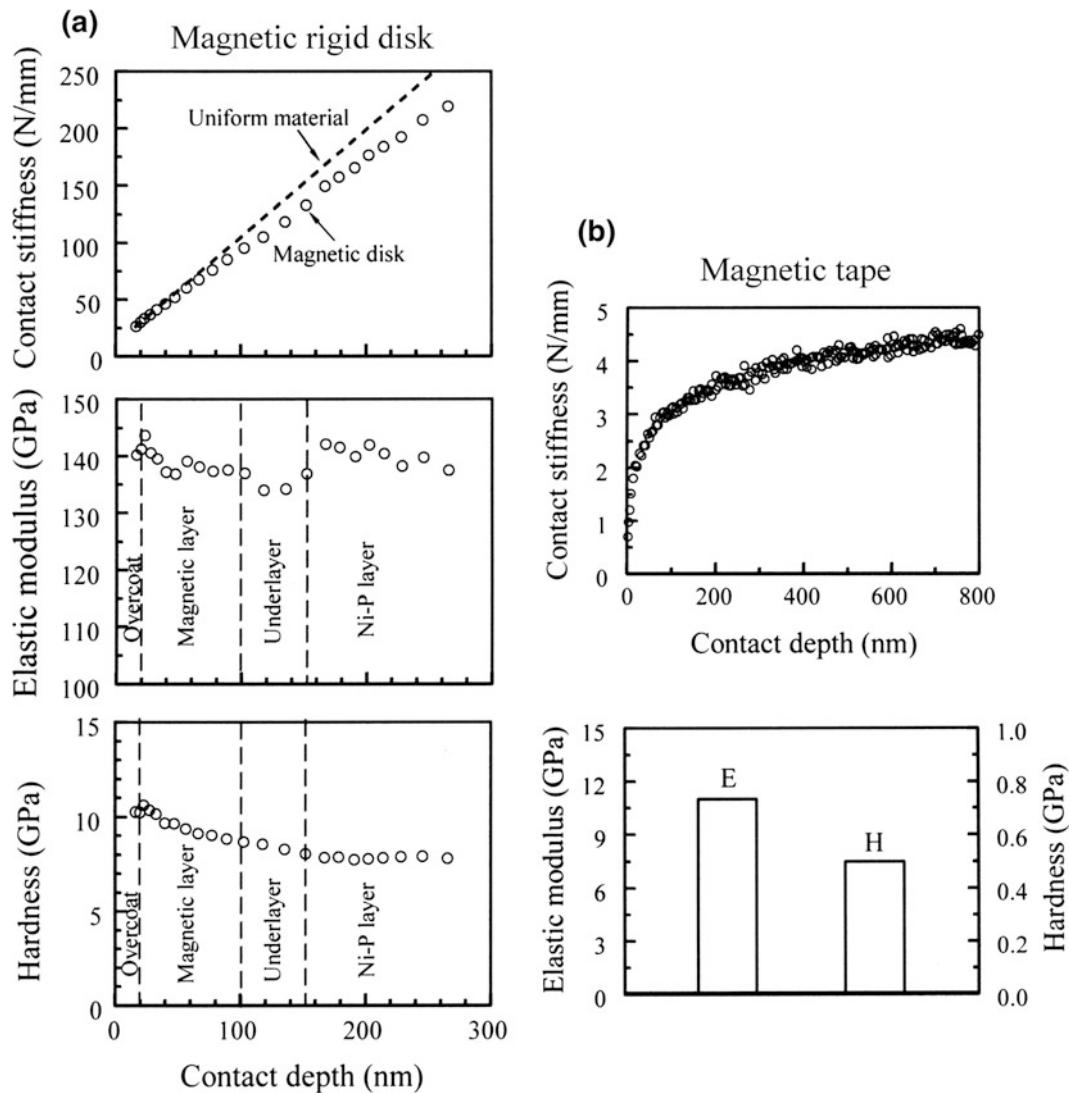


Fig. 5.36 **a** Contact stiffness, elastic modulus and hardness as a function of contact depth for a magnetic rigid disk with a multilayered structure, and **b** contact stiffness as a function of contact depth, elastic modulus, and hardness data at contact depth of 15 nm for magnetic tape (adapted from Li and Bhushan 2002a)

including gold (Au), graphene, iron oxide, polymer, and silica have been studied in targeted drug delivery for cancer treatment (Maharaj and Bhushan 2015; Bhushan 2017).

In tribological applications, nano-objects come into sliding contact with each other and the surfaces in which they are used. Contacts may occur locally on a nano-object or the entire nano-object may be compressed. To simulate local deformation of nano-objects, they may be

indented with a sharp indenter. For global deformation, the entire nano-object is compressed (Maharaj and Bhushan 2015).

Mechanical properties of nano-objects can be measured by using either a sharp nanoindenter tip to measure local deformation or a flat punch to measure global deformation (Fig. 5.38). Figure 5.39 shows TEM images of spherical gold nanoparticles (NPs) approximately 50 nm in diameter and gold nanorods (NRs) 50 nm in

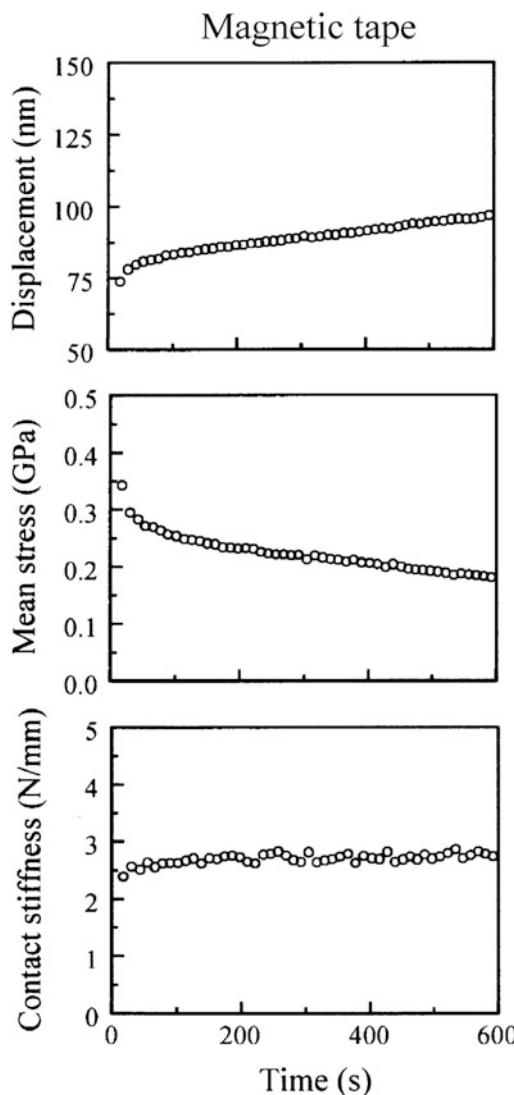


Fig. 5.37 Indentation displacement, mean stress, and contact stiffness as a function of time for magnetic tape (adapted from Li and Bhushan 2002a)

diameter and 200 nm in length for which nanoindentation data will be presented next (Maharaj and Bhushan 2015). Figure 5.40a shows a typical load-displacement curve for gold particles with approximately 500 nm in diameter. Pop-in events occur during nanoindentation and are shown by vertical arrows on the load-displacement curves. Pop-in events correspond to generation of new dislocations and multiplication of existing dislocations.

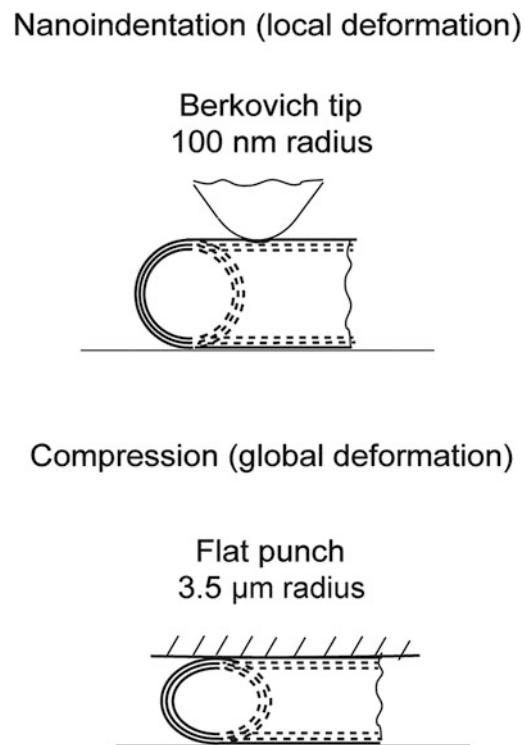


Fig. 5.38 Schematics of nanoindentation using a sharp tip to measure local deformation and compression using a flat punch to measure global deformation of nano-objects (adapted from Maharaj and Bhushan 2015)

Figure 5.40b shows nanoindentation hardness and elastic modulus as a function of contact depth. Scale effects in nanohardness can be observed.

Figure 5.41a shows typical load-displacement curves at a maximum load of 50 μ N with gold NPs and NRs with approximately 50 nm in diameter using a flat punch to study global deformation. To study recovery from deformation after the compression experiments, Fig. 5.41b shows the topography maps and 2-D profiles after 1 and 4 min of compression experiments. Two profiles were examined to observe whether there was recovery from the deformation. These profiles are essentially the same and suggest no recovery.

Micro/nanopillar structures represent a geometry of nanomaterials that can be used as components for various MEMS/NEMS devices (Palacio and Bhushan 2013; Bhushan 2017).

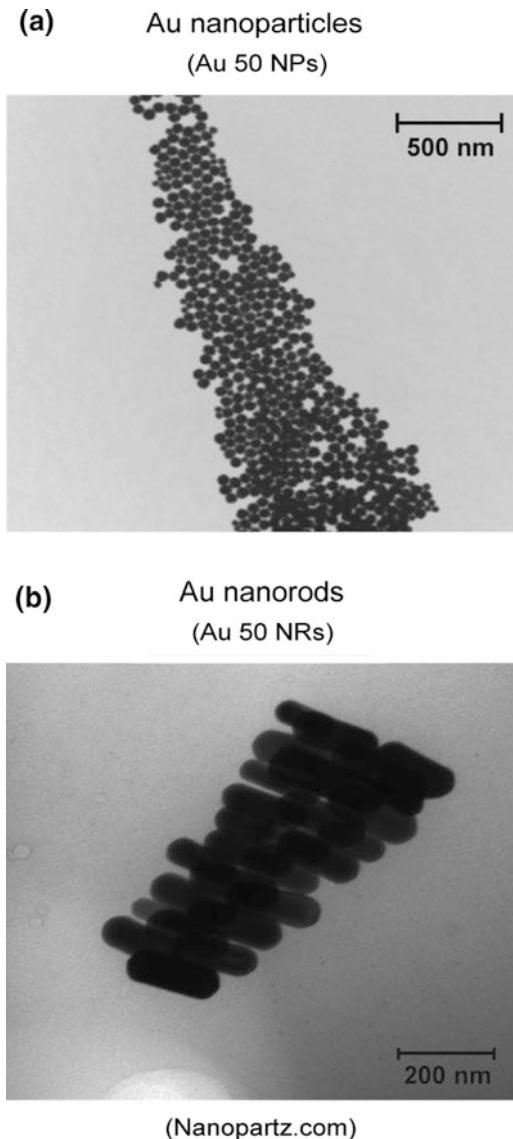


Fig. 5.39 TEM images of **a** spherical Au nanoparticles approximately 50 nm in diameter (Au 50 NPs) **b** Au nanorods 50 nm in diameter and 200 nm in length (Au 50 NRs) (adapted from Maharaj and Bhushan 2015)

Since submicron-sized pillar structures can be produced with well-defined geometries, they are used commonly to study scale effects on mechanical properties. Similar to nanoparticles, these structures have very small material volumes. As a result, the mechanical deformation of

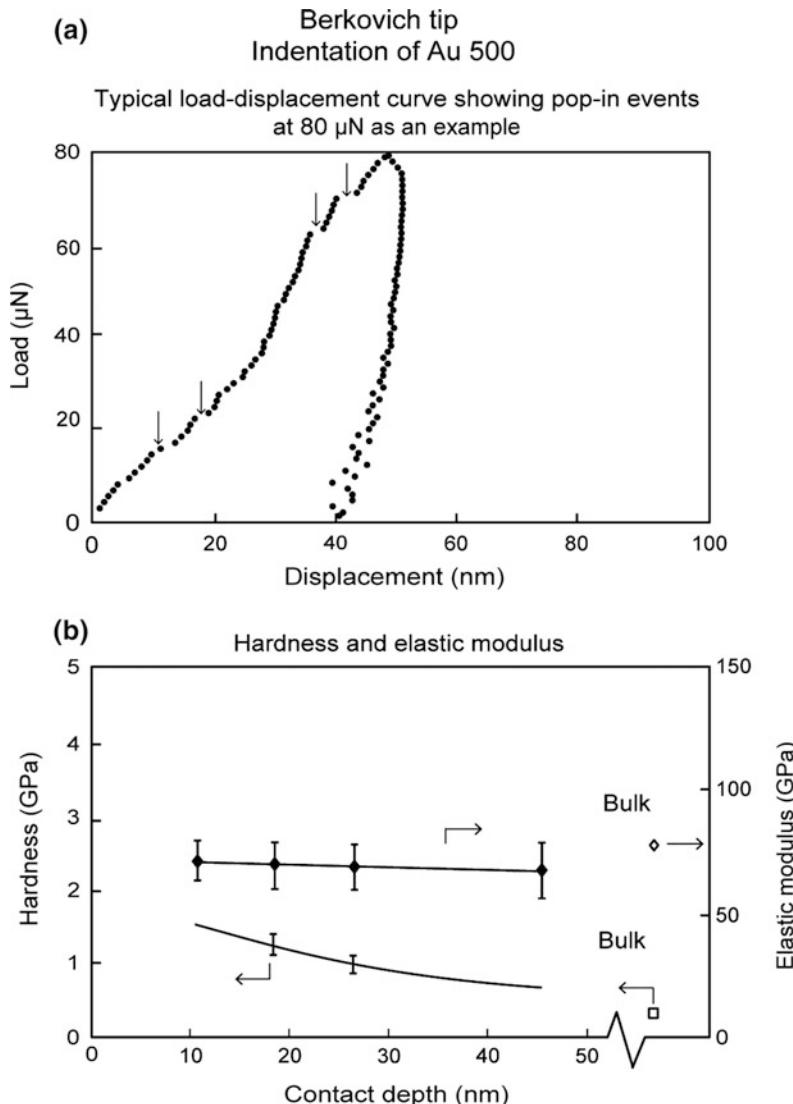
micro/nanopillars differs from their bulk counterparts.

These pillars are produced by various fabrication techniques including focused ion beam (FIB) micromachining. For deformation studies, these are loaded generally in compression using a flat punch. As an example, San Juan et al. (2009) performed compression studies on metal superalloy pillars composed of Cu-Al-Ni and with a diameter of 900 nm fabricated by FIB micromachining. Shape memory alloys are of scientific and technological interest as they can transform between the austenite phase and the martensite phase reversibly. The submicron diameter pillar such as the one shown in Fig. 5.42a was subjected to cyclic compression using a diamond spherically-conical tip. Two load-displacement curves taken after ten compression cycles are shown in Fig. 5.42b. At a load of about 95 μ N, an abrupt burst in displacement is observed as a result of the stress-induced martensitic transformation. During unloading, the reverse transformation is observed at about 45 μ N load, and continues gradually until a load of about 22 μ N is attained. The compression is completely elastic, as no residual deformation was observed. In this study, a size effect was observed such that the martensite phase has improved stability relative to its bulk form due to relaxation effects at the free surface. As a consequence, the reverse transformation takes place only when very low values of the stress are attained during the unloading. This is manifested by the large hysteresis seen in the load-displacement curve prior to elastic recovery.

5.4 Nanoscratch Technique for Adhesion Measurement and Nanoscratch/Nanowear Resistance

The nanoscratch technique is used commonly to measure relative adhesion strength of coated surfaces and to measure nanoscratch/nanowear resistance.

Fig. 5.40 **a** Typical load-displacement indentation curve of nanoparticles with 500 nm in diameter (Au 500 NPs) at a maximum load of 80 μ N with vertical arrows showing pop-in events, and **b** hardness and elastic modulus of Au 500 NPs as a function of contact depth represented by *bold square* and *diamond datum* points respectively with corresponding *open diamond* and *square datum* points representing bulk data (adapted from Maharaj and Bhushan 2015)



5.4.1 Adhesion Strength

Scratching a surface with a fingernail or a knife is probably one of the oldest methods for determining the adhesion of paints and other coatings. In 1822, Friedrich Mohs used resistance to scratch as a measure of hardness. Scratch tests to measure adhesion of films was introduced first by Heavens in 1950. A smoothly round, chrome-steel stylus with a tungsten carbide or Rockwell C diamond tip (in the form of 120° cone with a hemispherical tip of 200 μ m radius) (Perry 1983; Mehrotra and Quinto 1985; Valli

1986) or Vickers pyramidal indenter (Burnett and Rickerby 1987a; Bull and Rickerby 1990; Venkataraman et al. 1992) for macro- and microscratching a conical diamond indenter (with 1 or 5 μ m of tip radius and 60° of included angle) for nanoscratching (Wu 1991; Bhushan et al. 1995) is drawn across the coating surface. A normal load is applied to the scratch tip and is gradually increased during scratching until the coating is completely removed. The minimum or critical load at which the coating is detached or completely removed is used as a measure of adhesion (Benjamin and Weaver 1960; Campbell

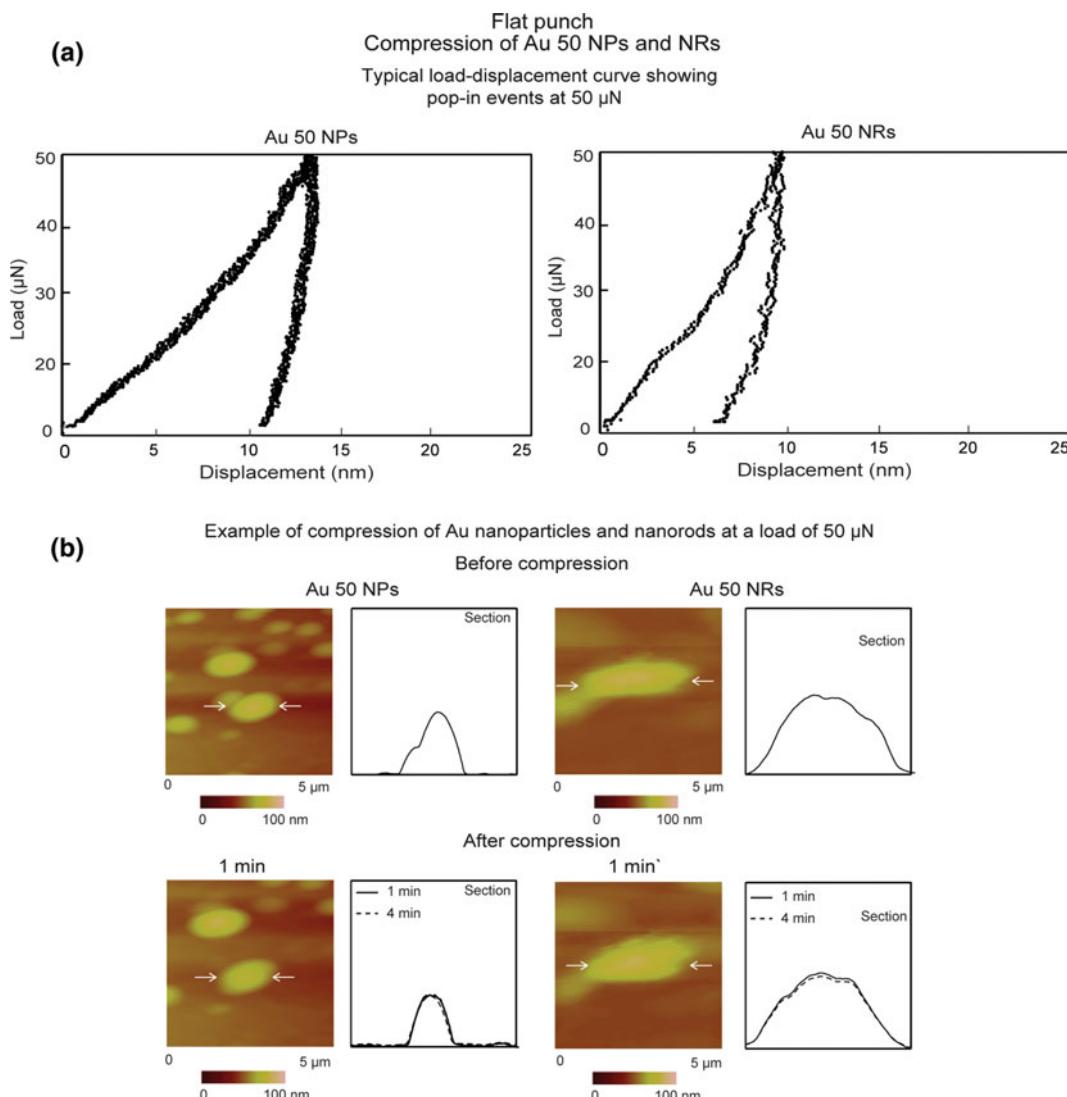


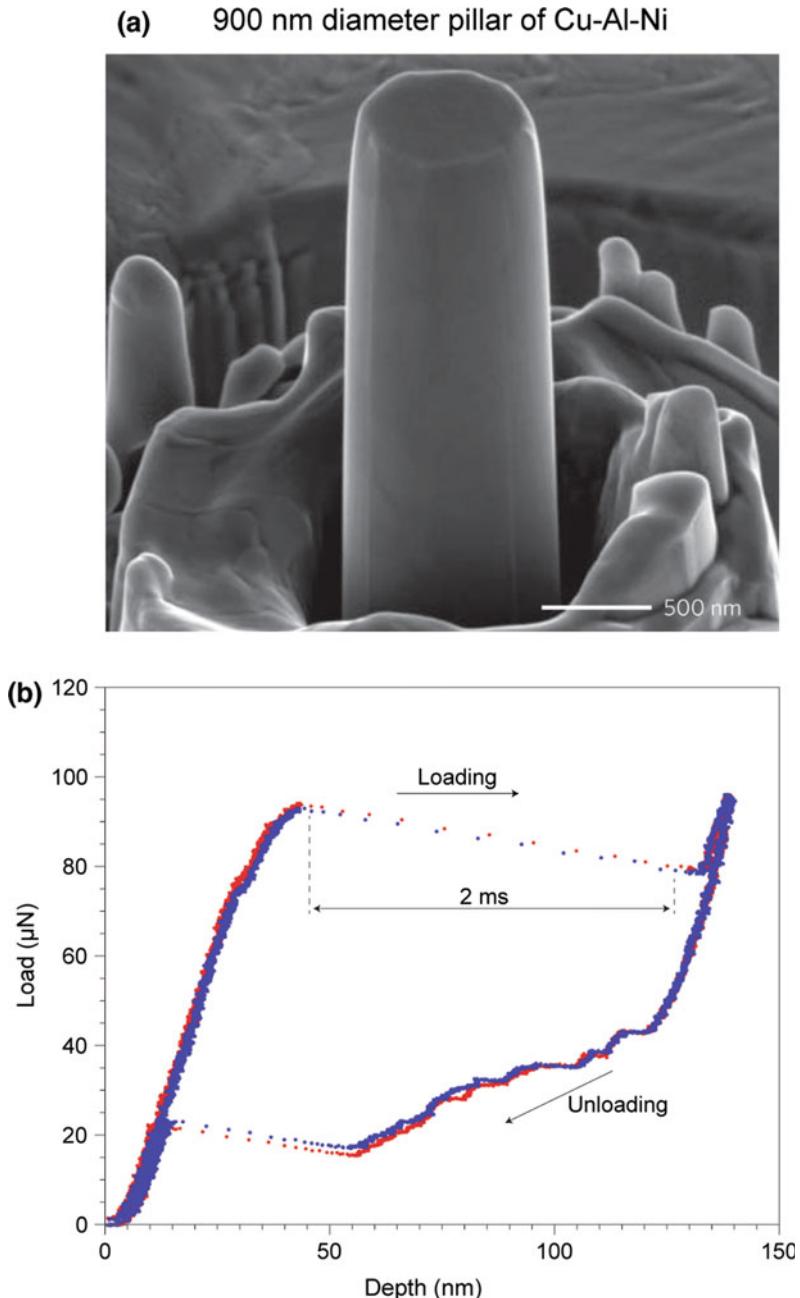
Fig. 5.41 **a** Typical load-displacement curve at a maximum load of 50 μN with Au nanoparticles with 50 nm in diameter (Au 50 NPs) and Au nanorods with 50 nm in diameter (Au 50 NRs) using a flat punch, and **b** topography maps and 2-D profiles at sections shown by the horizontal arrows before compression (*first row*) and after compression (*second row*). The *second row* shows topography maps 1 min. after compression and the *solid lines* and *dashed lines* show the 2-D profiles after 1 and 4 min respectively. The *horizontal white arrows* indicate the nano-objects of interest along with the section on which the profiles were taken (adapted from Maharaj and Bhushan 2015)

1970; Greene et al. 1974; Ahn et al. 1978; Mittal 1978; Laugier 1981; Mehrotra and Quinto 1985; Valli 1986; Burnett and Rickerby 1987a; Sekler et al. 1988; Wu 1991; Bull and Rickerby 1990; Bhushan et al. 1995; Gupta and Bhushan 1995a, b; Patton and Bhushan 1996; Bhushan and Li 1997; Li and Bhushan 1999a, b, c). It is a most

commonly used technique to measure adhesion of hard coatings with strong interfacial adhesion (>70 MPa).

For a scratch geometry produced using a stylus with a spherical tip, shown in Fig. 5.43, surface hardness H is given by

Fig. 5.42 **a** SEM image of a 900-nm diameter pillar composed of Cu-Al-Ni alloy, and **b** load-displacement curves from two consecutive nanocompression experiments (adapted from San Juan et al. 2009)



$$H = \frac{W_{cr}}{\pi a^2} \quad (5.21)$$

$$\tau = H \tan \theta \quad (5.22a)$$

and adhesion strength τ is given by (Benjamin and Weaver 1960)

$$= \frac{W_{cr}}{\pi a^2} \left[\frac{a}{(R^2 - a^2)^{1/2}} \right]$$

or

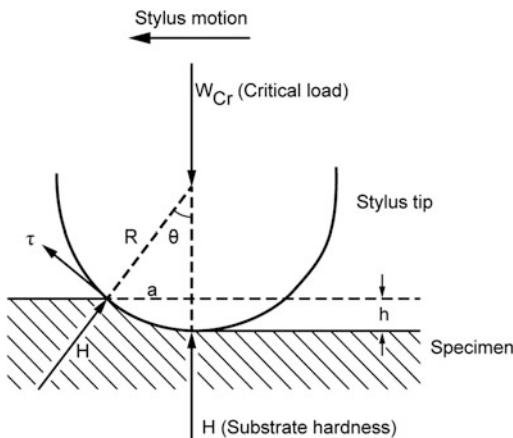


Fig. 5.43 Geometry of a scratch produced using a stylus with a spherical tip

$$\tau = \frac{W_{cr}}{\pi a R} \text{ if } R \gg a \quad (5.22b)$$

where W_{cr} is the critical normal load, a is the contact radius, and R is the stylus radius.

Burnett and Rickerby (1987a) and Bull and Rickerby (1990) analyzed the scratch test of a coated sample in terms of three contributions: (1) a plowing contribution, which will depend on the indentation stress field and the effective flow stress in the surface region, (2) an adhesive friction contribution due to interactions at the indenter-sample interface, and (3) an internal stress contribution, since any internal stress will oppose the passage of the indenter through the surface, thereby effectively modifying the surface flow stress. They derived a relationship between the critical normal load W_{cr} and the work of adhesion W_{ad}

$$W_{cr} = \frac{\pi a^2}{2} \left(\frac{2EW_{ad}}{t} \right)^{1/2} \quad (5.23)$$

where E is the elastic modulus and t is the coating thickness. Plotting of W_{cr} as a function of $a^2/t^{1/2}$ should give a straight line of the slope $\pi(2EW_{ad})^{1/2}/2$ from which W_{ad} can be calculated. Bull and Rickerby suggested that either the line slope (interface toughness) or W_{ad} could be used as a measure of adhesion.

An accurate determination of critical load W_{cr} sometimes is difficult. Several techniques have been used to obtain the critical load including (1) microscopic observation (optical or SEM) during the test, (2) chemical analysis of the bottom of the scratch channel (with electron microprobes) and (3) acoustic emission (Perry 1983; Valli 1986; Sekler et al. 1988; Wu 1991). In some instruments, tangential (or friction) force is measured during scratching to obtain the critical load (Jacobson et al. 1983; Valli 1986; Wu 1991; Bhushan et al. 1995, 1997; Bhushan and Li 1997; Gupta and Bhushan 1995a, b; Patton and Bhushan 1996; Li and Bhushan 1998b, 1999a, b, c). The AE and friction force techniques have been reported to be very sensitive in determining critical load. AE and friction force start to increase as soon as cracks begin to form perpendicular to the direction of the moving stylus.

5.4.2 Nanoscratch/Nanowear Resistance

In a nanoscratch test during scratching, normal load applied to the scratch tip is increased gradually until the material is damaged. Friction force generally is measured during the scratch test (Bhushan and Gupta 1995; Bhushan et al. 1996b; Palacio and Bhushan 2010; Kumar and Bhushan 2015; Cho and Bhushan 2016). After the scratch test, the morphology of the scratch region including debris is observed in an SEM. Based on the combination of the changes in the friction force as a function of normal load and SEM observations, the critical load is determined and the deformation mode is identified. Any damage to the material surface as a result of scratching at a critical ramp-up load results in an abrupt or gradual increase in friction.

Deformation modes can be identified by microscopic examination of scratch marks. The material may deform either by plastic deformation or fracture. Ductile materials (all metals) deform primarily by plastic deformation, resulting in significant plowing during scratching. The

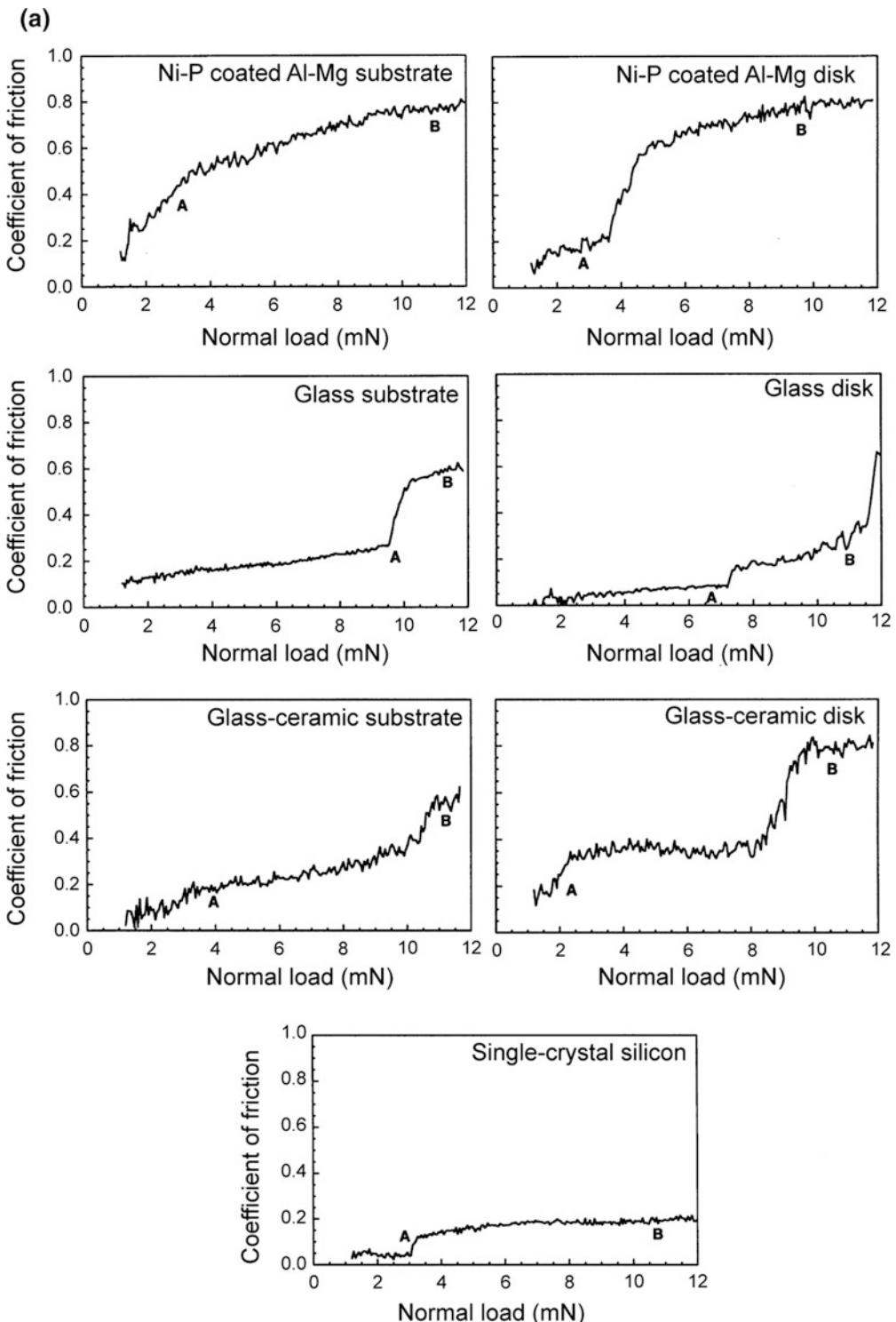


Fig. 5.44 **a** Coefficient of friction profiles as a function of normal load for 500- μm long scratches made using a diamond tip with 1- μm tip radius, at a normal load ranging from 1 to 12 mN on Ni-P coated Al-Mg substrate and various ceramic substrates, as well as corresponding magnetic disks, and single-crystal silicon, and **b** SEM images of two regions on 500- μm long scratches made at 1–12 mN load on various ceramic substrates and corresponding magnetic disks and single-crystal silicon. The scratching direction was from *left to right*. The *upper images* in the sets of two images for each sample correspond to a location or normal load where the friction increased abruptly and/or damage began to occur. These are the points indicated by ‘A.’ The *lower images* correspond to a location close to the end of the scratch (~ 11 mN). These are the points indicated by ‘B’ (adapted from Bhushan and Gupta 1995)

width and depth of the tracks produced increase with an increase in the normal load. Plowing results in a continuous increase in the coefficient of friction with an increase in the normal load during scratching and produces generally ribbon-like or curly debris. To compare, brittle materials deform primarily by brittle fracture with some plastic deformation. In the brittle fracture mode, the coefficient of friction increases very little until a critical load is reached at which the material fails catastrophically and produces fine debris, which is rounded, and the coefficient of friction increases rapidly above the critical load.

The profiles for the coefficient of friction as a function of normal load for scratches made on Ni-P coated Al-Mg substrate and various ceramic substrates, as well as corresponding magnetic disks (substrates coated with 75-nm-thick sputtered Co-Pt-Ni magnetic film and 20-nm thick sputtered DLC film) are compared in Fig. 5.44a (Bhushan and Gupta 1995). The figure also includes the friction force profile for a single-crystal silicon substrate for comparison. The SEM images of two regions of 500- μm -long scratches made at 1–12 mN normal load on various samples are compared in Fig. 5.44b (Bhushan and Gupta 1995). The upper images in the sets of two images for each sample correspond to a region where friction increased abruptly. These are the points indicated by “A” in Fig. 5.44b. The lower images in each set correspond to the region that is very close to the end of a scratch. These are the points indicated by “B”. The extent of a damage in a scratch is estimated by the width and depth of the scratch and by the amount of debris generated toward the end of the scratch.

Single-crystal silicon exhibited the lowest friction with little plowing at a low load and

cracking at higher loads. This observation suggests that scratching of the silicon took place primarily by brittle deformation. In the case of the Ni-P-coated Al-Mg substrate, friction increase was continuous from the beginning of the scratching, Fig. 5.44a. SEM images of the Ni-P-coated Al-Mg substrate presented in Fig. 5.44b, show the material removal occurred by plowing with formation of curly ductile chips. It is evident that scratching took place primarily by plastic deformation typical of ductile materials. Plowing is responsible for the continuous increase in the friction for this substrate. Glass and glass-ceramic substrates and corresponding disks and the Ni-P coated Al-Mg disk exhibited relatively low friction with a sudden increase at higher load. The glass substrate exhibited the lowest friction, followed by the glass-ceramic substrate. In the case of the Ni-P coated Al-Mg disk, the load at which friction increased was lower than that of the glass and glass-ceramic substrates. SEM images of these samples exhibit plowing in addition to the formation of fine debris. There is no evidence of cracking of ceramic substrates or the ceramic overcoats used in all disks at magnifications as high as 50,000 \times . Glass is chemically strengthened in order to produce significant compressive stresses in the glass surface. Glass-ceramic consists of fine-grained polycrystalline material in a glass matrix. The chemical strengthening and the crystals add to the fracture toughness of the material. Thus, both ceramic substrate materials are expected to deform with ductile and brittle deformation modes. Ductile deformation results in plowing, whereas brittle deformation aids in debris generation. Lower values of the coefficient of friction before a sudden increase, as compared to the Ni-P-coated Al-Mg substrate, suggest that brittle fracture contributes to overall deformation.

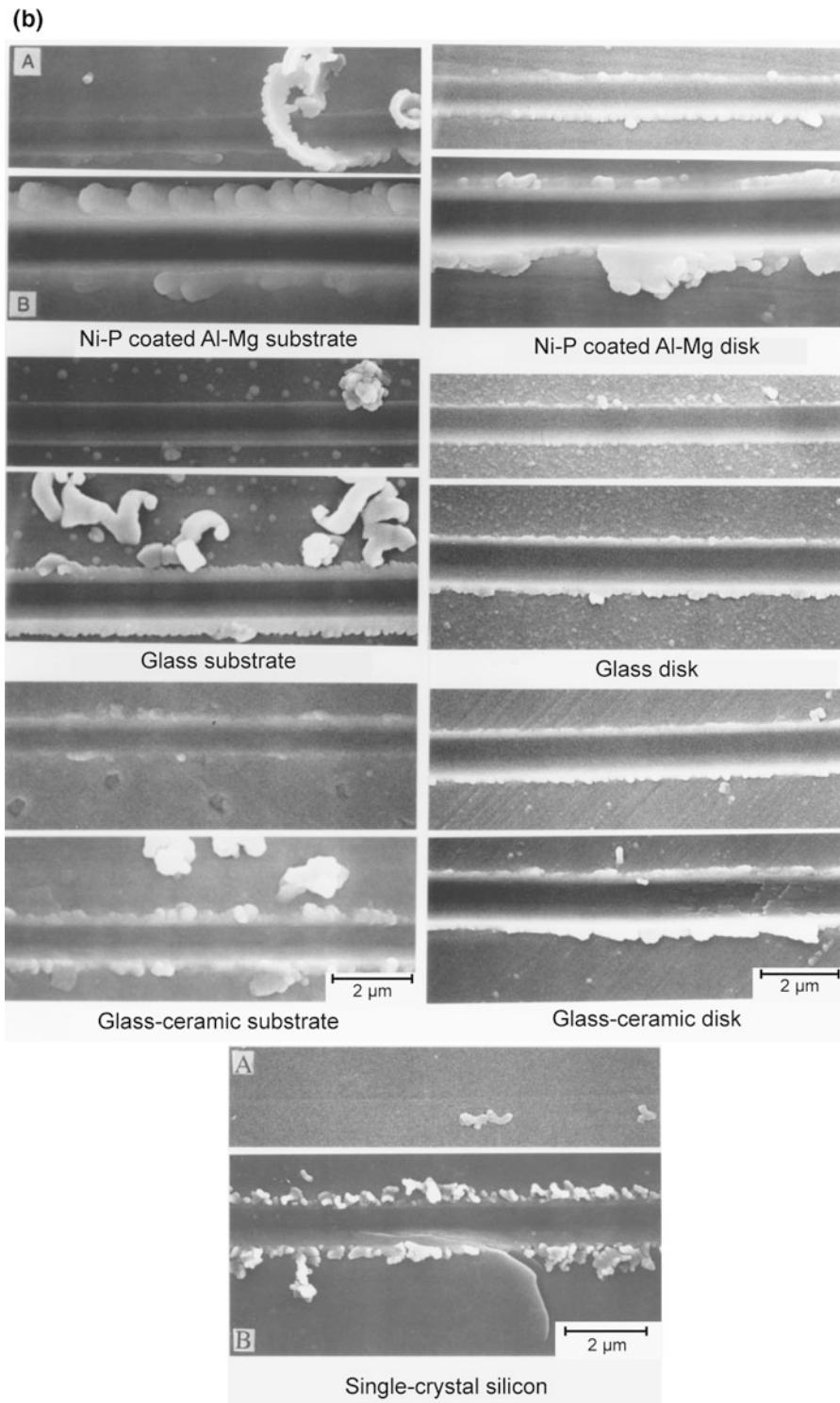
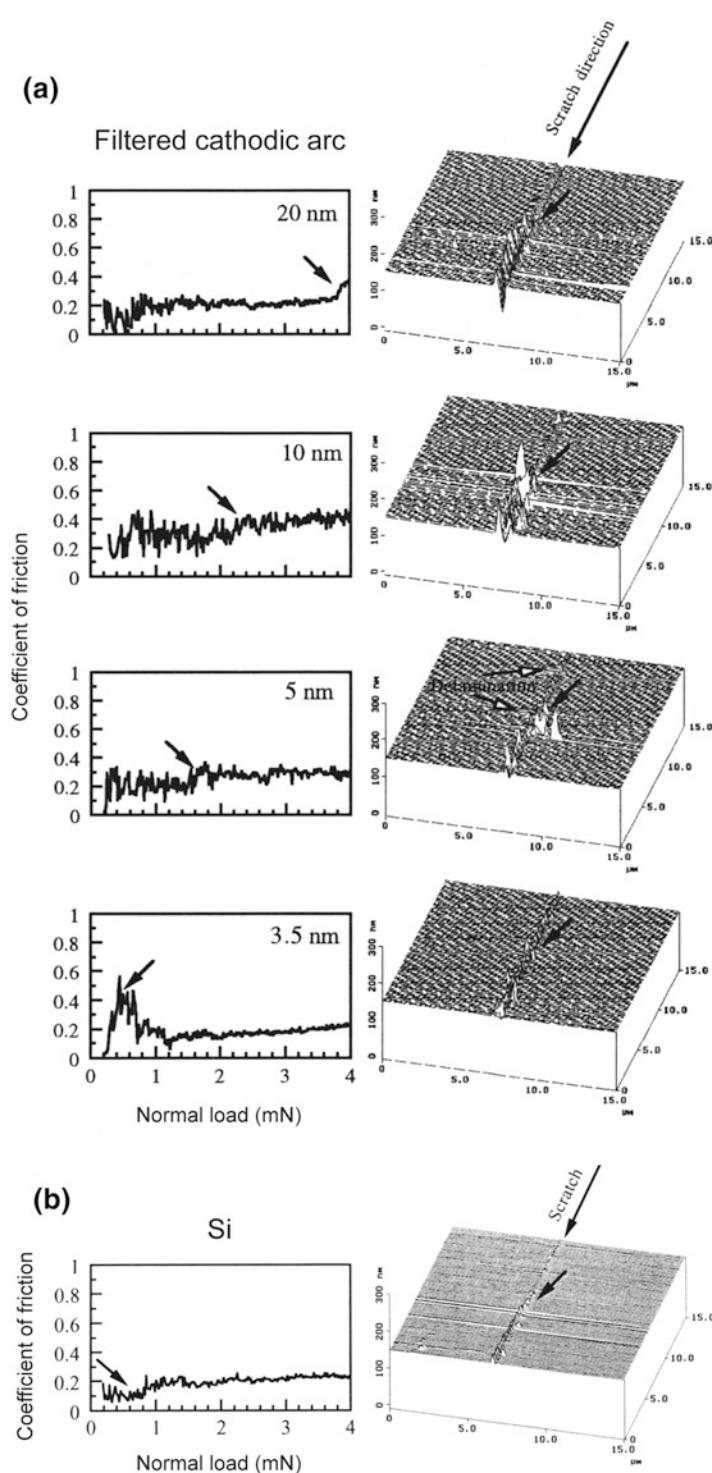


Fig. 5.44 (continued)

Fig. 5.45 Coefficient of friction profiles as a function of increasing normal load, three dimensional AFM images of regions over scratches at the respective critical loads (indicated by the arrows in the friction profiles and AFM images), made on the filtered cathodic arc deposited DLC coating of different thicknesses and an uncoated Si substrate (adapted from Li and Bhushan 1999a)



Hard overcoats generally consist of significant compressive residual stresses. It is these compressive stresses that allow ductile deformation with little cracking. We further note that a sudden increase in the coefficient of friction for ceramic substrates and for all disks at some load results from significant damage to the bulk material or to the coating surface (Fig. 5.44b).

Based on the friction data, the width and depth of scratches, the amount of debris generated, and scratch morphology, glass substrates and corresponding disks exhibit a lower coefficient of friction against a diamond tip and a superior resistance to scratch, followed by glass-ceramic substrates and corresponding disks. This example clearly suggests that deformation modes and critical load to failure can be identified using the scratch technique. The critical load to failure is a measure of scratch/wear resistance and adhesion in the case of coated surfaces.

Bhushan et al. (1995, 1997), Bhushan and Li (1997), Gupta and Bhushan (1995a, b), Patton and Bhushan (1996), and Li and Bhushan (1998b, 1999a, b, c) have used the scratch technique to study adhesion and scratch/wear resistance (a measure of mechanical durability) of various ceramic films. Scratch tests conducted with a sharp diamond tip simulate a sharp asperity contact. Bhushan and coworkers have also conducted accelerated friction and wear (ball-on-coated disk) and functional tests and have found a good correlation between the scratch resistance and wear resistance measured using accelerated tests (Bhushan et al. 1995, 1997; Gupta and Bhushan 1995b; Patton and Bhushan 1996; Li and Bhushan 1998b, 1999a, b) and functional tests (Patton and Bhushan 1996; Bhushan and Patton 1996; Bhushan et al. 1997). Based on this work, scratch tests can be successfully used to screen materials and coatings for wear applications.

Li and Bhushan (1999a) conducted scratch tests on ultrathin DLC coatings of thicknesses ranging from 20 down to 3.5 nm deposited by FCA technique on Si(100) substrates. Figure 5.45 shows the profiles of the coefficient of friction as a function of increasing normal load,

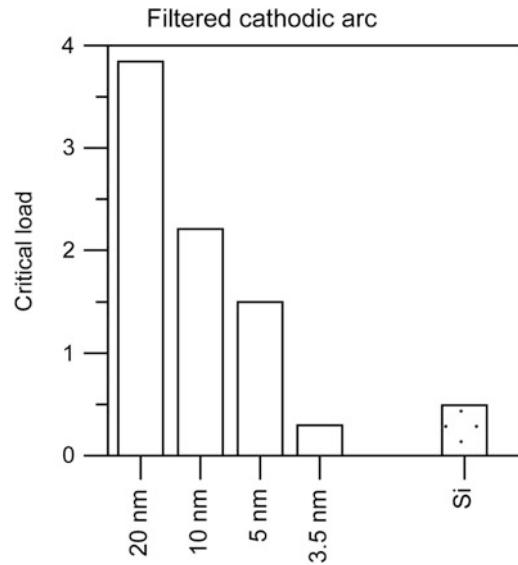
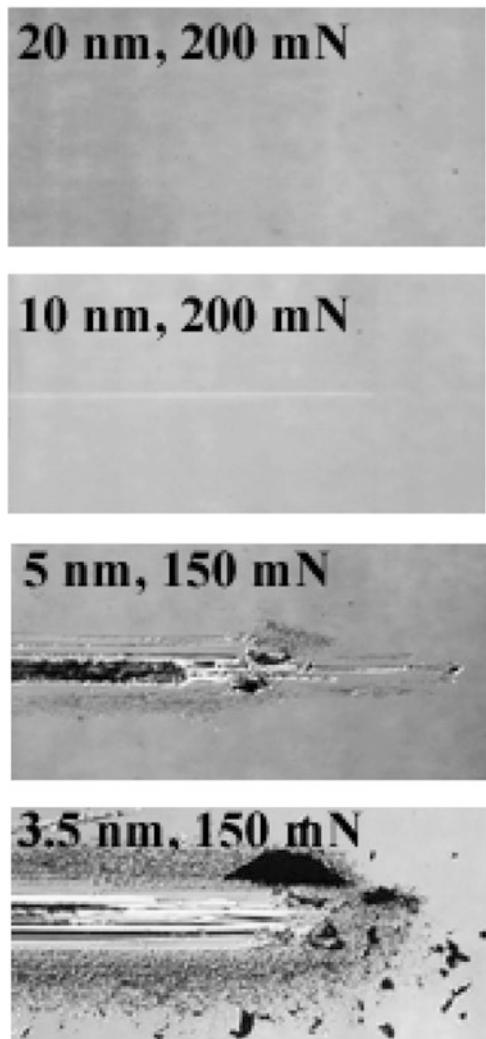


Fig. 5.46 Critical loads estimated from the friction profiles and AFM images for the filtered cathodic arc deposited DLC coatings of different thicknesses and an uncoated Si substrate (adapted from Li and Bhushan 1999a)

three-dimensional line AFM images of regions over scratches at the respective critical loads (indicated by the arrows in the friction profiles and AFM images) made on coatings of different thicknesses, and an uncoated Si substrate. From the AFM images, the debris can be seen easily, and the appearance of the first debris is indicated by the arrows, which correspond to critical loads. A well-defined critical load is observed for each coating. The AFM images clearly show that below the critical loads, the coatings were plowed by the scratch tip, associated with the plastic flow of materials. At and after the critical loads, debris (chips) or buckling were observed on the sides of scratches. Delamination or buckling can be observed around or after the critical loads. This suggests that the damage of these coatings starts from delamination and buckling. For the 3.5 and 5 nm thick FCA coatings before the critical loads, small debris is observed on the sides of scratches. This suggests that the thinner FCA coatings may be more brittle and more easily broken. Also, the 3.5 and 5 nm thick FCA coatings show more delamination and buckling

(a) Filtered cathodic arc



events than the other coatings. This indicates that the thinner FCA coatings have higher interfacial and residual stresses than the other coatings. The critical loads increase with increasing coating thickness. This indicates that the critical load is determined not only by the adhesive strength to the substrate but also by the coating thickness. It can be seen that the uncoated Si substrate exhibits a lower coefficient of friction of about 0.1 before the critical load. The AFM image shows that the uncoated Si substrate was damaged by plowing, associated with the plastic flow of materials. At and after the critical load, small and uniform debris is observed, and the amount of debris increases with increasing normal load.

The critical loads estimated from the friction profiles and AFM images for coatings of different thicknesses and an uncoated Si substrate are compared in Fig. 5.46. The thicker the coating, the higher the critical load. At 3.5 nm, the critical loads drop to about 0.4 mN, which is about the same critical load of the uncoated Si substrate.

The optical images of wear tracks and debris formed on all samples when slid against a 3-mm diameter sapphire ball after sliding distance of 5 m are compared in Fig. 5.47. The normal load used for the 20 and 10 nm thick coatings was 200 mN and the normal load used for the 5 and 3.5 nm thick coatings and Si substrate was 150 mN. No wear track and debris were found on the 20 nm thick coating. The low wear resistance of the 5 nm thick FCA coating is in good agreement with the low scratch critical load, which may be due to the higher interfacial and residual stresses as well as brittleness of the coating. At 3.5 nm, the FCA coating provides no wear resistance, failing instantly like the uncoated Si substrate. Large block-like debris is observed on the sides of the wear track of the coating. This indicates that large region delamination and buckling occurred during sliding, resulting in large block-like debris.

Fig. 5.47 Optical images of wear tracks and debris formed on all filtered cathodic arc deposited coatings of different thicknesses and an uncoated Si substrate when slid against a sapphire ball after sliding distance of 5 m (adapted from Li and Bhushan 1999a)

This example clearly suggests that scratch resistance is a powerful way of screening materials and the data correlates well with the wear performance (Li and Bhushan 1999a).

5.5 Nanoindentation for Adhesion Strength and Residual Stress Measurements of Thin Films

Adhesion describes the sticking together of two materials. Adhesion strength, in a practical sense, is the stress required to remove a coating from a substrate. Indentation and scratch (described earlier) on the micro- and nanoscales are the two commonly-used techniques to measure adhesion of thin hard films with good adhesion to the substrate (>70 MPa) (Campbell 1970; Mittal 1978; Blau and Lawn 1986; Bhushan 1987; Bhushan and Gupta 1997; Bhushan 1999a; Bhushan and Li 2003).

Nearly all coatings, by whatever means they are produced, and surface layers of treated parts are found to be in a state of residual (intrinsic or internal) stress. These are elastic stresses that exist in the absence of external forces and are produced through the differential action of plastic flow, thermal contraction, and/or changes in volume created by phase transformation. Microindentation and nanoindentation techniques are also used to measure residual stresses (Bhushan and Gupta 1997).

In this section, we describe adhesion strength measurements and residual stress measurements using nanoindentation.

5.5.1 Adhesion Strength Measurements

In the indentation test method at low loads, the coating deforms with the substrate. However, if the load is sufficiently high, a lateral crack is initiated and propagated along the coating-substrate interface. The lateral crack length increases with the indentation load. The minimum load at which the coating fracture is observed is called the *critical load* and is used as

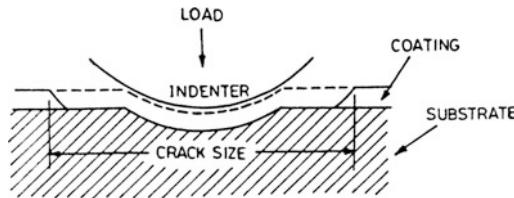


Fig. 5.48 Schematic illustration of the indentation method for adhesion measurement (adapted from Bhushan and Li 2003)

the measure of coating adhesion (Fig. 5.48) (Bhushan 1999a; Bhushan and Li 2003). For relatively thick films, the indentation is generally made using a Brinell hardness tester with a diamond sphere of 20 µm radius (Tangena and Hurkx 1986), Rockwell hardness tester with a Rockwell C 120° cone with a tip radius of 200 µm (Mehrotra and Quinto 1985) or a Vickers pyramidal indenter (Chiang et al. 1981; Lin et al. 1990; Alba et al. 1993). However, for extremely thin films, a Berkovich indenter (Stone et al. 1988) or a conical diamond indenter with a tip radius of 5 µm and 30° of included angle (Bhushan et al. 1995) is used in a nanoindenter.

It should be noted that the measured critical load W_{cr} is a function of hardness and fracture toughness, in addition to the adhesion of coatings. Chiang et al. (1981) have related the measured crack length during indentation, the applied load, and the critical load (at which coating fracture is observed) to the fracture toughness of the substrate-coating interface. A semi-analytical relationship derived between the measured crack length c and the applied load W ,

$$c = \alpha \left(1 - \frac{W_{cr}}{W} \right)^{1/2} W^{1/4} \quad (5.24)$$

where $\alpha^2 = \frac{\alpha_1 t_c^{3/2} H^{1/2}}{(K_{Ic})_{interface}}$, α_1 is a numerical constant, t_c is the coating thickness, H is the mean hardness, and $(K_{Ic})_{interface}$ is the fracture toughness of the substrate-coating interface. Mehrotra and Quinto (1985) used this analysis to calculate fracture toughness of the interface.

Bhushan (1987) estimated adhesion of composites by measuring the magnitude of shear

(friction) stresses at fiber/matrix interfaces in composites. Marshall and Oliver (1987) used a Berkovich indenter to push on the end of an individual fiber, and measured the resulting displacement of the surface of the fiber below the matrix surface (due to sliding). The shear stress was calculated from the force-displacement relation obtained by analysis of the frictional sliding. The force and displacement measurements were obtained only at the peak of the load cycle, and the sliding analysis was based on sliding at constant shear resistance at the interface. These experiments provided measurements of average shear stresses at individual fibers.

5.5.2 Residual Stress Measurements

Indentation measurements, similar to those used to determine the hardness and elastic modulus of a film, can be used to measure the residual stresses in it, also (Bhushan 1999a; Bhushan and Li 2003). When a compressive force on a biaxially stressed film during indentation is applied in a direction perpendicular to the film, yielding will occur at a smaller applied compressive force while a film is stressed in biaxial tension as compared with the unstressed film. Thus, the biaxial tension decreases hardness and the biaxial compression increases hardness (Swain et al. 1977; Vitovec 1986). LaFontaine et al. (1991) used the nanoindentation technique to measure the effect of residual stresses on the hardness of thin films. For samples that do not undergo large structural changes, changes in hardness with time reflect a change in residual stress in the film (LaFontaine et al. 1991). Bolshakov et al. (1996) and Tsui et al. (1996) studied the effect of residual stresses on hardness and elastic modulus. Based on indentation measurements and finite element analysis of the indentation process, they reported that, once pile up is accounted for, the residual stresses have little effect on hardness and elastic modulus.

Tsukamoto et al. (1987) measured the deflection at the center of the bent beam (bent as a result of residual stresses in the film) by pressing the beam flat with a nanoindenter. The

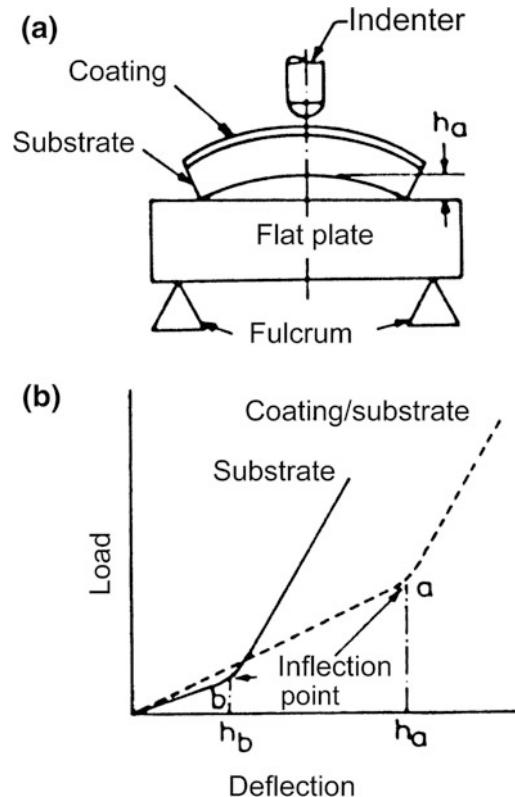


Fig. 5.49 **a** Schematic diagram of the deflection measurement of a bent beam using a nanoindenter, and **b** load-deflection curve for a warped composite-beam substrate

bent beam is placed on a flat glass surface supported by two fulcrums, and a load-deflection curve is generated (Fig. 5.49). The distance h_a can be estimated from the inflection point in the curve. Because of the limited flatness of most substrates, the film is removed from the substrate, then the initial deflection is measured. The true deflection resulting from residual stresses in the film equal to $h_a - h_b$. The curvature ($1/R$) of the substrate can then be calculated by the geometric relationship,

$$R = \frac{L^2}{8(h_a - h_b)} \quad (5.25)$$

where L is the span.

Hong et al. (1990) used another deflection measurement technique. In this technique, a circular section of the substrate is removed from

beneath the film to produce a drumhead-like membrane and the load is applied at its center. The stiffness of the membrane (film) is a sensitive function of the biaxial tension in it. The deflection h is related to load W as

$$h = \frac{Wa^2}{16\pi D} g(k) \quad (5.26)$$

where

$$D = \frac{Et^3}{12(1 - v^2)}$$

where a is the radius of the membrane, t is its thickness, and the function $g(k)$ depends on the membrane and its geometry. If the geometry and elastic constants of the membrane are known, the tension can be accurately evaluated. This technique can only be used to study tensile residual stresses since compressive stresses buckle the membrane when the substrate is removed.

5.6 Nanoindentation Techniques for Nanofracture Toughness and Nanofatigue

Nanoindentation techniques have been used for measurement of nanofracture toughness and nanofatigue (Bhushan 1999a; Bhushan and Li 2003).

5.6.1 Nanofracture Toughness

Fracture toughness, K_{Ic} , of a material is a measure of its resistance to the propagation of cracks and the ratio H/K_{Ic} is an index of brittleness, where H is the hardness (Bhushan 1999a; Bhushan and Li 2003). Resistance to fracture is a strong function of a crack pattern. It is measured typically in a tensile test in which a specimen containing a sharp crack of known length, c , is subjected to an applied stress, σ , which is increased during the test until the sample fractures (Lawn 1993), Fig. 5.50. The magnitude of the stresses near the crack tip are determined by

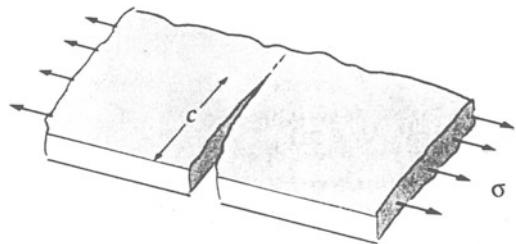


Fig. 5.50 Schematic of a standard specimen containing a sharp crack of known length used for measurement of fracture toughness of materials in tension

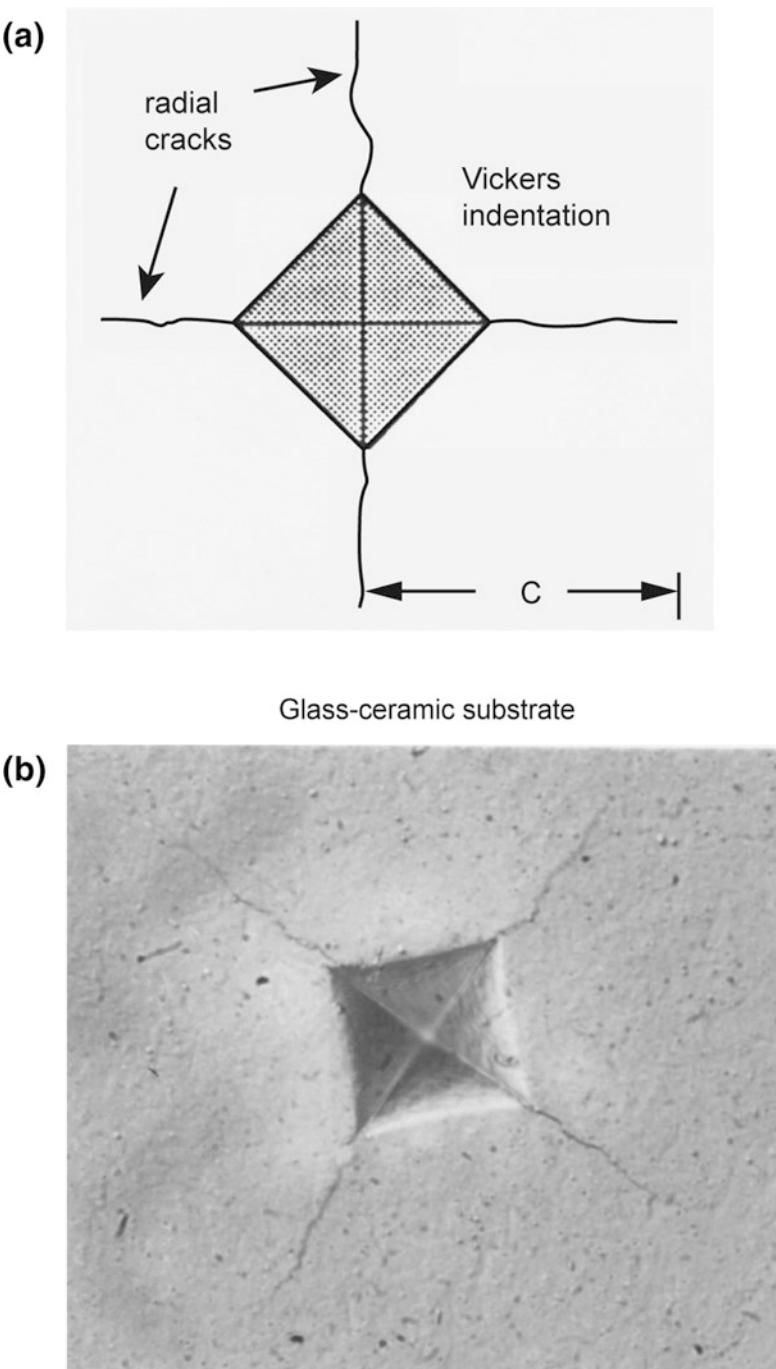
the stress intensity factor, K_I , which, in turn, depends on σ , c , and the specimen geometry

$$K_I = A\sigma\sqrt{\pi c} \quad (5.27)$$

The term A provides correction for the thickness-to-width ratio of the material. Units of stress-intensity factor are $MPa\sqrt{m}$. With more intense stress or with deeper cracks, the stress intensity becomes sufficient for the fracture to progress spontaneously. This threshold stress intensity is a property of the material and is called the critical stress intensity factor, K_{Ic} , or the fracture toughness of the material. Ceramics generally have relatively low fracture toughness, typically $1-2 MPa\sqrt{m}$; consequently, it is an important property to be considered for the selection of ceramics for industrial applications.

Indentation fracture toughness is another technique to determine fracture toughness (Palmquist 1957; Lawn and Wilshaw 1975; Antis et al. 1981; Chiang et al. 1982; Henshall and Brookes 1985; Cook and Pharr 1990; Lawn 1993; Pharr et al. 1993; Bhushan et al. 1996b). The indentation cracking method especially is useful for measurement of fracture toughness of thin films or small volumes. This method is quite different from conventional methods in that no special specimen geometry is required. Rather, the method relies on the fact that when indented with a sharp indenter, most brittle materials form radial cracks (Fig. 5.51b). The lengths of the surface traces of the radial cracks (for definition of crack length, see Fig. 5.51a) have been found to correlate reasonably well with fracture toughness. Then, by using simple empirical

Fig. 5.51 **a** Schematic of Vickers indentation with radial cracks, and **b** optical images of Vickers indentation made on a glass-ceramic substrate at 500 g load (adapted from Bhushan et al. 1996b)



equations, fracture toughness can be determined from simple measurement of crack length. Although this indentation fracture toughness measurement technique has been widely used in

practical applications, the stress/strain state is still, to a large extent, unknown. Stress/strain simulation on the indentation fracture with Vickers, Berkovich, and cube corner indenters is

needed for developing indentation fracture models and modifying the existing measurement methodology.

In microindentation, cracks at relatively high indentation loads of several hundred grams are on the order of 100 μm in length and can be measured optically. However, to measure toughness of very thin films or small volumes, much smaller indentations are required. However, a problem exists in extending the method to the nanoindentation regime in that there are well-defined loads, called cracking thresholds, below which indentation cracking does not occur in most brittle materials (Lankford 1981). For a Vickers indenter, cracking thresholds in most ceramics are about 25 g. Pharr et al. (1993), Li et al. (1997), and Li and Bhushan (1998a, 1999d) have found that the Berkovich indenter (a three-sided pyramid) with the same depth-to-area ratio as a Vickers indenter (a four-sided pyramid), has a cracking of the thresholds very similar to that of the Vickers indenter. They showed that cracking thresholds can be reduced substantially by using sharp indenters, i.e., indenters with smaller included tip angles such as a three-sided indenter with the geometry of the corner of a cube. Studies using a three-sided indenter with the geometry of the corner of a cube have revealed that cracking thresholds can be reduced to loads as small as 0.5 g, for which indentations and crack lengths in most materials are submicron in dimension.

Based on fracture mechanics analysis, Lawn et al. (1980) developed a mathematical relationship between fracture toughness and indentation crack length, given as

$$K_{Ic} = B \left(\frac{E}{H} \right)^{1/2} \left(\frac{W}{c^{3/2}} \right) \quad (5.28)$$

where W is the applied load and B is an empirical constant depending upon the geometry of the indenter (also see Lawn 1993; Pharr et al. 1993). Antis et al. (1981) conducted a study on a number of brittle materials chosen to span a wide range of toughnesses. They indented with a Vickers indenter at several loads and measured

crack length optically. They found a value of $B = 0.016$ to give good correlation between the toughness values measured from the crack length and the ones obtained using more conventional methods. Mehrotra and Quinto (1985) used a Vickers indenter to measure fracture toughness of the coatings. Bhushan et al. (1996b) measured fracture toughness of microcrystalline ceramic material (glass-ceramic) with Vickers indenter. They reported the formation of radial cracks, Fig. 5.51b. Note that cracks propagate in a zig-zag manner. The interlocked crystal morphology is responsible for propagation in a zigzag manner. Using Eq. 5.28, the fracture toughness for this material is calculated and presented in Table 5.2.

Pharr et al. (1993) tested several bulk ceramics listed in Table 5.3 using a cube corner indenter. They used an empirical constant, B , of about 0.032 for a cube corner indenter. Indentation cracking thresholds can be reduced significantly by using a cube corner indenter, which is of interest for hard ceramic materials requiring high loads to initiate cracks. Pharr et al. (1993) reported that predominant cracks formed with Vickers or Berkovich indenters were cone cracks, and the predominant cracks with a cube corner indenter were radial cracks, Fig. 5.52. The data for various ceramics are presented in Table 5.3.

Chantikul et al. (1981) developed a relationship between fracture toughness and the indentation fracture strength and the applied load, given as

$$K_{Ic} = c \left(\frac{E}{H} \right)^{1/8} \left(\sigma_f W^{1/3} \right)^{3/4} \quad (5.29)$$

where σ_f is the fracture strength after indentation at a given load and c is an empirical constant (0.59). The advantage of this analysis is that the measurement of crack length is not required.

For fracture toughness measurement of ultra-thin films ranging from 100 nm to few micrometers, because of the shallow indentation depths required in the indentation technique, it is difficult to measure a radial crack length even under

Fig. 5.52 Indentations in fused quartz made with the cube corner indenter showing radial cracking at indentation loads of **a** 12 g and **b** 0.45 g (adapted from Pharr et al. 1993)

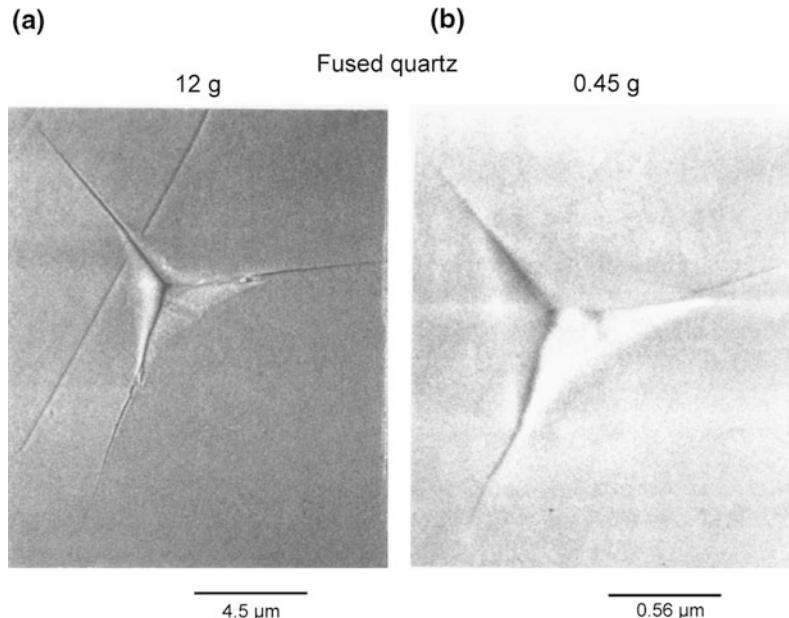


Table 5.3 Typical mechanical properties of materials (Pharr et al. 1993)

Material	E (GPa)	H (GPa)	K_{Ic} (MPa \sqrt{m})
Soda-lime glass	70	5.5	0.70
Fused quartz	72	8.9	0.58
Si(111)	168	9.3	0.70
Sapphire (111)	403	21.6	2.2
Si ₃ N ₄	300	16.3	4.0

SEM. Li et al. (1997) and Li and Bhushan (1998a, 1999d) developed a novel technique based on nanoindentation in which through-thickness cracking in the coating is detected from a discontinuity observed in the load-displacement curve and energy released during the cracking is obtained from the curve. Based on the energy released, fracture mechanics analysis is then used to calculate fracture toughness. A cube corner is preferred because the through-thickness cracking of hard films can be accomplished at lower loads (Li et al. 1997).

Load-displacement curves of indentations made at 30, 100, and 200 mN peak indentation loads together with the SEM micrographs of indentations on the 400 nm thick FCA carbon coating on silicon are shown in Fig. 5.53 (Li et al. 1997). Steps were found in all loading curves, as shown by arrows in Fig. 5.53a. In the

30-mN SEM micrograph, in addition to several radial cracks, ring-like through-thickness cracking was observed with small lips of material overhanging the edge of indentation. The step at about 23 mN in the loading curves of indentations made at 30 and 100 mN peak indentation loads resulted from the ring-like through thickness cracking. The step at 175 mN in the loading curve of indentation made at 200 mN peak indentation load was caused by spalling.

No steps were observed in the loading curve of indentation made at 20 mN peak indentation load (data not shown here), which suggests that the coating under the indenter was not separated instantaneously from the bulk coating via the ring-like through-thickness cracking but occurred over a period of time. At 30 mN peak indentation load, partial ring-like spalling was observed around the indenter and the other parts of the film

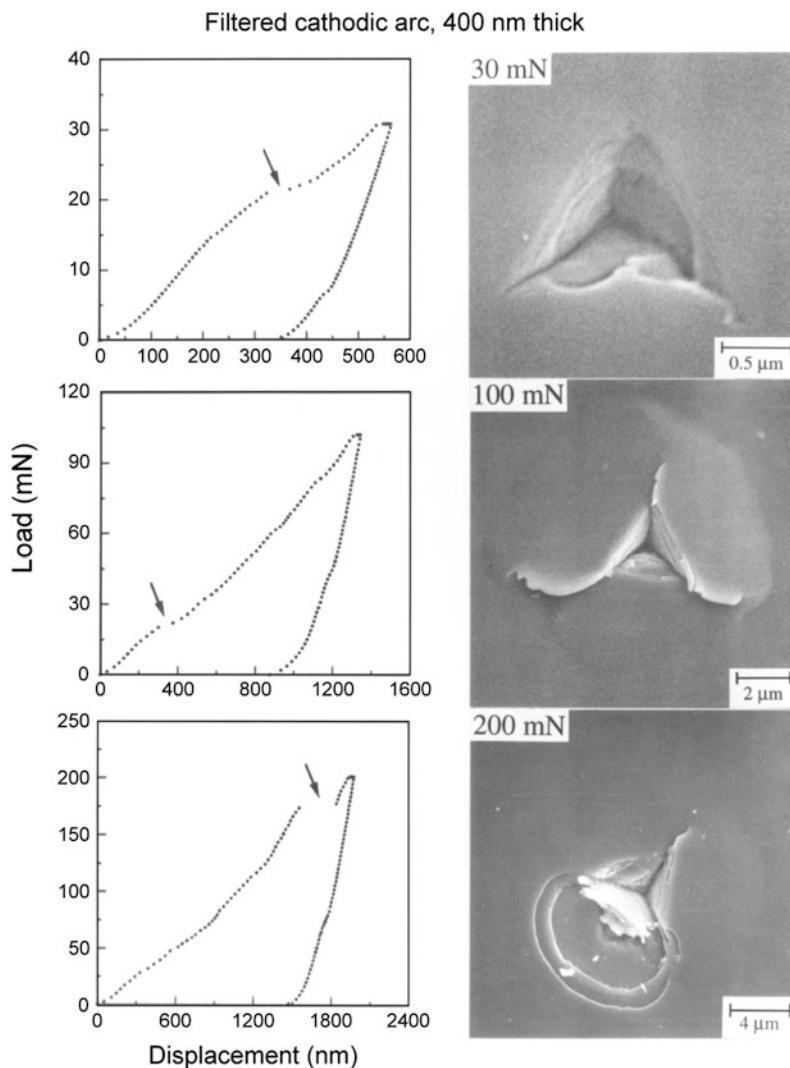


Fig. 5.53 **a** Load-displacement curves of indentations made at 30, 100 and 200 mN peak indentation loads using the cube corner indenter, and **b** the SEM micrographs of indentations on the 400 nm thick DLC coating deposited by filtered cathodic arc on silicon. Arrows indicate steps during loading portion of the load-displacement curve (adapted from Li et al. 1997)

bulged upward. This partial ring-like spalling was believed to result in the step in the loading curve. The absence of long steps in the loading curve for uncoated silicon reported, suggesting that the steps in the loading curve on the coating result from the film cracking (Li et al. 1997; Li and Bhushan 1998a). Based on their work, the fracture process progresses in three stages: (1) first ring-like through-thickness cracks form around the indenter by high stresses in the contact area, (2) delamination and buckling occur

around the contact area at the film/substrate interface by high lateral pressure, (3) second ring-like through-thickness cracks and spalling are generated by high bending stresses at the edges of the buckled film, see Fig. 5.54a (Li et al. 1997).

In the first stage, if the film under the indenter was separated from the bulk film via the first ring-like through-thickness cracking, a corresponding step will be present in the loading curve. If discontinuous cracks form and the film

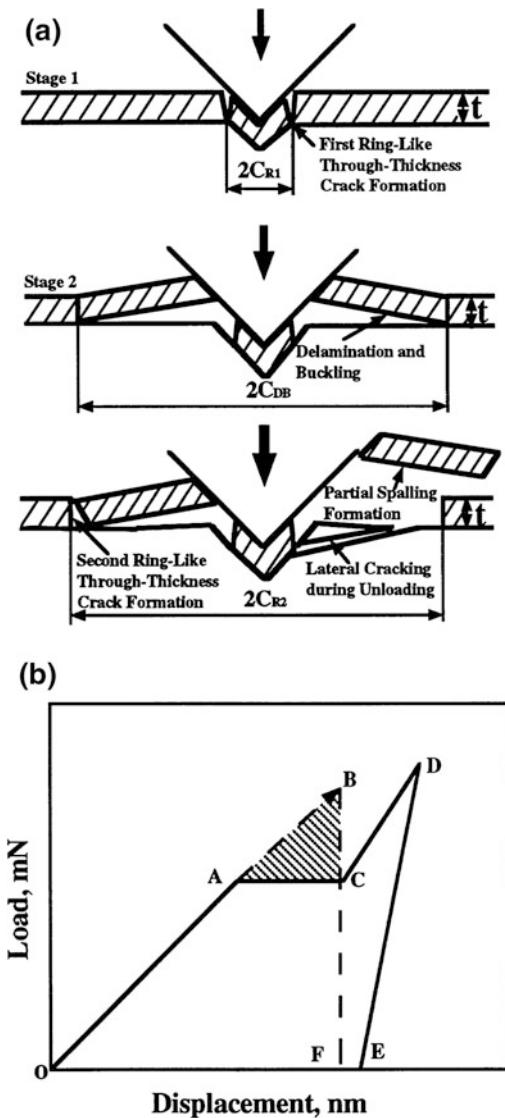


Fig. 5.54 **a** Schematic of various stages in nanoindentation fracture for the film/substrate system, and **b** schematic of a load-displacement curve, showing a step during the loading cycle and associated energy release (adapted from Li et al. 1997)

under the indenter is not separated from the remaining film, no step appeared in the loading curve because the film still supports the indenter and the indenter cannot suddenly advance into the material. In the second stage, the advance of the indenter during the radial cracking, delamination, and buckling was not big enough to form steps in the loading curve because the film

around the indenter still supported the indenter, but generated discontinuities that changed the slope of the loading curve with increasing indentation loads. In the third stage, the stress concentration at the end of the interfacial crack could not be relaxed by the propagation of the interfacial crack. With an increase in indentation depth, the height of the bulged film increased. When the height reached a critical value, the bending stresses caused by the bulged film around the indenter resulted in the second ring-like through-thickness crack formation and spalling at the edge of the buckled film, as shown in Fig. 5.54a, which leads to a step in the loading curve (Li et al. 1997). This was a single event and resulted in the separation of the part of the film around the indenter from the bulk film via cracking through films. The step in the loading curve was completely from the film cracking and not from the interfacial cracking or the substrate cracking.

The area under the load-displacement curve is the work performed by the indenter during elastic-plastic deformation of the film/substrate system. The strain energy release in the first/second ring-like cracking and spalling was calculated from the corresponding steps in the loading curve. Figure 5.54b shows a modeled load-displacement curve (Li et al. 1997). OACD is the loading curve. DE is the unloading. Since the first ring-like through-thickness cracking does not always lead to a step in the loading curve in some films, the second ring-like through-thickness crack should be considered. It should be emphasized that the edge of the buckled film is far from the indenter. Therefore, it does not matter if the indentation depth exceeds the film thickness or if deformation of the substrate occurs around the indenter when we measure fracture toughness of the film from the released energy during the second ring-like through-thickness cracking (spalling). Suppose that the second ring-like through-thickness cracking occurs at AC. Now, let us consider the loading curve OAC. If the second ring-like through-thickness crack does not occur, it can be understood that OA will be extended to OB to reach the same displacement as OC. This means

Table 5.4 Mechanical properties for the FCA, IB, ECR-CVD, and SP coatings

Coating	Hardness (Li and Bhushan 1999a) (GPa)	Elastic modulus (Li and Bhushan 1999a) (GPa)	Fracture toughness (Li and Bhushan 1999d) (MPa m ^{1/2})	Fatigue Life, N_f^a (Li and Bhushan 2002b) ($\times 10^4$)	Critical load during scratch (Li and Bhushan 1999a) (mN)
FCA	24	280	11.8	2.0	3.8
IB	19	140	4.3	0.8	2.3
ECR-CVD	22	180	6.4	1.2	5.3
SP	15	140	2.8	0.2	1.1
Single-crystal Silicon	11	220	0.75		0.6

Hardness, elastic modulus and fracture toughness were measured on 100-nm thick coatings. Fatigue life and critical load during scratch were measured on 20-nm thick coatings

^a N_f was obtained at a mean load of 10 μ N and a load amplitude of 8 μ N

that the crack formation changes the loading curve OAB into OAC. For point B, the elastic-plastic energy stored in the film/substrate system should be OBF. For point C, the elastic-plastic energy stored in the film/substrate system should be OACF. Therefore, the energy difference before and after the crack generation is the area of ABC; i.e., this energy stored in ABC will be released as strain energy to create the ring-like through-thickness crack. According to the theoretical analysis by Li et al. (1997), the fracture toughness of thin films can be written as

$$K_{Ic} = \left[\left(\frac{E}{(1 - v^2)2\pi C_R} \right) \left(\frac{U}{t} \right) \right]^{1/2} \quad (5.30)$$

where E is the elastic modulus, v is the Poisson's ratio, $2\pi C_R$ is the crack length in the film plane, U the strain energy difference before and after cracking, and t is the film thickness.

Using Eq. 5.30, the fracture toughness of the 0.4 μ m thick FCA carbon coating was calculated. The strain energy difference, U , of 7.1 nNm was assessed from the steps in Fig. 5.53a at the peak indentation loads of 200 mN (Li et al. 1997). The loading curve was extrapolated along the tangential direction of the loading curve from the starting point of the step up to reach the same displacement as the step. The area between the extrapolated line and the step was the estimated strain energy difference before and after cracking. C_R of 7.0 μ m is measured from the SEM

micrograph in Fig. 5.53b. The second ring-like crack is where the spalling occurs. For E of about 300 GPa measured using nanoindentation (Table 5.4) and an assumed value of 0.25 for v , fracture toughness values was calculated as about 11.8 $MPa\sqrt{m}$.

5.6.2 Nanofatigue

Fatigue failure progresses through a material via changes within the material at the tip of a crack, where there is a high stress intensity (Bhushan 1999a; Bhushan and Li 2003). Cyclic fatigue results from cyclic loading of machine components; e.g., the stresses cycle from tension and compression occurs in a loaded rotating shaft. Fatigue also can occur with fluctuating stresses of the same sign, similar to that in a leaf spring or in a dividing board. In a low-flying slider in a head-disk interface, isolated asperity contacts occur during use and the fatigue failure occurs in the multilayered thin-film structure of the magnetic disk (Bhushan 1996).

Li and Chu (1979) developed an indentation fatigue test, called impression fatigue. In this test, a cylindrical indenter with a flat end was pressed onto the surface of the test material with a cyclic load and the rate of plastic zone propagation was measured.

Li and Bhushan (2002b, c) used a depth-sensing nanoindenter using a harmonic

force and a conical diamond indenter typically with a nominal 1 μm radius tip. The CSM technique provides load cycles of a sinusoidal shape at high frequencies that can be used to perform nanoscale fatigue tests. The fatigue behavior of coatings can be studied by monitoring the change in contact stiffness since the contact stiffness is sensitive to damage formation. To obtain deformation and damage during fatigue loading, large amplitude oscillations were used. The numbers of cycles were determined from the elapsed time. Figure 5.55 shows the schematic of a fatigue test on a coating/substrate system using the CSM technique (Li and Bhushan 2002b). Load cycles are applied to the coating, resulting in a cyclic stress; P is the cyclic load, P_{mean} is the mean load, P_0 is the oscillation load amplitude, and ω is the oscillation frequency.

The following results can be obtained: (1) endurance limit; i.e., the maximum load below which there is no coating failure for a preset number of cycles; (2) number of cycles at which

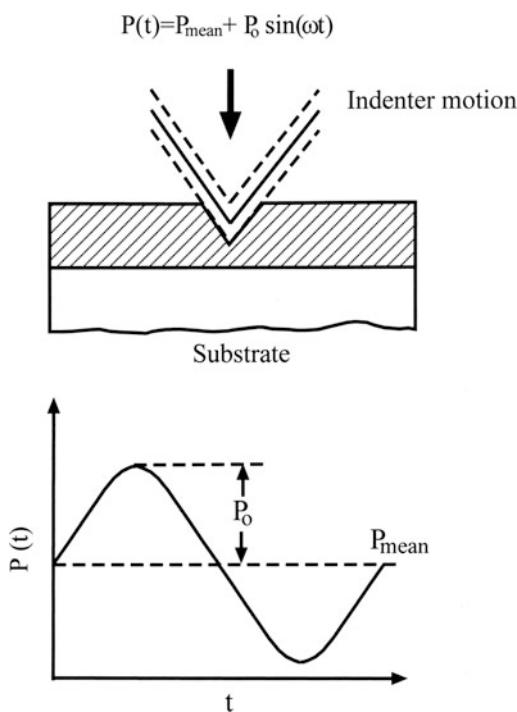


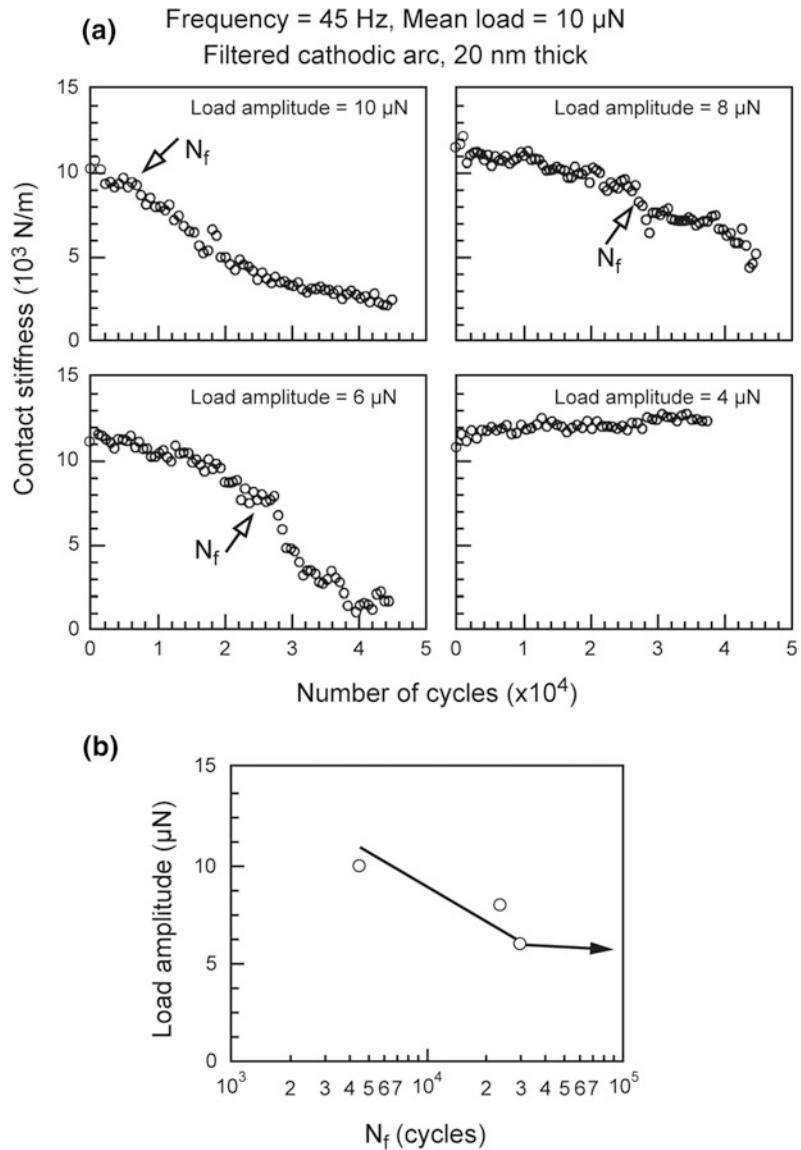
Fig. 5.55 Schematic of a fatigue test on a coating/substrate system using the CSM technique (adapted from Li and Bhushan 2002b)

the coating failure occurs; and (3) changes in contact stiffness measured using the unloading slope of each cycle, which can be used to monitor the propagation of the interfacial cracks during cyclic fatigue process.

Figure 5.56a shows the contact stiffness as a function of the number of cycles for 20 nm thick FCA coatings cyclically deformed by various oscillation load amplitudes with a mean load of 10 μN at a frequency of 45 Hz (Li and Bhushan 2002b). At 4 μN load amplitude, no change in contact stiffness was found for all coatings. This indicates that 4 μN load amplitude is not high enough to damage the coatings. At 6 μN load amplitude, an abrupt decrease in contact stiffness was found at a certain number of cycles for each coating, indicating that fatigue damage had occurred. With increasing load amplitude, the number of cycles to failure, N_f , decreased for all coatings. Load amplitude versus N_f , a so-called S-N curve, is plotted in Fig. 5.56b. The critical load amplitude, below which no fatigue damage occurs (an endurance limit), was identified for each coating. This critical load amplitude together with mean load is of critical importance to the design of head-disk interfaces or MEMS/NEMS device interfaces (Bhushan 1996, 2017).

Figure 5.57a shows the high magnification SEM images of 20 nm thick FCA coatings before, at, and after N_f (Li and Bhushan 2002b). In the SEM images, the net-like structure is the gold film coated on the DLC coating, which should be ignored in analyzing the indentation fatigue damage. Before the N_f , no delamination or buckling was found except the residual indentation mark at magnifications up to 1,200,000 \times using SEM. This suggests that only plastic deformation occurred before the N_f . At the N_f , the coating around the indenter bulged upwards, indicating delamination and buckling. Therefore, it is believed that the decrease in contact stiffness at the N_f results from the delamination and buckling of the coating from the substrate. After the N_f , the buckled coating was broken down around the edge of the buckled area, forming a ring-like crack. The remaining coating overhung at the edge of the buckled area.

Fig. 5.56 **a** Contact stiffness as a function of the number of cycles for 20-nm thick filtered cathodic arc coatings cyclically deformed by various oscillation load amplitudes with a mean load of 10 μN at a frequency of 45 Hz; and **b** plot of load amplitude versus N_f (adapted from Li and Bhushan 2002b)



It is noted that the indentation size increases with increasing number of cycles. This indicates that deformation, delamination and buckling, and ring-like crack formation occurred over a period of time.

The schematic in Fig. 5.57b shows various stages in the indentation fatigue damage for a coating/substrate system (Li and Bhushan 2002b). Based on this study, three stages in the indentation fatigue damage appear to exist: (1) indentation induced compression;

(2) delamination and buckling; (3) ring-like crack formation at the edge of the buckled coating. Residual stresses are often induced in coatings by the deposition process. The model shown in Fig. 5.57b considers a coating with a uniform biaxial residual compression σ_r . In the first stage, indentation induces elastic/plastic deformation, exerting an outward acting pressure on the coating around the indenter. Interfacial defects like voids and impurities act as original cracks. These cracks propagate and link up as the

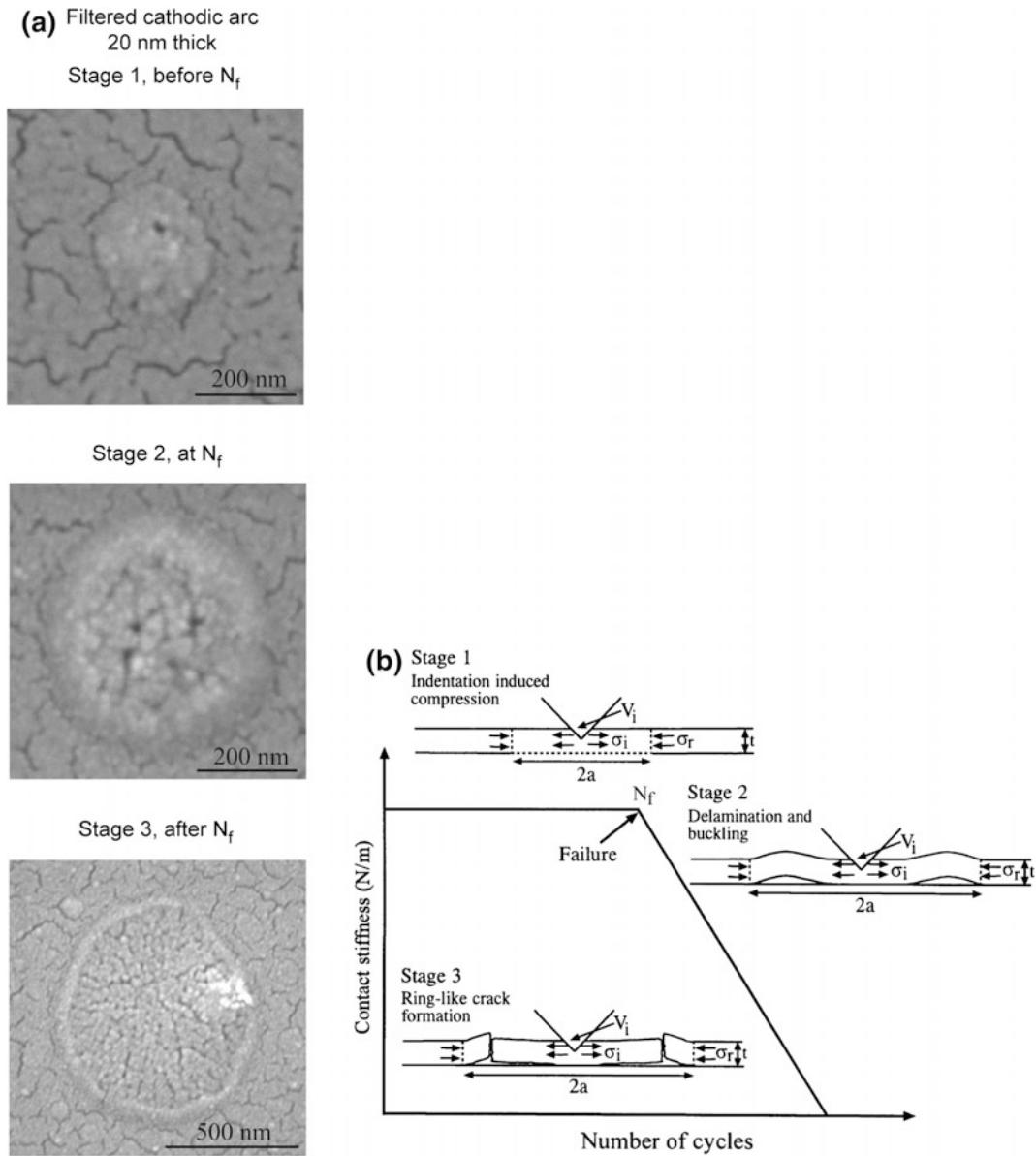


Fig. 5.57 **a** High magnification SEM images of a coatings before, at, and after N_f for 20-nm thick filtered cathodic arc coatings, and **b** schematic of various stages in the indentation fatigue damage for a coating/substrate system (adapted from Li and Bhushan 2002b)

indentation compressive stress increases. At this stage, the coating, which is under the indenter and above the interfacial crack (with a crack length of $2a$), still maintains a solid contact with the substrate; the substrate still fully supports the coating. Therefore, this interfacial crack does not lead to an abrupt decrease in contact stiffness, but

gives a rise to a slight decrease in contact stiffness as shown in Fig. 5.56. The coating above the interfacial crack is treated as a rigidly clamped disk. We assume that the crack radius, a , is large compared with the coating thickness t . Since the coating thickness ranges from 20 to 5 nm, this assumption is satisfied easily in this

study [radius of the delaminated and buckled area, shown in Fig. 5.57a, is on the order of 100 nm (Li and Bhushan 2002b)]. The compressive stress caused by indentation is given as (Marshall and Evans 1984)

$$\sigma_i = \frac{EV_i}{2\pi t a^2(1-v)} \quad (5.31)$$

where v and E are the Poisson's ratio and elastic modulus of the coating, V_i is the indentation volume, t is the coating thickness, and a is the crack radius. With increasing number of cycles, the indentation volume V_i increases. Therefore, the indentation compressive stress σ_i increases accordingly. In the second stage, buckling occurs during the unloading segment of fatigue testing cycle when the sum of indentation compressive stress σ_i and the residual stress σ_r exceed the critical buckling stress σ_b for the delaminated circular section as given by (Evans and Hutchinson 1984)

$$\sigma_b = \frac{\mu^2 E}{12(1-v^2)} \left(\frac{t}{a}\right)^2 \quad (5.32)$$

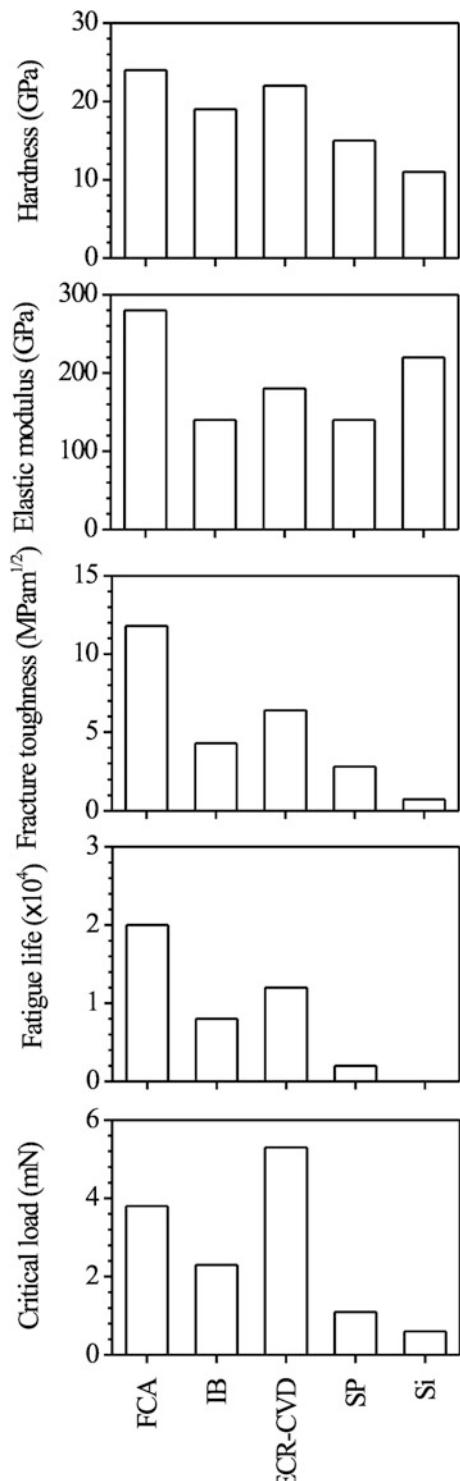
where the constant μ equals 42.67 for a circular clamped plate with a constrained center point and 14.68 when the center is unconstrained. The buckled coating acts as a cantilever. In this case, the indenter indents a cantilever rather than a coating/substrate system. This ultrathin coating cantilever has much less contact stiffness than the coating/substrate system. Therefore, the contact stiffness shows an abrupt decrease at the N_f . In the third stage, with an increased number of cycles, the delaminated and buckled size increases, resulting in a further decrease in contact stiffness since the cantilever beam length increases. On the other hand, a high bending stress acts at the edge of the buckled coating. The larger is the buckled size, the higher is the bending stress. The cyclically-bending stress causes fatigue damage at the end of the buckled coating, forming a ring-like crack. The coating under the indenter is separated from the bulk coating (caused by the ring-like crack at the edge of the buckled coating) and the substrate (caused

by the delamination and buckling in the second stage). Therefore, the coating under the indenter is not constrained, but is free to move with the indenter during fatigue testing. At this point, the sharp nature of the indenter is lost because the coating under the indenter gets stuck on the indenter. The indentation fatigue experiment results in the contact of a relatively huge, blunt tip with the substrate. This results in a low contact stiffness value.

Compressive residual stresses assist in delamination and buckling. A coating with higher adhesion strength and a less compressive residual stress is required for a higher fatigue life. Interfacial defects should be avoided in the coating deposition process. We know that the ring-like crack formation occurs in the coating. Formation of fatigue cracks in the coating depends upon the hardness and fracture toughness. Cracks are more difficult to form and propagate in the coating with higher strength and fracture toughness.

It is now accepted that long fatigue life in a coating/substrate almost always involves 'living with cracks,' that the threshold or limit condition is associated with the non-propagation of existing cracks or defects, even though these cracks may be undetectable (Suresh 1991). For all coatings studied, at 4 μ N, contact stiffness does not change much. This indicates that delamination and buckling did not occur within the number of cycles tested in this study. This is probably because the indentation induced compressive stress was not high enough to allow the cracks to propagate and link up under the indenter or the sum of indentation compressive stress σ_i and the residual stress σ_r did not exceed the critical buckling stress σ_b .

Figure 5.58 and Table 5.4 summarize the hardness, elastic modulus, fracture toughness, fatigue life, and critical load during scratch of the DLC coatings deposited by four deposition techniques—FCA, IB, ECR-CVD, and SP and single-crystal silicon substrate (Li and Bhushan 2002b). A good correlation exists between fatigue life and other mechanical properties. Higher mechanical properties result in a longer fatigue life. The mechanical properties of DLC coatings



are controlled by the sp^3 -to- sp^2 ratio. The sp^3 -bonded carbon exhibits the outstanding properties of diamond (Bhushan 1999c). A higher deposition kinetic energy will result in a larger fraction of sp^3 -bonded carbon in an amorphous network. Thus, the higher kinetic energy for the FCA could be responsible for its better carbon structure and higher mechanical properties (Bhushan 1999c). Higher adhesion strength between the FCA coating and substrate makes the FCA coating more difficult to delaminate from the substrate.

5.7 Summary

A review of the nanoindentation techniques and various measurement options are presented. Nanoindentation techniques can be used to measure elastic-plastic deformation behavior, hardness, elastic modulus, film-substrate adhesion, residual stresses, time-dependent creep and relaxation properties, fracture toughness, and fatigue of surface layers of bulk materials and thin coatings. The nanoindentation apparatus monitors the load and the position of the indenter relative to the surface of the specimen (displacement) during the indentation process. The load-displacement curve is used to calculate various mechanical properties. A motorized sample stage to move the sample in the lateral direction while loaded against the indenter allows nanoscratch and nanofriction and nanowear tests. A lateral force sensor and/or AE sensor generally are attached to the nanoindenter to monitor changes during the sliding process. The CSM technique *in situ* probes the mechanical property changes during indentation, and provides useful information for layered materials and nonhomogeneous (such as graded) composites. The CSM

Fig. 5.58 Bar charts summarizing the data of various DLC coatings and single-crystal silicon substrate. Hardness, elastic modulus and fracture toughness were measured on 100-nm thick coatings, and fatigue life and critical load during scratch were measured on 20-nm thick coatings (adapted from Li and Bhushan 2002b)

indentation creep tests can detect creep displacement and stress relaxation at small volumes. Indentation cracking method can be used for measurement of fracture toughness. Load cycles used in the CSM can be used to perform nanoscale fatigue tests.

Nanoindentation techniques can be used at indentation depths as low as about 10 nm. Indentation at an indentation depth less than 20 nm requires sharp indenters with a tip radius less than 20 nm and the use of ultralow loads.

References

- Ahn J, Mittal KL, Macqueen RH (1978) Hardness and adhesion of filmed structures as determined by the scratch technique. In: Mittal KL (ed) Adhesion measurement of thin films, thick films, and bulk coatings. STP 640, ASTM, Philadelphia, pp 134–157
- Alba S, Loubet JL, Vovelle L (1993) Evaluation of mechanical properties and adhesion of polymer coatings by continuous hardness measurements. *J Adhesion Sci Technol* 7:131–140
- Alekhin VP, Berlin GS, Isaev AV, Kalei GN, Merkulov VA, Skvortsov VN, Ternovskii AP, Krushchov MM, Shnyrev GD, Shorshorov MKh (1972) Micromechanical testing by micromechanical testing of materials by microcompression. *Zavod Lab* 38:619–621
- Anonymous (1979) Standard test method for microhardness of materials. ASME designation: E384–73. ASTM, Philadelphia, pp 359–379
- Anonymous (1988) Properties of silicon. EMIS data reviews series no. 4. INSPEC, The Institution of Electrical Engineers, London
- Anonymous (2014) Keysight Nano Indenter G200 data sheet. Keysight technologies, Santa Rosa, CA
- Antis GR, Chantikul P, Lawn BR, Marshall DB (1981) A critical evaluation of indentation techniques for measuring fracture toughness: I direct crack measurements. *J Am Ceram Soc* 64:533–538
- Atkins AG, Silverio A, Tabor D (1966) Indentation hardness and the creep of solids. *J Inst Metals* 94:369–378
- Bell TJ, Field JS, Swain MV (1992) Elastic-plastic characterization of thin films with spherical indentation. *Thin Solid Films* 220:289–294
- Benjamin P, Weaver C (1960) Measurement of adhesion of thin films. *Proc R Soc Lond A* 254:163–176
- Berkovich ES (1951) Three-faceted diamond pyramid for micro-hardness testing. *Indus Diamond Rev* 11:129–132
- Bhattacharya AK, Nix WE (1988a) Finite element simulation of indentation experiments. *Int J Solids Struct* 24:881–891
- Bhattacharya AK, Nix WD (1988b) Analysis of elastic and plastic deformation associated with indentation testing of the thin films on substrates. *Int J Solids Struct* 24:1287–1298
- Bhushan B (1987) Overview of coating materials, surface treatments, and screening techniques for tribological applications—Part 2: screening techniques. In: Hardening WB, DiBari GA (eds) Testing of metallic and inorganic coatings, STP 947. ASTM, Philadelphia, pp 310–319
- Bhushan B (1996) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer, New York
- Bhushan B (1999a) Nanomechanical properties of solid surfaces and thin films. *Handbook of micro/nanotribology*, 2nd edn. CRC Press, Boca Raton, pp 443–524
- Bhushan B (1999b) Wear and mechanical characterisation on micro- to picoscales using AFM. *Int Mater Rev* 44:105–117
- Bhushan B (1999c) Chemical, mechanical, and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments. *Diam Relat Mater* 8:1985–2015
- Bhushan B (2001) Modern tribology handbook, vol 1 and 2. CRC Press, Boca Raton
- Bhushan B (2011) *Nanotribology and nanomechanics I and II*. Springer, Heidelberg
- Bhushan B (2013a) Principles and applications of tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2017) *Springer handbook of nanotechnology*, 4th edn. Springer International, Switzerland
- Bhushan B, Gupta BK (1995) Micromechanical characterization of Ni-P coated aluminum-magnesium, glass, and glass ceramic substrates and finished magnetic thin-film rigid disks. *Adv Info Storage Syst* 6:193–208
- Bhushan B, Gupta BK (1997) Handbook of tribology: materials, coatings and surface treatments. McGraw Hill, New York (1991); Reprint edition (1997), Krieger Publishing Co., Malabar
- Bhushan B, Koinkar VN (1994) Nanoindentation hardness measurements using atomic force microscopy. *Appl Phys Lett* 64:1653–1655
- Bhushan B, Li X (1997) Micromechanical and tribological characterization of doped single-crystal silicon and polysilicon films for microelectromechanical system devices. *J Mater Res* 12:54–63
- Bhushan B, Li X (2003) Nanomechanical characterisation of solid surfaces and thin films. *Int Mater Rev* 48:125–164
- Bhushan B, Patton ST (1996) Pole tip recession studies of hard carbon-coated thin-film tape heads. *J Appl Phys* 79:5916–5918
- Bhushan B, Venkatesan S (2005) Effective mechanical properties of layered rough surfaces. *Thin Solid Films* 473:278–295
- Bhushan B, Landesman AL, Shack RV, Vukobratovich D, Walters VS (1985) Instrument for testing thin films

- such as magnetic tapes. IBM Tech Discl Bull 28:2975–2976
- Bhushan B, Williams VS, Shack RV (1988) In-situ nanoindentation hardness apparatus for mechanical characterization of extremely thin films. ASME J Tribol 110:563–571
- Bhushan B, Gupta BK, Azarian MH (1995) Nanoindentation, microscratch, friction and wear studies of coatings for contact recording applications. Wear 181–183:743–758
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996a) Nanoindentation and picoindentation measurements using a capacitive transducer system in atomic force microscopy. Philos Mag 74:1117–1128
- Bhushan B, Chyung K, Miller RA (1996b) Micromechanical property measurements of glass and glass-ceramic substrates for magnetic thin-film rigid disks for gigabit recording. Adv Info Storage Syst 7:3–16
- Bhushan B, Theunissen GSAM, Li X (1997) Tribological studies of chromium oxide films for magnetic recording applications. Thin Solid Films 311:67–80
- Bhushan B, Luo D, Schricker SR, Sigmund W, Zauscher S (2014) In: Handbook of nanomaterials properties, vol 1–2. Springer, Heidelberg
- Blau PJ, Lawn BR (eds) (1986) Microindentation techniques in materials science and engineering, STP 889. ASTM, Philadelphia
- Blau PJ, Oliver WC, Snead L (1997) The scanning micro-sclerometer: a new method for scratch hardness mapping. Tribol Int 30:483–490
- Bolshakov A, Oliver WC, Pharr GM (1996) Influences of stress on the measurement of mechanical properties using nanoindentation. II. Finite element simulations. J Mater Res 11:760–768
- Buckle H (1973) Use of the hardness test to determine material properties. In: *The Science of Hardness Testing and its Research Applications* (Westbrook JW, and Conrad H, eds.), pp. 453–491, Am. Soc. Metals, Metals Park, Ohio.
- Bull SJ, Rickerby DS (1990) New developments in the modelling of the hardness and scratch adhesion of thin films. Surf Coatings Technol 42:149–164
- Bulychev SI, Alekhin VP, Shorshorov MKh, Terновskii AP, Shnyrev GD (1975) Determining Young's Modulus from the indenter penetration diagram. Zavod Lab 41:11137–11140
- Bulychev SI, Alekhin VP, Shorshorov MKh (1979) Studies of physico-mechanical properties in surface layers and microvolumes of materials by the method of continuous application of an indenter. Fizika Khim Obrab Materialov, No 5
- Burnett PJ, Rickerby DS (1987a) The relationship between hardness and scratch adhesion. Thin Solid Films 154:403–416
- Burnett PJ, Rickerby DS (1987b) The mechanical properties of wear resistant coatings I: modelling of hardness behavior. Thin Solid Films 148:41–50
- Callahan DL, Morris JC (1992) The extent of phase transformation in silicon hardness indentations. J Mater Res 7:1614–1617
- Campbell DS (1970) Mechanical properties of thin films. In: Maissel LI, Glang R (eds) *Handbook of Thin Film Technology* (Chap. 12). McGraw-Hill, New York
- Chantikul P, Anstis GR, Lawn BR, Marshall DB (1981) A critical evaluation of indentation techniques for measuring fracture toughness: II, strength method. J Am Ceram Soc 64:539–543
- Chiang SS, Marshall DB, Evans AG (1981) Simple method for adhesion measurement. In: Pask J, Evans AG (eds) *Surfaces and interfaces in ceramics and ceramic-metal systems*. Plenum, New York, pp 603–612
- Chiang SS, Marshall DB, Evans AG (1982) The response of solids to elastic/plastic indentation: I. stresses and residual stresses. J Appl Phys 53:298–311
- Cho D, Bhushan B (2016) Nanofriction and nanowear of polypropylene, polyethylene terephthalate, and high-density polyethylene during sliding. Wear 352–353:18–23
- Chu SNG, Li JCM (1977) Impression creep: a new creep test. J Mater Sci 12:2200–2208
- Chu SNG, Li JCM (1980) Localized stress relaxation by impression testing. Mater Sci Eng 45:167–171
- Cook RF, Pharr GM (1990) Direct observation and analysis of indentation cracking in glasses and ceramics. J Am Ceram Soc 73:787–817
- Doerner MF, Nix WD (1986) A method for interpreting the data from depth-sensing indentation instruments. J Mater Res 1:601–609
- Doerner MF, Gardner DS, Nix WD (1986) Plastic properties of thin films on substrates as measured by submicron indentation hardness and substrate curvature techniques. J Mater Res 1:845–851
- Evans AG, Hutchinson JW (1984) On the mechanics of delamination and spalling in compressed films. Int J Solids Struct 20:455–466
- Fabes BD, Oliver WC, McKee RA, Walker FJ (1992) The determination of film hardness from the composite response of film and substrate to nanometer scale indentation. J Mater Res 7:3056–3064
- Fischer-Cripps AC (2002) *Nanoindentation*. Springer, New York
- Fleck NA, Muller GM, Ashby MF, Hutchinson JW (1994) Strain gradient plasticity: theory and experiments. Acta Metall Mater 42:475–487
- Gane N, Cox JM (1970) The micro-hardness of metals at very low loads. Philos Mag 22:881–891
- Goken M, Kempf M (2001) Pop-ins in Nanoindentation —The initial yield point. Z Metallkd 92:1061–1067
- Greene JE, Woodhouse J, Pestes M (1974) A technique for detecting critical loads in the scratch test for thin-film adhesion. Rev Sci Instrum 45:747–749
- Gupta BK, Bhushan B (1994) The nanoindentation studies of ion implanted silicon. Surf Coat Technol 68(69):564–570

- Gupta BK, Bhushan B (1995a) Micromechanical properties of amorphous carbon coatings deposited by different deposition techniques. *Thin Solid Films* 270:391–398
- Gupta BK, Bhushan B (1995b) Mechanical and tribological properties of hard carbon coatings for magnetic recording heads. *Wear* 190:110–122
- Gupta BK, Chevallier J, Bhushan B (1993) Tribology of ion bombarded silicon for micromechanical applications. *ASME J Tribol* 115:392–399
- Gupta BK, Bhushan B, Chevallier J (1994) Modification of tribological properties of silicon by boron ion implantation. *Tribol Trans* 37:601–607
- Hainsworth SV, Chandler HW, Page TF (1996) Analysis of nanoindentation load-displacement loading curves. *J Mater Res* 14:2283–2295
- Hainsworth SV, McGurk MR, Page TF (1998) The effect of coating cracking on the indentation response of thin hard-coated systems. *Surf Coat Technol* 102:97–107
- Hannula SP, Wanagel J, Li CY (1986) A miniaturized mechanical testing system for small scale specimen testing. In: Corwin WR, Lucas GE (eds) The use of small-scale specimens for testing irradiated material, STP 888. ASTM, Philadelphia, pp 233–251
- Hay JC, Bolshakov A, Pharr GM (1999) A critical examination of the fundamental relations used in the analysis of nanoindentation data. *J Mater Res* 14:2296–2305
- Heavens OS (1950) Some factors influencing the adhesion of films produced by vacuum evaporation. *J Phys Rad* 11:355–360
- Henshall JL, Brookes CA (1985) The measurement of K_{Ic} in single crystal SiC using the indentation method. *J Mater Sci Lett* 4:783–786
- Hong S, Weihs TP, Bravman JC, Nix WD (1990) Measuring stiffnesses and residual stresses of silicon nitride in thin films. *J Electron Mater* 19:903
- Jacobson S, Jonsson B, Sundquist B (1983) The use of fast heavy ions to improve thin film adhesion. *Thin Solid Films* 107:89–98
- Johnson KL (1985) Contact mechanics. Cambridge University Press, Cambridge
- Joslin DL, Oliver WC (1990) A new method for analyzing data from continuous depth-sensing microindentation tests. *J Mater Res* 5:123–126
- King RB (1987) Elastic analysis of some punch problems for layered medium. *Int J Solids Struct* 23:1657–1664
- Korsunsky AM, McGurk MR, Bull SJ, Page TF (1998) On the hardness of coated systems. *Surf Coat Technol* 99:171–183
- Kulkarni AV, Bhushan B (1996a) Nanoscale mechanical property measurements using modified atomic force microscopy. *Thin Solid Films* 290–291:206–210
- Kulkarni AV, Bhushan B (1996b) Nano/picoindentation measurements on single-crystal aluminum using modified atomic force microscopy. *Mater Lett* 29:221–227
- Kulkarni AV, Bhushan B (1997) Nanoindentation measurements of amorphous carbon coatings. *J Mater Res* 12:2707–2714
- Kumar A, Bhushan B (2015) Nanomechanical, nanotribological, and macrotribological characterization of hard coatings and surface treatment of H-13 steel. *Tribol Int* 81:149–158
- LaFontaine WR, Yost B, Black RD, Li CY (1990) Indentation load relaxation experiments with indentation depth in the submicron range. *J Mater Res* 5:1206–2100
- LaFontaine WR, Paszkiet CA, Korhonen MA, Li CY (1991) Residual stress measurements of thin aluminum metallizations by continuous indentation and X-ray stress measurement techniques. *J Mater Res* 6:2084–2090
- Lankford J (1981) Threshold-microfracture during elastic/plastic indentation of ceramics. *J Mater Sci* 16:1177–1182
- Laugier M (1981) The development of scratch test technique for the determination of the adhesion of coating. *Thin Solid Films* 76:289–294
- Laursen TA, Simo JC (1992) A study of the mechanics of microindentation using finite elements. *J Mater Res* 7:618–626
- Lawn B (1993) Fracture of brittle solids, 2nd edn. Cambridge University Press, Cambridge
- Lawn B, Wilshaw R (1975) Review indentation fracture: principles and applications. *J Mater Sci* 10:1049–1081
- Lawn BR, Evans AG, Marshall DB (1980) Elastic/plastic indentation damage in ceramics: the median/radial crack system. *J Am Ceram Soc* 63:574–581
- Li JCM, Chu SNG (1979) Impression fatigue. *Scr Metall* 13:1021–1026
- Li X, Bhushan B (1998a) Measurement of fracture toughness of ultra-thin amorphous carbon films. *Thin Solid Films* 315:214–221
- Li X, Bhushan B (1998b) Micromechanical and tribological characterization of hard amorphous carbon coatings as thin as 5 nm for magnetic recording heads. *Wear* 220:51–58
- Li X, Bhushan B (1999a) Micro/nanomechanical and tribological characterization of ultra-thin amorphous carbon coatings. *J Mater Res* 14:2328–2337
- Li X, Bhushan B (1999b) Mechanical and tribological studies of ultra-thin hard carbon overcoats for magnetic recording heads. *Z Metallkd* 90:820–830
- Li X, Bhushan B (1999c) Micro/nanomechanical characterization of ceramic films for microdevices. *Thin Solid Films* 340:210–217
- Li X, Bhushan B (1999d) Evaluation of fracture toughness of ultrathin and hard amorphous carbon coatings deposited by different deposition techniques. *Thin Solid Films* 355–356:330–336
- Li X, Bhushan B (2000a) Development of continuous stiffness measurement technique for composite magnetic tapes. *Scr Mater* 42:929–935
- Li X, Bhushan B (2000b) Continuous stiffness measurement and creep behavior of composite magnetic tapes. *Thin Solid Films* 377–378:401–406
- Li X, Bhushan B (2001a) Continuous stiffness measurements of layered materials used in magnetic storage devices. *J Info Storage Proc Syst* 3:131–142

- Li X, Bhushan B (2001b) Dynamic mechanical characterization of magnetic tapes using nanoindentation techniques. *IEEE Trans Mag* 37:1616–1619
- Li X, Bhushan B (2002a) A review of nanoindentation continuous stiffness measurement and its applications. *Mater Charact* 48:11–36
- Li X, Bhushan B (2002b) Development of a nanoscale fatigue measurement technique and its application to ultrathin amorphous carbon coatings. *Scr Mater* 47:473–479
- Li X, Bhushan B (2002c) Nanofatigue studies of ultrathin hard carbon overcoats used in magnetic storage devices. *J Appl Phys* 91:8334–8336
- Li X, Bhushan B (2003) Fatigue studies of nanoscale structures for MEMS/NEMS applications using nanoindentation techniques. *Surf and Coat Technol* 163–164:521–526
- Li WB, Henshall JL, Hooper RM, Easterling KE (1991) The mechanism of indentation creep. *Acta Metall Mater* 39:3099–3110
- Li X, Diao D, Bhushan B (1997) Fracture mechanisms of thin amorphous carbon films in nanoindentation. *Acta Mater* 45:4453–4461
- Li X, Bhushan B, Inoue M (2001) Time-dependent mechanical properties and tribological behavior of magnetic tapes. *Wear* 251:1150–1158
- Li X, Bhushan B, Takashima K, Baek CS, Kim YK (2003) Mechanical characterization of micro/nanoscale structures for MEMS/NEMS applications using nanoindentation techniques. *Ultramicroscopy* 97:481–494
- Lin MR, Ritter JE, Rosenfeld L, Lardner TJ (1990) Measuring the interfacial shear strength of thin polymer coatings on glass. *J Mater Res* 5:1110–1117
- Loubet JL, Georges JM, Marchesini O, Meille G (1984) Vickers indentation curves of magnesium oxide (MgO). *ASME J Tribol* 106:43–48
- Loubet JL, Bauer M, Tonck A, Bec S, Gauthier-Manuel B (1993) Nanoindentation with a surface force apparatus. In: Nastasi M, Parkin DM, Gleiter H (eds) *Mechanical properties and deformation behavior of materials having ultra-fine microstructures*. Kluwer Academic Pub., Dordrecht, pp 429–447
- Lysaght VE (1949) *Indentation Hardness testing*. Reinhold, New York
- Maharaj D, Bhushan B (2015) Friction, wear, and mechanical behavior of nano-objects on the nanoscale. *Mater Sci Eng R* 95:1–43
- Marshall DB, Evans AG (1984) Measurement of adherence of residual stresses in thin films by indentation. I. Mechanics of interface delamination. *J Appl Phys* 15:2632–2638
- Marshall DB, Lawn BR (1979) Residual stress effects in sharp contact cracking part 1 indentation fracture mechanics. *J Mater Sci* 14:2001–2012
- Marshall DB, Oliver WC (1987) Measurement of interfacial mechanical properties in fiber-reinforced ceramic composites. *J Am Ceram Soc* 70:542–548
- Mayo MJ, Nix WD (1988) A micro-indentation study of superplasticity in Pb, Sn, and Sn–38 Wt% Pb. *Acta Metall* 36:2183–2192
- McGurk MR, Page TF (1999) Using the P- δ analysis to deconvolute the nanoindentation response of hard-coated systems. *J Mater Res* 14:2283–2295
- Mehrotra PK, Qunito DT (1985) Techniques for evaluating mechanical properties of hard coatings. *J Vac Sci Technol A* 3:2401–2405
- Mittal KL (ed) (1978) *Adhesion measurements on thin coatings, thick coatings, and bulk coatings*, STP640. ASTM, Philadelphia
- Mott BW (1957) *Microindentation hardness testing*. Butterworths, London
- Mulhearn TO, Tabor D (1960) Creep and hardness of metals: a physical study. *J Inst Metals* 87:7–12
- Nastasi M, Parkin DM, Gleiter H (eds) (1993) *Mechanical properties and deformation behavior of materials having ultra-fine microstructures*. Kluwer Academic Pub., Dordrecht
- Newey D, Wilkins MA, Pollock HM (1982) An Ultra-low-load penetration hardness tester. *J Phys E Sci Instrum* 15:119–122
- Nix WD (1989) Mechanical properties of thin films. *Metall Trans A* 20:2217–2245
- Oliver WC (2001) Alternative technique for analyzing instrumented indentation data. *J Mater Res* 16:3202–3206
- Oliver WC, Pharr GM (1992) An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J Mater Res* 7:1564–1583
- Oliver WC, Hutchings R, Pethica JB (1986) Measurement of hardness at indentation depths as small as 20 nm. In: Blau PJ, Lawn BR (eds) *Microindentation techniques in materials science and engineering*, STP 889. ASTM, Philadelphia, pp 90–108
- O'Neill H (1967) *Hardness measurement of metals and alloys*. Chapman and Hall, London
- Page TF, Oliver WC, McHargue CJ (1992) The deformation behavior of ceramic crystals subjected to very low load (nano)indentations. *J Mater Res* 7:450–473
- Palacio M, Bhushan B (2010) Nanomechanical characterization of adaptive optics components in microprojectors. *J Micromech Microeng* 20:064002
- Palacio MLB, Bhushan B (2013) Depth-sensing indentation of nanomaterials and nanostructures. *Mater Charact* 78:1–20
- Palacio M, Bhushan B, Ferrell N, Hansford D (2007a) Nanomechanical characterization of polymer beam structures for BioMEMS applications. *Sens Actuators A* 135:637–650
- Palacio M, Bhushan B, Ferrell N, Hansford D (2007b) Adhesion properties of polymer/silicon interfaces for biological micro-/nanoelectromechanical systems applications. *J Vac Sci Technol A* 25:1275–1284
- Palmquist S (1957) Method of determining the toughness of brittle materials, particularly sintered carbides. *Jernkontorets Ann* 141:300–307

- Patton ST, Bhushan B (1996) Micromechanical and tribological characterization of alternate pole tip materials for magnetic recording heads. *Wear* 202:99–109
- Perry AJ (1981) The adhesion of chemically vapour-deposited hard coatings on steel—The scratch test. *Thin Solid Films* 78:77–93
- Perry AJ (1983) Scratch adhesion testing of hard coatings. *Thin Solid Films* 197:167–180
- Pethica JB, Oliver WC (1989) Mechanical properties of nanometer volumes of material: use of the elastic response of small area indentations. In: Bravman JC, Nix WD, Barnett DM, Smith DA (eds) *Thin films: stresses and mechanical properties*, vol 130. Mat. Res. Soc., Pittsburgh, pp 13–23
- Pethica JB, Hutchings R, Oliver WC (1983) Hardness measurements at penetration depths as small as 20 nm. *Philos Mag A* 48:593–606
- Pharr GM (1992) The anomalous behavior of silicon during nanoindentation. In: Nix WD, Bravman JC, Arzt E, Freund LB (eds) *Thin film: stresses and mechanical properties III*, vol 239. Mater. Res. Soc., Pittsburgh, pp 301–312
- Pharr GM (1998) Measurement of mechanical properties by ultra-low load indentation. *Mat Sci Eng A* 253:151–159
- Pharr GM, Oliver WC, Clarke DR (1989) Hysteresis and discontinuity in the indentation load-displacement behavior of silicon. *Scr Metall* 23:1949–1952
- Pharr GM, Oliver WC, Clarke DR (1990) The mechanical behavior of silicon during small-scale indentation. *J Electron Mater* 19:881–887
- Pharr GM, Oliver WC, Brotzen FR (1992) On the generality of the relationship among contact stiffness, contact area, and elastic modulus during indentation. *J Mater Res* 7:613–617
- Pharr GM, Harding DS, Oliver WC (1993) Measurement of fracture toughness in thin films and small volumes using nanoindentation methods. In: Nastasi M, Parkin DM, Gleiter H (eds) *Mechanical properties and deformation behavior of materials having ultra-fine microstructures*. Kluwer Academic, Dordrecht, pp 449–461
- Randall NX, Cristoph R, Droz S, Julia-Schmutz C (1996) Localised micro-hardness measurements with a combined scanning force microscope/nanoindentation system. *Thin Solid Films* 290–291:348–354
- Raman V, Berriche R (1992) An investigation of the creep processes in tin and aluminum using a depth-sensing indentation technique. *J Mater Res* 7:627–638
- San Juan J, No ML, Schuh CA (2009) Nanoscale shape-memory alloys for ultrahigh mechanical damping. *Nat Nanotechnol* 4:415–419
- Sargent PM (1986) Use of the indentation size effect on microhardness of materials characterization. In: Blau PJ, Lawn BR (eds) *Microindentation techniques in materials science and engineering*, STP 889. ASTM, Philadelphia, pp 160–174
- Scruby CB (1987) An introduction to acoustic emission. *J Phys E Sci Instrum* 20:946–953
- Sekler J, Steinmann PA, Hintermann HE (1988) The scratch test: different critical load determination techniques. *Surf Coat Tech* 36:519–529
- Shih CW, Yang M, Li JCM (1991) Effect of tip radius on nanoindentation. *J Mater Res* 6:2623–2628
- Sneddon IN (1965) The relation between load and penetration in the axisymmetric boussinesq problem for a punch of arbitrary profile. *Int J Eng Sci* 3:47–57
- Stilwell NA, Tabor D (1961) Elastic recovery of conical indentation. *Proc Phys Soc* 78:169–179
- Stone D, LaFontaine WR, Alexopoulos PS, Wu TW, Li CY (1988) An investigation of hardness and adhesion of sputter-deposited aluminum on silicon by utilizing a continuous indentation test. *J Mater Res* 3:141–147
- Sundararajan S, Bhushan B (2001) Development of a continuous microscratch technique in an atomic force microscope and its application to study scratch resistance of ultra-thin hard amorphous carbon coatings. *J Mater Res* 16:437–445
- Sundararajan S, Bhushan B, Namazu T, Isono Y (2002) Mechanical property measurements of nanoscale structures using an atomic force microscope. *Ultramicroscopy* 91:111–118
- Suresh S (1991) Fatigue of materials. Cambridge University Press, Cambridge
- Swadener JG, George EP, Pharr GM (2002) The correlation of the indentation size effect measured with indenters of various shapes. *J Mech Phys Solids* 50:681–694
- Swain MV, Hagan JT, Field JE (1977) Determination of the surface residual stresses in tempered glasses by indentation fracture mechanics. *J Mater Sci* 12:1914–1917
- Syed Asif SA, Pethica JB (1997) Nano scale creep and the role of defects. In: Gerberich WW, Gao H, Sundgren JE, Baker SP (eds) *Thin films: stresses and mechanical properties IV*. Mat Res. Soc. Pittsburgh, PA (MRS Symp Proc 436:201–206)
- Tabor D (1951) The hardness of metals. Clarendon Press, Oxford
- Tabor D (1970) The hardness of solids. *Rev Phys Technol* 1:145–179
- Tangena AG, Hurkx GAM (1986) The determination of stress-strain curves of thin layers using indentation tests. *ASME J Eng Mat Technol* 108:230–232
- Ternovskii AP, Alekhin VP, Shorshorov MKh, Khrushchov MM, Skvortsov VN (1973) Micromechanical testing of materials by depression. *Zavod Lab* 39:1620–1624
- Tsui TY, Pharr GM (1999) Substrate effects on nanoindentation mechanical property measurement of soft films on hard substrates. *J Mater Res* 14:292–301
- Tsui TY, Oliver WC, Pharr GM (1996) Influences of stress on the measurement of mechanical properties using nanoindentation. I. Experimental studies in an aluminum alloy. *J Mater Res* 11:752–759

- Tsukamoto Y, Yamaguchi H, Yanagisawa M (1987) Mechanical properties of thin films: measurements of ultramicroindentation hardness, Young's modulus and internal stresses. *Thin Solid Films* 154:171–181
- Valli J (1986) A review of adhesion test method for thin hard coatings. *J Vac Sci Technol A4*:3007–3014
- VanLandingham MR, Villarrubia JS, Guthrie WF, Meyers GF (2001) Nanoindentation of polymers: an overview. *Macromol Symp* 167:15–43
- Venkataraman S, Kohlstedt DL, Gerberich WW (1992) Microscratch analysis of the work of adhesion for Pt thin films on NiO. *J Mater Res* 1:1126–1132
- Vinci RP, Vlassak JJ (1996) Mechanical behavior of thin films. *Annu Rev Mater Sci* 26:431–462
- Vitovc FH (1986) Stress and load dependence of microindentation hardness. In: Blau PJ, Lawn BR (eds) *Microindentation techniques in materials science and engineering*, STP 889. ASTM, Philadelphia, pp. 175–185
- Walker WW (1973) Indentation creep at low homologous temperatures. In: *The Science of Hardness Testing and its Research Applications* (Westbrook JH, and Conrad H, eds.), pp. 258–273, Am. Soc. Metals, Metals Park, Ohio.
- Wei G, Bhushan B, Ferrell N, Hansford D (2005) Microfabrication and nanomechanical characterization of polymer mems for biological applications. *J Vac Sci Technol A* 23:811–819
- Weihs TP, Lawrence CW, Derby CB, Pethica JB (1992) Acoustic emissions during indentation tests. In: Nix WD, Bravman JC, Arzt E, Freund LB (eds) *Thin films: stresses and mechanical properties III*. Mat. Res. Soc., Pittsburgh (Symp Proc 239: 361–370)
- Westbrook JH (1957) Microhardness testing at high temperatures. *Proc Am Soc Test Mater* 57:873–895
- Westbrook JH, Conrad H (eds) (1973) *The science of hardness and its research applications*. Am Soc Metals, Metals Park
- Whitehead AJ, Page TF (1992) Nanoindentation studies of thin film coated systems. *Thin Solid Films* 220:277–283
- Wierenga PE, Franken AJJ (1984) Ultramicroindentation apparatus for the mechanical characterization of thin films. *J Appl Phys* 55:4244–4247
- Wierenga PE, van der Linden JHM (1986) Quasistatic and dynamic indentation measurements on magnetic tapes. In: Bhushan B, Eiss NS (eds) *Tribology and mechanics of magnetic storage systems*, vol 3 (SP-21). ASLE, Park Ridge, pp 31–37
- Williams VS, Landesman AL, Shack RV, Vukobratovich D, Bhushan B (1988) In situ microviscoelastic measurements by polarization-interferometric monitoring of indentation depth. *Appl Opt* 27:541–546
- Wu TW (1991) Microscratch and load relaxation tests for ultra-thin films. *J Mater Res* 6:407–426
- Wu TW, Hwang C, Lo J, Alexopoulos P (1988) Microhardness and microstructure of ion-beam-sputtered, nitrogen doped nife films. *Thin Solid Films* 166:299–308
- Wu TW, Shull AL, Berriche R (1991) Microindentation fatigue tests on submicron carbon films. *Surf Coat Tech* 47:696–709
- Yanagisawa M, Motomura Y (1987) An ultramicro indentation hardness tester and its application to thin films. *Lub Eng* 43:52–56
- Young WC, Budynas RG, Sadegh AM (2012) *Roark's formulas for stress and strain*, 8th edn. McGraw-Hill, New York

Nanomechanical Properties of Nanostructures and Scale Effects

Bharat Bhushan

Abstract

Structural integrity is of paramount importance in all devices. Load applied during the use of devices can result in component failure. Cracks can develop and propagate under tensile stresses, leading to failure. Knowledge of the mechanical properties of nanostructures is necessary for designing realistic micro/nanoelectromechanical systems (MEMS/NEMS) and biological micro/nanoelectromechanical systems (BioMEMS/Bio-NEMS) devices (Bhushan 2016). Elastic and inelastic properties are needed to predict deformation from an applied load in the elastic and inelastic regimes, respectively. The strength property is needed to predict the allowable operating limit. Some properties of interest are hardness, elastic modulus, bending strength, fracture toughness, and fatigue strength. Many of the mechanical properties are scale dependent. Therefore, these should be measured at relevant scales. Atomic force microscopy and nanoindenters can be used to evaluate the mechanical properties of micro/nanoscale structures. Commonly used materials in MEMS/NEMS are single-crystal silicon and silicon-based materials, e.g., SiO₂ and polysilicon films deposited by low-pressure chemical vapor deposition. Single-crystal SiC deposited on large-area silicon substrates is used for high-temperature micro/nanosensors and actuators. Amorphous alloys can be formed on both metal and silicon substrates by sputtering and plating techniques, providing more flexibility in surface-integration.

B. Bhushan (✉)

Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

Electroless deposited Ni-P amorphous thin films have been used to construct microdevices, especially using the so-called LIGA (Lithographie, Galvanoformung, Abformung) techniques. Micro/nanodevices need conductors to provide power, as well as electrical/magnetic signals to make them functional. Electroplated gold films have found wide application in electronic devices because of their ability to make thin films and process simply. Polymers, such as poly(methyl methacrylate) (PMMA), poly(dimethylsiloxane) (PDMS), and polystyrene are commonly used in BioMEMS/BioNEMS such as micro/nanofluidic devices because of the ease of manufacturing and reduced cost. Many polymers are biocompatible, so they may be integrated into biomedical devices. This chapter presents a review of mechanical property measurements on the micro/nanoscale of various materials of interest and stress and deformation analyses of nanostructures.

Keywords

Mechanical properties • Nanostructures • Cellular mechanics • Scale effects • Hardness • Elastic modulus • Bending strength • Fracture toughness • Fatigue • Creep

6.1 Introduction

Microelectromechanical systems (MEMS) refer to microscopic devices that have a characteristic length of less than 1 mm but more than 100 nm (or 1 μm), and nanoelectromechanical systems (NEMS) refer to nanoscopic devices that have a characteristic length of less than 100 nm (or 1 μm). These are referred to as an intelligent miniaturized system comprising of sensing, processing, and/or actuating functions and combine electrical and mechanical components. The acronym MEMS originated in the U.S.A. The term commonly used in Europe is micro/nanosystem technology (MST/NST), and in Japan it is micro/nanomachines. Another term generally used is micro/nanodevices. MEMS/NEMS terms are also now used in a broad sense and include electrical, mechanical, optical, biological, and/or fluidic functions. To put the dimensions in perspective, individual atoms are typically a fraction of a nanometer in diameter, DNA molecules are about 2.5 nm wide, biological cells are in the range of thousands of nm in

diameter, and human hair is about 75 μm in diameter. The mass of a micromachined silicon structure can be as low as 1 nN, and NEMS can be built with mass as low as 10^{-20} N with cross sections of about 10 nm. In comparison, the mass of an eyelash is about 100 nN, and the mass of a drop of water is about 10 μN .

A wide variety of MEMS, including Si-based devices, chemical and biological sensors and actuators, and miniature non-silicon structures (e.g., devices made from plastics or ceramics) have been fabricated with dimensions in the range of a couple to a few thousand microns (see e.g., Fujimasa 1996; Trimmer 1997; Bhushan 1998; Kovacs 1998; Senturia 2000; Elwenspoek and Wiegerink 2001; Gad-el-Hak 2002; Hsu 2002; Hierlemann 2005; Madou 2011; Bhushan 2016). A variety of NEMS have also been produced (see e.g., Drexler 1992; Timp 1999; Dresselhaus et al. 2001; Rietman 2001; Goddard et al. 2002; Nalwa 2002; Poole and Owens 2003). MEMS/NEMS technology and fabrication processes have found a variety of applications in biology and biomedicine, leading to the

establishment of an entirely new field known as BioMEMS/BioNEMS (Manz and Becker 1998; Cheng and Kricka 2001; Heller and Guttman 2001; San and Yap 2001; Mastrangelo and Becker 2001; Becker and Locascio 2002; Beebe et al. 2002; van der Berg 2003). The ability to use micro/nanofabrication processes to develop precision devices that can interface with biological environments at the cellular and molecular level has led to advances in the fields of biosensor technology (Zoval and Madou 2000; Raiteri et al. 2001; Tang and Lee 2001; Taylor et al. 2003; van der Berg 2003), drug delivery (Park 1997; Shawgo et al. 2002; Oeberg et al. 2004; Decuzzi et al. 2005), and tissue engineering (Bhatia and Chen 1999; Lanza et al. 2000; Leclerc et al. 2004). The miniaturization of fluidic systems using micro/nanofabrication techniques has led to new and more efficient devices for medical diagnostics and biochemical analysis (Schulte et al. 2002). The largest industrial applications of MEMS include accelerometers, pressure sensors, thermal and piezoelectric type inkjet printer heads, and digital micromirror devices. BioMEMS and BioNEMS are increasingly used in commercial applications. The largest applications of BioMEMS and BioNEMS include a variety of biosensors and drug delivery devices.

Structural integrity is of paramount importance in all devices. Load applied during the use of devices can result in component failure. Cracks can develop and propagate under tensile stresses leading to failure. Friction/stiction and wear limit the lifetimes and compromise the performance and reliability of devices involving relative motion (Bhushan 1998, 2001, 2013a, b). Most MEMS/NEMS applications demand extreme reliability. Stress and deformation analyses are carried out for an optimal design. MEMS/NEMS designers require information about mechanical properties at the nanoscale. Elastic and inelastic properties are needed to predict deformation from an applied load in the elastic and inelastic regimes, respectively. Understanding of the strength property is needed to predict the allowable operating limit. Some of

the properties of interest are hardness, elastic modulus, creep, bending strength (fracture stress), fracture toughness, and fatigue strength. Micro/nanostructures have some surface topography and local scratches dependent upon the manufacturing process. Surface roughness and local scratches may compromise the reliability of the devices, and their effect needs to be studied.

Most mechanical properties are scale dependent (Bhushan 2013a, b). Several researchers have measured the mechanical properties of silicon and silicon-based milli- to microscale structures. These include tensile tests and bending tests (Johansson et al. 1988; Ericsson and Schweitz 1990; Obermeier 1996; Wilson et al. 1996; Sharpe et al. 1997; Sato et al. 1998; Greek et al. 1999; LaVan and Buchheit 1999; Mazza and Dual 1999; Yi and Kim 1999), resonant structure tests for the measurement of elastic properties (Kahn et al. 1998), fracture toughness tests (Johansson et al. 1989; Ericson and Schweitz 1990; Ballarini et al. 1997; Wilson et al. 1996; Kahn et al. 1999; Fitzgerald et al. 2000; Tsuchiya et al. 2000), and fatigue tests (Connally and Brown 1993; Komai et al. 1998; Kahn et al. 1999). Nanomechanical properties of nanoscale structures have been measured using atomic force microscopy (AFM) (Namazu et al. 2000; Sundararajan and Bhushan 2002) and nanoin-dentation (Li and Bhushan 2003; Li et al. 2003; Wei et al. 2005; Palacio et al. 2007a). For stress and deformation analyses of simple geometries and boundary conditions, analytical models can be used. For analysis of complex geometries, numerical models are needed. Conventional finite element method (FEM) can be used down to a few tens of nanometer scale, although its applicability is questionable at nanoscale. FEM has been used for simulation and prediction of residual stresses and strains induced in MEMS devices during fabrication (Hsu and Sun 1998), to perform fault analysis in order to study MEMS faulty behavior (Kolpekar et al. 1998), to compute mechanical strain resulting from doping of silicon (Rueda and Law 1998), analyze micromechanical experimental data (Heinzelmann and Petzold 1994; Wilson and Beck 1996;

Wilson et al. 1996), and nanomechanical experimental data (Sundararajan and Bhushan 2002). FEM analysis of nanostructures has been performed to analyze the effect of types of surface roughness and scratches on stresses in nanostructures (Bhushan and Agrawal 2002, 2003).

Commonly used materials for MEMS/NEMS are single-crystal silicon and silicon-based materials (e.g., SiO₂ and polysilicon films deposited by low pressure chemical vapor deposition (LPCVD) process) (Madou 2011). An early study showed silicon to be a mechanically resilient material, in addition to its favorable electronic properties (Petersen 1982). Single-crystal 3C-SiC (cubic or β -SiC) films, deposited by atmospheric pressure chemical vapor deposition (APCVD) process on large-area silicon substrates, are produced for high-temperature micro/nanosensor and actuator applications (Bhushan et al. 1998; Sundararajan and Bhushan 1998; Li and Bhushan 1999). Amorphous alloys can be formed on both metal and silicon substrates by sputtering and plating techniques, providing more flexibilities in surface-integration. Electroless deposited Ni-P amorphous thin films have been used to construct microdevices, especially using the so-called LIGA techniques (Li et al. 2003; Madou 2011). Micro/nanodevices need conductors to provide power, as well as electrical/magnetic signals to make them functional. Electroplated gold films have found wide applications in electronic devices because of their ability to make thin films and process simplicity (Li et al. 2003).

As the field of MEMS/NEMS has progressed, alternative materials, especially polymers, have established an important role in their construction. This trend has been driven by the reduced cost associated with polymer materials. Polymer microfabrication processes, including micro-molding and hot embossing techniques (Becker and Gaertner 2000), can be orders of magnitude less expensive than traditional silicon photolithography processes. Many polymers are biocompatible, so these can be integrated into biomedical devices with minimal detrimental effects to the host or on the biofluids. An improvement in device functionality (relative to

silicon) is also possible due to the mechanical properties of the polymer used, which have properties much closer to biological tissues. Polymer BioMEMS structures involving microbeams have been designed to measure cellular forces (Wei et al. 2005; Palacio et al. 2007a). Polymer materials most commonly used for biomedical applications include poly(methyl methacrylate) (PMMA), poly(dimethylsiloxane) (PDMS), and polystyrene (Becker and Gaertner 2000; McDonald et al. 2000; Palacio et al. 2007a, b). Another material of interest due to ease of fabrication is poly(propyl methacrylate) (PPMA), which has lower glass transition temperature (T_g) (35–43 °C) (Ellis 2000) than PMMA (104–106 °C) (Brandrup et al. 1999; Mark 1999), permitting low temperature thermal processing (Wei et al. 2005; Palacio et al. 2007a, b).

This chapter presents a review of mechanical property measurements on the nanoscale of various materials of interest and stress and deformation analyses of nanostructures.

6.2 Experimental Techniques for Measurement of Mechanical Properties of Nanostructures

6.2.1 Indentation and Scratch Tests Using Micro/Nanoindenters

A nanoindenter is commonly used to measure hardness, elastic modulus, and fracture toughness, and to perform micro/nanoscratch studies to get a measure of scratch/wear resistance of materials (Bhushan and Li 2003).

6.2.1.1 Hardness and Elastic Modulus

The nanoindenter monitors and records the dynamic load and displacement of the three-sided pyramidal diamond (Berkovich) indenter during indentation with a force resolution of about 75 nN and displacement resolution of about 0.1 nm. Hardness and elastic modulus are calculated from the load-displacement data (Bhushan and Li 2003). The peak indentation load depends on the mechanical properties of the

specimen; a harder material requires a higher load for reasonable indentation depth.

6.2.1.2 Fracture Toughness

The indentation technique for fracture toughness measurement of brittle samples on the microscale is based on the measurement of the lengths of median-radial cracks produced by indentation. A Vickers indenter (a four-sided diamond pyramid) is used in a microhardness tester. A load on the order of 0.5 N is typically used in making the Vickers indentations. The indentation impressions are examined using an optical microscope with Nomarski interference contrast to measure the length of median-radial cracks, c . The fracture toughness (K_{IC}) is calculated by the following relation (Lawn et al. 1980),

$$K_{IC} = \alpha \left(\frac{E}{H} \right)^{1/2} \left(\frac{P}{c^{3/2}} \right) \quad (6.1)$$

where α is an empirical constant depending on the geometry of the indenter, H and E are hardness and elastic moduli, and P is the peak indentation load. For Vickers indenters, α has been empirically found based on experimental data and is equal to 0.016 (Bhushan and Li 2003). Both E and H values are obtained from the nanoindentation data. The crack length is measured from the center of the indent to the end of crack using an optical microscope. For one indent, all crack lengths are measured. The crack length c is obtained from the average value of several indents.

6.2.1.3 Indentation Creep

For indentation creep test of polymer samples, the test is performed using a continuous stiffness measurements (CSM) technique (Bhushan and Li 2003). In a study by Wei et al. (2005), the indentation load was typically 30 μN , and the loading rate was 3 $\mu\text{N}/\text{s}$. The tip was held typically for 600 s after the indentation load reached 30 μN . To measure the mean stress and contact stiffness, during the hold segment the indenter was oscillated at a peak-to-peak load amplitude of 1.2 μN and a frequency of 45 Hz.

6.2.1.4 Scratch Resistance

In micro/nanoscratch studies, in a nanoindenter, a conical diamond indenter having a tip radius of about 1 μm and an included angle of 60° is drawn over the sample surface, and the load is ramped up until substantial damage occurs (Bhushan and Li 2003). The coefficient of friction is monitored during scratching. In order to obtain scratch depths during scratching, the surface profile of the sample surface is first obtained by translating the sample at a low load of about 0.2 mN, which is insufficient to damage a hard sample surface. The 500 μm long scratches are made by translating the sample while ramping the loads on the conical tip over different loads dependent upon the material hardness. The actual depth during scratching is obtained by subtracting the initial profile from the scratch depth measured during scratching. In order to measure the scratch depth after the scratch, the scratched surface is profiled at a low load of 0.2 mN and is subtracted from the actual surface profile before scratching.

6.2.2 Bending Tests of Nanostructures Using an AFM

Quasi-static bending tests of fixed nanobeam arrays in the normal direction are carried out using an AFM (Sundararajan et al. 2002; Sundararajan and Bhushan 2002). A three-sided pyramidal diamond tip (with a radius of about 200 nm) mounted on a rectangular stainless steel cantilever is used for the bending tests. The beam stiffness is selected based on the desired load range. The stiffness of the cantilever beams for application of a normal load up to 100 μN is about 150–200 N/m.

For the bending test, the tip is brought over the nanobeam array with the help of the sample stage of the AFM and a built-in high magnification optical microscope (Fig. 6.1) (Sundararajan and Bhushan 2002). For the fine positioning of the tip over a chosen beam, the array is scanned in contact mode at a contact load of

about 2–4 μN , which results in negligible damage to the sample. After scanning, the tip is located at one end of a chosen beam. To position the tip at the center of the beam span, the tip is moved to the other end of the beam by giving the X-piezo an offset voltage. The value of this offset is determined after several such attempts have been made in order to minimize the effects of piezo drift. Half of this offset is then applied to the X-piezo after the tip is positioned at one end of the beam, which usually results in the tip being moved to the center of the span. Once the tip is positioned over the center of the beam span, the tip is held stationary without scanning, and the Z-piezo is extended by a known distance, typically about 2.5 μm , at a rate of 10 nm/s, as shown in Fig. 6.1. During this time, the vertical deflection signal (dV_{AFM}), which is proportional to the deflection of the cantilever (D_{tip}), is monitored. The displacement of the piezo is equal to the sum of the displacements of the

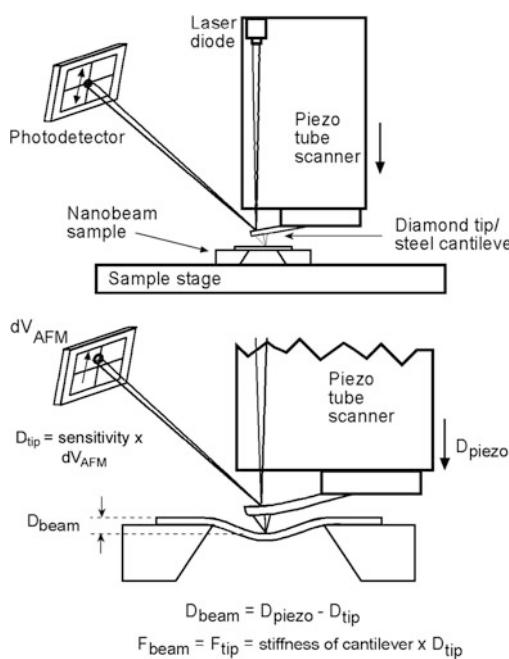


Fig. 6.1 Schematic showing the details of a nanoscale bending test using an AFM. The AFM tip is brought to the center of the nanobeam and the piezo is extended over a known distance. By measuring the tip displacement, a load displacement curve of the nanobeam can be obtained (Sundararajan and Bhushan 2002)

cantilever and the nanobeam. Hence the displacement of the nanobeam (D_{beam}) under the point of load can be determined as

$$D_{\text{beam}} = D_{\text{piezo}} - D_{\text{tip}} \quad (6.2)$$

The load (F_{beam}) on the nanobeam is the same as the load on the tip/cantilever (F_{tip}) and is given by

$$F_{\text{beam}} = F_{\text{tip}} = D_{\text{tip}} \times k \quad (6.3)$$

where k is the stiffness of the tip/cantilever. In this manner, a load displacement curve for each nanobeam can be obtained.

The photodetector sensitivity of the cantilever needs to be calibrated to obtain D_{tip} in nm. For this calibration, the tip is pushed against a smooth diamond sample by moving the Z-piezo over a known distance. For the hard diamond material, the actual deflection of the tip can be assumed to be the same as the Z-piezo travel (D_{piezo}), and the photodetector sensitivity (S) for the cantilever setup is determined as

$$S = D_{\text{piezo}} / dV_{\text{AFM}} \text{ nm/V} \quad (6.4)$$

In the measurements, D_{tip} is given as $dV_{\text{AFM}} \times S$.

Since a sharp tip would result in an undesirable large local indentation, Sundararajan and Bhushan (2002) used a diamond tip which was worn (blunt). Indentation experiments using this tip on a silicon substrate yielded a residual depth of less than 8 nm at a maximum load of 120 μN , which is negligible compared to displacements of the beams (several hundred nm). Hence we can assume that negligible local indentation or damage is created during the bending process of the beams and that the displacement calculated from Eq. 6.2 is from the beam structure.

6.2.2.1 Elastic Modulus and Bending Strength

Elastic modulus and bending strength (fracture stress) of the beams can be estimated by equations based on the assumption that the beams follow the linear elastic theory of an isotropic material. This is probably valid since the beams

have high length-to-width ℓ/w and length-to-thickness ℓ/t ratios and also since the length direction is along the principal stress direction during the test. For a fixed elastic beam loaded at the center of the span, the elastic modulus is expressed as

$$E = \frac{\ell^3}{192I} m \quad (6.5)$$

where ℓ is the beam length, I is the area moment of inertia for the beam cross-section, and m is the slope of the linear region of the load-displacement curve during bending (Young and Budynas 2002). The area moment of inertia for a beam with a trapezoidal cross section is calculated from the following equation

$$I = \frac{4w_1w_2 + w_1^2 + w_2^2}{36(w_1 + w_2)} t^3 \quad (6.6)$$

where w_1 and w_2 are the upper and lower widths, respectively, and t is the thickness of the beam. According to linear elastic theory, for a centrally loaded beam, the moment diagram is shown in Fig. 6.2. The maximum moments are generated at the ends (negative moment) and under the loading point (positive moment) as shown in Fig. 6.2. The bending stresses generated in the beam are proportional to the moments and are compressive or tensile about the neutral axis (line of zero stress). The maximum tensile stress (σ_b , which is the bending strength or fracture stress) is produced on the top surface at both the ends and is given by Young and Budynas (2002)

$$\sigma_b = \frac{F_{max}\ell e_1}{8I} \quad (6.7)$$

where F_{max} is the applied load at failure, and e_1 is the distance of the top surface from the neutral plane of the beam cross-section and is given by Young and Budynas (2002)

$$e_1 = \frac{t(w_1 + 2w_2)}{3(w_1 + w_2)} \quad (6.8)$$

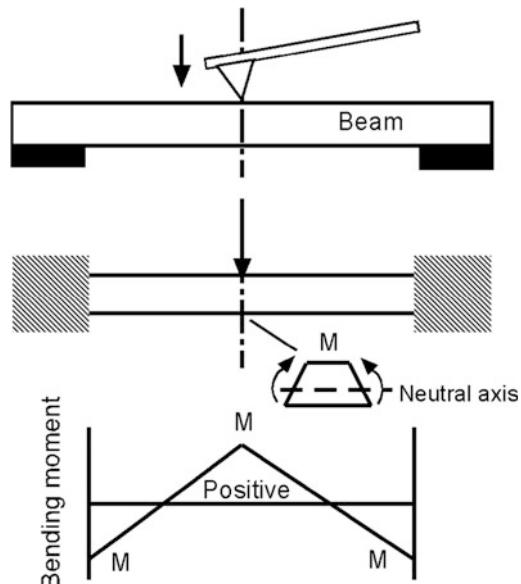
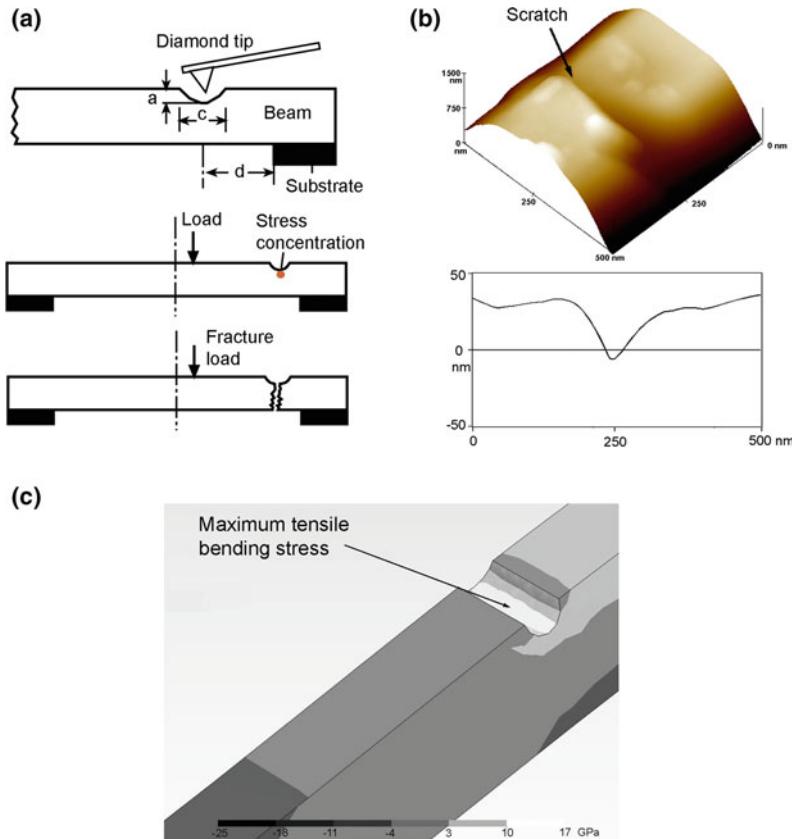


Fig. 6.2 Schematic of the bending moments generated in the beam during a quasi-static bending experiment, with the load at the center of the span. The maximum moments occur under the load and at the fixed ends. Due to the trapezoidal cross section, the maximum tensile bending stresses occur at the top surfaces at the fixed ends

Although the moment value at the center of the beam is the same as at the ends, the tensile stresses at the center (generated on the bottom surface) are less than those generated at the ends (per Eq. 6.7) because the distance from the neutral axis to the bottom surface is less than e_1 . This is because of the trapezoidal cross section of the beam, which results in the neutral axis being closer to the bottom surface than the top (Fig. 6.2).

In the preceding analysis, the beams were assumed to have fixed ends. However, in the nanobeams used by Sundararajan and Bhushan (2002), the underside of the beams was pinned over some distance on either side of the span. Hence a finite element model of the beams was created to see if the difference in the boundary conditions affected the stresses and displacements of the beams. It was found that the difference in the stresses was less than 1%. This indicates that the boundary conditions near the ends of the actual beams are not that different

Fig. 6.3 **a** Schematic of technique to generate a defect (crack) of known dimensions in order to estimate fracture toughness. A diamond tip is used to generate a scratch across the width of the beam. When the beam is loaded as shown, a stress concentration is formed at the bottom of the scratch. The fracture load is then used to evaluate the stresses using FEM. **b** AFM 3-D image and 2-D profile of a typical scratch. **c** Finite element model results verifying that the maximum bending stress occurs at the bottom of the scratch (Sundararajan and Bhushan 2002)



from that of fixed ends. Therefore the bending strength values can be calculated from Eq. 6.7.

6.2.2.2 Fracture Toughness

Fracture toughness is another important parameter for brittle materials such as silicon. In the case of the nanobeam arrays, these are not best suited for fracture toughness measurements because they do not possess regions of uniform stress during bending. Sundararajan and Bhushan (2002) developed a methodology, and its steps are outlined schematically in Fig. 6.3a. First, a crack of known geometry is introduced in the region of maximum tensile bending stress, i.e. on the top surface near the ends of the beam. This is achieved by generating a scratch at high normal load across the width (w_1) of the beam using a sharp diamond tip (radius < 100 nm). A typical scratch thus generated is shown in Fig. 6.3b. By bending the beam as shown, a stress concentration will be formed under the scratch. This will

lead to failure of the beam under the scratch once a critical load (fracture load) is attained. The fracture load and relevant dimensions of the scratch are input into the FEM model, which is used to generate the fracture stress plots. Figure 6.3c shows an FEM simulation of one such experiment, which reveals that the maximum stress does occur under the scratch.

If we assume that the scratch tip acts as a crack tip, a bending stress will tend to open the crack in Mode I. In this case, the stress field around the crack tip can be described by the stress intensity parameter K_I (for Mode I) for linear elastic materials (Hertzberg 1989). In particular the stresses corresponding to the bending stresses are described by

$$\sigma = \frac{K_I}{\sqrt{2\pi r}} \cos\left(\frac{\theta}{2}\right) \left[1 + \sin\left(\frac{\theta}{2}\right) \sin\left(\frac{3\theta}{2}\right) \right] \quad (6.9)$$

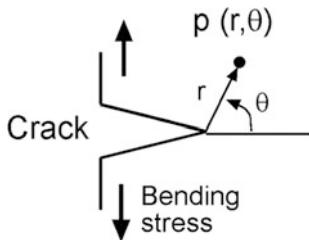


Fig. 6.4 Schematic of crack tip and coordinate systems used in Eq. 6.9 to describe a stress field around the crack tip in terms of the stress intensity parameter, K_I (Sundararajan and Bhushan 2002)

for every point $p(r, \theta)$ around the crack tip as shown in Fig. 6.4. If we substitute the fracture stress (σ_f) into the left hand side of Eq. 6.9, then the K_I value can be substituted by its critical value, which is the fracture toughness K_{IC} . Now, the fracture stress can be determined for the point ($r = 0, \theta = 0$), i.e. right under the crack tip as explained above. However, we cannot substitute $r = 0$ in Eq. 6.9. The alternative is to substitute a value for r which is as close to zero as possible. For silicon, a reasonable number is the distance between neighboring atoms in the (111) plane, the plane along which silicon exhibits the lowest fracture energy. This value was calculated from silicon unit cell dimensions of 0.5431 nm (Anonymous 1988) to be 0.4 nm (half of the face diagonal). This assumes that Si displays no plastic zone around the crack tip, which is reasonable since in tension, silicon is not known to display much plastic deformation at room temperature. Sundararajan and Bhushan (2002) used values $r = 0.4$ to 1.6 nm (i.e. distances up to 4 times the distance between the nearest neighboring atoms) to estimate the fracture toughness for both Si and SiO_2 according to the following equation

$$K_{IC} = \sigma_f \sqrt{2\pi r} \quad r = 0.4 \text{ to } 1.6 \text{ nm.} \quad (6.10)$$

6.2.2.3 Fatigue Strength

In addition to the properties mentioned so far that can be evaluated from quasi-static bending tests, the fatigue properties of nanostructures are also of interest. This is especially true for MEMS/NEMS

involving vibrating structures such as oscillators and comb drives (Nguyen and Howe 1999) and hinges in digital micromirror devices (Hornbeck 1999). To study the fatigue properties of the nanobeams, Sundararajan and Bhushan (2002) applied monotonic cyclic stresses using an AFM, Fig. 6.5a. Similar to the bending test, the diamond tip is first positioned at the center of the beam span. In order to ensure that the tip is always in contact with the beam (as opposed to impacting it), the piezo is first extended by a distance D_1 , which ensures a minimum stress on the beam. After this extension, a cyclic displacement of amplitude D_2 is applied continuously until failure of the beam occurs. This results in the application of a cyclic load to the beam. The maximum frequency of the cyclic load that could be attained using the AFM by Sundararajan and Bhushan (2002) was 4.2 Hz. The vertical deflection signal of the tip is monitored throughout the experiment. The signal follows the pattern of the piezo input up to failure, which is indicated by a sudden drop in the signal. During initial runs, piezo drift was observed that caused the piezo to gradually move away from the beam (i.e. to retract), resulting in a continuous decrease in the applied normal load. In order to compensate for this, the piezo is given a finite extension of 75 nm every 300 s as shown in Fig. 6.5a. This results in keeping the applied loads fairly constant. The normal load variation (calculated from the vertical deflection signal) from a fatigue test is shown in Fig. 6.5b. The cyclic stress amplitudes (corresponding to D_2) and fatigue lives are recorded for every sample tested. Values for D_1 are set such that minimum stress levels are about 20% of the bending strengths.

6.2.3 Bending Tests of Micro/Nanostructures Using a Nanoindenter

Quasi-static bending tests of micro/nanostructures in both normal and lateral directions are carried out using a nanoindenter (Fig. 6.6) (Li and Bhushan 2003; Li et al. 2003; Wei et al.

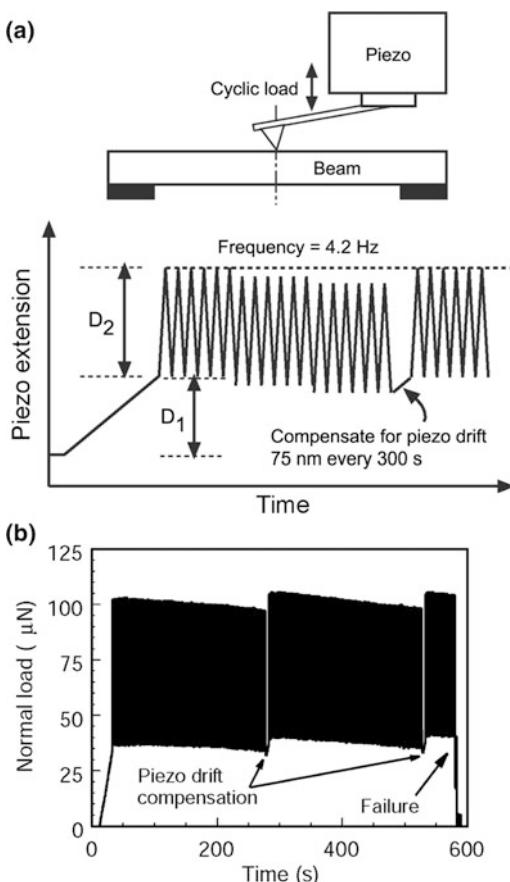


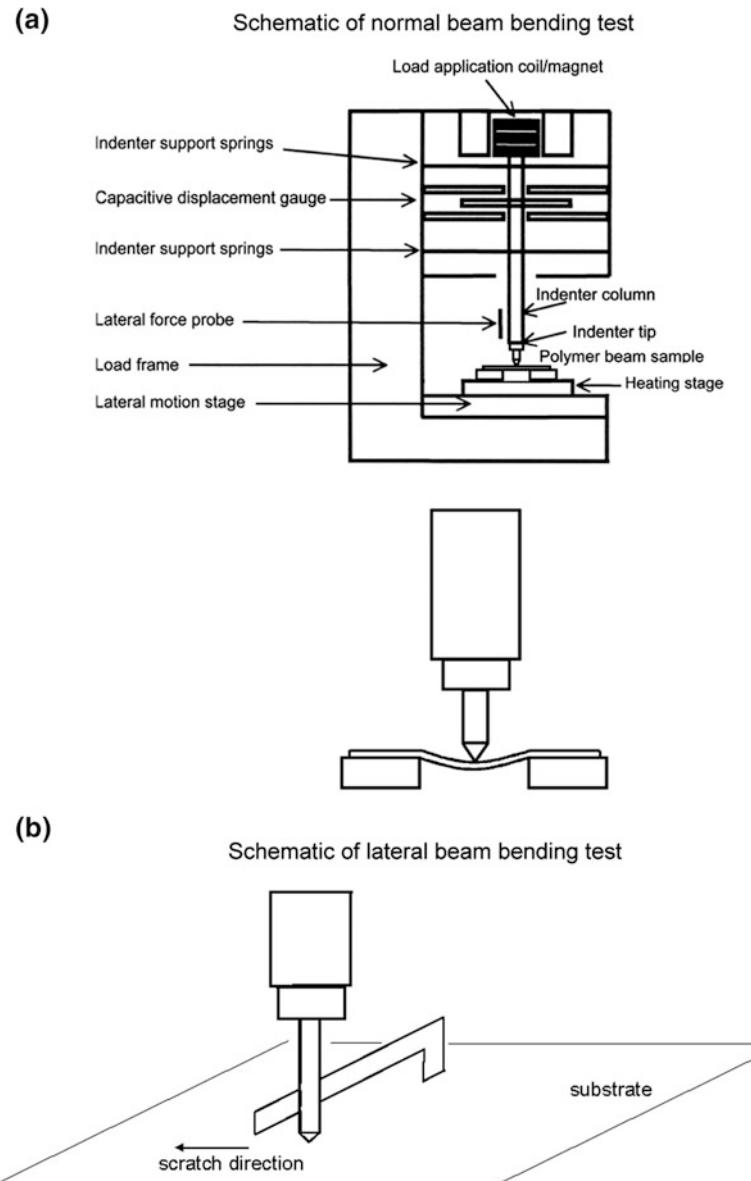
Fig. 6.5 **a** Schematic showing the details of the technique to study fatigue behavior of the nanobeams. The diamond tip is located at the middle of the span and a cyclic load at 4.2 Hz is applied to the beam by forcing the piezo to move in the pattern shown. An extension is made every 300 s to compensate for the piezo drift to ensure that the load on the beam is kept fairly constant. **b** Data from a fatigue experiment on a nanobeam until failure. The normal load is computed from the raw vertical deflection signal. The compensations for piezo drift keep the load fairly constant (Sundararajan and Bhushan 2002)

2005; Palacio et al. 2007a). The advantage of a nanoindenter is that loads up to about 400 mN, higher than that in AFM (up to about 100 μN), can be used for structures requiring high loads for experiments. Bending experiments in the normal direction have been carried out on the suspended beams using the nanoindenter fitted with a conical tip that has a 1 μm radius of curvature and a 60° included angle. Figure 6.6a

shows the schematic of the beam bending experimental setup. To avoid the indenter tip pushing into the specimen, a blunt tip is used in the bending and fatigue tests. For ceramic and metallic beam samples, Li et al. (2003) used a diamond conical indenter with a radius of 1 μm and an included angle of 60°. For polymer beam samples, Wei et al. (2005) and Palacio et al. (2007a) reported that the diamond tip penetrated the polymer beams easily and caused considerable plastic deformation during the bending test, which led to significant errors in the measurements. To avoid this issue, the diamond tip was dip-coated with PMMA (about 1–2 μm thick) by dipping the tip in the 2% PMMA (wt/wt) solution for about 5 s. Load position used was at the center of the span for the bridge beams and at 10 μm off from the free end of cantilever beams. An optical microscope with a magnification of 1500x or an in-situ AFM is used to locate the loading position. Then the specimen is moved under the indenter location with a resolution of about 200 nm in the longitudinal direction and less than 100 nm in the lateral direction. Using the analysis presented earlier, elastic modulus and bending strength of the beams can be obtained from the load displacement curves (Li et al. 2003; Wei et al. 2005; Palacio et al. 2007a). For fatigue tests, an oscillating load is applied and contact stiffness is measured during the tests. A significant drop in the contact stiffness during the test is a measure of the number of cycles to failure (Li and Bhushan 2003).

A schematic of the lateral bending experimental setup is shown in Fig. 6.6b. The nanoindenter was fitted with a conical tip which has a 1 μm radius of curvature and a 60° included angle. The tip was positioned at 200–300 μm away from the beam anchor and 20 μm from the edge of the beam length. The indenter was programmed to perform a scratch at a constant load of 400 μN , and the stage was moved at a rate of 10 $\mu\text{m}/\text{s}$. Since the lateral force contribution of the substrate surface is constant, any additional force recorded comes from the cantilever beam response. From the linear elastic regime in the

Fig. 6.6 Schematic of micro/nanoscale bending experimental setup using a nanoindenter **a** in the normal direction, and **b** in the lateral direction



measured lateral force as a function of lateral displacement, the elastic modulus can be evaluated. For a cantilever beam with one end clamped, it is expressed as (Young and Budynas 2002),

$$E = \frac{\ell^3}{3I} m \quad (6.11)$$

where, as described earlier, ℓ is the beam length, I the area moment of inertia for the beam cross section, and m is the slope of the linear region of the force displacement curve.

6.3 Experimental Results and Discussion

6.3.1 Indentation and Scratch Tests of Various Ceramic and Metals Using Micro/Nanoindenter

Studies have been conducted on six different materials: undoped single-crystal Si(100), undoped polysilicon film, SiO₂ film, SiC film,

electroless deposited Ni-11.5 wt% P amorphous film, and electroplated Au film (Sundararajan and Bhushan 1998; Li and Bhushan 1999; Li et al. 2003). A 3- μm thick polysilicon film was deposited by a low pressure chemical vapor deposition (LPCVD) process on an Si(100) substrate. The 1- μm thick SiO_2 film was deposited by a plasma enhanced chemical vapor deposition (PECVD) process on a Si(111) substrate. A 3- μm thick 3C-SiC film was epitaxially grown using an atmospheric pressure chemical vapor deposition (APCVD) process on Si(100) substrate. A 12- μm thick Ni-P film was electroless plated on a 0.8 mm thick Al-4.5 wt% Mg alloy substrate. A 3- μm thick Au film was electroplated on an Si(100) substrate.

6.3.1.1 Hardness and Elastic Modulus

Hardness and elastic modulus measurements are made using a nanoindenter (Li et al. 2003). The hardness and elastic modulus values of various materials at a peak indentation depth of 50 nm are summarized in Fig. 6.7 and Table 6.1. The SiC film exhibits the highest hardness of about 25 GPa, an elastic modulus of about 395 GPa among the samples examined, followed by the undoped Si(100), undoped polysilicon film, SiO_2 film, Ni-P film, and Au film. The hardness and elastic modulus data of the undoped Si(100) and undoped polysilicon film are comparable. For the metal alloy films, the Ni-P film exhibits higher hardness and elastic modulus than the Au film.

6.3.1.2 Fracture Toughness

The optical images of Vickers indentations made using a microindenter at a normal load of 0.5 N held for 15 s on the undoped Si(100), undoped polysilicon film, and SiC film are shown in Fig. 6.8 (Li and Bhushan 1999). The SiC film exhibits the smallest indentation mark, followed by the undoped polysilicon film and undoped Si(100). These Vickers indentation depths are smaller than one-third of the film thickness. Thus, the influence of substrate on the fracture toughness of the films can be ignored. In addition to the indentation marks, radial cracks are observed, emanating from the indentation corners. The SiC film shows the longest radial

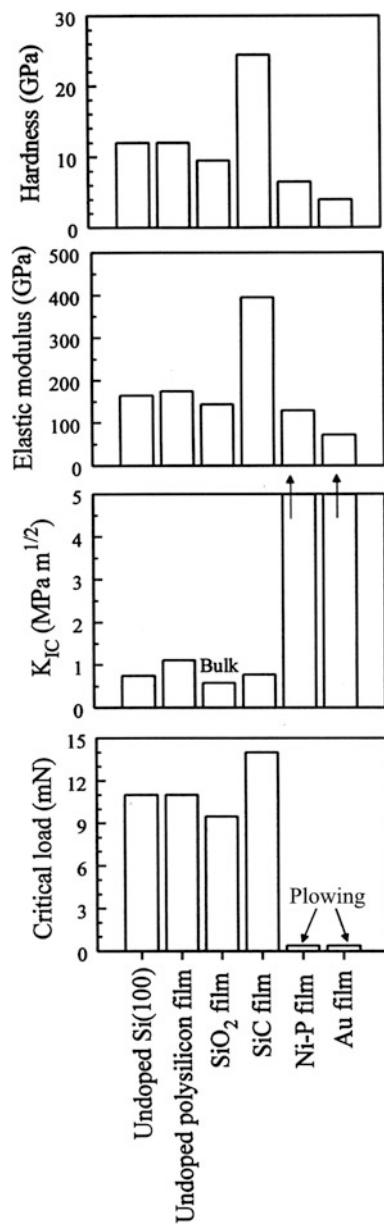


Fig. 6.7 Bar chart summarizing the hardness, elastic modulus, fracture toughness, and critical load (from scratch tests) results of the bulk undoped single-crystal Si(100) and thin films of undoped polysilicon, SiO_2 , SiC, Ni-P, and Au (Li et al. 2003)

crack length, followed by the undoped Si(100) and undoped polysilicon film. The radial cracks for the undoped Si(100) are straight whereas those for the SiC, undoped polysilicon film are not straight but go in a zigzag manner. The

Table 6.1 Hardness, elastic modulus, fracture toughness, and critical load results of the bulk single-crystal Si(100) and thin films of undoped polysilicon, SiO₂, SiC, Ni-P, and Au

Samples	Hardness (GPa)	Elastic modulus (GPa)	Fracture toughness (MPa m ^{1/2})	Critical load (mN)
Undoped Si(100)	12	165	0.75	11
Undoped polysilicon film	12	167	1.11	11
SiO ₂ film	9.5	144	0.58 (Bulk)	9.5
SiC film	24.5	395	0.78	14
Ni-P film	6.5	130		0.4 (Plowing)
Au film	4	72		0.4 (Plowing)

fracture toughness (K_{IC}) is calculated using Eq. 6.1.

The fracture toughness values of all samples are summarized in Fig. 6.7 and Table 6.1. The SiO₂ film used in this study is about 1 μm thick, which is not thick enough for fracture toughness measurement. The fracture toughness value of bulk silica is listed instead for a reference. The Ni-P and Au films exhibit very high fracture toughness values that cannot be measured by indentation methods. For other samples, the undoped polysilicon film has the highest value, followed by the undoped Si(100), SiC film, and SiO₂ film. For the undoped polysilicon film, the grain boundaries can stop the radial cracks and change the propagation directions of the radial cracks, making the propagation of these cracks more difficult. Values of fracture toughness for the undoped Si(100) and SiC film are comparable. Since the undoped Si(100) and SiC film are single crystal, no grain boundaries are present to stop the radial cracks and change the propagation directions of the radial cracks. This is why the SiC film shows a lower fracture toughness value than the bulk polycrystal SiC materials of 3.6 MPa m^{1/2} (Tanaka 1996).

6.3.1.3 Scratch Resistance

Scratch resistance of various materials have been studied using a nanoindenter by Li et al. (2003). Figure 6.9 compares the coefficient of friction and scratch depth profiles as a function of increasing normal load and optical images of three regions over scratches: at the beginning of

the scratch (indicated by A on the friction profile), at the point of initiation of damage at which the coefficient of friction increases to a high value or increases abruptly (indicated by B on the friction profile), and towards the end of the scratch (indicated by C on the friction profile) for all samples. Note that the ramp loads for the Ni-P and Au range are from 0.2 to 5 mN whereas the ramp loads for other samples are from 0.2 to 20 mN. All samples exhibit a continuous increase in the coefficient of friction with increasing normal load from the beginning of the scratch. The continuous increase in the coefficient of friction during scratching is attributed to the increasing plowing of the sample by the tip with increasing normal load, as shown in the scanning electron microscopy (SEM) images in Fig. 6.9. The abrupt increase in the coefficient of friction is associated with catastrophic failure as well as significant plowing of the tip into the sample. Before the critical load, the coefficient of friction of the undoped polysilicon, SiC, and SiO₂ films increased at a slower rate, and was smoother than that of the other samples. The undoped Si(100) exhibits some bursts in the friction profiles before the critical load. At the critical load, the SiC and undoped polysilicon films exhibit a small increase in the coefficient of friction whereas the undoped Si(100) and undoped polysilicon film exhibit a sudden increase in the coefficient of friction. The Ni-P and Au films show a continuous increase in the coefficient of friction, indicating the behavior of a ductile metal. The bursts in the friction profile

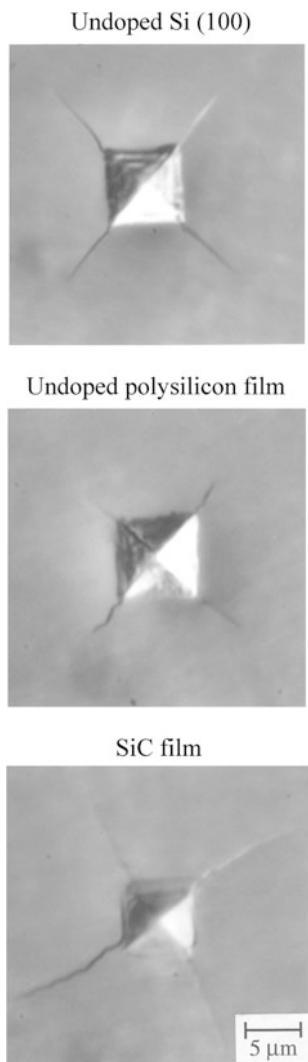


Fig. 6.8 Optical images of Vickers indentations made at a normal load of 0.5 N held for 15 s on the undoped Si(100), undoped polysilicon film, and SiC film (Li and Bhushan 1999)

might result from the plastic deformation and material pile-up in front of the scratch tip. The Au film exhibits a higher coefficient of friction than the Ni-P film. This is because the Au film has lower hardness and elastic modulus values than the Ni-P film.

The SEM images show that below the critical loads the undoped Si(100) and undoped polysilicon film were damaged by plowing, associated with the plastic flow of the material and formation of debris on the sides of the scratch. For the SiC

and SiO₂ films, in region A, a plowing scratch track was found without any debris on the side of the scratch, which is probably responsible for the smoother curve and slower increase in the coefficient of friction before the critical load. After the critical load, for the SiO₂ film, delamination of the film from the substrate occurred, followed by cracking along the scratch track. For the SiC film, only several small debris particles were found without any cracks on the side of the scratch, which is responsible for the small increase in the coefficient of friction at the critical load. For the undoped Si(100), cracks were found on the side of the scratch right from the critical load and up, which is probably responsible for the big bursts in the friction profile. For the undoped polysilicon film, no cracks were found on the side of the scratch at the critical load. This might result from grain boundaries which can stop the propagation of cracks. At the end of the scratch, some of the surface material was torn away and cracks were found on the side of the scratch in the undoped Si(100). A couple of small cracks were found in the undoped polysilicon and SiO₂ films. No crack was found in the SiC film. Even at the end of the scratch, less debris was found in the SiC film. A curly chip was found at the end of the scratch in both Ni-P and Au films. This is a typical characteristic of ductile metal alloys. The Ni-P and Au films were damaged by plowing right from the beginning of the scratch with material pile-up at the side of the scratch.

The scratch depth profiles obtained during and after the scratch on all samples with respect to initial profile, after the cylindrical curvature is removed, are plotted in Fig. 6.9. Reduction in scratch depth is observed after scratching as compared to that of during scratching. This reduction in scratch depth is attributed to an elastic recovery after removal of the normal load. The scratch depth after scratching indicates the final depth which reflects the extent of permanent damage and plowing of the tip into the sample surface, and is probably more relevant for visualizing the damage that can occur in real applications. For the undoped Si(100), undoped polysilicon film, and SiO₂ film, there is a large scatter in the scratch depth data after the critical

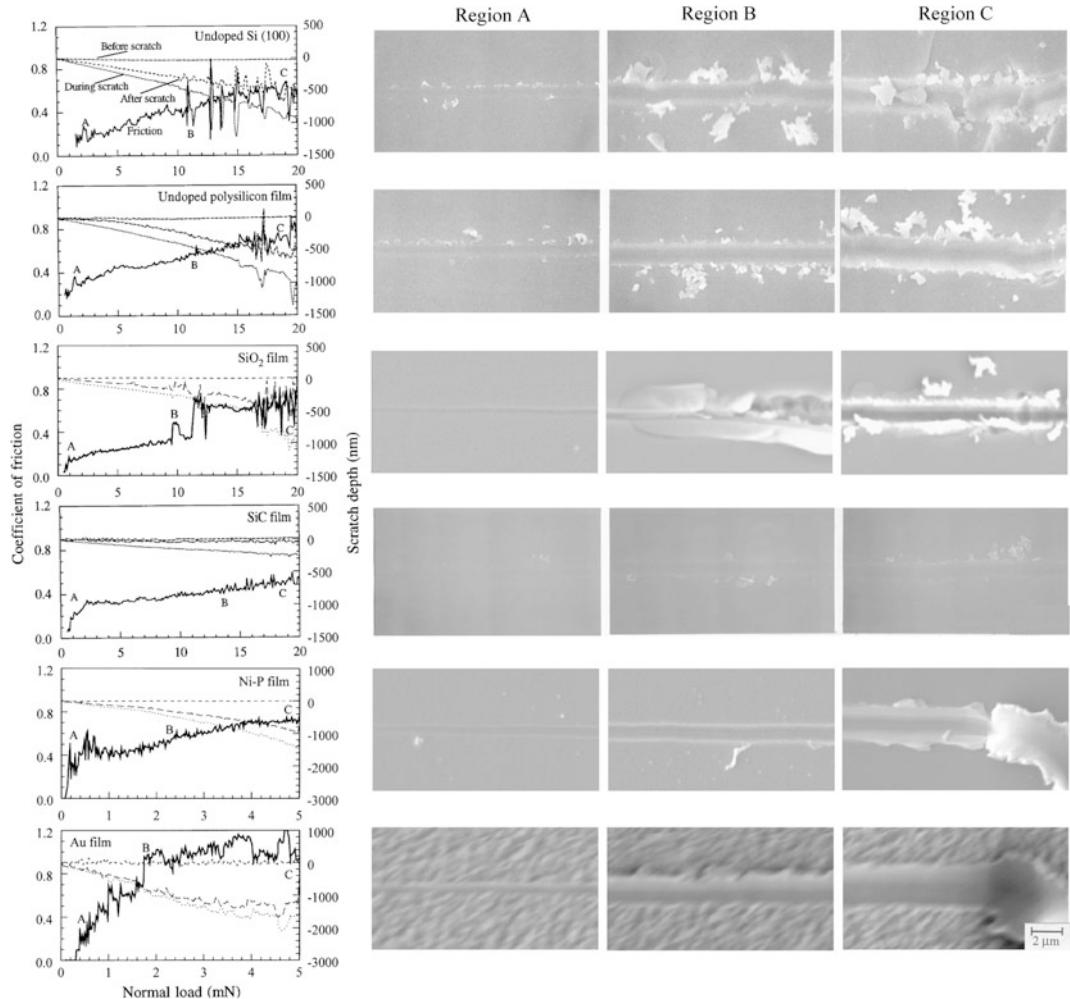


Fig. 6.9 Coefficient of friction and scratch depth profiles as a function of increasing normal load and optical images of three regions over scratches: at the beginning of the scratch (indicated by **a** on the friction profile), at the point of initiation of damage at which the coefficient of friction increases to a high value or increase abruptly (indicated by **b** on the friction profile), and towards the end of the scratch (indicated by **C** on the friction profile) for all samples (Li et al. 2003)

loads, which is associated with the generation of cracks, material removal, and debris. The scratch depth profile is smooth for the SiC film. It is noted that the SiC film exhibits the lowest scratch depth among the samples examined. The scratch depths of the undoped Si(100), undoped polysilicon film and SiO₂ film are comparable. The Ni-P and Au films exhibit much larger scratch depth than other samples. The scratch depth of the Ni-P film is smaller than that of the Au film.

The critical loads estimated from friction profiles for all samples are compared in Fig. 6.7 and Table 6.1. The SiC film exhibits the highest critical load of about 14 mN, as compared to other samples. The undoped Si(100) and undoped polysilicon film show comparable critical load of about 11 mN whereas the SiO₂ film shows a low critical load of about 9.5 mN. The Ni-P and Au films were damaged by plowing right from the beginning of the scratch.

6.3.2 Bending Tests of Ceramic Nanobeams Using an AFM

Bending tests have been performed on Si and SiO_2 nanobeam arrays (Sundararajan et al. 2002; Sundararajan and Bhushan 2002). The single-crystal silicon bridge nanobeams were fabricated by bulk micromachining incorporating enhanced-field anodization using an AFM (Namazu et al. 2000). The Si nanobeams are oriented along the [110] direction in the (001) plane. Subsequent thermal oxidation of the beams results in the formation of SiO_2 beams. The cross section of the nanobeams is trapezoidal owing to the anisotropic wet etching process. SEM micrographs of Si and SiO_2 nanobeam arrays and a schematic of the shape of a typical nanobeam are shown in Fig. 6.10. The actual widths and thicknesses of nanobeams were measured using an AFM in tapping mode prior to tests using a standard Si tapping mode tip (tip radius < 10 nm). Surface roughness measurements of the nanobeam surfaces in tapping mode yielded a σ of 0.7 ± 0.2 nm and peak-to-valley (P-V) distance of 4 ± 1.2 nm for Si and a σ of 0.8 ± 0.3 nm and a P-V distance of 3.1 ± 0.8 nm for SiO_2 . Prior to testing, the Si nanobeams were cleaned by immersing them in a “piranha etch” solution (3:1 solution by volume of 98% sulphuric acid and 30% hydrogen peroxide) for 10 min to remove any organic contaminants.

6.3.2.1 Bending Strength

Figure 6.11 shows typical load displacement curves for Si and SiO_2 beams that were bent to failure (Sundararajan et al. 2002; Sundararajan and Bhushan 2002). The upper width (w_1) of the beams is indicated in the figure. Also indicated in the figure are the elastic modulus values obtained from the slope of the load displacement curve (Eq. 6.5). All the beams tested showed linear elastic behavior followed by abrupt failure, which is suggestive of brittle fracture. Figure 6.12 shows the scatter in the values of elastic modulus obtained for both Si and SiO_2 along with the average values (\pm standard deviation). The scatter in the values may be due to

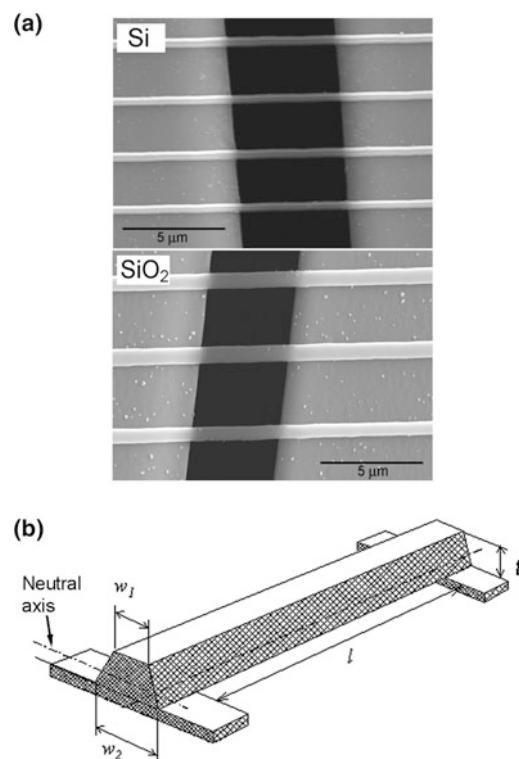


Fig. 6.10 **a** SEM micrographs of nanobeam arrays, and **b** a schematic of the shape of a typical nanobeam. The trapezoidal cross-section is due to the anisotropic wet etching during the fabrication (Sundararajan et al. 2002)

differences in orientation of the beams with respect to the trench and the loading point being a little off-center with respect to the beam span. The average values are a little higher than the bulk values (169 GPa for Si[110] and 73 GPa for SiO_2 in Table 6.2). However the values of E obtained from Eq. 6.5 have an error of about 20% due to the uncertainties in beam dimensions and spring constant of the tip/cantilever (which affects the measured load). Hence the elastic modulus values on the nanoscale can be considered to be comparable to bulk values.

Most of the beams when loaded quasi-statically at the center of the span broke at the ends as shown in Fig. 6.13a, which is consistent with the fact that maximum tensile stresses occur on the top surfaces near the ends. (See FEM stress distribution results in Fig. 6.13b.) Figure 6.14 shows the values of bending strength

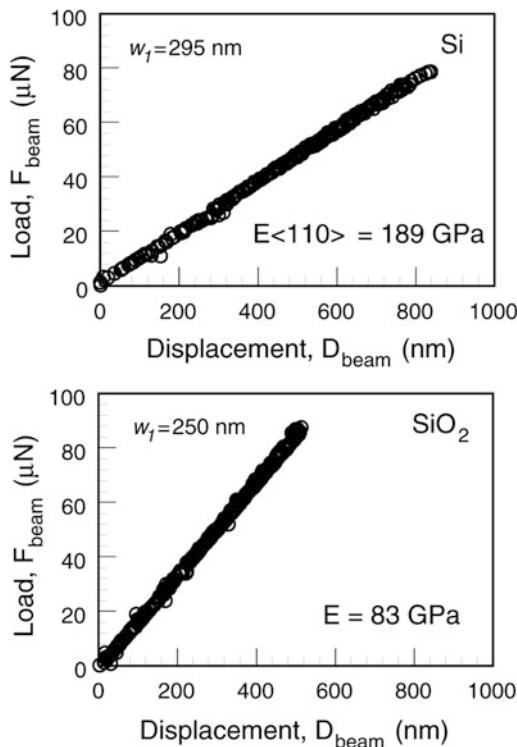


Fig. 6.11 a Typical load displacement curves of silicon and SiO_2 nanobeams. The curves are linear until sudden failure, indicative of brittle fracture of the beams. The elastic modulus (E) values calculated from the curves are shown. The dimensions of the Si beam were $w_1 = 295 \text{ nm}$, $w_2 = 484 \text{ nm}$ and $t = 255 \text{ nm}$, while those of the SiO_2 beam were $w_1 = 250 \text{ nm}$, $w_2 = 560 \text{ nm}$ and $t = 425 \text{ nm}$ (Sundararajan et al. 2002)

obtained for different beams. There appears to be no trend in bending strength with the upper width (w_1) of the beams. The large scatter is expected for the strength of brittle materials, since they are dependent on pre-existing flaw population in the material and hence are statistical in nature. The Weibull distribution, a statistical analysis, can be used to describe the scatter in the bending strength values. The means of the Weibull distributions were found to be 17.9 and 7.6 GPa for Si and SiO_2 , respectively. Previously reported numbers of strengths range from 1 to 6 GPa for silicon (Johansson et al. 1988; Ericson and Schweitz 1990; Wilson and Beck 1996; Wilson et al. 1995; Sharpe et al. 1997; Sato et al. 1998; Tsuchiya et al. 1998; Greek et al. 1999; Mazza and

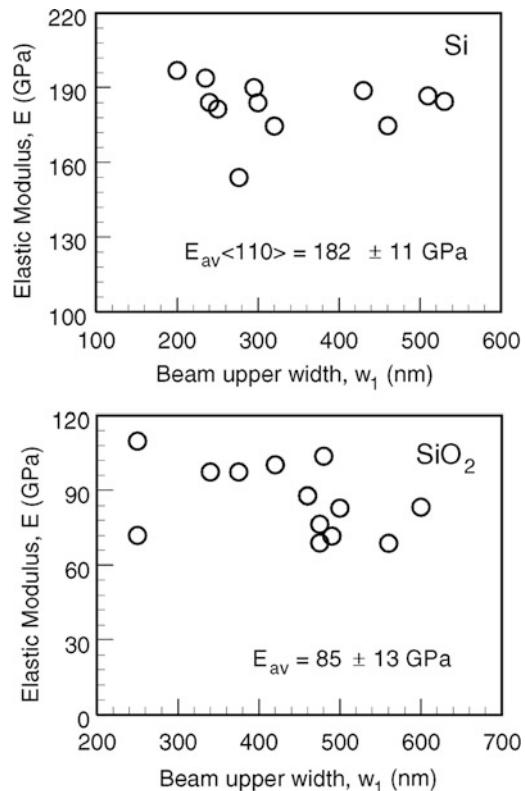


Fig. 6.12 Elastic modulus values measured for Si and SiO_2 . The average values are shown. These are comparable to bulk values, which shows that elastic modulus shows no specimen size dependence (Sundararajan and Bhushan 2002)

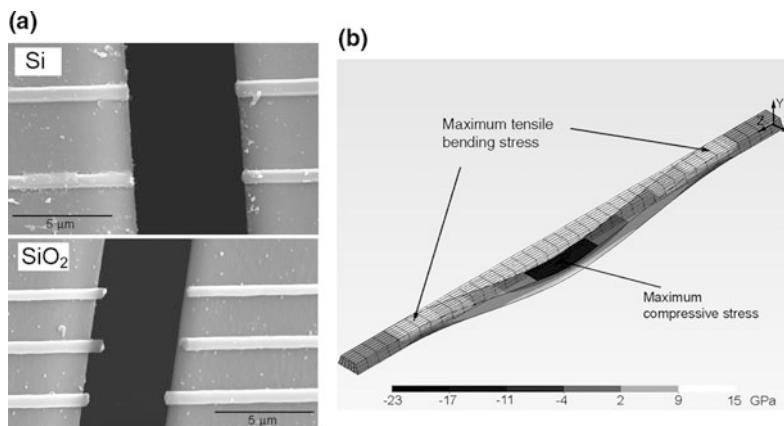
Dual 1999; Yi et al. 2000) and about 1 GPa for SiO_2 (Tsuchiya et al. 2000) microscale specimens. This clearly indicates that bending strength shows a specimen size dependence. Strength of brittle materials is dependent on pre-existing flaws in the material. Since for nanoscale specimens, the volume is smaller than for micro and macroscale specimens, the flaw population will be smaller as well, resulting in higher values of strength.

6.3.2.2 Fracture Toughness

Estimates of fracture toughness calculated using Eq. 6.10 for Si and SiO_2 are shown in Fig. 6.15 (Sundararajan and Bhushan 2002). The results show that the K_{IC} estimate for Si is about $1 - 2 \text{ MPa} \sqrt{m}$, whereas for SiO_2 the estimate is

Table 6.2 Summary of measured parameters from quasi-static bending tests

Sample	Elastic modulus E (GPa)		Bending strength σ_b (GPa)		Fracture toughness K_{IC} (MPa \sqrt{m})		
	Measured	Bulk value	Measured	Reported (microscale)	Estimated	Reported (microscale)	Bulk value
Si	182 \pm 11	169 ^a	18 \pm 3	<10 ^c	1.67 \pm 0.4	0.6–1.65 ^e	0.9 ^f
SiO ₂	85 \pm 13	73 ^b	7.6 \pm 2	<2 ^d	0.60 \pm 0.2	0.5–0.9 ^d	–

^aSi[110], Bhushan and Venkatesan (1993)^bBhushan and Gupta (1997)^cJohansson et al. (1988), Ericson et al. (1990), Wilson and Beck (1996), Wilson et al. (1996), Sharpe et al. (1997), Sato et al. (1998), Tsuchiya et al. (1998), Greek et al. (1999), Yi et al. (2000)^dTsuchiya et al. (2000)^eJohansson et al. (1989), Ballarini et al. (1997), Kahn et al. (1999), Fitzgerald et al. (2000)^fAnonymous (1988)**Fig. 6.13** **a** SEM micrographs of nanobeams that failed during quasi-static bending experiments. The beams failed at or near the ends, which is the location of maximum tensile bending stress (Sundararajan et al. 2002), and **b** bending stress distribution for silicon nanobeam indicating that the maximum tensile stresses occur on the top surfaces near the fixed ends

about 0.5–0.9 MPa \sqrt{m} . These values are comparable to values reported by others on larger specimens for Si (Johansson et al. 1989; Ballarini et al. 1997; Kahn et al. 1999; Fitzgerald et al. 2000) and SiO₂ (Tsuchiya et al. 2000). The high values obtained for Si could be due to the fact that the scratches, despite being quite sharp, still have a finite radius of about 100 nm. The bulk value for silicon is about 0.9 MPa \sqrt{m} (Table 6.2). Fracture toughness is considered to be a material property and is believed to be independent of specimen size. The values obtained in this study, given its limitations,

appear to show that fracture toughness is comparable, if not a little higher on the nanoscale.

6.3.2.3 Fatigue Strength

Fatigue strength measurements of Si nanobeams have been carried out by Sundararajan and Bhushan (2002) using an AFM and Li and Bhushan (2003) using a nanoindenter. Various stress levels were applied to nanobeams by Sundararajan and Bhushan (2002). The minimum stress was 3.5 GPa for Si beams and 2.2 GPa for SiO₂ beams. The frequency of applied load was 4.2 Hz. In general, the fatigue

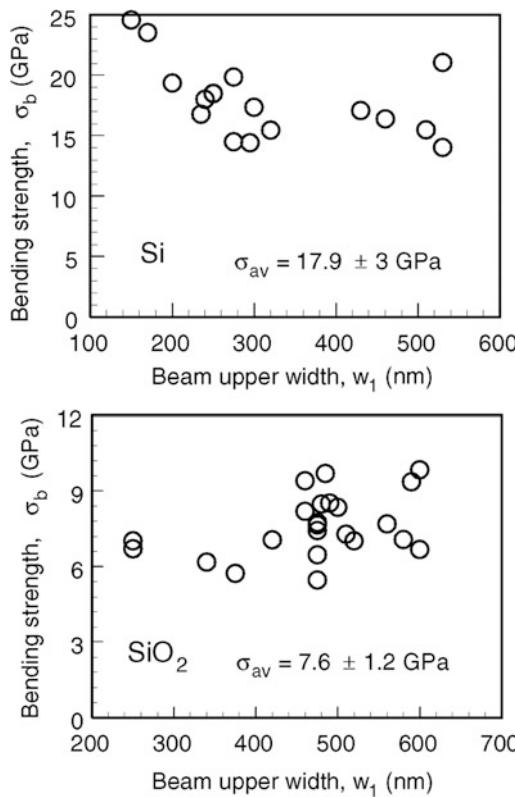


Fig. 6.14 Bending strength values obtained from bending experiments. Average values are indicated. These values are much higher than values reported for microscale specimens, indicating that bending strength shows a specimen size effect (Sundararajan and Bhushan 2002)

life decreased with increasing mean stress as well as increasing stress amplitude. When the stress amplitude was less than 15% of the bending strength, the fatigue life was greater than 30,000 cycles for both Si and SiO₂. However, the mean stress had to be less than 30% of the bending strength for a life of greater than 30,000 cycles for Si whereas even at a mean stress of 43% of the bending strength, SiO₂ beams showed a life greater than 30,000 cycles. During fatigue, the beams broke under the loading point or at the ends, when loaded at the center of the span. This was different from the quasi-static bending tests, where the beams broke at the ends almost every time. This could be due to the fact that the stress levels under the load and at the ends are not that different and fatigue crack propagation could

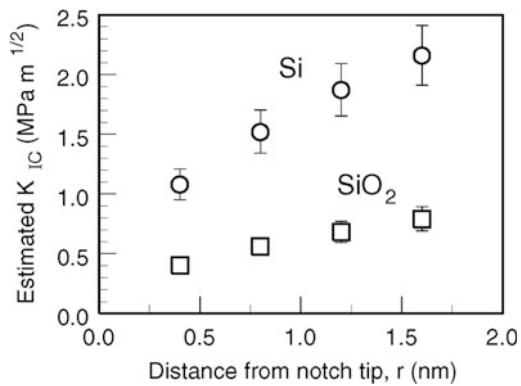


Fig. 6.15 Fracture toughness (K_{IC}) values of for increasing values of r corresponding to distance between neighboring atoms in {111} planes of silicon (0.4 nm). Hence r values between 0.4 and 1.6 nm are chosen. The K_{IC} values thus estimated are comparable to values reported by others for both Si and SiO₂ (Sundararajan and Bhushan 2002)

occur at either location. Figure 6.16 shows a nanoscale S-N curve, with bending stress (S) as a function of fatigue in number of cycles (N) with an apparent endurance life at lower stress. This study clearly demonstrates that fatigue properties of nanoscale specimens can be studied.

6.3.2.4 SEM Observations of Fracture Surfaces

Figure 6.17 shows SEM images of the fracture surfaces of nanobeams broken during quasi-static bending as well as fatigue (Sundararajan and Bhushan 2002). In the quasi-static cases, the maximum tensile stresses occur on the top surface, so it is reasonable to assume that fracture initiated at or near the top surface and propagated downward. The fracture surfaces of the beams suggest a cleavage type of fracture. Silicon beam surfaces show various ledges or facets, which is typical for crystalline brittle materials. Silicon usually fractures along the (111) plane due to this plane having the lowest surface energy to overcome by a propagating crack. However, failure has also been known to occur along the (110) planes in microscale specimens, despite the higher energy required as compared to the (111) planes (Wilson and Beck 1996). The plane normal to the beam direction in these samples is

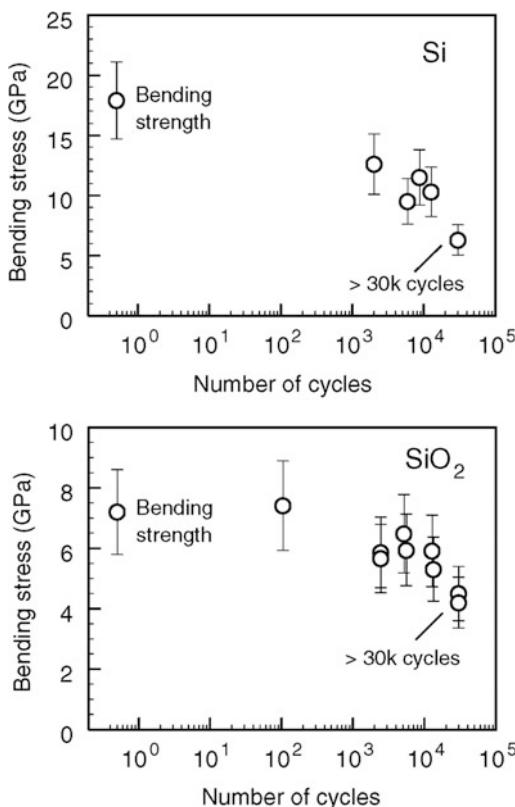


Fig. 6.16 Fatigue test data showing applied bending stress as a function of number of cycles. A single load-unload sequence is considered as 1 cycle. The bending strength data points are therefore associated with $\frac{1}{2}$ cycle, since failure occurs upon loading (Sundararajan and Bhushan 2002)

the (110) plane while (111) planes will be oriented at 35° from the (110) plane. The presence of facets and irregularities on the silicon surface in Fig. 6.17a suggest that it is a combination of these two types of fractures that has occurred. Since the stress levels are very high for these specimens, it is reasonable to assume that crack propagating forces will be high enough to result in (110) type failures.

In contrast, the silicon fracture surfaces under fatigue, shown in Fig. 6.17b, appear very smooth without facets or irregularities. This is suggestive of low energy fracture, i.e. of (111) type fracture. We do not see evidence of fatigue crack propagation in the form of steps or striations on the fracture surface. We believe that for the stress

levels applied in these fatigue experiments, failure in silicon occurred via cleavage associated with ‘static fatigue’ type of failures.

SiO_2 shows very smooth fracture surfaces for both quasi-static bending and fatigue. This is in contrast to the irregular surface one might expect for the brittle failure of an amorphous material on the macroscale. However, in larger scale fracture surfaces for such materials, the region near the crack initiation usually appears smooth or mirror-like. Since the fracture surface here is so small and very near the crack initiation site, it is not unreasonable to see such a smooth surface for SiO_2 on this scale. There appears to be no difference between the fracture surfaces obtained by quasi-static bending and fatigue for SiO_2 .

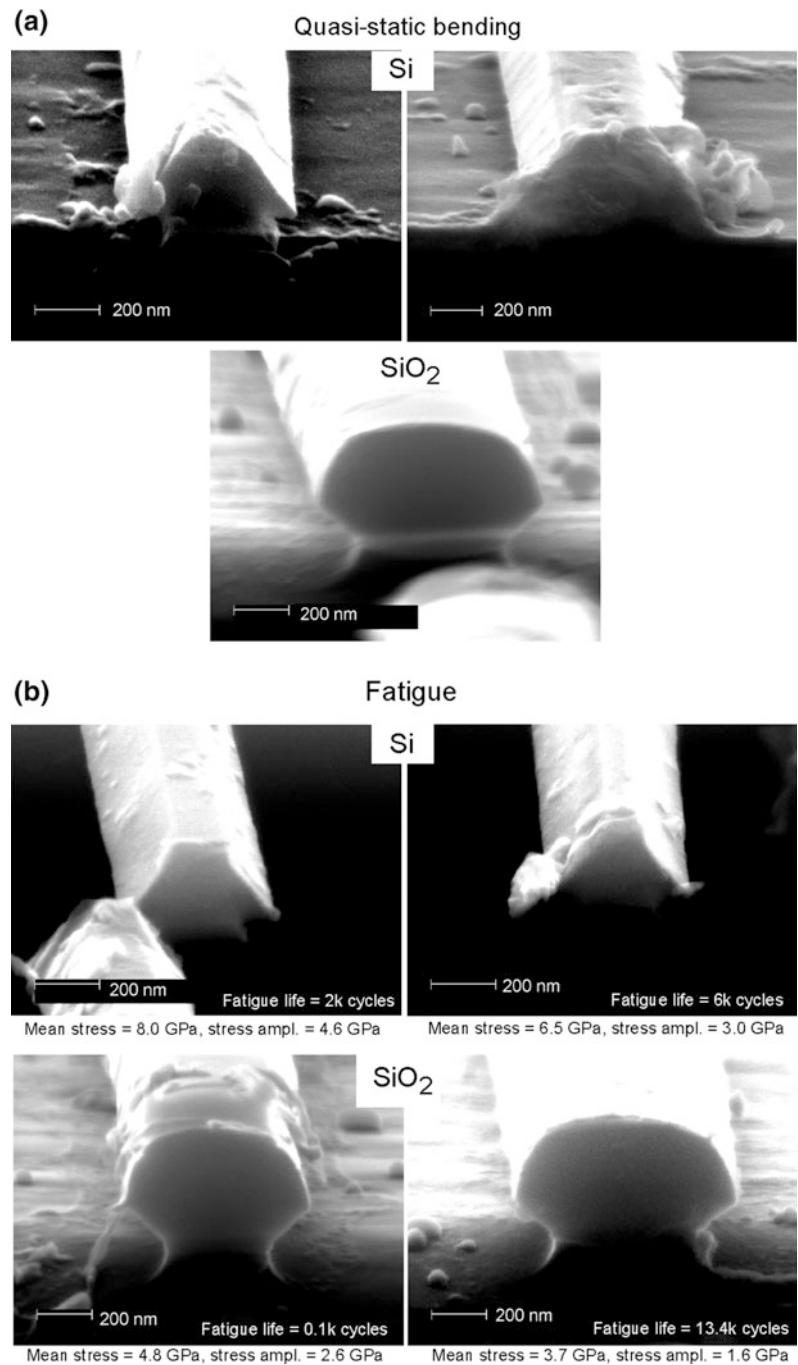
6.3.2.5 Summary of Mechanical Properties Measured Using Quasi-Static Bending Tests

Table 6.2 summarizes the various properties measured via quasi-static bending in this study (Sundararajan and Bhushan 2002). Also shown are bulk values of the parameters along with values reported on larger scale specimens by other researchers. Elastic modulus and fracture toughness values appear to be comparable to bulk values and show no dependence on specimen size. However bending strength shows a clear specimen size dependence with nanoscale numbers being twice as large as numbers reported for larger scale specimens.

6.3.3 Bending Tests of Metallic Microbeams Using a Nanoindenter

Bending tests have been performed on Ni-P and Au microbeams (Li et al. 2003). The Ni-P cantilever microbeams were fabricated by a focused ion beam machining technique. The dimensions were $10 \times 12 \times 50 \mu\text{m}^3$. Notches with a depth of $3 \mu\text{m}$ and a tip radius of $0.25 \mu\text{m}$ were introduced in the microbeams to facilitate failure at a lower load in the bending tests. The Au

Fig. 6.17 SEM micrographs of fracture surfaces of silicon and SiO_2 beams subjected to **a** quasi-static bending and **b** fatigue (Sundararajan and Bhushan 2002)



bridge microbeams were fabricated by electroplating technique.

Figure 6.18 shows the SEM images, load displacement curve, and FEM stress contour for the notched Ni-P cantilever microbeam that was bent to failure (Li et al. 2003). The distance between the loading position and the fixed end is 40 μm . The 3 μm deep notch is 10 μm from the fixed end. The notched beam showed linear behavior followed by abrupt failure. The FEM stress contour shows that there is a higher stress concentration at the notch tip. The maximum tensile stress σ_m at the notch tip can be analyzed

by using the Griffith fracture theory as follows (Lawn et al. 1980),

$$\sigma_m \approx 2\sigma_0 \left(\frac{c}{\rho} \right)^{1/2} \quad (6.12)$$

where σ_0 is the average applied tensile stress on the beam, c is the crack length, and ρ is the crack tip radius. Therefore, elastic-plastic deformation will first occur locally at the end of the notch tip, followed by abrupt fracture failure after σ_m reaches the ultimate tensile strength of Ni-P, even though the rest of the beam is still in the

Fig. 6.18 SEM micrographs of the new and broken beams, load displacement curve and FEM stress contour for the notched Ni-P cantilever microbeam (Li et al. 2003)

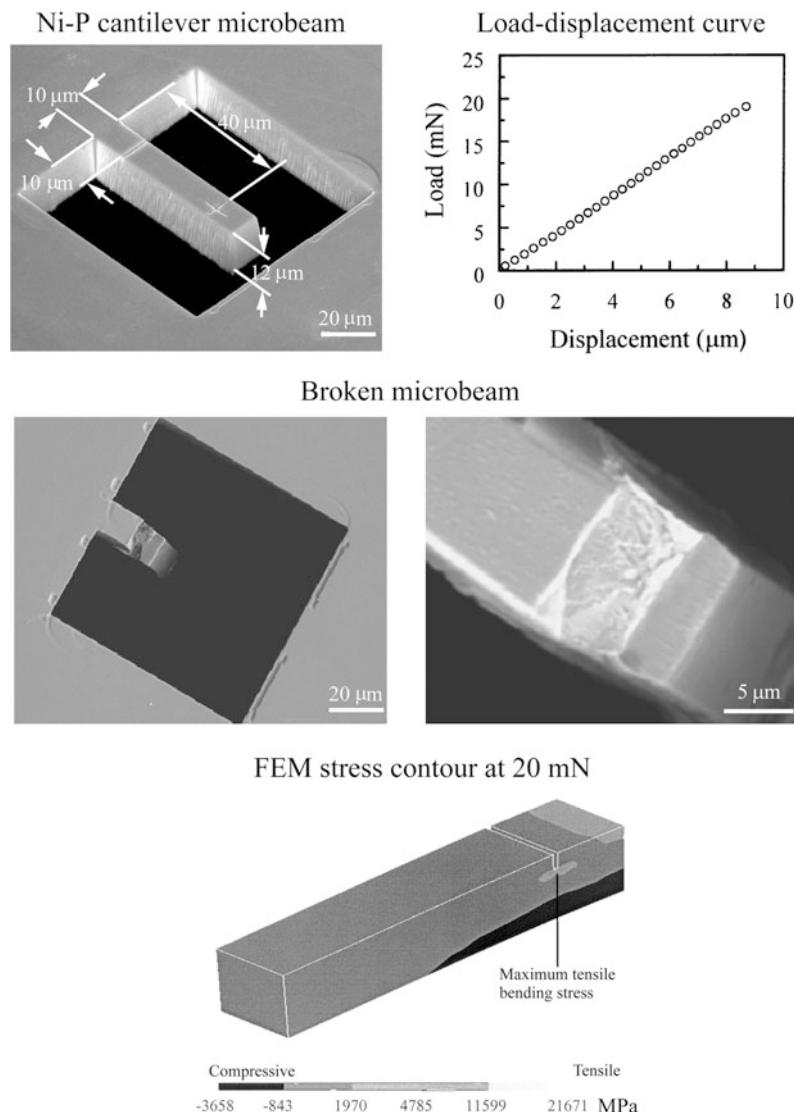
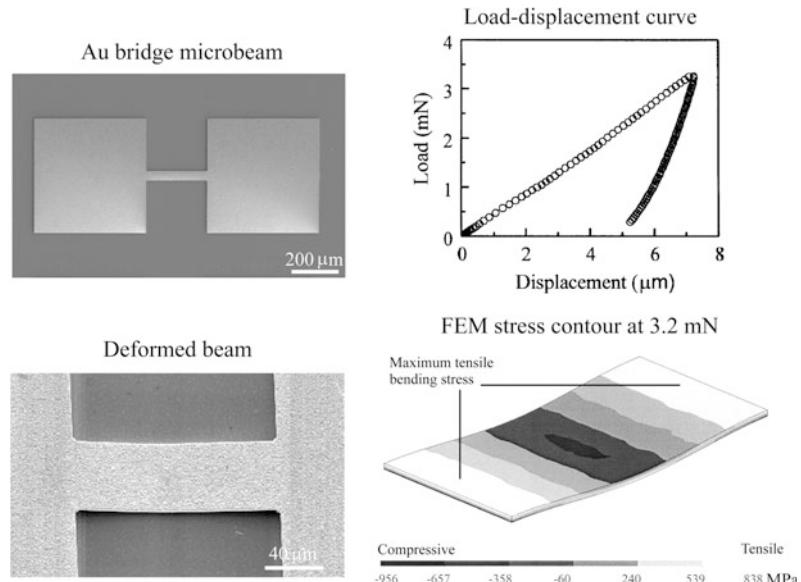


Fig. 6.19 SEM micrographs of the new and deformed beams, load displacement curve and FEM stress contour for the Au bridge microbeam (Li et al. 2003)



elastic regime. The SEM image of the fracture surface shows that the fracture started right from the notch tip with plastic deformation characteristics. This indicates that although local plastic deformation occurred at the notch tip area, the whole beam failed catastrophically. The present study shows that FEM simulation can predict well the stress concentration, and helps in understanding the failure mechanism of the notched beams.

Figure 6.19 shows the SEM images, load displacement curve, and FEM stress contour for the Au bridge microbeam that was deformed by the indenter (Li et al. 2003). The recession gap between the beam and substrate is about 7 μm, which is not large enough to break the beam at the load applied. From the load-displacement curve, we note that the beam experienced elastic-plastic deformation. The FEM stress contour shows that the maximum tensile stress is located at the fixed ends whereas the minimum compressive stress is located around the center of the beam. The SEM image shows that the beam has been permanently deformed. No crack was found on the beam surface. The present study shows a possibility for mechanically forming the Au film into the shape as needed. This may help in designing/fabricating functionally complex

smart micro/nanodevices which need conductors for power supply and input/output signals.

6.3.4 Indentation and Scratch Tests of Polymeric Microbeams Using a Nanoindenter

Four potential structural materials for BioMEMS devices were selected for this study, namely poly (propyl methacrylate) (PPMA), poly(methyl methacrylate) (PMMA), polystyrene (PS) and a polystyrene-nanoclay composite (PS/Clay) (Palacio et al. 2007a). The physical properties of PPMA, PMMA and PS are summarized in Table 6.3. PMMA and PS are widely used engineering polymers for a number of commercial applications. Historically, PMMA was the polymer of choice in ophthalmologic devices due to its high refractive index, hardness, and biocompatibility. Its surface can be functionalized with proteins, which promotes the bonding of tissues for in vivo implants (Loh et al. 1997). PMMA is also employed in chips and valve components for immunosensors and other lab-on-a-chip applications (Holt et al. 2002; van der Berg 2003). PPMA has a lower glass transition temperature (T_g) (35–43 °C) (Ellis 2000)

Table 6.3 Summary of the physical properties of PPMA, PMMA and PS

	Structure	Density (g/cm ³)	T _g (°C)	Water absorption after 24 h (%)	Thermal expansion coefficient (K ⁻¹)	Thermal conductivity (0–50 °C) (W m ⁻¹ K ⁻¹)	Heat capacity (0–100 °C) (kJ kg ⁻¹ K ⁻¹)
PPMA		—	35–43 ^a	—	—	—	—
PMMA		1.188 ^b	104–106 ^{b,c}	0.3–0.4 ^d 0.1–0.3 ^e	2.0–3.0 × 10 ^{-4c} (<T _g) 6.0 × 10 ^{-4c} (>T _g)	0.193 ^b	1.255–1.720 ^b
PS		1.040–1.065 ^c (amorphous) 1.111–1.127 ^c (crystalline)	100 ^{b,c}	0.03–0.1 ^d	1.7–2.1 × 10 ^{-4c} (<T _g) 5.1–6.0 × 10 ^{-4c} (>T _g)	0.105–0.116 ^b	1.185–1.838 ^b

^aEllis (2000)^bBrandrup et al. (1999)^cMark (1999)^dBilimeyer (1984)^eAnonymous

than PMMA (104–106 °C) (Brandrup et al. 1999; Mark 1999), which allows for easier processing at a lower temperature.

Polystyrene is particularly desirable for BioMEMS due to its ubiquitous use in tissue culture applications. Plasma treated polystyrene is the most commonly used material for *in vitro* cell biology studies of adherent cells. Thus, there is a wealth of knowledge about cellular behavior on polystyrene. In addition to tissue culture polystyrene, various surface modification techniques can be employed for functionalizing the PS surface in order to promote cell attachment and proliferation (Loh et al. 1997). Oxygen plasma modified polystyrene has been shown to improve cell growth, proliferation, and expression of cellular adhesion protein proportional to the surface oxygen concentration (van Kooten et al. 2004). The previous knowledge of cellular interactions with polystyrene makes it a logical choice to use in BioMEMS devices for cellular interactions.

Clay nanoparticles in a polymer matrix act to improve the mechanical and thermal properties as compared to the native polymer (Alexandre and Dubois 2000; Ray and Okamoto 2003). Nanoclay composites can also be used to improve barrier resistance and improve ionic conductivity. The extent to which polymer properties are affected is determined by the clay content, polymer/clay interfacial strength, and dispersion of the clay particles. The fabrication procedure for the double-anchored and cantilever beams is presented in App. A. SEM images of the two types of beams are shown in Fig. 6.20.

The hardness (H), elastic modulus (E) and creep of the beams were measured at the supported region of the beams (Wei et al. 2005; Palacio et al. 2007a). The indentation location, where the H, E and creep were measured, is indicated by an arrow in Fig. 6.20a.

Since the fabricated polymer beams were too narrow (width about 5 μm) to perform nanoscratch tests, the nanoscratch experiments were conducted on thin films of polymeric materials. PPMA, PMMA, PS and PS/Clay were dissolved in anisole at various concentration and spin coated for 1 min at 3000 RPM to achieve the desired film thickness of 500 nm. All film

samples were then baked at 95 °C for 2 min to remove any residual solvent.

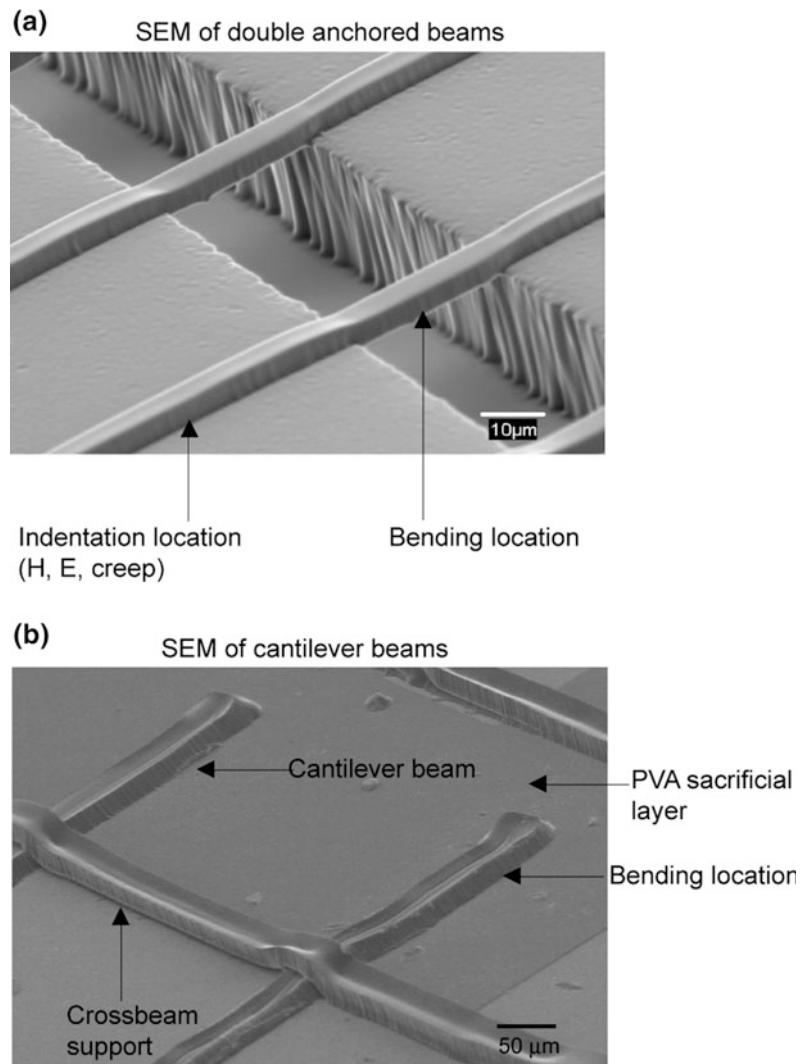
6.3.4.1 Hardness and Elastic Modulus

Figure 6.21a shows the H and E of various beams as a function of contact depth (Palacio et al. 2007a). Five beams in one wafer sample were indented once, and the average values of H and E at a contact depth of 100 nm are reported in Table 6.4. Among the materials examined, the PS/Clay nanocomposite exhibited the highest hardness and elastic modulus (of 390 ± 50 MPa and 5.1 ± 0.4 GPa, respectively), followed by PMMA (340 ± 30 MPa, 4.8 ± 0.5 GPa), PS (290 ± 20 MPa, 3.6 ± 0.4 GPa) and PPMA (110 ± 30 MPa, 1.7 ± 0.5 GPa). The standard deviation was calculated from the five indents performed.

6.3.4.2 Indentation Creep

Creep is an important aspect of polymer mechanical behavior. Figure 6.21b shows representative creep data for the supported polymer beams using CSM indentation, where the change in displacement, mean stress and contact stiffness was monitored while holding the tip at the maximum imposed load (Palacio et al. 2007a). Conclusions on the creep response of the materials are based mainly on the displacement change at constant loading, which is directly measured in the experiment. For each material, three beams on one sample were tested, and it was found that the variation in creep response was no more than 15%. As expected for polymers, the indentation displacement increases with time. The material with the lowest hardness and modulus (PPMA) exhibited the fastest rate of displacement change (i.e., mostly creep), while the PS/Clay nanocomposite (which has the highest H and E value) showed minimal creep. This shows a direct correlation between the hardness, modulus, and creep resistance. The mean stress exerted by the tip decreases with time as the polymer deforms viscoelastically underneath the tip during the hold segment. For the PS/Clay nanocomposite, it is observed that this quantity decreases more rapidly during the first few seconds of the measurement compared

Fig. 6.20 SEM images of polystyrene **a** double anchored beams, and **b** cantilever beams (Palacio et al. 2007a)



to the other materials. This could be attributed to the presence of an additional phase (the clay filler). The filler-matrix interface could exhibit yield without necessarily causing further penetration into the material.

6.3.4.3 Scratch Resistance

The scratch behavior of thin films of the four polymers studied is presented in Fig. 6.21c. Plots on the left side show the depth profile before, during, and after a scratch, along with the coefficient of friction. SEM images were taken at three areas: at the beginning of the scratch (indicated by “A” on the friction profile), middle

of the scratch (“B”) and towards the end of the scratch (“C”). Among the three homogeneous films, PMMA had the greatest scratch resistance as it exhibited the least penetration as well as the absence of bursts in the scratch depth profile. PMMA is followed by PS, then PPMA. This order is consistent with the observed trend for the elastic modulus and hardness. The features in the scratch depth profiles can be correlated to the morphology of the scratches. In the PPMA depth profile, the tip jumped up and down during the scratch instead of continuously ramping down. This is accounted for by the SEM images, which reveal deformation bands that are convex with

respect to the scratch direction. The bands indicate plastic deformation and material pile-up in front of the tip. In PS, the oscillation in the depth profile occurs only towards the end of the scratch, and the deformation bands are observed only in Region C. For PMMA, the smooth depth profile corresponds to a clear scratch track,

indicating that minimal pile-up occurred. The nanocomposite exhibits enhanced scratch resistance relative to the unfilled PS. The peaks and valleys in the depth profile come from the roughness of the surface and not from scratch-induced material displacement, which is confirmed from the SEM images.

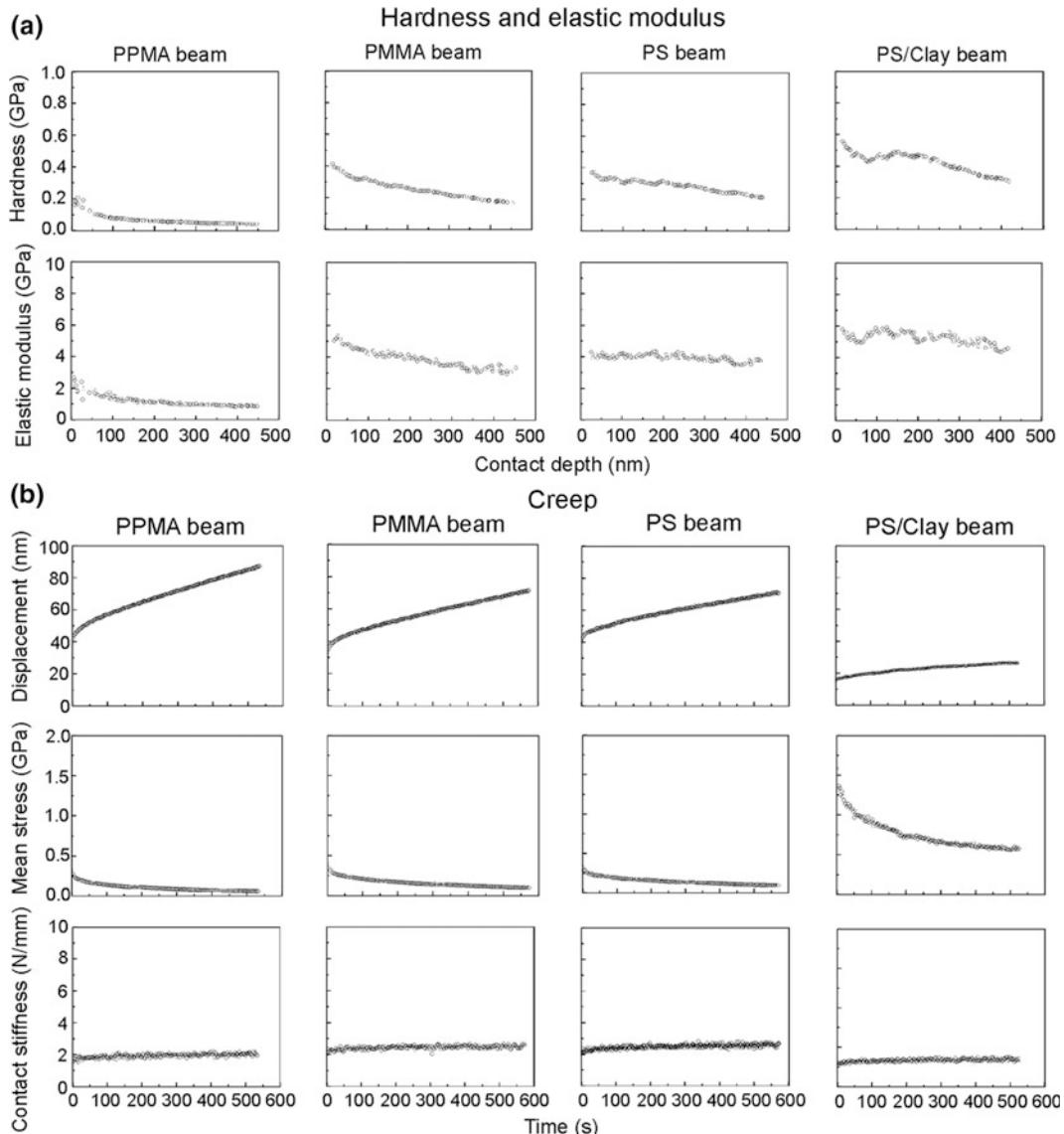


Fig. 6.21 **a** Hardness and elastic modulus of PPMA, PMMA, PS and PS/Clay beams as a function of contact depth, **b** creep displacement, mean stress and contact stiffness as a function of time for PPMA, PMMA, PS and PS/Clay, and **c** scratch depth profiles and coefficient of friction as a function of increasing normal loads for thin films of PPMA, PMMA, PS and PS/Clay. SEM images were taken at three regions: at the beginning of the scratch (“A”), in the middle of the scratch (“B”) and at the end of the scratch (“C”) (Palacio et al. 2007a)

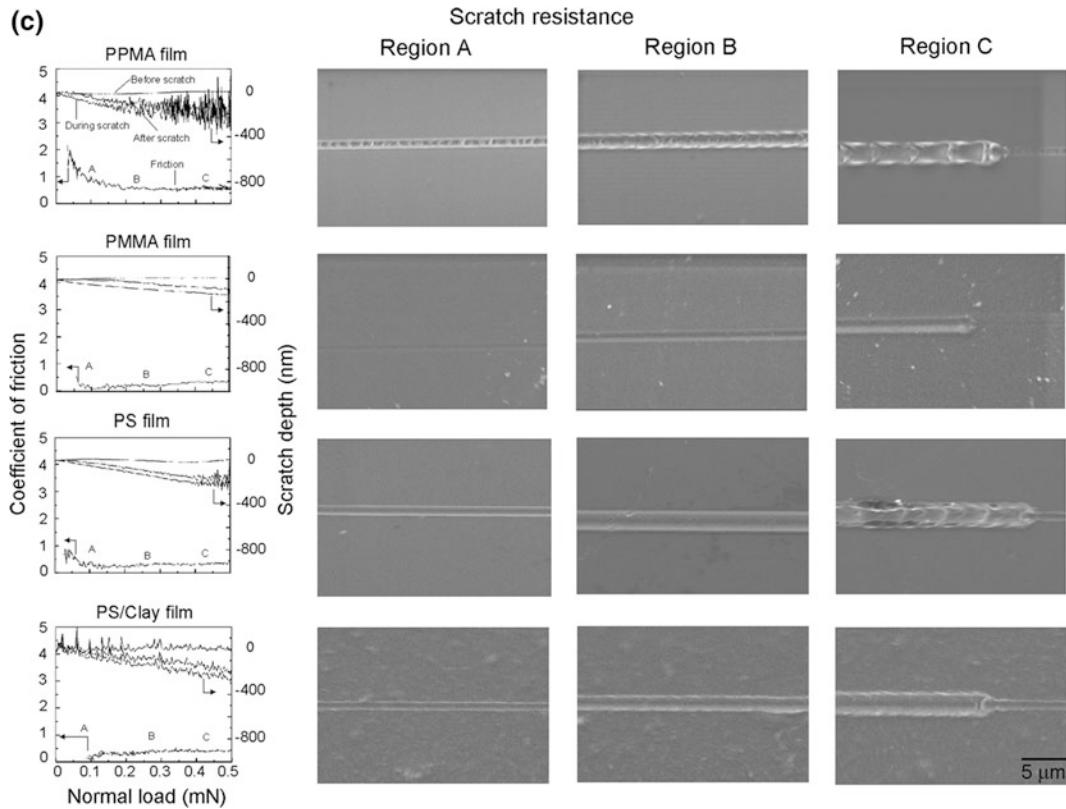


Fig. 6.21 (continued)

6.3.5 Bending Tests of Polymeric Microbeams Using a Nanoindenter

6.3.5.1 Suspended Beams

For each material, five suspended beams in one sample were measured, and representative load-displacement data are presented in Fig. 6.22. In this case, the displacement is due to beam bending. The loading profile can be fit into a straight line, indicating that bending at this load range is fully elastic. From the slope of this line and Eq. 6.5, the calculated elastic moduli are 0.7 ± 0.2 , 1.9 ± 0.3 , 1.9 ± 0.1 and 4.6 ± 0.3 GPa for PPMA, PMMA, PS and PS/Clay, respectively. The standard deviation was based on five measurements.

The elastic moduli from beam bending are compared to values from nanoindentation and the available bulk literature values. The consistency in the obtained trend in E for the four materials using the three methods provides validation for the beam bending technique. In addition, the beam bending method has been successfully employed on various metal and ceramic materials, as presented earlier (Li et al. 2003). The slightly lower elastic moduli calculated from beam bending tests (relative to nanoindentation and bulk literature data) can be attributed to several reasons. Some deformation may have occurred in the boundary between the supported and the suspended region of the beam due to stress concentration at this junction. A small amount of plastic deformation of the beam surface may occur even

though the indenter tip was coated with PMMA. This could reduce the slope of the loading curve.

Effect of Soaking and Temperature

The effects of aqueous medium and human body temperature (37.5°C) were investigated on both the suspended and supported beams. Figure 6.23a is a summary of the effects of beam soaking and elevated temperature on the

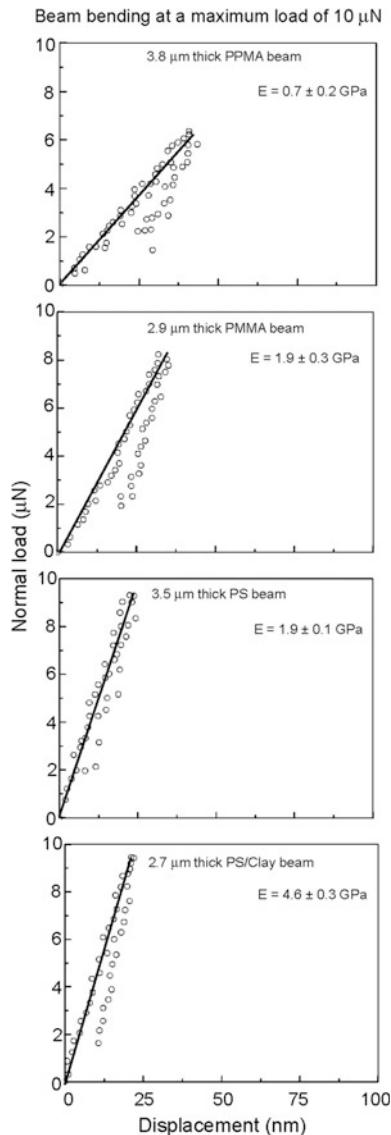


Fig. 6.22 Normal beam bending results for PPMA, PMMA, PS and PS/Clay at a maximum load of $10\ \mu\text{N}$ (Palacio et al. 2007a)

hardness and elastic modulus as obtained by nanoindentation on the supported part of the beam. Among the four materials studied, soaking had the greatest effect on PPMA. On the other hand, the H and E of the other materials (PMMA, PS and PS/Clay) were not significantly affected. The properties of the PPMA beam were adversely affected when the indentation temperature was increased to 37.5°C . This is because the glass transition temperature (T_g) of PPMA is within $35\text{--}43^{\circ}\text{C}$ (Ellis 2000), such that the glassy to rubbery transition is being observed here. On the other hand, no significant decrease in H and E was observed on the other three materials since their T_g 's are much higher, such that the indent was performed well within the glassy regime.

In Fig. 6.23b, beam bending data are shown for each of the four polymers investigated. The PPMA beam is affected by soaking in deionized (DI) water, as shown by the appearance of a large amount of hysteresis in the load-displacement data. Human body temperature has an adverse effect on the stiffness of the PPMA beam. These are consistent with results from indentation. From a materials selection standpoint, these are important findings as these imply that the performance of PPMA structures will be compromised if it were in a device subjected to either soaking in an aqueous medium or above ambient temperatures.

Normal Beam Bending at Elevated Loads

Bending experiments were performed with applied loads in the millinewton range in order to investigate the behavior of polymer beams beyond the elastic deformation regime. The upper half of Fig. 6.24 shows normal load-displacement profiles at maximum imposed loads of 0.5, 1.5 and $3.0\ \text{mN}$ applied on three different beams in one wafer sample. For PPMA, PMMA and PS, the loading portion overlaps, indicating uniform beam quality. Slight variability in the loading profile is observed in the PS/Clay composite, which is attributed to variable interfacial adhesion between the filler and the matrix and inhomogeneities in the dispersion of the clay particles.

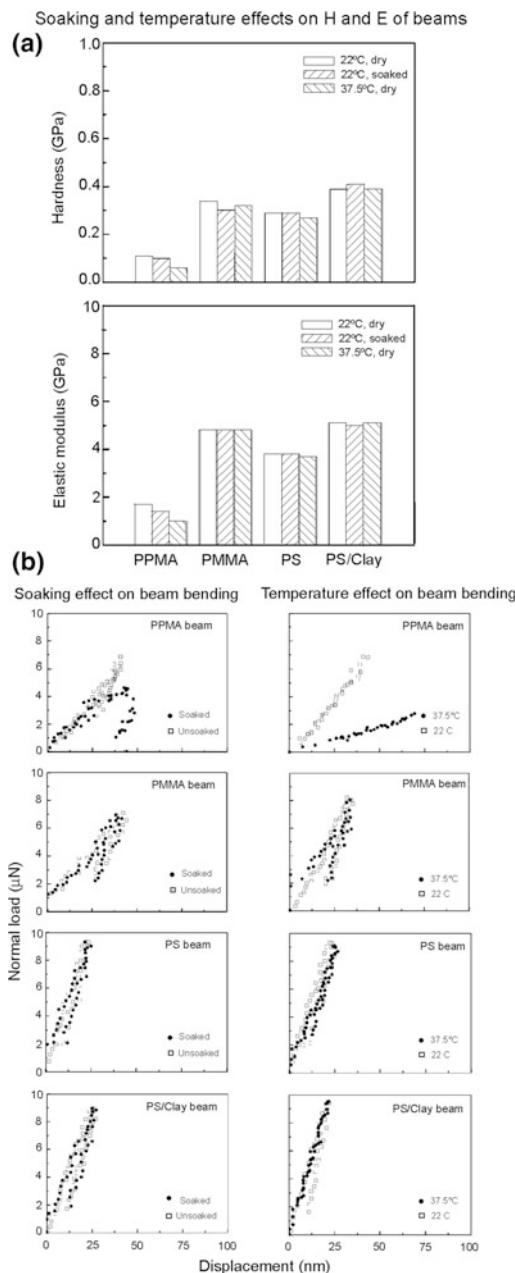


Fig. 6.23 **a** Effect of soaking and elevated temperature on suspended beams and **b** effect of soaking and elevated temperature on the hardness and elastic modulus of supported beams (Palacio et al. 2007a)

The first inflection in the load-displacement profiles implies the onset of yield, and its location is indicated by arrows. The three unfilled polymers (PPMA, PMMA and PS) exhibit

ductility, as seen in the SEM images for the beams tested with an applied load of 3 mN. The beams stretch at the ends and underneath the center of the beam. These correspond to regions subjected to maximum tensile stress (Li et al. 2003). It is worth noting that the beams deformed symmetrically, indicating that the load was applied equidistant from the clamping points. The PS/Clay nanocomposite beam exhibited slight yielding, followed by breaking, as indicated by an arrow in the load-displacement profile. After breaking, the load continued to increase because a pre-defined load is imposed by the nanoindenter. The corresponding SEM image for PS/Clay shows a flat fracture surface, indicating that the crack propagated perpendicular to the direction of the applied stress.

From the load-displacement curves, the yield strength was evaluated by using the following equation (Young and Budynas 2002):

$$\sigma = \frac{3}{4} \frac{Fl}{bh^2} \quad (6.13)$$

where F is the load corresponding to yield of the beam (as indicated by arrows in the load-displacement profile), and l, b and h are the beam's length, width and height, respectively. It is assumed that the beam ends are clamped. The yield strength values listed in Table 6.4 are in the same order of magnitude as the bulk yield strength data in the literature.

Bending data at a maximum load of 10 mN are shown in the lower half of Fig. 6.24. As with the bending tests performed at 3 mN, symmetric deformation and failure is observed in the beams. The three unfilled polymers stretched considerably prior to fracture, and the PS/Clay nanocomposite beam exhibited brittle failure as observed in the lower load experiment.

For brittle materials, the elastic regime is immediately followed by breaking, and the breaking strength can be evaluated using Eq. 6.13. Even though plastic deformation was observed in all materials, Eq. 6.13 was assumed valid, and the force corresponding to breaking is taken as the load observed prior to the abrupt jump in displacement. The calculated breaking strength

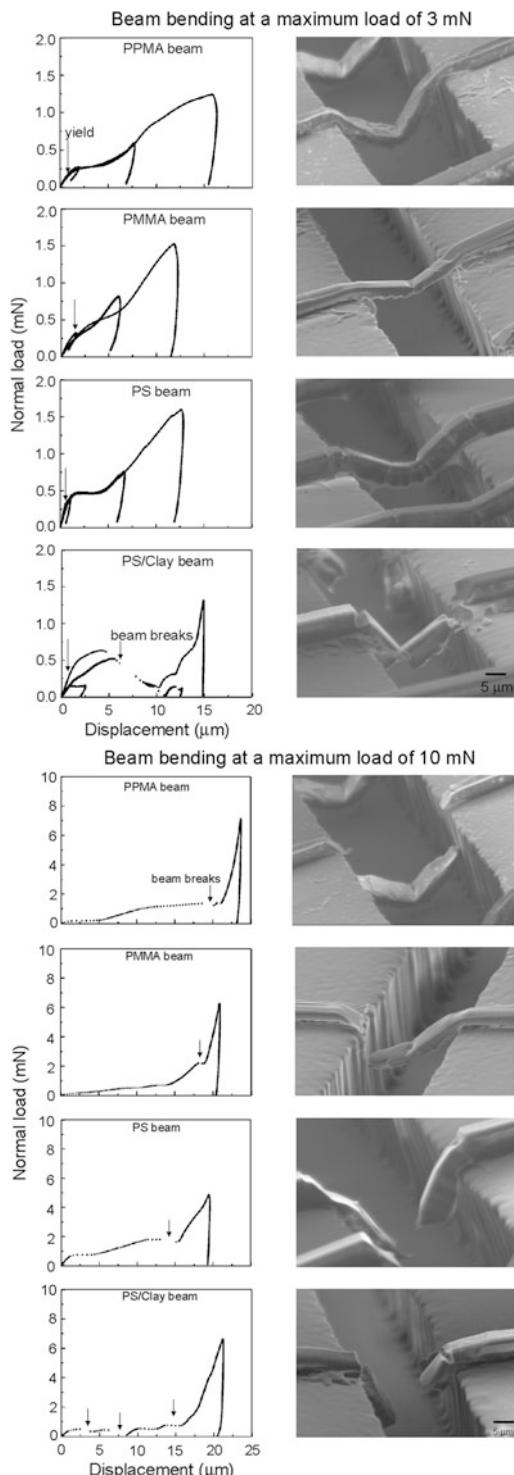


Fig. 6.24 Normal beam bending at elevated loads. The upper half shows normal load-displacement profiles at a maximum load of 0.5, 1.5 and 3.0 mN, where arrows indicate the onset of yield and breaking of the beams. For plots with multiple profiles (*upper half*), the SEM image corresponds to the beam subjected to the highest load (3 mN). The lower half shows normal load-displacement profiles and SEM images for beams subjected to 10 mN maximum load (Palacio et al. 2007a)

(as shown in Table 6.4) is higher and does not appear to correlate with the trends observed for the other mechanical properties reported. This can be attributed to a number of reasons. The load-displacement response may not purely be from beam bending. A higher load can be imposed without beam breaking due to plastic deformation on the beam surface. At elevated loads, a larger contact area between the tip and beam is expected such that permanent deformation on the surface can be significant. In addition, the compliant trench material (SU-8) can help to accommodate a higher stress prior to breaking of the beam.

6.3.5.2 Lateral Bending of the Cantilever Beams

Figure 6.25 presents the results of the lateral bending experiments on PS and PS/Clay cantilever beams. On the left column are the plots for the lateral force recorded as a function of the lateral displacement. From the linear elastic regime (as indicated by the straight line drawn over the data points), the modulus can be evaluated using Eq. 6.11. The resulting elastic modulus of PS/Clay is three times that of PS, as shown in Table 6.4. This was obtained from measurements on two beams for each material. However, the elastic modulus values obtained from this experiment are lower than that obtained from nanoindentation and beam bending in the normal direction. This is attributed to the following: the applied pressure coming from the addition of the crossbeam reinforcement induces deformation on the anchor. In addition, a small curvature is observed on the beam. This diminishes the beam width, which is the dominant term

Table 6.4 Mechanical properties of the polymer beams under investigation

Polymer	Poisson's ratio	Elastic modulus (E) (GPa)			Hardness (H) (GPa)			Yield strength (σ_{ys}) (MPa)			Breaking strength (σ_b) (MPa)			
		Bulk (literature)	Nano-indentation (measured)	Normal beam bending (measured)	Bulk (literature)	Nano-indentation (measured)	Bulk (literature)	Normal beam bending (measured)	Lateral beam bending (measured)	Tensile strength (literature)	Flexural strength (literature)	Normal beam bending (measured) ^f	Lateral beam bending (measured) ^f	
PPMA	—	—	1.7	0.7	—	110	—	54	—	—	—	600	—	
PMMA	0.35 ^a	3.1–3.3 ^{c,d}	4.8	1.9	—	195 ^e	340	53.8–73.1 ^e	71	—	48–76 ^d	—	944	—
PS	0.325 ^b	3.2–3.4 ^c	3.6	1.9	0.2	110 ^d	290	—	66	21	30–60 ^d	95 ^d	271	—
PS/clay	—	—	5.1	4.6	0.6	—	390	—	95	64	—	—	128	110

^aRohm and Haas General Information on PMMA^bBoundy and Boyer (1952)^cBrandup et al. (1999)^dMark (1999)^eAnonymous (1996)

^fValues may be an overestimate due to limitations of the technique and the geometry of the beams
All values have $\pm \sigma$ of about 10% based on measurements on 5 samples for nanoindentation and 3–5 samples for beam bending except for the 10–40% variation in breaking strength

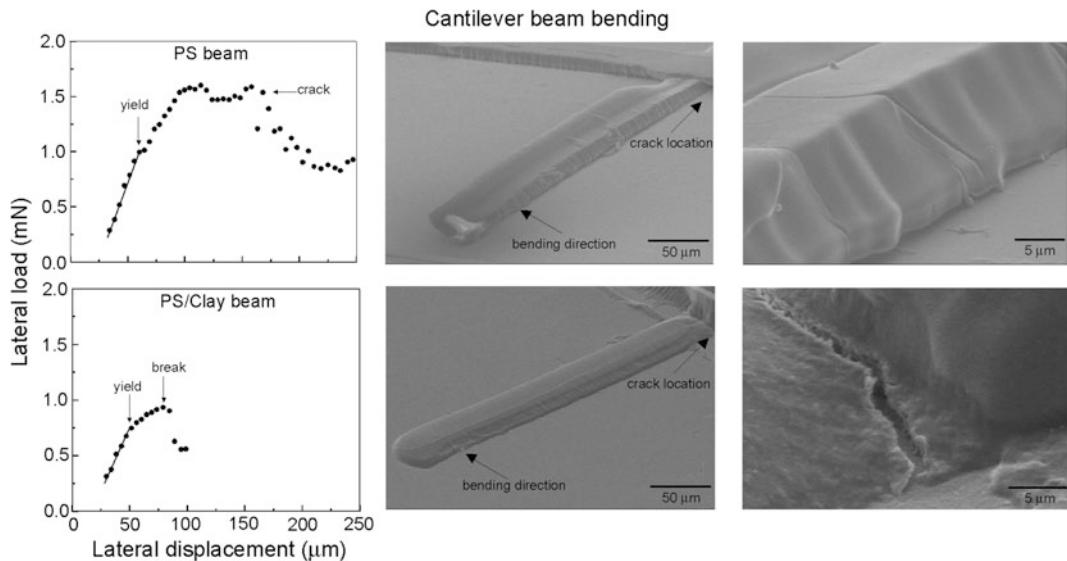


Fig. 6.25 Lateral bending of PS and PS/Clay cantilever beams. Lateral force is plotted as a function of the lateral displacement, where the arrows indicate the onset of yield for both PS and PS/Clay, cracking for PS and breaking for PS/Clay. SEM images of the cantilever beams indicate the bending direction and the location of the cracks. The images at the right hand column show the cracks at a higher magnification (Palacio et al. 2007a)

in the equation for evaluating the modulus (Eqs. 6.5 and 6.11).

From the load-displacement curves, the yield strength was evaluated by using (Young and Budynas 2002):

$$\sigma = \frac{6F\ell}{bh^2} \quad (6.14)$$

where F , ℓ , b and h are defined similarly as in Eq. 6.13. The force used for calculating the yield strength was taken at the point where a change in slope was observed (as indicated by the arrows on the left). The results are listed in Table 6.4. They are lower than the values obtained from normal bending, and this could be attributed to the same causes mentioned above.

For PS, a steady decrease in the lateral load was observed at 160 μm , while for the PS/Clay beam, the drop at 80 μm was abrupt (as indicated by the arrows on the right). SEM images reveal that for PS, this corresponds to crack formation at the base of the beam that did not run throughout the entire thickness of the beam. However, for PS/Clay, the beam broke. This implies that the

addition of the nanoclay filler induces embrittlement of the beam.

Equation 6.14 is assumed valid for evaluating the breaking strength of this beam. The force used in the calculation was the load prior to the abrupt decrease. The breaking strength for PS/Clay obtained from this technique (listed in Table 6.4) is lower than the value from the normal bending experiment. This is possibly due to the beam geometry and processing conditions, as described earlier. However, the large difference in the calculated yield strength between the unfilled and filled PS demonstrates that the technique can adequately differentiate between material compositions.

6.4 Finite Element Analysis of Nanostructures with Roughness and Scratches

Micro/nanostructures have some surface topography and local scratches dependent upon the manufacturing process. Surface roughness and local scratches may compromise the reliability of

the devices, and their effect needs to be studied. Finite element modeling is used to perform parametric analysis to study the effect of surface roughness and scratches in different well defined forms on tensile stresses which are responsible for crack propagation (Bhushan and Agrawal 2002, 2003). The analysis has been carried out on trapezoidal beams supported at the bottom whose data (on Si and SiO_2 nanobeams) have been presented earlier.

The finite element analysis has been carried out by using the static analysis of ANSYS 5.7 which calculates the deflections and stresses produced by applied loading. The type of element selected for the study was SOLID95 type which allows the use of different shapes without much loss of accuracy. This element is 3-D with 20 nodes, each node having three degrees of freedom which implies translation in the x, y and z directions. The nanobeam cross section is divided into six elements each along the width and the thickness and forty elements along the length. SOLID95 has plasticity, creep, stress stiffening, large deflection and large strain capabilities. The large displacement analysis is used for large loads. The mesh is kept finer near the asperities and the scratches in order to take into account variation in the bending stresses. The beam materials studied are made of single-crystal silicon (110) and SiO_2 films whose data have been presented earlier. Based on the bending experiments presented earlier, the beam materials can be assumed to be linearly elastic isotropic materials. Young's modulus of elasticity (E) and Poisson's ratio (ν) for Si and SiO_2 are 169 GPa (Bhushan and Venkatesan 1993) and 0.28 (Anonymous 1988), and 73 GPa (Bhushan and Gupta 1997) and 0.17 (Bhushan and Gupta 1997), respectively. A sample nanobeam of silicon was chosen for performing most of the analysis as silicon is the most widely used MEMS/NEMS material. The cross section of the fabricated beams used in the experiment is trapezoidal and supported at the bottom so nanobeams with trapezoidal cross section are modeled, Fig. 6.10. The following dimensions are used $w_1 = 200$ nm, $w_2 = 370$ nm, $t = 255$ nm, and $P = 6 \mu\text{m}$. In the boundary conditions, the displacements are

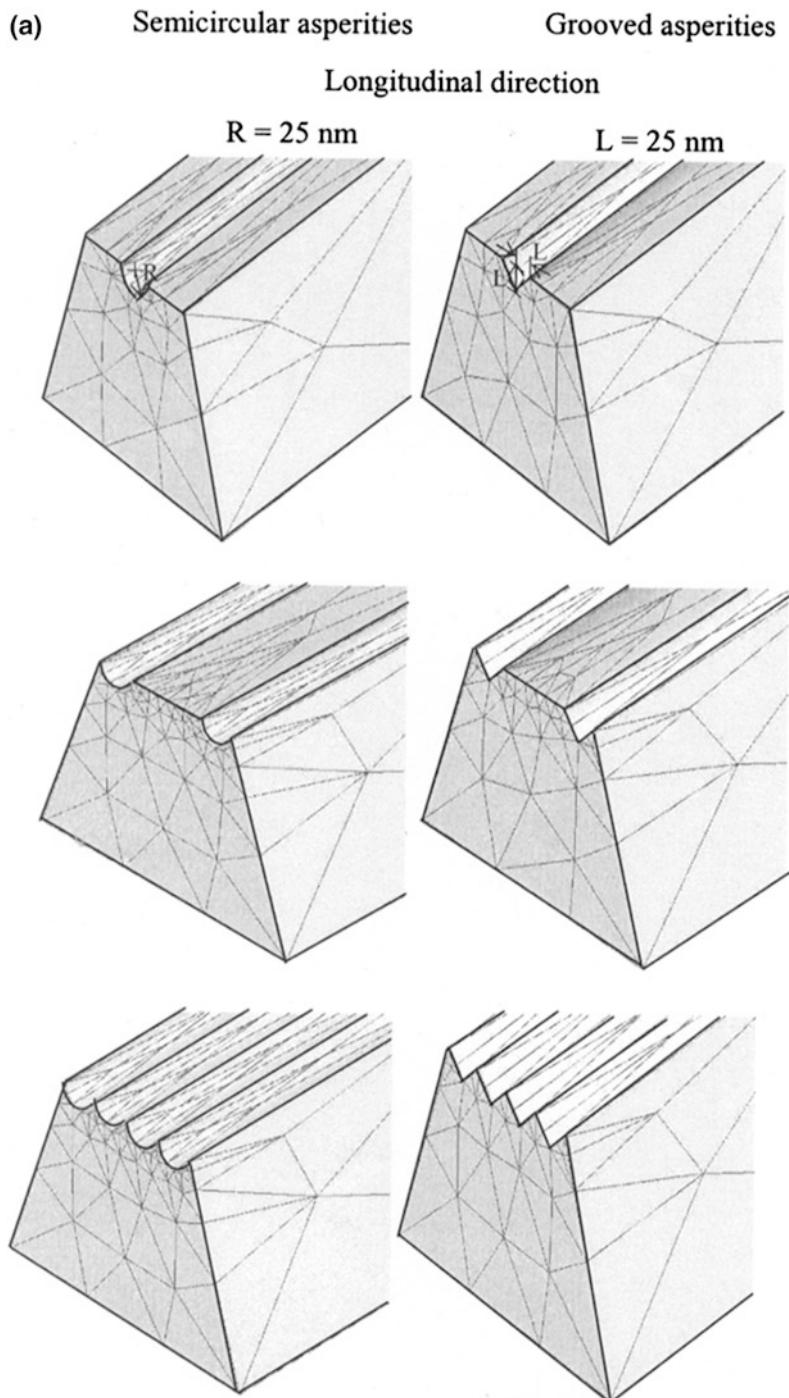
constrained in all directions on the bottom surface for 1 μm from each end. A point load applied at the center of the beam is simulated with the load being applied at three closely located central nodes on the beam used. It has been observed from the experimental results that the Si nanobeam breaks at around 80 μN . Therefore, in this analysis, a nominal load of 70 μN is selected. At this load, deformations are large, and large displacement option is used.

To study the effect of surface roughness and scratches on the maximum bending stresses the following cases were studied. First the semicircular and grooved asperities in the longitudinal direction with defined geometrical parameters are analyzed, Fig. 6.26a. Next semicircular asperities and scratches placed along the transverse direction at a distance c from the end and separated by pitch p from each other are analyzed, Fig. 6.26b. Lastly the beam material is assumed to be either purely elastic, elastic-plastic, or elastic-perfectly plastic. In the following, we begin with the stress distribution in smooth nanobeams followed by the effect of surface roughness in the longitudinal and transverse directions and scratches in the transverse direction.

6.4.1 Stress Distribution in a Smooth Nanobeam

Figure 6.27 shows the stress and vertical displacement contours for a nanobeam supported at the bottom and loaded at the center, (Bhushan and Agrawal 2002, 2003). As expected, the maximum tensile stress occurs at the ends while the maximum compressive stress occurs under the load at the center. Stress contours obtained at a section of the beam from the front and side are also shown. In the beam cross section, the stresses remain constant at a given vertical distance from one side to another and change with a change in vertical location. This can be explained due to the fact that the bending moment is constant at a particular cross section so the stress is only dependent on the distance from the neutral axis. However, in cross section A-A the high

Fig. 6.26 **a** Plots showing the geometries of modeled roughness—semicircular and grooved asperities along the nanobeam length with defined geometrical parameters, **b** Schematic showing semicircular asperities and scratches in the transverse direction followed by the illustration of the mesh created on the beam with fine mesh near the asperities and the scratches. Also shown are the semicircular asperities and scratches at different pitch values



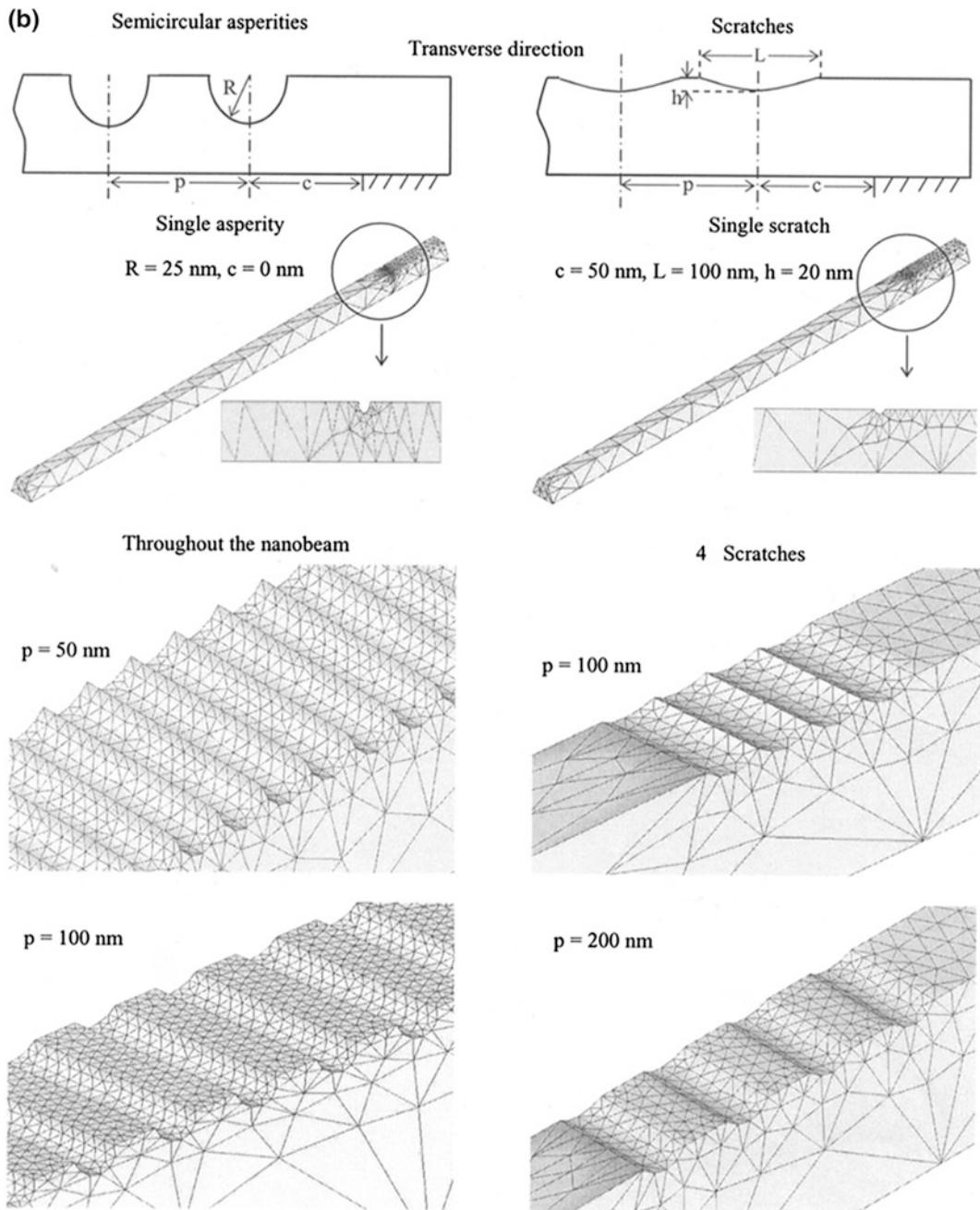


Fig. 6.26 (continued)

tensile and compressive stresses are localized near the end of the beam at top and bottom, respectively, whereas the lower values are spread

out away from the ends. High value of tensile stresses occurs near the ends because of high bending moment.

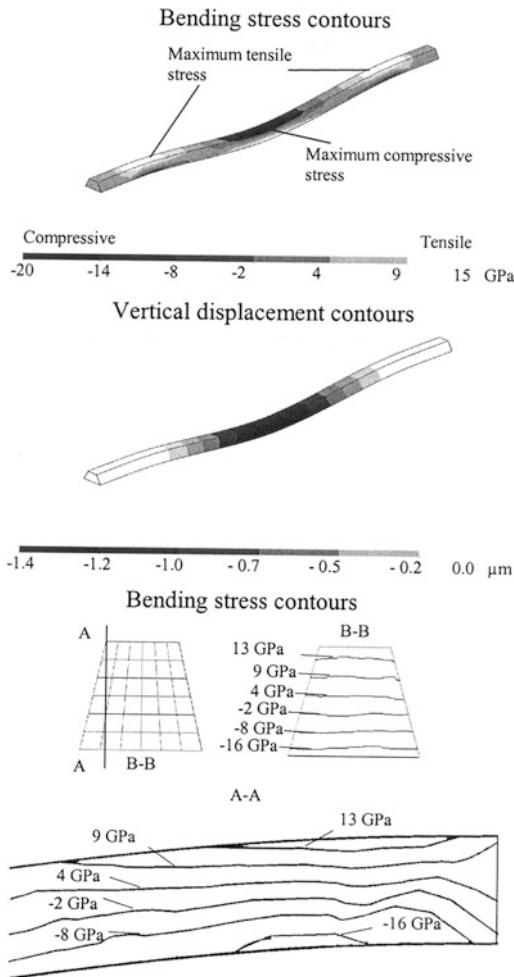


Fig. 6.27 Bending stress contours, vertical displacement contours, and bending stress contours after loading trapezoidal Si nanobeam ($w_1 = 200$ nm, $w_2 = 370$ nm, $t = 255$ nm, $\ell = 6$ μm , $E = 169$ GPa, $v = 0.28$) at 70 μN load (Bhushan and Agrawal 2003)

6.4.2 Effect of Roughness in the Longitudinal Direction

The roughness in the form of semicircular and grooved asperities in the longitudinal direction on the maximum bending stresses are analyzed (Bhushan and Agrawal 2003). The radius R and depth L are kept fixed at 25 nm while the number of asperities is varied, and their effect is observed on the maximum bending stresses. Figure 6.28 shows the variation of maximum bending

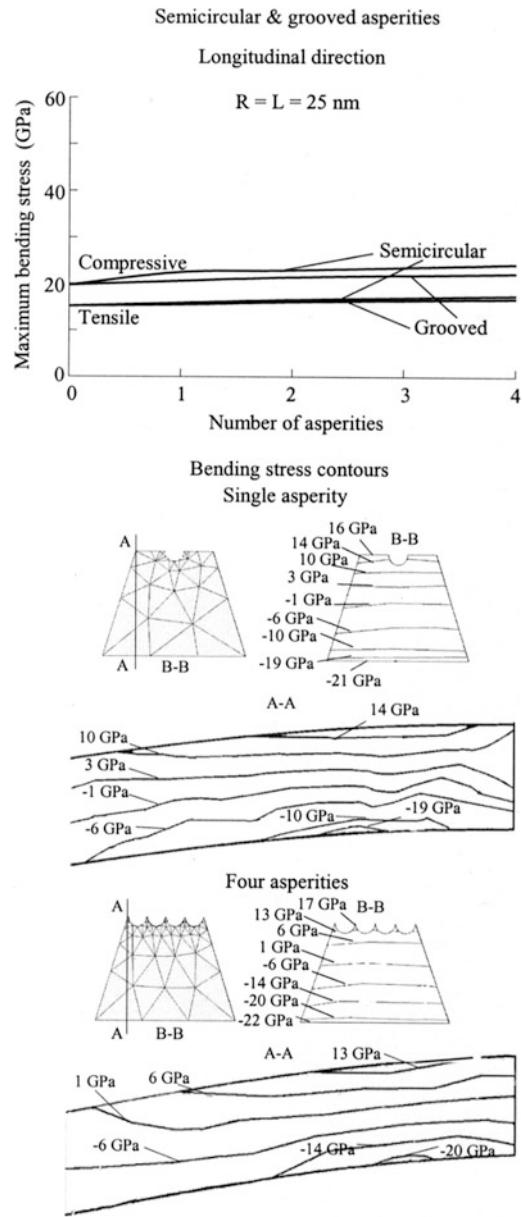


Fig. 6.28 Effect of longitudinal semicircular and grooved asperities in different numbers on maximum bending stresses after loading trapezoidal Si nanobeams ($w_1 = 200$ nm, $w_2 = 370$ nm, $t = 255$ nm, $\ell = 6$ μm , $E = 169$ GPa, $v = 0.28$, Load = 70 μN). Bending stress contours obtained in the beam with semicircular single asperity and four adjacent asperities of $R = 25$ nm (Bhushan and Agrawal 2003)

stresses as a function of asperity shape and the number of asperities. The maximum bending stresses increase as the asperity number increases

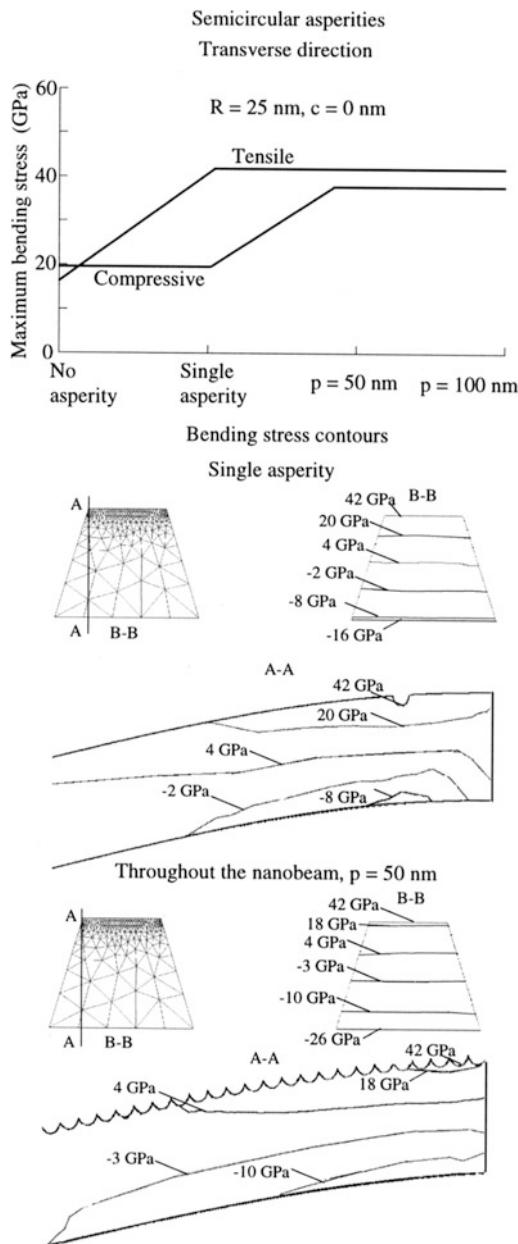


Fig. 6.29 Effect of transverse semicircular asperities located at different pitch values on the maximum bending stresses after loading trapezoidal Si nanobeams ($w_1 = 200 \text{ nm}$, $w_2 = 370 \text{ nm}$, $t = 255 \text{ nm}$, $\ell = 6 \mu\text{m}$, $E = 169 \text{ GPa}$, $v = 0.28$, Load = $70 \mu\text{N}$). Bending stress contours obtained in the beam with semicircular single asperity and semicircular asperities throughout the nanobeam surface at $p = 50 \text{ nm}$ (Bhushan and Agrawal 2003)

for both semicircular and grooved asperities. This can be attributed to the fact that as asperity number increases, the moment of inertia decreases for that cross section. Also the distance from the neutral axis increases because the neutral axis shifts downwards. Both these factors lead to the increase in the maximum bending stresses, and this effect is more pronounced in the case of a semicircular asperity as it exhibits a higher value of maximum bending stress than that in a grooved asperity. Figure 6.28 shows the stress contours obtained at a section of the beam from the front and from the side for both cases when we have a single semicircular asperity and when four adjacent semicircular asperities are present. Trends are similar to that observed earlier for a smooth nanobeam (Fig. 6.27).

6.4.3 Effect of Roughness in the Transverse Direction and Scratches

We analyze semicircular asperities when placed along the transverse direction followed by the effect of scratches on the maximum bending stresses in varying numbers and different pitch (Bhushan and Agrawal 2003). In the analysis of semicircular transverse asperities three cases were considered which included a single asperity, asperities throughout the nanobeam surface separated by pitch equal to 50 nm and pitch equal to 100 nm. In all of these cases, c value was kept equal to 0 nm. Figure 6.29 shows that the value of maximum tensile stress is 42 GPa which is much larger than the maximum tensile stress value with no asperity of 16 GPa or when the semicircular asperity is present in the longitudinal direction. It is also observed that the maximum tensile stress does not vary with the number of asperities or the pitch, while the maximum compressive stress does increase dramatically for the asperities present throughout the beam surface from its value when a single asperity is present. Maximum tensile stress occurs at the ends. An increase in p does not add any asperities at the ends whereas

asperities are added in the central region where compressive stresses are maximum. The semicircular asperities present at the center cause the local perturbation in the stress distribution at the center of the asperity where load is being applied leading to a high value of maximum compressive stress (Timoshenko and Goodier 1970). Figure 6.29 also shows the stress contours obtained at a section of the beam from front and from the side for both cases when there is a single semi-circular asperity and when asperities are present throughout the beam surface at a pitch equal to 50 nm. Trends are similar to that observed earlier for a smooth nanobeam (Fig. 6.27).

In the study pertaining to scratches, the number of scratches are varied along with the variation in the pitch as well. Furthermore the load is applied at the center of the beam and at the center of the scratch near the end as well for all the cases. In all of these cases, c value was kept equal to 50 nm, and L value was equal to 100 nm with h value being 20 nm. Figure 6.30 shows that the value of maximum tensile stress remains almost the same with the number of scratches for both types of loading, that is, when load is applied at the center of the beam and at the center of the scratch near the end. This is because the maximum tensile stress occurs at the beam ends, no matter where the load gets applied. But the presence of a scratch does increase the maximum tensile stress as compared to its value for a smooth nanobeam, although the number of scratches no longer matters as the maximum tensile stress occurring at the nanobeam end is unaffected by the presence of more scratches beyond the first scratch in the direction towards the center. The value of the tensile stress is much lower when the load is applied at the center of the scratch, and it can be explained as follows. The negative bending moment at the end near the load applied decreases with load offset after two-thirds of the length of the beam (Shigley and Mitchell 1993). Since this negative bending moment is responsible for tensile stresses, their behavior with the load offset is the same as the negative bending moment. Also, the value of maximum compressive stress when load is applied at the center of the nanobeam remains

almost the same, as the center geometry is unchanged due to the number of scratches, and hence the maximum compressive stress occurring below the load at the center is same. On the other hand when the load is applied at the center of the scratch we observe that the maximum compressive stress increases dramatically because the local perturbation in the stress distribution at the center of the scratch where load is being applied leads to a high value of maximum compressive stress (Timoshenko and Goodier 1970). It increases further with the number of scratches and then levels off. This can be attributed to the fact that when there is another scratch present close to the scratch near the end, the stress concentration is more, as the effect of local perturbation in the stress distribution is more significant. However, this effect is insignificant when more than two scratches are present.

Now we address the effect of pitch on the maximum compressive stress when the load is applied at the center of the scratch near the end. When the pitch is up to a value of 200 nm the maximum compressive stress increases with the number of scratches, as discussed earlier. On the other hand, when the pitch value goes beyond 225 nm this effect is reversed. This is because the presence of another scratch no longer affects the local perturbation in the stress distribution at the scratch near the end. Instead more scratches at a fair distance distribute the maximum compressive stress at the scratch near the end, and the stress starts going down. Such observations of maximum bending stresses can help in identifying the number of asperities and scratches allowed separated by an optimum distance from each other.

6.4.4 Effect on Stresses and Displacements for Materials Which Are Elastic, Elastic-Plastic or Elastic-Perfectly Plastic

This section deals with the beam modeled as elastic, elastic-plastic and elastic-perfectly plastic to observe the variation in the stresses and

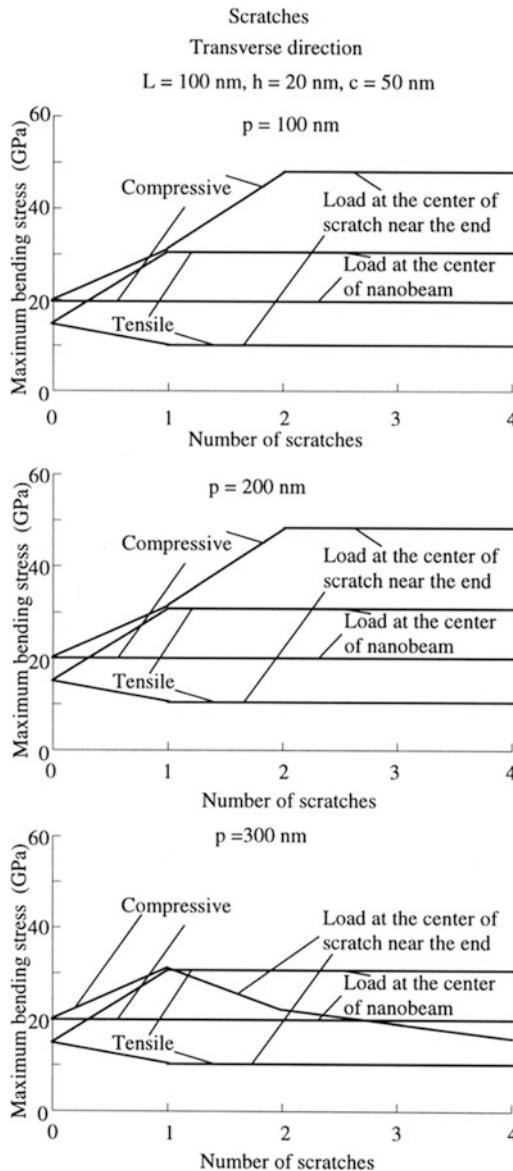


Fig. 6.30 Effect of number of scratches along with the variation in the pitch on the maximum bending stresses after loading trapezoidal Si nanobeams ($w_1 = 200 \text{ nm}$, $w_2 = 370 \text{ nm}$, $t = 255 \text{ nm}$, $\ell = 6 \mu\text{m}$, $E = 169 \text{ GPa}$, $v = 0.28$, Load = 70 μN) Also shown is the effect of load when applied at the center of the beam and at the center of the scratch near the end (Bhushan and Agrawal 2003)

displacements from an elastic model used so far (Bhushan and Agrawal 2003). Figure 6.31 shows the typical stress-strain curves for the three types

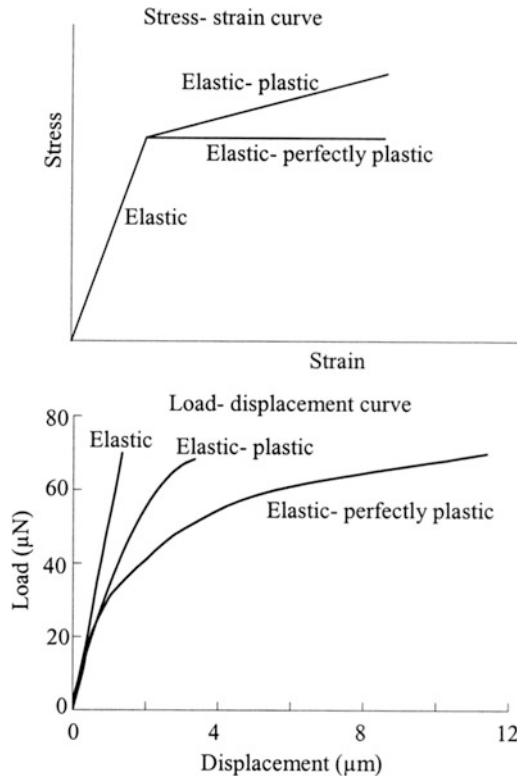


Fig. 6.31 Schematic representation of stress-strain curves and load-displacement curves for material when it is elastic, elastic-plastic, or elastic-perfectly plastic for a Si nanobeam ($w_1 = 200 \text{ nm}$, $w_2 = 370 \text{ nm}$, $t = 255 \text{ nm}$, $\ell = 6 \mu\text{m}$, $E = 169 \text{ GPa}$, tangent modulus in plastic range = 0.5 E, $v = 0.28$) (Bhushan and Agrawal 2003)

of deformation regimes and their corresponding load-displacement curves obtained from the model of an Si nanobeam which are found to exhibit same trends.

Table 6.5 shows the comparison of the maximum von Mises stress and maximum displacements for both a smooth nanobeam and a nanobeam with a defined roughness, which is a single semicircular longitudinal asperity of R value equal to 25 nm for the three different models. It is observed that the maximum value of stress is obtained at a given load for elastic material whereas the displacement is maximum for elastic-perfectly plastic material. Also the pattern that the maximum bending stress value increases for a rough nanobeam still holds true in the other models as well.

Table 6.5 Stresses and displacements for materials which are elastic, elastic-plastic, or elastic-perfectly plastic (load = 70 μ N, w_1 = 200 nm, w_2 = 370 nm, t = 255 nm, ℓ = 6 μ m, R = 25 nm, E = 169 GPa, tangent modulus in plastic range = 0.5 E, v = 0.28)

	Elastic		Elastic-plastic		Elastic-perfectly plastic	
	Smooth nanobeam	Single semicircular longitudinal asperity	Smooth nanobeam	Single semicircular longitudinal asperity	Smooth nanobeam	Single semicircular longitudinal asperity
Max. von mises stress (GPa)	18.2	19.3	13.5	15.2	7.8	9.1
Max. displacement (μ m)	1.34	1.40	3.35	3.65	11.5	12.3

6.5 Closure

Mechanical properties of nanostructures at relevant scales are required to design realistic MEMS/NEMS and BioMEMS/BioNEMS devices. Most mechanical properties are scale dependent. Micro/nanomechanical properties, hardness, elastic modulus, and scratch resistance of bulk materials of undoped single-crystal silicon (Si) and thin films of undoped polysilicon, SiO_2 , SiC, Ni-P, and Au were presented. It was found that the SiC film exhibits higher hardness, elastic modulus, and scratch resistance as compared to other materials.

Bending tests have been performed on the Si and SiO_2 nanobeams and Ni-P, Au, PPMA and PMMA, PS and PS/clay microbeams using an AFM and a depth-sensing nanoindenter, respectively. The bending tests were used to evaluate elastic modulus, bending strength (fracture stress), fracture toughness (K_{IC}), and fatigue strength of the beam materials. The Si and SiO_2 nanobeams exhibited an elastic linear response with sudden brittle fracture. The notched Ni-P beam showed linear deformation behavior followed by abrupt failure. The Au beam showed elastic-plastic deformation behavior. Measured elastic modulus values for Si(110) and SiO_2 were comparable to bulk values. Measured bending strength for Si and SiO_2 were twice as large as the values reported on larger scale specimens. This indicates that

bending strength shows a specimen size dependence. Measured fracture toughness value for Si and SiO_2 were also comparable to values obtained on larger specimens. At stress amplitudes less than 15% of their bending strength and at mean stresses of less than 30% of the bending strength, Si and SiO_2 displayed an apparent endurance life of greater than 30,000 cycles. SEM observations of the fracture surfaces revealed a cleavage type of fracture for both materials when subjected to bending as well as fatigue.

The hardness, elastic modulus, and creep behavior of the PPMA, PMMA, PS and a PS/Clay nanocomposite microbeams were evaluated using nanoindentation in continuous stiffness mode. The H, E, and creep resistance decreased in the following order: PS/Clay, PMMA, PS, PPMA. The scratch resistance was observed to decrease in the following order: PMMA, PS/Clay, PS, PPMA.

The elastic modulus of the double anchored beams obtained from normal beam bending was lower than the E measured from nanoindentation due to stress concentration at the beam ends. The test environment affected the mechanical properties as well. After 36 h of soaking in DI water, the hardness and elastic modulus of PPMA was affected, while the properties of the other three polymers were unaffected due to their low water absorption properties. Similarly, for tests conducted at human body temperature (37.5 °C), only PPMA exhibited a decrease in the hardness

and modulus. The other materials did not exhibit any change because while this temperature is within the glass transition temperature range of PMMA, it is well below the T_g of the other polymers. By performing bending experiments at higher imposed loads, permanent deformation can be attained such that the yield and breaking strength can be evaluated. Adding nanoclay filler improved bending strength. It also modified the fracture behavior of the polymer microbeams. The unfilled polymers deformed in a ductile manner while the PS/Clay nanocomposite exhibited minimal yielding followed by brittle failure.

Lateral bending has been demonstrated to provide a measure for the elastic, plastic and fracture properties of cantilever beams.

The AFM and nanoindenters used in this study can be satisfactorily used to evaluate the mechanical properties of micro/nanoscale structures for use in MEMS/NEMS.

FEM simulations were used to predict stress and deformation in nanostructures. FEM has been used to analyze the effect of the type of surface roughness and scratches on stresses and deformation of nanostructures. We found that roughness affects the maximum bending stresses. The maximum bending stresses increased as the asperity number increased for both semicircular and grooved asperities in longitudinal direction. When the semicircular asperity was present in the transverse direction, the value of the maximum tensile stress was much larger than the maximum tensile stress value with no asperity or when the semicircular asperity was present in the longitudinal direction. This observation suggests that an asperity in the transverse direction is more detrimental. The presence of scratches increased the maximum tensile stress. The maximum tensile stress remained almost the same with the number of scratches for two types of loading, that is, when load was applied at the center of the beam or at the center of the scratch near the end, although the value of the tensile stress was much lower when the load was applied at the center of the scratch. This means that the load applied at the ends was less damaging. This analysis shows that FEM

simulations can be useful to designers to develop the most suitable geometry for nanostructures.

Appendix A: Fabrication Procedure for the Double Anchored and Cantilever Beams

The starting materials for the polymer microbeams are PPMA (molecular weight 250,000, Scientific Polymer Products), PMMA (molecular weight 75,000, Sigma-Aldrich), PS (melt flow index 4.0, Sigma-Aldrich) and PS/Clay solutions in anisole (Acros Organics). The clay additive in PS/Clay is Cloisite® 20A surface modified natural montmorillonite (Southern Clay Products, Inc.) with a thickness of approximately 1 nm and lateral dimensions of 70–150 nm. The surface modification of the clay additive improves the dispersion of the nanoparticles in the polymer matrix, thus improving the properties of the composite. To prepare the nanocomposite, the clay was first dispersed in the PS matrix by melt compounding at a concentration of 10% wt/wt (clay/PS). The composite was then dissolved in anisole and sonicated for at least 8 h to dissolve the polymer and redisperse the particles.

Double anchored polymer microbeams were fabricated using a soft lithography based micro-molding process along with standard photolithography (Palacio et al. 2007a). The process involves selectively filling a poly(dimethyl siloxane) (PDMS) mold with the polymer of interest followed by transfer of the resulting structures to a prefabricated substrate. This substrate is a silicon wafer with a layer of SU-8 25 negative tone photoresist (MicroChem Corp.) patterned by photolithography to create 25 μm wide channels, which is the resulting length of the suspended beams. (The second number 25 in the SU-8 designation relates to a viscosity appropriate for a given thickness.)

A patterned PDMS mold with the desired polymer beam geometry was fabricated from a photoresist master. Briefly, a layer of SU-8 5 photoresist was spin coated on silicon, and photolithography was used to define 5 μm wide

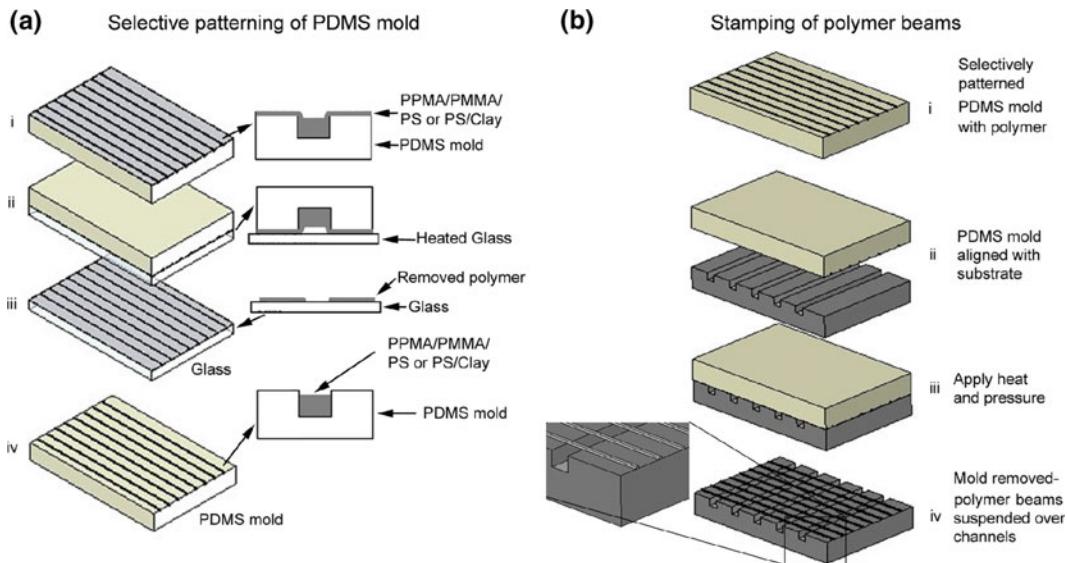


Fig. 6.32 Schematic of the micromolding process used in polymer double anchored beam fabrication. **a** Selective patterning of PDMS mold (i) mold is spin coated with a polymer layer. (ii) mold is inverted and brought into contact with a heated glass plate (iii) mold is removed from the plate, transferring the surface polymer onto the glass (iv) mold is left with polymer only in the recessed features. **b** Stamping of the polymer beams (i) Selectively patterned mold (ii) mold is inverted and aligned with the photolithographically patterned substrate (iii) mold and substrate are brought into contact under heat and pressure (iv) polymer is transferred onto the photoresist and mold is removed (Palacio et al. 2007a)

photoresist features separated by 45 μm gaps. A 10:1 ratio of T-2 PDMS translucent base and curing agent (Dow Corning) were mixed thoroughly and poured over the photoresist master to transfer the pattern into the PDMS. The mold was then placed in a vacuum dessicator to remove bubbles. The sample was removed from the vacuum periodically, and a razor blade was used to remove surface bubbles. After the bubbles were completely removed, the PDMS mold was allowed to cure at room temperature for 48 h before removing it from the wafer.

Next, the PDMS mold was selectively coated with the polymer solution to form the microbeams and then transferred to the substrate. Figure 6.32 is a schematic of the fabrication process used for making the polymer microbeams. As shown in Fig. 6.32a, the polymer solutions were spin coated on the PDMS mold for one minute. A 10% solution of polymer was spin coated on a mold with 5.3 μm deep features at 3000 RPM for fabrication of the beams in the bending experiments. After spin coating, the

mold was brought into contact with a heated glass plate to promote adhesion of the contacting polymer materials. This process removed the polymer material from the raised surfaces of the mold, resulting in the polymer remaining only in the recessed portions of the PDMS mold. The glass plate was heated to 175 °C for all four materials. As shown in Fig. 6.32b, the selectively coated mold was then aligned with the photolithographically patterned silicon substrate so that the beams of interest (PPMA, PMMA, PS, PS/Clay) ran perpendicular to the channels defined in the photoresist. The substrate was then heated and pressure was applied to the top of the mold to transfer the material onto the substrate. The transfer temperature for PPMA, PS, PS/Clay and PMMA were 95, 125, 125, and 175 °C, respectively, and the transfer pressure for all materials was around 0.21 MPa.

After removal of the mold, two types of polymer beams were transferred onto the wafer sample. The first type is supported beams, which was used to determine the hardness, elastic

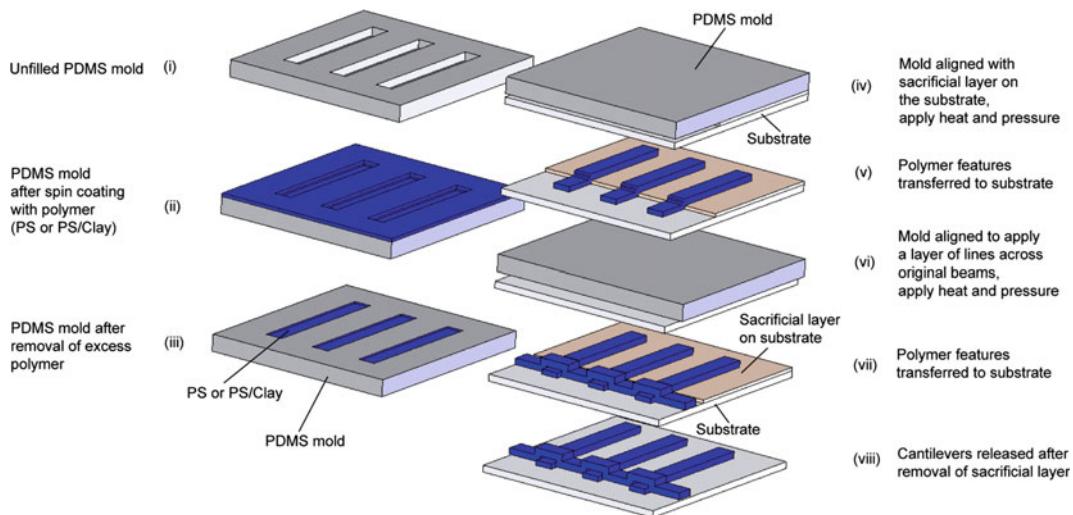


Fig. 6.33 Schematic of the polymer cantilever beam fabrication process (i) PDMS mold (ii) mold spin coated with the polymer of interest (iii) selectively filled mold after removal of surface polymer (iv) mold aligned with sacrificial layer (v) polymer features transfer to the substrate using heat and pressure (vi) mold aligned for a second time to add another layer of lines across original beams for reinforcement (vii) polymer features transfer to the substrate using heat and pressure (viii) cantilever structures after removal of the sacrificial layer (Palacio et al. 2007a)

modulus, and creep response. The second type is suspended beams, on which the bending experiments were performed. The double anchored beam samples studied by Palacio et al. (2007a) were 3–5 μm thick, 5 μm wide and nominally 25 μm long.

The process for fabrication of the polymer cantilever for lateral bending tests is shown in Fig. 6.33 (Palacio et al. 2007a). A PDMS mold was first cast from a photolithographically patterned SU-8/silicon master. The resulting molds consisted of 50 μm wide channels that were approximately 27 μm deep. The PDMS mold was then coated with the polymer of interest (PS or PS/Clay). The polymers were spin coated on the mold at 3000 RPM for one minute at concentrations of 15% and 10% for PS and PS/Clay, respectively. The polymer on the raised surface of the mold was removed by contacting the surface with a glass slide heated to 180 °C. The mold was then inverted and manually aligned with a silicon substrate coated with patches of polyvinyl alcohol (PVA), which acts as the sacrificial layer. The PVA was patterned using

photolithography and reactive ion etching with an oxygen plasma. The substrate was heated and pressure was applied to the top side of the mold to transfer the polymer from the recessed features of the mold onto the substrate such that a 350 μm portion of it is attached to the PVA and the remainder is attached to the bare silicon surface. Transfer temperatures for PS and PS/Clay were 150 and 175 °C, respectively. The process was repeated to apply another layer of beams across the length of the original layer. This was performed to provide reinforcement as it was observed that the one-layer design was not robust. The cantilever beam samples studied by Palacio et al. (2007a) were 12–27 μm thick, 60–80 μm wide and 350 μm long.

References

- Alexandre M, Dubois P (2000) Polymer-layered silicate nanocomposites: preparation, properties and uses of a new class of materials. Mater Sci Eng 28:1–63
- Anonymous, Rohm and Haas General Information on PMMA, Philadelphia, Pennsylvania

- Anonymous (1988) "Properties of Silicon," *EMIS Databeviews Series No. 4*, INSPEC. Institution of Electrical Engineers, London
- Anonymous (1996) Modern plastics encyclopedia. McGraw-Hill, New York
- Ballarini R, Mullen RL, Yin Y, Kahn H, Stemmer S, Heuer AH (1997) The fracture toughness of polysilicon microdevices: a first report. *J Mater Res* 12:915–922
- Becker H, Gaertner C (2000) Polymer microfabrication methods for microfluidic analytical applications. *Electrophoresis* 21:12–26
- Becker H, Lacascio LE (2002) Polymer microfluidic devices. *Talanta* 56:267–287
- Beebe DJ, Mensing GA, Walker GM (2002) Physics and applications of microfluidics in biology. *Annu Rev Biomed Eng* 4:261–286
- Bhatia SN, Chen CS (1999) Tissue engineering at the micro-scale. *Biomed Microdevices* 2:131–144
- Bhushan B (1998) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht
- Bhushan B (2001) Macro- and microtribology of MEMS materials. In: Bhushan B (ed) *Modern tribology handbook*. CRC Press, Boca Raton, FL, pp 1515–1548
- Bhushan B (2013a) Principles and applications of tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2016) Encyclopedia of nanotechnology, 2nd edn. Springer, Switzerland
- Bhushan B, Agrawal GB (2002) Stress analysis of nanostructures using a finite element method. *Nanotechnology* 13:515–523
- Bhushan B, Agrawal GB (2003) Finite element analysis of nanostructures with roughness and scratches. *Ultramicroscopy* 97:495–507
- Bhushan B, Gupta BK (1997) *Handbook of Tribology: Materials, Coatings, and Surface Treatments*. McGraw Hill, New York (1991); Reprint w/corrections, Krieger Publishing Co., Malabar, Florida (1997)
- Bhushan B, Li X (2003) Nanomechanical characterization of solid surfaces and thin films. *Int Mater Rev* 48: 125–164
- Bhushan B, Venkatesan S (1993) Mechanical and tribological properties of silicon for micromechanical applications: a review. *Adv Info Storage Syst* 5:211–239
- Bhushan B, Sundararajan S, Li X, Zorman CA, Mehregany M (1998) Micro/nanotribological studies of single-crystal silicon and polysilicon and SiC films for use in MEMS Devices. In: Bhushan B (ed) *Tribology issues and opportunities in MEMS*. Kluwer Academic Publishers, Dordrecht, pp 407–430
- Billmeyer FWJ (1984) *Textbook of polymer science*. Wiley, New York
- Boundy RH, Boyer RF (1952) (eds) *Styrene, its polymers, copolymers and derivatives*. Reinhold, New York
- Brandrup J, Immergut EH, Grulke EA (eds) (1999) *Polymer handbook*, 4th edn. Wiley, New York
- Cheng J, Kricka LJ (eds) (2001) *Biochip technology*. Harwood Academic Publishers, Philadelphia
- Connally JA, Brown SB (1993) Micromechanical fatigue testing. *Exp Mech* 33:81–90
- Decuzzi P, Lee S, Bhushan B, Ferrari M (2005) A theoretical model for the margination of particles with blood vessels. *Annals Biomed Eng* 33:179–190
- Dresselhaus MS, Dresselhaus G, Avouris Ph (2001) *Carbon Nanotubes—synthesis, structure, properties and applications*. Springer, Berlin
- Drexler KE (1992) *Nanosystems: molecular machinery, manufacturing and computation*. Wiley, New York
- Ellis B (2000) Polymers: a property database. Available on compact disk. CRC Press, Boca Raton, FL. Also see www.polymersdatabase.com
- Elwenspoek M, Wiegerink R (2001) *Mechanical microsensors*. Springer, Berlin
- Ericson F, Schweitz JA (1990) Micromechanical fracture strength of silicon. *J Appl Phys* 68:5840–5844
- Fitzgerald AM, Dauskardt RH, Kenny TW (2000) Fracture toughness and crack growth phenomena of plasma-etched single crystal silicon. *Sens Actuators A* 83:194–199
- Fujimasa I (1996) *Micromachines: a new era in mechanical engineering*. Oxford University Press, Oxford
- Gad-el-Hak M (2002) *The MEMS handbook*. CRC Press, Boca Raton, FL
- Goddard WA, Brenner DW, Lyshevski SE, Iafrate GJ (eds) (2002) *Handbook of nanoscience, engineering, and technology*. CRC Press, Boca Raton, FL
- Greek S, Ericson F, Johansson S, Furtsch M, Rump A (1999) Mechanical characterization of thick polysilicon films: young's modulus and fracture strength evaluated with microstructures. *J Micromech Microeng* 9:245–251
- Heinzelmann M, Petzold M (1994) FEM analysis of microbeam bending experiments using ultra-micro indentation. *Comput Mater Sci* 3:169–176
- Heller MJ, Guttman A (eds) (2001) *Integrated microfabricated biodevices*. Marcel Dekker, New York
- Hertzberg RW (1989) *Deformation and fracture mechanics of engineering materials*, 3edn. Wiley, New York, pp 277–278
- Hierlemann A (2005) *Integrated chemical microsensor systems in CMOS technology*. Springer, Berlin
- Holt DB, Gauger PR, Kusterbeck AW, Ligler FS (2002) Fabrication of a capillary immunosensor in polymethyl methacrylate. *Biosens Bioelec* 17:95–103
- Hornbeck LJ (1999) A digital light processing™ update—status and future applications. *Proc Soc Photo-Opt Eng* 3634:158–170 (*Projection Displays V*)
- Hsu TR (2002) *MEMS and microsystems: design and manufacture*. McGraw-Hill, Boston, MA
- Hsu T, Sun N (1998) Residual stresses/strains analysis of MEMS. In: Laudon M, Romanowicz B (eds) *Proceedings of the international conference on modeling and simulation of microsystems, semiconductors, sensors and actuators*. Computational Publications, Cambridge, MA, pp 82–87
- Johansson S, Schweitz JA, Tenerz L, Tiren J (1988) Fracture testing of silicon microelements in-situ in a

- scanning electron microscope. *J Appl Phys* 63:4799–4803
- Johansson S, Ericson F, Schweitz JA (1989) Influence of surface-coatings on elasticity, residual-stresses, and fracture properties of silicon microelements. *J Appl Phys* 65:122–128
- Kahn H, Huff MA, Heuer AH (1998) Heating effects on the young's modulus of films sputtered onto micro-machined resonators. *Microelectromechanical Structures for Materials Research, Symposium Proc*, vol 518. Materials Research Society, Pittsburgh, Pennsylvania, pp 33–38
- Kahn H, Ballarini R, Mullen RL, Heuer AH (1999) Electrostatically actuated failure of microfabricated polysilicon fracture mechanics specimens. *Proc R Soc Lond A* 455:3807–3823
- Kolpekwar A, Kellen C, Blanton RD (Shawn) (1998) Fault model generation for MEMS. In: Laudon M, Romanowicz B (eds) *Proceedings of the international conference on modeling and simulation of microsystems, semiconductors, sensors and actuators*. Computational Publications, Cambridge, MA, pp 111–116
- Komai K, Minoshima K, Inoue S (1998) Fracture and fatigue behavior of single-crystal silicon microelements and nanoscopic AFM damage evaluation. *Microsyst Technol* 5:30–37
- Kovacs GTA (1998) *Micromachined transducers sourcebook*. WCB McGraw-Hill, Boston, MA
- Lanza RP, Langer R, Vacanti J (eds) (2000) *Principles of tissue engineering*. Academic Press, San Diego
- LaVan DA, Buchheit TE (1999) Strength of polysilicon for MEMS devices. In: *Proceedings of SPIE 3880—MEMS reliability for critical and space applications*, pp 40–44
- Lawn BR, Evans AG, Marshall DB (1980) Elastic/plastic indentation damage in ceramics: the median/radial system. *J Am Ceram Soc* 63:574
- Leclerc E, Furukawa KS, Miyata F, Sakai T, Ushida T, Fujii T (2004) Fabrication of microstructures in photosensitive biodegradable polymers for tissue engineering applications. *Biomaterials* 25:4683–4690
- Li X, Bhushan B (1999) Micro/nanomechanical characterization of ceramic films for microdevices. *Thin Solid Films* 340:210–217
- Li X, Bhushan B (2003) Fatigue studies of nanoscale structures for MEMS/NEMS applications using nanoindentation techniques. *Surface Coatings Technol* 163–164:521–526
- Li X, Bhushan B, Takashima K, Baek CW, Kim YK (2003) Mechanical characterization of micro/nanoscale structures for MEMS/NEMS applications using nanoindentation techniques. *Ultramicroscopy* 97:481–494
- Loh IH, Sheu MS, Fisher AB (1997) Biocompatible polymer surfaces. In: Arshady R (ed) *Desk reference of functional polymers: syntheses and applications*. American Chemical Society, Washington, D.C.
- Madou M (2011) *Fundamentals of microfabrication and nanotechnology*, 3rd edn. CRC Press, Boca Raton, FL
- Manz A, Becker H (eds) (1998) *Microsystem technology in chemistry and life sciences, topics in current chemistry* 194. Springer, Heidelberg
- Mark JE (1999) *Polymers data handbook*. Oxford University Press, New York
- Mastrangelo CH, Becker H (eds) (2001) *Microfluidics and BioMEMS*. *Proceedings of SPIE*, vol 4560. SPIE, Bellingham, Washington
- Mazza E, Dual J (1999) Mechanical behavior of a μm -sized single crystal silicon structure with sharp notches. *J Mech Phys Solids* 47:1795–1821
- McDonald JC, Duffy DC, Anderson JR, Chiu DT, Wu H, Schueller OJA, Whitesides GM (2000) Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis* 21:27–40
- Nalwa HS (ed) (2002) *Nanostructures materials and nanotechnology*. Academic Press, San Diego, CA
- Namazu T, Isono Y, Tanaka T (2000) Evaluation of size effect on mechanical properties of single-crystal silicon by nanoscale bending test using AFM. *J Microelectromech Syst* 9:450–459
- Nguyen CTC, Howe RT (1999) An integrated CMOS micromechanical resonator high-Q oscillator. *IEEE J Solid-St Circ* 34:440–455
- Obermeier E (1996) Mechanical and thermophysical properties of thin film materials for mems: techniques and devices. *Micromechanical structures for materials research, Symposium Proceedings*, vol 444. Materials Research Society, Pittsburgh, Pennsylvania, pp 39–57
- Oeberg PA, Togawa T, Spelman FA (2004) *Sensors in medicine and health care*. Wiley, New York
- Palacio M, Bhushan B, Ferrell N, Hansford D (2007a) Nanomechanical characterization of polymer beam structures for BioMEMS applications. *Sens Actuators A* 135:637–650
- Palacio M, Bhushan B, Ferrell N, Hansford D (2007b) Adhesion properties of polymer/silicon interfaces for biological micro-/nanoelectromechanical applications. *J Vac Sci Technol A* 25:1275–1284
- Park K (ed) (1997) *Controlled drug delivery: challenges and strategies*. American Chemical Society, Washington, D.C.
- Petersen KE (1982) Silicon as a mechanical material. *Proc IEEE* 70:420–457
- Poole CP, Owens FJ (2003) *Introduction to nanotechnology*. Wiley, Hoboken, New Jersey
- Raiteri R, Grattarola M, Butt H, Skladal P (2001) Micromechanical cantilever-based biosensors. *Sens Actuators B: Chem* 79:115–126
- Ray SS, Okamoto M (2003) Polymer/layered silicate nanocomposites: a review from preparation to processing. *Prog Polym Sci* 28:1539–1641
- Rietman EA (2001) *Molecular engineering of nanosystems*. Springer, New York
- Rueda HA, Law ME (1998) Modeling of strain in boron-doped silicon cantilevers. In: Laudon M, Romanowicz B (eds) *Proceedings of the international conference on modeling and simulation of*

- microsystems, semiconductors, sensors and actuators. Computational Publications, Cambridge, MA, pp 94–99
- San CLP, Yap EPH (eds) (2001) Frontiers in human genetics. World Scientific, Singapore
- Sato K, Yoshioka T, Anso T, Shikida M, Kawabata T (1998) Tensile testing of silicon film having different crystallographic orientations carried out on a silicon chip. *Sens Actuators A* 70:148–152
- Schulte TH, Bardell RL, Weigl BH (2002) Microfluidic technologies in clinical diagnostics. *Clin Chim Acta* 321:1–10
- Senturia SD (2000) Microsystem design. Kluwer Academic, Boston, MA
- Sharpe WN Jr, Yuan B, Edwards RL (1997) A new technique for measuring the mechanical properties of thin films. *J Microelectromech Syst* 6:193–199
- Shawgo RS, Grayson ACR, Li Y, Cima MJ (2002) BioMEMS for drug delivery. *Current Opinion Solid State Mat Sci* 6:329–334
- Shigley JE, Mitchell LD (1993) Mechanical engineering design, 4th edn. McGraw-Hill, New York
- Sundararajan S, Bhushan B (1998) Micro/nanotribological studies of polysilicon and SiC films for MEMS applications. *Wear* 217:251–261
- Sundararajan S, Bhushan B (2002) Development of AFM-based techniques to measure mechanical properties of nanoscale structures. *Sens Actuators A* 101:338–351
- Sundararajan S, Bhushan B, Namazu T, Isono Y (2002) Mechanical property measurements of nanoscale structures using an atomic force microscope. *Ultramicroscopy* 91:111–118
- Tanaka M (1996) Fracture toughness and crack morphology in indentation fracture of brittle materials. *J Mater Sci* 31:749
- Tang WC, Lee AP (2001) Defense applications of MEMS. *MRS Bull* 26:318–319
- Taylor MR, Nguyen P, Ching J, Peterson KE (2003) Simulation of microfluidic pumping in a genomic DNA blood-processing cassette. *J Micromech Microeng* 13:201–208
- Timoshenko SP, Goodier JN (1970) Theory of elasticity, 3rd edn. McGraw-Hill, New York
- Timp G (ed) (1999) Nanotechnology. Springer, New York
- Trimmer WS (ed) (1997) Micromachines and MEMS, classic and seminal papers to 1990. IEEE Press, New York
- Tsuchiya T, Tabata O, Sakata J, Taga Y (1998) Specimen size effect on tensile strength of surface-micromachined polycrystalline silicon thin films. *J Microelectromech Syst* 7:106–113
- Tsuchiya T, Inoue A, Sakata J (2000) Tensile testing of insulating thin films: humidity effect on tensile strength of SiO₂ films. *Sens Actuators A* 82:286–290
- van der Berg A (ed) (2003) Lab-on-a-chip: chemistry in miniaturized synthesis and analysis systems. Elsevier, Amsterdam
- van Kooten TG, Spijkerv HT, Busscher HH (2004) Plasma-treated polystyrene surfaces: model surface for studying cell-biomaterial interactions. *Biomaterials* 25:1735–1747
- Wei G, Bhushan B, Ferrell N, Hansford D (2005) Microfabrication and nanomechanical characterization of polymer MEMS for biological applications. *J Vac Sci Technol A* 23:811–819
- Wilson CJ, Beck PA (1996) Fracture testing of bulk silicon microcantilever beams subjected to a side load. *J Microelectromech Syst* 5:142–150
- Wilson CJ, Ormeggi A, Narbutovskikh M (1996) Fracture testing of silicon microcantilever beams. *J Appl Phys* 79:2386–2393
- Yi T, Kim CJ (1999) Measurement of mechanical properties for MEMS materials. *Measur Sci Technol* 10:706–716
- Yi T, Li L, Kim CJ (2000) Microscale material testing of single crystalline silicon: process effects on surface morphology and tensile strength. *Sens Actuators A* 83:172–178
- Young WC, Budynas RG (2002) Roark's formulas for stress and strain. McGraw-Hill, New York
- Zoval JV, Madou MJ (2000) Centrifuge-based fluidic platforms. *Proc IEEE* 92:140–153

Computer Simulations of Nanometer-Scale Indentation and Friction

Susan B. Sinnott, Seong-Jun Heo, Donald W. Brenner,
Judith A. Harrison and Douglas L. Irving

Abstract

Engines and other machines with moving parts are often limited in their design and operational lifetime by friction and wear. This limitation has motivated the study of fundamental tribological processes with the ultimate aim of controlling and minimizing their impact. The recent development of miniature apparatus, such as microelectromechanical systems and nanometer-scale devices, has increased interest in atomic-scale friction, which has been found to, in some cases, be due to mechanisms that are significantly distinct from the mechanisms that dominate in macroscale friction. Presented in this chapter is a review of computational studies of tribological processes at the atomic and nanometer scale. In particular, a review of the findings of computational studies of nanometer-scale indentation, friction, and lubrication is presented, along with a review of the salient computational methods that are used in these studies, and the conditions under which they are best applied.

Keywords

Molecular dynamics · Nanomechanics · Atomic friction

S.B. Sinnott (✉)

Department of Materials Science and Engineering,
The Pennsylvania State University, 221A Steidle
Building, University Park, PA 16802, USA
e-mail: sinnott@mse.psu.edu

S.-J. Heo

Department of Materials Science and Engineering,
University of Florida, Gainesville, FL 32611-6400,
USA

D.W. Brenner · D.L. Irving

Department of Materials Science and Engineering,
North Carolina State University, Raleigh, NC
27695-7907, USA

J.A. Harrison

Department of Chemistry, U.S. Naval Academy,
Annapolis, MD 21402, USA

7.1 Introduction

Engines and other machines with moving parts are often limited in their design and operational lifetime by friction and wear. This limitation has motivated the study of tribological processes with the aim of controlling and minimizing the impact of these processes. There are numerous historical examples that illustrate the importance of friction to the development of civilizations, including the ancient Egyptians who invented technologies to move the stones used to build the pyramids (Dowson 1979); Coulomb, who was motivated to study friction by the need to move ships easily and without wear from land to the water (Dowson 1979; Johnson et al. 1971), who developed an improved understanding of contact mechanics and surface energies through the study of automobile windshield wipers. At present, substantial research and development is aimed at microscale and nanoscale machines with moving parts that at times challenge our fundamental understanding of friction and wear. This has motivated the study of atomic-scale friction and has, consequently, led to new discoveries such as self-lubricating surfaces and wear-resistant materials. While there are similarities between friction at the macroscale and the atomic scale, in many instances the mechanisms that lead to friction at these two scales are quite different. Thus, as devices such as magnetic storage disks and microelectromechanical systems (MEMS) (Gad-el-Hak 2002) continue to shrink in size, it is expected that new phenomena associated with atomic-scale friction, adhesion and wear will dominate the functioning of these devices.

The last two decades have seen considerable scientific effort expended on the study of atomic-scale friction (Krim 1995, 1996a, b; Sutton 1996; Mate 1995; Stoneham et al. 1993; Singer 1994; Bhushan et al. 1995; Harrison and Brenner 1995; Sokoloff 1992; Zhong et al. 1991; Israelachvili 1992a; Sinnott 2000; Heo et al. 2005). This effort has been facilitated by the development of advanced new experimental tools to measure friction over nanometer-scale distances at low loads, rapid improvements in computer power, and the maturation of computational methodologies for the modeling of materials at the atomic

scale. For example, friction-force and atomic-force microscopes (FFM and AFM) allow the frictional properties of solids to be characterized with atomic-scale resolution under single asperity indentation and sliding conditions (Binnig et al. 1986; Mate et al. 1987; Germann et al. 1993; Carpick and Salmeron 1997). In addition, the surface force apparatus (SFA) provides data about the tribological and lubrication responses of many liquid and solid systems with atomic resolution Israelachvili (1992b) and the quartz crystal microbalance (QCM) provides information about the atomic-scale origins of friction (Krim 1996a; Krim et al. 1991). These and related experimental methods allow researchers to study sliding surfaces at the atomic scale and relate the observed phenomena to macroscopically-observed friction, lubrication and wear.

Analytic models and computational simulations have played an important role in characterizing and understanding friction. They can, for example, assist in the interpretation of experimental data or provide predictions that subsequent experiments can confirm or refute. Analytic models have long been used to study friction, including early studies by Tomlinson (1929) and Frenkel and Kontorova (1938) and more recent studies by McClelland et al. (1992), Sokoloff (1984, 1990, 1992, 1993, 1995, 1996, 2001, 2002), Persson et al. (1991), Volokitin and Persson (2003a, b), Persson (2005), Helman et al. (1994), Kawaguchi and Matsukawa (1997), Muser (2002a, b), Reimann and Evstigneev (2004), Ritter et al. (2005), Fusco and Fasolino (2005). Most of these idealized models divide the complex motions that create friction into more fundamental components defined by quantities such as spring constants, the curvature and magnitude of potential wells, and bulk phonon frequencies. While these simplifications provide these approaches with some predictive capabilities, many assumptions must be made to apply these models to study friction that may lead to incorrect or incomplete results.

In atomic scale molecular dynamics (MD) simulations, atom trajectories are calculated by numerically integrating coupled classical equations of motion. Interatomic forces that enter these equations are typically calculated either from total energy methods that include electronic degrees of

freedom, or from simplified mathematical expressions that give the potential energy as a function of interatomic displacements. MD simulations can be considered numerical experiments that provide a link between analytic models and experiments. The main strength of MD simulations is that they can reveal unanticipated phenomena or unexpected mechanisms for well-known observations. Weaknesses include a lack of quantum effects in classical atomistic dynamics, and perhaps more importantly, the fact that meaningless results can be obtained if the simulation conditions are incorrectly chosen. The next section contains a review of MD simulations, including the approximations that are inherent in their application to the study of friction, and the conditions under which they should and should not be applied.

7.2 Computational Details

Molecular dynamics simulations are straight-forward to describe: given a set of initial conditions and a way of mathematically modeling interatomic forces, Newton's (or equivalent) classical equation of motion is numerically integrated (Gear 1971)

$$F = ma \quad (7.1a)$$

$$-\nabla E = m(\partial^2 r / \partial t^2) \quad (7.1b)$$

where F is the force on each atom, m is the atomic mass, a is the atomic acceleration, E is the potential energy felt by each atom, r is the atomic position, and t is time. The forces acting on any given atom are calculated, and then the atoms move a short increment Δt (called a time step) forward in time in response to these applied forces. This is accompanied by a change in atomic positions, velocities, and accelerations. The process is then repeated for some specified number of time steps.

The output of these simulations includes new atomic positions, velocities, and forces that allow additional quantities such as temperature and pressure to be determined. As the size of the system increases, it is useful to render the atomic positions in animated movies that reveal the responses of the system in a qualitative manner.

Quantitative data can be obtained by analyzing the numerical output directly.

The following sections review the way in which energies and forces are calculated in MD simulations and the important approximations that are used to realistically model the friction that occurs in experiments with smaller systems of only a few tens of thousands of atoms in simulations. The reader is referred to additional sources (Allen and Tildesley 1987; Brenner 1989a, 2000; Finnis 2003; Haile 1992; Harrison et al. 2008; Heermann 1986; Hoover 1986) for a more comprehensive overview of MD simulations (including computer algorithms), analysis methods, and the potentials that are used to calculate energies and forces in MD simulations.

7.2.1 Energies and Forces

There are several different approaches by which interatomic energies and forces are determined in MD simulations. The most theoretically rigorous methods are those that are classified as ab initio or first principles. These techniques, which include density functional theory (Parr and Yang 1989; Car and Parrinello 1985) and quantum chemical ab initio (Cramer 2004) methods, are derived from quantum mechanical principles and are generally both the most accurate and the most computationally intensive. They are therefore limited to a small number of atoms (<500), which has limited their use in the study of friction. Alternatively, empirical methods are functions containing parameters that are determined by fitting to experimental data or the results of ab initio calculations (Finnis 2003). These techniques can usually be relied on to correctly describe qualitative trends and are often the only choice available for modeling systems containing tens of thousands, millions, or billions of atoms. Empirical methods have therefore been widely used in studies of friction. Semi-empirical methods, including tight-binding methods, include some elements of both empirical methods and ab initio methods. For instance, they require quantum mechanical information in the form of, for example, on-site and hopping matrix elements, and include fits to experimental data (Sutton 1993).

Empirical methods simplify the modeling of materials by treating the atoms as spheres that interact with each other via repulsive and attractive terms that can be either pairwise additive or many-body in nature. In this approach electrons are not treated explicitly, although it is understood that the interatomic interactions are ultimately dependent on them. As discussed in this section, some empirical methods explicitly include charge through classical electrostatic interactions although most methods assume charge-neutral systems. The repulsive and attractive functional forms generally depend on interatomic distances and/or angles and contain adjustable parameters that are fit to ab initio results and/or experimental data.

The main strength of empirical potentials is their computational speed. Recent simulations with these approaches have modeled billions of atoms (Kadau et al. 2004), something that is not possible with ab initio or semi-empirical approaches at this time. The main weakness of empirical potentials is their lack of quantitative accuracy, especially if they are poorly formulated or applied to systems that are too far removed from the fitting database used in their construction. Furthermore, because of the differences in the nature of chemical bonding in various materials, e.g., covalent bonding in carbon versus metallic bonding in gold, empirical methods have been historically derived for particular classes of materials. They are therefore generally non-transferable, although some methods have been shown to be theoretically equivalent (Brenner 1989b; Thijssse 2002) and in recent years there has been progress towards the development of empirical methods that can model heterogeneous material systems (Baskes et al. 1989a; Ohira et al. 2001; Streitz and Mintmire 1994; Yasukawa 1996; Iwasaki and Miura 2001; Lee and Baskes 2000).

Several of the most important and common general classes of empirical methods for calculating interatomic energies and forces in materials, the so-called potentials, are reviewed here. The first to be considered are the potentials that are used to model covalently bound materials,

including the bond-order potential and the Stillinger-Weber potential.

The bond-order potential was first formulated by Abell (1985) and subsequently developed and parameterized by Tersoff for silicon and germanium (Tersoff 1988, 1989), Brenner and co-workers for hydrocarbons (Brenner 1990a, 2000; Brenner et al. 2002), Dyson and Smith for carbon–silicon–hydrogen systems (Dyson and Smith 1996), Sinnott and co-workers for carbon–oxygen–hydrogen systems (Ni et al. 2004), and Graves and co-workers (Tanaka et al. 2000), Sinnott and co-workers (Jang and Sinnott 2004) for fluorocarbons, Schall and coworkers for pure Si (Schall et al. 2008), and Hu and coworkers on C–N–O–H (Hu 2008).

The bond-order potential has the general functional form

$$E = \sum_i \sum_{j(>i)} [V_R(r_{ij}) - b_{ij} V_A(r_{ij})] \quad (7.2)$$

where $V_R(r)$ and $V_A(r)$ are pair-additive interactions that model the interatomic repulsion and electron-nuclear attraction, respectively. The quantity r_{ij} is the distance between pairs of nearest-neighbor atoms i and j , and b_{ij} is a bond order-term that takes into account the many-body interactions between atoms i and j , including those due to nearest neighbors and angle effects. The potential is short-ranged and only considers nearest neighbor bonds. To model long-range non-bonded interactions, the bond-order potential is combined with pair-wise potentials either directly through splines (Sinnott et al. 1998) or indirectly with more sophisticated functions (Stuart et al. 2000).

The Stillinger-Weber potential (Stillinger and Weber 1985) potential was formulated to model silicon, with a particular emphasis on the liquid phases of silicon. It includes many-body interactions in the form of a sum of two and three-body interactions

$$E = \sum_{ij} V_{ij}^2(r_{ij}) + \sum_{jik} V_{jik}^3(r_{ij}, r_{ik}) \quad (7.3)$$

where V^2 is a pair-additive interaction and V^3 is a three-body term. The three-body term includes

an angular interaction that minimizes the potential energy for tetrahedral angles. This term favors the formation of open structures, such as the diamond cubic crystal structure.

The second potential is the embedded atom method (EAM) approach (Foiles 1985; Daw and Baskes 1983) and related methods (Raeker and Depristo 1991), which were initially developed for modeling metals and alloys. The functional form in the EAM is

$$E = \sum_i F(\rho_i) + \sum_{i>j} \Phi(r_{ij}) \quad (7.4)$$

where F is called the embedding energy. This term models the energy due to embedding an atom into a uniform electron gas with a uniform compensating positive background (jellium) of density ρ_i that is equal to the actual electron density of the system. The term $\Phi(r_{ij})$ is a pairwise functional form that corrects for the jellium approximation. Several parameterizations of the EAM exist [see, e.g., (Foiles 1985; Daw and Baskes 1983; Smith and Was 1989; Pasianot et al. 1991; Pasianot and Savino 1992)] and it has recently been extended to model non-metallic systems. For example, the modified EAM (MEAM) approach (Lee and Baskes 2000; Baskes et al. 1989b; Baskes 1992) was developed so that EAM could be applied to metal oxides (Ohira et al. 2001) and covalently bound materials (Baskes 1992).

The third method is the general class of Coulomb or multipole interaction potentials used to model charged ionic materials or molecules (Ohno et al. 1999). In this formalism, an energy term is given as

$$E = \sum_i \sum_{j(>i)} [(q(r_i)q(r_j)/r_{ij})] \quad (7.5)$$

where $q(r_i)$ is the charge on atom i and r_{ij} is the distance between atoms i and j . More complex formalisms that take into account, e.g., the Madelung constant in the case of ionic crystals, are used in practice. In general, the charges are held fixed, but methods that allow charge to vary in a realistic manner have been developed

(Streitz and Mintmire 1994; Rappe and Goddard 1991).

Lastly, long-range van der Waals or related forces are typically modeled with pairwise additive potentials. A widely used approximation is the Lennard-Jones (LJ) potential (Frenkel and Smit 1996), which has the following functional form:

$$E = 4\epsilon \sum_i \sum_{j(>i)} [(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6] \quad (7.6)$$

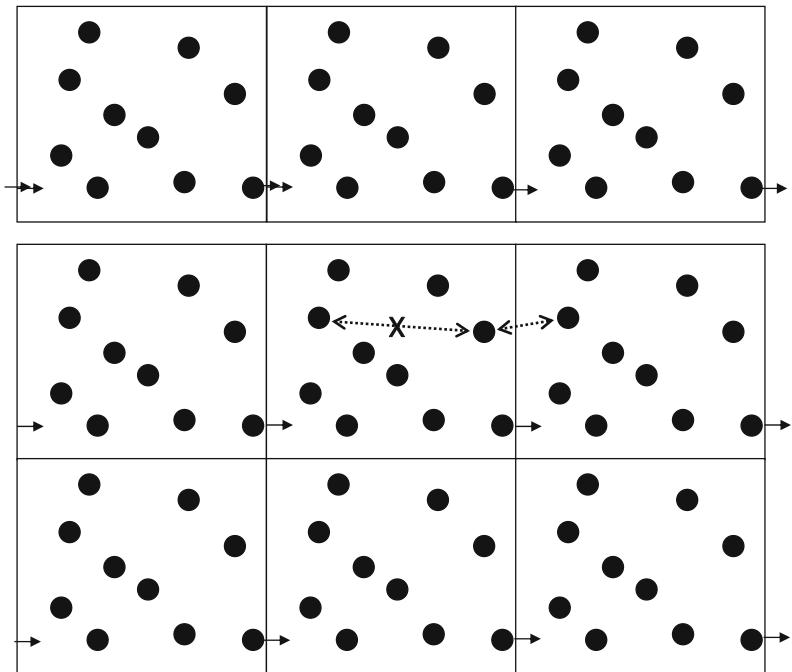
In this approach ϵ and σ are parameters and r_{ij} is the distance between atoms i and j .

All of these potentials are widely used in MD simulations of materials, including studies of friction, lubrication, and wear.

7.2.2 Important Approximations

Several approximations are typically used in MD simulations of friction. The first is the use of periodic boundary conditions (PBCs) and the minimum image convention for interatomic interactions (Allen and Tildesley 1987). In both cases the simulation supercell is surrounded by replicas of itself so that atoms (or phonons, etc.) that exit one side of the supercell remerge into the simulation through the opposite side of the supercell. In the minimum image convention an atom interacts either with another atom in the supercell or its equivalent atom in a surrounding cell depending on which distance to the atom is shortest. This process is illustrated in Fig. 7.1. In this convention supercells must be large enough that atoms do not interact with themselves over the periodic boundaries. In computational studies of friction and wear, PBCs are usually applied in the two dimensions within the plane(s) of the sliding surface(s). The strength of this approach is that it allows a finite number of atoms to model an infinite system. However, the influence of boundaries on system dynamics is not completely eliminated; for example, phonon scattering due to the periodic boundaries can influence heat transport and therefore frictional properties of sliding interfaces.

Fig. 7.1 Illustration of periodic boundary conditions consisting of a central simulation cell surrounded by replica systems. The *solid arrows* indicated an atom leaving the central box and re-entering on the opposite side. The *dotted arrows* illustrate the minimum image convention



Another important tool that is often used in MD simulations of friction is thermostats to regulate system temperature. In macroscopic systems, heat that is generated from friction is dissipated rapidly from the surface to the bulk phonon modes. Because atomistic computer simulations are limited systems that are many orders of magnitude smaller than systems that are generally studied experimentally, thermostats are needed to prevent the system temperature from rising in a non-physical manner. Typically in simulations of indentation or friction, the thermostat is applied to a region of the simulation cell that is well removed from the interface where friction and indentation is taking place. In this way, local heating of the interface that occurs as work is done on the system, but excess heat is efficiently dissipated from the system as a whole. In this manner the adjustment of atomic temperatures occurs away from the processes of interest, and simplified approximations for the friction term can be used without unduly influencing the dynamics produced by the interatomic forces.

There are several different formalisms for atomistic thermostats. The simplest of these controls the temperature by intermittently

rescaling the atomic velocities to values corresponding to the desired temperature (Woodcock 1971) such that

$$\left(\frac{\mathbf{v}_{new}}{\mathbf{v}_{old}}\right)^2 = \frac{T}{T_{ins}} \quad (7.7)$$

where \mathbf{v}_{new} is the rescaled velocity, and \mathbf{v}_{old} is the velocity before the rescaling. This approach, which is called the velocity rescaling method, is both simple to implement and effective at maintaining a given temperature over the course of an MD simulation. It was consequently widely used in early MD simulations. The velocity rescaling approach does have some significant disadvantages, however. First, there is little theoretical basis for the adjustment of atomic velocities, and the system dynamics are not time reversible, which is inconsistent with classical mechanics. Second, the rate and mode of heat dissipation are disconnected from system properties, which may affect system dynamics. Lastly, for typical MD simulation system sizes, the averaged quantities that are obtained, such as pressure for instance, do not correspond to values in any thermodynamic ensemble.

For these reasons more sophisticated methods for maintaining system temperatures in MD simulations have been developed. The Langevin dynamics approach (Allen and Tildesley 1987), which was originally developed from the theory of Brownian motion, falls into this category. In this approach, terms are added to the interatomic forces that correspond to a random force and a frictional term (Hoover 1986; Schneider and Stoll 1978; Kremer and Grest 1990). Therefore, Newton's equation of motion for atoms subjected to Langevin thermostats is given by the following equation rather than Eqs. (7.1a) and (7.1b):

$$m\ddot{\mathbf{a}} = \mathbf{F} - m\xi\dot{\mathbf{v}} + \mathbf{R}(t) \quad (7.8)$$

where \mathbf{F} are the forces due to the interatomic potential, ξ is a friction coefficient, m and \mathbf{v} are the particle's mass and velocity, respectively, and $\mathbf{R}(t)$ is a random force that acts as "white noise". The friction term can be formulated in terms of a memory kernel, typically for harmonic solids (Adelman and Doll 1976; Adelman 1980; Tully 1980), or a friction coefficient can be approximated using the Debye frequency. The random force can be given by a Gaussian distribution where the width, which is chosen to satisfy the fluctuation-dissipation theorem, is determined from the equation

$$\langle \mathbf{R}(0) \cdot \mathbf{R}(t) \rangle = 2mkT\xi\delta(t). \quad (7.9)$$

Here, the function \mathbf{R} is the random force in Eq. (7.8), m is the particle mass, T is the desired temperature, k is Boltzmann's constant, t is time, and ξ is the friction coefficient. It should be noted that the random forces are uncoupled from those at previous steps, which is denoted by the delta function. Additionally, the width of the Gaussian distribution from which the random force is obtained varies with temperature. Thus, the Langevin approach does not require any feedback from the current temperature of the system as the random forces are determined solely from Eq. (7.9).

In the early 1980s, Nosé developed a new thermostat that corresponds directly to a canonical ensemble (system with constant temperature,

volume and number of atoms) (Nosé 1984a, b), which is a significant advance from the methods described so far. In this approach, Nosé introduces a degree of freedom s that corresponds to the heat bath and acts as a time scaling factor, and adds a parameter Q that may be regarded as the heat bath "mass". A simplified form of Nosé's method was subsequently implemented by Hoover (Hoover 1986) that eliminated the time scaling factor while introducing a thermodynamic friction coefficient ζ . Hoover's formulation of Nosé's method is therefore easy to use and is commonly referred to as the Nosé-Hoover thermostat.

When this thermostat is applied to a system containing N atoms, the equations of motion are written as (dots denote time derivatives):

$$\begin{aligned} \dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \zeta\mathbf{p}_i \\ \dot{\zeta} &= \frac{1}{Q} \left(\sum_{i=1}^N \frac{p_i^2}{m_i} - N_f k_B T \right) \end{aligned} \quad (7.10)$$

where \mathbf{r}_i is the position of atom i , \mathbf{p}_i is the momentum and \mathbf{F}_i is the force applied to each atom. The last equation in Eq. (7.10) contains the temperature control mechanism in the Nosé-Hoover thermostat. In particular, the term between the parentheses on the right-hand side of this equation is the difference between the system's instantaneous kinetic energy and the kinetic energy at the desired temperature. If the instantaneous value is higher than the desired one, the friction force will increase to lower it and vice versa.

It should be pointed out that the choice of the heat bath "mass" Q is arbitrary but crucial to the successful performance of the thermostat. For example, a small value of Q leads to rapid temperature fluctuation while large Q values result in inefficient sampling of phase space. Nosé recommended that Q should be proportional to $N_f k_B T$ and should allow the added degree of freedom s to oscillate around its averaged value at a frequency of the same order as the characteristic frequency of the physical system (Nosé

1984a, b). If ergodic dynamic behavior is assumed, the Nosé-Hoover thermostat will maintain a well-defined canonical distribution in both momentum and coordinate space. However, for small systems where the dynamic is not ergodic, the Nosé-Hoover thermostat fails to generate a canonical distribution. Therefore, more sophisticated algorithms based on the Nosé-Hoover thermostat have been proposed to fix its ergodicity problem, for example, the “Nosé-Hoover chain” method of Martyna et al. (1992). However, these complex thermostats are not as easy to apply as the Nosé-Hoover thermostat due to the difficult evaluation of the coupling parameters for each different case and the significant computational cost (D’Alessandro et al. 2002). From a practical point of view, if the molecular system is large enough so that the movements of the atoms are sufficiently chaotic, ergodicity is guaranteed and the performance of the Nosé-Hoover thermostat is satisfactory (Frenkel and Kontorova 1938).

In an alternative approach, Schall et al. recently introduced a hybrid continuum-atomistic thermostat (Schall et al. 2005). In this method an MD system is divided into grid regions, and the average kinetic energy in the atomistic simulation is used to define a temperature for each region. A continuum heat transfer equation is then solved stepwise on the grid using a finite difference approximation, and the velocities of the atoms in each grid region are scaled to match the solution of the continuum equation. To help account for a time lag in the transfer of kinetic to potential energy, Hoover constraining forces are added to those from the interatomic potential. This process is continued, leading to an ad hoc feedback between the continuum and atomistic simulations. The main advantage of this approach is that the experimental thermal diffusivity can be used in the continuum expression, leading to heat transfer behavior that matches experimental data. For example, in metals the majority of the thermal properties at room temperature arise from electronic degrees of freedom that are neglected with strictly classical potentials. This thermostat is relatively straight forward to implement, and requires only the

interatomic potential and the bulk thermal diffusivity as input. It is also appropriate for nonequilibrium heat transfer, such as occurs as heat is dissipated from surfaces sliding moving at high relative velocities. Other localized phenomena, such as Joule heating and melting in current carrying applications, can also be simulated by using a recent extension to the hybrid thermostat (Padgett and Brenner 2005). This modification allows the ability to model degradation of interfaces under high electrical load at the atomic level. Relevant examples are hot switched radio frequency MEMS and metal/metal contacts in electromagnetic launchers.

Schoen et al. (1989) and Curry et al. (1994) developed a unique alternative to the grand canonical ensemble by performing a series of grand canonical Monte Carlo simulations (Allen and Tildesley 1987; Adams 1975) at various points along a hypothetical sliding trajectory. The results from these simulations are then used to calculate the correct particle numbers at a fixed chemical potential, which are then used as inputs to non-sliding, constant-*NVE* MD simulations at each of the chosen trajectory points. The sliding speed can be assumed to be infinitely slow because the system is fully equilibrated at each step along the sliding trajectory. This approach offers a useful alternative to continuous MD simulations that are restricted to sliding speeds that are orders of magnitude larger than most experimental studies (about 1 m/s or greater).

To summarize, this section provides a brief review and description of components that are used in atomistic, molecular dynamics simulation of many of the processes related to friction, such as indentation, sliding, and wear. The components discussed here include the potential energy expression used to calculate energies and forces in the simulations, periodic boundary conditions and thermostats. Each of these components has their own strengths and weaknesses that should be well-understood both prior to their use and in the interpretation of results. For example, general principles related to liquid lubrication in confined areas may be most easily understood and generalized from simulations that use pair potentials

and may not require a thermostat. On the other hand, if one wants to study the wear or indentation of a surface of a particular metal, then EAM or other semiempirical potentials, together with a thermostat, would be expected to yield more reliable results. If one requires information on electronic effects, ab initio or semi-empirical approaches that include the evaluation of electronic degrees of freedom must be used. Thus, the best combination of components for a particular study depends on the chemical nature of the system of interest, the processes being simulated, the type of information desired, and the available computational resources.

7.3 Indentation

It is critical to understand the nanometer-scale properties of materials that are being considered for use as new coatings with specific friction and wear behavior. The experimental determination of these properties is most frequently done with the AFM, which provides a variety of data related to the interaction of the microscope tips with the sample surface (Burnham and Colton 1993; Meyer 1998; Totten and Liang 2004). In AFM experiments, the tip has a radius of about 1–100 nm and is pressed against the surface under ambient conditions (in air), ultra-high vacuum (UHV) conditions, or in a liquid. The microscope tip can either move in the direction normal to the surface, which is the case in nanoindentation studies, or raster across the surface, which is the case in surface imaging or friction studies. Sliding rates of 1 nm/s–1 μm/s are typically used, which are many orders of magnitude slower than the rates used in MD simulations of sliding or indentation of around 1–100 m/s. As discussed in the previous section, the higher rates used in computational simulations are a consequence of modeling full atomic motion, which occurs on a femto- to picosecond timescale, and the step-wise solution of the classical equations of motion, which makes the large number of simulation steps needed to reach experimental timescales computationally impossible with current processor speeds.

As the tip moves either normal to, or across the surface, the forces acting upon it as a result of its interactions with the surface are measured. When the tip is moved in the surface normal direction, it can penetrate the surface on the nanometer scale and provide information on the nanometer-scale mechanical properties of the surface (Burnham and Colton 1989; Burnham et al. 1990). The indentation process also causes the force on the tip to increase, and the rate of increase is related to both the depth of indentation and the properties of the surface. The region of the force curve that reflects this high force is known as the repulsive wall region (Burnham and Colton 1993), or, when considered without any lateral motion of the tip, an indentation curve. When the tip is retracted after indentation, enhanced adhesion between the tip and surface relative to the initial contact can result. This phenomenon is indicated by hysteresis in the force curve.

Tip-surface adhesion can result from the formation of chemical bonds between the tip and the sample, or from the formation of liquid capillaries between the microscope tip and the surface caused by the interaction of the tip with a layer of liquid contamination on the surface. The latter case is especially prevalent in AFM studies conducted in ambient environments. In the case of clean metallic systems, the sample can wet the tip or the tip can wet the sample in the form of a connective “neck” of metal atoms between the surface and the tip that can lead to adhesion. In the case of polymeric or molecular systems, entanglement of molecules that are anchored on the tip with molecules anchored on the sample can be responsible for force curve hysteresis.

In the case of horizontal rastering of AFM tips across surfaces, the force curve data provide a map of the surface that is indicative of the surface topography (Meyer et al. 1992). If the deflection of the tip in the lateral direction is recorded while the tip is being rastered, a friction map of the surface (Germann et al. 1993) is produced.

The rest of this section discusses some of the important insights and findings that have been obtained from MD simulations of nanoindentation. These studies have not only provided

insight into the physical phenomena responsible for the qualitative shapes of AFM force curves, they have also revealed a wealth of atomic-scale phenomena that occur during nanoindentation that was not previously known.

7.3.1 Surfaces

The nature of adhesive interactions between clean, deformable metal tips indenting metal surfaces have been identified and clarified over the course of the last decade through the use of MD simulations (Totten and Liang 2004; Sutton et al. 1992; Raffi-Tabar and Sutton 1991; Landman et al. 1990, 1992; Tomagnini et al. 1991). In particular, the high surface energies associated with clean metal surfaces can lead to strongly attractive interactions between surfaces in contact. The strength of this attraction can be so large that when the tip gets close enough to the surface to interact with it, surface atoms “jump” upwards to wet the tip in a phenomena known as jump-to-contact (JC). This wetting mechanism was first discovered in MD simulations (Landman et al. 1990) and has been confirmed experimentally (Burnham and Colton 1989; Ohmae 1996; Burnham et al. 1993; Agrait et al. 1996) using the AFM, as shown in Fig. 7.2.

The MD simulations of Landman et al. (1989a, 1990, 1993) and Landman and Luedtke 1991) using EAM potentials revealed that the JC phenomenon in metallic systems is driven by the need of the atoms at the tip-surface interface to optimize their interaction energies while maintaining their individual material cohesive binding. When the tip advances past the JC point it indents the surface, which causes the force to increase. This behavior is indicated in Fig. 7.3, points D to M. This region of the computer-generated force curve has a maximum not present in the force curve generated from experimental data (see Fig. 7.3, point L). This is due to tip-induced flow of the metal atoms in the surface that causes “piling-up” of the surface atoms around the edges of the indenter. Hysteresis on the withdrawal of the tip, shown in Fig. 7.3, points M to X, is present due to

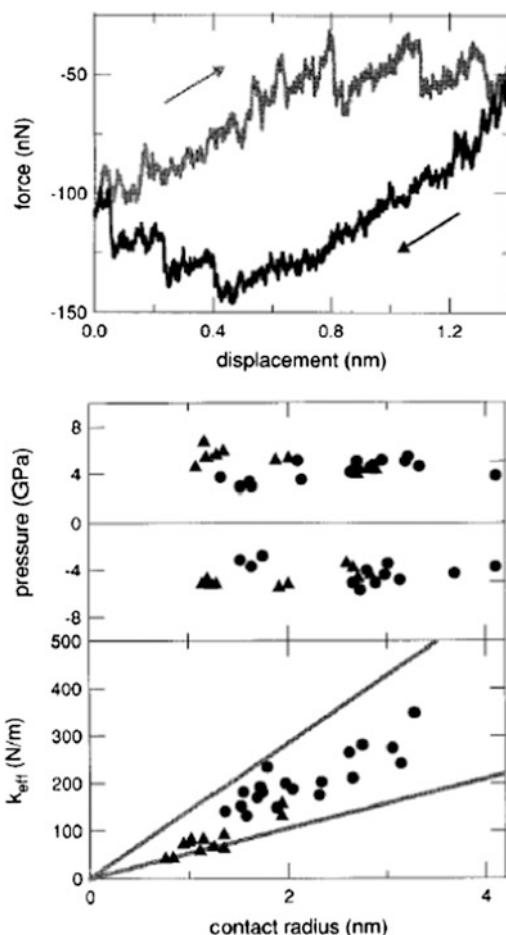


Fig. 7.2 *Top* The experimental values for the force between a tip and a surface that have a connective neck between them. The neck contracts and extends without breaking on the scales shown. *Bottom* The effective spring constant, k_{eff} , determined experimentally for the connective necks and corresponding maximum pressures, versus contact radius of the tip. The triangles indicate measurements taken at room temperature; the circles are the measurements taken at liquid He temperatures. Reprinted with permission from Agrait et al. (1996). Copyright 1996 American Chemical Society

adhesion between the tip and the substrate. In particular, as the tip retracts from the sample, a connective “neck” or nanowire of atoms forms between the tip and the substrate that is primarily composed of metal atoms from the surface with some atoms from the metal indenter that have diffused into the structure. A snapshot from the MD simulations that illustrates this behavior is shown in Fig. 7.4.

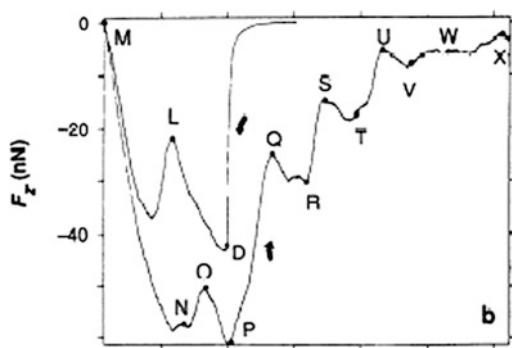


Fig. 7.3 Computationally derived force F_z versus tip-to-sample distance d_{hs} curves for approach, contact, indentation, then separation using the same tip-sample system shown in Fig. 7.4. These data were calculated from an MD simulation. Reprinted with permission from Landman et al. (1990). Copyright 1990 AAAS

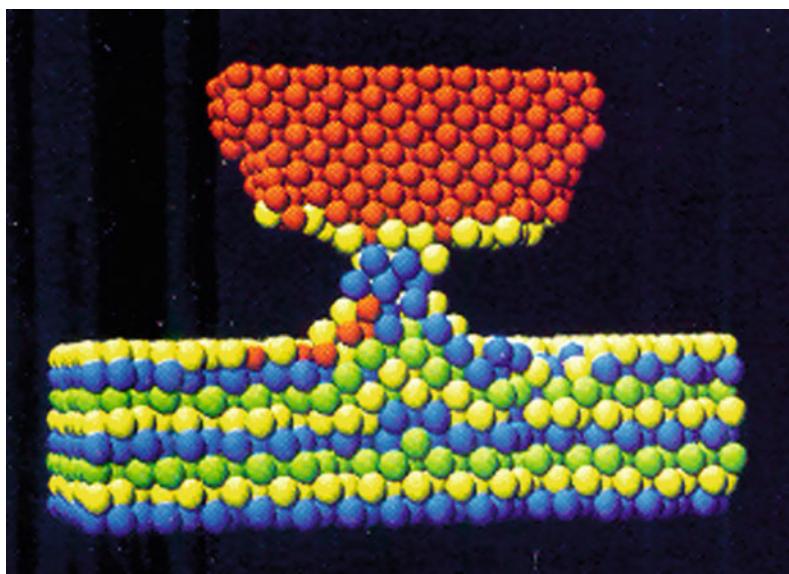
As the tip is withdrawn farther, the magnitude of the force increases (i.e., become more negative) until, at a critical force, the atoms in adjacent layers of the connective nanowire rearrange so that an additional row of atoms is created. This process causes elongation of the connective nanowire and is responsible for the fine structure (apparent as a series of maxima) present in the retraction portion of the force curve. These elongation and rearrangement steps are repeated until the connection between the tip and the surface is broken. Similar elongation events have

been observed experimentally. For example, scanning tunneling microscopy (STM) experiments demonstrate that the metal nanowires between metal tips and surfaces can elongate approximately 2500 Å without breaking (Frenken et al. 1993).

The JC process has been shown to affect the temperature at the tip-surface interface. For instance, the constant-energy MD simulations of Tomagnini et al. (1993) predicted that the energy released due to the wetting of the tip by surface atoms increases the temperature of the tip of about 15 K at room temperature and is accompanied by significant structural rearrangement. At temperatures high enough to cause the first few metal surface layers to be liquid, the distance at which the JC occurs increases, as does the contact area between the tip and the surface and the amount of nanowire elongation prior to breakage.

Simulations by Komvopoulos and Yan (1997) using LJ potentials showed how metallic surfaces respond to single and repeated indentation by metallic, or covalently bound, rigid tips. The simulations predicted that a single indentation event produces hysteresis in the force curve as a result of surface plastic deformation and heating. The repulsive force decreases abruptly during surface penetration by the tip and surface plastic deformation. Repeated indentation results in the

Fig. 7.4 Illustration of atoms in the MD simulation of a Ni tip being pulled back from an Au substrate. This causes the formation of a connective neck of atoms between the tip and the surface. Reprinted with permission from Landman et al. (1990). Copyright 1990 AAAS



continuous decrease of the elastic stiffness, surface heating, and mean contact pressures at maximum penetration depths to produce behaviors that are similar to cyclic work hardening and softening by annealing observed in metals at the macroscale.

When the tip is much stiffer than the surface, pile-up of surface atoms around the tip occurs to relieve the stresses induced by nanoindentation. In contrast, when the surface is much stiffer than the tip, the tip can be damaged or destroyed. Simulations by Belak et al. (1990) using perfectly rigid tips showed the mechanism by which the surface yields plastically after its elastic threshold is exceeded. The simulations showed how nanoindentation causes surface atoms to move on to the surface but under the tip and thus cause atomic pileup. In this study, variations in the indentation rate reveal that point defects created as a result of nanoindentation relax by moving through the surface if the rate of indentation is slow enough. If the indentation rate is too high, there is no time for the point defects to relax and move away from the indentation area and so strain builds up more rapidly. The rigid indenters considered in these MD simulations are analogous to experiments that use surface passivation to prevent JC between the tip and the surface (Yokohata and Kato 1993; Fournel et al.

1996), the results of which agree with the predicted results of pile-up and crater formation, as shown in Fig. 7.5 (Yokohata and Kato 1993).

In short, MD simulations are able to explain the atomic scale mechanisms behind measured experimental force curves produced when metal tips indent homogeneous metal surfaces to nanometer-scale depths. This preliminary work has spawned much of the current interest in using the JC to produce metal nanowires (Costakramer et al. 1995; Yanson et al. 2001a; b).

MD simulations have also been used to examine the relationship between nanoindentation and surface structure. This is most apparent in a series of computational studies that consider the indentation of a surface with a “virtual” hard-sphere indenter in a manner that is independent of the rate of indentation, as shown in Fig. 7.6. The virtual indenter is modeled through the application of a repulsive force to the surface rather than through the presence of an actual atomic tip. Kelchner et al. (1998), rather than use MD, pushed the indenter against the surface a short distance and then allowed the system to relax using standard energy minimization methods in combination with EAM potentials. The system is fully relaxed when the energy of the surface system is minimized. After relaxation, the tip is pushed further into the surface and the

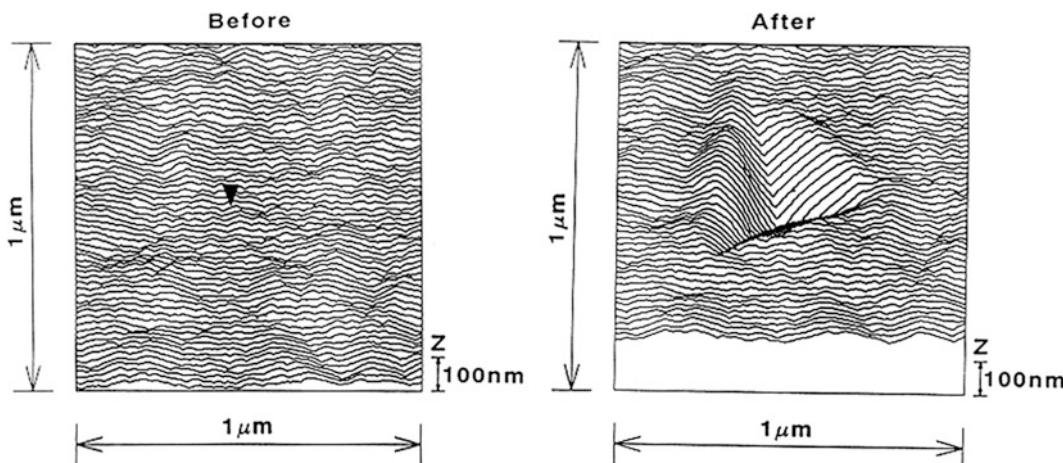
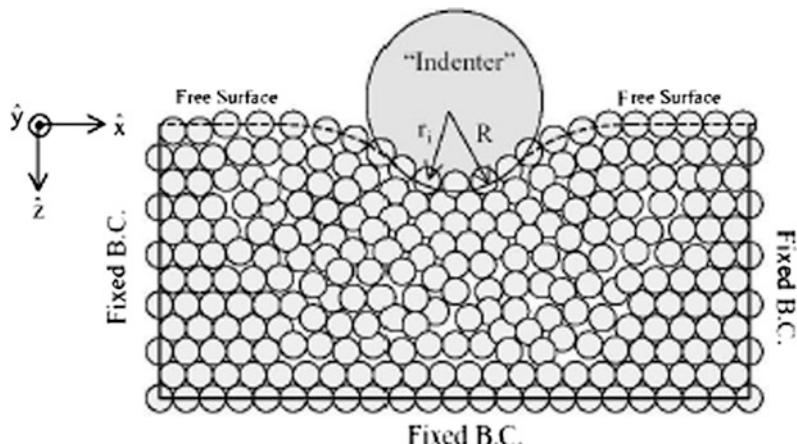


Fig. 7.5 Images of a gold surface before and after being indented with a pyramidal shaped diamond tip in air. The indentation created a surface crater. Note the pileup around the crater edges. Reprinted from Yokohata and Kato (1993). Copyright 1993 with permission from Elsevier

Fig. 7.6 A schematic of a spherical, virtual tip indenting a metal surface. Reprinted from Lilleodden et al. (2003). Copyright 2003 with permission from Elsevier



process is repeated. As the tip generates more stresses in the surface, dislocations are generated and plastic deformation occurs. If the tip is pulled back after indenting less than a specific critical value, the atoms that were plastically deformed are healed during the retraction and the surface recovers its original structure. In contrast, if the tip is indented past the critical depth, additional dislocations are created that interfere with the surface healing process on tip withdrawal. In this case, a surface crater is left on the surface following nano-indentation.

A similar study by Lilleodden et al. (2003) considered the generation of dislocations in

perfect crystals and near grain boundaries in gold. Analysis of the relationship between the load and the tip displacement in the perfect crystal shows discrete load drops followed by elastic behavior. These load drops are shown to correspond to the homogeneous nucleation of dislocations, as illustrated in Fig. 7.7, which is a snapshot taken just after the load drop. When nanoindentation occurs close to a grain boundary, similar relationships between the load and tip displacement are predicted to occur as were seen for the perfect crystal. However, the dislocations responsible for the load drop are

Fig. 7.7 Snapshot of two partial dislocations separated by a stacking fault. The dark red spheres in the center of the structure indicate atoms in perfect crystal positions after both partial dislocations have passed. Reprinted from Lilleodden et al. (2003). Copyright 2003 with permission from Elsevier

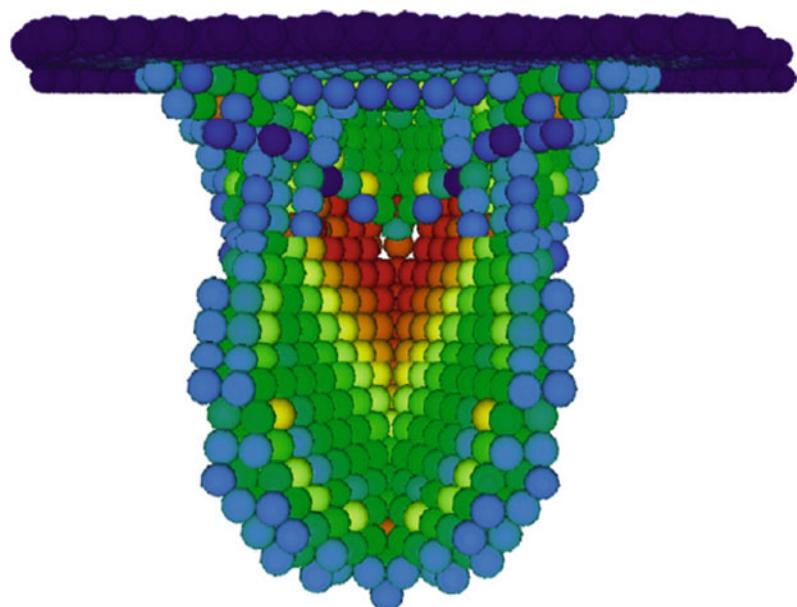
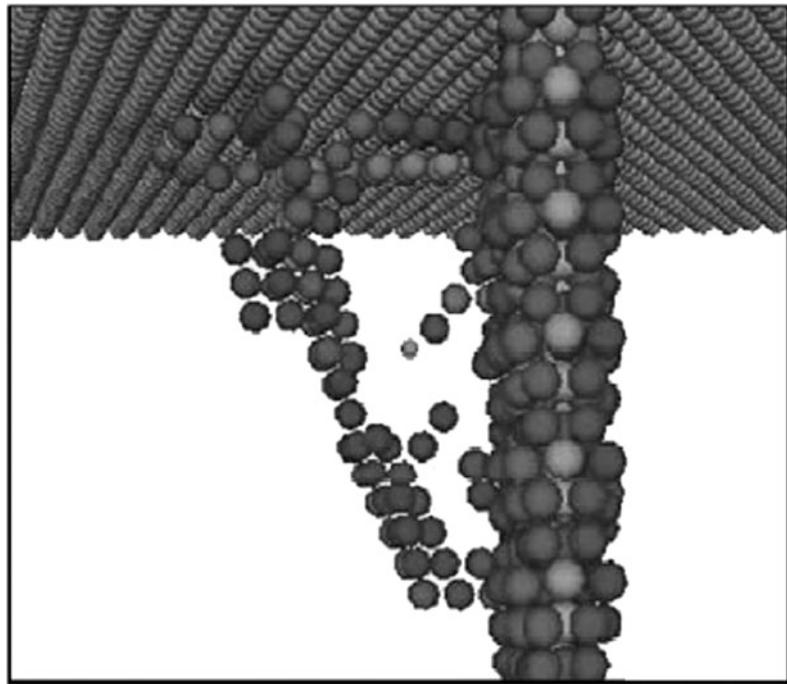


Fig. 7.8 Snapshot of the high-energy atoms only after a load drop caused by dislocation generation during the nanoindentation of gold near a grain boundary.

Reprinted from Lilleodden et al. (2003). Copyright 2003 with permission from Elsevier



preferentially emitted from the grain boundaries, as illustrated in Fig. 7.8.

Simulations can also show how atomic structure and stresses are affected by nanoindentation. For instance, MD simulations with a virtual indenter by Hasnaoui et al. (2004) using semi-empirical tight-binding methods showed the interaction between the grain boundaries under the indenter and the dislocations generated by the indentation, as illustrated in Fig. 7.9. This study shows that if the size of indenter is smaller than the grain size, the grain boundaries can emit, absorb, and reflect the dislocations in a manner that depends on atomic structure and the distribution of stresses.

Zimmerman et al. considered the indentation of a single crystal gold substrate both near and far from a surface step (de la Fuente et al. 2002). The results of these simulations, which used EAM potentials, showed that the onset of plastic deformation depends to a significant degree on the distance of indentation from the step, and whether the indentation is on the plane above or below the step. In a related set of simulations, Shenderova et al. (2000) examined whether ultra shallow elastic nanoindentation can

nondestructively probe surface stress distributions associated with surface structures such as a trench and a dislocation intersecting a surface. The simulations carried out the nanoindentation to a constant depth. They predicted maximum loads that reflect the in-plane stresses at the point of contact between the indenter and the substrate, as illustrated in Fig. 7.10.

Since the 1930s, studies have been performed using hardness measurement techniques (Kokubo 1932; Sines and Calson 1952; Oppel 1964; Simes et al. 1984) and indentation methods (Tsui et al. 1995) that suggest that the hardness of a material depends on applied in-plane uni- and bi-axial strain. In general, tensile strain appeared to decrease hardness while increases in hardness under compressive in-plane strain were reported. This behavior had traditionally been attributed to the contribution of stresses from the local strain from the indentation to the resolved shear stresses and the in-plane strain (Sines and Calson 1952; Simes et al. 1984). However, in 1996, Pharr and co-workers determined that changes in elastic modulus determined from unloading curves of strained substrates using contact areas estimated via an elastic model are

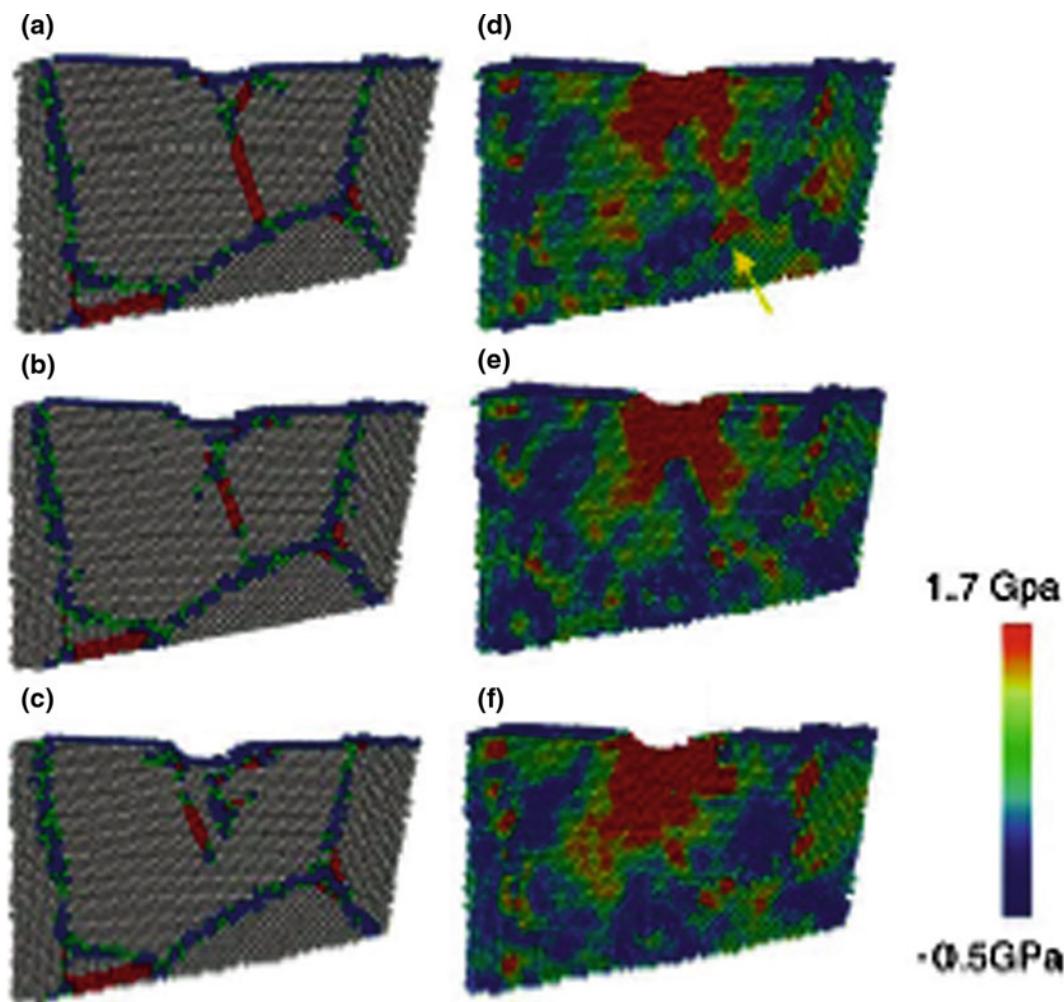
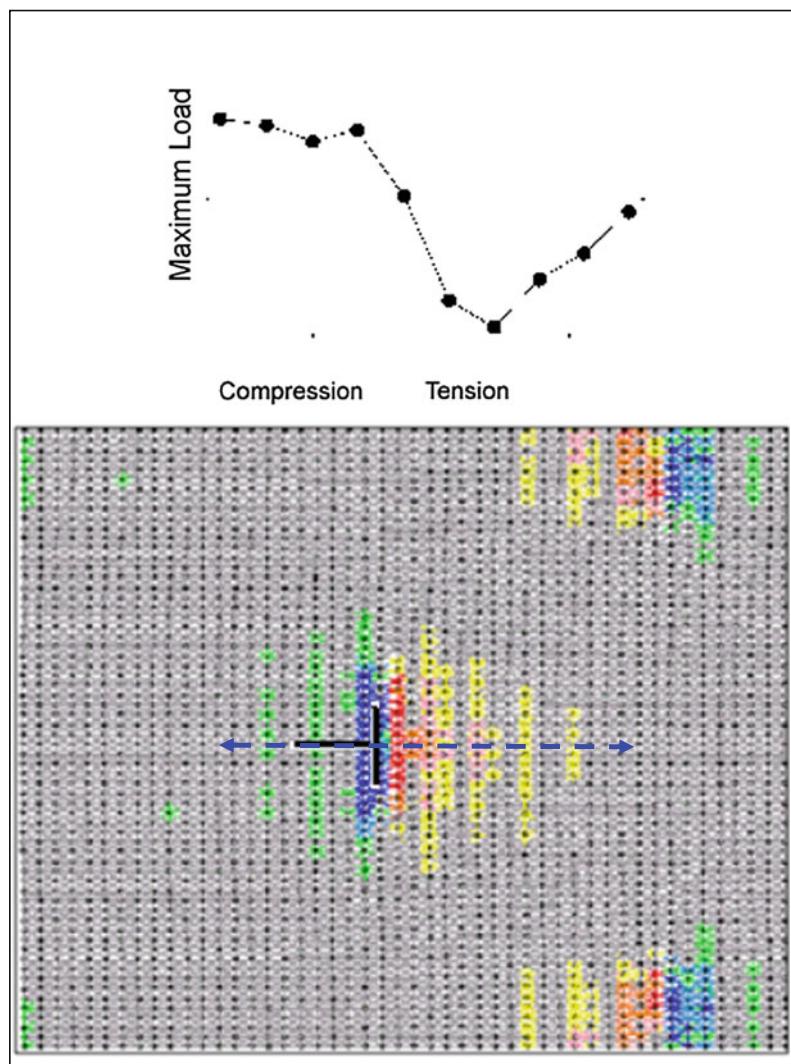


Fig. 7.9 Snapshots showing the atomic stress distribution and atomic structures in a gold surface. Figures **a–c** show the atomic structure at indentation depths of 7.9, 8.6, and 9.6 Å, respectively with a virtual spherical indenter. A dislocation is represented by the two parallel {111} planes (two red/dark lines) that show the stacking fault left behind after the leading partial dislocation has passed. Figures **d–f** show the atomic stress distribution of the same system at the same indentation depths. Here the red/dark color indicates compressive hydrostatic pressures of 1.7 GPa and higher while the blue/gray color indicates tensile pressures of -0.5 GPa and lower. The arrow in **d** shows the region of the system where a dislocation interacts with a grain boundary. Reprinted from Hasnaoui et al. (2004). Copyright 2004 with permission from Elsevier

too large to have physical significance, a result that brought into question the interpretation of prior hardness data (Tsui et al. 1995; Bolshakov et al. 1996). They hypothesized that the apparent change in modulus (and hardness) with in-plane strain is mainly due to changes in contact area that are not typically taken into consideration in elastic half-space models. This hypothesis was based on experimental nanoindentation studies of

a strained polycrystalline aluminum alloy and finite element calculations on an isotropic solid (Tsui et al. 1995; Bolshakov et al. 1996). They further suggested that in-plane compression increases pile-up around the indenter that, when not taken into account in the analysis of unloading curves, implies a nonphysical increase in modulus. Likewise, they suggested that in-plane tensile strain reduces the amount of

Fig. 7.10 Data and system illustration from a simulation of a gold surface containing a dislocation. *Top* Maximum load for simulated shallow indentation at several points along the dotted line in the bottom illustration. *Bottom* Top view of the simulated surface. The dislocation is denoted by the solid black lines



material that is pile-up around an indenter, which leads to a corresponding reduced (nonphysical) modulus when interpreting unloading curves using elastic models.

To explore in more detail the issue of pile-up and its influence on the interpretation of loading curves, Schall and Brenner used MD simulations and EAM potentials to model the plastic nanoinindentation of a single-crystal gold surface under an applied in-plane strain (Schall and Brenner 2004). These simulations predicted that the mean pressure calculated from true contact areas that take into account plastic pile-up around the indenter, varies only slightly with applied pre-stress. They

also predicted that the higher values occur in compression rather than in tension, and that the modulus calculated from the true contact area is essentially independent of the pre-stress level in the substrate. In contrast, if the contact area is estimated from approximate elastic formulae, the contact area is under estimated, which leads to a strong, incorrect dependence of apparent modulus on the pre-stress level. The simulations also showed larger pile-up in compression than in tension, in agreement with the Pharr model, and both regimes produced contact areas larger than those typically assumed in elastic analyses. These findings are illustrated in Fig. 7.11.

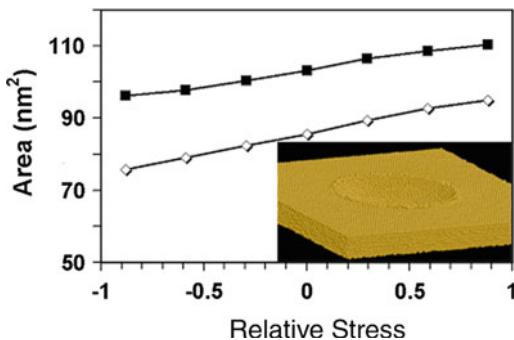


Fig. 7.11 Contact area projected in the plane at a maximum load for simulated indentation of a gold surface as a function of in-plane biaxial stress. The stress is normalized to the theoretical yield stress. The *top curve* is from an atomistic simulation; the *bottom curve* is from an elastic model. *Inset* Illustration of the region near the indentation from the simulation. The tip is not shown for clarity. Initial formation of pile-up around the edge of the indentation is apparent

Nanometer-scale indentation of ceramic systems has also been investigated with MD simulations. Ceramics are stiffer and more brittle than metals at the macroscale and examining the nanoindentation of ceramic surfaces provides information about the nanometer-scale properties. They also reveal the manner by which defects form in covalent and ionic materials. For example, Landman et al. (1989b, 1992) considered the interaction of a CaF_2 tip with a CaF_2 substrate in MD simulations using empirical potentials. As the tip approaches the surface, the attractive force between them steadily increases. This attractive force increases dramatically at the critical distance of 2.3 \AA as the interlayer spacing of the tip increases (i.e., the tip is elongated) in a process that is similar to the JC phenomenon observed in metals. An important difference, however, is the amount of elongation which is 0.35 \AA in the case of the ionic ceramics and several angstroms in the case of metals. As the distance between the tip and the surface decreases further, the attractive nature of their interaction increases until a maximum value is reached. Indentation beyond this point results in a repulsive tip/substrate interaction, compression of the tip, and ionic bonding between the tip and substrate. These bonds are responsible for the

hysteresis predicted to occur in the force curve on retraction, which ultimately leads to plastic deformation of the tip followed by fracture.

The responses of covalently bound ceramics such as diamond and silicon to nanoindentation have been heavily studied with MD simulations. One of the first of these computational studies was carried out by Kallman et al. who used the Stillinger-Weber potential to examine the indentation of amorphous and crystalline silicon (Kallman et al. 1993). The motivation for this study came from experimental data that indicated a large change in electrical resistivity during indentation of silicon, which led to the suggestion of a load-induced phase transition below the indenter. Clarke et al., for example, reported forming an Ohmic contact under load, and using transmission electron microscopy they observed an amorphous phase at the point of contact after indentation (Clarke et al. 1988). Using micro-Raman microscopy, Kailer et al. identified a metallic $\beta\text{-Sn}$ phase in silicon near the interface of a diamond indenter during hardness loading (Kailer et al. 1999). Furthermore, upon rapid unloading they detected amorphous silicon as in the Clarke et al. experiments, while slow unloading resulted in a mixture of high-pressure polymorphs near the indent point. At the highest indentation rate and the lowest temperature, the simulations by Kallman et al. showed that amorphous and crystalline silicon have similar yield strengths of 138 and 179 kbar, respectively. In contrast, at temperatures near the melting temperature and at the slowest indentation rate, both amorphous and crystalline silicon are predicted to have lower yield strengths of 30 kbar. The simulations thus show how the predicted yield strength of silicon at the nanometer scale depends on structure, rate of deformation, and surface temperature.

Interestingly, Kallman et al. (1993) found that amorphous silicon does not crystallize upon indentation, but indentation of crystalline silicon at temperatures near the melting point transforms the surface structure near the indenter to the amorphous phase. The simulations do not predict transformation to the $\beta\text{-Sn}$ structure under any of the conditions considered. These results agree

with the outcomes of scratching experiments (Minowa and Sumino 1992) that showed that amorphous silicon emerges from room-temperature scratching of crystalline silicon.

Kaxiras and co-workers revisited the silicon nanoindentation issue using a quasi-continuum model that couples interatomic forces from the Stillinger-Weber potential to a finite element grid (Smith et al. 2000). They report good agreement between simulated loading curves and experiment provided that the curves are scaled by the indenter size. Rather than the β -Sn structure, however, atomic displacements suggest formation of a metallic structure with five-fold coordination below the indenter upon loading, and a residual simple cubic phase near the indentation site after the load is released rather than the mix of high-pressure phases characterized experimentally. Smith et al. attribute this discrepancy to shortcomings of the Stillinger-Weber potential in adequately describing the high-pressure phases of silicon. They also used a simple model for changes in electrical resistivity with loading involving contributions from both a Schottky barrier and spreading resistance. Simulated resistance-versus-loading curves agree well with experiment despite possible discrepancies between the high-pressure phases under the indenter, suggesting that the salient features of the experiment are not dependent on the details of the high-pressure phases produced.

Additional MD simulations of the indentation of silicon were carried out by Cheong and Zhang (2000). Their simulations provide further details about the phase transformations that occur in silicon as a result of nanoindentation. In particular, they find that the diamond cubic silicon is transformed into a body-centered tetragonal structure (β -silicon) upon loading of the indenter as illustrated in Fig. 7.12. Figure 7.13 shows that the coordination numbers of silicon atoms also coincide with that of the theoretical β -silicon structure. The body centered tetragonal structure is transformed into amorphous silicon during the unloading stage. A second indentation simulation again predicted that this is a reversible process. Atomistic simulations by Sanz-Navarro et al. (2004) shows the relation between the

indentation of silicon and the hydrostatic pressure on surface cells due to the nanoindentation as illustrated in Fig. 7.14. These simulations further predict that the transformation of diamond silicon into the β -silicon structure can occur if the hydrostatic pressure is somewhat over 12 GPa.

Multimillion atom simulations of the indentation of silicon nitride was recently carried out by Walsh et al. (2003). The elastic modulus and hardness of the surface was calculated using load-displacement relationships. Snapshots from the simulations, illustrated in Fig. 7.15, show that pile-up occurs on the surface along the edges of the tip. Plastic deformation of the surface is predicted to extend a significant distance beyond the actual contact area of the indenter, as illustrated in Fig. 7.15.

The indentation of bare and hydrogen-terminated diamond (111) surfaces beyond the elastic limit was investigated by Harrison et al. (1992a) using a hydrogen-terminated sp^3 -bonded tip in MD simulations that utilized bond-order potentials. The simulations identified the depth and applied force at which the diamond (111) substrate incurred plastic deformation due to indentation. At low indentation forces, the tip-surface interaction is purely elastic, as illustrated in Fig. 7.16. This finding agrees with the findings of Cho and Joannopoulos (1995), who examined the atomic-scale mechanical hysteresis experienced by an AFM tip indenting Si(100) with density functional theory. The calculations predicted that at low rates it is possible to cycle repeatedly between two buckled configurations of the surface without adhesion.

When the nanoindentation process of diamond (111) is plastic, connective strings of atoms are formed between the tip and the surface, as illustrated in Fig. 7.17. These strings break as the distance between the tip and crystal increases and each break is accompanied by a sudden drop in the potential energy at large positive values of tip/substrate separation. The simulations further predict that the tip end twists to minimize interatomic repulsive interactions between the hydrogen atoms on the surface and the tip. This behavior is predicted to lead to new covalent

Fig. 7.12 Snapshots of a silicon sample during indentation. The smaller dots are diamond atoms.

a Crystalline silicon prior to indentation. **b** Atoms beneath the indenter are displaced as a result of indentation. **c** The system at maximum indentation. Some of the atoms are in a crystalline arrangement (*circled region*) that is different from the diamond structure. **d** The surface structure is largely amorphous as the tip is withdrawn. **e** The surface after indentation. Note the amorphous region at the site of the indentation process.

Reprinted with permission from Cheong and Zhang (2000). Copyright 2000 Institute of Physics and IOP Publishing Ltd

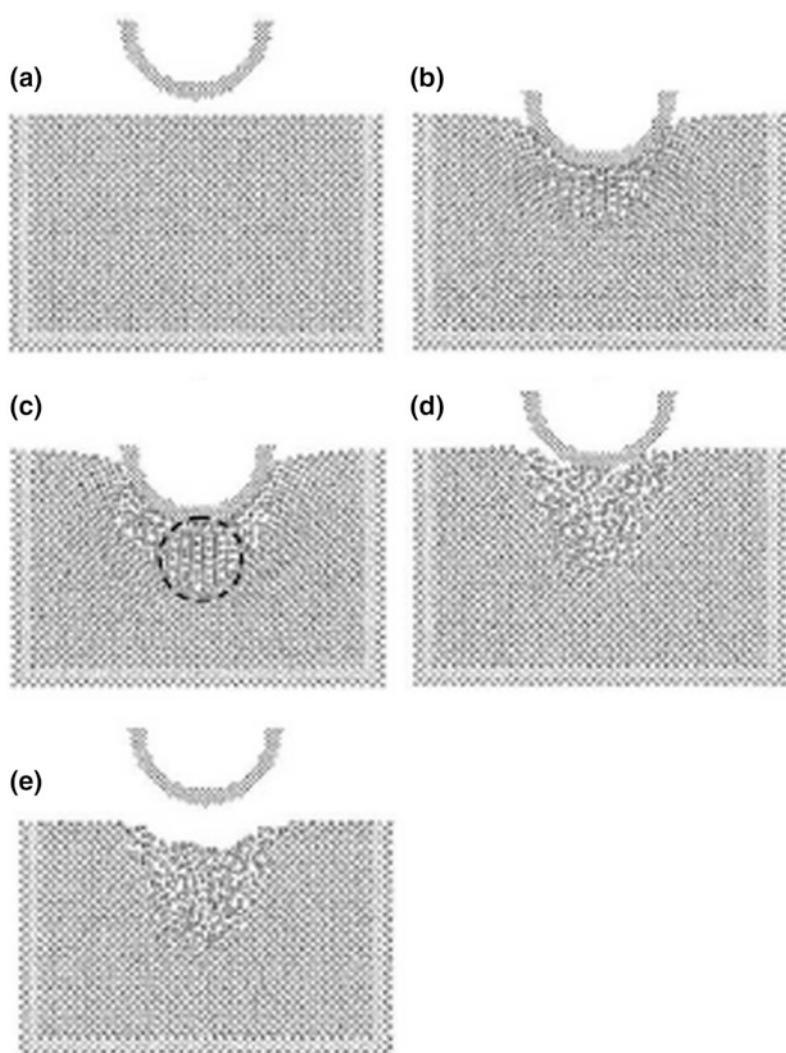
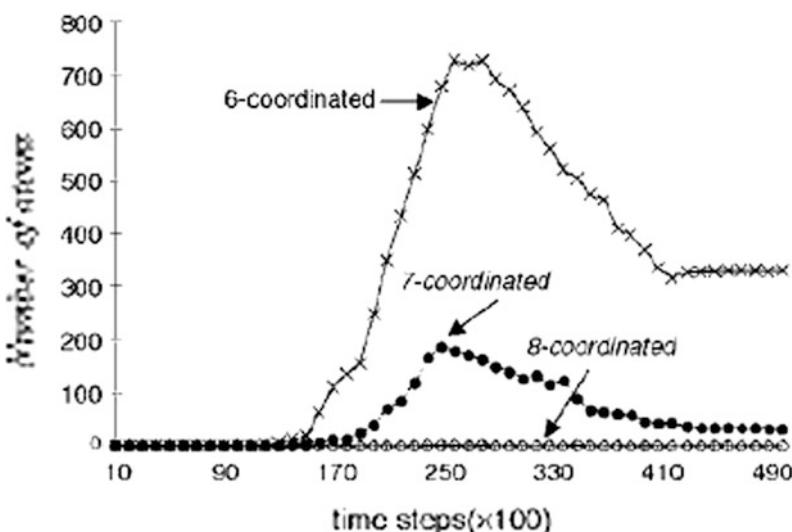


Fig. 7.13 The coordination of the silicon atoms shown in Fig. 7.12 as a function of time during nanoindentation.

Reprinted with permission from Cheong and Zhang (2000). Copyright 2000 Institute of Physics and IOP Publishing Ltd



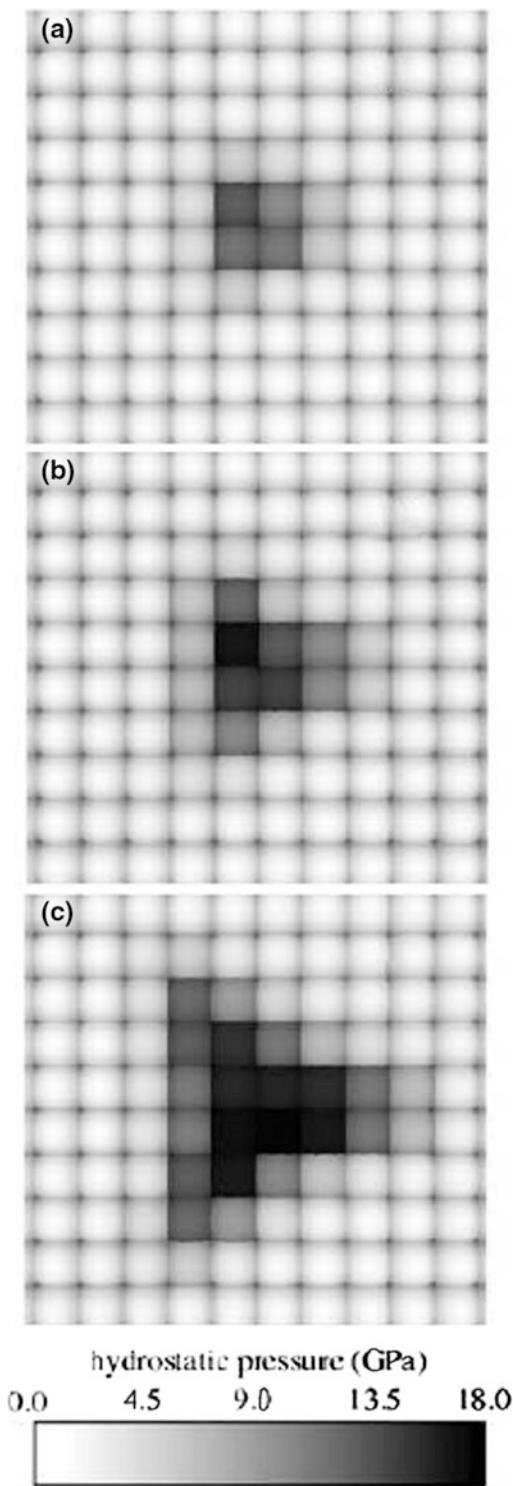


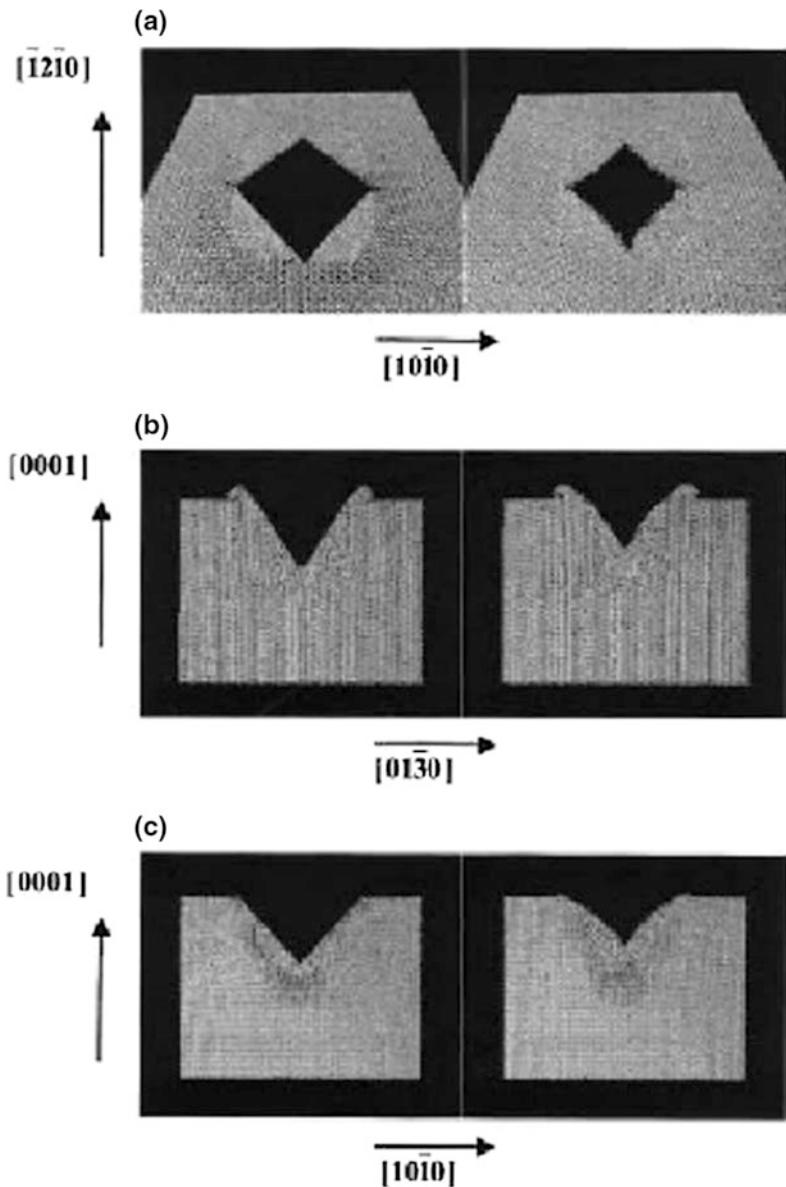
Fig. 7.14 Calculated hydrostatic pressure of surface cells at indentation depths of **a** 8.9 Å, **b** 15.7 Å, and **c** 25.3 Å. Reprinted with permission from Sanz-Navarro et al. (2004). Copyright 2000 Institute of Physics and IOP Publishing Ltd

bond formation between the tip and the carbon atoms below the first layer of the surface and connective strings of atoms between the tip and the surface when the tip is retracted. Not surprisingly, when the surface is bare and not terminated with hydrogen atoms, the repulsive interactions between the tip and the surface are minimized and the tip indents the substrate without twisting (Harrison et al. 1992a). Because carbon–carbon bonds are formed between the tip and the first layer of the substrate, the indentation is ordered (i.e., the surface is not disrupted as much by interacting with the tip) and the eventual fracture of the tip during retraction results in minimal damage to the substrate. The concerted fracture of all bonds in the tip give rise to a single maximum in the potential versus distance curve at large distances.

Harrison et al. (1997, 2001), Garg et al. (1998) and Garg and Sinnott (1999) considered the indentation of diamond and graphene surfaces with AFM tips of carbon nanotubes and nanotube bundles using MD simulations and bond-order potentials. Tips consisting of both single-walled nanotubes and multi-walled nanotubes were considered. The simulations predicted that nanotubes do not plastically deform during tip crashes on these surfaces. Rather, they elastically deform, buckle, and slip as shown in Fig. 7.18. However, as is the case for diamond tips indenting reactive diamond surfaces discussed above, in the case of highly reactive surfaces strong adhesion can occur between the nanotube and the surface that destroys the nanotube, as illustrated in Fig. 7.19.

To summarize, MD simulations reveal the properties of ceramic tips and surfaces with covalent or ceramic bonding that are most important for nanometer-scale indentation. They predict that brittle fracture of the tip can occur that is sometimes accompanied by strong adhesion with the surface. They also reveal the conditions under which neither the tip nor the surface is affected by the nanoindentation process. The insight gained from these simulations helps in the interpretation of experimental data as well as revealing the nanometer-scale mechanisms by which, for example, tip buckling and permanent modification of the surface occur.

Fig. 7.15 Snapshots of the silicon nitride **a** surface, **b** slide parallel to the edges of the indenter, and **c** slide across the indenter diagonal. The *left-hand side* shows the surface when it is fully loaded, while the *right-hand side* shows the surface after the tip has been withdrawn. Reprinted with permission from Walsh et al. (2003). Copyright 2003 American Institute of Physics



7.3.2 Thin Films

In many instances, surfaces are covered with thin films that can range in thickness from a few atomic layers to several μm . These films are more likely to have properties that differ from the properties of bulk materials of similar composition, and the likelihood of this increases as the film thickness decreases. Nanoindentation is one of the best approaches for determining the

properties of these films. Consequently, numerous computational simulations of this process have been carried out.

For example, MD simulations have been used to study the indentation of metal surfaces covered with liquid n-hexadecane films, as illustrated in Fig. 7.20. As the metal tip touches the film, some of the molecules from the surface transfer to the tip and this causes the film to “swell”. As the tip continues to push against the

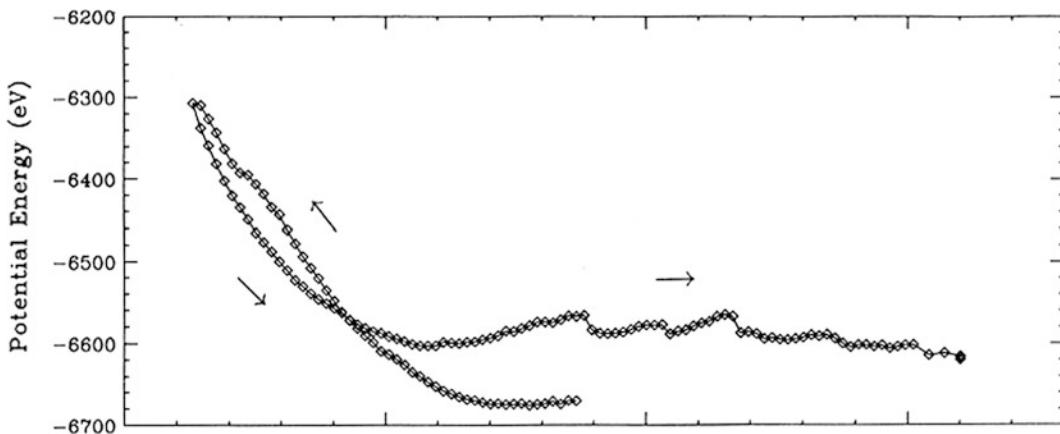


Fig. 7.16 Potential energy as a function of rigid-layer separation generated from an MD simulation of an elastic (nonadhesive) indentation of a hydrogen-terminated diamond (111) surface using a hydrogen-terminated, sp^3 -hybridized tip. Reprinted from Harrison et al. (1992a). Copyright 1992 with permission from Elsevier

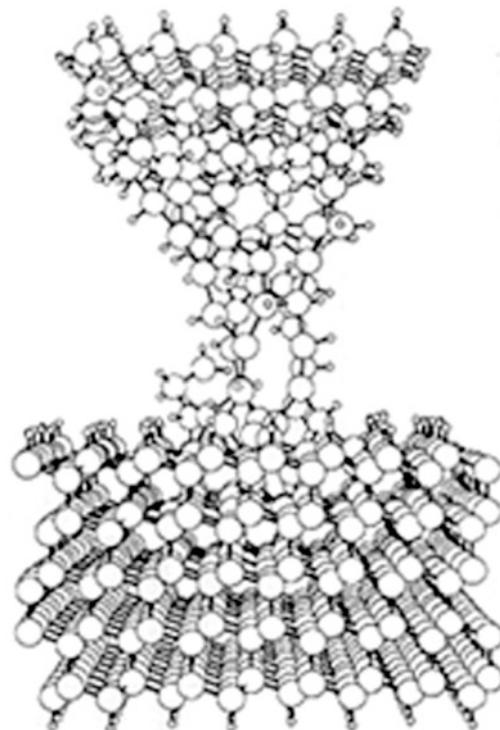


Fig. 7.17 Illustration of atoms in the MD simulation of the indentation of a hydrogen-terminated diamond (111) substrate with a hydrogen-terminated, sp^3 -hybridized tip at selected time intervals. The tip-substrate system as the tip was withdrawn from the sample. Large and small spheres represent carbon and hydrogen atoms, respectively. Reprinted from Harrison et al. (1992a). Copyright 1992 with permission from Elsevier

surface the hydrocarbon film wets the side of the tip. The simulations show how the hydrocarbon film passivates the surface and prevents the strong attractive interactions discussed above for clean metal surfaces and tips from occurring.

In a series of MD simulations, Tupper and Brenner modeled the compression of a thiol self-assembled monolayer (SAM) on a rigid gold surface using both a smooth compressing surface (Tupper and Brenner 1994) and a compressing surface with an asperity (Tupper et al. 1994). These simulations showed that compression with the smooth surface produced a compression-induced structural change that led to a change in slope of the simulated force versus compression curve. This transition is reversible and involves a change in the ordered arrangement of the sulfur head-groups on the gold surface. A similar change in slope seemed to be present in the experimental indentation curves of Houston and co-workers (Joyce et al. 1992) but was not discussed by the authors. The simulations with the asperity showed that the asperity is able to penetrate the tail groups of the self-assembled monolayer, as illustrated in Fig. 7.21, before an appreciable load is apparent on the compressing surface. This result indicates that it is possible to image the head groups of a thiol self-assembled monolayer that are adsorbed onto the surface of a

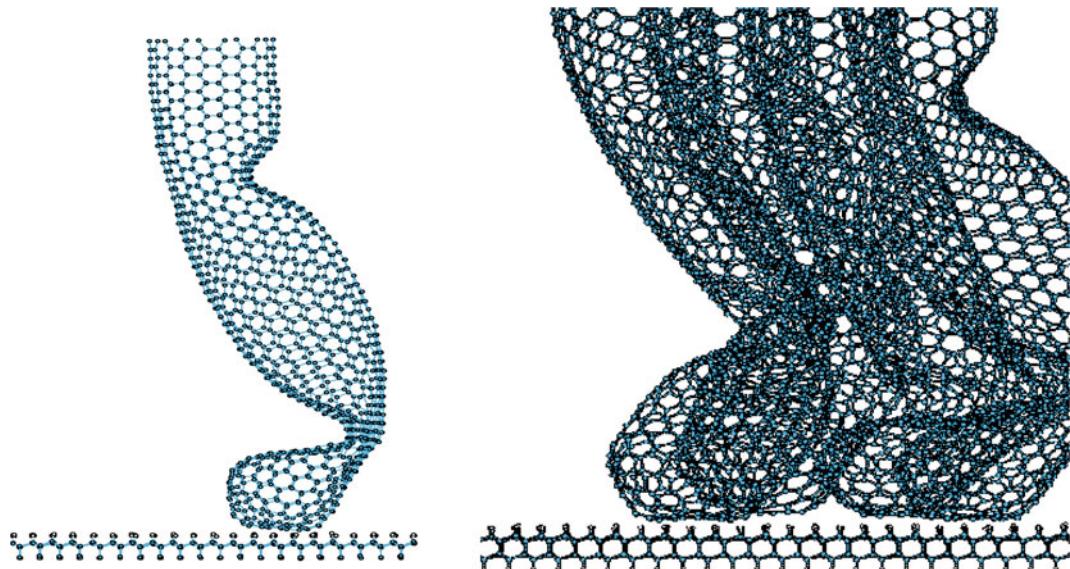


Fig. 7.18 Snapshots of the indentation of a single-walled nanotube (*left-hand image*) and a bundle of nanotubes (*right-hand image*) on hydrogen-terminated diamond (111)

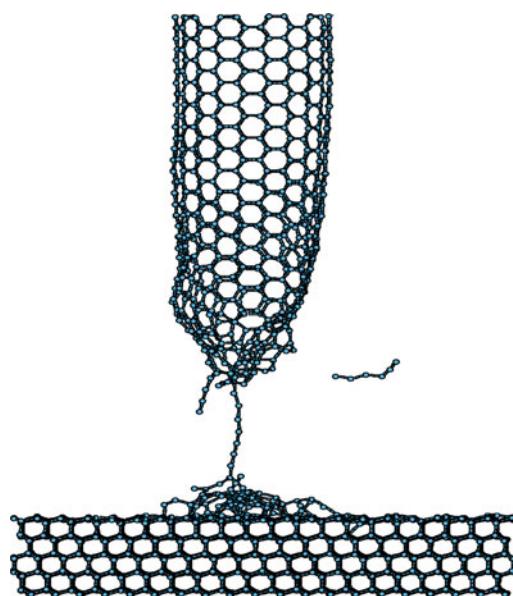


Fig. 7.19 Snapshot of a single-walled carbon nanotube as it is withdrawn following indentation on a bare, diamond (111) surface. Reprinted with permission from Garg et al. (1998). Copyright 1998, 1999 American Physical Society

gold substrate using STM, and consequently ordered images of these systems may not be indicative of the arrangement of the tails groups.

Zhang et al. (2003a) used a hybrid MD simulation approach, where a dynamic element model for the AFM cantilever was merged with a MD relaxation approach for the rest of the system, to study the frictional properties of alkanethiol SAMs on gold. They investigated the effect of several variables like chain length, terminal group, scan direction, and scan velocity. Their results show that friction forces decrease as the chain length of the SAMs increase. In the case of shorter chains such as C_7CH_3 , the SAMs near the tip can be deformed by indentation as illustrated in Fig. 7.22. This behavior is predicted to be the cause of higher friction that occurs for the short length chains.

Harrison and coworkers have used classical MD simulations (Harrison et al. 2001; Tutein et al. 1999) to examine the indentation of monolayers composed of linear hydrocarbon chains that are chemically bound (or anchored) to a diamond substrate. Both flexible and rigid single-wall, capped nanotubes were used as tips. The simulations showed that indentation causes the ordering of the monolayer to be disrupted regardless of the type of tip used. Indentation results in the formation of gauche defects within the monolayer and, for deep indents, results in

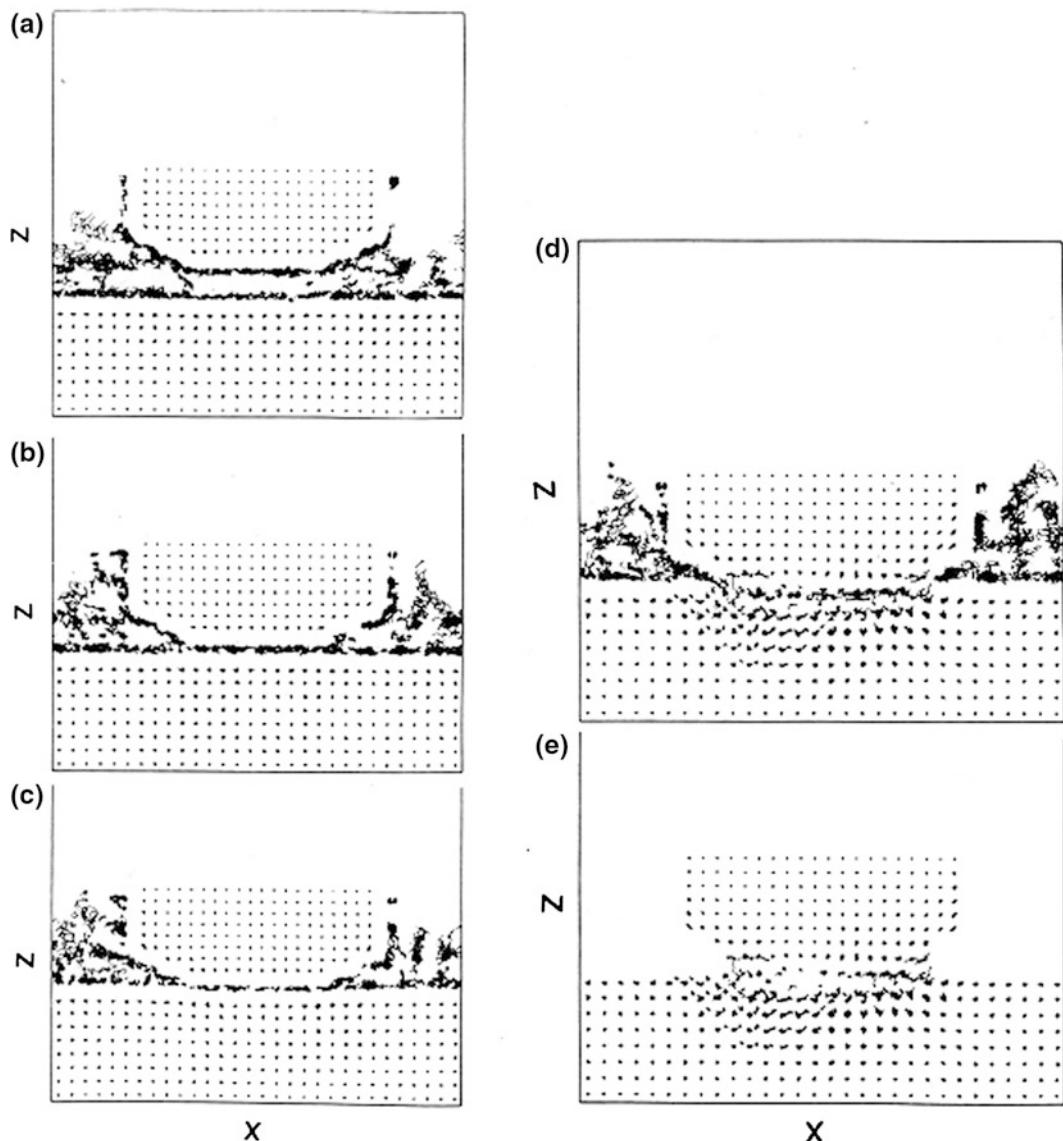
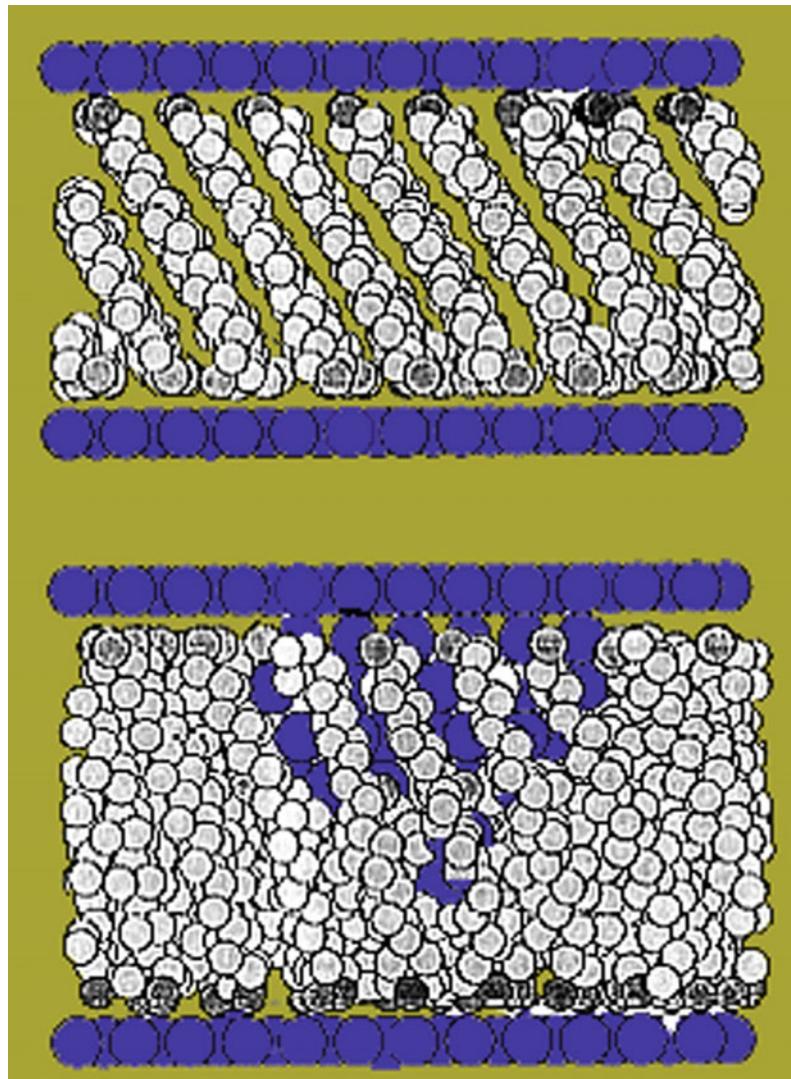


Fig. 7.20 Cutaways of the side view from molecular dynamics simulations of a Ni tip indenting a Au (001) surface covered with a hexadecane film. In e only the metal atoms are shown. Note how the hexadecane is forced out from between the metal surfaces. Reprinted with permission from Landman et al. (1992). Copyright 1995 with permission from Elsevier

the pinning of selected hydrocarbon chains beneath the tube. Flexible nanotubes tilt slightly as they begin to indent the softer monolayers. This small distortion is due to the fact that nanotubes are stiff along their axial direction and more flexible in the transverse direction. In contrast, when the nanotubes encounter the hard diamond substrate, after “pushing” through the

monolayer, they buckle. This process is illustrated in Fig. 7.23 and the force curves are shown in Fig. 7.24. The buckling of the nanotube was previously observed when single-walled, capped nanotubes were brought into contact with hydrogen-terminated diamond (111) surfaces (Harrison et al. 1997; Garg et al. 1998). In the absence of the monolayer, the nanotube tips

Fig. 7.21 Snapshots illustrating the compression of a self-assembled thiol film on gold for a smooth surface (*top*) and a surface containing an asperity (*bottom*). The asperity can penetrate and disorder the film tail groups before appreciable load occurs



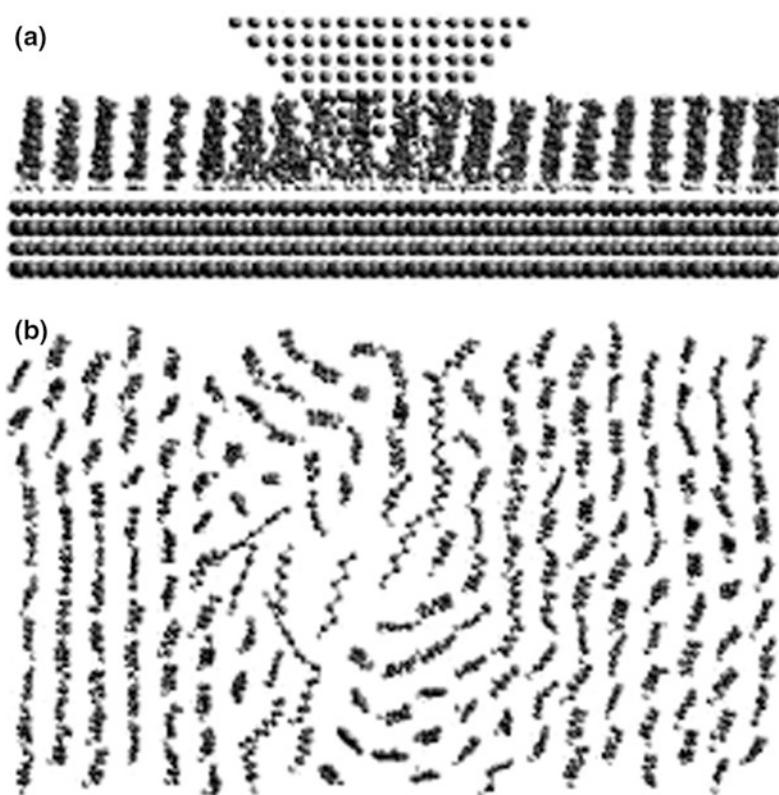
encounter the hard substrate in a nearly vertical position. This interaction with the diamond substrate causes the cap of the nanotubes to be “pushed” inside the nanotube, i.e., invert. Increasing the load on the nanotubes causes the walls of the tube to buckle. Both the cap inversion and the buckling are reversible processes. That is, when the load on the tube is removed, it recovers its original shape.

Deep indents of the hydrocarbon monolayers using rigid nanotubes result in rupture of chemical bonds. The simulations also show that the number of gauche defects generated by the indentation is a linear function of penetration

depth and equal for C_{13} and C_{22} monolayers. Thus, it is the tip that governs the number of gauche defects generated.

Leng and Jiang (2002) investigated the effect of using tips coated with SAMs containing hydrophobic methyl (CH_3) or hydrophilic hydroxyl (OH) terminal groups to nanoindent gold surfaces that also are covered with SAMs with the identical terminal groups as the tip. Figure 7.25 contains snapshots for the indentation process predicted to occur for terminal OH/OH interactions during compression and the pull-off. The adhesion force of OH/OH pairs is calculated to be about four times larger

Fig. 7.22 **a** Side and **b** top views of the final configuration of a C_7CH_3 self-assembled monolayer on Au(111) under a high normal load of 1.2 nN at 300 K. The tip is not shown in (b) for clarity. Reprinted with permission from Zhang et al. (2003a). Copyright 2003 American Chemical Society



than that of CH_3/CH_3 pairs, as shown in Fig. 7.26. This is due to the formation of hydrogen bonding between OH/OH pairs. This interaction is also expected to increase the frictional force between monolayers with OH terminations.

Related MD simulations by Mate predict that the end groups on polymer lubricants have a significant influence on the lubrication properties of polymers (Mate 1992). For instance, fluorinated end groups are predicted to be less reactive than regular alcohol end groups. When fluorinated films are indented, the normal force becomes more attractive as the distance between the tip and film decreases until the hard wall limit is reached and the interactions become repulsive. In contrast, when AFM tips indent hydrogenated films, the forces become increasingly repulsive as the distance between them decreases, as shown in Figs. 7.27 and 7.28. This predicted behavior is due to the compression of the end group beneath the tip. For the lubricant

molecules to be squeezed out from between the tip and the surface, the hydrogen-bonding between the two must first be broken and this increases the force needed to indent the system. As a result, a major effect of the presence of alcohol end groups is to dramatically increase the load that a liquid lubricant can support before failure (solid–solid contact) occurs.

When atomically-sharp tips are used to indent solid-state thin films where there is a large mismatch in the mechanical properties of the film and the substrate, it is difficult to determine the true contact area between the tip and the surface during nanoindentation. In the case of soft films on hard substrates, pileup can occur around the tip that effectively increases the contact area. In contrast, with hard films on soft substrates, “sink-in” is experienced around the tip that decreases the true contact area.

A class of coatings that has received much attention is diamond-like, amorphous-carbon (DLC) coatings. DLC coatings are almost has

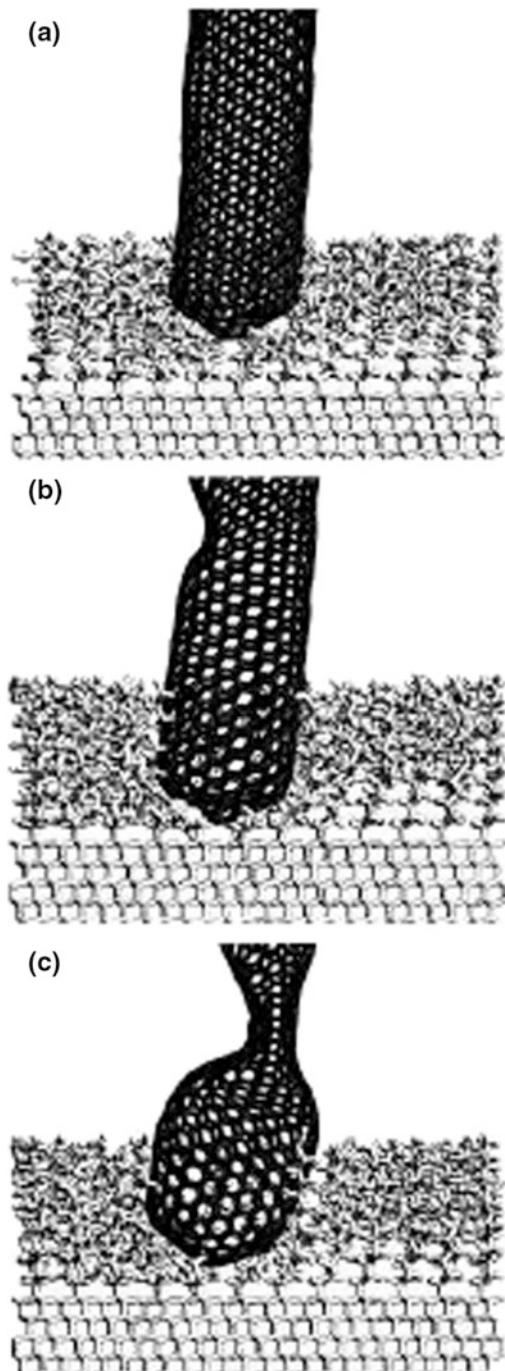


Fig. 7.23 Snapshots from the simulation of the interaction of a flexible single-walled carbon nanotube with a monolayer of C_{13} chains on diamond. The loads are **a** 19.8 nN, **b** 41.2 nN, and **c** 36.0 nN. Reprinted with permission from Tutein et al. (1999). Copyright 1999 American Chemical Society

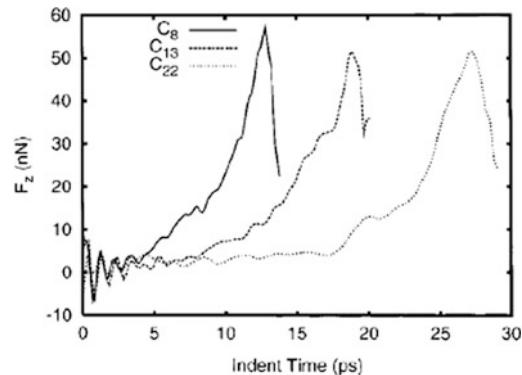


Fig. 7.24 The load on the upper two layers of the flexible carbon nanotube indenter shown in Fig. 7.23 as a function of indentation time for the nanoindentation of the indicated hydrocarbon monolayers on diamond. Reprinted with permission from Tutein et al. (1999). Copyright 1999 American Chemical Society

hard as crystalline diamond and may have very low friction coefficients (<0.01) depending upon the growth conditions (Enke et al. 1980; Enke 1981; Miyake et al. 1987; Erdemir and Donnet 2000). They have therefore generated much interest in the tribological community and there have been several MD simulation studies to determine the mechanical and atomic-scale frictional properties of DLC coatings. MD simulations with bond-order potentials by Sinnott et al. (1997) examined the differences in indentation behavior of a hydrogen-terminated diamond tip on hydrogen-terminated single-crystal diamond surfaces and diamond surfaces covered with DLC. In the former case, the tip goes through shear and twist deformations at low loads that change to plastic deformation and adhesion with the surface at high loads. When the surface is covered with the DLC film, the tip easily penetrates the film, as illustrated in Fig. 7.29, which “heals” easily when the tip is retracted so that no crater or other evidence of the indentation is left behind.

MD simulations by Glosli et al. (1995) of the indentation of DLC films that are about 20 nm thick give similar results. In this case a larger, rigid diamond tip was used in the indentations and was also slid across the surface. During sliding the tip plows the surface which causes

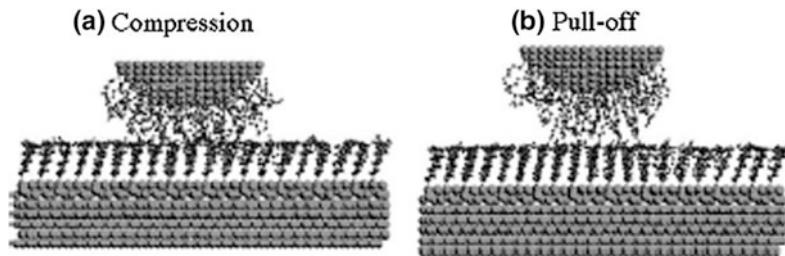


Fig. 7.25 Snapshots from the OH/OH pair interaction during **a** compression and **b** pulloff. Reprinted with permission from Leng and Jiang (2002). Copyright 2002 American Chemical Society

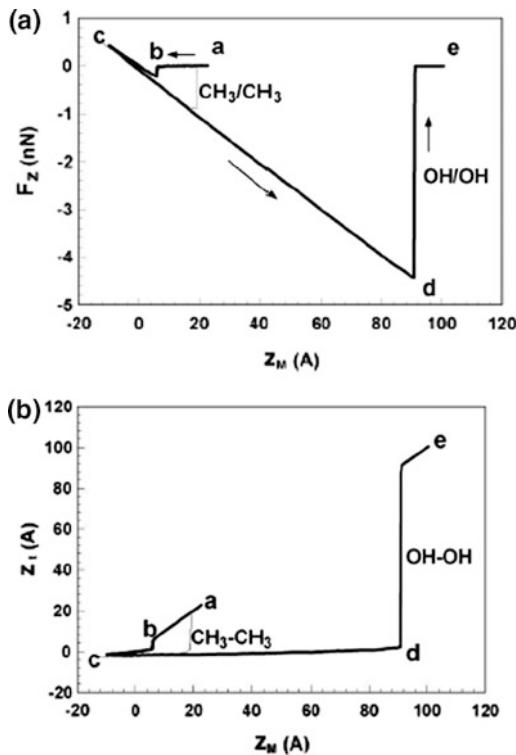


Fig. 7.26 **a** Force distance curves and **b** tip position (z_t) versus support position (z_M) for the OH/OH contact pair and the CH_3/CH_3 contact pair. Reprinted with permission from Leng and Jiang (2002). Copyright 2002 American Chemical Society

some changes to the film not seen during indentation. However, because the tip is perfectly rigid, adhesion between the film and surface is not allowed which influences the results.

This section shows that repulsive interactions between surfaces covered with molecular films and proximal probe tips are minimized relative to interactions between bare surfaces and

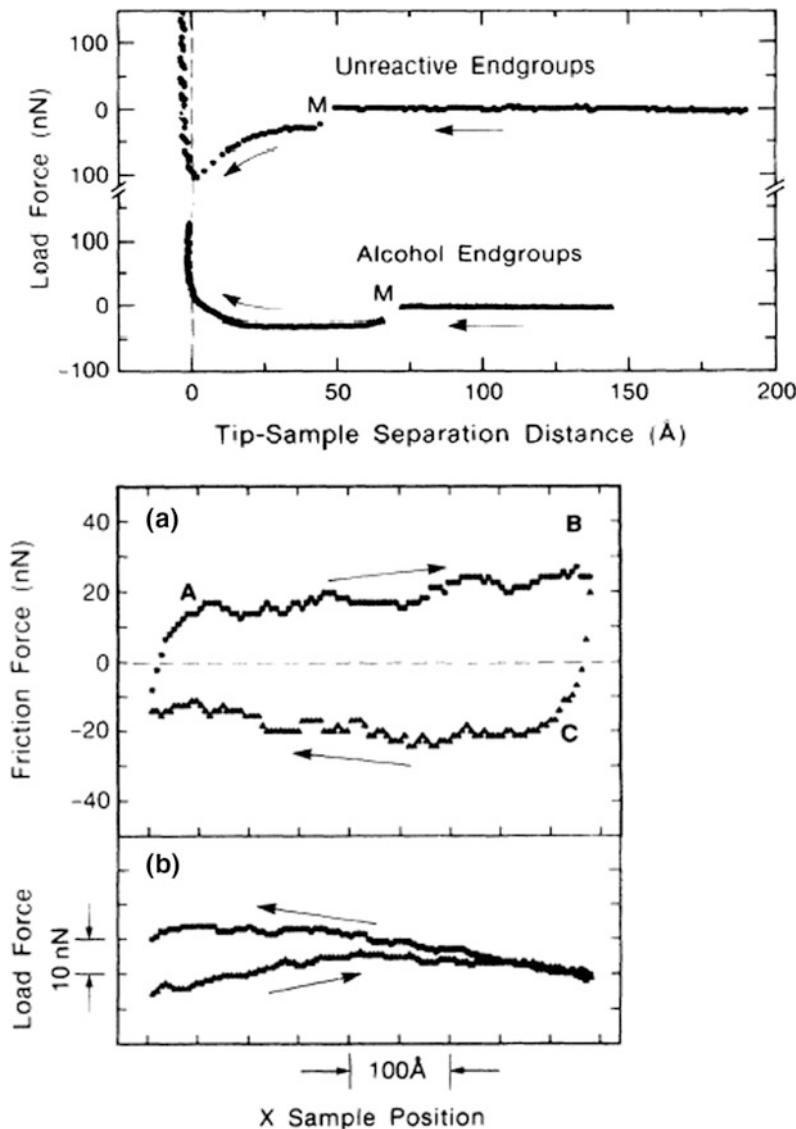
indentation tips. The lubrication properties of polymers and SAMs can vary with chain length, the rigidity of the tip, and the chemical properties of the end groups. In some cases, indentations can disrupt the initial ordering of polymers and SAMs, which affects their responses to nanoindentation and friction.

7.4 Friction and Lubrication

Work is required to slide two surfaces against one another. When the work of sliding is converted to a less ordered form, as required by the first law of thermodynamics, friction will occur. For instance, if the two surfaces are strongly adhering to one another, the work of sliding can be converted to damage that extends beyond the surfaces and into the bulk. If the adhesive force between the two surfaces is weaker, the conversion of work results in damage that is limited to the area at or near the surface and produces transfer films or wear debris (Singer 1991; Singer et al. 1991). While the thermodynamic principles of the conversion of work to heat are well-known, the mechanisms by which this takes place at sliding surfaces are much less well established despite their obvious importance for a wide variety of technological applications.

Atomic-scale simulations of friction are therefore important tools for achieving this understanding. They have consequently been applied to numerous materials in a wide variety of structures and configurations, including atomically flat and atomically rough diamond surfaces (Harrison et al. 1992b, 1993a, b), rigid

Fig. 7.27 *Top* The force versus distance curve (indentation part only) for unbonded perfluoropolyether on Si(100). The unreactive end groups were from a 10 Å thick film; the reactive alcohol end groups were from a 30 Å thick film. The negative forces represent attractive interactions between the tip and the surface. *Bottom* Measured plots of friction and load forces of the tip as it slides over the sample with the alcohol end groups. Reprinted with permission from Mate (1992). Copyright 1992 American Physical Society



substrates covered with monolayers of alkane chains (Glosli and McClelland 1993), perfluorocarboxylic acid and hydrocarboxylic Langmuir-Blodgett (LB) monolayers (Koike and Yoneya 1996), between contacting copper surfaces (Hammerberg et al. 1995; Sorensen et al. 1996), between a silicon tip and a silicon substrate (Landman et al. 1989a, b), and between contacting diamond surfaces that have organic molecules absorbed on them (Perry and Harrison 1997). These and several other studies are discussed below.

7.4.1 Bare Surfaces

Sliding friction that takes place between two surfaces in the absence of lubricant is termed “dry” friction even if the process occurs in an ambient environment. Simple models have been developed to model dry sliding friction that, for example, consider the motion of a single atom over a monoatomic chain (Buldum and Ciraci 1997). Results from these models reveal how elastic deformation of the substrate from the sliding atom affects energy dissipation and how

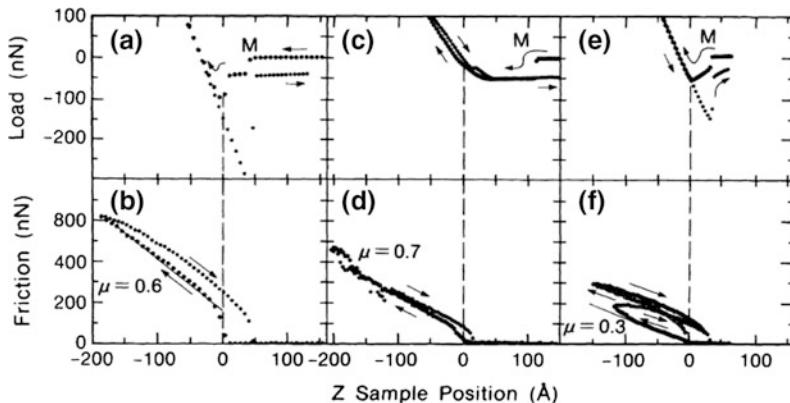


Fig. 7.28 Measured values for friction and load as an atomic-force microscope tip is scanned across a 30 Å thick sample of perfluoropolyether on Si(100). **a** and **b** The unbonded polymer with unreactive end groups. **c** and **d** The unbonded polymer with alcohol end groups. **e** and **f** A bonded polymer. Reprinted with permission from Mate (1992). Copyright 1992 American Physical Society

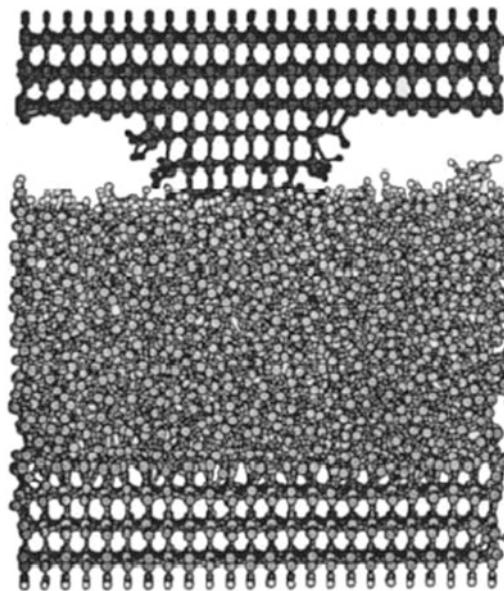


Fig. 7.29 Snapshot from a molecular dynamics simulation where a pyramidal diamond tip indented an amorphous carbon thin film that is 20 layers thick. The simulation took place at room temperature and the carbon atoms in the film were 21% sp^3 -hybridized and 58% sp^2 -hybridized (the remaining atoms were on the surface and were not counted). Reprinted with permission from Sinnott et al. (1997). Copyright 1997 American Institute of Physics

the average frictional force varies with changes in the force constant of the substrate in the direction normal to the scan direction. Much of

the correct behavior involved in dry sliding friction is captured by these types of simple models. However, more detailed models and simulations, such as MD simulations, are required to provide information about more complex phenomena.

MD simulations have been used to study the sliding of metal tips across clean metal surfaces by numerous groups (Sorensen et al. 1996; Sutton and Pithica 1990; Akamine et al. 1990; Nieminen et al. 1992a, b; Pokropivny et al. 1997). An illustrative case is shown in Fig. 7.30 for a copper tip sliding across a copper surface (Sorensen et al. 1996). Adhesion and wear occur when the attractive force between the atoms on the tip and the atoms at the surface becomes greater than the attractive forces within the tip itself. Atomic-scale stick and slip can occur through nucleation and subsequent motion of dislocations, and wear can occur if part of the tip gets left behind on the surface (see Fig. 7.30). The simulations can further provide data on how the characteristic ‘stick-slip’ friction motion can depend on the area of contact, the rate of sliding, and the sliding direction (see Fig. 7.31).

An additional study of stick-slip in the sliding of much larger, square-shaped metal tips across metal surfaces was carried out by Li et al. (2001) using EAM potentials. The initial structure of a NiAl tip and surface system is shown in Fig. 7.32.

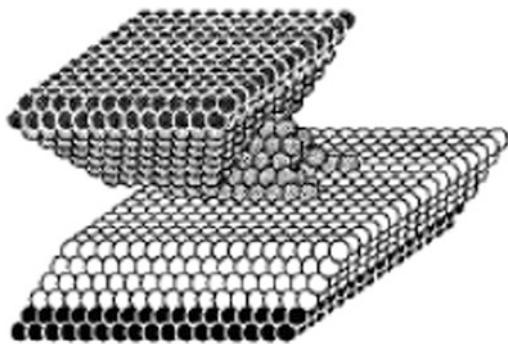


Fig. 7.30 Snapshot from a molecular dynamics simulation of a copper tip sliding across a Cu(100) surface. A connective neck between the two is sheared during the sliding, leading to wear of the tip. The simulation was performed at a temperature of 0 K. Reprinted with permission from Sorensen et al. (1996). Copyright 1996 American Physical Society

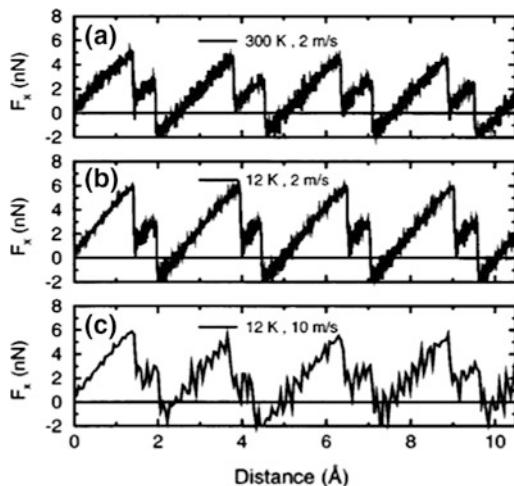


Fig. 7.31 Plots of the lateral force versus distance from a simulation similar to that shown in Fig. 7.30. The plots illustrate the dependence of the force on temperature and sliding velocity. Reprinted with permission from Sorensen et al. (1996). Copyright 1996 American Physical Society

This study predicted that collective elastic deformation of the surface layers in response to sliding is the main cause of the stick-slip behavior shown in Fig. 7.33. The simulations also predicted that stick-slip produces phonons that propagate through the surface slab.

Large scale simulations using pairwise, Morse potentials that are similar in form to Eq. (7.6)

were used to study the wear of metal surfaces caused by metal tips that plow the surface, as illustrated in Fig. 7.34. They provide insight into the wear track dependence of the sliding rate (Fang et al. 2002) and how variations in the scratching force, friction coefficient, and other quantities depend on the scratch depth (Komanduri and Chandrasekaran 2000), as illustrated in Fig. 7.35.

On the whole, the results of experimental studies show good agreement with the results of the computational studies described above. This is true despite the fact that all these MD simulations use empirical potentials that do not include electronic effects and thus effectively assume that the electronic contributions to friction on metal surfaces are negligible. However, experiments have measured a non-negligible contribution of conduction electrons to friction (Dayo et al. 1998). Thus, future simulations of metal tip/metal substrate interactions using more sophisticated tight-binding or first principles methods that include electronic effects are encouraged.

Layered ceramics, such as mica, graphite and MoS₂ that have structures that include strongly bound layers that interact with one another through weak van der Waals bonds, have long been known to have good lubricating properties because of the ease with which the layers slide over one another. They have, therefore, been the focus of some of the earliest experimental studies of nanometer-scale friction (Mate et al. 1987; Erlandsson et al. 1988). The results of these early studies lead researchers to hypothesize that at high loads measured friction forces were related to “incipient sliding” (Johnson and 1985; Pethica 1986) caused by a small flake from the surface becoming attached to the end of the tip. If true, this would mean that all measured interactions were between the surface and the flake, which has a larger contact area than the clean tip. However, subsequent simulations of constant force AFM images of graphite by Tang et al. (1993) showed that there is no need for the assumption of a graphite flake under the tip to reproduce the experimental images of a graphite surface.

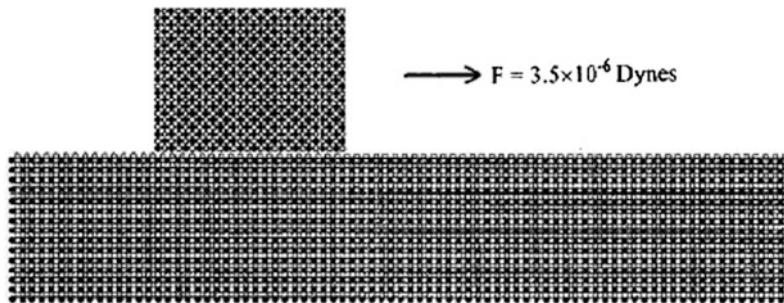


Fig. 7.32 Starting configuration of sliding NiAl tip on a NiAl surface. Reprinted with permission from Li et al. (2001). Copyright 2001 American Institute of Physics

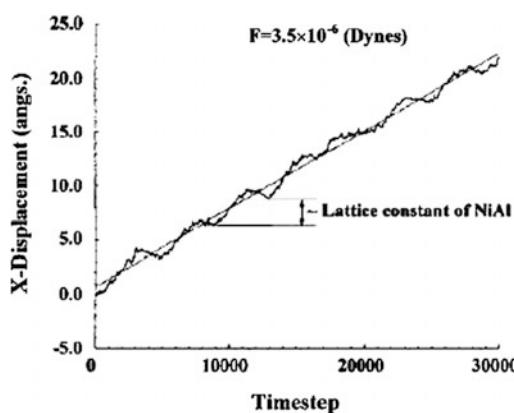


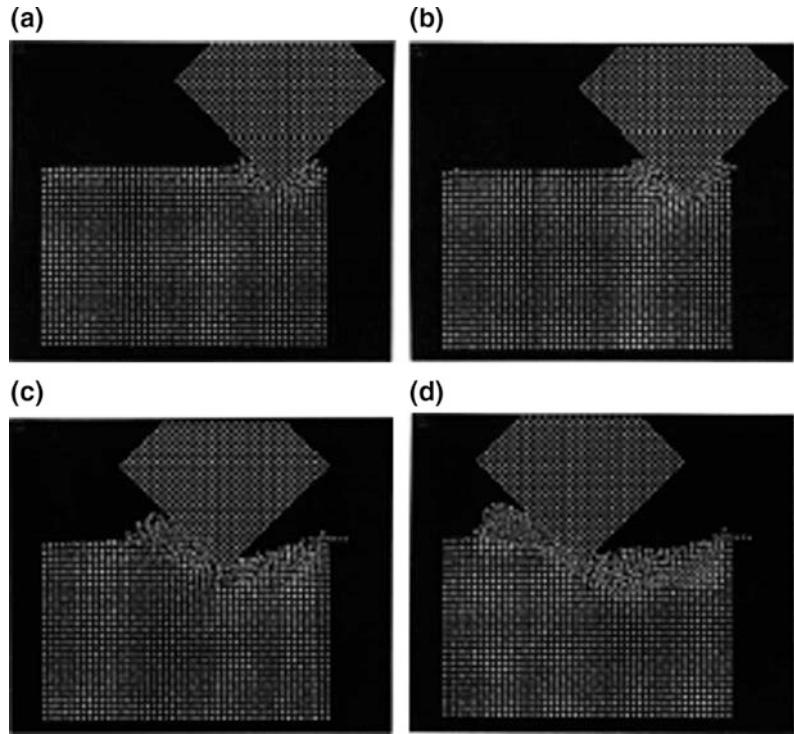
Fig. 7.33 A structured curve of frictional dynamics of an atom in the upper right corner that is indicative of stick-slip. Reprinted with permission from Li et al. (2001). Copyright 2001 American Institute of Physics

Surprisingly strong localized fluctuations in atomic-scale friction are displayed by layered ceramics (Fujisawa et al. 1993, 1994, 1996; Morita et al. 1996). For instance, square-well signals with sub-angstrom lateral width are obtained in FFM scans on MoS₂(001) in the direction across the scan direction, while saw-tooth signals are detected along the scan direction, as shown in Fig. 7.36. This finding can be explained by a stick-slip model by Mate et al. (1987) and Erlandsson et al. (1988) that assumes that the tip does a zigzag walk along the scan. Measured variations in the frictional force with the periodicity of cleavage planes (Erlandsson et al. 1988) are consistent with the results of this simple model. However, additional experiments indicate a more complex tip-surface interaction,

such as changes in the intrinsic lateral force between the substrate and the AFM tip (Ruan and Bhushan 1994) or sliding-induced chemistry between the tip and the surface (Carpick et al. 1996a, b).

MoS₂ has proven to be a very successful solid lubricant for applications that operate in vacuum but its performance quickly deteriorates when exposed to ambient air. Recently, Liang et al. used ab initio DFT methods to examine the potential energy surfaces between sliding MoS₂ (001) | MoS₂ (001), MoS₂ (001) | MoO₃ (001), and MoO₃ (001) | MoO₃ (001) interface systems in an effort to understand the deterioration in lubricity due to oxidation (Liang et al. 2008). The potential energy surfaces then give information on the minimum energy path along particular sliding directions from which lateral forces needed to slide the interface can be calculated. In this work a normal force of 500 MPa was applied to all three interfacial systems before the energy surface calculations were performed. It was found that the minimum energy path for self-mated MoS₂ (001) | MoS₂ (001) was a zig-zag path that avoided direct overlap of sulfur atoms in the topmost layers of the top and bottom surfaces. From this PE surface a lateral frictional force of 0.058 nN/atom was predicted for this interface. The lowest frictional force of 0.011 nN/atom predicted in these calculations was for the MoS₂ (001) | MoO₃ (001) interface along the channel direction formed by S atoms at the sliding surface. Although this doesn't explain the mechanism by which degradation occurs experimentally, it is in general agreement with

Fig. 7.34 Snapshots of the scratching of an aluminum surface with a rigid tip at a depth of 0.8 nm. Reprinted with permission from Komanduri and Chandrasekaran (2000). Copyright 2000 American Physical Society

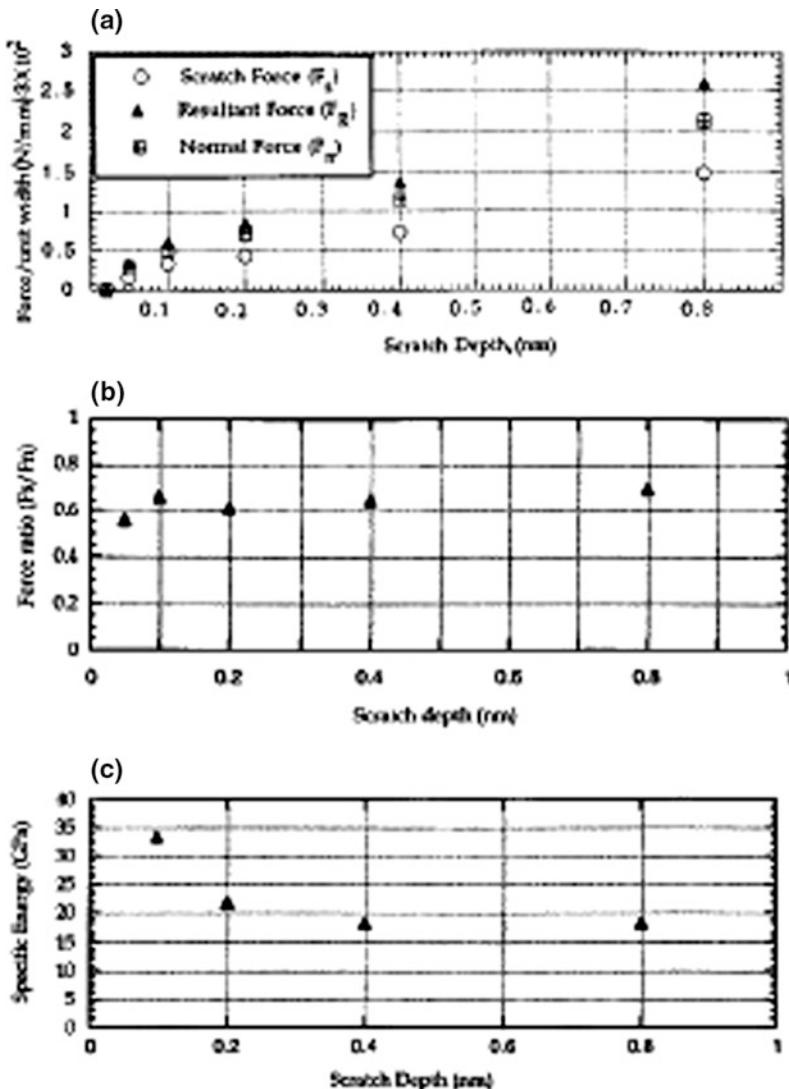


experimental results that show that this interface can produce lower frictional coefficients than pure MoS₂. As suggested by the authors, it may be a surface defect driven process that is currently inaccessible to first principles methods employed in these calculations. The last interface of MoO₃ (001) | MoO₃ (001) was found to have the largest frictional force of 0.352 nN/atom.

Crystalline ceramics differ from layered ceramics in that they are held together by relatively strong covalent or ionic bonds. In the case of ionic systems, Shluger et al. (1995) used a mixture of atomistic and macroscopic modeling methods to study the interaction of a MgO tip and a LiF surface. In particular, the tip-surface interaction was treated atomistically and the cantilever deflection was treated with a macroscopic approach. The results, shown in Fig. 7.37, show that if the tip is charged and in hard contact with the surface, tip and surface distortions are possible that can lead to motion of the surface ions within the surface plane and the transfer of some of the ions onto the tip.

In the case of covalently bound ceramics, there is an extensive literature related to friction of diamond (Samuels and Wilks 1988; Cagin et al. 1999) because, while it is the hardest material known, it also exhibits relatively low friction. The “ratchet mechanism” has been proposed for energy dissipation during friction on the macro-scale in diamond, where energy is released by the transfer of normal force from one surface asperity to another. The elastic mechanism is another mechanism that has been proposed, where the released energy comes from elastic strain in an asperity. Atomic-scale friction has been measured experimentally (Germann et al. 1993) for diamond tips with near atomic-scale radii sliding over hydrogen-terminated diamond surfaces. These experiments are sensitive enough to detect the 2 × 1 reconstruction on the diamond (100) surface. Furthermore, the average friction coefficient determined with an AFM on H-terminated diamond (111) surfaces is about two orders of magnitude smaller than the value measured on bare, 2 × 1 diamond (111) surfaces,

Fig. 7.35 Variation in **a** the scratching force, the normal force, and the resultant force, and **b** the friction coefficient, and **c** the specific energy during scratch processes similar to those shown in Fig. 7.34 at scratch depths ranging from 0.8 nm to almost 0 nm. Reprinted with permission from Komanduri and Chandrasekaran (2000). Copyright 2000 American Physical Society



indicating greater adhesion in the latter case (vanden Oetelaar and Flipse 1997). More recently the friction between a tungsten carbide tip and hydrogen-terminated diamond (111) was examined with AFM in UHV by Enachescu et al. (1998). The friction between these two hard surfaces was shown to obey Derjaguin-Muller-Toporov, or DMT, (Derjaguin et al. 1975) contact mechanics and the shear strength of the interface was determined to be 246 MPa.

Extensive MD simulations have been carried out by Harrison and coworkers that examine the friction between hydrogen-terminated diamond

(111) surfaces (Harrison et al. 1992b, 1995; Gao et al. 2007) and diamond (100) surfaces (Gao et al. 2007; Perry and Harrison 1995) in sliding contact and its temperature dependence (Brukman et al. 2008). The simulations of sliding between the diamond (111) surfaces reveal that the potential energy, load, and friction are all periodic functions of the sliding distance (see Fig. 7.38). Maxima in these quantities occur when the hydrogen atoms on opposing surfaces interact strongly. Recent ab initio studies by Neitola and Park of the friction between hydrogen-terminated diamond (111) surfaces

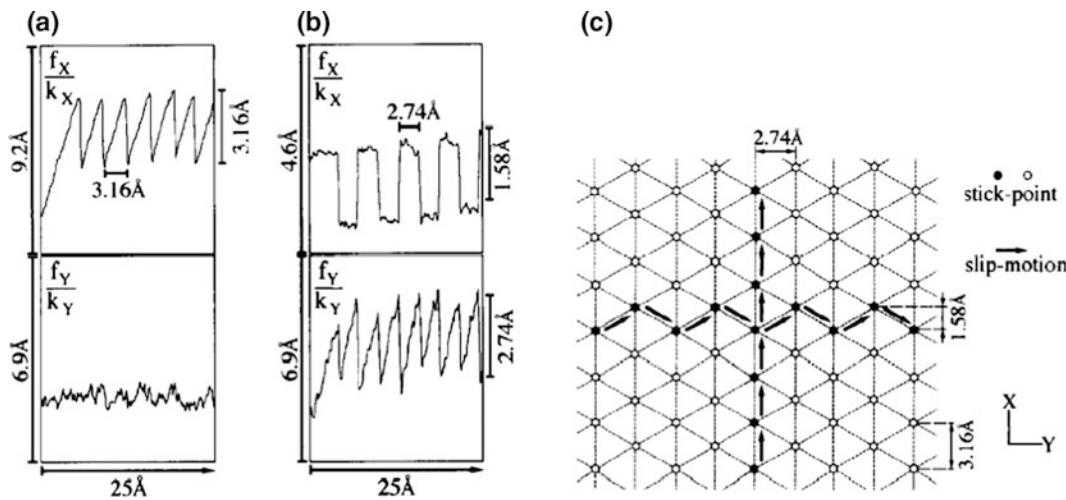


Fig. 7.36 Displacement data from a scan across a MoS₂(001) surface. The data in **a** and **b** are from scans along the x and y directions, respectively, on the surface shown in **c**. Reprinted with permission from Morita et al. (1996). Copyright 1996 American Physical Society

also show that the potential energy is periodic with sliding distance (see Fig. 7.39) (Neitola and Pakkanen 2001). Because the results of the ab initio studies and the MD simulations are in good agreement, Neitola and Park conclude that the potential model used in the MD studies is accurate.

As mentioned previously, the maxima in the load and the friction values during sliding are caused by the interactions of hydrogen atoms on opposing surfaces. When sliding in the [11̄2] direction, the H atoms “revolve” around one another, thus decreasing the repulsive interaction between the sliding surfaces because the hydrogen atoms are not forced to pass directly over one another (Harrison et al. 1992b). Increasing the load causes increased stress at the interface. The opposing hydrogen atoms become “stuck”. Once the stress at the interface becomes large enough to overcome the hydrogen-hydrogen interaction between opposing surfaces, the hydrogen atoms “slip” past one another with the same “revolving” motion observed at low loads. This phenomenon is known as atomic-scale stick-slip and has the periodicity of the diamond lattice. It should be noted that due to the alignment of the opposing surfaces, the hydrogen atoms are directly in line with each other when sliding in

the [11̄2] direction. However, the hydrogen atoms are not “aligned” with each other when sliding in the [1̄10] so the friction in this direction is lower than in the [11̄2] direction. It should be noted, however, that experimentally all initial alignments are likely to be probed.

Harrison and coworkers have further shown that the peaks in the frictional force are correlated with peaks in the temperature of the atoms at the interface when two hydrogen-terminated diamond (111) surfaces are in sliding contact (Harrison et al. 1995). Figure 7.40 shows the vibrational energy (or temperature) between diamond layers as a function of sliding distance. These data clearly show that layers close to the sliding interface can be vibrationally excited during sliding. When the hydrogen atoms are “stuck” or interacting with each other strongly, the stress and friction force at the interface build up. When the hydrogen atoms “slip” past one another, the stress at the interface is relieved and the energy is transferred to the diamond in the form of vibration or heat. Thus, the peaks in the temperature occur slightly after the peaks in the frictional force.

It should be noted that atomic-scale stick-slip is observed in other systems. Harrison and coworkers used MD simulations to demonstrate

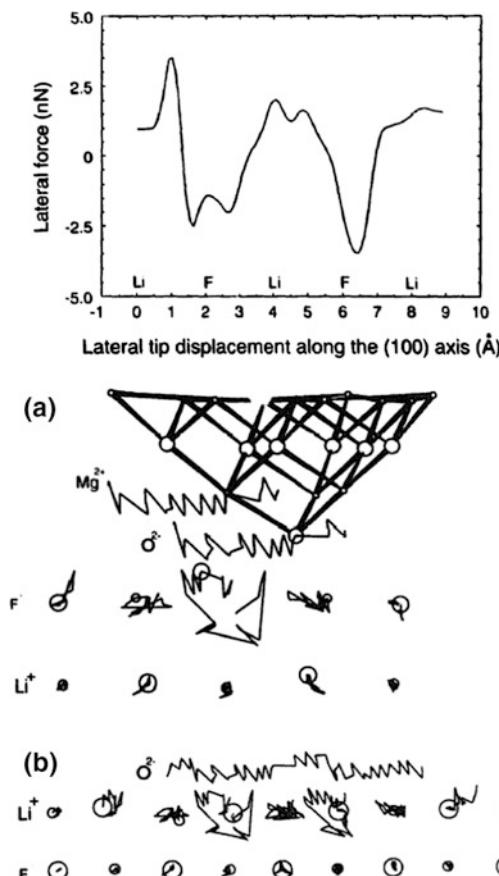


Fig. 7.37 *Top* The lateral force calculated for a MgO tip scanning in the $\langle 001 \rangle$ direction on LiF(001). *Bottom* A view of the side of the surface plane along the scan direction. The surface Li^+ and F^- atoms are seen to relax to relieve the frictional energy and this relaxation motions is indicated in the figure by the category lines. **a** How a F^- ion on the surface can be moved into an interstitial site by the tip and then returns to its original position. **b** How the relaxation of the surface atoms is reversible. Reprinted from Shluger et al. (1995). Copyright 1995 with permission from Elsevier

that two hydrogen-terminated diamond (100) (2×1) surfaces in sliding contact also exhibit stick-slip (Perry and Harrison 1995). In addition, it was shown that the shape of the friction versus sliding distance curves is influenced slightly by the speed of the sliding with features in the curves becoming more pronounced at slower speeds. Stick-slip behavior was also observed in AFM studies of diamond (100) (2×1) surfaces (vanden Oetelaar and Flipse 1997). However, in this case the stick-slip

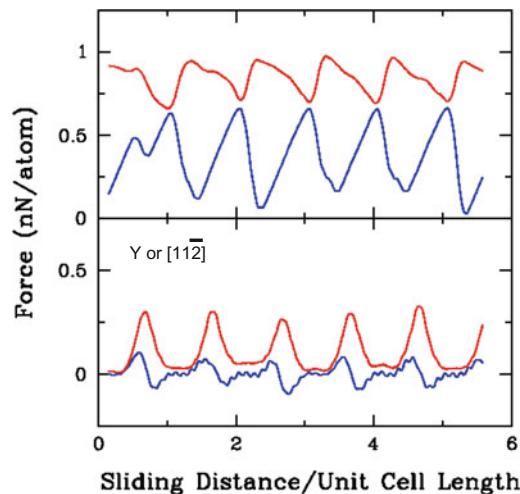


Fig. 7.38 Calculated frictional force (blue/lower lines) and normal force (red/upper lines) felt by a hydrogen-terminated (111) surface as it slides against another hydrogen terminated diamond (111) surface in a MD simulation. The sliding direction is given in the legend. The sliding speed is 1 $\text{\AA}/\text{ps}$ and the simulation temp is 300 K. The two plots show how the simulated stick-slip motion changes as a function of the applied load. The load is high and low in the upper and lower panels, respectively. Reprinted with permission from Perry and Harrison (1995). Copyright 1995 American Chemical Society

was over a much longer length scale and may be due to the fact that the surfaces were not hydrogen terminated.

Mulliah et al. (2004) used MD simulations with bond-order potentials (Brenner 1990b) to model interactions between indenter atoms, EAM potentials (Ackland et al. 1987) to model interactions between substrates atoms, and the Ziegler-Biersack-Littmark potential (Biersack et al. 1985) to model interactions between indenter and substrate atoms to study the atomic-scale stick-slip phenomenon of a pyramidal diamond tip interacting with a silver surface at several sliding rates and vertical support displacements. These simulations showed that dislocations are related to the stick events emitting a dislocation in the substrate near the tip. The scratch in the substrate is discrete due to the tip jumping over the surface in the case of small vertical displacements. In contrast, large displacements of 15 \AA or more result in a

Fig. 7.39 Distance between hydrogen-terminated (111) crystals as a function of sliding distance. Reprinted with permission from Neitola and Pakkanen (2001). Copyright 2001 American Chemical Society

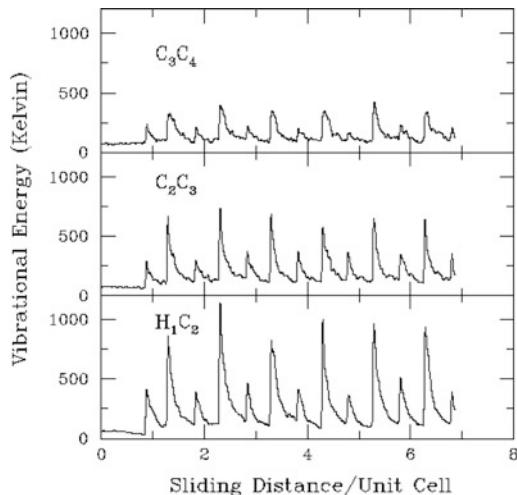
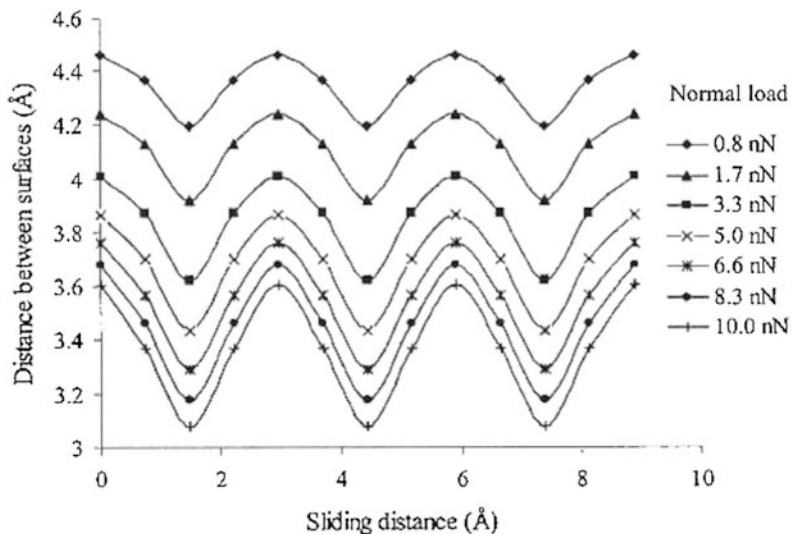


Fig. 7.40 Average vibrational energy of oscillators between diamond layers as a function of sliding distance. These energies are derived from a molecular dynamics simulation of the sliding of a hydrogen-terminated diamond (111) surface over another hydrogen-terminated diamond (111) surface. The vibrational energy between the first and second layers of the lower diamond surface is shown in the lower panel, between the second and third layers in the middle panel, and between the third and fourth layers in the upper panel. Reprinted from Harrison et al. (1995). Copyright 1995 with permission from Elsevier

continuous scratch. These simulations also showed how the dynamic friction coefficient and the static friction coefficient increase with

increasing tip depth. The tip moves continuously through a stick and slip motion at large depths, whereas it comes to a halt in the case of shallow indents. Although the sliding rate can change the exact points of stick and slip, the range of sliding rates over the range of values considered in this study (1.0–5.0 m/s) has no influence on the damage to the substrate, the atomistic stick-slip mechanisms, or the calculated friction coefficients.

The effect of the way in which the tip is rastered across the surface in MD simulations was considered by Cai and Wang (2001, 2002) using bond-order potentials. In particular, they dragged silicon tips across several silicon surfaces, as illustrated in Fig. 7.41, in two different ways. In the first, they moved the tip every MD step while in the other they advanced the tip every 1000 steps. In both cases the overall sliding rate is the same and equals 1.67 m/s. In both cases, wear of the tip such as is illustrated in Fig. 7.41 occurs. However, the mechanisms by which the wear occurs are found to depend on the approach used, and the latter approach is found to be in better agreement with experimental data.

In many studies diamond tips or diamond decorated tips are used in friction measurements. Diamond is an attractive material for an FFM tip because of its high mechanical strength and the belief that such tips are wear resistant. However,

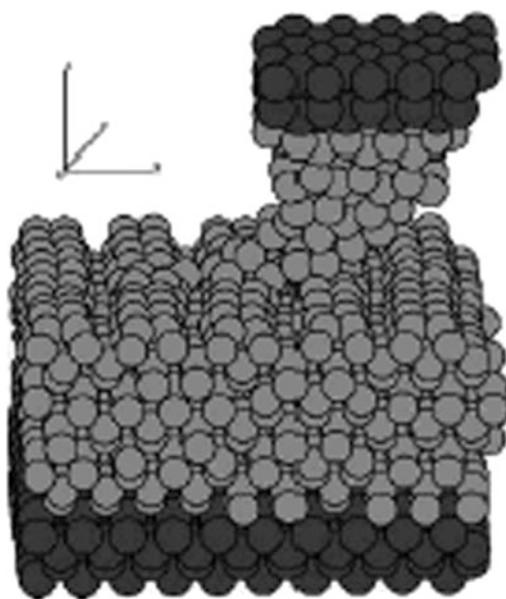


Fig. 7.41 Snapshots of a Si(111) tip interacting with a Si(001) 2×1 surface. The tip is rastering along the surface in the x-direction and starts off at a distance of 9 Å from the surface. Reprinted with permission from Cai and Wang (2002). Copyright 2002 Council for the Central Laboratory of the Research Councils

diamond tips that were used to scratch diamond and silicon surfaces and then imaged showed significant wear that increased with the increasing hardness of the tested material (Khurshudov et al. 1996; Khurshudov and Kato 1995). This wear altered the shape of the tip and hence influences the contact area that is used to determine friction coefficients.

In summary, MD simulations provide insight into dry sliding friction and the sliding of metal tips across clean metal, crystalline ceramics, and layered ceramics surfaces. Stick-slip friction or wear can occur depending on the sliding conditions. The good lubricating properties of layered ceramics are observed in the simulations along with localized fluctuations in atomic-scale friction. Crystalline ceramics, such as diamond, exhibit relatively low friction and the simulations show how stick-slip atomic-scale motion changes with the conditions of sliding and the way in which the simulation is performed.

7.4.2 Decorated Surfaces

While dry sliding friction in vacuum assumes that ambient gas particles have no direct effect on the results, MD simulations show that free particles between two surfaces in sliding contact influence friction to a surprisingly large degree. These so-called third-body molecules have been studied extensively by Perry and Harrison (1996a, b, 1997) using MD simulations with bond-order and LJ potentials. These simulations focus on the effect of trapped small hydrocarbon molecules on the atomic-scale friction of two (111) crystal faces of diamond with hydrogen termination. These molecules might represent hydrocarbon contamination trapped between contacting surfaces prior to a sliding experiment in dry friction, or hydrocarbon debris formed during sliding.

In particular, the effects on friction of methane (CH_4), ethane (C_2H_6), and isobutane ($(\text{CH}_3)_3\text{CH}$) trapped between diamond (111) surfaces in sliding contact were examined in separate studies (see Fig. 7.42). The frictional force for all these systems generally increases as the load increases, as illustrated in Fig. 7.43. The simulations predict that the third-body molecules markedly reduce the average frictional force compared to the results for pristine hydrogen-terminated surfaces. This is particularly true at high loads where the third-body molecules act as a boundary layer between the two diamond surfaces. That is, the third-body molecules reduce the interaction of hydrogen atoms on opposing surfaces (Perry and Harrison 1996b). This is demonstrated by examining the vibrational energy excited in the diamond lattice during the sliding (Fig. 7.44). Significant vibrational excitation of the diamond outer layer (C–H) occurs in the absence of the methane molecules. Thus, the friction is approximately 3.5 times larger when methane is not present. The application of load to the diamond surfaces causes the normal mode vibrations of the trapped methane molecules to change. Power spectra calculated from MD simulations (Perry and Harrison 1996a, b) show

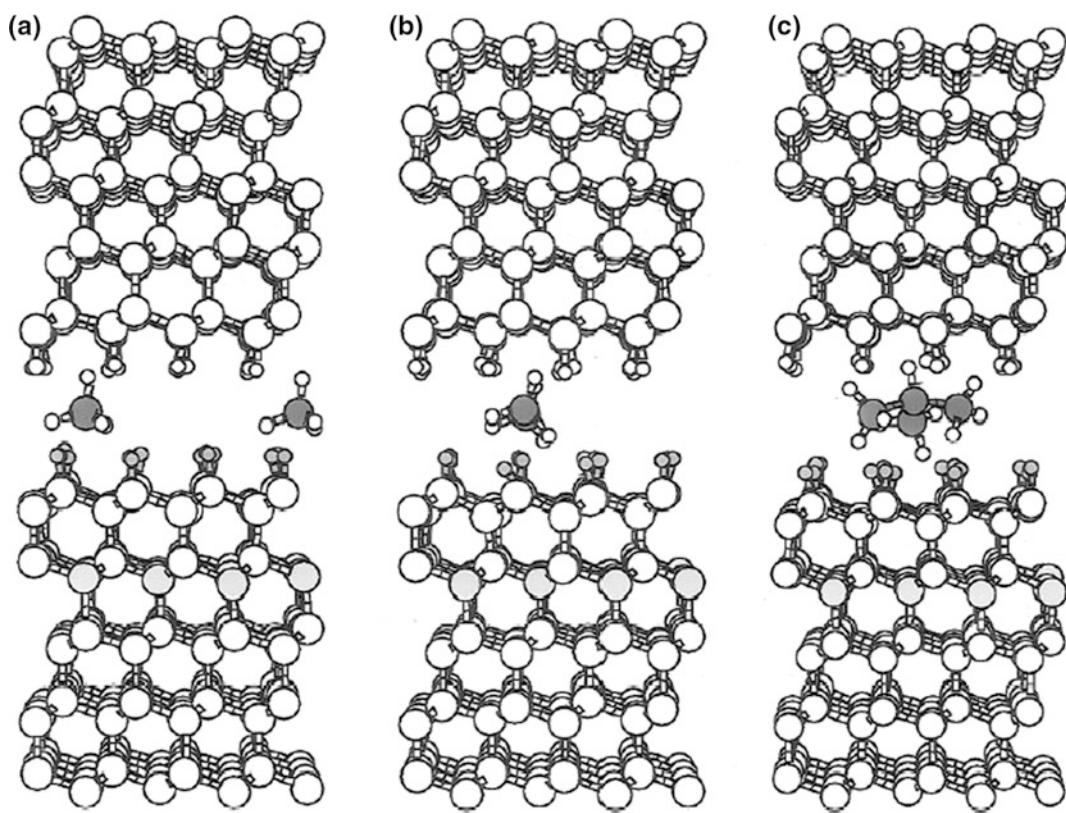


Fig. 7.42 Initial configuration at low load for the diamond plus third-body molecules systems. These systems are composed of two diamond surfaces, viewed along the $\langle\bar{1}10\rangle$ direction, and two methane molecules in (a), one ethane molecules in (b), and one isobutane molecule in (c). Large white and dark gray spheres represent carbon atoms of the diamond surfaces and the third-body molecules, respectively. Small gray spheres represent hydrogen atoms of the lower diamond surface. Hydrogen atoms of the upper diamond surface and the third body molecules are both represented by small white spheres. Sliding is achieved by moving the rigid layers of the upper surface from left to right in the figure. Reprinted with permission from Perry and Harrison (1997). Copyright 1997 American Chemical Society

that even under low loads, the peaks in the power spectra are significantly broadened. Peaks in the low energy region of the spectrum almost disappear with the additional application of load.

The size of the methane molecules allows them to be “pushed” in between hydrogen atoms on the diamond surfaces while sliding (Perry and Harrison 1996b). However, steric considerations cause the larger ethane and isobutane molecules to change orientation during sliding. Conformations that lead to increased interactions with the diamond surfaces increase the average frictional forces. Thus, despite the fact that the two diamond surfaces are farther apart when ethane and isobutane are present compared to when methane

is present, the friction is larger because these molecules do not “fit” nicely into potential energy valleys between hydrogen atoms when sliding.

When similar hydrocarbon molecules (methyl, ethyl, and n-propyl groups) are chemisorbed to one of the sliding diamond surfaces, instead of trapped between the surfaces, different behavior is observed by Harrison et al. (1993a, b, c, 1995) Simulations show that methyl-termination does not decrease friction significantly but results in frictional forces that are nearly the same as they are for hydrogen-terminated diamond surfaces (Harrison et al. 1992c). While the methane third-body molecules decrease the frictional force

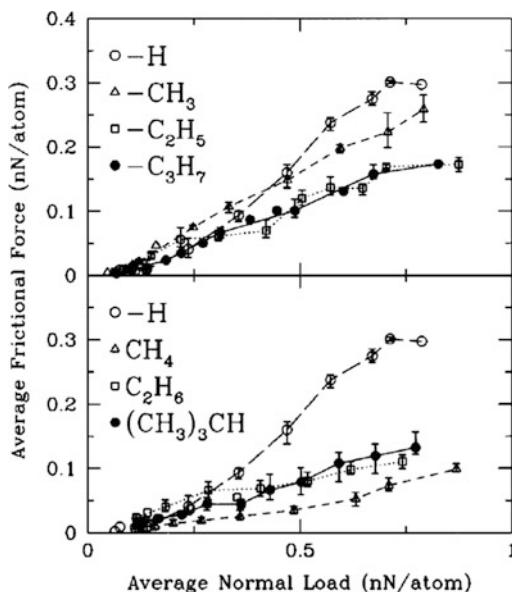


Fig. 7.43 Average frictional force per rigid-layer atom as a function of average normal load per rigid-layer atom for sliding the upper diamond surface in the $[11\bar{2}]$ crystallographic direction. Data for the methane (CH_4) system (open triangles), the ethane (C_2H_6) system (open squares), the isobutane ($(\text{CH}_3)_3\text{CH}$) system (filled circles), and diamond surfaces in the absence of third-body molecules (open circles) are shown in the lower panel. Data for the methyl-terminated ($-\text{CH}_3$) system (open triangles), the ethyl-terminated ($-\text{C}_2\text{H}_5$) system (open squares), the *n*-propyl-terminated ($-\text{C}_3\text{H}_7$) system (filled circles), and diamond surfaces in the absence of third-body molecules (open circles) are shown in the upper panel. Lines have been drawn to aid the eye. Reprinted with permission from Perry and Harrison (1997). Copyright 1997 American Chemical Society

to a greater extent than the chemisorbed methyl groups, friction as a function of load is comparable for the ethyl-terminated and ethane systems, with the former giving slightly higher frictional forces. Attaching the hydrocarbon groups to the diamond surfaces causes them to have less freedom to move between hydrogen atoms on opposing diamond surfaces during sliding. This generally increases their repulsive interaction with the diamond counterpart.

MD simulations can also provide insight into the rich, nonequilibrium tribochemistry that occurs between surfaces in sliding contact. Harrison and Brenner examined tribochemistry that occurs when ethane molecules are trapped

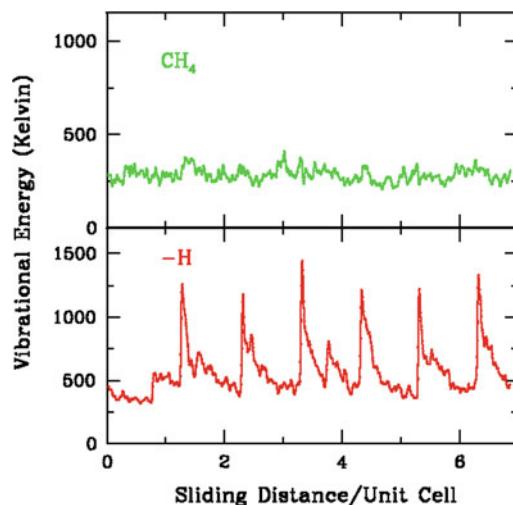


Fig. 7.44 Average vibrational energy between the (C–H) bonds of the upper diamond (111) surface versus sliding distance for hydrogen-terminated diamond (111) surfaces, with (CH_4) and without methane (H), trapped between them. The average normal load is approximately the same in both simulations and is in the range 0.8–0.85 nN/atom. The average frictional force on the upper surface is about 3.5 times smaller in the presence of the methane third-body molecules. Reprinted from Perry and Harrison (1996b). Copyright 1996 with permission from Elsevier

between diamond surfaces in sliding contact, as illustrated in Fig. 7.45 (Harrison and Brenner 1994). This simulation was the first to show the atomic-scale mechanisms for the degradation of lubricant molecules due to friction. The type of debris formed during the sliding simulation is similar to the types of debris molecules that were observed in macroscopic experiments that examined the friction between diamond surfaces (Feng and Field 1991).

In the case of sliding metal surfaces, impurity molecules or atoms, both adsorbed and absorbed, on thin metal films can be expected to affect the film's properties. For example, calculations have shown that resistivity changes in the metal are strongly dependent on the nature of the adsorption bond (Persson 1993). When this result is used to interpret atomic-scale friction results obtained with the QCM, the sliding of adsorbate structures on metal surfaces are shown to be a combination of electron excitation and lattice vibrations. Additionally, other interesting

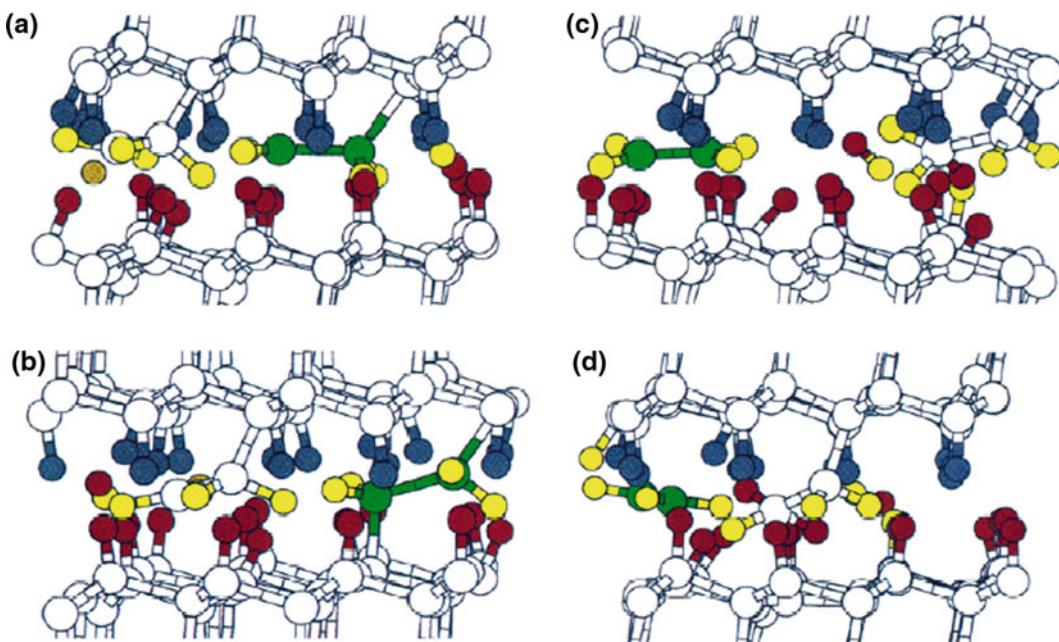


Fig. 7.45 Snapshots from a molecular dynamics simulation of the sliding of two hydrogen-terminated diamond (111) surfaces against one another in the [112] direction. The upper surface has two ethyl fragments chemisorbed to it. The simulation shows how sliding can induce chemistry at the interface. Reprinted with permission from Perry and Harrison (1997). Copyright 1994 American Chemical Society

quantum effects can come into play when the adsorbate is very different chemically from the surface on which it is sliding. For instance, the electronic frictional forces acting on small, inert atoms and molecules, such as C₂H₆ and Xe, sliding on metal surfaces have been calculated by Persson and Volokitin (1995), where the metal surface was approximated by an electron gas (jellium) model. The calculations showed that the Pauli repulsive and attractive van der Waals forces are of similar magnitudes. In addition, the calculated electronic friction contributions agree well with the values derived from surface resistivity by Grabhorn et al. (1992) and QCM measurements. These studies showed that parallel friction is mainly due to electronic effects while perpendicular friction is phononic in nature in this system.

In summary, MD simulations show that the average frictional force decreases significantly in systems with third-body molecules, especially at high loads. Simulations also provide information

about the details of tribochemical interactions that can occur between lubricants and sliding surfaces. Additionally, the effect of the presence of small molecules on thin metal films can influence film properties, such as resistivity.

7.4.3 Thin Films

As discussed at the beginning of this section, the conversion of the work of sliding into some other less ordered form is responsible for friction at sliding solid interfaces. In the case of adhering systems, the work of sliding may be converted into damage within the bulk (plastic deformation), while in the case of weakly adhesive forces, friction can occur through the conversion of work to heat at the interface that causes no permanent damage to the surface (wearless friction). The latter case, when it is achieved through the presence of lubricating thin films, is the topic of this subsection.

There are several types of lubricating thin films, the simplest of which consist of small molecules that are analogous to wear debris that can “roll” between the sliding surfaces or that represent very short-chain bonded lubricants. These thin films were discussed in the previous subsection. The rest of this subsection will, therefore, focus on the effects of liquids, larger nanoparticles, self-assembled monolayers, and solid thin films on lubrication and friction.

7.4.3.1 Liquids

Liquids are common lubricants and have therefore been much studied at the macroscale. At the nanoscale, the tribological response of spherical liquid molecules has been well-characterized experimentally using the SFA and computationally with MD simulations by Berman et al. (1996). The SFA experiments considered one to three liquid layers and the stick-slip motion at the interface is found to increase in a quantized fashion as the number of lubricant layers decreased. When no external forces are applied to the system, the sliding stops and the solid-lubricant interactions are strong enough to force the liquid molecules to form a close-packed structure that is ordered. The transformation of the liquid into this solid-like structure causes the two surfaces to effectively bond to each other through the lubricant. When the surfaces start to slide again, lateral shear forces are introduced that steadily increase, which causes the molecules in the liquid to undergo small lateral displacements that change the film thickness. If these shear forces become greater than a critical value, the film disorders in a manner that is analogous to melting. This allows the surfaces to slide easily past each other in a manner that is still quantized. This sequence of events is nicely illustrated in Fig. 7.46 and can be reproduced multiple times for the same system.

Persson et al. (1995) used MD simulations with pairwise potentials similar to those in Eq. (7.6) to examine the mechanism by which this sharp transition occurs. They find that in the case of sliding on insulating crystal surfaces, the solid-state lubricant may be in a “superlubric” state where the friction is negligible. It is clear

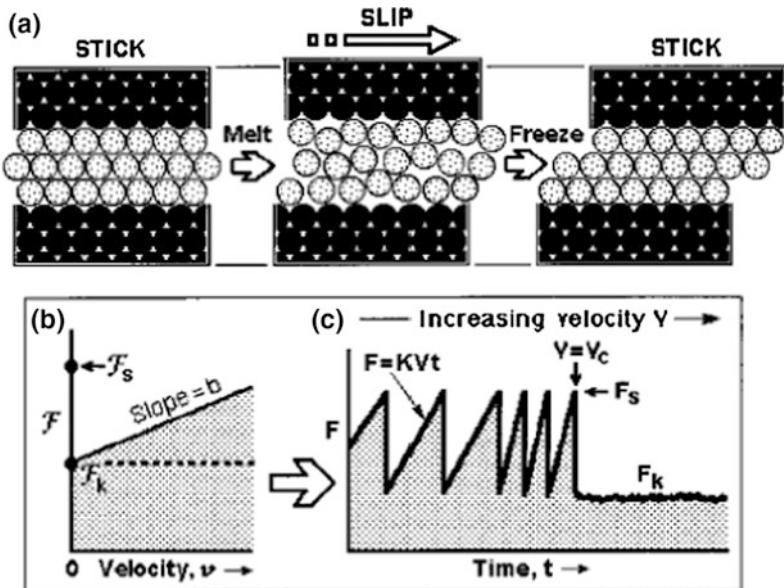
from the simulations, however, that any surface defects, even in low concentrations, will disrupt this state and transform the lubricant back into a fluid. In addition, when sliding occurs on metallic surfaces above cryogenic temperatures, the electronic contributions to friction are no longer zero and no superlubric state is possible.

High applied pressures can force the fluid molecules out from between the two confining surfaces (Persson and Tosatti 1994). The fact that liquid molecules close to a stiff surface are strongly layered in the direction perpendicular to the surface explains the experimental observation of a ($n \rightarrow n-1$) layer transition, where n is number of monolayers, that is observed as the normal load increases (Yoshizawa and Israelachvili 1993). Nucleation theory is used to calculate the critical pressure and determine the spreading dynamics of the ($n-1$) island.

The reactivity of the liquid molecules are also critically important to boundary layer friction. MD studies by Persson (1997) show that inert molecules interact weakly with sliding surfaces. Consequently, as the rate of sliding increases, the molecular conversion from the solid state to the liquid state occurs in an abrupt manner. However, when the molecules interact strongly with the surfaces, they undergo a more gradual transition from the solid to the liquid state. Persson et al. also considered systems where the molecules are attached to one of the surfaces, which causes the transitions to be abrupt. This is especially true if there are large separations between the chains.

While the studies discussed so far have focused on spherical liquids, most widely used liquid lubricants consist of long-chain hydrocarbons. Non-spherical liquid molecules have more difficulty aligning and solidifying. This is borne out in MD simulations by Thompson and Robbins (1990) that show that spherical molecules have higher critical velocities than branched molecules. In particular, the simulations show that when the molecules are branched, the amount of time various parts of the system spend in the sticking and sliding modes changes with sliding rate. The critical velocity can also depend on the number of liquid layers in the film, the

Fig. 7.46 The stick/slip transition that occurs for thin films of liquid between two sliding solid surfaces. F is the intrinsic friction and F_s is the friction where the liquid is in the rigid state; F_k is the friction where the liquid is in the liquidlike state. Reprinted with permission from Berman et al. (1996). Copyright 1996 American Chemical Society



structure and relative orientation of the two sliding surfaces, the applied load and the stiffness of the surfaces.

Additional studies by Landman et al. (1996) used MD simulations with bond-order and EAM potentials coupled with pair-wise potentials similar to Eq. (7.6) to study the sliding of two gold surfaces with pyramidal asperities that have straight chain C₁₆H₃₄ lubricant molecules trapped between them, as illustrated in Fig. 7.47. An important aspect of this study is that the sliding rate in the simulations is about 10 m/s, which is the same order of magnitude as the scanning speed in a computer disk. As the asperities approach each other, the hydrocarbon molecules begin to form layers. This is reflected in the oscillations in the frictional force shown in Fig. 7.48. When the asperities overlap in height and approach each other laterally, the pressure of the lubricant molecules increases to about 4 GPa which leads to the deformation of the gold asperities.

Glosli and McClelland (1993) modeled the sliding of two ordered monolayers of alkane chains that are attached to two rigid substrates. This system is shown schematically in Fig. 7.49. The simulations predicted that energy dissipation occurs by a discontinuous plucking mechanism (sudden release of shear strain) or a viscous

mechanism (continuous collisions of atoms of opposite films). The specific mechanism that occurs depends on the interfacial interaction strength. In particular, the “pluck” occurs when mechanical energy stored as strain is converted into thermal energy that leads to low friction forces at low temperatures. On the other hand, at higher temperatures some of the energy of sliding is dissipated through phonon excitations, which results in higher frictional forces. Interestingly, this trend reverses again at the highest temperatures considered when the molecules move so much that they slide easily over the surfaces, which decreases the frictional force. These results are summarized in Figs. 7.50 and 7.51.

Other studies of sliding surfaces with attached organic chains include MD simulations with LJ potentials by Muser and co-workers, (Muser 2002b; Kreer et al. 2001, 2003) which considered friction between polymer “brushes” in sliding contact with one another. In particular, they considered the effect of sliding rate on the tilting of polymers and the effect of steady-state sliding versus non-steady-state, or transient, sliding. The simulations find that shear forces are lower for chains that tilt in a direction that is parallel to the shear direction. This tilting effect is significant for grafted polymers, as illustrated in

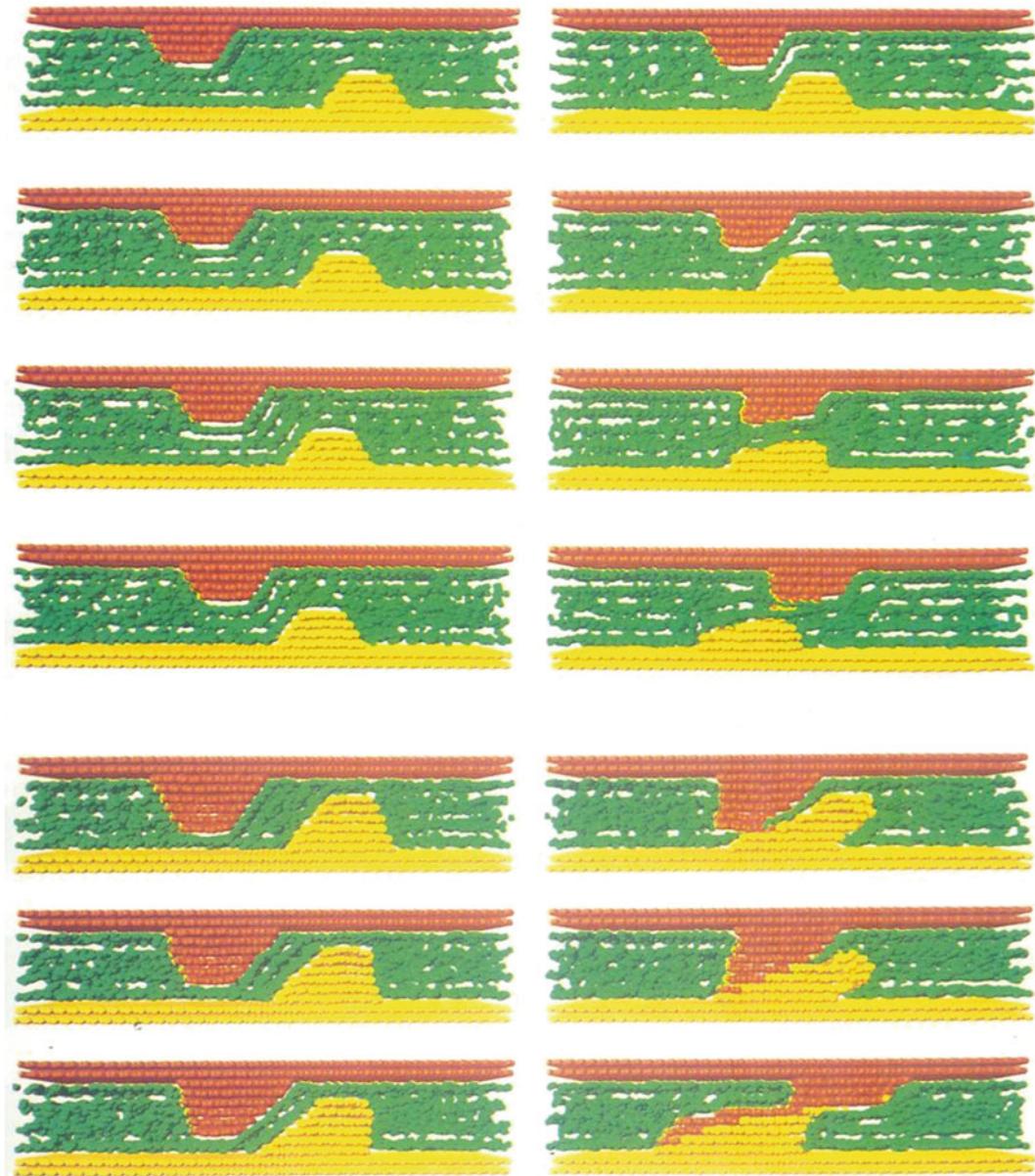


Fig. 7.47 Stills from a molecular dynamics simulation where Au(111) surfaces with surface roughness slide over one another while separated by hexadecane molecules. The scanning velocity is 10 m/s. Layering of the lubricant and asperity deformation occurs as the sliding continues. The top three rows show the results when the asperity heights are separated by 4.6 Å. The bottom three rows show the results when the asperity heights are separated by -6.7 Å. Reprinted with permission from Landman et al. (1996). Copyright 1996 American Chemical Society

Fig. 7.52, and less significant for absorbed polymers. This is due to the decrease of the differential frictional coefficient for the grafted polymers but the increase in the friction coefficient for absorbed polymers under shear. The

tilting is also affected by the rate of sliding and is much larger at high sliding rates than small rates, as indicated in Fig. 7.52. The simulations further show that the inclination angle of the chains decreases much more slowly than the shear

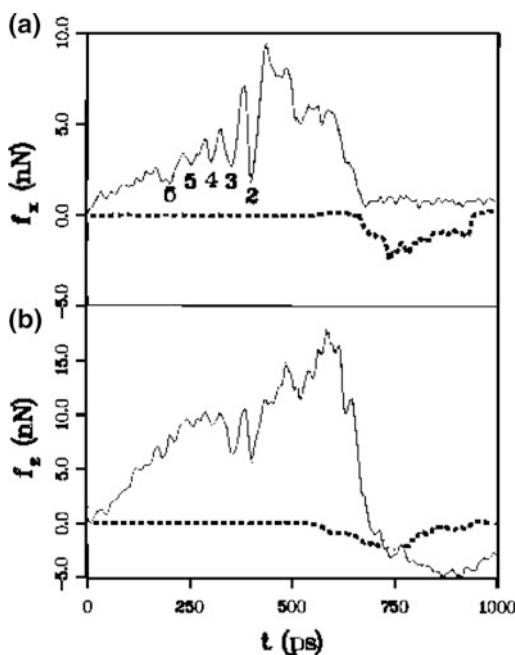


Fig. 7.48 The lateral force (f_x) and normal force (f_z) from the molecular dynamics simulations shown in Fig. 7.35 as a function of time. The forces between the two metal surfaces are shown by the dashed line. The force oscillations correspond to the structural changes of the lubricant in Fig. 7.35. Reprinted with permission from Landman et al. (1996). Copyright 1996 American Chemical Society

stress, and the shear stress maximum is more pronounced if there is hysteresis in the chain orientations.

Typical friction loops for tips that are functionalized and sliding against surfaces that are functionalized in the same manner as illustrated in Fig. 7.25 are shown in Fig. 7.53. The friction force between the OH/OH pairs is significantly

larger than the friction force between the CH_3/CH_3 pairs. This is due to the formation and breaking of hydrogen bonds during the shearing for the OH/OH pairs. The mean forces vs. load forces for the OH/OH and CH_3/CH_3 pairs given in Fig. 7.54 are reduced by the tip radius.

MD simulations by Manias et al. considered the shearing of entangled oligomer chains that are attached to sliding surfaces, as illustrated in Fig. 7.55 (Manias et al. 1996). They find that slip takes place within the film and that this occurs through changes in the chain conformations. Increased viscosity is predicted at the film-surface interface compared to the middle of the film which results in a range of viscosities across the film as one moves away from the points of sliding contact.

To summarize this section, experiments and MD simulations show similar stick/slip transitions that occur for thin films of liquid between two sliding solid surfaces. Frictional properties are found to depend to a significant degree on molecular shape, whether the molecules are grafted on the surfaces or merely absorbed on them, and the degree of tilting in the case of molecular chains. In the case of long-chain molecules, temperature is found to affect the frictional force because the mechanical energy stored in long-chains can be converted into the thermal energy by friction.

7.4.3.2 Self-Assembled and Polymer Thin Film Structures

There have been numerous experimental studies of friction on SAMs on solid surfaces with AFM and FFM. The experimental results reveal

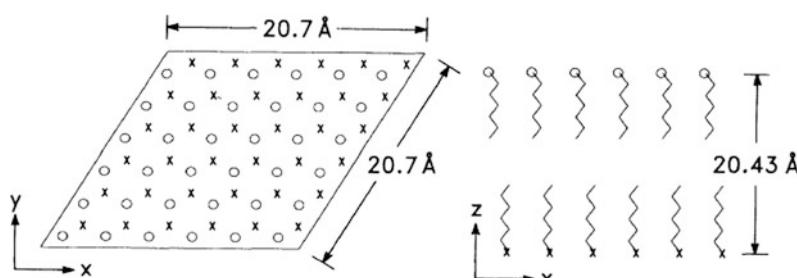


Fig. 7.49 Top and side views of the alkane chains attached to surfaces that are sliding against each other. Reprinted with permission from Glosli and McClelland (1993). Copyright 1993 American Physical Society

Fig. 7.50 Data from molecular dynamics simulations of the sliding of the surfaces shown in Fig. 7.32. **a–f** The shear stress and **g–j** the heat flow as a function of sliding for normal and reduced interfacial strengths. The plots show how the calculated values change with system temperature. Reprinted with permission from Glosli and McClelland (1993). Copyright 1993 American Physical Society

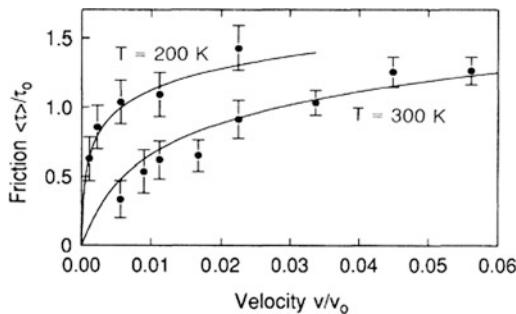
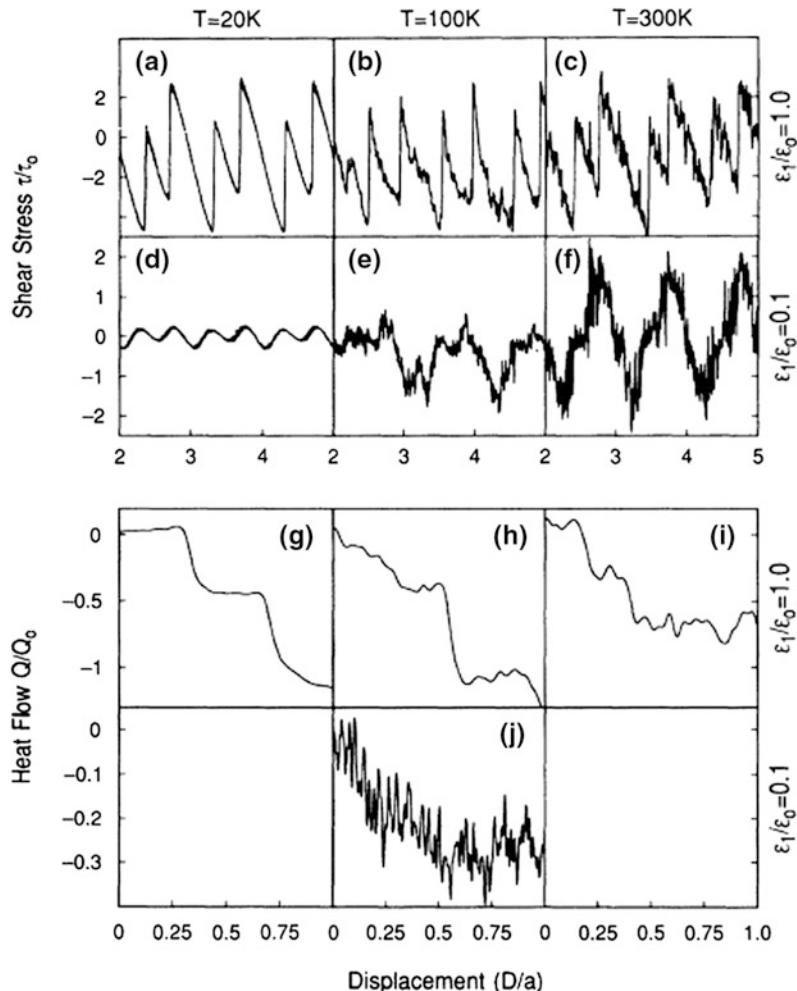


Fig. 7.51 A plot of calculated values of the average interfacial shear stress as a function of the velocity of sliding of the two surfaces shown in Fig. 7.33. Reprinted with permission from Glosli and McClelland (1993). Copyright 1993 American Physical Society

relationships among elastic compliance, topography and friction on thin LB films (Overney et al. 1994a). For example, they have detected differences in the adhesive interactions between the microscope tips and CH_3 and CF_3 end groups (Burnham et al. 1990). Fluorocarbon domains generally exhibit higher friction than the hydrocarbon films, which the authors attribute to the lower elasticity modulus of the fluorocarbon films that results in a larger contact area between the tip and the sample (Overney et al. 1992, 1994a, b). Perry and coworkers examined the friction of alkanethiols terminated with $-\text{CF}_3$ and $-\text{CH}_3$ (Li et al. 2005). The lattice constants for both films

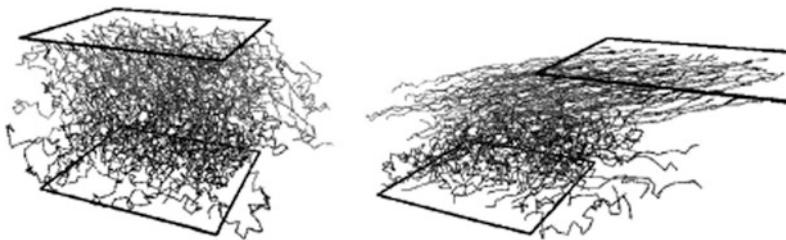


Fig. 7.52 Snapshots of sliding walls with attached polymers in a solvent. Right-hand figure illustrates the sliding process at low sliding rates while the left-hand figure illustrates the sliding process at high sliding rates Reprinted with permission from Müser (2002b), from Kreer et al. (2001). Copyright 2002 Council for the Central Laboratory of the Research Councils

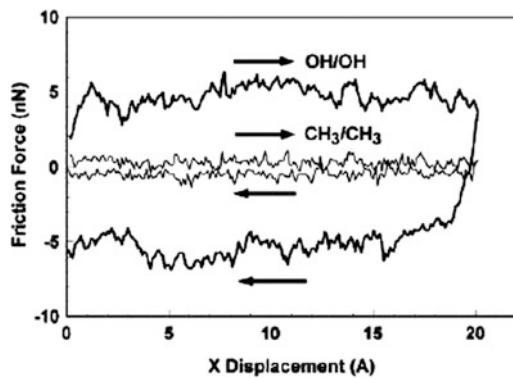


Fig. 7.53 Typical friction loops for the systems shown in Fig. 7.25 for CH₃/CH₃ and OH/OH pairs under a contact load of 0.2 nN. Reprinted with permission from Leng and Jiang (2002). Copyright 2002 American Chemical Society

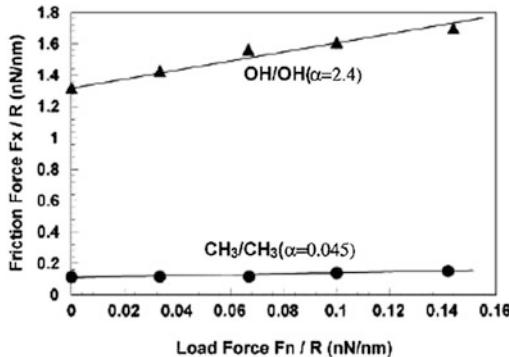


Fig. 7.54 Friction force versus contact load from the systems shown in Fig. 7.25 for CH₃/CH₃ and OH/OH. Reprinted with permission from Leng and Jiang (2002). Copyright 2002 American Chemical Society

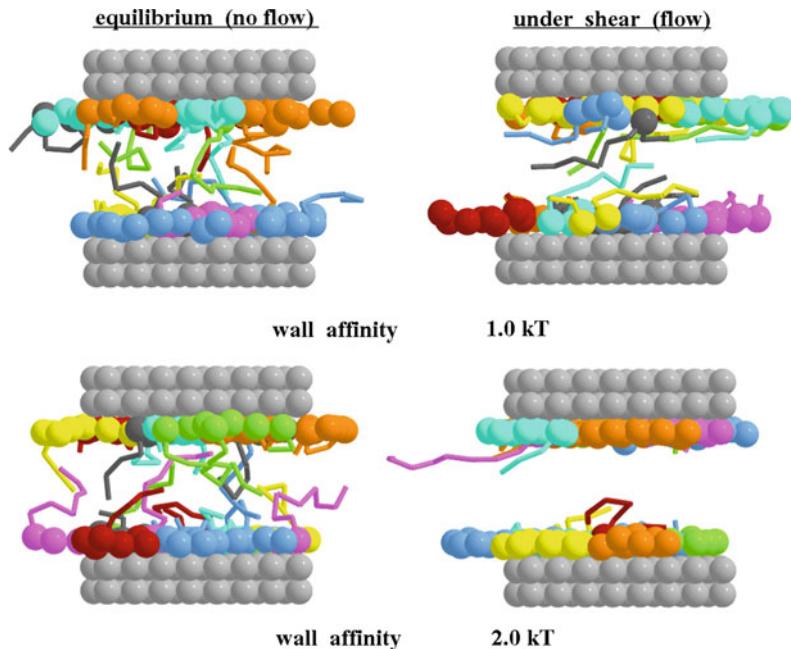
are similar and the films are well ordered. The friction of the SAMs with chains that are terminated with fluorine end groups is larger than the

friction of the SAMs with chains that are terminated with hydrogen end groups. However, the pull off force is similar in both systems, which implies that these end groups have similar contact areas. The authors speculate that the larger $-CF_3$ groups interact more strongly with adjacent chains than the $-CH_3$ -terminated chains. Therefore, the fluorinated chains have more modes of energy dissipation within the plane of the monolayer and, thus, have larger friction.

Molecular disorder of the alkyl chains at the surface can also affect the frictional properties of self-assembled films if the layers are not packed too closely together (GarciaParajo et al. 1997). Indentation can induce disorder in the chains that then compress as the tip continues to press against them. If the tip presses hard enough, the film hardens as a result of the repulsive forces between the chains. However, if the chains are tilted, they bend or deform when the tip pushes on them in a mostly elastic fashion that produces long lubrication lifetimes. At low contact loads of about 10^{-8} N, wear usually occurs at defect sites, such as steps. Wear can also occur if there are strong adhesive forces between the film and the surface (Overney et al. 1994c).

The friction of model SAMs composed of alkane chains was examined using MD simulations with bond-order and LJ potentials by Mikulski and Harrison (2001a, b). These simulations show that periodicities observed in a number of system quantities are the result of the synchronized motion of the chains when they are in sliding contact with the diamond counterpart. The tight packing of the monolayer and commensurability of the counterpart are both needed

Fig. 7.55 Changes in the conformation of adsorbed hydrocarbon chains on weakly (top) and strongly (bottom) physisorbing surfaces at equilibrium and under shear. Reprinted with permission from Manias et al. (1996). Copyright 1996 American Chemical Society



to achieve synchronized motion when sliding in the direction of chain tilt. The tightly packed monolayer is composed of alkane chains attached to diamond (111) in the (2×2) arrangement and the loosely packed system has approximately 30% fewer chains. The average friction at low loads is similar in both the tightly and loosely packed systems at low loads. Increasing the load, however, causes the tightly packed monolayer to have significantly lower friction than the loosely packed monolayer (see Fig. 7.56). While the movement of chains is somewhat restricted in both systems, the tightly packed monolayer under high loads is more constrained with respect to the movement of individual chains than the loosely packed monolayer, as illustrated in Fig. 7.57. Therefore, sliding initiates larger bond-length fluctuations in the loosely packed system, which ultimately led to more energy dissipation via vibration and, thus, higher friction. Thus, the efficient packing of the chains is responsible for the lower friction observed for tightly packed monolayers under high loads.

Several AFM experiments have examined the friction of SAMs composed of chains of mixed lengths (Barrena et al. 2001). For example, the friction of SAMs composed of

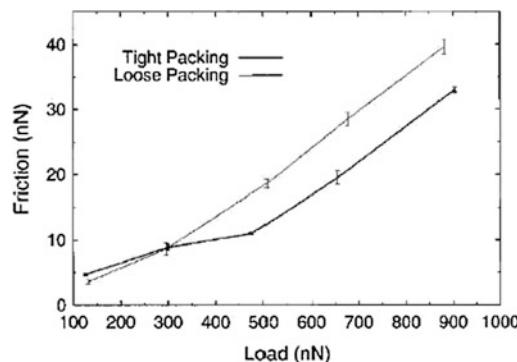


Fig. 7.56 Friction as a function of load when a hydrogen-terminated counterface is in sliding contact with C_{18} alkane monolayers. Reprinted with permission from Mikulski and Harrison (2001b). Copyright 2001 American Chemical Society

spiroalkanedithiols was examined Perry and coworkers (Shon et al. 2000). The effects of crystalline order at the sliding interface were examined by systematically shortening some of the chains. The resulting increase in disorder at the sliding interface causes an increase in friction.

The link between friction and disorder in monolayers composed of n-alkane chains was recently examined using MD simulations by

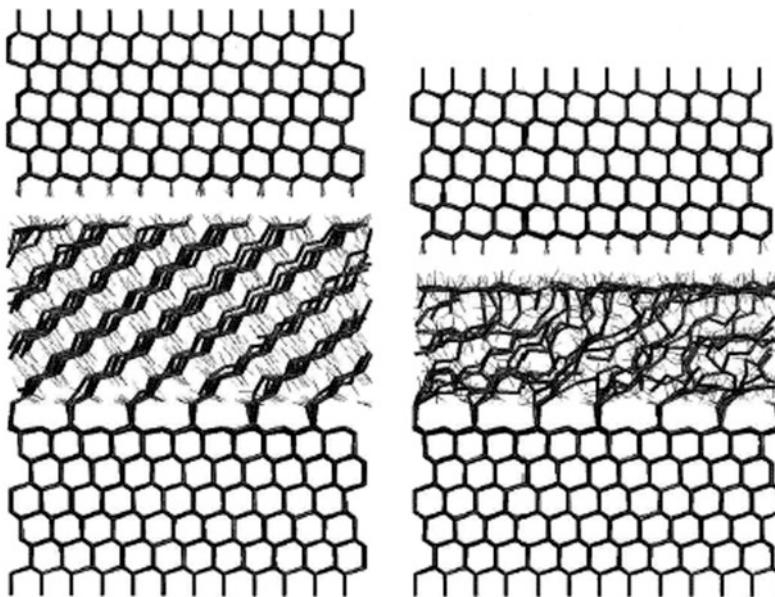


Fig. 7.57 Snapshots of tightly packed C_{18} alkane monolayers on the *left*, and loosely packed monolayers on the *right* under a load of about 500 nN. The chains in both systems are arranged in a (2×2) arrangement on diamond (111). The loosely packed system has 30% of the chains randomly removed. The sliding direction is from left-to-right. Reprinted with permission from Mikulski and Harrison (2001b). Copyright 2001 American Chemical Society

Harrison and coworkers (Mikulski et al. 2005). The tribological behavior of monolayers of 14 carbon atom-containing alkane chains, or pure monolayers, was compared to monolayers that randomly combine equal amounts of 12 and 16 carbon-atom chains, or mixed monolayers. Pure monolayers consistently show lower friction than mixed monolayers when sliding under repulsive (positive) loads in the direction of chain tilt. These MD simulations reproduce trends observed in AFM experiments of mixed-length alkanethiols (Barrena et al. 2001) and spiroalkanedithiols on Au (Lee et al. 2000). Harrison and coworkers (Mikulski et al. 2005) have also examined the “odd-even” effect noted in experiment (Wong et al. 1998), where friction is found to be larger for SAMs differing by one methylene group. The MD simulations demonstrated that the effect was due to conformational differences in the chains of different length and became more pronounced at higher loads.

Because the force on individual atoms is known as a function of time in MD simulations, it is possible to calculate the contact forces

between individual monolayer chain groups and the tip, where contact force is defined as the force between the tip and a $-CH_3$ or a $-CH_2$ -group in the alkane chains. The distribution of contact forces between individual monolayer chain groups and the tip are shown in Fig. 7.58. It is clear from these contact force data that the magnitude, or scale of the forces, is similar in both the pure and the mixed monolayers. In addition, it is also apparent that the pure and mixed monolayers resist tip motion in the same way. That is, the shape of the histograms in the positive force intervals is similar. In contrast, the contact forces “pushing” the tip along differ in the two monolayers. The pure monolayers exhibit a high level of symmetry between resisting and pushing forces. Because the net friction is the sum of the resisting and pushing forces, the symmetry in these distributions of the pure monolayers results in a lower net friction than the mixed monolayers. Thus, the ordered, densely packed nature of the pure monolayers allows the energy stored when the monolayer is resisting tip motion (positive forces) to be regained efficiently

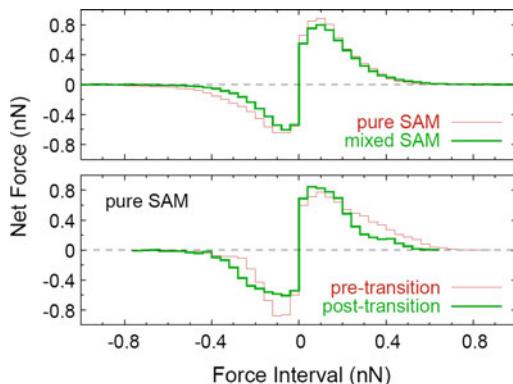


Fig. 7.58 The distribution of contact forces along the sliding direction (friction force). In the upper panel, the forces for the mixed and pure system sliding in the direction of chain tilt are shown. The forces for the pure system sliding in the transverse direction to the chain tilt are shown in the lower panel. Positive force intervals correspond to chain groups that resist tip motion while negative intervals correspond to chain groups that “push” the tip in the sliding direction. Forces from four runs with independent starting configurations are binned for all sets of data

when the monolayer “pushes” on the tip (negative forces). The distribution of negative contact forces in the mixed monolayers is different from the distribution of the positive forces. For this reason, mechanical energy is not efficiently channeled back into the mixed monolayer as the tip passes over the chains and, as a result, the friction is higher. The range of motion of the chains is monitored by computing the deviation in a chain group’s position compared to its starting position, as illustrated in Fig. 7.59. It is clear from analysis of these data that the increased range of motion is linked to large contact forces. The increased range of motion of the protruding tails in the mixed system prevents the efficient recovery of energy during sliding (negative contact force distribution) and allows for the dissipation of energy.

The pure monolayers exhibit marked friction anisotropy. The contact force distribution changes dramatically as a result of the change in sliding direction, resulting in an increase in friction (see Fig. 7.58). Sliding in the direction perpendicular to chain tilt can cause both types of monolayers to transition to a state where the

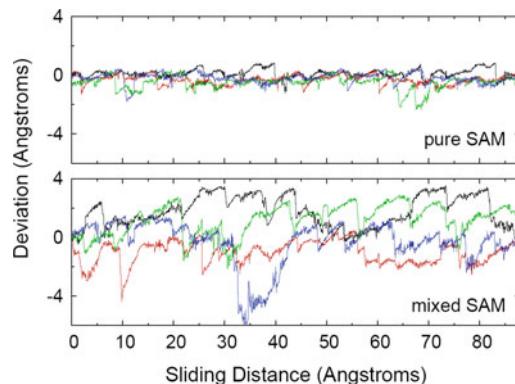


Fig. 7.59 Trajectories of individual chain groups that generate the largest contact forces when sliding in the direction of chain tilt for both the pure and mixed monolayer systems. The deviation is defined as the change in position along the sliding direction relative to the chain group’s starting position (The positions are averaged over 2000 simulation steps)

chains are primarily tilted along the sliding direction. This transition is accompanied by a large change in the distribution of contact forces and a reduction in friction.

Recently, the response of monolayers composed of alkyne chains, which contain diacetylene moieties, to compression and shear (Chateauneuf et al. 2004) was examined using MD simulations. These are the only simulations to date that show that compression and shear can result in cross-linking, or polymerization, between chains. The vertical positioning of the diacetylene moieties within the alkyne chains (spacer length) and the sliding direction both have an influence on the pattern of cross-linking and friction. Compression and shear cause irregular polymerization patterns to be formed among the carbon backbones, as illustrated in Fig. 7.60. When diacetylene moieties are located at the ends of the chains closest to the tip, chemical reactions between the chains of the monolayer and the amorphous carbon tip occur causing the friction to increase 100 times, as indicated in Fig. 7.61. The friction between the amorphous carbon tip and all the diacetylene-containing chains is larger than the friction between a hydrogen-terminated diamond counterface and tightly packed monolayers

Fig. 7.60 **a** Perpendicular-, **b** tilted-, and **c** end-chain monolayer systems after compression to 200 nN and pull-back of the hydrogen-terminated tip. Large, dark spheres in the hydrocarbon monolayers represent cross-linked atoms with sp^2 hybridization. Dark, small spheres represent hydrogen atoms that are initially on the hydrogen-terminated amorphous carbon tip. Reprinted with permission from Chateauneuf et al. (2004). Copyright 2004 American Chemical Society

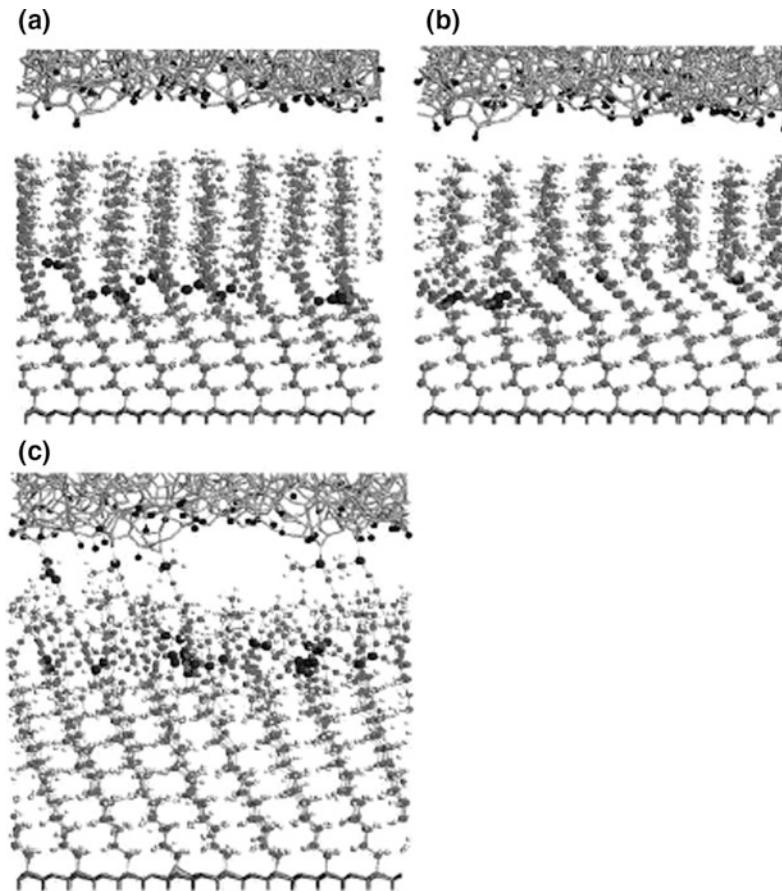
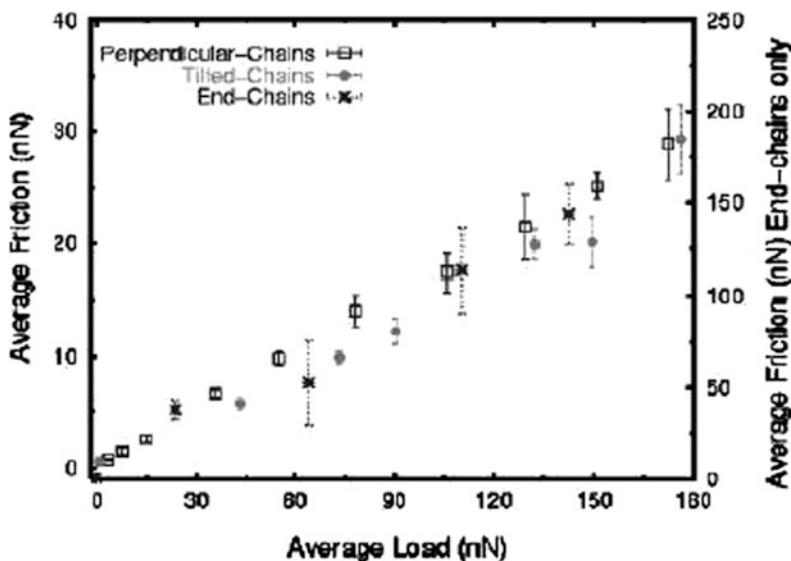


Fig. 7.61 Average friction on the tip as a function of load for the monolayer systems shown in Fig. 7.60. The scale for the average friction in the end-chain system is shown on the right-hand side of the figure. Reprinted with permission from Chateauneuf et al. (2004). Copyright 2004 American Chemical Society



composed of n-alkane chains (Mikulski and Harrison 2001b). This is attributed to the disorder at the interface caused by the irregular counterface.

Zhang and Jiang (2002) used MD simulations to study the effect of confined water between alkyl monolayers terminated with $-\text{CH}_3$ (hydrophobic) and $-\text{OH}$ (hydrophilic) groups on Si (111), as illustrated in Fig. 7.62. For the hydrophobic molecules, the friction coefficient is almost constant independent of the number of water molecules. For the hydrophilic molecules, the friction coefficient decreases rapidly with the increase of the number of water molecules as shown in Fig. 7.63. These results are in good agreement with surface force microscopy (SFM) experimental results. Zhang et al. (2003b) also studied the friction of alkanethiol SAMs on gold using hybrid molecular simulations at the same times scales as are used in AFM and FFM experiments. Various quantities were varied in the simulations, including chain length, terminal group, scan direction, and scan velocity. The simulations showed that the frictional force decreases as the chain lengths increase and is smallest when scanned along the tilt direction. They also predicted a maximum friction coefficient for hydrophobic $-\text{CH}_3$ -terminated SAMs and low friction coefficients for hydrophilic $-\text{OH}$ -terminated SAMs as the scan velocity increases. The simulations further predicted a saturated constant value at high scan rates for both surfaces. These results are summarized in Figs. 7.64 and 7.65.

The work of Chandross (2004) and Chandross et al. (2002) illustrates the effects of chain length on friction and stick-slip behavior between two ordered SAMs consisting of alkylsilane chains over a range of shear rates at various separation distances or pressure, as illustrated in Fig. 7.66. The adhesion forces between the two SAMs at the same separation distance decreases as the chain length increases from 6 to 18 carbon atoms. However, the friction forces are independent of the chain length and the shear velocity. The system size is shown to have an effect on the sharpness of the slip transitions but not on the dynamical events as shown in

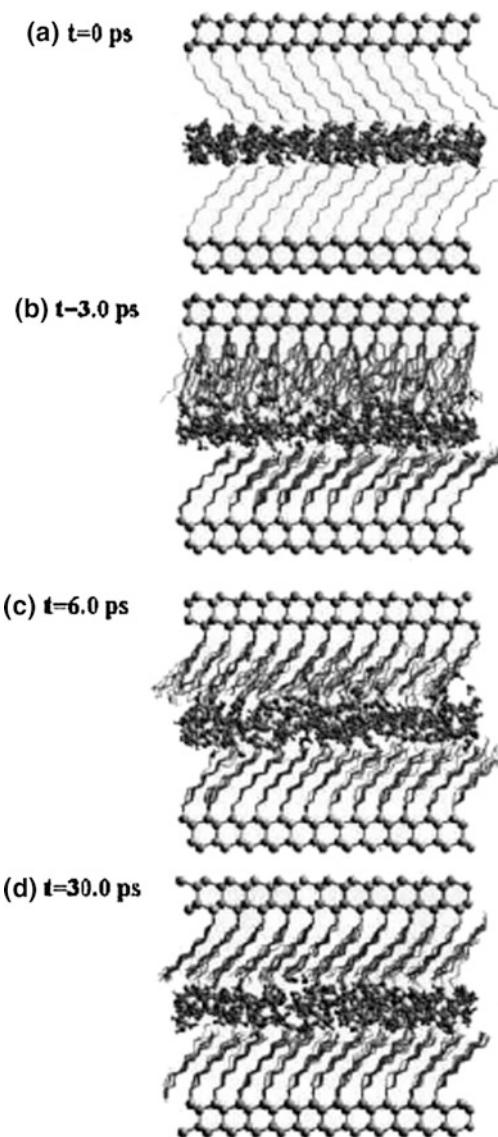


Fig. 7.62 Snapshots of hydrophilic monolayers and confined water molecules from MD simulations at 300 K. The tilt direction of monolayers on the top plate changed after $t = 10.0$ ps. Reprinted with permission from Zhang and Jiang (2002). Copyright 2005 American Institute of Physics

Fig. 7.67. In a later paper, Chandross et al. used SiO_2 tips with radius of curvatures in the range of 3–30 nm to interact with fully physisorbed, chemisorbed, and mix bonded alkylsilane SAMs on amorphous Si (Chandross et al. 2008). This tip based geometry allows for the exploration of actual contact area as a function of load, which

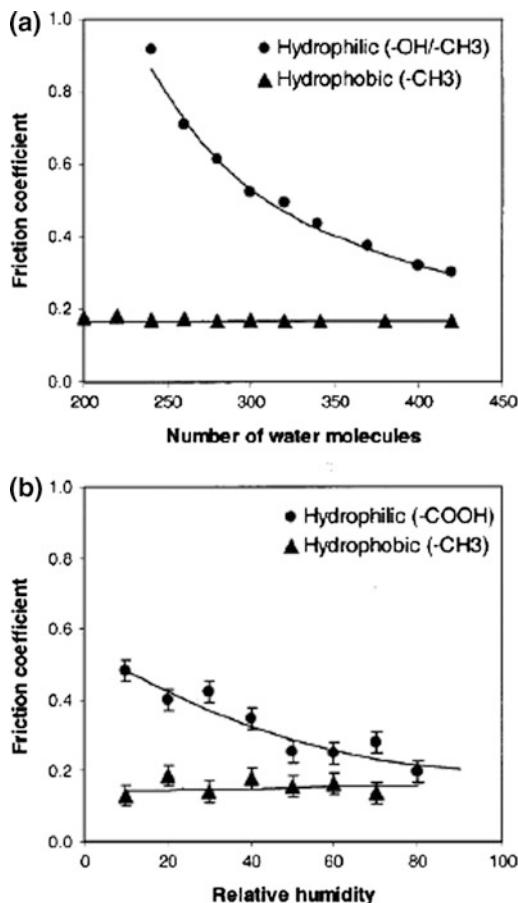


Fig. 7.63 **a** Friction coefficients for hydrophobic ($-\text{CH}_3$) and hydrophilic (50% mixed $-\text{CH}_3/-\text{OH}$) monolayers as a function of water molecules from MD simulations at 300 K ($H = 6.0 \text{ \AA}$), and **b** scanning force microscopy measurements of frictional forces of difference surfaces under various relative humidities. Reprinted with permission from Zhang and Jiang (2002). Copyright 2005 American Institute of Physics

was found to be inversely proportional to the inverse square root of the load.

SAMs have been very successful in the lubrication of surfaces that infrequently come into sliding or normal contact. Their inability to lubricate a reciprocating contact is due, in part, to the inability to replenish the coating in situ. One proposed solution is to use a chemically bound SAM in conjunction with a physisorbed mobile molecule in a “bound + mobile” lubricant scheme, which is similar to the lubricants used to mitigate head crashes in hard drives.

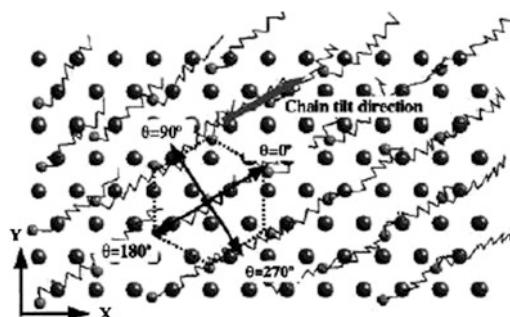


Fig. 7.64 Schematic illustration of the chain tilt and scan directions on alkanethiol SAMs/Au(111) in hybrid molecular simulations. θ is the angle between the tip moving direction and the chain tilt direction. The larger spheres represent substrate Au atoms, smaller spheres sulfur atoms in molecular chains, and zigzag lines molecular chains. Reprinted with permission from Zhang et al. (2003a). Copyright 2003 American Chemical Society

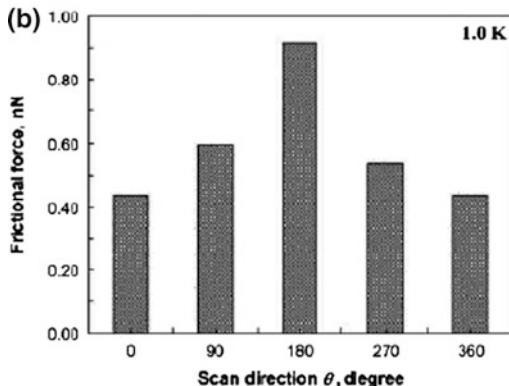
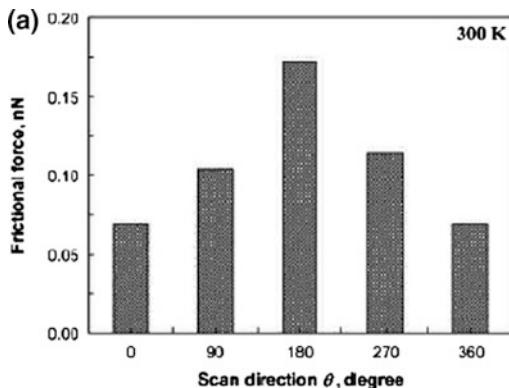


Fig. 7.65 Frictional force as a function of scan direction from hybrid simulations for C_{11}CH_3 SAMs on Au(111) at **a** 300 K and **b** 1.0 K. Frictional force is the smallest when scanned along the tilt direction, the largest when scanned against the tilt direction, and between when scanned perpendicular to the tilt direction at both temperatures. Reprinted with permission from Zhang et al. (2003a). Copyright 2003 American Chemical Society

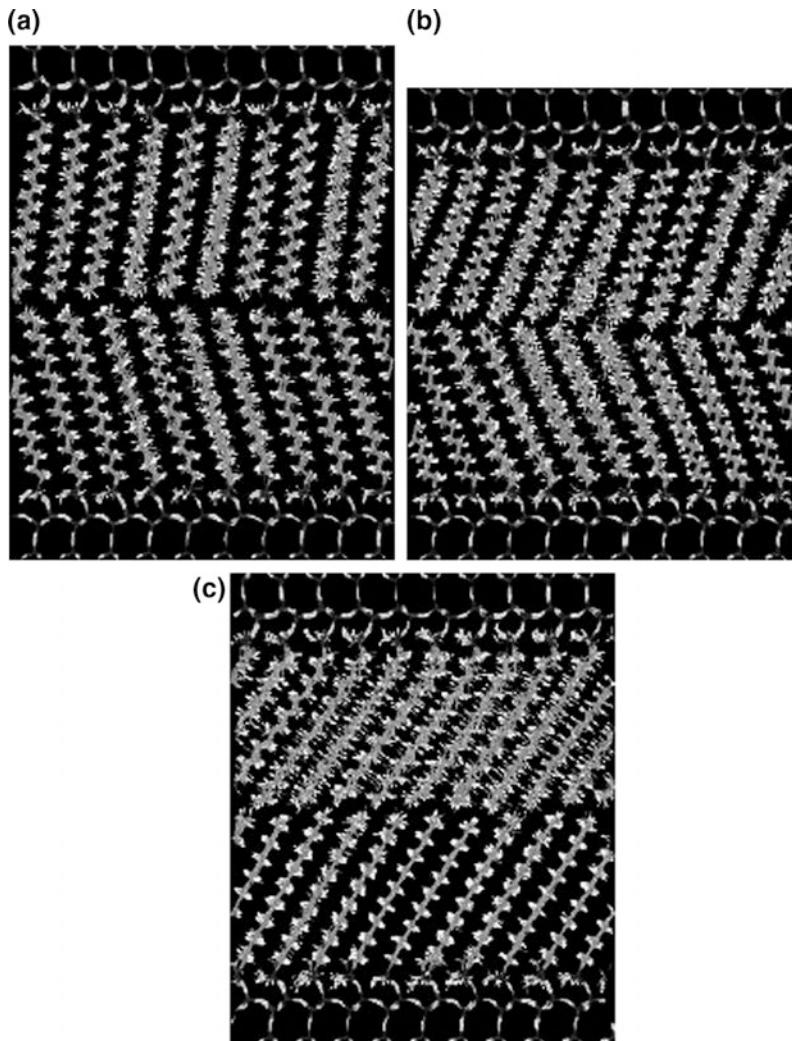


Fig. 7.66 Wireframe images of $n = 18$ SAMs at fixed separations of **a** $d = -5.2 \text{ \AA}$ (low pressure, under compression only) **b** $d = -10.2 \text{ \AA}$ (high pressure, under compression only) and **c** $d = -10.2 \text{ \AA}$ (high pressure, under shear). Reprinted with permission from Chandross et al. (2002). Copyright 2002 American Chemical Society

Irving and Brenner used molecular dynamics to study the interfacial structure, self-diffusion, and ability of the mobile phase to incorporate into defected sites (Irving and Brenner 2006). A potential bound + mobile lubricant combination of a chemically bound octadecyltrichlorosilane (ODTS) SAM together with mobile tricresyl phosphate (TCP) molecules was studied. The simulations showed that the TCP did not incorporate into the interior of the close packed ODTS SAM. The TCP molecules on the surface of the SAM were also not tightly bound to

particular surface site but instead were found to readily diffuse across the surface in a random walk fashion. An estimated diffusion barrier of 0.0937 eV with an Arrhenius prefactor of $26.47 \times 10^{-4} \text{ cm}^2/\text{s}$ was calculated for single-molecule diffusion. It was also found that the TCP molecules would localize in the vicinity of methylene groups ($-\text{CH}_2-$) along the backbone of the ODTS chain, which were exposed when cylindrical defect was created in the SAM. However, to get to this lower energy state meant that the molecules needed to overcome an anisotropic

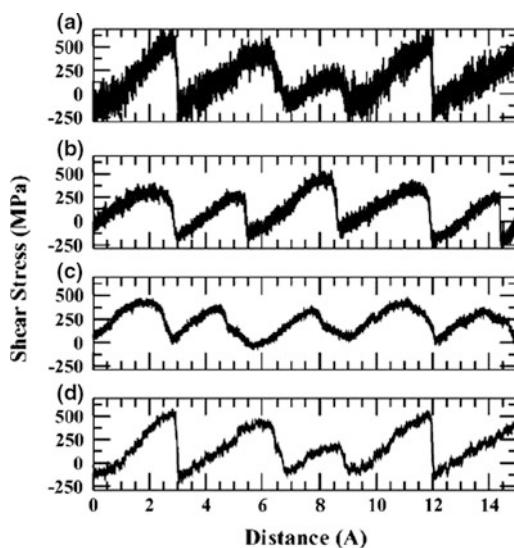


Fig. 7.67 Shear stress σ_s as a function of system size for $n = 6$ SAMs corresponding to a pressure of 200 MPa at $v = 1.0$ m/s: **a** 100 chains per surface, **b** 400 chains per surface, **c** 1600 chains per surface, and **d** 16 point box average of system with 100 chains per surface. Reprinted with permission from Chandross et al. (2002). Copyright 2002 American Chemical Society

energy barrier for inclusion in the defect that was related to the direction of molecular approach and ODTs chain tilt. A later study used the diffusion information and multiscale methods to analyze the conditions under which this scheme would be successful (Brenner et al. 2007).

Polymer thin films are also a widely studied for their lubricating properties. An example is polytetrafluoroethylene (PTFE), which has been used in a wide range of applications from satellites to frying pans. In a joint computational and experimental work Jang et al. examined the molecular origins of friction using classical molecular dynamics as well as an AFM and microtribometer (Jang et al. 2007). The simulations predicted an anisotropic behavior of the friction coefficient depending on whether the sliding direction was parallel or perpendicular to the PTFE chains lying on the surface. Sliding directions parallel produced lower friction coefficients and wear while sliding perpendicular to alignment produced higher coefficients and wear. The microtribometer results were in agreement with these findings. Also of interest was transfer

films in the AFM work always produced transfer films parallel to sliding direction.

Similar experimental findings for anisotropic tribological behavior of polyethylene (PE) in the literature motivated Heo et al. to examine this system using classical molecular dynamics to study crystalline PE interfaces (Heo et al. 2008). Like the findings for PTFE it was found that friction and wear had an anisotropic behavior that depended on molecular orientation and sliding direction. Unlike the findings for PTFE, the PE system exhibited a stick-slip motion as the interfaces passed by one another. The reason for the differences between the two systems was attributed to increases bond scission seen in PTFE as compared to PE under sliding conditions. This scission allows collections of molecules in PTFE to move at the interface, which does not occur as readily in the PE system.

In short, atomic-scale simulations show the relationship among elastic properties, degree of molecular disorder, and friction of self-assembled thin films that illuminates the origin of the properties that are measured experimentally.

7.4.3.3 Nanoparticles

Nanoparticles are being considered for a wide variety of applications, including as fillers for nanocomposite materials, novel catalysts or catalytic supports, and components for nanometer-scale electronic devices (Sinnott and Andrews 2001). They have also generated considerable interest as possible new lubricant materials that have the potential to function as “nano-ball bearings” with exceptionally low friction coefficients. Examples of nanoparticles of most interest for tribological applications are C₆₀ (Bhushan et al. 1993a; Thundat et al. 1993; Mate 1993; Lüthi et al. 1993, 1994b; Xue et al. 1994; Allers et al. 1995; Schwarz et al. 1995a; Ruan and Bhushan 1996; Schwarz et al. 1997; Okita et al. 1999; Okita and Miura 2001; Miura et al. 2003), carbon nanotubes (Buldum and Lu 1999; Falvo et al. 1999, 2000; Schall and Brenner 2000; Ni and Sinnott 2001a, b; Miura et al. 2001a; b), and MoO₃ nanoparticles (Sheehan and Lieber 1996; Wang et al. 1999), among others (Ouyang and Okada 1994).

The experiments report wide variations in frictional coefficients (for instance, values of 0.06–0.9 have been measured for C₆₀) that may be caused by differences in the experimental methods used, the thickness of the nanoparticle layer or island, atmosphere (argon versus air, levels of humidity), and the transfer of nanoparticles to the FFM tips. As a result, there is much that remains to be clarified about the tribological behavior of nanoparticle films.

In the case of C₆₀, the mechanistic response to applied shear forces has not been definitively determined. For example, some experimental studies show evidence of C₆₀ molecules rolling against the substrate, each other, or the sliding surfaces (Bhushan et al. 1993a; Xue et al. 1994; Schwarz et al. 1995a; Okita et al. 1999; Miura et al. 2003) while others hypothesize that the low friction of C₆₀ films is due in part to blunting of the tip by transfer of fullerene molecules to the tip apex. Fullerene films are found experimentally to have dissipation energies and shear strengths that are a full order of magnitude lower than the values that are typical for boundary lubricants (Lüthi et al. 1994a). Experimental testing of the frictional properties of fullerenes reveal low mechanical stability accompanied by progressive wear and transfer of fullerene materials when they are only physisorbed on a solid surface (Bhushan et al. 1993b). Furthermore, measurements with a FFM show that under certain

conditions, adsorbed fullerene films deteriorate at pressures as low as about 0.1 GPa (Schwarz et al. 1995b). The challenge is therefore to obtain mechanically stable, ordered molecular films of fullerenes firmly attached to a solid substrate.

There have been several MD simulation studies to investigate the tribological properties of fullerenes. A representative study by Legoas et al. (2004) investigated the experimentally observed low-friction system of C₆₀ molecules positioned on highly oriented pyrolytic graphite. The results show that decreasing the van der Waals interaction between a C₆₀ monolayer and graphite sheets, and the characteristic movements of graphite flakes over C₆₀ monolayers, explains the measured ultra-low friction of C₆₀ molecules and graphite sheets.

Several MD simulation studies have also been carried out on the tribological properties of carbon nanotubes. For example, simulations by Buldum and Lu (1999) and Schall and Brenner (2000) indicate that single-walled carbon nanotubes roll when their honeycomb lattice is “in registry” with the honeycomb lattice of the graphite. If this registry is not present, the carbon nanotubes respond to applied forces from an AFM by sliding. This behavior is nicely summarized in Fig. 7.68. These MD simulation findings were simultaneously confirmed in experimental studies by Falvo et al. (2000). Experimental studies of multi-walled carbon

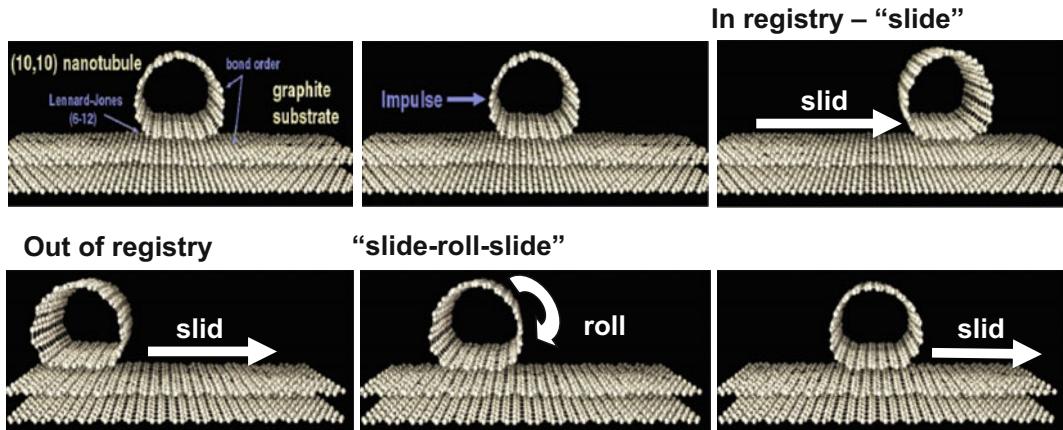


Fig. 7.68 Dynamics of a nanotube on a graphite surface. When the nanotube and graphite plane are out of registry, the nanotube slides as it slows down from an initial impulse (*upper right panel*). When the nanotube is oriented such that it is in registry with the graphite, it slows by a combination of rolling and sliding

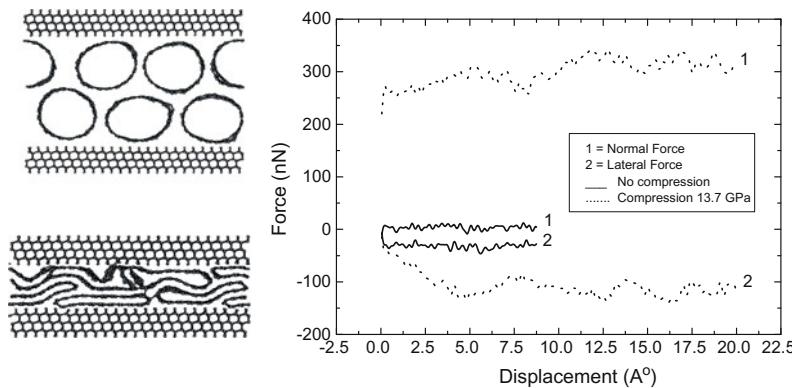


Fig. 7.69 *Left* Snapshots from simulations that examine the sliding of the potential of the topmost diamond surface on horizontally arranged nanotubes at different compressions; **a** is at a pressure of ~ 0 GPa; **b** is with a pressure of 13.7 GPa. *Right* Plots of the normal and lateral components of force during sliding of the top diamonds surface on horizontally arranged nanotubes as a function of the displacement of the top diamond surface with respect to the diamond surface on the bottom

nanotubes on graphite (Miura et al. 2001a) show similar evidence of nanotube rolling when the outer tube is pushed.

The tribological properties of nanotube bundles are important, as it is well-known that carbon nanotubes agglomerate together very readily to form bundles and are often grown in bundle form (Sinnott and Andrews 2001). An experimental study by Miura et al. (2001b) of carbon nanotube bundles being pushed around on a KCl surface with an AFM tip indicates that bundles of single-walled carbon nanotubes can be induced to roll in a manner that is similar to the rolling observed for multi-walled nanotubes.

MD simulations by Ni and Sinnott (2001a, b) considered the responses of horizontally and vertically aligned single-walled carbon nanotubes between two hydrogen-terminated diamond surfaces, where the top surface is slid relative to the bottom surface. The movement of the carbon nanotubes in response to the shear forces was predicted to be simple sliding for both orientations. Interestingly, the simulations do not predict rolling of the horizontally arranged carbon nanotubes even when they are aligned with each other in two-layer and three-layer structures. Instead, at low compressive forces, illustrated in Fig. 7.69, the nanotube bundles slide as a single unit, and at high compressive forces, also

illustrated in Fig. 7.69, the deformed carbon nanotubes closest to the topmost moving diamond surface start to slide in a motion reminiscent of the movement of a tank or bulldozer wheel belt. However, when these moving carbon nanotube atoms would have turned the first corner at the top of the ellipse, they encounter the neighboring nanotube and cannot slide past it. This causes them to deform even further, form cross-links with one another, and, in some cases, move in the reverse direction to the sliding motion of the diamond surface. This causes the large oscillations in the normal and lateral forces plotted in Fig. 7.69.

A later study by Heo et al. examined the frictional properties of single wall, double wall, and filled carbon nanotubes contained between two hydrogen terminated diamond like carbon layers (Heo and Sinnott 2007). It was shown that over a wide range of loads the simulations predicted a friction coefficient of 0.13 for filled as well as unfilled nanotubes. This friction coefficient was found to be constant for pressures below 5 GPa, which should include most practical applications. At pressures above 5 GPa the friction coefficient was found to increase to 0.2. Unlike the work of Ni et al. it was demonstrated that the nanotubes would roll in response to the applied shear. This difference was attributed to

the finite size of the nanotubes used in this simulation as compared to the infinite tubes used by Ni et al. Although the frictional properties were similar for all systems considered, it was found that filled nanotubes, whether they are double-walled nanotubes or nano-peapods, were more adept at sustaining higher load than the unfilled single-walled nanotubes. Surprisingly, it was also found that the addition of lubricating benzene molecules did little to alter the friction in the system. Rather, it was found that the addition of benzene altered the mechanism by which the system responded to the applied shear stress.

The responses of the horizontally arranged carbon nanotubes are substantially different from the responses of the vertically arranged nanotubes at high compression, as can be seen by comparing Figs. 7.69 and 7.70. The vertical, capped carbon nanotubes are quite flexible and bend and buckle in response to applied forces. As the buckle is forming, the normal force decreases than stabilizes in the buckled structure as illustrated in Fig. 7.70. As the topmost diamond surface slides, the buckled nanotubes swing around the buckle “neck” which helps dissipate the applied stresses. For this reason, the magnitudes of the lateral forces are not significantly different for the vertical nanotubes at low and high compression, as indicated in Fig. 7.70.

When the ratio of the frictional (lateral) force to the normal force is taken to calculate friction coefficients for these systems, high, non-intuitive values were obtained. As outlined by Ni and Sinnott (2001a), this is because the actual contact area of the nanotubes are not proportional to the sliding force. In the case of the horizontal nanotube bundles, the tubes are able to deform and significantly change their contact area with the sliding surface with minimal change in the normal force, as shown in Fig. 7.69. In the case of the vertical nanotubes, the contact area remains approximately the same regardless of the initial loading force because of the flexibility of the nanotubes. This causes the lateral forces to change only slightly with significant changes in the normal force, as shown in Fig. 7.70. This analysis indicates that care must be taken in

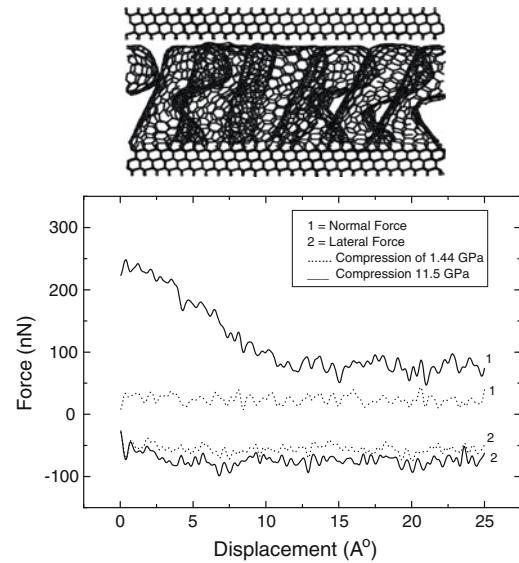


Fig. 7.70 *Top* Snapshots from simulations that examine the sliding of the topmost diamond surface on vertically arranged nanotubes with one set of capped ends compressed at a pressure of 11.5 GPa. *Bottom* Plots of the normal and lateral components of force during sliding of the top diamonds surface on vertically arranged nanotubes as a function of the displacement of the top diamond surface with respect to the diamond surface on the bottom

calculating friction coefficients for nanotubes systems. Recent experiments by Dickrell et al. (2005) show good agreement with these predictions, as shown in Fig. 7.71.

To summarize, this subsection shows that nanoparticles show some promise as lubricating materials due to their exceptionally low friction coefficients in experiments and simulations. Some nanoparticles show lattice-directed sliding on substrates due to their unique atomic structures. However, there is much that remains to be done before the nanometer-scale friction of these materials are well understood.

7.4.3.4 Solid State

Surfaces are able to slide over each other at high loads with a minimum of resistance from friction in the presence of liquid lubricants. Some solid thin films can also fulfill these functions and, when they do, are termed solid lubricants. Solid lubricants are generally defined as having friction coefficients of 0.3 or less and low wear.

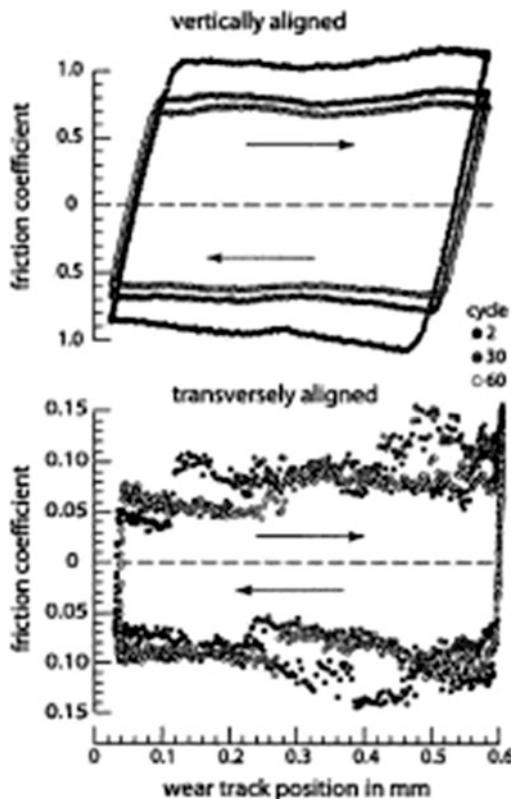


Fig. 7.71 Coefficient of friction data versus track position collected for one full cycle of reciprocating sliding for nanotubes that are vertically and transversely aligned. Reprinted with permission from Dickrell et al. (2005). Copyright 2005 Springer Science + Business Media B.V.

Bowden and Tabor showed how thin solid films can reduce friction as follows (Bowden and Tabor 1964). The total friction force F_f is given as

$$F_f = A * F_S + F_p \quad (7.11)$$

where F_p is the plowing term, A is the area of contact and F_S is the shear strength of the interface. If the surfaces are soft, F_S will be reduced while the other parameters will increase. However, if the surfaces under the solid film are very stiff, A and F_p will decrease thereby decreasing friction. The properties specific to the film will also have an effect on friction. For instance, if the films are less than 1 μm thick, the surface asperities will be able to break through the film to eventually cause wear between the surfaces under normal circumstances. On the other hand,

if the lubricant film is too thick, there will be increased plowing and wear that causes the frictional forces to increase. It is important that the lubricant not delaminate in response to the frictional forces, so strong bonds between the lubricant and the surface are required for a solid state lubricant to be effective.

The most common materials used as solid lubricants have layered structures such as graphite or MoS_2 , that, as discussed above, experience low friction. It is not necessary for the lubricant film to have a layered structure to give low friction. For example, diamond-like carbon has some of the lowest coefficients of friction measured and yet does not have a layered structure. Similarly, not all layered structures are lubricants. For instance, mica gives a relatively high coefficient of friction (>1).

The atomic-scale tribological behavior when a hydrogen-terminated diamond (111) counterface is in sliding contact with amorphous, hydrogen-free, DLC films was examined using MD simulations by Gao et al. (2002). Two films, with approximately the same ratio of sp^3 -to- sp^2 carbon but different thicknesses, were examined. Similar average friction is obtained from both films in the load range examined. A series of tribochemical reactions occur above a critical load that result in a significant restructuring of the film, which is analogous to the “run-in” observed in macroscopic friction experiments, and reduces the friction. The contribution of adhesion between the counterface and the sample to friction is examined by varying the saturation of the counterface. The friction increases when the degree of saturation of the diamond counterface is reduced by randomly removing hydrogen atoms. Lastly, two potential energy functions that differ only in their long-range forces are used to examine the contribution of long-range interactions to friction in the same system (as illustrated in Figs. 7.72 and 7.73).

MD simulations were also recently used by Gao et al. (2003) to examine the effects of the sp^2 -to- sp^3 carbon ratio and surface hydrogen on the mechanical and tribological properties of amorphous carbon films. This work showed that, in addition to the sp^2 -to- sp^3 ratio of carbon, the

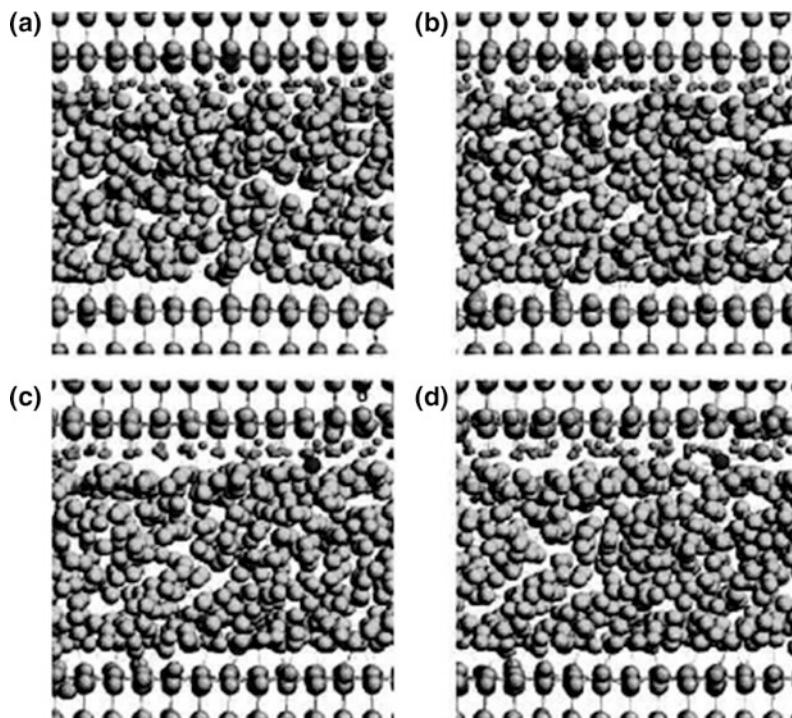


Fig. 7.72 A series of chemical reactions induced by sliding of the counterface over the thin film under an average load of 300 nN. **a** The sliding causes the rupture of a carbon-hydrogen bond in the counterface. **b** The hydrogen atom is incorporated into the film and forms a bond to a carbon atom in the film. **c** A bond is formed between the unsaturated carbon atoms in the film and the carbon that suffered the bond rupture in the counterface, and continued sliding causes this carbon being transferred into the film. **d** The transferred carbon forms a bond with another carbon in the counterface. The counterface has slid 0.0 (**a**), 15.9 (**b**), 26.1 (**c**), and 30.5 Å (**d**). Reprinted with permission from Gao et al. (2002). Copyright 2002 American Chemical Society

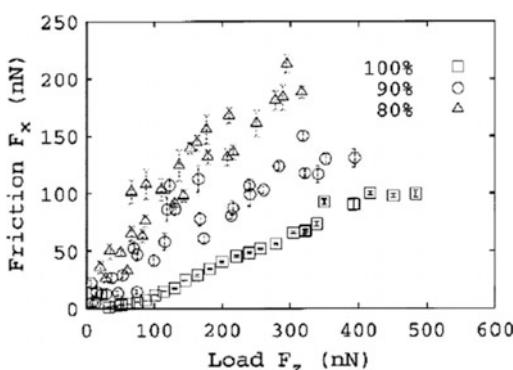


Fig. 7.73 Friction curves for the thin-film system with a counterface that is 100% hydrogen terminated (open squares), 90% hydrogen terminated (filled squares), and 80% hydrogen terminated (open circles). Reprinted with permission from Gao et al. (2002). Copyright 2002 American Chemical Society

three-dimensional structure of the films is important when determining the mechanical properties of the films. For example, it is possible to have high sp^2 -carbon content, which is normally associated with softer films, and large elastic constants. This occurs when sp^2 -ringlike structures are oriented perpendicular to the compression direction. The layered nature of the amorphous films examined leads to novel mechanical behavior that influences the shape of the friction versus load data, as illustrated in Fig. 7.74. When load is applied to the films, the film layer closest to the interface is compressed. This results in the very low friction of films I and II up to approximately 300 nN and the response of films IV and V up to 100 nN. Once the outer film layers have been compressed, additional

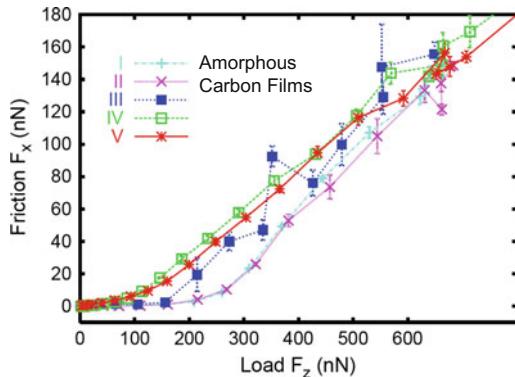


Fig. 7.74 Average friction versus load for five amorphous carbon films. Films I–III are hydrogen free and contain various ratios of sp^2 -to- sp^3 carbon. Films IV and V are both over 90% sp^3 carbon and have surface hydrogenation

application of load causes an almost linear increase in friction for both Films I & II and IV & V. Film III has an erratic friction versus load response due to the early onset of tribochemical reactions between the tip and the film.

7.5 Conclusions

This chapter provides a wide-ranging discussion of the background of MD and related simulation methods, their role in the study of nanometer-scale indentation and friction, and their contributions to these fields. Specific, illustrative examples are presented that show how these approaches are providing new and exciting insights into mechanisms responsible for nanoindentation, atomic-scale friction, wear, and related atomic-scale and molecular scale processes. The examples also illustrate how the results from MD and related simulations are complementary to experimental studies, serve to guide experimental work, and assist in the interpretation of experimental data. The ability of these simulations and experimental techniques such as the surface force apparatus and proximal probe microscopes to study nanometer-scale indentation and friction at approximately the same scale is revolutionizing our understanding of the origin of friction at its most fundamental atomic level.

Acknowledgements S.B.S. and S.-J.H. acknowledge support from the Air Force through grant FA9550-04-1-0367 and from the National Science Foundation supported Network for Computational Nanotechnology (ECC-0228390). D.W.B. and D.L.I. acknowledge support from the Office of Naval Research through grant N00014-04-2006, the National Science Foundation through grant DMR-0304299, the Army Research Office, and the Air Force Office of Scientific Research. JAH acknowledges support from the Air Force Office of Scientific Research under contracts F1ATA08018G001 (Extreme Friction MURI) and F1ATA07351G001. JAH also acknowledges support from The Office of Naval Research under contract N0001408WR20106.

References

- Abell GC (1985) Empirical chemical pseudopotential theory of molecular and metallic bonding. *Phys Rev B* 31(10):6184–6196
- Ackland GJ, Tichy G, Vitek V, Finnis MW (1987) Simple N-body potentials for the noble-metals and nickel. *Philos Mag A Phys Condens Matter Struct Defects Mech Prop* 56(6):735–756
- Adams DJ (1975) Grand canonical ensemble Monte-Carlo for a Lennard-Jones fluid. *Mol Phys* 29(1):307–311
- Adelman SA (1980) Generalized Langevin equations and many-body problems in chemical dynamics. *Adv Chem Phys* 44:143–253
- Adelman SA, Doll JD (1976) Generalized Langevin equation approach for atom-solid-surface scattering—general formulation for classical scattering off harmonic solids. *J Chem Phys* 64(6):2375–2388
- Agrait N, Rubio G, Vieira S (1996) Plastic deformation in nanometer scale contacts. *Langmuir* 12(19):4505–4509
- Akamine S, Barrett RC, Quate CF (1990) Improved atomic force microscope images using microcantilevers with sharp tips. *Appl Phys Lett* 57(3):316–318
- Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Clarendon Press, Oxford
- Allers W, Schwarz UD, Gensterblum G, Wiesendanger R (1995) Low-load friction behavior of epitaxial C_{60} monolayers. *Z Phys B* 99:1–2
- Barrena E, Ocal C, Salmeron M (2001) A comparative AFM study of the structural and frictional properties of mixed and single component films of alkanethiols on Au(111). *Surf Sci* 482:1216–1221
- Baskes MI (1992) Modified embedded-atom potentials for cubic materials and impurities. *Phys Rev B* 46 (5):2727–2742
- Baskes MI, Nelson JS, Wright AF (1989a) Semiempirical modified embedded atom potentials for silicon and germanium. *Phys Rev B* 40(9):6085–6100
- Baskes MI, Nelson JS, Wright AF (1989b) Semiempirical modified embedded-atom potentials for silicon and germanium. *Phys Rev B* 40(9):6085–6100

- Belak J, Stowers IF (1990) Proceedings of the American society for precision engineering, p 76
- Berman AD, Ducker WA, Israelachvili JN (1996) Origin and characterization of different stick-slip friction mechanisms. *Langmuir* 12(19):4559–4563
- Bhushan B, Gupta BK, Van Cleef GW, Capp C, Coe JV (1993a) Sublimed C₆₀ films for tribology. *Appl Phys Lett* 62(25):3253–3255
- Bhushan B, Gupta BK, Vancleef GW, Capp C, Coe JV (1993b) Fullerene (C-60) films for solid lubrication. *Tribol Trans* 36(4):573–580
- Bhushan B, Israelachvili JN, Landman U (1995) Nanotribology—friction, wear and lubrication at the atomic-scale. *Nature* 374(6523):607–616
- Biersack JP, Ziegler J, Littmack U (1985) The stopping and range of ions in solids. Pergamon, Oxford
- Binnig G, Quate CF, Gerber C (1986) Atomic force microscope. *Phys Rev Lett* 56(9):930–933
- Bolshakov A, Oliver WC, Pharr GM (1996) Influences of stress on the measurement of mechanical properties using nanoindentation. 2. Finite element simulations. *J Mater Res* 11(3):760–768
- Bowden FP, Tabor D (1964) The friction and lubrication of solids, part 2. Clarendon Press, Oxford
- Brenner DW (1989a) Tersoff-type potentials for carbon, hydrogen, and oxygen. *Mater Res Soc Symp Proc* 141:59–65
- Brenner DW (1989b) Relationship between the embedded-atom method and tersoff potentials. *Phys Rev Lett* 63(9):1022
- Brenner DW (1990a) Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. *Phys Rev B* 42:9458–9471
- Brenner DW (1990b) Empirical potential for hydrocarbons for use in simulating the chemical vapor-deposition of diamond films. *Phys Rev B* 42(15):9458–9471
- Brenner DW (2000) The art and science of an analytic potential. *Phys Status Solidi B* 217:23–40
- Brenner DW, Shenderova OA, Harrison JA, Stuart SJ, Ni B, Sinnott SB (2002) Second generation reactive empirical bond order (REBO) potential energy expression for hydrocarbons. *J Phys Condens Matter* 14:783–802
- Brenner DW, Irving DL, Kingon AI, Krim J (2007) Multiscale analysis of liquid lubrication trends from industrial machines to micro-electrical-mechanical systems. *Langmuir* 23(18):9253–9257
- Bruckman MJ, Nemanich GGRJ, Harrison JA (2008) Temperature dependence of single asperity diamond-diamond friction elucidated using AFM and MD simulations. *J Phys Chem C* (in press)
- Buldum A, Ciraci S (1997) Atomic-scale study of dry sliding friction. *Phys Rev B* 55(4):2606–2611
- Buldum A, Lu JP (1999) Atomic scale sliding and rolling of carbon nanotubes. *Phys Rev Lett* 83(24):5050–5053
- Burnham NA, Colton RJ (1989) Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *J Vac Sci Technol A* 7:2906–2913
- Burnham NA, Colton RJ (1993) Force microscopy. In: Bonnell DA (eds) Scanning tunneling microscopy and spectroscopy: theory, techniques, and applications. VCH Publishers, New York, pp 191–249
- Burnham NA, Dominguez DD, Mowery RL, Colton RJ (1990) Probing the surface forces of monolayer films with an atomic-force microscope. *Phys Rev Lett* 64(16):1931–1934
- Burnham NA, Colton RJ, Pollock HM (1993) Work-function anisotropies as an origin of long-range surface forces. *Phys Rev Lett* 69:144–147
- Cagin T, Che JW, Gaidos MN, Fijany A, Goddard WA (1999) Simulation and experiments on friction and wear of diamond: a material for MEMS and NEMS application. *Nanotechnology* 10(3):278–284
- Cai J, Wang JS (2001) Friction between a Ge tip and the (001)-2 × 1 surface: a molecular-dynamics simulation. *Phys Rev B* 6411(11):113313
- Cai J, Wang J-S (2002) Friction between Si tip and (001)-2 × 1 surface: a molecular dynamics simulation. *Comput Phys Commun* 147:145–148
- Car R, Parrinello M (1985) Unified approach for molecular-dynamics and density-functional theory. *Phys Rev Lett* 55(22):2471–2474
- Carpick RW, Salmeron M (1997) Scratching the surface: fundamental investigations of tribology with atomic force microscopy. *Chem Rev* 97(4):1163–1194
- Carpick RW, Agrait N, Ogletree DF, Salmeron M (1996a) Variation of the interfacial shear strength and adhesion of a nanometer-sized contact. *Langmuir* 12(13):3334–3340
- Carpick RW, Agrait N, Ogletree DF, Salmeron M (1996b) Measurement of interfacial shear (friction) with an ultrahigh vacuum atomic force microscope (vol 14, pg 1289, 1996). *J Vac Sci Technol B* 14(4):2772
- Chandross M, Grest GS, Stevens MJ (2002) Friction between alkylsilane monolayers: molecular simulation of ordered monolayers. *Langmuir* 18:8392–8399
- Chandross M, Webb EB III, Stevens MJ, Grest GS (2004) Systematic study of the effect of disorder on nanotribology of self-assembled monolayers. *Phys Rev Lett* 93:166103
- Chandross M, Lorenz CD, Stevens MJ, Grest GS (2008) Simulations of nanotribology with realistic probe tip models. *Langmuir* 24(4):1240–1246
- Chateauneuf GM, Mikulski PT, Gao GT, Harrison JA (2004) Compression- and shear-induced polymerization in model diacetylene-containing monolayers. *J Phys Chem B* 108(43):16626–16635
- Cheong WCD, Zhang LC (2000) Molecular dynamics simulation of phase transformations in silicon monocrystals due to nano-indentation. *Nano Tech* 11:173–180
- Cho K, Joannopoulos JD (1995) Mechanical hysteresis on an atomic-scale. *Surf Sci* 328(3):320–324
- Clarke DR, Kroll MC, Kirchner PD, Cook RF, Hockey BJ (1988) Amorphization and conductivity of silicon and

- germanium induced by indentation. *Phys Rev Lett* 60:2156–2159
- Costakramer JL, Garcia N, Garciamochales P, Serena PA (1995) Nanowire formation in macroscopic metallic contacts—quantum-mechanical conductance tapping a table top. *Surf Sci* 342(1–3):L1144–L1149
- Cramer C (2004) Essentials of computational chemistry, theories and models, 2nd edn. John Wiley & Sons, West Sussex, England
- Curry JE, Zhang FS, Cushman JH, Schoen M, Diestler DJ (1994) Transient coexisting nanophases in ultrathin films confined between corrugated walls. *J Chem Phys* 101(12):10824–10832
- D'Alessandro M, D'Abromo M, Brancato G, Di Nola A, Amadei A (2002) Statistical mechanics and thermodynamics of simulated ionic solutions. *J Phys Chem B* 106(45):11843–11848
- Daw MS, Baskes MI (1983) Semiempirical, quantum-mechanical calculation of hydrogen embrittlement in metals. *Phys Rev Lett* 50(17):1285–1288
- Dayo A, Alnasrallah W, Krim J (1998) Superconductivity-dependent sliding friction. *Phys Rev Lett* 80(8):1690–1693
- de la Fuente OR, Zimmerman JA, Gonzalez MA, de la Figuera J, Hamilton JC, Pai WW, Rojo JM (2002) Dislocation emission around nanoindentations on a (001) fcc metal surface studied by scanning tunneling microscopy and atomistic simulations. *Phys Rev Lett* 88(3):036101
- Derjaguin BV, Muller VM, Toporov Y (1975) Effect of contact deformations on adhesion of particles. *J Colloid Interface Sci* 53:314–326
- Dickrell PL, Sinnott SB, Hahn DW, Raravikar NR, Schadler LS, Ajayan PM, Sawyer WG (2005) Frictional anisotropy of oriented carbon nanotube surfaces. *Trib Lett* 18(1):59–62
- Dowson D (1979) History of tribology. Longman, London, p 215
- Dyson AJ, Smith PV (1996) Extension of the Brenner empirical interatomic potential to C–Si–H. *Surf Sci* 355:140–150
- Enachescu M, van den Oetelaar RJA, Carpick RW, Ogletree DF, Flipse CFJ, Salmeron M (1998) Atomic force microscopy study of an ideally hard contact: The diamond (111) tungsten carbide interface. *Phys Rev Lett* 81(9):1877–1880
- Enke K (1981) Some new results on the fabrication of and the mechanical, electrical and optical-properties of I-carbon layers. *Thin Solid Films* 80(1–3):227–234
- Enke K, Dimigen H, Hubsch H (1980) Frictional-properties of diamond-like carbon layers. *Appl Phys Lett* 36(4):291–292
- Erdemir A, Donnet C (2000) Tribology of diamond, diamond-like carbon, and related films. In: Bhushan B (ed) Modern tribology handbook. CRC Press, Boca Raton, FL, pp 871–908
- Erlandsson R, Hadzioannou G, Mate CM, McClelland GM, Chiang S (1988) Atomic scale friction between the muscovite mica cleavage plane and a tungsten tip. *J Chem Phys* 89(8):5190–5193
- Falvo MR, Taylor RM, Helser A, Chi V, Brooks FP, Washburn S, Superfine R (1999) Nanometer-scale rolling and sliding of carbon nanotubes. *Nature* 397:236–238
- Falvo MR, Steele J, Taylor RM II, Superfine R (2000) Gearlike rolling motion mediated by commensurate contact: carbon nanotubes on HOPG. *Phys Rev B* 62 (16):R10664–R10667
- Fang T-H, Weng C-I, Chang J-G (2002) Molecular dynamics simulation of nano-lithography process using atomic force microscopy. *Surf Sci* 501:138–147
- Feng Z, Field JE (1991) Friction of diamond on diamond and chemical vapor-deposition diamond coatings. *Surf Coat Technol* 47(1–3):631–645
- Finnis M (2003) Interatomic forces in condensed matter. Oxford series on materials. Oxford
- Foiles SM (1985) Application of the embedded-atom method to liquid transition metals. *Phys Rev B* 32:3409–3415
- Fournel M, Lacaze E, Schott M (1996) Tip-surface interactions in STM experiments on Au(111): atomic-scale metal friction. *Europhys Lett* 34 (7):489–494
- Frenkel FC, Kontorova T (1938) On the theory of plastic demortation and twinning. *Zh Eksp Teor Fiz* 8:1340
- Frenkel D, Smit B (1996) Understanding molecular simulation: from algorithms to applications. Academic Press, San Diego
- Frenken JWM, Vanpintxteren HM, Kuipers L (1993) New views on surface melting obtained with STM and ion-scattering. *Surf Sci* 283(1–3):283–289
- Fujisawa S, Sugawara Y, Ito S, Mishima S, Okada T, Morita S (1993) The two-dimensional stick-slip phenomenon with atomic resolution. *Nanotechnology* 4(3):138–142
- Fujisawa S, Sugawara Y, Morita S, Ito S, Mishima S, Okada T (1994) Study on the stick-slip phenomenon on a cleaved surface of the muscovite mica using an atomic-force lateral force microscope. *J Vac Sci Technol B* 12(3):1635–1637
- Fujisawa S, Sugawara Y, Morita S (1996) Localized fluctuation of a two-dimensional atomic-scale friction. *Jpn J Appl Phys Part 1 Regul Pap Short Notes Rev Pap* 35(11):5909–5913
- Fusco C, Fasolino A (2005) Velocity dependence of atomic-scale friction: A comparative study of the one- and two-dimensional Tomlinson model. *Phys Rev B* 71:045413
- Gad-el-Hak M (ed) (2002) The MEMS handbook. The mechanical engineering handbook series. CRC Press, Boca Raton, FL
- Gao GT, Mikulski PT, Harrison JA (2002) Molecular-scale tribology of amorphous carbon coatings: effects of film thickness, adhesion, and long-range interactions. *J Am Chem Soc* 124 (24):7202–7209
- Gao GT, Mikulski PT, Chateauneuf GM, Harrison JA (2003) The effects of film structure and surface hydrogen on the properties of amorphous carbon films. *J Phys Chem B* 107(40):11082–11090

- Gao GT, Cannara RJ, Carpick RW, Harrison JA (2007) Atomic-scale friction on diamond: a comparison of different sliding directions on (001) and (111) surfaces using MD and AFM. *Langmuir* 23(10):5394–5405
- GarciaParajo M, Longo C, Servat J, Gorostiza P, Sanz F (1997) Nanotribological properties of octadecyltrichlorosilane self-assembled ultrathin films studied by atomic force microscopy: contact and tapping modes. *Langmuir* 13(8):2333–2339
- Garg A, Sinnott SB (1999) Molecular dynamics of carbon nanotube proximal probe tip-surface contacts. *Phys Rev B* 60(19):13786–13791
- Garg A, Han J, Sinnott SB (1998) Interactions of carbon-nanotubule proximal probe tips with diamond and graphene. *Phys Rev Lett* 81(11):2260–2263
- Gear CW (1971) Numerical initial value problems in ordinary differential equations. Prentice-Hall, Inc., Englewood Cliffs, NJ
- Germann GJ, Cohen SR, Neubauer G, McClelland GM, Seki H, Coulman D (1993) Atomic scale friction of a diamond tip on diamond (100)-surface and (111)-surface. *J Appl Phys* 73(1):163–167
- Glosli JN, McClelland GM (1993) Molecular-dynamics study of sliding friction of ordered organic monolayers. *Phys Rev Lett* 70(13):1960–1963
- Glosli JN, Philpott MR, McClelland GM (1995) Molecular dynamics simulation of mechanical deformation of ultra-thin amorphous carbon films. *Math Res Soc Symp Proc* 383:431–435
- Grabhorn H, Otto A, Schumacher D, Persson BNJ (1992) Variation of the DC-resistance of smooth and atomically rough silver films during exposure to C_2H_6 and C_2H_4 . *Surf Sci* 264(3):327–340
- Haile JM (1992) Molecular dynamics simulation: elementary methods. John Wiley and Sons Inc., New York
- Hammerberg JE, Holian BL, Zhuo SJ (1995) Studies of sliding friction in compressed copper. In: Conference of the American physical society topical group on shock compression of condensed matter. AIP Press, Seattle, WA
- Harrison JA, Brenner DW (1994) Simulated tribochemistry—an atomic-scale view of the wear of diamond. *J Am Chem Soc* 116(23):10399–10402
- Harrison JA, Brenner DW (1995) In: Bhushan B (ed) Handbook of micro/nanotechnology. Chemical Rubber, Boca Raton, FL
- Harrison JA, White CT, Colton RJ, Brenner DW (1992a) Nanoscale investigation of indentation, adhesion and fracture of diamond (111) surfaces. *Surf Sci* 271(1–2):57–67
- Harrison JA, White CT, Colton RJ, Brenner DW (1992b) Molecular-dynamics simulations of atomic-scale friction of diamond surfaces. *Phys Rev B* 46(15):9700–9708
- Harrison JA, Colton RJ, White CT, Brenner DW (1992c) Atomistic simulation of the nanoindentation of diamond and graphite surfaces. *Math Res Soc Symp Proc* 239:573–578
- Harrison JA, Colton RJ, White CT, Brenner DW (1993a) Effect of atomic-scale surface-roughness on friction—a molecular-dynamics study of diamond surfaces. *Wear* 168(1–2):127–133
- Harrison JA, White CT, Colton RJ, Brenner DW (1993b) Atomistic simulations of friction at sliding diamond interfaces. *MRS Bull* 18(5):50–53
- Harrison JA, White CT, Colton RJ, Brenner DW (1993c) Effects of chemically-bound, flexible hydrocarbon species on the frictional-properties of diamond surfaces. *J Phys Chem* 97(25):6573–6576
- Harrison JA, White CT, Colton RJ, Brenner DW (1995) Investigation of the atomic-scale friction and energy-dissipation in diamond using molecular-dynamics. *Thin Solid Films* 260(2):205–211
- Harrison JA, Stuart SJ, Robertson DH, White CT (1997) Properties of capped nanotubes when used as SPM tips. *J Phys Chem B* 101(47):9682–9685
- Harrison JA, Stuart SJ, Tutein AB (2001) A new, reactive potential energy function to study indentation and friction of C_{13} n-alkane monolayers. In: Frommer JE, Overney R (eds) Interfacial properties on the submicron scale. ACS Press, Washington, DC pp 216–229
- Harrison JA, Schall JD, Kinippenberg MT, Gao G, Mikulski PT (2008) Elucidating atomic-scale friction using molecular dynamics and specialized analysis techniques. *J Phys Condens Matter* (in press)
- Hasnaoui A, Derlet PM, Swygenhoven HV (2004) Interaction between dislocations and grain boundaries under an indentor—a molecular dynamics simulation. *Acta Mater* 52:2251–2258
- Heermann DW (1986) Computer simulation methods in theoretical physics. Springer, Berlin
- Helman JS, Baltensperger W, Holyst JA (1994) Simple-model for dry friction. *Phys Rev B* 49(6):3831–3838
- Heo S, Sinnott SB (2007) Effect of molecular interactions on carbon nanotube friction. *J Appl Phys* 102(6)
- Heo S-J, Sinnott SB, Brenner DW, Harrison JA (2005) Computational modeling of nanometer-scale tribology. In: Bhushan B (ed) Nanotribology and nanomechanics. Springer, Heidelberg, Germany
- Heo SJ, Jang I, Barry PR, Phillipot SR, Perry SS, Sawyer WG, Sinnott SB (2008) Effect of the sliding orientation on the tribological properties of polyethylene in molecular dynamics simulations. *J Appl Phys* 103(8)
- Hoover WG (1986) Molecular dynamics. Springer, Berlin
- Hu Y (2008) Personal communication
- Irving DL, Brenner DW (2006) Diffusion on a self-assembled monolayer: molecular modeling of a bound plus mobile lubricant. *J Phys Chem B* 110(31):15426–15431
- Israelachvili JN (1992a) Fundamentals of friction. In: Singer IL, Pollock HM (eds) Macroscopic and microscopic processes. Kluwer Academic Publishers, The Netherlands, p 351
- Israelachvili JN (1992) Intermolecular and surface forces. Academic Press, London, pp 169–172
- Iwasaki T, Miura H (2001) Molecular dynamics analysis of adhesion strength of interfaces between thin films. *J Mater Res* 16(6):1789–1794

- Jang I, Sinnott SB (2004) Molecular dynamics simulations of the chemical modification of polystyrene through $C_xF_y^+$ beam deposition. *J Phys Chem B* 108:9656–9664
- Jang I, Burris DL, Dickrell PL, Barry PR, Santos C, Perry SS, Phillipot SR, Sinnott SB, Sawyer WG (2007) Sliding orientation effects on the tribological properties of polytetrafluoroethylene. *J Appl Phys* 102(12)
- Johnson KL (1985) Contact mechanics. Cambridge University Press, Cambridge (Chapter 7)
- Johnson KL, Kendall K, Roberts AD (1971) Surface energy and the contact of elastic solids. *Proc R Soc Lond A* 324:301–313
- Joyce SA, Thomas RC, Houston JE, Michalske TA, Crooks RM (1992) Mechanical relaxation of organic monolayer films measured by force microscopy. *Phys Rev Lett* 68(18):2790–2793
- Kadau K, Germann TC, Lomdahl PS (2004) Large-scale molecular-dynamics simulation of 19 billion particles. *Int J Mod Phys C* 15:193–201
- Kailer A, Nickel KG, Gogotsi YG (1999) Raman microspectroscopy of nanocrystalline and amorphous phases in hardness indentations. *J Raman Spectrosc* 30:939–961
- Kallman JS, Hoover WG, Hoover CG, Degroot AJ, Lee SM, Wooten F (1993) Molecular-dynamics of silicon indentation. *Phys Rev B* 47(13):7705–7709
- Kawaguchi T, Matsukawa H (1997) Dynamical frictional phenomena in an incommensurate two-chain model. *Phys Rev B* 56:13932–13942
- Kelchner CL, Plimpton SJ, Hamilton JC (1998) Dislocation nucleation and defect structure during surface indentation. *Phys Rev B* 58(17):11085–11088
- Khurshudov A, Kato K (1995) Volume increase phenomena in reciprocal scratching of polycarbonate studied by atomic-force microscopy. *J Vac Sci Technol B* 13(5):1938–1944
- Khurshudov AG, Kato K, Koide H (1996) Nano-wear of the diamond AFM probing tip under scratching of silicon, studied by AFM. *Tribol Lett* 2:345
- Koike A, Yoneya M (1996) Molecular dynamics simulations of sliding friction of Langmuir-Blodgett monolayers. *J Chem Phys* 105(14):6060–6067
- Kokubo S (1932) Science reports of the Tohoku Imperial University, vol 21, p 256
- Komanduri R, Chandrasekaran N (2000) Molecular dynamics simulation of atomic-scale friction. *Phys Rev B* 61:14007–14019
- Komvopoulos K, Yan W (1997) Molecular dynamics simulation of single and repeated indentation. *J Appl Phys* 82(10):4823–4830
- Kreer T, Müser MH, Binder K, Klein J (2001) Frictional drag mechanisms between polymer-bearing surfaces. *Langmuir* 17:7804–7813
- Kreer T, Binder K, Müser MH (2003) Friction between polymer brushes in good solvent conditions: steady-state sliding versus transient behavior. *Langmuir* 19:7551–7559
- Kremer K, Grest GS (1990) Dynamics of entangled linear polymer melts—a molecular-dynamics simulation. *J Chem Phys* 92(8):5057–5086
- Krim J (1995) Progress in nanotribology: experimental probes of atomic scale friction. *Comments Cond Mat Phys* 17:263
- Krim J (1996a) Friction at the atomic scale. *Sci Am* 275 (4):74–80
- Krim J (1996b) Atomic-scale origins of friction. *Langmuir* 12(19):4564–4566
- Krim J, Solina DH, Chiarello R (1991) Nanotribology of a Kr monolayer—a quartz-crystal microbalance study of atomic-scale friction. *Phys Rev Lett* 66(2):181–184
- Landman U, Luedtke WD (1991) Nanomechanics and dynamics of tip substrate interactions. *J Vac Sci Technol B* 9(2):414–423
- Landman U, Luedtke WD, Nitzan A (1989a) Dynamics of tip substrate interactions in atomic force microscopy. *Surf Sci* 210(3):L177–L184
- Landman U, Luedtke WD, Ribarsky MW (1989b) Structural and dynamical consequences of interactions in interfacial systems. *J Vac Sci Tech A* 7(4):2829–2839
- Landman U, Luedtke WD, Burnham NA, Colton RJ (1990) Atomistic mechanisms and dynamics of adhesion, nanoindentation, and fracture. *Science* 248 (4954):454–461
- Landman U, Luedtke WD, Ringer EM (1992) Atomistic mechanisms of adhesive contact formation and interfacial processes. *Wear* 153(1):3–30
- Landman U, Luedtke WD, Ouyang J, Xia TK (1993) Nanotribology and the stability of nanostructures. *Jpn J Appl Phys Part 1 Regul Pap Short Notes Rev Pap* 32(3B):1444–1462
- Landman U, Luedtke WD, Gao JP (1996) Atomic-scale issues in tribology: Interfacial junctions and nano-elastohydrodynamics. *Langmuir* 12(19):4514–4528
- Lee B-J, Baskes MI (2000) Second nearest-neighbor modified embedded-atom-method potential. *Phys Rev B* 62(13):8564–8567
- Lee S, Shon YS, Colorado R, Guenard RL, Lee TR, Perry SS (2000) The influence of packing densities and surface order on the frictional properties of alkanethiol self-assembled monolayers (SAMs) on gold: A comparison of SAMS derived from normal and spiroalkanedithiols. *Langmuir* 16(5):2220–2224
- Legoas SB, Giro R, Galvao DS (2004) Molecular dynamics simulations of C-60 nanobearings. *Chem Phys Lett* 386(4–6):425–429
- Leng Y, Jiang S (2002) Dynamic simulations of adhesion and friction in chemical force microscopy. *J Am Chem Soc* 124:11764–11770
- Li B, Clapp PC, Rifkin JA, Zhang XM (2001) Molecular dynamics simulation of stick-slip. *J Appl Phys* 90:3090–3094
- Li S, Cao P, Ramon Colorado J, Yan X, Wenzl I, Shmakova OE, Graupe M, Lee TR, Perry SS (2005) Local packing environment strongly influences the

- frictional properties of mixed CH_3- and CF_3 -terminated alkanethiol SAMs on Au(111). *Langmuir* 21:933–936
- Liang T, Sawyer WG, Perry SS, Sinnott SB, Phillpot SR (2008) First-principles determination of static potential energy surfaces for atomic friction in MoS_2 and MoO_3 . *Phys Rev B* 77(10)
- Lilleodden ET, Zimmerman JA, Foiles SM, Nix WD (2003) Atomistic simulations of elastic deformation and dislocation nucleation during nanoindentation. *J Mech Phys Solids* 51:901–920
- Lüthi R, Meyer E, Haefke H (1993) Sled-type motion on the nanometer scale: determination of dissipation and cohesive energies of C_{60} . *Science* 266:1979–1981
- Lüthi R, Meyer E, Haefke H, Howald L, Gutmannsbauer W, Güntherodt HJ (1994a) Sled-type motion on the nanometer-scale—determination of dissipation and cohesive energies of C-60. *Science* 266(5193):1979–1981
- Lüthi R, Haefke H, Meyer E, Howald L, Lang H-P, Gerth G, Güntherodt HJ (1994b) Frictional and atomic-scale study of C_{60} thin films by scanning force microscopy. *Z Phys B* 95:1–3
- Manias E, Hadzioannou G, ten Brinke G (1996) Inhomogeneities in sheared ultrathin lubricating films. *Langmuir* 12(19):4587–4593
- Martyna GJ, Klein ML, Tuckerman M (1992) Nose-Hoover chains—the canonical ensemble via continuous dynamics. *J Chem Phys* 97(4):2635–2643
- Mate CM (1992) Atomic-force-microscope study of polymer lubricants on silicon surfaces. *Phys Rev Lett* 68(22):3323–3326
- Mate CM (1993) Nanotribology studies of carbon surfaces by force microscopy. *Wear* 168(1–2):17–20
- Mate CM (1995) Force microscopy studies of the molecular origins of friction and lubrication. *IBM J Res Dev* 39(6):617–627
- Mate CM, McClelland GM, Erlandsson R, Chiang S (1987) Atomic-scale friction of a tungsten tip on a graphite surface. *Phys Rev Lett* 59(17):1942–1945
- McClelland GM, Glosli JN (1992) Friction at the atomic scale. In: Singer IL, Pollock HM (eds) *Fundamentals of friction: macroscopic and microscopic processes*. Kluwer Academic Publishers, Dordrecht, pp 405–422
- Meyer E (1998) *Nanoscience: friction and rheology on the nanometer scale*. World Scientific, New Jersey
- Meyer E, Overney R, Brodbeck D, Howald L, Luthi R, Frommer J, Güntherodt HJ (1992) Friction and wear of Langmuir-Blodgett-films observed by friction force microscopy. *Phys Rev Lett* 69(12):1777–1780
- Mikulski PT, Harrison JA (2001a) Periodicities in the properties associated with the friction of model self-assembled monolayers. *Tribol Lett* 10(1–2):29–35
- Mikulski PT, Harrison JA (2001b) Packing-density effects on the friction of n-alkane monolayers. *J Am Chem Soc* 123(28):6873–6881
- Mikulski PT, Gao G, Chateauneuf GM, Harrison JA (2005a) Contact forces at the sliding interface: mixed versus pure model alkane monolayers. *J Chem Phys* 122:024701
- Mikulski PT, Herman LA, Harrison JA (2005b) Odd and even model self-assembled monolayers: Links between friction and structure. *Langmuir* 21(26):12197–12206
- Minowa K, Sumino K (1992) Stress-induced amorphization of a silicon crystal by mechanical scratching. *Phys Rev Lett* 69(2):320–322
- Miura K, Takagi T, Kamiya S, Sahashi T, Yamauchi M (2001a) Natural rolling of zigzag multiwalled carbon nanotubes on graphite. *Nano Lett* 1(3):161–163
- Miura K, Ishikawa M, Kitanishi R, Yoshimura M, Ueda K, Tatsumi Y, Minami N (2001b) Bundle structure and sliding of single-walled carbon nanotubes observed by friction-force microscopy. *Appl Phys Lett* 78(6):832–834
- Miura K, Kamiya S, Sasaki N (2003) C_{60} molecular bearings. *Phys Rev Lett* 90(5):055509
- Miyake S, Takahashi S, Watanabe I, Yoshihara H (1987) Friction and wear behavior of hard carbon-films. *Asle Trans* 30(1):121–127
- Morita S, Fujisawa S, Sugawara Y (1996) Spatially quantized friction with a lattice periodicity. *Surf Sci Rep* 23(1):1–41
- Mulliah D, Kenny SD, Smith R (2004) Modeling of stick-slip phenomena using molecular dynamics. *Phys Rev B* 69(20):205407
- Müser MH (2002a) Nature of mechanical instabilities and their effect on kinetic friction. *Phys Rev Lett* 89(22):224301
- Müser MH (2002b) Towards an atomistic understanding of solid friction by computer simulations. *Comput Phys Commun* 146:54–62
- Neitola R, Pakkanen TA (2001) Ab initio studies on the atomic-scale origin of friction between diamond (111) surfaces. *J Phys Chem B* 105:1338–1343
- Ni B, Sinnott SB (2001a) Tribological properties of carbon nanotube bundles. *Surf Sci* 487:87–96
- Ni B, Sinnott SB (2001b) Mechanical and tribological properties of carbon nanotubes investigated with atomistic simulations. In: *Nanotubes and related materials*. Materials Research Society, Pittsburgh, PA, pp A17.3.1–A17.3.5
- Ni B, Lee K-H, Sinnott SB (2004) Development of a reactive empirical bond order potential for hydrocarbon-oxygen interactions. *J Phys C Condens Matter* 16:7261–7275
- Nieminan JA, Sutton AP, Pethica JB, Kaski K (1992) Mechanism of lubrication by a thin solid film on a metal-surface. *Model Simul Math Sci Eng* 1:83
- Nieminan JA, Sutton AP, Pethica JB (1992) Static junction growth during frictional sliding of metals. *Acta Metall Mater* 40(10):2503–2509

- Nose S (1984a) A unified formulation of the constant temperature molecular-dynamics methods. *J Chem Phys* 81(1):511–519
- Nose S (1984b) A molecular-dynamics method for simulations in the canonical ensemble. *Mol Phys* 52 (2):255–268
- Ohira T, Inoue Y, Murata K, Murayama J (2001) Magnetite scale cluster adhesion on metal oxide surfaces: atomistic simulation study. *Appl Surf Sci* 171:175–188
- Ohmae N (1996) Field ion microscopy of microdeformation induced by metallic contacts. *Philos Mag A Phys Condens Matter Struct Defects Mech Prop* 74 (5):1319–1327
- Ohno K, Esfarjani K, Kawazoe Y (1999) Computational materials science from ab initio to Monte Carlo methods. Springer, New York
- Okita S, Miura K (2001) Molecular arrangement in C₆₀ and C₇₀ films on graphite and their nanotribological behavior. *Nano Lett* 1(2):101–103
- Okita S, Ishikawa M, Miura K (1999) Nanotribological behavior of C₆₀ films at an extremely low load. *Surf Sci* 442:L959–L963
- Oppel GU (1964) Biaxial elasto-plastic analysis of load and residual stresses. *Exp Mech* 4:135
- Ouyang Q, Okada K (1994) Nano-ball bearing effect of ultra-fine particles of cluster diamond. *Appl Surf Sci* 78:309–313
- Overney RM, Meyer E, Frommer J, Brodbeck D, Luthi R, Howald L, Guntherodt HJ, Fujihira M, Takano H, Gotoh Y (1992) Friction measurements on phase-separated thin-films with a modified atomic force microscope. *Nature* 359(6391):133–135
- Overney RM, Bonner T, Meyer E, Reutschi M, Luthi R, Howald L, Frommer J, Guntherodt HJ, Fujihara M, Takano H (1994a) Elasticity, wear, and friction properties of thin organic films observed with atomic-force microscopy. *J Vac Sci Technol, B* 12 (3):1973–1976
- Overney RM, Meyer E, Frommer J, Guntherodt HJ, Fujihira M, Takano H, Gotoh Y (1994b) Force microscopy study of friction and elastic compliance of phase-separated organic thin-films. *Langmuir* 10 (4):1281–1286
- Overney RM, Takano H, Fujihira M, Meyer E, Guntherodt HJ (1994c) Wear, friction and sliding speed correlations on langmuir-blodgett-films observed by atomic-force microscopy. *Thin Solid Films* 240(1–2):105–109
- Padgett CW, Brenner DW (2005) A continuum-atomistic method for incorporating Joule heating into classical molecular dynamics simulations. *Mol Simul* 31 (11):749–757
- Parr RG, Yang W (1989) Density functional theory of atoms and molecules. Oxford University Press, New York
- Pasianot R, Savino EJ (1992) Embedded-atom-method interatomic potentials for hcp metals. *Phys Rev B* 45:12704–12710
- Pasianot R, Farkas D, Savino EJ (1991) Empirical many-body interatomic potential for bcc transition metals. *Phys Rev B* 43:6952–6961
- Perry MD, Harrison JA (1995) Universal aspects of the atomic-scale friction of diamond surfaces. *J Phys Chem B* 99:9960–9965
- Perry MD, Harrison JA (1996a) Molecular dynamics studies of the frictional properties of hydrocarbon materials. *Langmuir* 12(19):4552–4556
- Perry MD, Harrison JA (1996b) Molecular dynamics investigations of the effects of debris molecules on the friction and wear of diamond. *Thin Solid Films* 291:211–215
- Perry MD, Harrison JA (1997) Friction between diamond surfaces in the presence of small third-body molecules. *J Phys Chem B* 101(8):1364–1373
- Persson BNJ (1993) Applications of surface resistivity to atomic scale friction, to the migration of hot adatoms, and to electrochemistry. *J Chem Phys* 98(2):1659–1672
- Persson BNJ (1995) Theory of friction—dynamical phase-transitions in adsorbed layers. *J Chem Phys* 103(9):3849–3860
- Persson BNJ (1997) Theory of friction: friction dynamics for boundary lubricated surfaces. *Phys Rev B* 55 (12):8004–8012
- Persson BNJ, Tosatti E (1994) Layering transition in confined molecular thin-films—nucleation and growth. *Phys Rev B* 50(8):5590–5599
- Persson BNJ, Volokitin AI (1995) Electronic friction of physisorbed molecules. *J Chem Phys* 103(19):8679–8683
- Persson BNJ, Schumacher D, Otto A (1991) Surface resistivity and vibrational damping in adsorbed layers. *Chem Phys Lett* 178(2–3):204–212
- Pethica JB (1986) Interatomic forces in scanning tunneling microscopy—giant corrugations of the graphite surface—comment. *Phys Rev Lett* 57(25):3235
- Pokropivny VV, Skorokhod VV, Pokropivny AV (1997) Atomistic mechanism of adhesive wear during friction of atomic sharp tungsten asperity over (114) bcc-iron surface. *Mater Lett* 31(1–2):49–54
- Raecker TJ, Depristo AE (1991) Theory of chemical bonding based on the atom-homogeneous electron-gas system. *Int Rev Phys Chem* 10(1):1–54
- Raffi-Tabar H, Sutton AP (1991) Long-range Finnis-Sinclair potentials for FCC metallic alloys. *Phil Mag Lett* 63:217–224
- Rappe AK, Goddard WA III (1991) Charge equilibration for molecular dynamics simulations. *J Phys Chem* 95:3358–3363
- Reimann P, Evstigneev M (2004) Nonmonotonic velocity dependence of atomic friction. *Phys Rev Lett* 93:230802
- Ritter C, Heyde M, Stegemann B, Rademann K, Schwarz UD (2005) Contact-area dependence of frictional forces: Moving adsorbed antimony nanoparticles. *Phys Rev B* 71:085405

- Ruan JA, Bhushan B (1994) Atomic-scale and microscale friction studies of graphite and diamond using friction force microscopy. *J Appl Phys* 76(9):5022–5035
- Ruan J, Bhushan B (1996) Nanoindentation studies of sublimed fullerene films using atomic force microscopy. *J Mater Res* 8(12):3019–3022
- Samuels B, Wilks J (1988) The Friction of Diamond Sliding on Diamond. *Journal of Materials Science* 23 (8):2846–2864
- Sanz-Navarro CF, Kenny SD, Smith R (2004) Atomistic simulations of structural transformations. *Nanotechnology* 15:692–697
- Schall JD, Brenner DW (2000) Molecular dynamics simulations of carbon nanotube rolling and sliding on graphite. *Mol Simul* 25(1):73–80
- Schall JD, Brenner DW (2004) Atomistic simulation of the influence of pre-existing stress on the interpretation of nanoindentation data. *J Mater Res* 19:3172–3180
- Schall JD, Padgett CW, Brenner DW (2005) Ad hoc continuum-atomistic thermostat for modeling heat flow in molecular dynamics simulations. *Mol Simul* 31:283–288
- Schall JD, Gao G, Harrison JA (2008) Elastic constants of silicon materials calculated as a function of temperature using a parametrization of the second-generation reactive empirical bond-order potential. *Phys Rev B* 77(11)
- Schneider T, Stoll E (1978) Molecular-dynamics study of a 3-dimensional one-component model for distortive phase-transitions. *Phys Rev B* 17(3):1302–1322
- Schoen M, Rhykerd CL, Diestler DJ, Cushman JH (1989) Shear forces in molecularly thin-films. *Science* 245 (4923):1223–1225
- Schwarz UD, Allers W, Gensterblum G, Wiesendanger R (1995a) Low-load friction behavior of epitaxial C₆₀ monolayers under Hertzian contact. *Phys Rev B* 52:14976–14984
- Schwarz UD, Allers W, Gensterblum G, Wiesendanger R (1995b) Low-load friction behavior of epitaxial C-60 monolayers under Hertzian contact. *Phys Rev B* 52 (20):14976–14984
- Schwarz UD, Zworer O, Koster P, Wiesendanger R (1997) Quantitative analysis of the frictional properties of solid materials at low loads. I. Carbon compounds. *Phys Rev B* 56(11):6987–6996
- Sheehan PE, Lieber CM (1996) Nanotribology and nanofabrication of MoO₃ structures by atomic force microscopy. *Science* 272:1158–1161
- Shenderova O, Mewkill J, Brenner DW (2000) Nanoindentation as a probe of nanoscale residual stresses: atomistic simulation results. *Mol Simul* 25(1–2):81
- Shluger AL, Williams RT, Rohl AL (1995) Lateral and friction forces originating during force microscope scanning of ionic surfaces. *Surf Sci* 343(3):273–287
- Shon Y-S, Lee S, Colorado R, Perry SS, Lee TR (2000) Spiroalkanedithiol-based SAMs reveal unique insight into the wettabilities and frictional properties of organic thin films. *J Am Chem Soc* 122:7556–7563
- Simes TR, Mellor SG, Hills DA (1984) A note on the influence of residual-stress on measured hardness. *J Strain Anal Eng Des* 19(2):135–137
- Sines G, Calson R (1952) Hardness measurements for determination of residual stresses. *ASTM Bull* 180:35
- Singer IL (1991) A thermochemical model for analyzing low wear-rate materials. *Surf Coat Technol* 49(1–3):474–481
- Singer IL (1994) Friction and energy dissipation at the atomic scale: a review. *J Vac Sci Technol A* 12:2605–2616
- Singer IL, Fayeulle S, Ehni PD (1991) Friction and wear behavior of tin in air—the chemistry of transfer films and debris formation. *Wear* 149(1–2):375–394
- Sinnott SB (2000) Theory of atomic-scale friction. In: Nalwa H (ed) *Handbook of nanostructured materials and nanotechnology*. Academic Press, San Diego, CA, pp 571–618
- Sinnott SB, Andrews R (2001) Carbon nanotubes: synthesis, properties and applications. *Crit Rev Solid State Math Sci* 26(3):145–249
- Sinnott SB, Colton RJ, White CT, Shenderova OA, Brenner DW, Harrison JA (1997) Atomistic simulations of the nanometer-scale indentation of amorphous-carbon thin films. *J Vac Sci Technol A Vac Surf Films* 15(3):936–940
- Sinnott SB, Shenderova OA, White CT, Brenner DW (1998) Mechanical properties of nanotubule fibers and composites determined from theoretical calculations and simulations. *Carbon* 36(1–2):1–9
- Smith RW, Was GS (1989) Application of molecular dynamics to the study of hydrogen embrittlement in Ni–Cr–Fe alloys. *Phys Rev B* 40:10322–10336
- Smith GS, Tadmor EB, Kaxiras E (2000) Multiscale simulation of loading and electrical resistance in silicon nanoindentation. *Phys Rev Lett* 84(6):1260–1263
- Sokoloff JB (1984) Theory of dynamical friction between idealized sliding surfaces. *Surf Sci* 144(1):267–272
- Sokoloff JB (1990) Theory of energy-dissipation in sliding crystal-surfaces. *Phys Rev B* 42(1):760–765
- Sokoloff JB (1992) Theory of atomic level sliding friction between ideal crystal interfaces. *J Appl Phys* 72 (4):1262–1270
- Sokoloff JB (1993) Possible nearly frictionless sliding for mesoscopic solids. *Phys Rev Lett* 71(21):3450–3453
- Sokoloff JB (1995) Microscopic mechanisms for kinetic friction: nearly frictionless sliding for small solids. *Phys Rev B* 52:7205–7214
- Sokoloff JB (1996) Theory of electron and phonon contributions to sliding friction. In: Persson BNJ, Tosatti E (eds) *Physics of sliding friction*. Kluwer Academic Publishers, Dordrecht, pp 217–229
- Sokoloff JB (2001) Static friction between elastic solids due to random asperities. *Phys Rev Lett* 86(15):3312–3315
- Sokoloff JB (2002) Possible microscopic explanation of the virtually universal occurrence of static friction. *Phys Rev B* 65:115415

- Sorensen MR, Jacobsen KW, Stoltze P (1996) Simulations of atomic-scale sliding friction. *Phys Rev B* 53(4):2101–2113
- Stillinger FH, Weber TA (1985) Computer-simulation of local order in condensed phases of silicon. *Phys Rev B* 31(8):5262–5271
- Stoneham AM, Ramos MMD, Sutton AP (1993) How do they stick together—the statics and dynamics of interfaces. *Philos Mag A Phys Condens Matter Struct Defects Mech Prop* 67(4):797–811
- Streitz FH, Mintmire JW (1994) Electrostatic potentials for metal-oxide surfaces and interfaces. *Phys Rev B* 50(16):11996–12003
- Stuart SJ, Tutein AB, Harrison JA (2000) A reactive potential for hydrocarbons with intermolecular interactions. *J Chem Phys* 112:6472–6486
- Sutton AP (1993) Electronic Structure of materials. Clarendon, Oxford
- Sutton AP (1996) Deformation mechanisms, electronic conductance and friction of metallic nanocontacts. *Curr Opin Solid State Mater Sci* 1(6):827–833
- Sutton AP, Pethica JB (1990) Inelastic flow processes in nanometre volumes of solids. *J Phys Condens Matter* 2:5317–5326
- Sutton AP, Pethica JB, Rafii-Tabar H, Nieminen JA (1992) In: Pettifor DG, Cottrell AH (eds) Electron theory in alloy design. Institute of Materials, London (Chapter 7)
- Tanaka J, Abrams CF, Graves DB (2000) New C–F interatomic potential for molecular dynamics simulation of fluorocarbon film formation. *Nucl Inst Meth B* 18(3):938–945
- Tang H, Joachim C, Devillers J (1993) Interpretation of Afm images—the graphite surface with a diamond tip. *Surf Sci* 291(3):439–450
- Tersoff J (1988) New empirical approach for the structure and energy of covalent systems. *Phys Rev B* 37:6991–7000
- Tersoff J (1989) Modeling solid-state chemistry: Interatomic potentials for multicomponent systems. *Phys Rev B* 39:5566–5569
- Thijssse BJ (2002) Relationship between the modified embedded-atom method and Stillinger-Weber potentials in calculating the structure of silicon. *Phys Rev B* 65:195207
- Thompson PA, Robbins MO (1990) Origin of stick-slip motion in boundary lubrication. *Science* 250(4982):792–794
- Thundat T, Warmack RJ, Ding D, Compton RN (1993) Atomic force microscope investigation of C_{60} adsorbed on silicon and mica. *Appl Phys Lett* 63(7):891–893
- Tomagnini O, Ercolessi F, Tosatti E (1991) Microscopic interaction between a gold tip and a Pb(110) surface. *Surf Sci* 287(288):1041–1045
- Tomagnini O, Ercolessi F, Tosatti E (1993) Microscopic interaction between a gold tip and a Pb(110) surface. *Surf Sci* 287:1041–1045
- Tomlinson GA (1929) A molecular theory of friction. *Philos Mag Ser* 7(7):905–939
- Totten GE, Liang H (2004) Mechanical tribology: materials characterization, and applications. Marcel Dekker, New York
- Tsui TY, Pharr GM, Oliver WC, Bhatia CS, White RL, Anders S, Anders A, Brown IG (1995) Nanoindentation and nanoscratching of hard carbon coatings for magnetic disks. *Math Res Soc Symp Proc* 383:447
- Tully JC (1980) Dynamics of gas-surface interactions—3d generalized Langevin model applied to Fcc and Bcc surfaces. *J Chem Phys* 73(4):1975–1985
- Tupper KJ, Brenner DW (1994) Compression-induced structural transition in a self-assembled monolayer. *Langmuir* 10(7):2335–2338
- Tupper KJ, Colton RJ, Brenner DW (1994) Simulations of self-assembled monolayers under compression—effect of surface asperities. *Langmuir* 10(7):2041–2043
- Tutein AB, Stuart SJ, Harrison JA (1999) Indentation analysis of linear-chain hydrocarbon monolayers anchored to diamond. *J Phys Chem B* 103(51):11357–11365
- vanden Oetelaar RJA, Flipse CFJ (1997) Atomic-scale friction on diamond (111) studied by ultra-high vacuum atomic force microscopy. *Surf Sci* 384(1–3):L828–L835
- Volokitin AI, Persson BNJ (2003a) Resonant photon tunneling enhancement of the van der Waals friction. *Phys Rev Lett* 91(10):106101
- Volokitin AI, Persson BNJ (2003b) Noncontact friction between nanostructures. *Phys Rev B* 68(15):155420
- Volokitin AI, Persson BNJ (2005) Adsorbate-induced enhancement of electrostatic noncontact friction. *Phys Rev Lett* 94:086104
- Walsh P, Omelchenko A, Kalia RK, Nakano A, Vashishta P, Saini S (2003) Nanoindentation of silicon nitride: a multimillion-atom molecular dynamics study. *Appl Phys Lett* 82:118–120
- Wang J, Rose KC, Lieber CM (1999) Load-independent friction: MoO_3 nanocrystal lubricants. *J Phys Chem B* 103(40):8405–8408
- Wong SS, Takano H, Porter MD (1998) Mapping orientation differences of terminal functional groups by friction force microscopy. *Anal Chem* 70(24):5209–5212
- Woodcock LV (1971) Isothermal molecular dynamics calculations for liquid salts. *Chem Phys Lett* 10(3):257–261
- Xue Q-J, Zhang X-S, Yan F-Y (1994) Study of the structural transformations of C_{60}/C_{70} crystals during friction. *Chin Sci Bull* 39(10):819–822
- Yanson AI, van Ruitenbeek JM, Yanson IK (2001a) Shell effects in alkali metal nanowires. *Low Temp Phys* 27(9–10):807–820
- Yanson AI, Yanson IK, van Ruitenbeek JM (2001b) Crossover from electronic to atomic shell structure in alkali metal nanowires. *Phys Rev Lett* 8721(21):216805
- Yasukawa A (1996) Using an extended tersoff interatomic potential to analyze the static-fatigue strength of SiO_2 under atmospheric influence. *JSME Int J A* 39(3):313–320

- Yokohata T, Kato K (1993) Mechanism of nanoscale indentation. *Wear* 168(1–2):109–114
- Yoshizawa H, Israelachvili J (1993) Fundamental mechanisms of interfacial friction. 2. Stick-slip friction of spherical and chain molecules. *J Phys Chem* 97(43):11300–11313
- Zhang L, Jiang S (2002) Molecular simulation study of nanoscale friction for alkyl monolayers on Si(111). *J Chem Phys* 117:1804–1811
- Zhang L, Leng Y, Jiang S (2003a) Tip-based hybrid simulation study of frictional properties of self-assembled monolayers: Effects of chain length, terminal group, scan direction, and scan velocity. *Langmuir* 19:9742–9747
- Zhang LZ, Leng YS, Jiang SY (2003b) Tip-based hybrid simulation study of frictional properties of self-assembled monolayers: effects of chain length, terminal group, scan direction, and scan velocity. *Langmuir* 19(23):9742–9747
- Zhong W, Overney G, Tomanek D (1991) In: Van Hove MA, Tong SY, Xie XD, Takayanagi K (eds) *The structure of surfaces III*. Springer, Berlin, p 243

Part III
Nanotribology

Nanotribology, Nanomechanics and Materials Characterization Studies Using Scanning Probe Microscopy

Bharat Bhushan

Abstract

Nanotribology and nanomechanics studies are needed to develop a fundamental understanding of interfacial phenomena on a small scale, and to study interfacial phenomena in micro/nanoelectromechanical systems (MEMS/NEMS), magnetic storage devices, and many other applications. Friction and wear of lightly loaded micro/nanocomponents are highly dependent on surface interactions (a few atomic layers). These structures generally are coated with molecularly thin films. Nanotribology and nanomechanics studies are also valuable in the fundamental understanding of interfacial phenomena in macrostructures, and provide a bridge between science and engineering. An atomic force microscope (AFM) tip is used to simulate a single asperity contact with a solid or lubricated surface. AFMs are used to study the various tribological phenomena, which include surface roughness, adhesion, friction, scratching, wear, detection of material transfer, and boundary lubrication. In situ surface characterization of local deformation of materials and thin coatings can be carried out using a tensile stage inside an AFM. Mechanical properties such as hardness, Young's modulus of elasticity, and creep/relaxation behavior can be determined on micro- to picoscales using a depth-sensing indentation system in an AFM. Localized surface elasticity and viscoelastic mapping near surface regions can be obtained of with nanoscale lateral resolution. Finally, an AFM can be used for nanofabrication/nanomachining.

B. Bhushan (✉)

Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

Keywords

Nanotribology · Nanofriction · Nanowear · Nanolubrication · Materials characterization · AFM · SPM · FFM · Nanoprobe microscopy

8.1 Introduction

The mechanisms and dynamics of the interactions of two contacting solids during relative motion, ranging from atomic-scale to microscale, need to be understood in order to develop a fundamental understanding of adhesion, friction, wear, indentation, and lubrication processes. For most solid-solid interfaces of technological relevance, contact occurs at multiple asperities. Consequently the importance of investigating single asperity contacts in studies of the fundamental micro/nanomechanical and micro/nanotribological properties of surfaces and interfaces has long been recognized. The emergence and proliferation of proximal probes, in particular scanning probe microscopies (the scanning tunneling microscope and the atomic force microscope), the surface force apparatus, and computational techniques for simulating tip-surface interactions and interfacial properties, allowed systematic investigations of interfacial problems with high resolution as well as ways and means for modifying and manipulating nanoscale structures. These advances led to the appearance of the field of nanotribology. The word nanotribology was used for the first time in the title of a research paper by Bhushan et al. (1995a), and a first book on the field by Bhushan in 1995 (first edition of Bhushan 1999a). The first dedicated international meeting on the topic was organized in 1996 in Sesimbra, Portugal (Bhushan 1997).

Nanotribology pertains to experimental and theoretical investigations of interfacial processes on scales ranging from the atomic- and molecular-scale to the microscale, occurring during adhesion, friction, scratching, wear, indentation, and thin-film lubrication at sliding surfaces (Singer and Pollock 1992; Bhushan et al. 1995a;

Guntherodt et al. 1995; Persson and Tosatti 1996; Bhushan 1997, 1999a, b, c, 2001a, b, c, 2005, 2008, 2011). Proximal probes have also been used for mechanical and electrical characterization, *in situ* characterization of local deformation, and other nanomechanics studies (Bhushan 1999c, 2001c, 2011, 2016).

Nanotribological and nanomechanical studies are needed to develop a fundamental understanding of interfacial phenomena on a small scale, and to study interfacial phenomena in nanostructures used in magnetic storage devices, micro/nanoelectromechanical systems (MEMS/NEMS), and other applications (Bhushan 1996, 1997, 1998, 1999a, b, c, d, 2001a, b, c, 2003, 2005, 2008, 2011, 2016; Bhushan et al. 1995a, b). Friction and wear of lightly loaded micro/nanocomponents are highly dependent on surface interactions (a few atomic layers). These structures generally are coated with molecularly thin films. Nanotribological and nanomechanical studies also are valuable in the fundamental understanding of interfacial phenomena in macrostructures, and provide a bridge between science and engineering.

The surface force apparatus (SFA), the scanning tunneling microscopes (STM), and atomic force and friction force microscopes (AFM and FFM) are used widely in nanotribological and nanomechanical studies. Typical operating parameters are compared in Table 8.1. The SFA, developed in 1968, is commonly employed to study both static and dynamic properties of molecularly thin films sandwiched between two molecularly smooth surfaces. The STM, developed in 1981, allows the imaging of electrically conducting surfaces with atomic resolution, and has been used for the imaging of clean surfaces as

Table 8.1 Comparison of typical operating parameters in SFA, STM and AFM/FFM used for micro/nanotribological studies

Operating parameter	SFA	STM ^a	AFM/FFM
Radius of mating surface/tip	~ 10 mm ^b	5–100 nm	5–100 nm
Radius of contact area	10–40 μm	N/A	0.05–0.5 nm
Normal load	10–100 mN	N/A	<0.1 nN–500 nN
Sliding velocity	0.001–100 μm/s	0.02–200 μm/s (scan size ~ 1 nm × 1 nm to 125 μm × 125 μm; scan rate <1–122 Hz)	0.02–200 μm/s (scan size ~ 1 nm × 1 nm to 125 μm × 125 μm; scan rate <1–122 Hz)
Sample limitations	Typically atomically smooth, optically transparent mica; opaque ceramic, smooth surfaces can also be used	Electrically conducting samples	None of the above

^aCan be used for atomic-scale imaging

^bSince stresses scale inverse of tip radius, SFA can provide very low stress measurement capabilities

well as of lubricant molecules. The introduction of the AFM in 1985 provided a method for measuring ultra-small forces between a probe tip and an engineering (electrically conducting or insulating) surface. AFM has been used for morphological and surface roughness measurements of surfaces on the nanoscale, as well as for adhesion measurements. Subsequent modifications of the AFM led to the development of the FFM, designed for atomic- and microscale studies of friction. This instrument measures forces in the scanning direction. The AFM is being used for various investigations including scratching, wear, indentation, detection of transfer of material, boundary lubrication, and fabrication and machining (Bhushan et al. 1995a; Bhushan 1999a, 2011, 2016). Meanwhile, significant progress in understanding the fundamental nature of bonding and interactions in materials, combined with advances in computer-based modeling and simulation methods, has allowed theoretical studies of complex interfacial phenomena with high resolution in space and time. Such simulations provide insights into atomic-scale energetics, structure, dynamics, thermodynamics, transport, and rheological aspects of tribological processes.

The nature of interactions between two surfaces brought together closely, and those between two contacting surfaces as they are separated, have been studied experimentally with the SFA. This has led to a basic understanding of the normal forces between surfaces and the way in which these are modified by the presence of a thin liquid or a polymer film. The frictional properties of such systems have been studied by moving the surfaces laterally. Such experiments have provided insights into the molecular-scale operation of lubricants such as thin liquid or polymer films. Complementary to these studies are those in which the AFM tip is used to simulate a single asperity contact with a solid or lubricated surface, Fig. 8.1. These experiments have demonstrated that the relationship between friction and surface roughness is not always simple or obvious. AFM studies have also revealed much about the nanoscale nature of intimate contact during wear, indentation, and lubrication.

In this chapter, we present a review of significant aspects of nanotribological, nanomechanical, and materials characterization studies conducted using AFM/FFM.

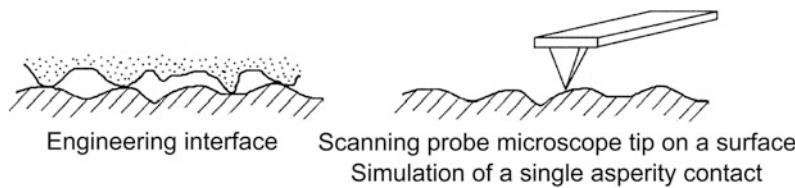


Fig. 8.1 Schematics of an engineering interface and scanning probe microscope tip in contact with an engineering interface

8.2 Description of AFM/FFM and Various Measurement Techniques

An AFM was developed by Gerd Binnig and his colleagues in 1985. It is capable of investigating surfaces of scientific and engineering interest on an atomic scale (Binnig et al. 1986, 1987). The AFM relies on a scanning technique to produce very high-resolution, three-dimensional images of sample surfaces. It measures ultra-small forces (less than 1 nN) present between the AFM tip surface mounted on a flexible cantilever beam and a sample surface. These small forces are obtained by measuring the motion of a very flexible cantilever beam having an ultrasmall mass, by a variety of measurement techniques including optical deflection, optical interference, capacitance, and tunneling current. The deflection can be measured to within 0.02 nm, so for a typical cantilever spring constant of 10 N/m, a force as low as 0.2 nN can be detected. To put these numbers in perspective, individual atoms and human hair are typically a fraction of a nanometer and about 75 μm in diameter, respectively, and a drop of water and an eyelash have a mass of about 10 and 100 nN, respectively. In the operation of high-resolution AFM, the sample is generally scanned rather than the tip because any cantilever movement would add vibrations. AFMs are available for the measurement of large samples, where the tip is scanned and the sample is stationary. To obtain an atomic resolution with the AFM, the spring constant of the cantilever should be weaker than the equivalent spring between atoms. A cantilever beam with a spring constant of about 1 N/m or lower is desirable. For high lateral

resolution, tips should be as sharp as possible. Tips with a radius ranging from 5 to 50 nm are commonly available. Interfacial forces, adhesion, and surface roughness, including atomic-scale imaging, are routinely measured using the AFM.

A modification to the AFM providing a sensor to measure the lateral force led to the development of the friction force microscope (FFM) or the lateral force microscope (LFM), designed for atomic-scale and microscale studies of friction (Mate et al. 1987; Bhushan and Ruan 1994; Ruan and Bhushan 1994a, b, c; Bhushan et al. 1994, 1995a; Bhushan and Kulkarni 1996; Bhushan 1997, 1999a, b, 2001a, b, 2011; Bhushan and Sundararajan 1998; Scherer et al. 1998, 1999; Reinstaedtler et al. 2003, 2005a, b; Bhushan and Kasai 2004; Tambe and Bhushan 2005a) and lubrication (Bhushan et al. 1995b, 2005, 2006, 2007; Koinkar and Bhushan 1996a, b; Bhushan and Liu 2001; Liu et al. 2001; Liu and Bhushan 2002, 2003a; Bhushan et al. 2005; Kasai et al. 2005; Lee et al. 2005; Tambe and Bhushan 2005h; Tao and Bhushan 2005a, b; Palacio and Bhushan 2007a, b). This instrument measures lateral or friction forces (in the plane of sample surface and in the scanning direction). By using a standard or a sharp diamond tip mounted on a stiff cantilever beam, AFM is used in investigations of scratching and wear (Bhushan et al. 1994, 1995a; Bhushan and Koinkar 1994a; Koinkar and Bhushan 1996c, 1997a; Bhushan 1999a, b, c, 2001c; Sundararajan and Bhushan 2001), nanoindentation (Ruan and Bhushan 1993; Bhushan et al. 1994, 1995a, 1996; Bhushan and Koinkar 1994b; Bhushan 1999d, 2001c; Li and Bhushan 2002; Bhushan and Li 2003), and nanofabrication/machining (Bhushan et al. 1994, 1995a; Bhushan 1999a). An

oscillating cantilever is used for localized surface elasticity and viscoelastic mapping, referred to as dynamic AFM (Maivald et al. 1991; Anczkowski et al. 1996; DeVecchio and Bhushan 1997; Scherer et al. 1997; Amelio et al. 2001; Scott and Bhushan 2003; Bhushan and Qi 2003; Kasai et al. 2004; Chen and Bhushan 2005; Reinstaedler et al. 2005b). In situ surface characterization of local deformation of materials and thin coatings has been carried out by imaging the sample surfaces using an AFM, during tensile deformation using a tensile stage (Bobji and Bhushan 2001a, b; Tambe and Bhushan 2004b).

8.2.1 Surface Roughness and Friction Force Measurements

Surface height imaging down to atomic resolution of electrically-conducting surfaces is carried out using an STM. An AFM is also used for surface height imaging and roughness characterization down to nanoscale. Commercial AFM/FFM are routinely used for simultaneous measurements of surface roughness and friction force (Bhushan 1999a, 2011). These instruments are available for the measurement of small samples and large samples. In a small sample AFM shown in

Fig. 8.2 Schematics **a** of a commercial small sample atomic force microscope/friction force microscope (AFM/FFM), and **b** of a large sample AFM/FFM

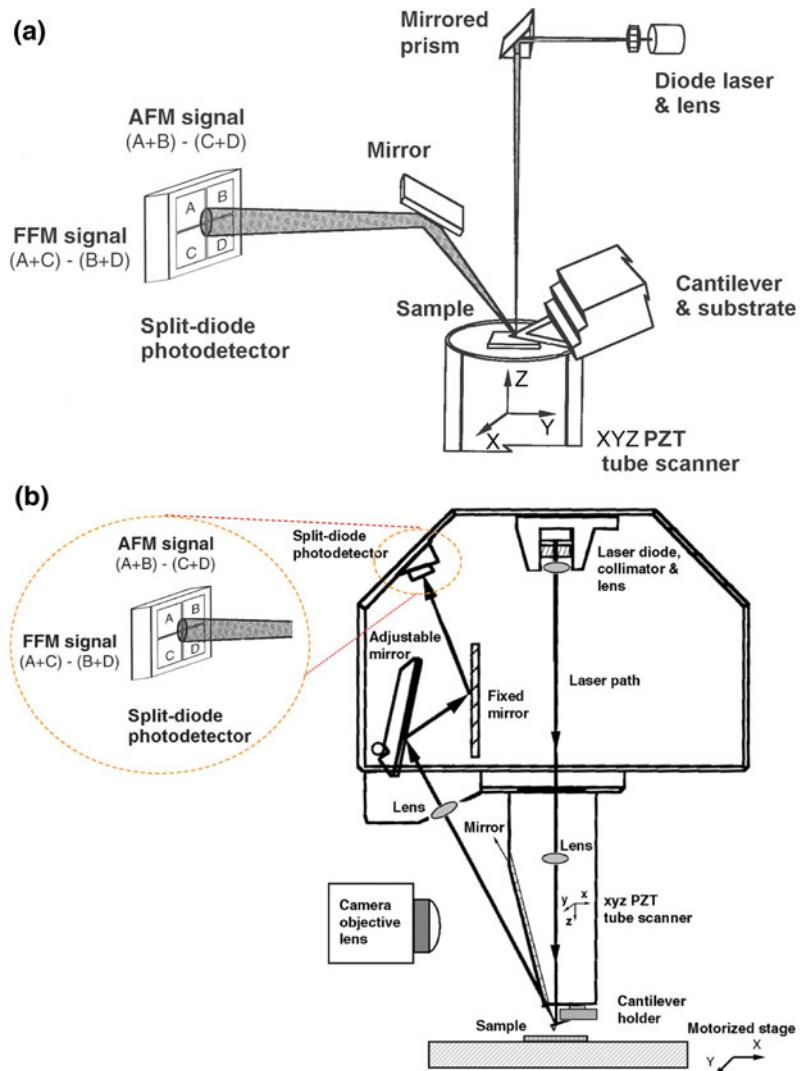


Fig. 8.2a, the sample, generally no larger than $10\text{ mm} \times 10\text{ mm}$, is mounted on a piezoelectric crystal in the form of a cylindrical tube (referred to as a PZT tube scanner) which consists of separate electrodes to scan the sample precisely in the X-Y plane in a raster pattern and to move the sample in the vertical (Z) direction. A sharp tip at the free end of a flexible cantilever is brought into contact with the sample. The normal and frictional forces being applied at the tip-sample interface are measured using a laser beam deflection technique. A laser beam from a diode laser is directed by a prism onto the back of a cantilever near its free end, tilted downward at about 10° with respect to the horizontal plane. The reflected beam from the vertex of the cantilever is directed through a mirror onto a quad photodetector (a split photodetector with four quadrants). The differential signal from the top and bottom photodiodes provides the AFM signal which is a sensitive measure of the cantilever vertical deflection. Topographic features of the sample cause the tip to deflect in the vertical direction as the sample is scanned under the tip. This tip deflection will change the direction of the reflected laser beam, changing the intensity difference between the top and bottom sets of photodetectors (the AFM signal). In the AFM operating mode called the height mode, for topographic imaging or for any other operation in which the applied normal force is to be kept constant, a feedback circuit is used to modulate the voltage applied to the PZT scanner to adjust the height of the PZT, so that the cantilever vertical deflection (given by the intensity difference between the top and bottom detector) will remain constant during scanning. The PZT height variation is thus a direct measure of the surface roughness of the sample.

In a large sample AFM, both force sensors using optical deflection method and scanning unit are mounted on the microscope head, Fig. 8.2b. Because of vibrations caused by the cantilever movement, lateral resolution of this design can be somewhat poorer than the design in Fig. 8.2a in which the sample is scanned instead of cantilever beam. The advantage of the large sample AFM is that large samples can be measured readily.

Most AFMs can be used for surface roughness measurements in the so-called tapping mode (intermittent contact mode), also referred to as dynamic (atomic) force microscopy. In the tapping mode, during scanning over the surface, the cantilever/tip assembly with a normal stiffness of 20–100 N/m (DI tapping mode etched Si probe or TESP) is sinusoidally vibrated at its resonance frequency (350–400 kHz) by a piezo mounted above it, and the oscillating tip slightly taps the surface. The piezo is adjusted using the feedback control in the Z direction to maintain a constant (20–100 nm) oscillating amplitude (setpoint) and constant average normal force, Fig. 8.3 (Bhushan 1999a, 2011). The feedback signal to the Z-direction sample piezo (to keep the setpoint constant) is a measure of surface roughness. The cantilever/tip assembly is vibrated at some amplitude, here referred to as the free amplitude, before the tip engages the sample. The tip engages the sample at some setpoint, which may be thought of as the amplitude of the cantilever as influenced by contact with the sample. The setpoint is defined as a ratio of the vibration amplitude after engagement to the vibration amplitude in free air before engagement. A lower setpoint gives a reduced amplitude and closer mean tip-to-sample distance. The amplitude should be kept large enough so that the tip does not get stuck to the sample because of adhesive attractions. Also the oscillating amplitude applies less average (normal) load as compared to the contact mode and reduces the sample damage. The tapping mode is used in topography measurements to minimize the effects of friction and other lateral forces and to measure topography of soft surfaces.

To measure the friction force at the tip surface during sliding, the left hand and right hand sets of quadrants of the photodetector are used. In the so-called friction mode, the sample is scanned back and forth in a direction orthogonal to the long axis of the cantilever beam. A friction force between the sample and the tip will produce a twisting of the cantilever. As a result, the laser beam will be reflected out of the plane defined by the incident beam and the beam reflected vertically from an untwisted cantilever. This produces

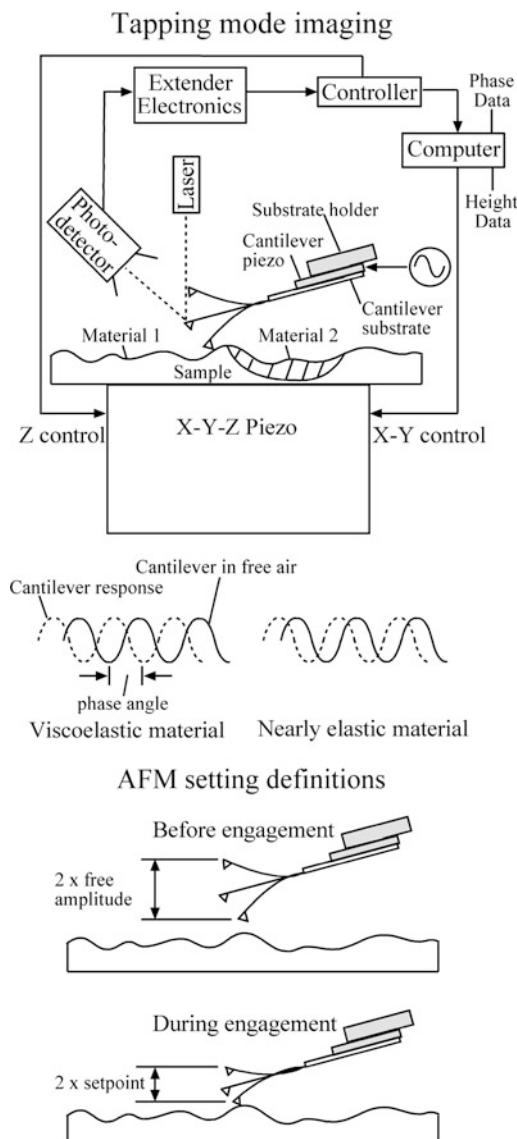


Fig. 8.3 Schematic of tapping mode used to obtain height and phase data and definitions of free amplitude and setpoint. During scanning, the cantilever is vibrated at its resonance frequency and the sample X-Y-Z piezo is adjusted by feedback control in the Z-direction to maintain a constant setpoint. The computer records height (which is a measure of surface roughness) and phase angle (which is a function of the viscoelastic properties of the sample) data

an intensity difference of the laser beam received in the left hand and right hand sets of quadrants of the photodetector. The intensity difference between the two sets of detectors (FFM signal) is

directly related to the degree of twisting and hence to the magnitude of the friction force. One problem associated with this method is that any misalignment between the laser beam and the photodetector axis would introduce error in the measurement. However, by following the procedures developed by Ruan and Bhushan (1994a), in which the average FFM signal for the sample scanned in two opposite directions is subtracted from the friction profiles of each of the two scans, the misalignment effect is eliminated. This method provides three-dimensional maps of friction force. By following the friction force calibration procedures developed by Ruan and Bhushan (1994a), voltages corresponding to friction forces can be converted to force units (Palacio and Bhushan 2010). The coefficient of friction is obtained from the slope of friction force data measured as a function of normal loads typically ranging from 10 to 150 nN. This approach eliminates any contributions due to the adhesive forces (Bhushan et al. 1994). To calculate the coefficient of friction based on a single point measurement, friction force should be divided by the sum of the applied normal load and intrinsic adhesive force. Furthermore it should be pointed out that for a single asperity contact, the coefficient of friction is not independent of load (see discussion later).

Surface roughness measurements in the contact mode are typically made using a sharp, microfabricated square-pyramidal Si_3N_4 tip with a radius of 30–50 nm on a triangular cantilever beam (Fig. 8.4a) with normal stiffness on the order of 0.06–0.58 N/m with a normal natural frequency of 13–40 kHz (DI silicon nitride probe or NP) at a normal load of about 10 nN, and friction measurements are carried out in the load range of 1–100 nN. Surface roughness measurements in the tapping mode utilize a stiff cantilever with high resonance frequency; typically a square-pyramidal etched single-crystal silicon tip, with a tip radius of 5–10 nm, integrated with a stiff rectangular silicon cantilever beam (Fig. 8.4a) with a normal stiffness on the order of 17–60 N/m and a normal resonance frequency of 250–400 kHz (DI TESP), is used. Multiwalled carbon nanotube tips having a small

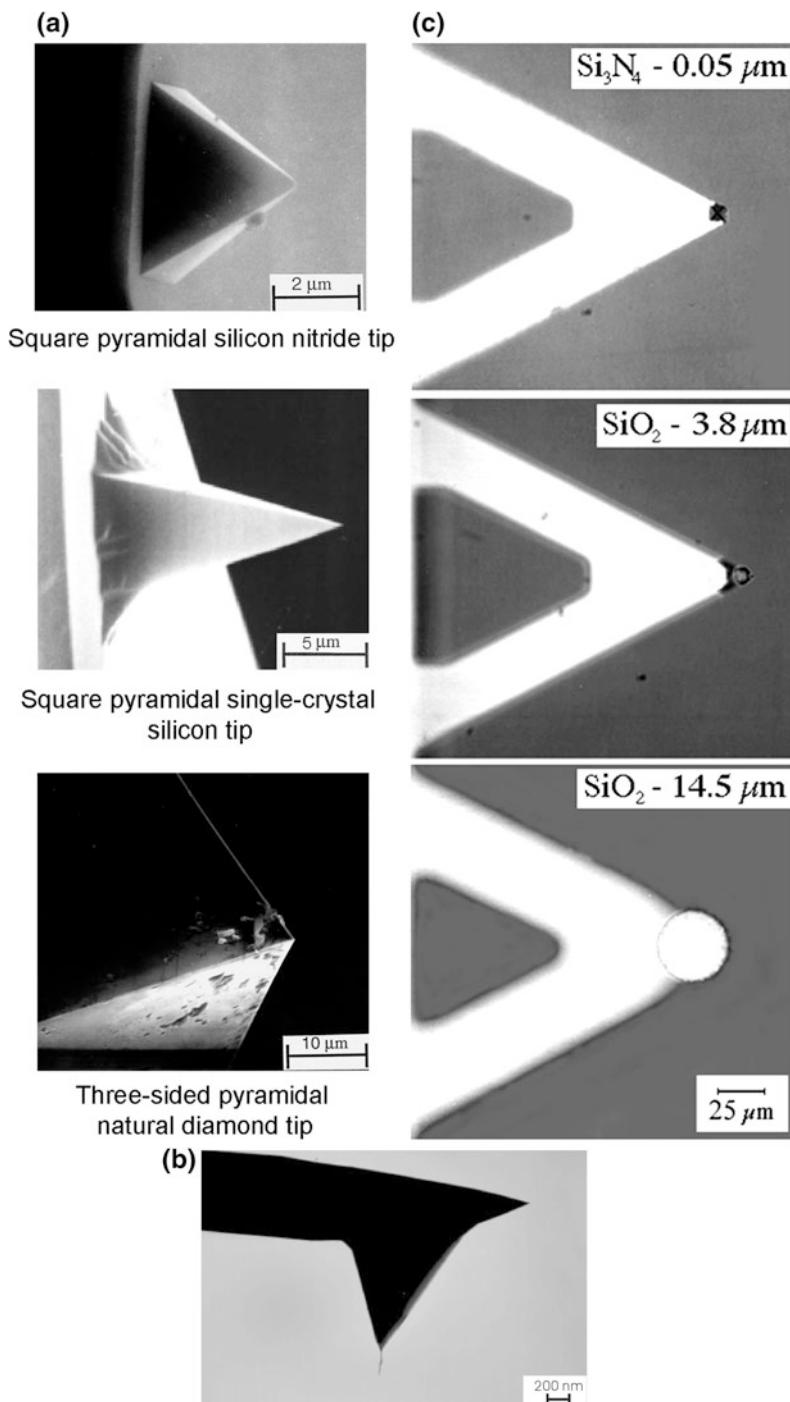


Fig. 8.4 **a** SEM micrographs of a square-pyramidal plasma-enhanced chemical vapor deposition (PECVD) Si_3N_4 tip with a triangular cantilever beam, a square-pyramidal etched single-crystal silicon tip with a rectangular silicon cantilever beam, and a three-sided pyramidal natural diamond tip with a square stainless steel cantilever beam, **b** SEM micrograph of a multi-walled carbon nanotube (MWNT) physically attached on the single-crystal silicon, square pyramidal tip, and **c** optical micrographs of commercial Si_3N_4 tip and two modified tips showing SiO_2 spheres mounted over the sharp tip, at the end of the triangular Si_3N_4 cantilever beams (radii of the tips are given in the figure)

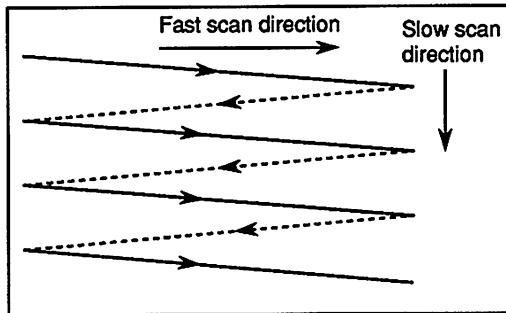


Fig. 8.5 Schematic of triangular pattern trajectory of the tip as the sample (or the tip) is scanned in two dimensions. During scanning, data are recorded only during scans along the solid scan lines

diameter (few nm) and a length of about 1 μm (high aspect ratio) attached on the single-crystal silicon, square pyramidal tips are used for the high resolution imaging of surfaces and of deep trenches in the tapping mode (noncontact mode) (Fig. 8.4b) (Bhushan et al. 2004a). The MWNT tips are hydrophobic. To study the effect of radius of a single asperity (tip) on adhesion and friction, microspheres of silica with radii ranging from about 4 to 15 μm are attached at the end of cantilever beams. Optical micrographs of two of the microspheres at the ends of triangular cantilever beams are shown in Fig. 8.4c.

The tip is scanned in such a way that its trajectory on the sample forms a triangular pattern, Fig. 8.5. Scanning speeds in the fast and slow scan directions depend on the scan area and scan frequency. Scan sizes ranging from less than 1 $\text{nm} \times 1 \text{ nm}$ to 125 $\mu\text{m} \times 125 \mu\text{m}$ and scan rates from less than 0.5–122 Hz typically can be used. Higher scan rates are used for smaller scan lengths. For example, scan rates in the fast and slow scan directions for an area of 10 $\mu\text{m} \times 10 \mu\text{m}$ scanned at 0.5 Hz are 10 and 20 nm/s, respectively.

8.2.2 Adhesion Measurements

Adhesive force measurements are performed in the so-called force calibration mode. In this

mode, force-distance curves are obtained, for an example see Fig. 8.6. The horizontal axis gives the distance that the piezo (and hence the sample) travels, and the vertical axis gives the tip deflection. As the piezo extends, it approaches the tip, which is at this point in free air and hence shows no deflection. This is indicated by the flat portion of the curve. As the tip approaches the sample within a few nanometers (point A), an attractive force exists between the atoms of the tip surface and the atoms of the sample surface. The tip is pulled towards the sample and contact occurs at point B on the graph. From this point on, the tip is in contact with the surface and as the piezo further extends, the tip gets further deflected. This is represented by the sloped portion of the curve. As the piezo retracts, the tip goes beyond the zero deflection (flat) line because of attractive forces (van der Waals forces and long-range meniscus forces), into the adhesive regime. At point C in the graph, the tip snaps free of the adhesive forces and is again in free air. The horizontal distance between points B and C along the retrace line gives the distance moved by the tip in the adhesive regime. This distance multiplied by the stiffness of the cantilever gives the adhesive force. Incidentally, the horizontal shift between the loading and unloading curves results from the hysteresis in the PZT tube (Bhushan 1999a, 2011).

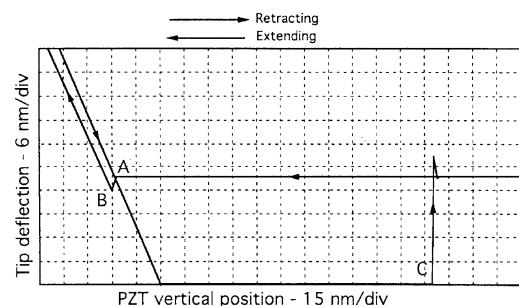


Fig. 8.6 Typical force-distance curve for a contact between Si_3N_4 tip and single-crystal silicon surface in measurements made in the ambient environment. Snap-in occurs at point A; contact between the tip and silicon occurs at point B; tip breaks free of adhesive forces at point C as the sample moves away from the tip

8.2.3 Scratching, Wear and Fabrication/ Machining

For microscale scratching, microscale wear, nanofabrication/nanomachining, and nanoindentation hardness measurements, an extremely hard tip is required. A three-sided pyramidal single-crystal natural diamond tip with an apex angle of 80° and a radius of about 100 nm mounted on a stainless steel cantilever beam with normal stiffness of about 25 N/m is used at relatively higher loads (1–150 μN), Fig. 8.4a. For scratching and wear studies, the sample is generally scanned in a direction orthogonal to the long axis of the cantilever beam (typically at a rate of 0.5 Hz) so that friction can be measured during scratching and wear. The tip is mounted on the cantilever such that one of its edges is orthogonal to the long axis of the beam; therefore, wear during scanning along the beam axis is higher (about 2x to 3) than that during scanning orthogonal to the beam axis. For wear studies, an area on the order of $2 \times 2 \mu\text{m}$ is scanned at various normal loads (ranging from 1 to 100 μN) for a selected number of cycles (Bhushan et al. 1994; Bhushan 1999a, 2011).

Scratching can also be performed at ramped loads and the coefficient of friction can be measured during scratching (Sundararajan and Bhushan 2001). A linear increase in the normal load approximated by a large number of normal load increments of small magnitude is applied using a software interface (lithography module in Nanoscope III) that allows the user to generate controlled movement of the tip with respect to the sample. The friction signal is tapped out of the AFM and is recorded on a computer. A scratch length on the order of 25 μm and a velocity on the order of 0.5 $\mu\text{m/s}$ are used and the number of loading steps is usually taken to be 50.

Nanofabrication/nanomachining is conducted by scratching the sample surface with a diamond tip at specified locations and scratching angles. The normal load used for scratching (writing) is on the order of 1–100 μN with a writing speed on the order of 0.1–200 $\mu\text{m/s}$ (Bhushan et al. 1994, 1995a; Bhushan 1995, 1999a, c, 2011).

8.2.4 Surface Potential Measurements

Various electrical measurements on the nanoscale can be made using a conducting tip. The conducting tip acts as a voltage electrode to provide stimuli and monitor surface electrical properties (Avila and Bhushan 2010). Various electrical techniques include scanning Kelvin probe microscopy (SKPM), scanning capacitance microscopy (SCM), and scanning spreading resistance microscopy (SSRM) (Bhushan and Goldade 2000b; Lee et al. 2002; Palacio and Bhushan 2007a).

Kelvin probe microscopy can be used to detect wear precursors and to study the early stages of localized wear (DeVecchio and Bhushan 1998; Bhushan and Goldade 2000a, b). In this technique, the potential difference is measured between the tip and the sample by applying a DC bias potential and an oscillating (AC) potential to a conducting tip over a grounded substrate, also a so-called “nano-Kelvin probe” technique. Mapping of the surface potential is made in the so-called “lift mode” (Fig. 8.7). These measurements are made simultaneously with the topography scan in the tapping mode, using an electrically-conducting (nickel-coated single-crystal silicon) tip. After each line of the topography scan is completed, the feedback loop controlling the vertical piezo is turned off, and the tip is lifted from the surface and traced over the same topography at a constant distance of 100 nm. During the lift mode, a DC bias potential and an oscillating potential (3–7 V) is applied to the tip. The frequency of oscillation is chosen to be equal to the resonance frequency of the cantilever ($\sim 80 \text{ kHz}$). When a DC bias potential equal to the negative value of surface potential of the sample (on the order of $\pm 2 \text{ V}$) is applied to the tip, it does not vibrate. During scanning, a difference between the DC bias potential applied to the tip and the potential of the surface will create DC electric fields that interact with the oscillating charges (as a result of the AC potential), causing the cantilever to oscillate at its resonance frequency, as in tapping mode. However, a feedback loop is used to adjust the DC bias on the tip to exactly nullify the electric field, and

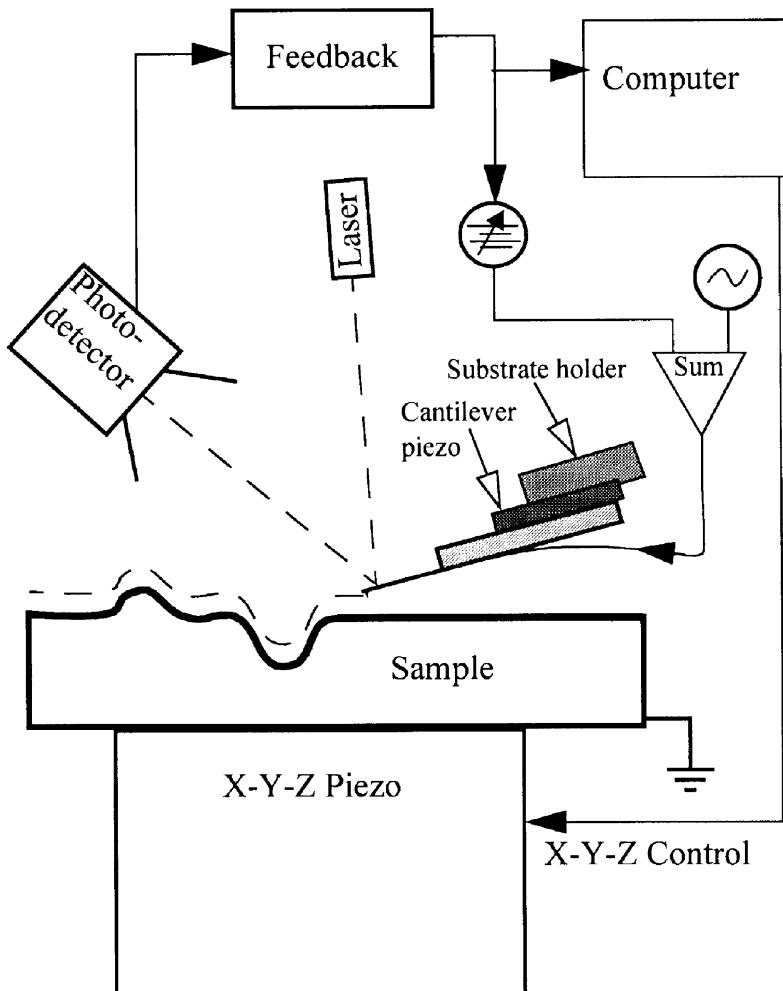


Fig. 8.7 Schematic of lift mode used to make surface potential measurement. The topography is collected in tapping mode in the primary scan. The cantilever piezo is de-activated. Using topography information of the primary scan, the cantilever is scanned across the surface at a constant height above the sample. An oscillating voltage at the resonant frequency is applied to the tip and a feedback loop adjusts the DC bias of the tip to maintain the cantilever amplitude at zero. The output of the feedback loop is recorded by the computer and becomes the surface potential map

thus the vibrations of the cantilever. The required bias voltage follows the localized potential of the surface. The surface potential was obtained by reversing the sign of the bias potential provided by the electronics (Bhushan and Goldade 2000a, b).

Surface and subsurface changes of structure and/or chemistry can cause changes in the measured potential of a surface. Thus, mapping of the surface potential after sliding can be used for detecting wear precursors and studying the early

stages of localized wear (DeVecchio and Bhushan 1998; Bhushan and Goldade 2009a, b).

8.2.5 In Situ Characterization of Local Deformation Studies

In situ characterization of local deformation of materials can be carried out by performing tensile, bending, or compression experiments inside

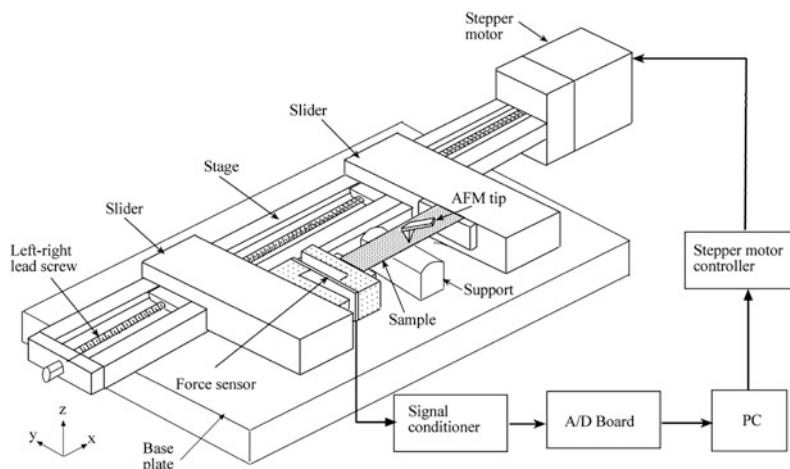


Fig. 8.8 Schematic of the tensile stage to conduct in situ tensile testing of the polymeric films in AFM

an AFM and by observing nanoscale changes during the deformation experiment (Bhushan 1999d). In these experiments, small deformation stages are used to deform the samples inside an AFM. For tensile testing of the polymeric films carried out by Bobji and Bhushan (2001a, b) and Tambe and Bhushan (2004b) and of the hair by Seshadri and Bhushan (2008a, b), a tensile stage was used (Fig. 8.8). The stage with a left-right combination lead screw (that helps to move the slider in the opposite direction) was used to stretch the sample to minimize the movement of the scanning area, which was kept close to the center of the tensile specimen. One end of the sample was mounted on the slider via a force sensor to monitor the tensile load. The samples were stretched for various strains using a stepper motor and the same control area at different strains was imaged. In order to better locate the control area for imaging, a set of four markers was created at the corners of a $30 \times 30 \mu\text{m}$ square at the center of the sample by scratching the sample with a sharp silicon tip. The scratching depth was controlled such that it did not affect cracking behavior of the coating. A minimum displacement of $1.6 \mu\text{m}$ could be obtained. This corresponded to a strain increment of $8 \times 10^{-3}\%$ for a sample length of 38 mm. The maximum travel distance was about

100 mm. The resolution of the force sensor was 10 mN with a capacity of 45 N. During stretching, a stress-strain curve was obtained during the experiment to study any correlation between the degree of plastic strain and propensity of cracks.

8.2.6 Nanoindentation Measurements

For nanoindentation hardness measurements the scan size is set to zero, and then a normal load is applied to make the indents using the diamond tip. During this procedure, the tip is continuously pressed against the sample surface for about two seconds at various indentation loads. The sample surface is scanned before and after the scratching, wear, or indentation to obtain the initial and the final surface topography, at a low normal load of about $0.3 \mu\text{N}$ using the same diamond tip. An area larger than the indentation region is scanned to observe the indentation marks. Nanohardness is calculated by dividing the indentation load by the projected residual area of the indents (Bhushan and Koinkar 1994b).

Direct imaging of the indent allows one to quantify piling up of ductile material around the indenter. However, it becomes difficult to identify the boundary of the indentation mark with great

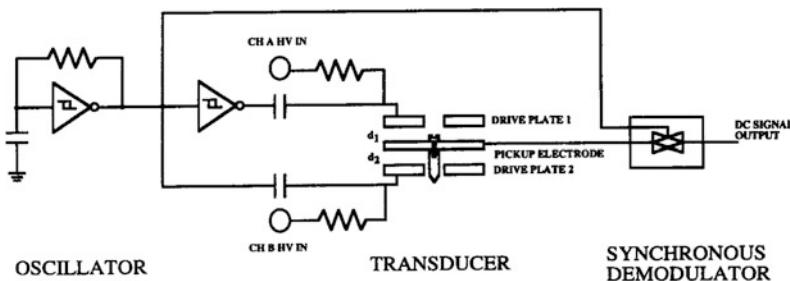


Fig. 8.9 Schematic of a nano/picoindentation system with three-plate transducer with electrostatic actuation hardware and capacitance sensor (Bhushan et al. 1996)

accuracy. This makes the direct measurement of contact area somewhat inaccurate. A technique with the dual capability of depth-sensing as well as in situ imaging, which is most appropriate in nanomechanical property studies, is used for accurate measurement of hardness with shallow depths (Bhushan et al. 1996; Bhushan 1999a, 2011). This nano/picoindentation system is used to make load-displacement measurement and subsequently carry out in situ imaging of the indent, if required. The indentation system, shown in Fig. 8.9, consists of a three-plate transducer with electrostatic actuation hardware used for the direct application of a normal load and a capacitive sensor used for the measurement of vertical displacement. The AFM head is replaced with this transducer assembly while the specimen is mounted on the PZT scanner, which remains stationary during indentation experiments. The transducer consists of a three (Be-Cu) plate capacitive structure, and the tip is mounted on the center plate. The upper and lower plates serve as drive electrodes, and the load is applied by applying appropriate voltage to the drive electrodes. Vertical displacement of the tip (indentation depth) is measured by measuring the displacement of the center plate relative to the two outer electrodes using the capacitance technique. Indent area and consequently hardness value can be obtained from the load-displacement data. The Young's modulus of elasticity is obtained from the slope of the unloading curve.

8.2.7 Localized Surface Elasticity and Viscoelasticity Mapping

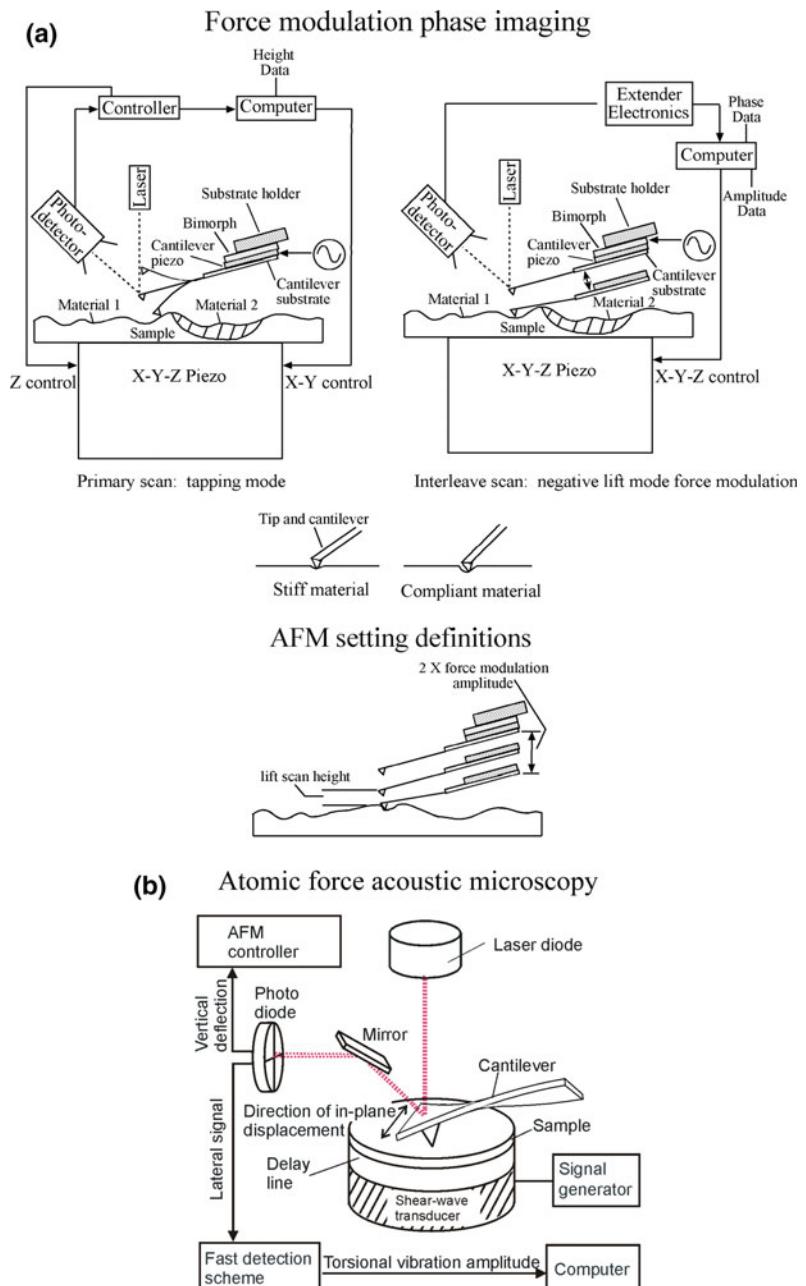
8.2.7.1 Localized Surface Elasticity

Indentation experiments provide a single-point measurement of the Young's modulus of elasticity calculated from the slope of the indentation curve during unloading. Localized surface elasticity maps can be obtained using dynamic force microscopy, in which an oscillating tip is scanned over the sample surface in contact under steady and oscillating load. Lower frequency operation mode in the kHz range, such as force modulation mode (Maivald et al. 1991; DeVecchio and Bhushan 1997) or pulsed force mode (Krotil et al. 1999), are well suited for soft samples such as polymers. However, if the tip-sample contact stiffness becomes significantly higher than the cantilever stiffness, the sensitivity of these techniques strongly decreases. In this case, sensitivity of the measurement of stiff materials can be improved by using high-frequency operation modes in the MHz range with a lateral motion, such as acoustic (ultrasonic) force microscopy, referred to as atomic force acoustic microscopy (AFAM) or contact resonance spectroscopy (Rabe et al. 1996; Scherer et al. 1997; Amelio et al. 2001). Inclusion of vibration frequencies other than only the first cantilever flexural or torsional resonance frequency also allows additional information to be obtained.

In negative lift mode force modulation technique, during primary scanning height data is recorded in tapping mode as described earlier. During interleave scanning, the entire cantilever/tip assembly is moved up and down at the force modulation holder's bimorph resonance frequency (about 24 kHz) at some amplitude, here

referred to as the force modulation amplitude, and the Z-direction feedback control for the sample X-Y-Z piezo is deactivated, Fig. 8.10a (Maivald et al. 1991; DeVecchio and Bhushan 1997; Scott and Bhushan 2003). During this scanning, height information from the primary scan is used to maintain a constant lift scan height. This

Fig. 8.10 **a** Schematic of force modulation mode used to obtain amplitude (stiffness) and definitions of force modulation amplitude and lift scan height. During primary scanning, height data is recorded in tapping mode. During interleave scanning, the entire cantilever/tip assembly is vibrated at the bimorph's resonance frequency and the Z-direction feedback control for the sample X-Y-Z piezo is deactivated. During this scanning, height information from the primary scan is used to maintain a constant lift scan height. The computer records amplitude (which is a function of material stiffness) during the interleave scan, and **b** Schematic of an AFM incorporating shear wave transducer which generates in-plane lateral sample surface vibrations. Because of the forces between the tip and the surface, torsional vibrations of the cantilever are excited (Reinstaedtler et al. 2003). The shift in contact resonance frequency is a measure of contact stiffness



eliminates the influence of height on the measured signals during the interleave scan. Lift scan height is the mean tip-to-sample distance between the tip and sample during the interleave scan. The lift scan height is set such that the tip is in constant contact with the sample, i.e. a constant static load is applied. (A higher lift scan height gives a closer mean tip-to-sample distance.) In addition, the tip motion caused by the bimorph vibration results in a modulating periodic force. The sample surface resists the oscillations of the tip to a greater or lesser extent depending upon the sample's stiffness. The computer records amplitude (which is a function of elastic stiffness of the material). Contact analyses can be used to obtain a quantitative measure of localized elasticity of soft surfaces (DeVecchio and Bhushan 1997). Etched single-crystal silicon cantilevers with integrated tips (DI force modulation etched Si probe or FESP) with a radius of 25–50 nm, a stiffness of 1–5 N/m, and a natural frequency of 60–100 kHz, are commonly used for the measurements. Scanning is normally set to a rate of 0.5 Hz along the fast axis.

In the AFAM technique (Rabe et al. 1996; Scherer et al. 1997; Amelio et al. 2001), the cantilever/tip assembly is moved either in the normal or lateral mode, and the contact stiffness is evaluated by comparing the resonance frequency of the cantilever in contact with the sample surface to those of the free vibrations of the cantilever. Several free resonance frequencies are measured. Based on the shift of the measured frequencies, the contact stiffness is determined by solving the characteristic equation for the tip vibrating in contact with the sample surface. The elastic modulus is calculated from contact stiffness using Hertz analysis for a spherical tip indenting a plane. Contact stiffness is equal to $8x$ contact radius x reduced shear modulus in shear mode.

In the lateral mode using the AFAM technique, the sample is glued on cylindrical pieces of aluminum which serves as ultrasonic delay lines coupled to an ultrasonic shear wave transducer, Fig. 8.10b (Scherer et al. 1997; Amelio

et al. 2001; Reinstaedtler et al. 2003). The transducer is driven with frequency sweeps to generate in-plane lateral sample surface vibrations. These couple to the cantilever via the tip-sample contact. To measure torsional vibrations of the cantilever at frequencies up to 3 MHz, the original electronic circuit of the lateral channel of the AFM (using a lowpass filter with limited bandwidth to few hundred kHz) was replaced by a high-speed scheme which bypasses the lowpass filter. The high frequency signal was fed to a lock-in amplifier, digitized using a fast A/D card, and fed into a broadband amplifier, followed by an rms-to-dc converter and read by a computer. Etched single-crystal silicon cantilevers (normal stiffness of 3.8–40 N/m) integrated tips are used.

8.2.7.2 Viscoelastic Mapping

Another form of dynamic force microscopy, phase contrast microscopy, is used to detect the contrast in viscoelastic (viscous energy dissipation) properties of the different materials across the surface (Anczykowski et al. 1996; Tamayo and Garcia 1996; Garcia et al. 1998; Bhushan and Qi 2003; Scott and Bhushan 2003; Kasai et al. 2004; Chen and Bhushan 2005). In these techniques, both deflection amplitude and phase angle contrasts are measured, which are measures of the relative stiffness and viscoelastic properties, respectively. Two phase measurement techniques—tapping mode and torsional resonance (TR) mode—have been developed. We describe them next.

In the tapping mode (TM) technique, as described earlier, the cantilever/tip assembly is sinusoidally vibrated at its resonant frequency, and the sample X-Y-Z piezo is adjusted using feedback control in the z-direction to maintain a constant setpoint, Fig. 8.3 (Bhushan and Qi 2003; Scott and Bhushan 2003). The feedback signal to the Z-direction sample piezo (to keep the setpoint constant) is a measure of surface roughness. The extender electronics is used to measure the phase angle lag between the cantilever piezo drive signal and the cantilever

response during sample engagement. As illustrated in Fig. 8.3, the phase angle lag (at least partially) is a function of the viscoelastic properties of the sample material. A range of tapping amplitudes and setpoints can be used for measurements. Commercially etched single-crystal silicon tip (DI TESP) used for tapping mode, with a radius of 5–10 nm, a stiffness of 20–100 N/m, and a natural frequency of 350–400 kHz, is normally used. Scanning is normally set to a rate of 1 Hz along the fast axis.

In the torsional mode (TR mode), a tip is vibrated in the torsional mode at high frequency at the resonance frequency of the cantilever beam. An etched single-crystal silicon cantilever with integrated tip (DI FESP) with a radius of about 5–10 nm, normal stiffness of 1–5 N/m, torsional stiffness of about 30 times normal stiffness and torsional natural frequency of 800 kHz is normally used. A major difference between the TM and TR modes is the directionality of the applied oscillation—a normal (compressive) amplitude exerted for the TM and a torsional amplitude for the TR mode. The TR mode is expected to provide good contrast in the tribological and mechanical properties of the near surface region as compared to the TM. Two of the reasons are as follows. (1) In the TM, the interaction is dominated by the vertical properties of the sample, so the tip spends a small fraction of its time in the near field interaction with the sample. Furthermore, the distance between the tip and the sample changes during the measurements, which changes interaction time and forces, and affects measured data. In the TR mode, the distance remains nearly constant. (2) The lateral stiffness of a cantilever is typically about two orders of magnitude larger than the normal (flexural) stiffness. Therefore, in the TM, if the sample is relatively rigid, much of the deformation occurs in the cantilever beam, whereas in the TR mode, much of the deformation occurs in the sample. A few comments on the special applications of the TR mode are made next. Since most of the deformation occurs in the sample, the TR mode can be used to measure stiff and hard samples. Furthermore, properties of thin

films can be measured more readily with the TR mode. For both the TM and TR modes, if the cantilever is driven to vibrate at frequencies above resonance, it would have less motion (high apparent stiffness), leading to higher sample deformation and better contrast. It should be further noted that the TM exerts a compressive force, whereas the TR mode exerts torsional force, therefore normal and shear properties are measured in the TM and TR modes, respectively.

In the TR mode, the torsional vibration of the cantilever beam is achieved using a specially designed cantilever holder. It is equipped with a piezo system mounted in a cantilever holder, in which two piezos vibrate out-of-phase with respect to each other. A tuning process prior to scanning is used to select the torsional vibration frequency. The piezo system excites torsional vibration at the cantilever's resonance frequency. The torsional vibration amplitude of the tip (TR amplitude) is detected by the lateral segments of the split-diode photodetector, Fig. 8.11 (Kasai et al. 2004). The TR mode measures surface roughness and phase angle as follows. During the measurement, the cantilever/tip assembly is first vibrated at its resonance at some amplitude dependent upon the excitation voltage, before the tip engages the sample. Next, the tip engages the sample at some setpoint. A feedback system coupled to a piezo stage is used to keep a constant TR amplitude during scanning. This is done by controlling the vertical position of the sample using a piezo moving in the Z direction, which changes the degree of tip interaction. The displacement of the sample Z piezo gives a roughness image of the sample. A phase angle image can be obtained by measuring the phase lag of the cantilever vibration response in the torsional mode during engagement with respect to the cantilever vibration response in free air before engagement. The control feedback of the TR mode is similar to that of tapping, except that the torsional resonance amplitude replaces flexural resonance amplitude (Kasai et al. 2004).

Chen and Bhushan (2005) used a variation to the approach just described (referred to as mode I

here). They performed measurements at constant normal cantilever deflection (constant load) (mode II) instead of using constant setpoint in the Kasai et al. (2004) approach. Their approach

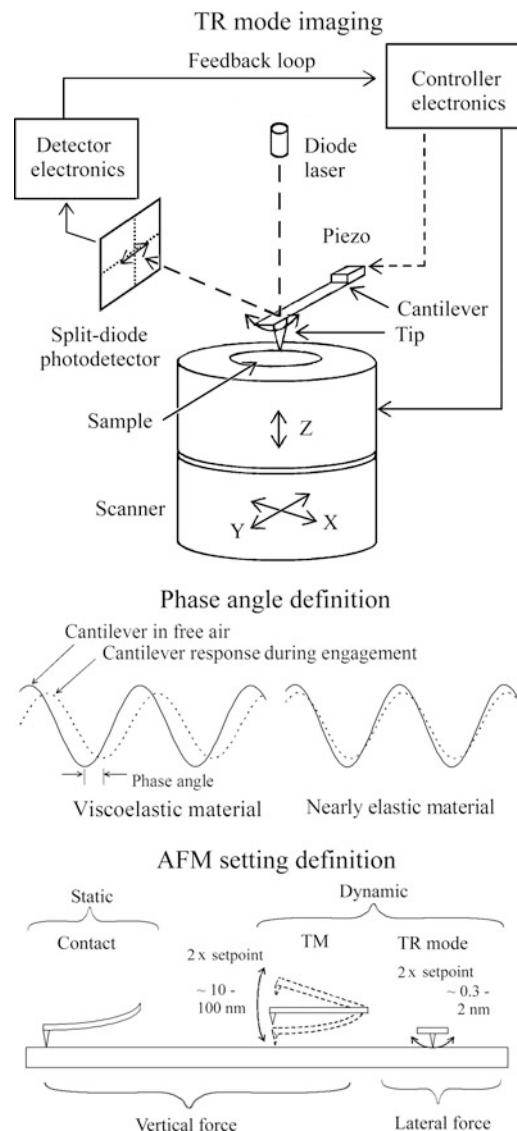


Fig. 8.11 Schematic of torsional resonance mode shown at the *top*. Two examples of the phase angle response are shown in the *middle*. One is for materials exhibiting viscoelastic (*left*) and the other nearly elastic properties (*right*). Three AFM settings are compared at the *bottom* contact, tapping mode (TM), and TR modes. The TR mode is a dynamic approach with a laterally vibrating cantilever tip that can interact with the surface more intensively than other modes. Therefore, more detailed near surface information is available

overcomes the meniscus adhesion problem present in mode I and reveals true surface properties.

Song and Bhushan (2005) presented a forced torsional vibration model for a tip-cantilever assembly under viscoelastic tip-sample interaction. This model provides a relationship of torsional amplitude and phase shift with lateral contact stiffness and viscosity which can be used to extract in-plane interfacial mechanical properties.

Various operating modes of AFM used for surface roughness, localized surface elasticity, and viscoelastic mapping and friction force measurements (to be discussed later) are summarized in Table 8.2.

8.2.8 Boundary Lubrication Measurements

To study nanoscale boundary lubrication properties, adhesive forces are measured in the force calibration mode, as previously described. The adhesive forces are also calculated from the horizontal intercept of friction versus normal load curves at a zero value of friction force. For friction measurements, the samples are typically scanned using a Si_3N_4 tip over an area of $2 \times 2 \mu\text{m}$ at the normal load ranging from 5 to 130 nN. The samples are generally scanned with a scan rate of 0.5 Hz resulting in a scanning speed of 2 $\mu\text{m}/\text{s}$. Velocity effects on friction are studied by changing the scan frequency from 0.1 to 60 Hz, while the scan size is maintained at $2 \times 2 \mu\text{m}$, which allows velocity to vary from 0.4 to 240 $\mu\text{m}/\text{s}$. To study the durability properties, the friction force and coefficient of friction are monitored during scanning at a normal load of 70 nN and a scanning speed of 0.8 $\mu\text{m}/\text{s}$, for a desired number of cycles (Koinkar and Bhushan 1996a, b; Liu and Bhushan 2003a).

8.2.9 AFM Tip Wear

AFM tips wear during use. Wear can result in an increase in tip radius and adhesion between the tip and the sample, thus reducing the image resolution

Table 8.2 Summary of various operating modes of AFM for surface roughness, stiffness, phase angle, and friction

Operating mode	Direction of cantilever vibration	Vibration frequency of cantilever (kHz)	Vibration amplitude (nm)	Feedback control	Data obtained
Contact	n/a			Constant normal load	Surface height, friction
Tapping	Vertical	350–400	10–100	Setpoint (constant tip amplitude)	Surface height, phase angle (normal viscoelasticity)
Force modulation	Vertical	10–20 (bimorph)	10–100	Constant normal load	Surface height, amplitude (normal stiffness)
Lateral	Lateral (AAFM)	100–3000 (sample)	~5 (sample)	Constant normal load	Shift in contact resonance (normal stiffness, friction)
TR mode I	Torsional	~800	0.3–2	Setpoint (constant tip amplitude)	Surface height, phase angle (lateral viscoelasticity)
TR mode II	Torsional	~800	0.3–2	Constant normal load	Surface height, amplitude and phase angle (lateral stiffness and lateral viscoelasticity)
TR mode III	Torsional	>800 in contact	0.3–2	Constant normal load	Shift in contact resonance (friction)

and introducing artifacts. The tip geometry can be measured by AFM imaging using an ultrasharp geometry (sharper than the AFM tip), known as tip characterizer, and analyzing the data for tip reconstruction using a tip deconvolution algorithm (Tao and Bhushan 2006a). A grating sample with an array of sharp tips with a radius less than 10 nm is commonly used.

Tao and Bhushan (2006a) developed fluoropolymer coatings deposited on the silicon probe to reduce their tip wear.

8.3 Surface Imaging, Friction and Adhesion

8.3.1 Atomic-Scale Imaging and Friction

Surface height imaging down to atomic resolution of electrically conducting surfaces can be carried out using an STM. An AFM can also be used for surface height imaging and roughness characterization down to the nanoscale.

Figure 8.12 shows a sequence of STM images at various scan sizes of solvent deposited C₆₀ film on 200-nm thick gold coated freshly cleaved mica (Bhushan et al. 1993). The film consists of clusters of C₆₀ molecules of 8 nm in diameter. The C₆₀ molecules within a cluster appear to pack into a hexagonal array with a spacing of about 1 nm, however, they do not follow any long range order. The measured cage diameter of the C₆₀ molecule is about 0.7 nm, very close to the projected diameter of 0.71 nm.

In an AFM measurement during surface imaging, the tip comes in intimate contact with the sample surface and leads to surface deformation with finite tip-sample contact area (typically few atoms). The finite size of the contact area prevents the imaging of individual point defects, and only the periodicity of the atomic lattice can be imaged. Figure 8.13a shows the topography image of a freshly-cleaved surface of highly oriented pyrolytic graphite (HOPG) (Ruan and Bhushan 1994b). The periodicity of the graphite is clearly observed.

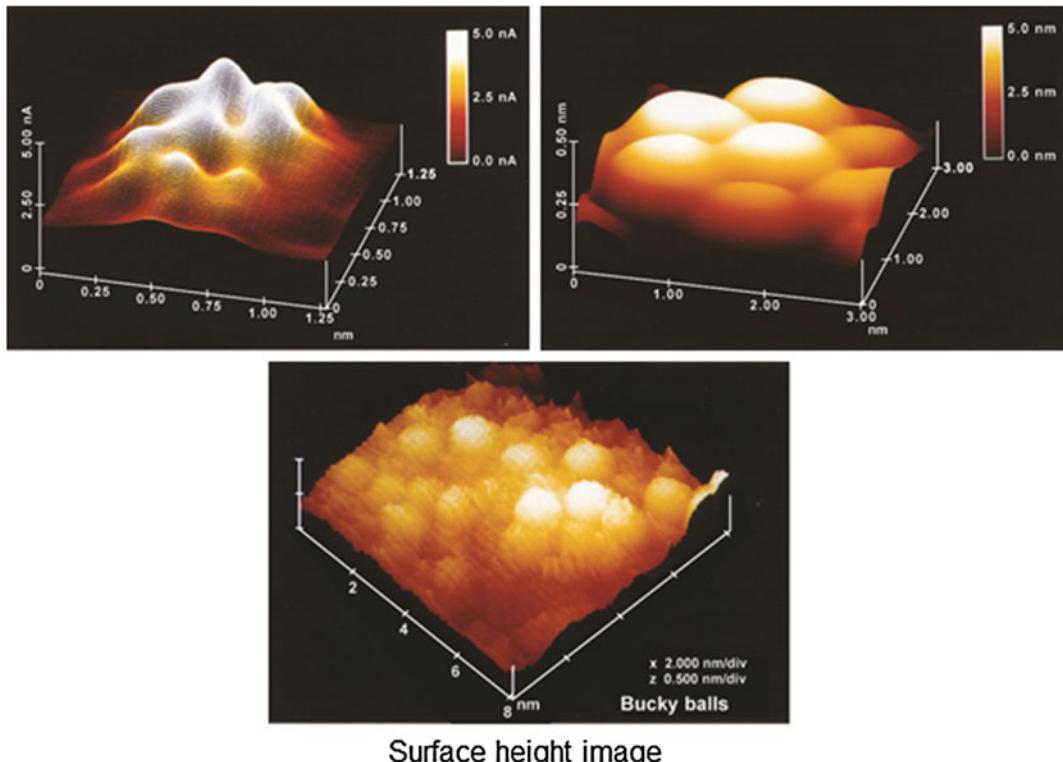


Fig. 8.12 STM images of solvent deposited C_{60} film on a gold-coated freshly-cleaved mica at various scan sizes (Bhushan et al. 2003)

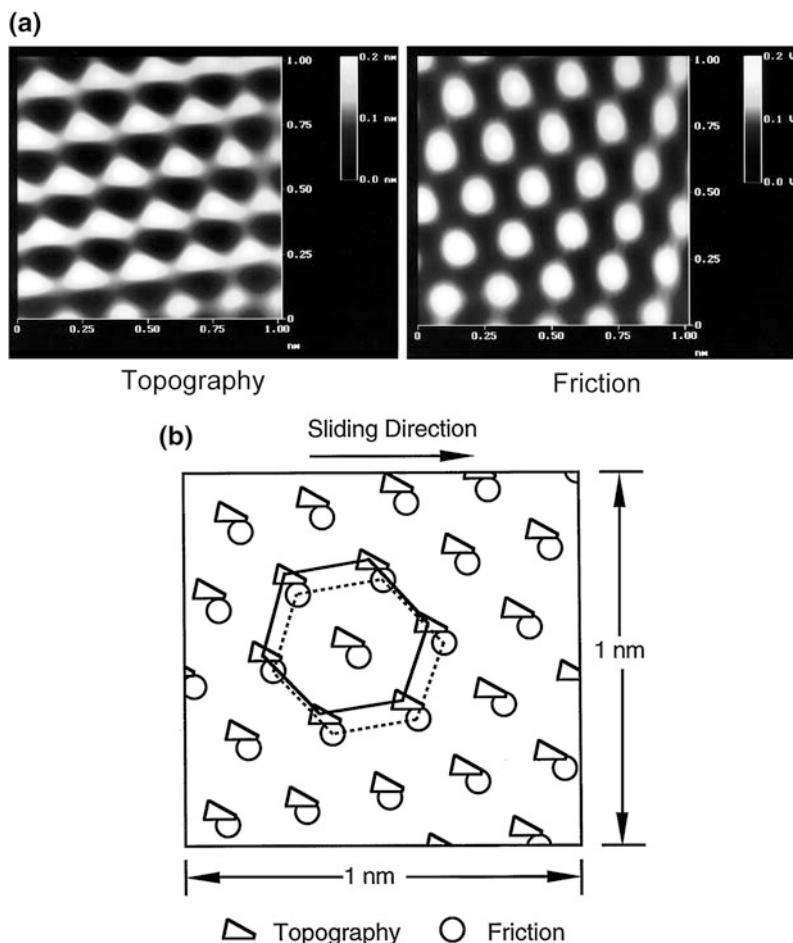
To study the friction mechanisms on an atomic scale, a freshly cleaved HOPG was studied by Mate et al. (1987) and Ruan and Bhushan (1994b). Figure 8.14a shows the atomic-scale friction force map (raw data) and Fig. 8.13a shows the friction force maps (after 2D spectrum filtering with high frequency noise truncated) (Ruan and Bhushan 1994b). Figure 8.14a also shows a line plot of friction force profiles along some crystallographic direction. The actual shape of the friction profile depends upon the spatial location of axis of tip motion. Note that a portion of atomic-scale lateral force is conservative. Mate et al. (1987) and Ruan and Bhushan (1994b) reported that the average friction force linearly increased with normal load and was reversible with load. Friction profiles were similar during the sliding of the tip in either direction.

During scanning, the tip moves discontinuously over the sample surface and jumps with

discrete steps from one potential minimum (well) to the next. This leads to a saw-tooth-like pattern for the lateral motion (force) with a periodicity of the lattice constant. This motion is called the stick-slip movement of the tip (Mate et al. 1987; Ruan and Bhushan 1994b; Bhushan 1999a). The observed friction force includes two components—conservative and periodic, and nonconservative and constant. If the relative motion of the sample and tip were simply that of two rigid collections of atoms, the effective force would be a conservative force oscillating about zero. Slow reversible elastic deformation would also contribute to conservative force. The origin of the nonconservative direction-dependent force component would be phonon generation, viscous dissipation, or plastic deformation.

Stick-slip on the atomic scale, discussed above, is the result of the energy barrier required to be overcome to jump over the atomic corrugations on the sample surface. It corresponds to the energy

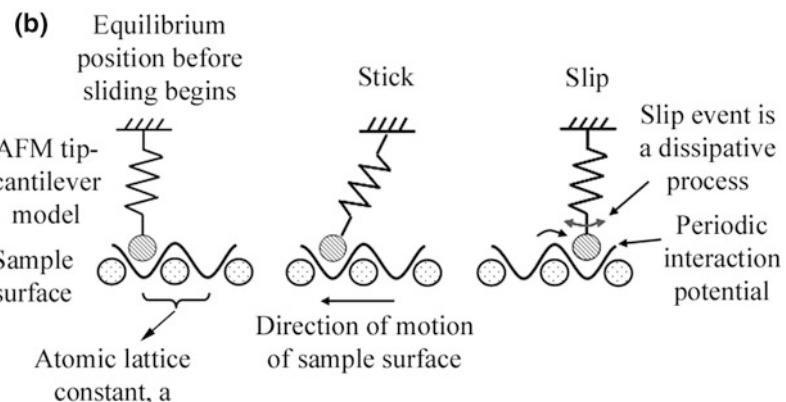
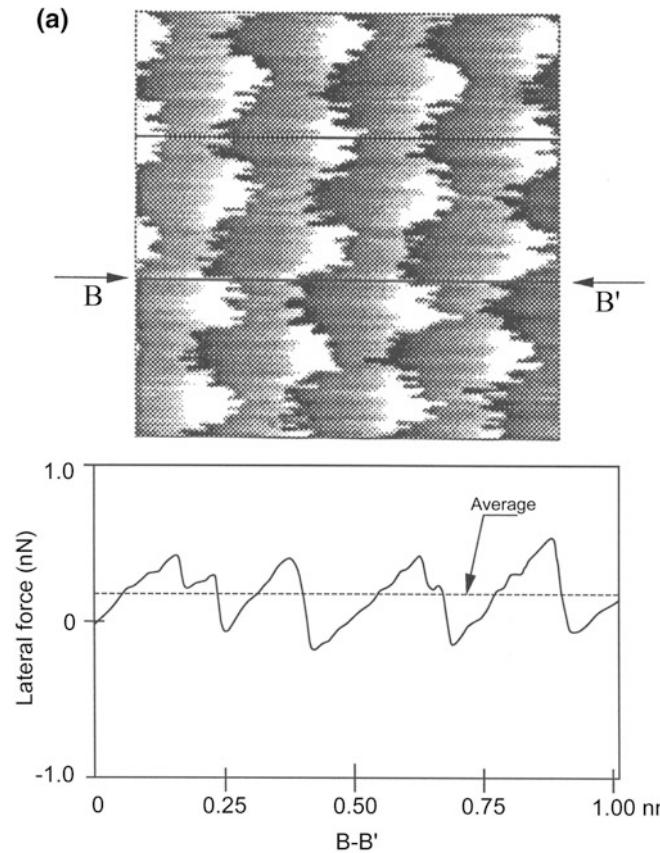
Fig. 8.13 **a** Gray-scale plots of surface topography and friction force maps (2D spectrum filtered), measured simultaneously, of a $1 \text{ nm} \times 1 \text{ nm}$ area of freshly cleaved HOPG, showing the atomic-scale variation of topography and friction, and **b** schematic of superimposed topography and friction maps from (a); the symbols correspond to maxima. Note the spatial shift between the two plots (Ruan and Bhushan 1994b)



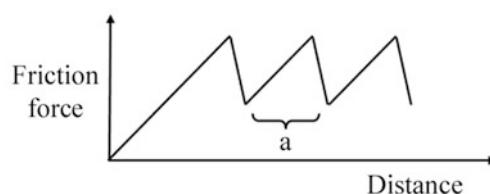
required for the jump of the tip from a stable equilibrium position on the surface to a neighboring position. The perfect atomic regularity of the surface guarantees the periodicity of the lateral force signal, independent of the actual atomic structure of the tip apex. A few atoms (based on the magnitude of the friction force, less than 10) on a tip sliding over an array of atoms on the sample are expected to undergo the stick-slip. For simplicity, Fig. 8.14b shows a simplified model for one atom on a tip with a one-dimensional spring mass system. As the sample surface slides against the AFM tip, the tip remains "stuck" initially until it can overcome the energy (potential) barrier, which is illustrated by a sinusoidal interaction potential as experienced by the tip. After some motion, there is enough energy stored in the

spring which leads to "slip" into the neighboring stable equilibrium position. During the slip and before attaining stable equilibrium, stored energy is converted into vibrational energy of the surface atoms in the range of 1013 Hz (phonon generation) and decays within the range of 10^{-11} s into heat. (A wave of atoms vibrating in concert are termed a phonon). The stick-slip phenomenon, resulting from irreversible atomic jumps, can be theoretically modeled with classical mechanical models (Tomlinson 1929; Tomanek et al. 1991). The Tomanek-Zhong-Thomas model (Tomanek et al. 1991) is the starting point for determining friction force during atomic scale stick-slip. The AFM model describes the total potential as the sum of the potential acting on the tip due to interaction with the sample and the elastic energy

Fig. 8.14 **a** Gray scale plot of friction force map (raw data) of a $1 \times 1 \text{ nm}^2$ area of freshly cleaved HOPG, showing atomic-scale variation of friction force. High points are shown by lighter color. Also shown is line plot of friction force profile along the line indicated by arrows. The normal load was 25 nN and the cantilever normal stiffness was 0.4 N/m (Ruan and Bhushan 1994b), and **b** Schematic of a model for a tip atom sliding on an atomically flat periodic surface. The schematic shows the tip jumping from one potential minimum to another, resulting in stick-slip behavior



Sawtooth pattern of friction force arising from atomic scale stick-slip



stored in the cantilever. Thermally activated stick-slip behavior can explain the velocity effects on friction, to be presented later.

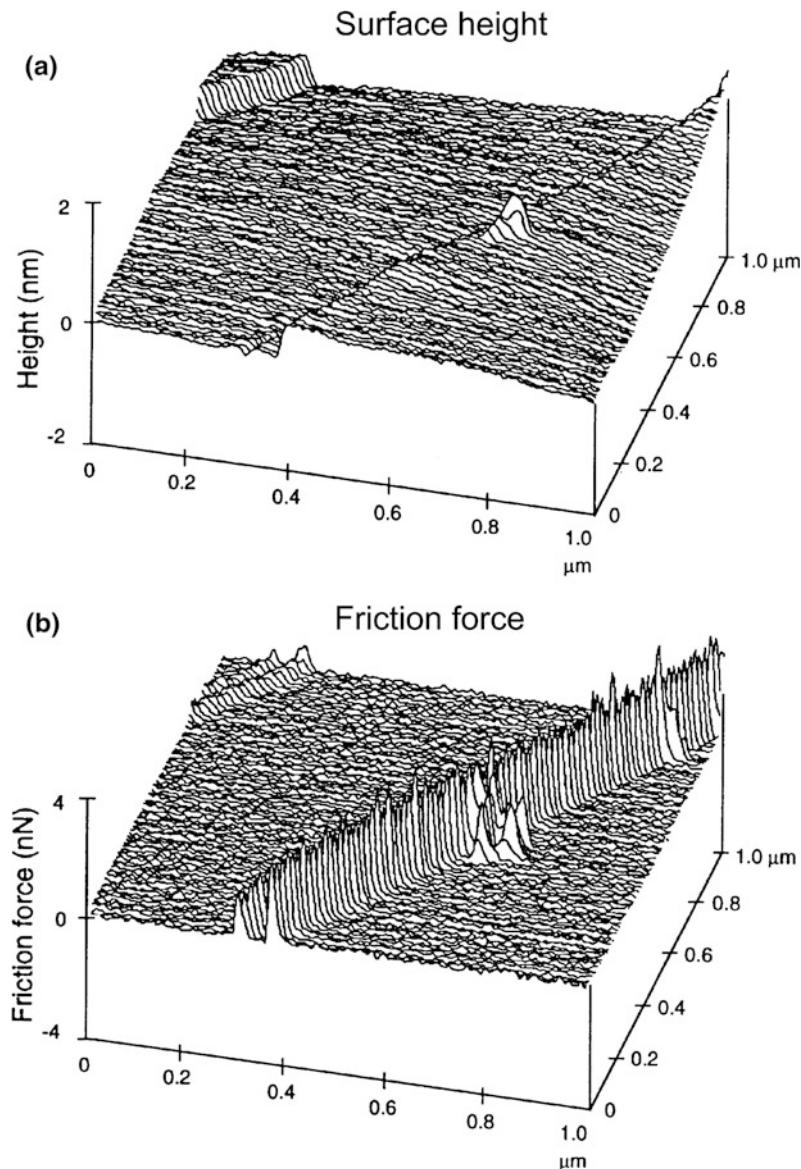
Finally, based on Fig. 8.13a, the atomic-scale friction force of HOPG exhibited the same periodicity as that of the corresponding topography, but the peaks in friction and those in the topography are displaced relative to each other (Fig. 8.13b). A Fourier expansion of the interatomic potential was used by Ruan and Bhushan (1994b) to calculate the conservative interatomic forces between the atoms of the FFM tip and those of the graphite surface. Maxima in the

interatomic forces in the normal and lateral directions do not occur at the same location, which explains the observed shift between the peaks in the lateral force and those in the corresponding topography.

8.3.2 Microscale Friction

Local variations in the microscale friction of cleaved graphite are observed, Fig. 8.15. Microscale friction is defined as the friction measured with a scan size equal to or larger than 1 μm

Fig. 8.15 **a** Surface roughness and **b** friction force maps at a normal load of 42 nN of freshly cleaved HOPG surface against an Si_3N_4 FFM tip. Friction in the *line-shaped* region is over an order of magnitude larger than the smooth areas (Ruan and Bhushan 1994b)



1 μm . These arise from structural changes that occur during the cleaving process (Ruan and Bhushan 1994c). The cleaved HOPG surface is largely atomically smooth but exhibits line-shaped regions in which the coefficient of friction is more than an order of magnitude larger. Transmission electron microscopy indicates that the line-shaped regions consist of graphite planes of different orientation, as well as of amorphous carbon. Differences in friction have also been observed for multiphase ceramic materials (Koinkar and Bhushan 1996c). Figure 8.16 shows the surface roughness and friction force maps of Al_2O_3 -TiC (70–30 wt%). TiC grains have a Knoop hardness of about 2800 kg/mm^2 and Al_2O_3 has 2100 kg/mm^2 , therefore, TiC grains do not polish as much and result in a slightly higher elevation (about 2–3 nm higher than that of Al_2O_3 grains). TiC grains exhibit higher friction force than Al_2O_3 grains. The coefficients of friction of TiC and Al_2O_3 grains are 0.034 and 0.026, respectively, and the coefficient of friction of Al_2O_3 -TiC composite is 0.03. Local variation in friction force also arises from the scratches present on the Al_2O_3 -TiC surface. Meyer et al. (1992) also used FFM to measure structural variations of organic mono- and multi-layer films. All of these measurements suggest that the FFM can be used for structural mapping of the surfaces. FFM measurements can also be used to map chemical variations, as indicated by the use of the FFM with a modified probe tip to map the spatial arrangement of chemical functional groups in mixed organic monolayer films (Frisbie et al. 1994). Here, sample regions that had stronger interactions with the functionalized probe tip exhibited larger friction.

Local variations in the microscale friction of nominally rough, homogeneous-material surfaces can be significant, and are seen to depend on the local surface slope rather than the surface height distribution, Fig. 8.17. This dependence was first reported by Bhushan and Ruan (1994), Bhushan et al. (1994), and Bhushan (1995) and later discussed in more detail by Koinkar and Bhushan (1997b) and Sundararajan and Bhushan (2000). In order to elegantly show any correlation between local values of friction and surface

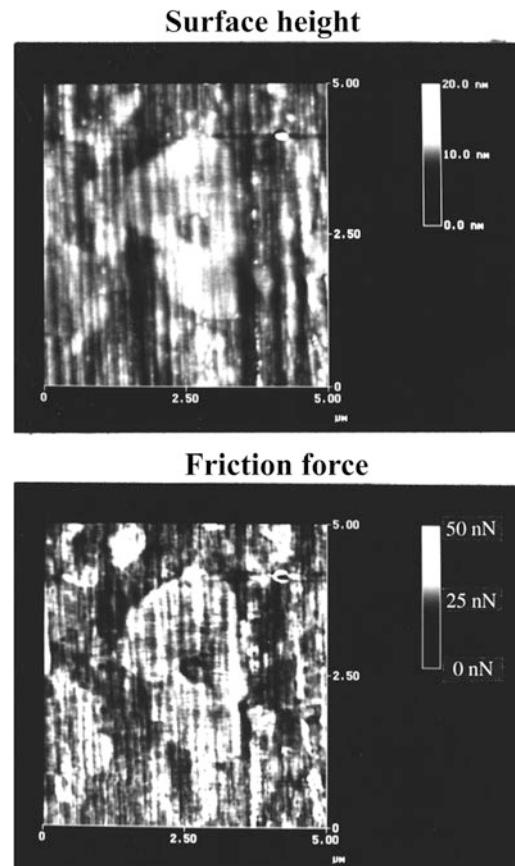
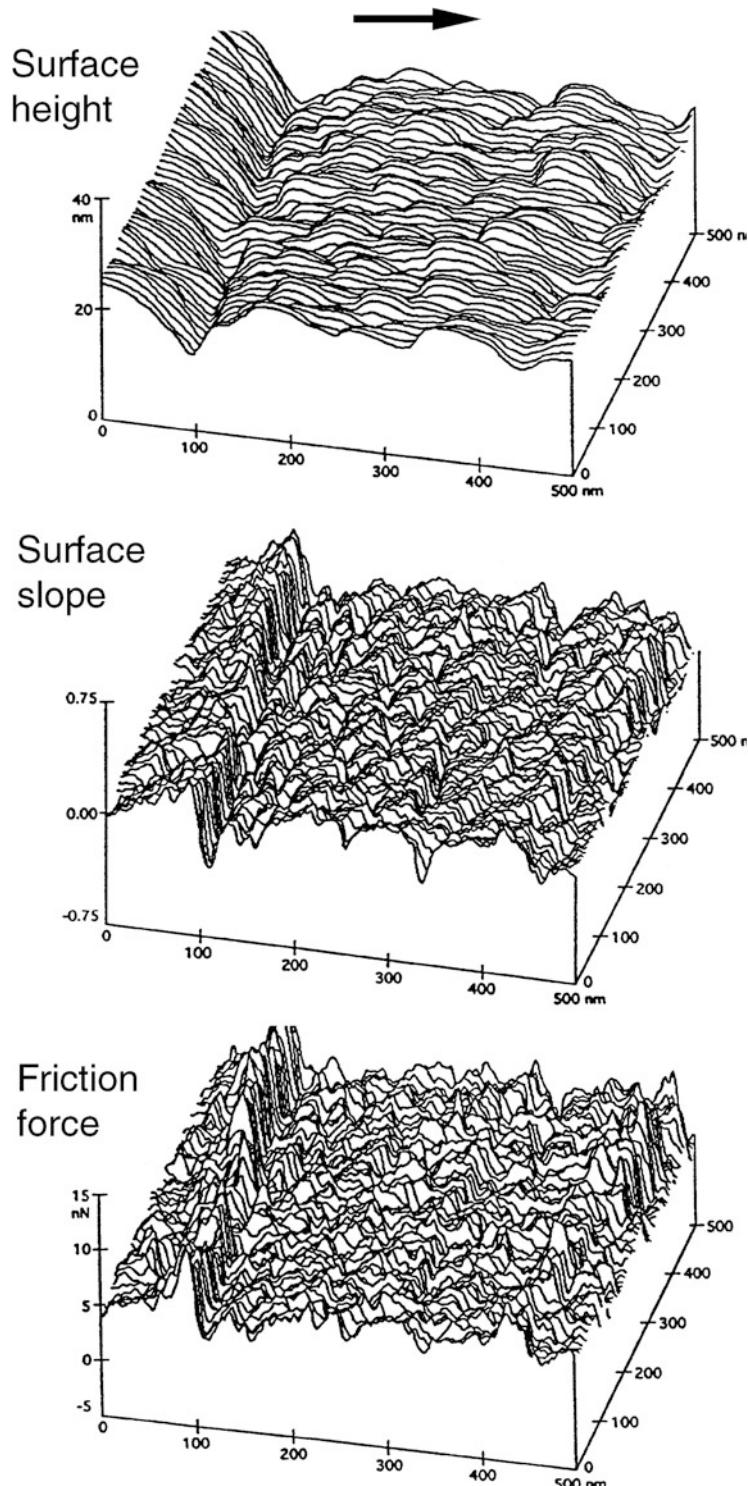


Fig. 8.16 Gray scale surface roughness ($\sigma = 0.80 \text{ nm}$) and friction force maps (mean = 7.0 nN, $\sigma = 0.90 \text{ nN}$) for Al_2O_3 -TiC (70–30 wt%) at a normal load of 138 nN (Koinkar and Bhushan 1996c)

roughness, surface roughness and friction force maps of a gold-coated ruler with somewhat rectangular grids and a silicon grid with square pits were obtained, Fig. 8.18 [Sundararajan and Bhushan (2000)]. Figures 8.17 and 8.18 show the surface roughness map, the slopes of the roughness map taken along the sliding direction (surface slope map) and the friction force map for various samples. There is a strong correlation between the surface slopes and friction forces. For example, in Fig. 8.18, the friction force is high locally at the edge of the grids and pits with a positive slope and is low at the edges with a negative slope.

We now examine the mechanism of microscale friction, which may explain the

Fig. 8.17 Surface roughness map ($\sigma = 4.4$ nm), surface slope map taken in the sample sliding direction (the horizontal axis; mean = 0.023, $\sigma = 0.197$), and friction force map (mean = 6.2 nN, $\sigma = 2.1$ nN) for a lubricated thin-film magnetic rigid disk for a normal load of 160 nN (Bhushan et al. 1994)



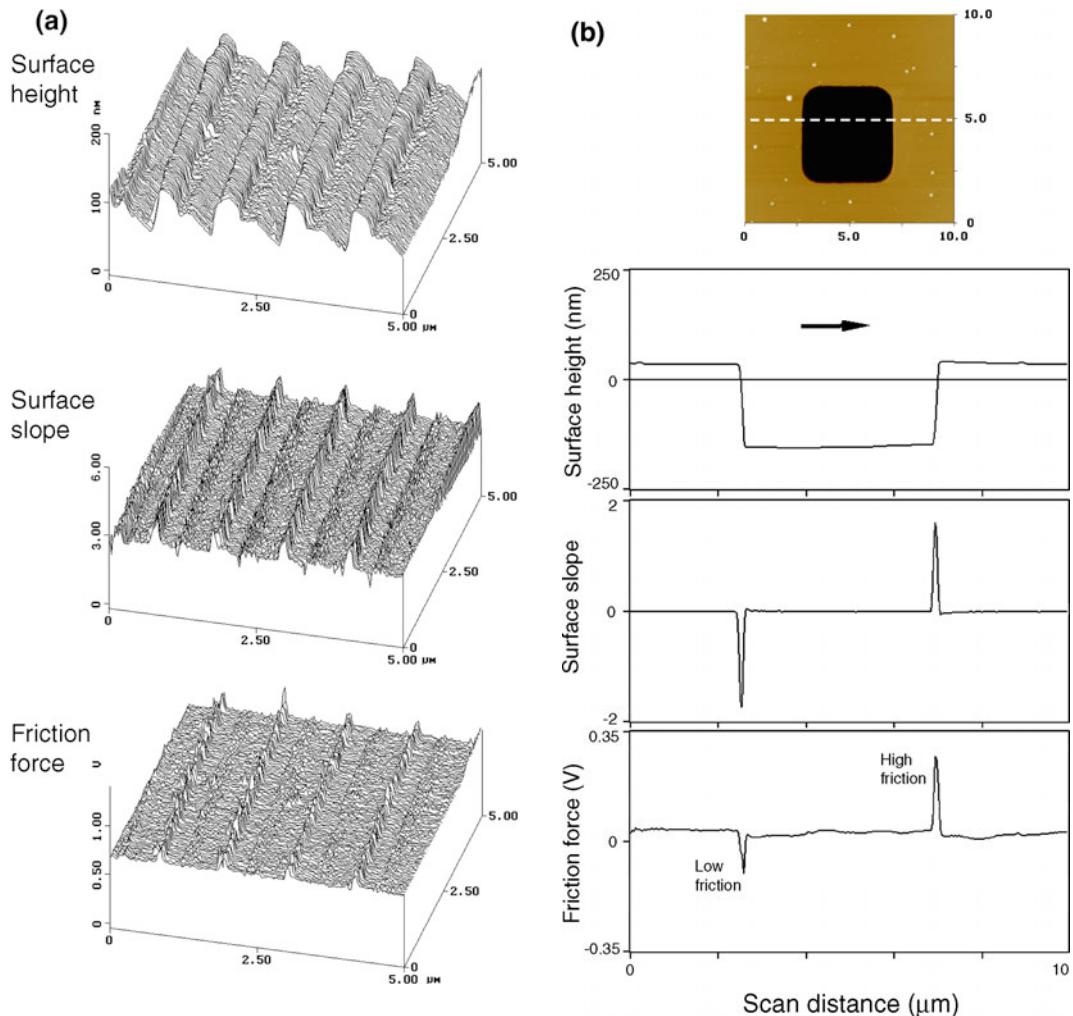


Fig. 8.18 Surface roughness map, surface slope map taken in the sample sliding direction (the horizontal axis), and friction force map for **a** a gold-coated ruler (with somewhat rectangular grids with a pitch of 1 μm and a ruling step height of about 70 nm) at a normal load of 25 nN and **b** a silicon grid (with 5 μm square pits of depth 180 nm and a pitch of 10 μm) (Sundararajan and Bhushan 2000)

resemblance between the slope of surface roughness maps and the corresponding friction force maps (Bhushan and Ruan 1994; Bhushan et al. 1994; Ruan and Bhushan 1994b, c; Koinkar and Bhushan 1997b; Bhushan 1999a, b, 2008, 2011; Sundararajan and Bhushan 2000). There are three dominant mechanisms of friction; adhesive, ratchet, and plowing (Bhushan 2013a, b). As a first order, we may

Fig. 8.18 (continued)

assume these to be additive. The adhesive mechanism cannot explain the local variation in friction. Next we consider the ratchet mechanism. We consider a small tip sliding over an asperity making an angle θ with the horizontal plane, Fig. 8.19. The normal force W (normal to the general surface) applied by the tip to the sample surface is constant. The friction force F on the sample would be a constant for a smooth surface if the friction mechanism does not change. For a rough surface shown in Fig. 8.19, if the adhesive mechanism does not change during sliding, the local value of the coefficient of friction remains constant,

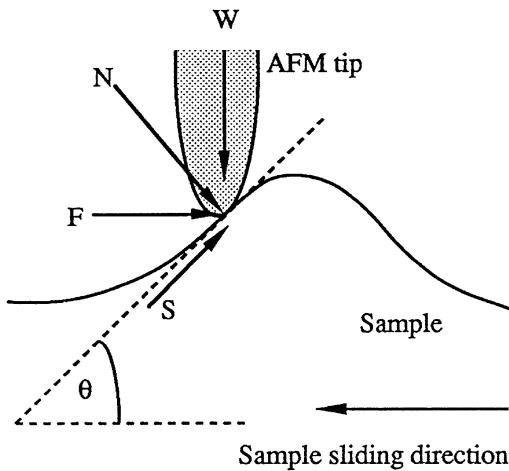


Fig. 8.19 Schematic illustration showing the effect of an asperity (making an angle θ with the *horizontal plane*) on the surface in contact with the tip on local friction in the presence of adhesive friction mechanism. W and F are the normal and friction forces, respectively, and S and N are the force components along and perpendicular to the local surface of the sample at the contact point, respectively

$$\mu_0 = S/N \quad (8.1)$$

where S is the local friction force and N is the local normal force. However, the friction and normal forces are measured with respect to global horizontal and normal axes, respectively. The measured local coefficient of friction μ_1 in the ascending part is

$$\begin{aligned} \mu_1 &= F/W \\ &= (\mu_0 + \tan \theta)/(1 - \mu_0 \tan \theta) \sim \mu_0 + \tan \theta \\ &\text{for small } \mu_0 \tan \theta \end{aligned} \quad (8.2)$$

indicating that in the ascending part of the asperity one may simply add the friction force and the asperity slope to one another. Similarly, on the right-hand side (descending part) of the asperity,

$$\begin{aligned} \mu_2 &= (\mu_0 - \tan \theta)/(1 + \mu_0 \tan \theta) \sim \mu_0 \\ &- \tan \theta, \text{ for small } \mu_0 \tan \theta \end{aligned} \quad (8.3)$$

For a symmetrical asperity, the average coefficient of friction experienced by the FFM tip traveling across the whole asperity is

$$\begin{aligned} \mu_{\text{ave}} &= (\mu_1 + \mu_2)/2 \\ &= \mu_0(1 + \tan^2 \theta)/(1 - \mu_0^2 \tan^2 \theta) \sim \mu_0(1 + \tan^2 \theta), \\ &\text{for small } \mu_0 \tan \theta \end{aligned} \quad (8.4)$$

Finally we consider the plowing component of friction with tip sliding in either direction, which is (Bhushan 1995a, 2011)

$$\mu_p \sim \tan \theta \quad (8.5)$$

Because in FFM measurements we notice little damage of the sample surface, the contribution by plowing is expected to be small, and the ratchet mechanism is believed to be the dominant mechanism for the local variations in the friction force map. With the tip sliding over the leading (ascending) edge of an asperity, the surface slope is positive; it is negative during sliding over the trailing (descending) edge of an asperity. Thus, measured friction is high at the leading edge of asperities and low at the trailing edge. In addition to the slope effect, the collision of the tip when encountering an asperity with a positive slope produces additional torsion of the cantilever beam leading to higher measured friction force. When encountering an asperity with the same negative slope, however, there is no collision effect and hence no effect on torsion. This effect also contributes to the difference in friction forces when the tip scans up and down on the same topography feature. The ratchet mechanism and the collision effects thus semi-quantitatively explain the correlation between the slopes of the roughness maps and friction force maps observed in Figs. 8.17 and 8.18. We note that in the ratchet mechanism, the FFM tip is assumed to be small compared to the size of asperities. This is valid since the typical radius of curvature of the tips is about 10–50 nm. The radii of curvature of the asperities of the samples measured here (the asperities that produce most of the friction variation) are found to typically be about 100–200 nm, which is larger than that of the FFM tip (Bhushan and Blackman 1991). It is important to note that the measured local values of friction and normal forces are measured with

respect to global (and not local) horizontal and vertical axes, which are believed to be relevant in applications.

8.3.3 Directionality Effect on Microfriction

During friction measurements, the friction force data from both the forward (trace) and backward (retrace) scans are useful in understanding the origins of the observed friction forces. Magnitudes of material-induced effects are independent of the scanning direction whereas topography-induced effects are different between forward and backward scanning directions. Since the sign of the friction force changes as the scanning direction is reversed (because of the reversal of torque applied to the end of the tip), addition of the friction force data of the forward and backward scan eliminates the material-induced effects while topography-induced effects still remain. Subtraction of the data between forward and backward scans does not eliminate either effect, Fig. 8.20 (Sundararajan and Bhushan 2000).

Owing to the reversal of the sign of the retrace (R) friction force with respect to the trace (T) data, the friction force variations due to topography are in the same direction (peaks in trace correspond to peaks in retrace). However, the magnitudes of the peaks in trace and retrace at a given location are different. An increase in the friction force experienced by the tip when scanning up a sharp change in topography is more than the decrease in the friction force experienced when scanning down the same topography change, partly because of collision effects discussed earlier. Asperities on engineering surfaces are asymmetrical, which also affect the magnitude of friction force in the two directions. Asymmetry in tip shape may also have an effect on the directionality effect of friction. We will note later that the magnitude of surface slopes are virtually identical, therefore, the tip shape asymmetry should not have much effect.

Figure 8.21 shows the surface height and friction force data for a gold ruler and a silicon grid in the trace and retrace directions. Subtraction of

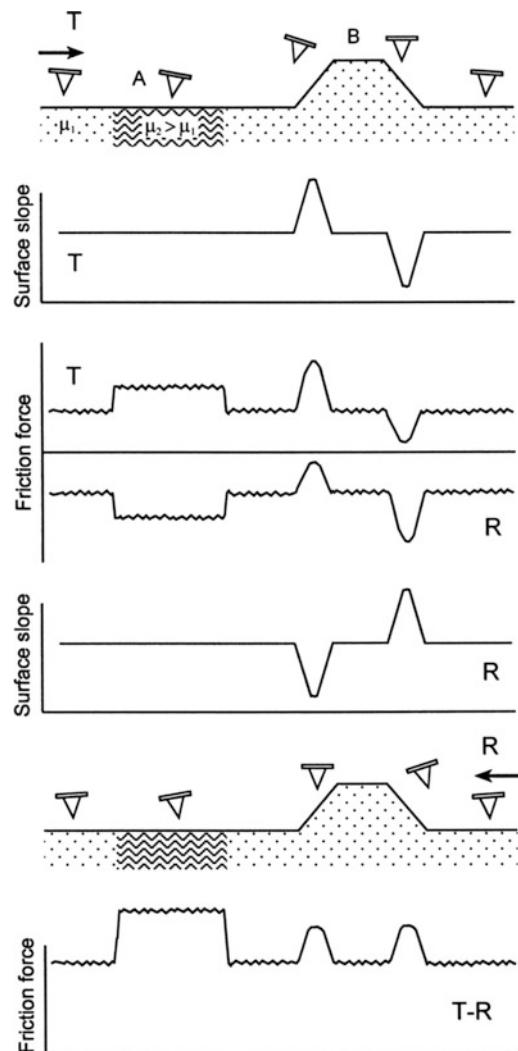


Fig. 8.20 Schematic of friction forces expected when a tip traverses a sample that is composed of different materials and sharp changes in topography. A schematic of surface slope is also shown

two friction data yields a residual peak because of the differences in the magnitudes of friction forces in the two directions. This effect is observed at all locations of significant changes in topography.

In order to facilitate comparison of the directionality effect on friction, it is important to take into account the sign change of the surface slope and friction force in the trace and retrace directions. Figure 8.22 shows surface height, surface slope, and friction force data for the two samples in the trace and retrace directions. The correlations

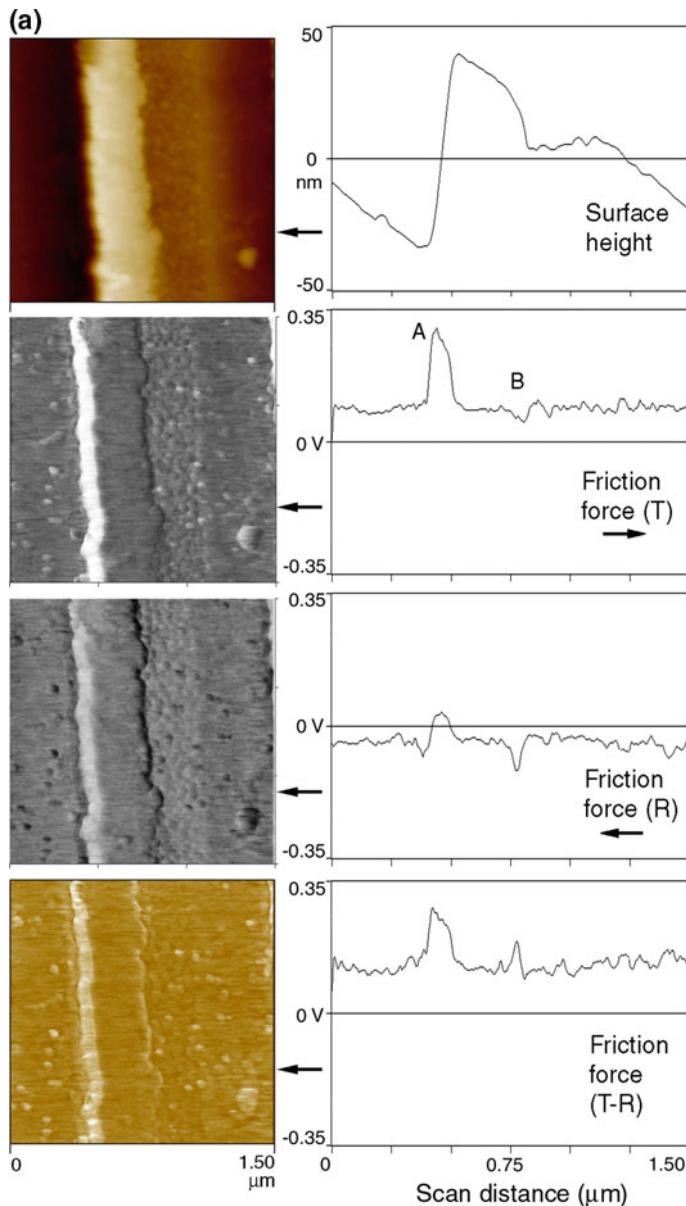


Fig. 8.21 **a** Gray scale images and two dimensional profiles of surface height and friction forces across a single ruling of the gold-coated ruling, and **b** two dimensional profiles of surface height and friction forces across a silicon grid pit. Friction force data in trace and retrace directions, and substrated force data are presented (Sundararajan and Bhushan 2000)

between surface slope and friction forces are clear. The third column in the figures shows retrace slope and friction data with an inverted sign (retrace). Now we can compare trace data with retrace data. It is clear that the friction experienced by the tip is dependent upon the scanning

direction because of surface topography. In addition to the effect of topographical changes discussed earlier, during surface-finishing processes, material can be transferred preferentially onto one side of the asperities, which also causes asymmetry and direction dependence. Reduction in

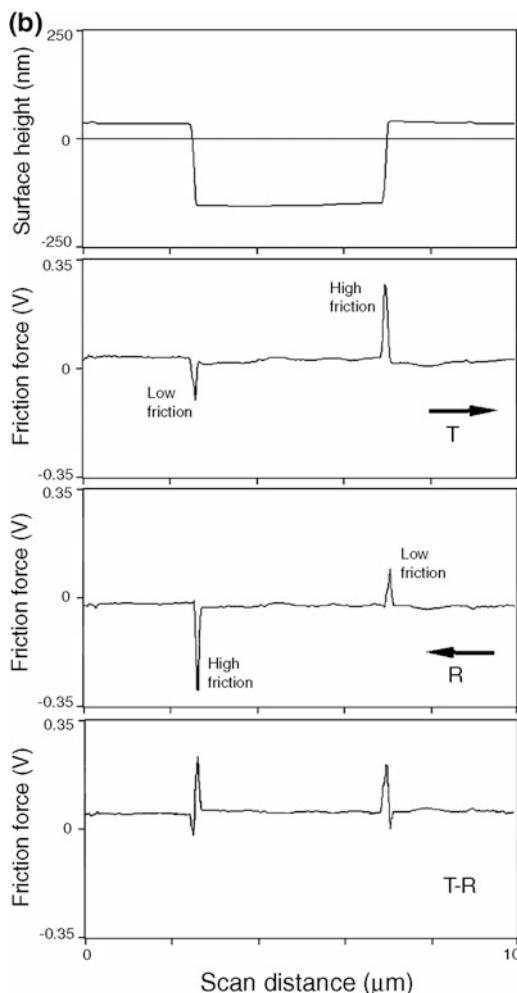


Fig. 8.21 (continued)

local variations and in the directionality of friction properties requires careful optimization of surface roughness distributions and of surface-finishing processes.

The directionality as a result of surface asperities effect will be also manifested in macroscopic friction data, i.e., the coefficient of friction may be different in one sliding direction than that in the other direction. The asymmetrical shape of the asperities accentuates this effect. The frictional directionality can also exist in materials with particles having a preferred orientation. The directionality effect in friction on a macroscale is observed in some magnetic tapes. In a macroscale test, a 12.7-mm wide polymeric magnetic tape

was wrapped over an aluminum drum and slid in a reciprocating motion with a normal load of 0.5 N and a sliding speed of about 60 mm/s (Bhushan 1995). The coefficient of friction as a function of sliding distance in either direction is shown in Fig. 8.23. We note that the coefficient of friction on a macroscale for this tape is different in different directions. Directionality in friction is sometimes observed on the macroscale; on the microscale this is the norm (Bhushan 1996, 1999a). On the macroscale, the effect of surface asperities normally is averaged out over a large number of contacting asperities.

8.3.4 Surface Roughness–Independent Microscale Friction

As just reported, the friction contrast in conventional friction measurements is based on interactions dependent upon interfacial material properties superimposed by roughness-induced lateral forces, and the cantilever twist is dependent on the sliding direction because of the local surface slope. Hence it is difficult to separate the friction-induced from the roughness-induced cantilever twist in the image. To obtain the roughness-independent friction, lateral or torsional modulation techniques are used in which the tip is oscillated in-plane with a small amplitude at a constant normal load, and change in shape and magnitude of cantilever resonance is used as a measure of friction force (Yamanaka and Tomita 1995; Scherer et al. 1998, 1999; Reinstaedtler et al. 2003, 2005a, b; Bhushan and Kasai 2004). These techniques also allow measurements over a very small area (few nm to few μm).

Scherer et al. (1999) and Reinstaedtler et al. (2003, 2005a) used lateral mode for friction measurements (Fig. 8.10b). Bhushan and Kasai (2004) used the TR mode for these measurements (Fig. 8.11). Before engagement, the cantilever is driven into torsional motion of the cantilever/tip assembly with a given normal vibration amplitude (vibration amplitude in free air). After engagement, the vibration amplitude decreases due to the interaction between the tip

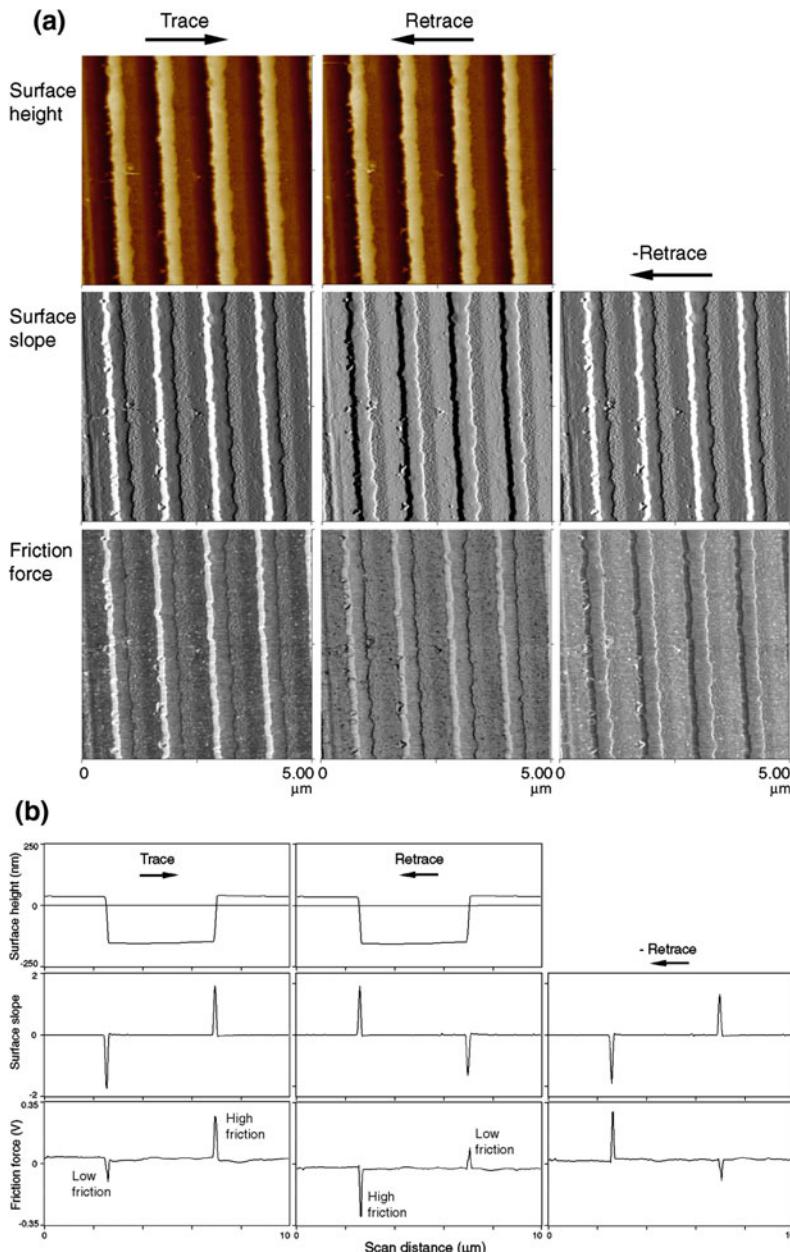


Fig. 8.22 **a** Gray scale images of surface heights, surface slopes and friction forces for scans across a gold-coated ruling, and **b** two dimensional profiles of surface heights, surface slopes and friction forces for scans across the silicon grid pit. Arrows indicate the tip sliding direction (Sundararajan and Bhushan 2000)

and the sample, the vibration frequency increases, and phase shift occurs. During scanning, the normal load is kept constant, and the vibration amplitude of the cantilever is measured at the contact frequency.

As mentioned earlier, the shift in contact resonance frequency both in lateral and in TR modes is a measure of contact stiffness as shown schematically in Fig. 8.24. At excitation voltage above a certain value, as a result of micro-slip at

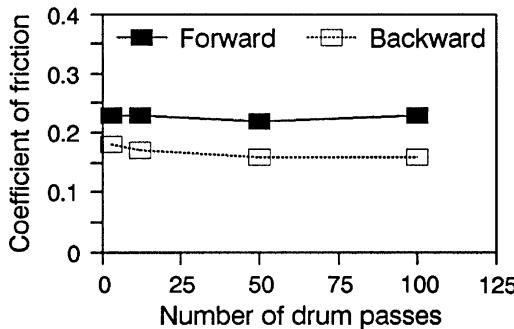


Fig. 8.23 Coefficient of macroscale friction as a function of drum passes for a polymeric magnetic tape sliding over an aluminum drum in a reciprocating mode in both directions. Normal load = 0.5 N over 12.7-mm wide tape, sliding speed = 60 mm/s (Bhushan 1995)

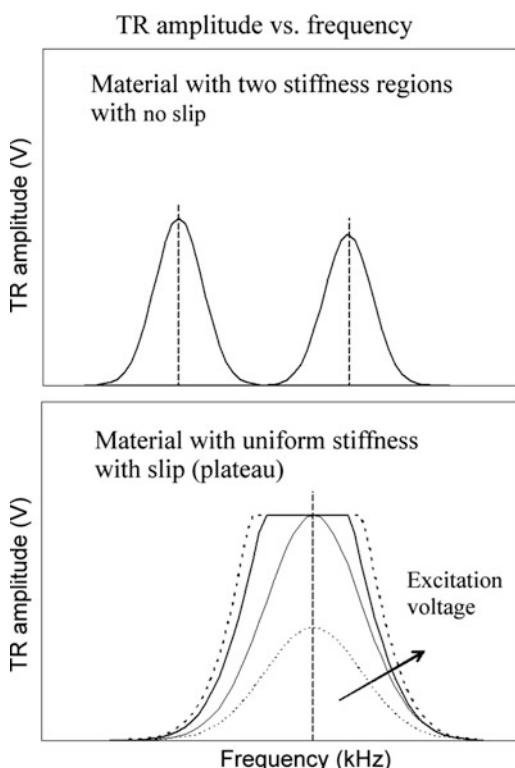


Fig. 8.24 Schematic showing frequency profiles of the TR amplitude for materials with two phases and a single phase. The maximum TR amplitude at the contact resonance frequency of the resonance curve with a flattened top, resulting from slip, can be used for friction force measurement

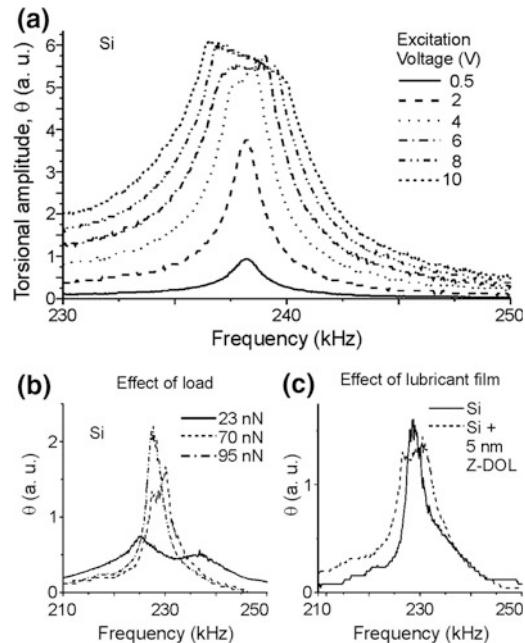
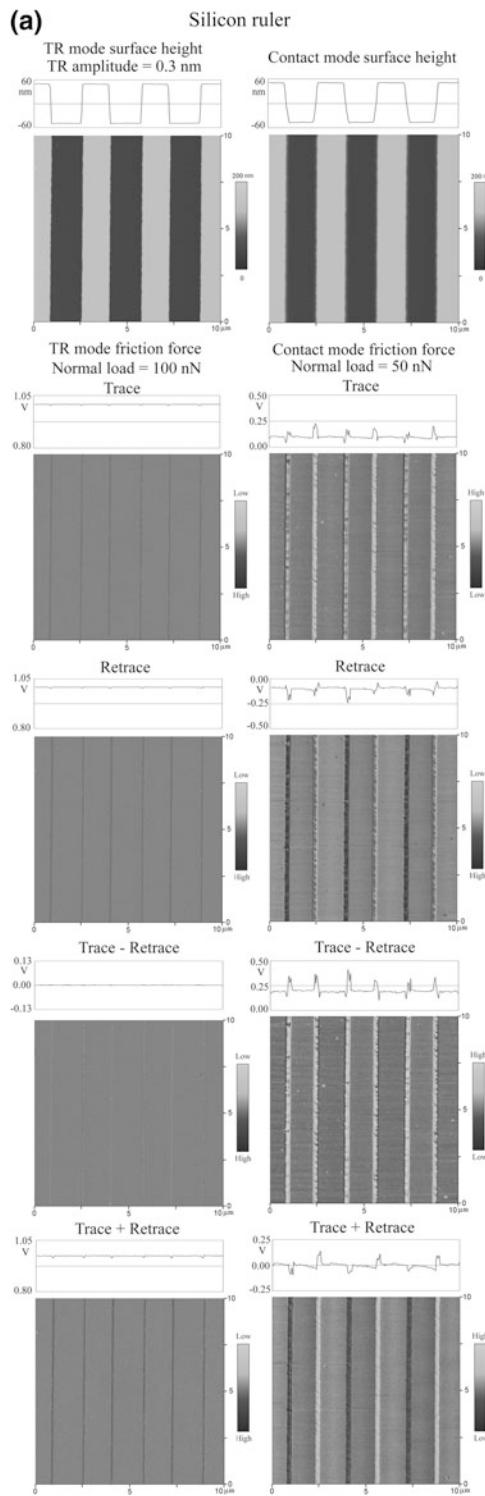


Fig. 8.25 Torsional vibration amplitude of the cantilever as a function of excitation frequency. **a** Measurement on bare silicon. The different curves correspond to increasing excitation voltages applied to the transducer and, hence, increasing surface amplitudes. **b** Measurement on silicon lubricated with a 5-nm-thick Z-DOL layer. Curves for three different static loads are shown. The transducer was excited with 5 V of amplitude. **c** Measurement with a static load of 70 nN and 7 V excitation amplitude. The two curves correspond to bare silicon and lubricated silicon, respectively (Reinstaedtler et al. 2003)

the interface, a flattening of the resonance frequency spectra occurs (Fig. 8.22). At lowest excitation voltage, the AFM tip sticks to the sample surface and follows the motion like an elastic contact with viscous damping, and the resonance curve is Lorentzian with a well-defined maximum. The excitation voltage should be high enough to initiate micro-slip. The maximum torsional amplitude at a given resonance frequency is a function of friction force and sample stiffness, so the technique is not valid for inhomogeneous samples. If the torsional stiffness of the cantilever is very high compared to the sample stiffness, the technique should work.

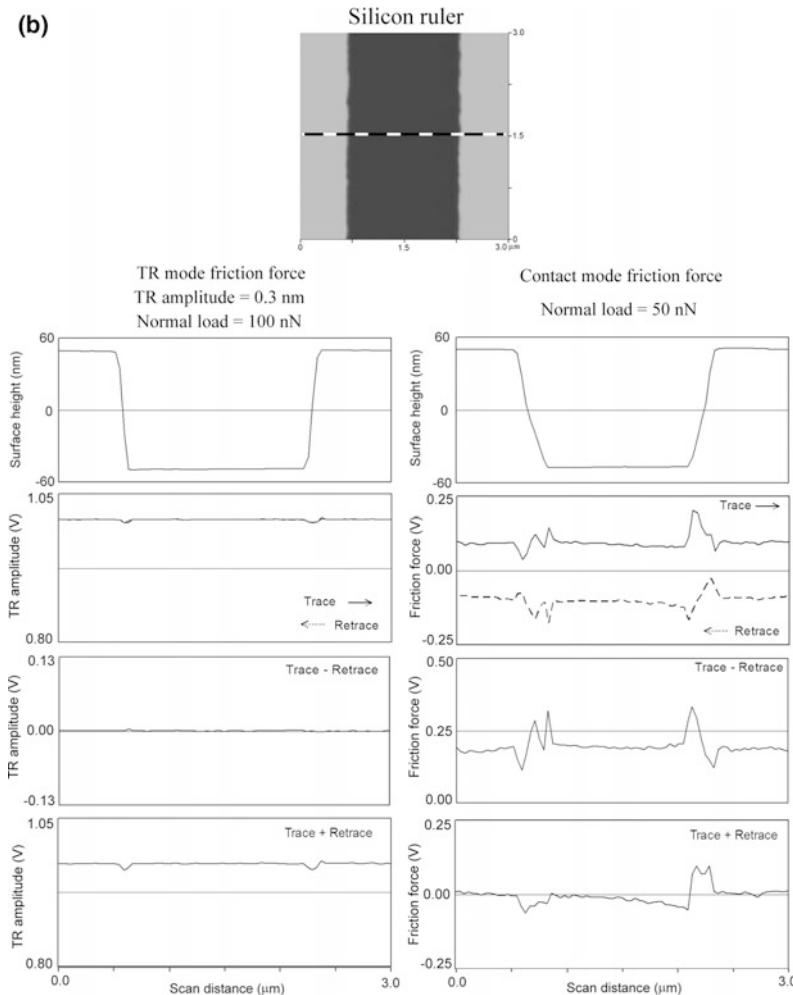
Reinstaedtler et al. (2003) performed lateral mode experiments on bare Si and Si lubricated with 5-nm thick chemically bonded



perfluoropolyether (Z-DOL) lubricant film. Figure 8.25a shows the amplitude of the cantilever torsional vibration as a function of frequency on a bare silicon sample. The frequency sweep was adjusted such that a contact resonance frequency was covered. The different curves correspond to different excitation voltages applied to the shear wave transducer. At low amplitudes, the shape of the resonance curve is Lorentzian. Above a critical excitation amplitude of the transducer (excitation voltage = 4 V corresponding to ≈ 0.2 nm lateral surface amplitude as measured by interferometry), the resonance curve flattens out, and the frequency-span of the flattened part increases further with the excitation amplitude. Here, the static force applied was 47 nN and the adhesion force was 15 nN. The resonance behavior of the tip-cantilever system in contact with the lubricated silicon sample (Fig. 8.25b) was similar to that of the bare silicon sample. By increasing the static load, the critical amplitude for the appearance of the flattening increases. The deviations from the Lorentzian resonance curve became visible at static loads lower than 95 nN. As shown in Fig. 8.25c, the resonance curve obtained at the same normal load of 70 nN and at the same excitation voltage (7 V) is more flattened on the lubricated sample than on the bare sample which led us to conclude that the critical amplitude is lower on the lubricated sample than on the bare sample. These experiments clearly demonstrate that torsional vibration of an AFM cantilever at ultrasonic frequencies leads to stick-slip phenomena and sliding friction. Above a critical vibration amplitude, sliding friction sets in.

Bhushan and Kasai (2004) performed friction measurements on a silicon ruler and demonstrated that friction data in TR mode is essentially independent of surface-roughness and sliding direction. Figure 8.26a shows surface height and

Fig. 8.26 **a** A comparison between the TR mode friction and contact mode friction maps together with line scans, on the silicon ruler. The TR mode surface height and contact mode surface height images are also shown. **b** A comparison of the line scans of the TR mode friction and contact mode friction on a selected pitch of the silicon ruler (Bhushan and Kasai 2004)

**Fig. 8.26** (continued)

friction force maps on a silicon ruler obtained using the TR mode and contact mode techniques. A comparison is made between the TR mode and contact mode friction force maps. For easy comparison, the line scan profiles near the central area are shown on top of the grey scale maps. The vertical scales of the friction force profiles in the two graphs are selected to cover the same range of friction force so that direct comparison can be made, i.e., 0.25 V in full scale for the TR mode corresponds to 0.5 V for the contact mode in these measurements. As expected, for the trace scan, small downward peaks in the TR mode map and large upward and downward peaks in the contact mode map are observed. The positions of these peaks coincide with those of the surface

slope; therefore, the peaks in the friction signals are attributed to a topography-induced effect. For the retrace scan, the peak pattern for the TR mode stays similar, but for the contact mode, the pattern becomes reversed.

The subtraction image for the TR mode shows almost flat contrast, since the trace and retrace friction data profiles are almost identical. For the contact mode, the subtraction image shows that the topography-induced contribution still exists. As stated earlier, the addition image of the TR mode and the addition image of the contact mode enhance the topography-induced effect, which is observed in the figure.

A closer look at the silicon ruler images at one pitch was taken, and the associated images are

shown in Fig. 8.26b. The surface height profiles in the TR mode and contact mode are somewhat different. The TR mode shows sharper edges than those in contact mode. The ratios of the change in amplitude at the steps to the change in the mean amplitude in the TR mode and in the contact mode are a measure of topography effects. The ratio in the contact mode ($\sim 85\%$) is about seven times larger than that of the TR mode ($\sim 12\%$).

8.3.5 Velocity Dependence on Micro/Nanoscale Friction

AFM/FFM experiments can be generally conducted at relative velocities as high as about 100–250 $\mu\text{m/s}$. To simulate applications, it is of interest to conduct friction experiments at higher velocities (up to 1 m/s). Furthermore, high velocity experiments would be useful to study velocity dependence on friction and wear. One approach has been to mount samples on a shear wave transducer (ultrasonic transducer) and then drive it at very high frequencies (in the MHz range) as reported earlier, see Fig. 8.10 (Yamanaka and Tomita 1995; Scherer et al. 1998, 1999; Marti and Krotli 2001; Reinstaedtler et al. 2003, 2005a, b). The coefficient of friction on the nanoscale is estimated based on the contact resonance frequency and requires the solution of the characteristic equations for the tip vibrating in contact with the sample surface. The approach is complex and is dependent upon various assumptions.

An alternative approach is to utilize piezo stages with large amplitude (~ 10 –100 μm) and relatively low resonance frequency (few kHz) and directly measure the friction force on microscale using the FFM signal without any analysis with assumptions used in the previous approaches using shear wave transducers. The commercial AFM setup modified with this approach can yield sliding velocities up to 200 mm/s (Tambe and Bhushan 2005a; Tao and Bhushan 2006b). In the high velocity piezo stage shown in Fig. 8.27a, the single axis piezo stage is oriented such that the scanning axis is perpendicular to the long axis of the AFM cantilever (which corresponds to the 90°

scan angle mode of the commercial AFM). The displacement is monitored using an integrated capacitive feedback sensor, located diametrically opposite to the piezo crystal. The capacitive change, corresponding to the stage displacement, gives a measure of the amount of displacement and can be used as feedback to the piezo controller for better guiding and tracking accuracy during scanning. The closed loop position control of the piezoelectric driven stages using capacitive feedback sensors provides linearity of motion better than 0.01% with nanometer resolution and a stable drift free motion (Tambe and Bhushan 2005a). In the ultrahigh velocity piezo stage shown in Fig. 8.27a, a rectangular monolithic piezoceramic plate (the stator) with two excitation electrodes is resonated using a 12 V power supply. Depending on the desired direction of the motion, the left or right electrode is excited to produce high-frequency eigenmode oscillations up to 200 kHz. Simultaneous eigenmodes result in quasi-elliptical motion. An alumina friction tip (pusher) attached to the plate pushes a slider with a glued friction bar which rests on a set of bearings. Through its contact with the friction bar, the piezoceramic plate provides micro-impulses and drives the slider forward or backward. While the longitudinal oscillation component provides the energy as driving force, the transverse component serves to change the pressure of the friction tip against the friction bar. The transverse oscillation energy determines the maximum frictional force and hence the holding and driving force of the stage. An optical position reference photosensor is located approximately in the middle of the travel range and is used to reference the absolute position of the stage within 1 μm repeatability. During motion, the increments of the linear scale from a home (reference) position point are converted to determine the position using a linear optical encoder.

The block diagram of the high speed data collection and processing system used for the friction force measurement is shown in Fig. 8.27b. During the experiments, the AFM cantilever is held stationary by maintaining a scan size of zero. The mounted sample is scanned below the AFM tip by moving stages, and the normal and torsional

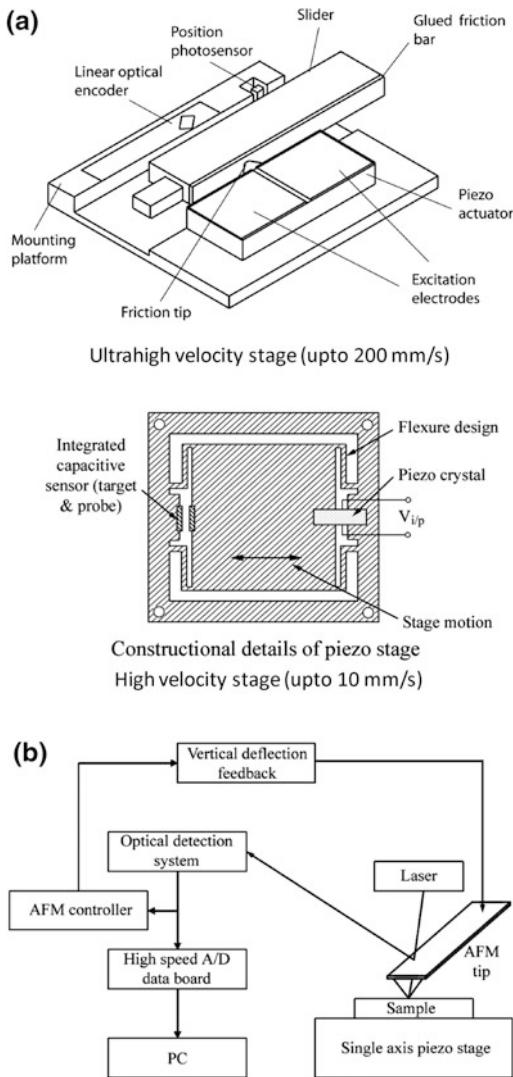


Fig. 8.27 Schematics of **a** an ultrahigh velocity piezo stage, **b** high velocity piezo stage, and **c** a block diagram of the high speed data collection and processing system used for friction force measurement (Tambe and Bhushan 2005a; Tao and Bhushan 2006b)

deflections of the tip are recorded by a photodiode detector. The raw deflection signals from the optical detection system are directly routed to a high speed data acquisition A/D board. Raw friction data is acquired at a high sampling rate up to 80 kilosamples/s.

Velocity dependence on friction for Si(100), diamond like carbon (DLC), self-assembled monolayer, and perfluoropolyether lubricant films were studied by Tambe and Bhushan (2004a, 2005a, b, c, e) and Tao and Bhushan (2006b, 2007). The friction force as a function of velocity for Si (100) and DLC (deposited by filtered cathodic arc) is shown in Fig. 8.28 on a logarithm velocity scale (middle column). The solid lines in the figure represent the results on a scan length of 1000 μm with velocity ranging from 1000 $\mu\text{m/s}$ to $2 \times 10^5 \mu\text{m/s}$ using the ultrahigh velocity stage. The dotted lines represent results on a 25 μm scan length with velocity ranging from 5 to 500 $\mu\text{m/s}$ using the high velocity stage. To clearly show the friction force dependence to velocity in the lower range, the test results with velocity varying from 5 to 500 $\mu\text{m/s}$ on 25 μm are shown on a magnified scale in the left column of Fig. 8.28.

On the Si (100) sample, the friction force decreased with velocity at low velocities ($v < 10 \mu\text{m/s}$) and then increased linearly with $\log(v)$ on a 25 μm scan length. On the 1000 μm scan length, the friction force increased linearly with $\log(v)$ when velocity was lower than $2 \times 10^4 \mu\text{m/s}$. When velocities were higher than $2 \times 10^4 \mu\text{m/s}$, the friction force increased linearly with velocity. For DLC, the friction force increased linearly with $\log(v)$ from 5 to 500 $\mu\text{m/s}$ on 25 μm scan length. On 1000 μm scan length, the friction force increases with velocity until about $2 \times 10^4 \mu\text{m/s}$ where the friction force reaches a maximum, then the friction force decreased with velocity.

For different samples, the change in the friction force with velocity involves different mechanisms due to the sample surface conditions. The silicon surface is hydrophilic, and DLC surface is nearly hydrophobic. Under the ambient condition, a thin layer of water film is condensed on a hydrophilic sample surface. On a hydrophobic surface with high contact angle, the water film would be difficult to form on the sample surface, and the effect of the water film

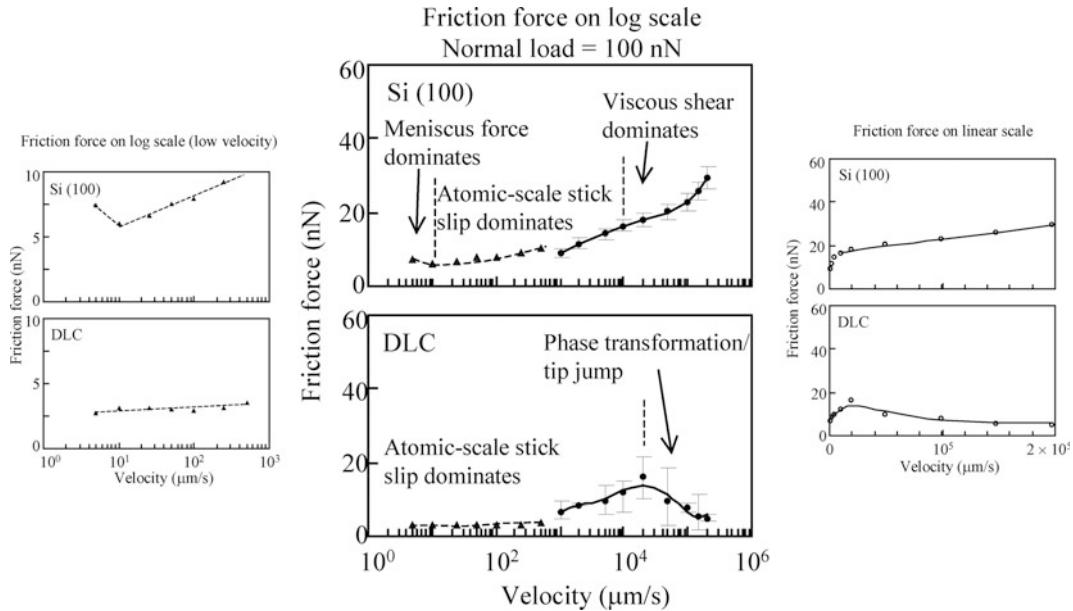


Fig. 8.28 Friction force as a function of sliding velocity obtained on 25 μm scan length using a high velocity stage (dotted line) and on 1000 μm scan length using an ultrahigh velocity stage (solid line). In the left and middle graphs, velocity is plotted on log scale. Left column shows at lower range of the velocity—between 1 and 500 $\mu\text{m}/\text{s}$. Right column shows the data at higher range of velocity on the linear scale (Tao and Bhushan 2007)

on the adhesive force and friction force could be neglected.

On the silicon surface, when the velocity is lower than 10 $\mu\text{m}/\text{s}$, the friction force decreased with velocity. This can be explained as follows. The water meniscus bridges develop as a function of time around the tip until reaching the equilibrium condition and are the dominant contributor to the friction force (Bhushan 1999a, 2011, 2013a, b). The motion of the tip results in continuous breaking and reforming of the meniscus bridges. As the tip sliding velocity exceeds a critical velocity (10 $\mu\text{m}/\text{s}$), there is not sufficient time for the menisci to reform, and the meniscus force would no longer play a dominant role. Between 10 and 2×10^4 $\mu\text{m}/\text{s}$, the friction increases linearly with $\log(v)$ on both 25 and 1000 μm scan lengths. This logarithmic dependence can be explained by the atomic-scale stick slip (Tambe and Bhushan 2005b; Tao and Bhushan 2007). At a velocity larger than 2×10^4 $\mu\text{m}/\text{s}$, the friction increases linearly with the velocity, and this trend can be explained by viscous shear; see the friction force plotted as a

function of velocity on a linear scale on a magnified scale in the right column of Fig. 8.28.

To explain atomic-scale stick slip mechanism of friction, the motion for the tip is expressed by a spring mass model (Zwerner et al. 1998) as follows:

$$m\ddot{x}_t = -\eta\dot{x}_t - k(x_M - x_t) - F \quad (8.6)$$

where m is the effective mass of the system, η the viscous damping coefficient, k the spring constant of the cantilever, $x_M = v_{MT}$ is the equilibrium position of the cantilever, x_t is the position of the tip, and F is the external force. The lateral force is expressed as $F_l = k(x_M - x_t)$, and the friction force F_{fric} is the lateral force averaged over time.

For velocities lower than 2×10^4 $\mu\text{m}/\text{s}$, the damping part ($\eta\dot{x}_t$) in Eq. 8.6 is comparatively low, and the atomic-scale stick slip is dominant. To investigate the stick slip, Tomlinson (1929) assumed the periodic surface with potential.

$$V(x) = V_0 \left(1 - \cos \frac{2\pi x}{a} \right) \quad (8.7)$$

where V_0 is surface barrier potential height, and a is the lattice constant of the surface. Then the force (F) in Eq. 8.6 is expressed as

$$F = V'(x) = \frac{2\pi V_0}{a} \sin \left(\frac{2\pi x}{a} \right) \quad (8.8)$$

Based on the Tomlinson model, and taking into account the effect of thermal activation, or the elastic energy stored in the cantilever during sliding, Gnecco et al. (2000) derived the relationship between the friction force and velocity, which is expressed as

$$F_{\text{stick-slip}} = F_0 + c \ln v \quad (8.9)$$

where F_0 and c are constants.

When the tip slides on a solid surface covered by viscous film such as water film at high velocities, the friction force (F_{fric}) is related to the velocity and viscosity of the film given by (Bhushan 2013a, b)

$$F_{\text{fric}} = \mu N + \eta \dot{\gamma} A \approx \mu N + \frac{\eta v A}{d} \quad (8.10)$$

where μ is the coefficient of friction between the dry sliding bodies, N is the applied load, τ is the shear stress, A is the real contact area, η is the viscosity of the film, $\dot{\gamma}$ is the velocity gradient, v is the sliding velocity, and d is the thickness of the film. From Eq. 8.10, the relationship between the friction force and the sliding velocity is linear when sliding on a viscous coating. The relationship is consistent with the conclusion by Helman et al. (1994) about the linear relationship between the friction force and the sliding velocity ($F_{\text{fric}} \approx \eta v M$) at high sliding velocities for a spring mass model in Eq. 8.6 which simulates the AFM tip sliding on a viscous liquid.

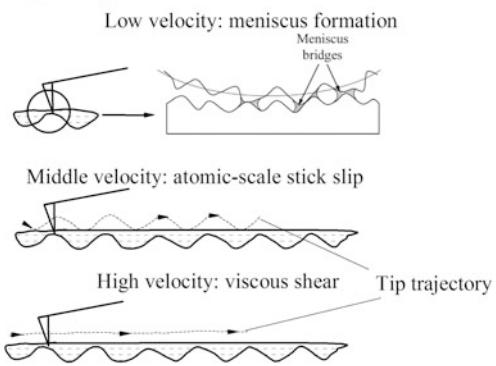
The sliding of the tip on a hydrophilic surface with water films at low velocities, middle velocities, and high velocities is schematically illustrated in Fig. 8.29a (Tao and Bhushan 2007). It should be noted that the stick slip mechanism

considered by Gnecco et al (2000) was based on the investigation on a dry surface. In this study, although the water was condensed on the Si (100) surface, the water film on the surface would not have significant effect on the energy dissipation due to the surface variation at relatively low velocities. Thus the linear relationship between friction and $\log(v)$ could be maintained. When the velocity increases above a certain value, the tip would lose direct contact with the sample surface and shear the water film. At velocities higher than $2 \times 10^4 \mu\text{m/s}$, the asperity deformation from the high velocity impact could be another mechanism as proposed by Tambe and Bhushan (2005b).

For the DLC film, since the surface is nearly hydrophobic, a uniform water film would not form on the surface. When sliding at a velocity lower than $1000 \mu\text{m/s}$, the friction force increased linearly with $\log(v)$, which could also be explained by atomic-scale stick slip. At velocities higher than $1000 \mu\text{m/s}$, the friction force increased with velocity until the local maximum at the velocity of $2 \times 10^4 \mu\text{m/s}$, then

(a)

A Tip sliding on a surface covered with water film



(b) A Tip sliding on a dry solid surface

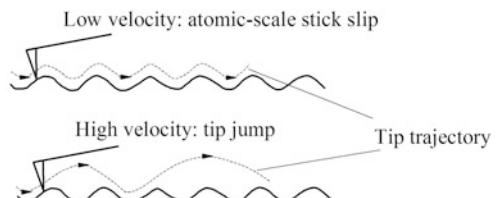


Fig. 8.29 Schematics of a tip sliding at different velocities on a water covered surface, and b dry surface (Tao and Bhushan 2007)

decreased with velocity. The decreasing trend in friction at higher velocities could be due to tip jump during sliding. The tip jump is illustrated in Fig. 8.29b. The tip jump results in the reduction of lateral force during sliding. Variation of the friction force with distance, indicative of the tip jump, was observed from the lateral force signal (not shown). When damping is low and velocity is high, the tip could jump several periodical cycles or several peaks (Fusco and Fasolino 2005). At a given low damping coefficient, the slip results in a low transient lateral force, as discussed by Fusco and Fasolino (2005). Thus the average lateral force (friction force) over the scan length is low. The tip jump could also cause high velocity impact of asperities on DLC surface, resulting in the phase transformation of DLC from sp^3 to sp^2 , as explained by Tambe and Bhushan (2005b). The layer of sp^2 phase can act as lubricant and reduce the interfacial friction.

8.3.6 Nanoscale Friction and Wear Mapping

Contrary to the classical friction laws postulated by Amontons and Coulomb centuries ago, nanoscale friction force is found to be strongly dependent on the normal load and sliding velocity. Many materials, coatings, and lubricants that have wide applications show reversals in friction behavior corresponding to transitions between different friction mechanisms (Tambe and Bhushan 2004a, 2005a, b, 2008, 2011). Most of the analytical models developed to explain nanoscale friction behavior have remained limited in their focus and have left investigators at a loss when trying to explain friction behavior scaling multiple regimes. Nanoscale friction maps provide fundamental insights into friction behavior. They help identify and classify the dominant friction mechanisms, as well as determine the critical operating parameters that influence transitions between different mechanisms (Tambe and Bhushan 2005b, c). Figure 8.30 shows a nanoscale friction map for DLC with the friction mapped as a function of the normal load and the sliding velocity (Tambe and Bhushan

2005f). The contours represent constant friction force lines. The friction force is seen to increase with the normal load as well as velocity. The increase in friction force with velocity is the result of atomic scale stick-slip. This is a result of thermal activation of the irreversible jumps of the AFM tip that arise from overcoming of the energy barrier between the adjacent atomic positions, as described earlier. The concentric contour lines corresponding to the constant friction force predicts a peak point, a point where the friction force reaches maxima and beyond which any further increase in the normal load or the sliding velocity results in a decrease in friction force. This characteristic behavior for DLC is the result of phase transformation of DLC into a graphite-like phase by sp^3 to sp^2 phase transition, as described earlier. During the AFM experiments, the Si_3N_4 tip gives rise to contact pressures in the range of 1.8–4.4 GPa for DLC for normal loads of 10–150 nN (Tambe and Bhushan 2005d). A combination of the high contact pressures that are encountered on the nanoscale and the high frictional energy dissipation arising from the asperity impacts at the tip-sample interface due to the high sliding velocities accelerates a phase transition process whereby a low shear strength graphite-like layer is formed at the sliding interface.

Friction force contour map for DLC
(contour lines are constant friction force lines)

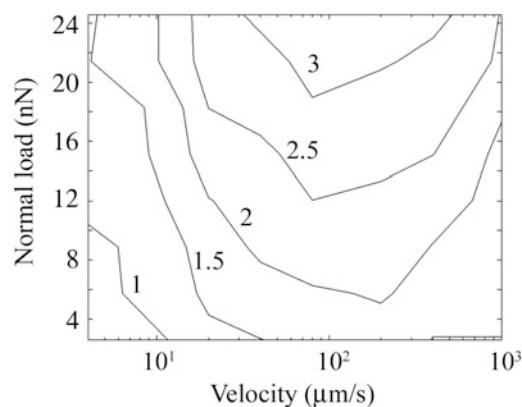


Fig. 8.30 Contour map showing friction force dependence on normal load and sliding velocity for DLC (Tambe and Bhushan 2005f)

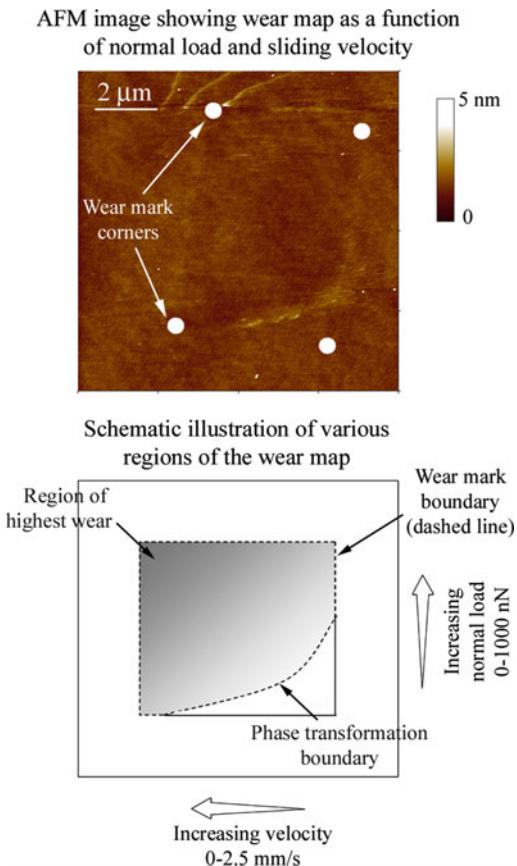


Fig. 8.31 Nanowear map (AFM image and schematic) illustrating the effect of sliding velocity and normal load on the wear of DLC resulting from phase transformation. Curved area shows debris lining and is indicative of the minimum frictional energy needed for phase transformation. For clarity, the wear mark corners are indicated by white dots in the AFM image (top) and the various zones of interest over the entire wear mark are schematically illustrated (Tambe and Bhushan 2005g)

Similar to friction mapping, one way of exploring the broader wear patterns is to construct wear mechanism maps that summarize data and models for wear, thereby showing mechanisms for any given set of conditions to be identified (Lim and Ashby 1987; Lim et al. 1987; Tambe and Bhushan 2005g, 2008). Wear of sliding surfaces can occur through one or more wear mechanisms, including adhesive, abrasive, fatigue, impact, corrosive, and fretting (Bhushan 2013a, b). Tambe and Bhushan (2005d, g) performed AFM experiments to develop nanoscale wear maps. Figure 8.31 shows a nanowear map

generated for a DLC sample by simultaneously varying the normal load and the sliding velocity over the entire scan area. The wear map was generated for a normal load range of 0–1000 nN and sliding velocity range of 0–2.5 mm/s. Wear debris, believed to be resulting from phase transformation of DLC by sp^3 to sp^2 phase transition, was seen to form only for a high value of sliding velocities times normal loads, i.e., only beyond a certain threshold of friction energy dissipation (Tambe and Bhushan 2005d, g). Hence the wear region exhibits a transition line indicating that for low velocities and low normal loads there is no phase transformation. For clarity, the wear mark corners are indicated by white dots in the AFM image (top) and the two zones of interest over the entire wear mark are schematically illustrated in Fig. 8.31 (top).

Nanoscale friction and wear mapping are novel techniques for investigating friction and wear behavior on the nanoscale over a range of operating parameters. By simultaneously varying the sliding velocity and normal load over a large range of values, nanoscale friction and wear behavior can be mapped, and the transitions between different wear mechanisms can be investigated. These maps help identify and demarcate critical operating parameters for different wear mechanisms and are very important tools in the process of design and selection of materials/coatings.

8.3.7 Adhesion and Friction in Wet Environment

8.3.7.1 Experimental Observations

Relative humidity affects adhesion and friction for dry and lubricated surfaces (Bhushan and Sundararajan 1998; Bhushan and Dandavate 2000; Bhushan 2003). Figure 8.32 shows the variation of single point adhesive force measurements as a function of tip radius on a Si(100) sample for several humidities. The adhesive force data are also plotted as a function of relative humidity for several tip radii. The general trend at humidities up to the ambient is that a 50-nm radius Si₃N₄ tip exhibits a lower adhesive force as compared to the other microtips of larger

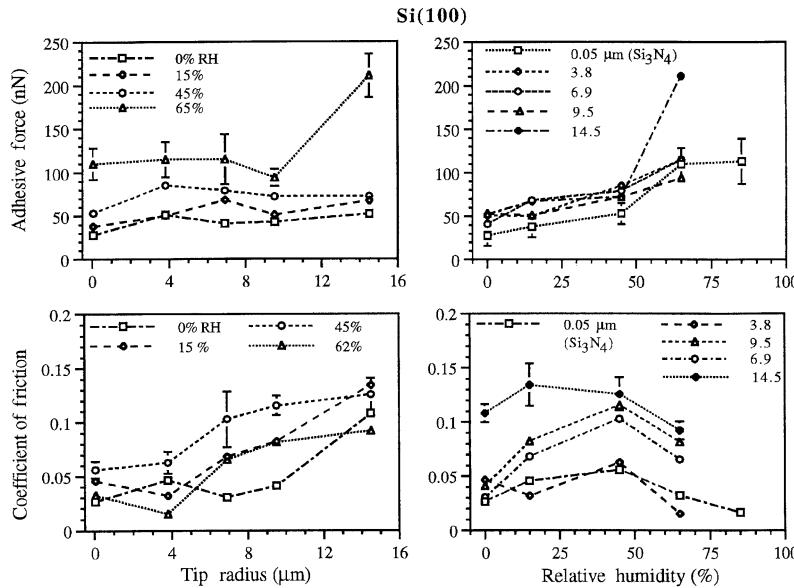


Fig. 8.32 Adhesive force and coefficient of friction as a function of tip radius at several humidities and as a function of relative humidity at several tip radii on Si(100) (Bhushan and Sundararajan 1998)

radii however; in the latter case, values are similar. Thus, for the microtips there is no appreciable variation in adhesive force with tip radius at a given humidity up to the ambient. The adhesive force increases as relative humidity increases for all tips.

The sources of adhesive force between a tip and a sample surface are van der Waals attraction and meniscus formation (Bhushan 2003, 2013a, b). The relative magnitudes of the forces from the two sources are dependent upon various factors including the distance between the tip and the sample surface, their surface roughness, their hydrophobicity, and the relative humidity (Stifter et al. 2000). For most rough surfaces, the meniscus contribution dominates at moderate to high humidities which arise from the capillary condensation of water vapor from the environment. If enough liquid is present to form a meniscus bridge, the meniscus force should increase with an increase in tip radius (proportional to tip radius for a spherical tip). In addition, an increase in the tip radius results in an increased contact area leading to higher values of van der Waals forces. However, if nanoasperities on the tip and the sample are considered then the

number of contacting and near-contacting asperities forming meniscus bridges increases with an increase of humidity leading to an increase in meniscus forces. These explain the trends observed in Fig. 8.32. From the data, the tip radius has little effect on the adhesive forces at low humidities but increases with tip radius at high humidity. Adhesive force also increases with an increase in humidity for all tips. This observation suggests that thickness of the liquid film at low humidities is insufficient to form continuous meniscus bridges to affect adhesive forces in the case of all tips.

Figure 8.32 also shows the variation in the coefficient of friction as a function of the tip radius at a given humidity, and as a function of the relative humidity for a given tip radius for Si(100). It can be observed that for 0% RH, the coefficient of friction is about the same for the tip radii except for the largest tip, which shows a higher value. At all other humidities, the trend consistently shows that the coefficient of friction increases with the tip radius. An increase in friction with tip radius at low to moderate humidities arises from the increased contact area (due to higher van der Waals forces) and higher

values of the shear forces required for a larger contact area. At high humidities, similar to adhesive force data, an increase with tip radius occurs because of both contact area and meniscus effects. Although the AFM/FFM measurements are able to measure the combined effect of the contribution of van der Waals and meniscus forces towards the friction force or adhesive force, it is difficult to measure their individual contributions separately. It can be seen that for all tips, the coefficient of friction increases with humidity to about ambient, beyond which it starts to decrease. The initial increase in the coefficient of friction with humidity arises from the fact that the thickness of the water film increases with an increase in the humidity, which results in a larger number of nanoasperities forming meniscus bridges and leads to higher friction (larger shear force). The same trend is expected with the microtips beyond 65% RH. This is attributed to the fact that at higher humidities, the adsorbed water film on the surface acts as a lubricant between the two surfaces. Thus the interface is changed at higher humidities, resulting in lower shear strength and hence lower friction force and coefficient of friction.

8.3.7.2 Adhesion and Friction Force Expressions for a Single Asperity Contact

We now obtain the expressions for the adhesive force and coefficient of friction for a single asperity contact with a meniscus formed at the interface, Fig. 8.33. For a spherical asperity of radius R in contact with a flat and smooth surface with the composite modulus of elasticity E^* and in the presence of liquid with a concave meniscus, the attractive meniscus force (adhesive force), designated as F_m or W_{ad} , is given as (Bhushan 2003, 2013a, b).

$$W_{ad} = 2\pi R \gamma (\cos \theta_1 + \cos \theta_2) \quad (8.11)$$

where γ is the surface tension of the liquid, and θ_1 and θ_2 are the contact angles of the liquid with surfaces 1 and 2, respectively. For an elastic contact for both extrinsic (W) and intrinsic (W_{ad}) normal load, the friction force is given as,

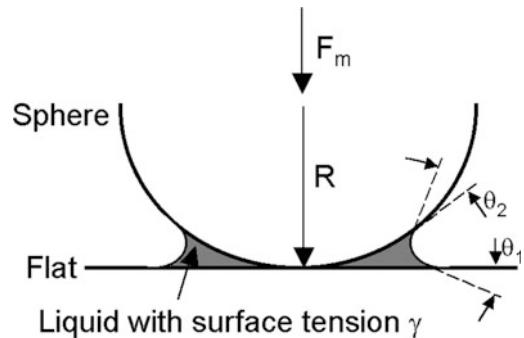


Fig. 8.33 Meniscus formation from a liquid condensate at the interface for a sphere in contact with a plane surface

$$F_e = \pi \tau \left[\frac{3(W + W_{ad})R}{4E^*} \right]^{2/3} \quad (8.12)$$

where W is the external load, and τ is the average shear strength of the contacts. (The surface energy effects are not considered here.) Note that adhesive force increases linearly with an increase in the tip radius, and the friction force increases with an increase in the tip radius as $R^{2/3}$ and with normal load as $(W + W_{ad})^{2/3}$. The experimental data in support of $W^{2/3}$ dependence on the friction force can be found in various references (see e.g., Schwarz et al. 1997). The coefficient of friction μ_e is obtained from Eq. 8.12 as

$$\mu_e = \frac{F_e}{(W + W_{ad})} = \pi \tau \left[\frac{3R}{4E^*} \right]^{2/3} \frac{1}{(W + W_{ad})^{1/3}} \quad (8.13)$$

In the plastic contact regime (Bhushan 2013a, b), the coefficient of friction μ_p is obtained as

$$\mu_p = \frac{F_p}{(W + W_{ad})} = \frac{\tau}{H_s} \quad (8.14)$$

where H_s is the hardness of the softer material. Note that in the plastic contact regime, the coefficient of friction is independent of external load, adhesive contributions and surface geometry.

For comparisons, for multiple asperity contacts in the elastic contact regime the total

adhesive force W_{ad} is the summation of adhesive forces at n individual contacts,

$$W_{ad} = \sum_{i=1}^n (W_{ad})_i \quad (8.15)$$

and

$$\mu_e \approx \frac{3.2\tau}{E^*(\sigma_p/R_p)^{1/2} + (W_{ad}/W)}$$

where σ_p and R_p are the standard deviation of summit heights and average summit radius, respectively. Note that the coefficient of friction depends upon the surface roughness. In the plastic contact regime, the expression for μ_p in Eq. 8.14 does not change.

The source of the adhesive force in a wet contact in the AFM experiments being performed in an ambient environment includes mainly attractive meniscus force due to capillary condensation of water vapor from the environment. The meniscus force for a single contact increases with an increase in the tip radius. A sharp AFM tip in contact with a smooth surface at low loads (on the order of a few nN) for most materials can be simulated as a single-asperity contact. At higher loads, for rough and soft surfaces, multiple contacts would occur. Furthermore, at low loads (nN range) for most materials, the local deformation would be primarily elastic. Assuming that the shear strength of contacts does not change, the adhesive force for smooth and hard surfaces at low normal load (on the order of few nN) (for a single asperity contact in the elastic contact regime) would increase with an increase in tip radius, and the coefficient of friction would decrease with an increase in total normal load as $(W + W_{ad})^{-1/3}$ and would increase with an increase of tip radius as $R^{2/3}$. In this case, Amontons law of friction which states that the coefficient of friction is independent of the normal load and is independent of the apparent area of contact does not hold. For a single-asperity plastic contact and multiple-asperity plastic contacts, neither the normal load nor tip radius comes into play in the calculation of the coefficient of friction. In the case of multiple-asperity

contacts, the number of contacts increase with an increase of normal load, therefore the adhesive force increases with an increase in load.

In the data presented earlier in this section, the effect of tip radius and humidity on the adhesive forces and the coefficient of friction is investigated for experiments with Si(100) surface at loads in the range of 10–100 nN. The multiple asperity elastic-contact regime is relevant for this study involving large tip radii. An increase in humidity generally results in an increase in the number of meniscus bridges, which would increase the adhesive force. As it was suggested earlier, that increase in humidity also may decrease the shear strength of contacts. A combination of an increase in adhesive force and a decrease in shear strength would affect the coefficient of friction. An increase in the tip radius would increase the meniscus force (adhesive force). A substantial increase in the tip radius may also increase interatomic forces. These effects influence the coefficient of friction with an increase in the tip radius.

8.3.8 Separation Distance Dependence of Meniscus and van der Waals Forces

When two surfaces are in close proximity, sources of adhesive forces are weak van der Waals attraction and meniscus formation. Relative magnitudes of the forces from the two sources are dependent upon various factors including the interplanar separation, their surface roughness, their hydrophobicity, and relative humidity (liquid volume) (Stifter et al. 2000). Meniscus contribution dominates at moderate to high humidities, and van der Waals forces dominate at asperities few nm apart. In some micro/nanocomponents, it is important to know the relative contribution of two sources as a function of given interplanar separation so as to design an interface for low adhesion. For example, if two ultrasmooth surfaces come in close proximity with interplanar separation on the order of a nm, van der Waals forces may

dominate, and their magnitude may be reduced by creating bumps on one of the interfaces. This analysis is also of interest in AFM studies to understand distance dependence of adhesive forces as the tip goes in and out of contact.

Stifter et al. (2000) modeled the contact of a parabolic shaped tip and a flat, smooth sample surface. The tip may represent a surface asperity on an interface or an AFM tip in an AFM experiment. They calculated van der Waals and meniscus forces as a function of different parameters, namely, tip geometry, tip-sample starting distance, relative humidity, surface tension, and contact angles. They compared the meniscus forces with van der Waals forces to understand their relative importance in various operating conditions.

The interacting force between tip and sample in dry conditions is the Lennard-Jones force derived from Lennard-Jones potential. Lennard-Jones potential is composed of two interactions—the van der Waals attraction and the Pauli repulsion. van der Waals forces are significant because they are always present. For a parabolic tip above a half plane with a distance D between tip and plane, the Lennard-Jones potential is obtained by integrating the atomic potential over the volume of the tip and sample. It is given as (Stifter et al. 2000).

$$V(D) = \frac{c}{12} \left(-\frac{A}{D} + \frac{B}{210D^7} \right) \quad (8.16)$$

where c is the width of the parabolic tip (= diameter in the case of a spherical tip), and A and B are two potential parameters, where A is Hamakar constant. This equation provides expressions for attractive and repulsive parts. The calculations were made for Lennard-Jones force (total) and van der Waals force (attractive part) for two Hamakar constants: 0.04×10^{-19} J (representative of polymers) and 3.0×10^{-19} J (representative of ceramics), and meniscus force for a water film ($\gamma_\ell = 72.5$ N/m). Figure 8.34 shows various forces as a function of separation distance. The effect of two relative humidities and three tip radii was also studied which affect meniscus forces. The two dashed curves indicate

the spread of possible van der Waals forces for two Hamakar constants. The figure shows that meniscus forces exhibit weaker distance dependence. The meniscus forces can be stronger or weaker than van der Waals forces for distances smaller than about 0.5 nm. For longer distances, the meniscus forces are stronger than the van der Waals forces. van der Waals forces must be considered for a tip-sample distance up to few nm ($D < 5$ nm). The meniscus forces operate up to break of the meniscus to range from 5 to 20 nm (Stifter et al. 2000).

8.3.9 Scale Dependence in Friction

Table 8.3 presents the adhesive force and the coefficient of friction data obtained on the nanoscale and microscale (Ruan and Bhushan 1994a; Liu and Bhushan 2003b; Bhushan et al. 2004b; Tambe and Bhushan 2004a). Adhesive force and coefficient of friction values on the nanoscale are about half to one order of magnitude lower than that on the microscale. Scale dependence is clearly observed in this data. As a further evidence of scale dependence, Table 8.4 shows the coefficient of friction measured for Si (100), HOPG, natural diamond, and DLC on the nanoscale and microscales. It is clearly observed that friction values are scale dependent.

To estimate the scale length, the apparent contact radius at test loads is calculated and presented in Table 8.3. Mean apparent pressures are also calculated and presented. For nanoscale AFM experiments, it is assumed that an AFM tip coming in contact with a flat surface represents a single asperity and elastic contact, and Hertz analysis was used for the calculations. In the microscale experiments, a ball coming in contact with a flat surface represents multiple asperity contacts due to the roughness, and the contact pressure of the asperity contacts is higher than the apparent pressure. For calculation of a characteristic scale length for multiple asperity contacts, which is equal to the apparent length of contact, Hertz analysis was also used. This analysis provides an upper limit on apparent radius and lower limit on the mean contact pressure.

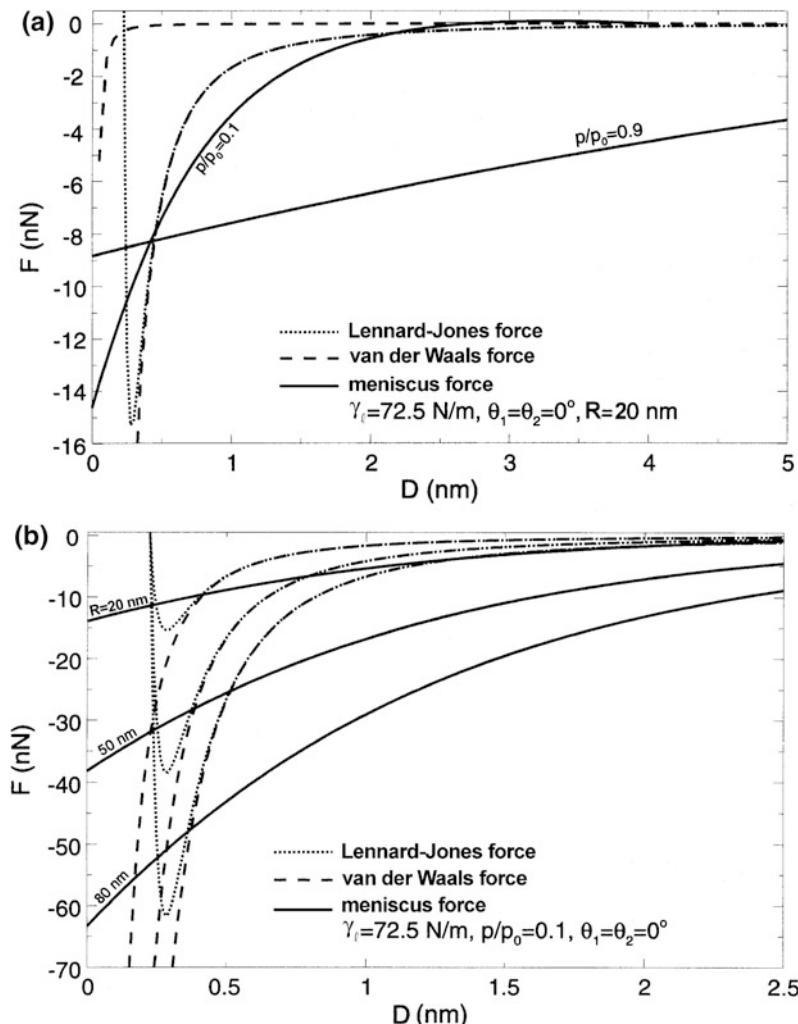


Fig. 8.34 Relative contribution of meniscus, van der Waals- and Lennard-Jones forces (F) as a function of separation distance (D) and at **a** two values of relative humidity (p/p_0) for tip radius of 20 nm and Hamaker constants of $0.04 \times 10^{-19} \text{ J}$ and $3.0 \times 10^{-19} \text{ J}$, and **b** three tip radii (R) and Hamaker constant of $3.0 \times 10^{-19} \text{ J}$ (Stifter et al. 2000)

Table 8.3 Micro- and nanoscale values of adhesive force and coefficient of friction in micro- and nanoscale measurements (Bhushan et al. 2004b)

Sample	Adhesive force		Coefficient of friction	
	Microscale ^a (μN)	Nanoscale ^b (nN)	Microscale ^a	Nanoscale ^b
Si(100)	685	52	0.47	0.06
DLC	325	44	0.19	0.03
Z-DOL	315	35	0.23	0.04
HDT	180	14	0.15	0.006

^aVersus 500-μm radius Si(100) ball

^bVersus 50-nm radius Si_3N_4 tip

Table 8.4 Micro- and nanoscale values of the coefficient of friction, typical physical properties of specimen, and calculated apparent contact radii and apparent contact pressures at loads used in micro- and nanoscale measurements

Sample	Coefficient of friction	Elastic modulus (GPa)		Poisson's ratio	Hardness (GPa)	Apparent contact radius at test load for		Mean apparent pressure at test load for Microscale (GPa) (lower limit) Nanoscale (GPa)
		Microscale	Nanoscale			Microscale (μm) (upper limit)	Nanoscale (nm)	
Si(100) wafer	0.47 ^a	0.06 ^c	130 ^{e,f}	0.28 ^f	9–10 ^{e,f}	0.8–2.2 ^a	1.6–3.4 ^e	0.05–0.13 ^a
Graphite (HOPG)	0.1 ^b	0.006 ^c	9–15 ^e (9)	–(0.25)	0.01 ^j	62 ^b	3.4–7.4 ^e	0.082 ^b
Natural diamond	0.2 ^b	0.05 ^c	1140 ^h	0.07 ^h	80–104 ^{g,h}	21 ^b	1.1–2.5 ^e	0.74 ^b
DLC film	0.19 ^a	0.03 ^d	280 ⁱ	0.25 ⁱ	20–30 ⁱ	0.7–2.0 ^a	1.3–2.9 ^d	0.06–0.16 ^a

^a500 μm radius Si(100) ball at 100–2000 μN and 720 μm/s in dry air, Bhushan et al. (2004b)

^b3 mm radius Si₃N₄ ball (Elastic modulus 310 GPa, Poisson's ratio 0.22 (Bhushan and Gupta, 1991)) at 1 N and 800 μm/s, Ruan and Bhushan (1994a)

^c50 nm radius Si₃N₄ tip at load range from 10–100 nN and 0.5 nm/s, in dry air, Ruan and Bhushan (1994a)

^d50 nm radius Si₃N₄ tip at load range from 10–100 nN in dry air, Bhushan et al. (2004b)

^eBhushan and Venkatesan (1993)

^fAnonymous (2002)

^gBhushan and Gupta (1991)

^hField (1992)

ⁱBhushan (1999d)

^jAnonymous (1959)

For calculation purposes it is assumed that contacts on micro- and nanoscale are single asperity elastic contacts. (Nosonovsky and Bhushan 2005)

There are several factors responsible for the differences in the coefficients of friction at the micro- and nanoscale. Among them are the contributions from wear and contaminant particles, transition from elasticity to plasticity, and meniscus effect (Bhushan and Nosonovsky 2003, 2004a, b; Nosonovsky and Bhushan 2005). The contribution of wear and contaminant particles is more significant at the macro/microscale because of the larger number of trapped particles, referred to as third body contribution. It can be argued that for the nanoscale AFM experiments the asperity contacts are predominantly elastic (with average real pressure being less than the hardness of the softer material), and adhesion is the main contribution to the friction, whereas for the microscale experiments the asperity contacts are predominantly plastic, and deformation is an important factor. It will be shown later that hardness has scale effect; it increases with decreasing scale and is responsible for less deformation on a smaller scale. The meniscus effect results in an increase of friction with increasing tip radius (Fig. 8.32). Therefore, third body contribution, scale dependent hardness, and other properties transition from elastic contacts in nanoscale contacts to plastic deformation in microscale contacts, and meniscus contribution plays an important role.

Friction is a complex phenomenon, which involves asperity interactions involving adhesion and deformation (plowing). Adhesion and plastic deformation imply energy dissipation, which is responsible for friction (Fig. 8.35) (Bhushan 2013a, b). A contact between two bodies takes place on high asperities, and the real area of contact (A_r) is a small fraction of the apparent area of contact. During contact of two asperities, a lateral force may be required for asperities of a given slope to climb against each other. This mechanism is known as the ratchet mechanism, and it also contributes to the friction. Wear and contaminant particles present at the interface, referred as the “third body,” also contribute to the friction (Fig. 8.35). In addition, during contact even at low humidity, a meniscus is formed (Fig. 8.33). Generally any liquid that wets or has a small contact angle on surfaces will condense

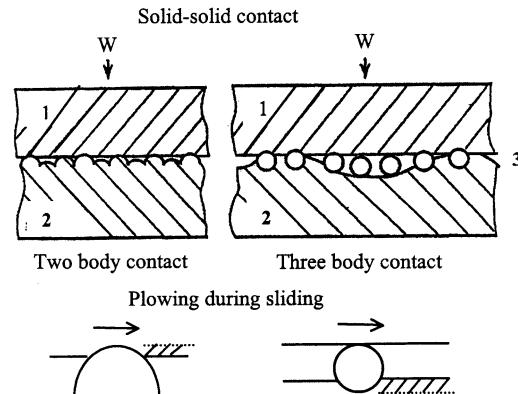


Fig. 8.35 Schematic of two-bodies and three-bodies during dry contact of rough surfaces

from vapor into cracks and pores on surfaces as bulk liquid and in the form of annular-shaped capillary condensate in the contact zone. A quantitative theory of scale effects in friction should consider the scale effect on physical properties relevant to various contributions.

According to the adhesion and deformation model of friction, the coefficient of dry friction μ is a sum of the adhesion component μ_a and the deformation (plowing) component μ_d . The latter, in the presence of particles, is a sum of the asperity summits deformation component μ_{ds} and particles deformation component μ_{dp} , so that the total coefficient of friction is (Bhushan and Nosonovsky 2004a)

$$\begin{aligned}\mu &= \mu_a + \mu_{ds} + \mu_{dp} = \frac{F_a + F_{ds} + F_{dp}}{W} \\ &= \frac{A_{ra}\tau_a + A_{ds}\tau_{ds} + A_{dp}\tau_{dp}}{W}\end{aligned}\quad (8.17)$$

where W is the normal load, F is the friction force, A_{ra} , A_{ds} , A_{dp} are the real areas of contact during adhesion, two body deformation and with particles, respectively, and τ is the shear strength. The subscripts a , ds , and dp correspond to adhesion, summit deformation, and particle deformation, respectively.

The adhesional component of friction depends on the real area of contact and adhesion shear strength. The real area of contact is scale dependent due to the scale dependence of the surface roughness (for elastic and plastic contact) and due

to the scale dependence of hardness (for plastic contact) (Bhushan and Nosonovsky 2004a). We limit the analysis here for multiple-asperity contact. For this case, scale L is defined as apparent size of contact between two bodies. (For completeness, for single-asperity contact, scale is defined as the contact diameter.) It is suggested by Bhushan and Nosonovsky (2003) that for many materials, dislocation-assisted sliding (microslip) is the main mechanism which is responsible for the shear strength. They considered dislocation assisted sliding based on the assumption that contributing dislocations are located in a subsurface volume. The thickness of this volume is limited by the distance which dislocations can climb ℓ_s (material parameter) and by the radius of contact a. They showed that τ_a is scale dependent. Based on this, the adhesional components of the coefficient of friction in the case of elastic contact μ_{ae} and in the case of plastic contact μ_{ap} are given as (Bhushan and Nosonovsky 2003)

$$\mu_{ae} = \frac{\mu_{ae0}}{\sqrt{\ell + (\ell_s \sqrt{a_0})}} \left(\frac{L}{L_{lc}} \right)^{m-n} \sqrt{1 + (L_s/L)^m}, L < L_{lc} \quad (8.18)$$

$$\mu_{ap} = \mu_{ap0} \sqrt{\frac{1 + (\ell_d \sqrt{a_0})}{1 + (\ell_s \sqrt{a_0})}} \sqrt{\frac{1 + (L_s/L)^m}{1 + (L_d/L)^m}}, L < L_{lc} \quad (8.19)$$

where μ_{ae0} and μ_{ap0} are values of the coefficient of friction at the macroscale ($L \geq L_{lc}$), m and n are indices which characterize the scale dependence of surface parameters, $\overline{a_0}$ is the macroscale value of mean contact radius, L_{lc} is long wavelength limit for scale dependence of the contact parameters, ℓ_s and ℓ_d are material specific characteristic length parameters, and L_s and L_d are length parameters related to ℓ_s and ℓ_d . Scale dependence of the adhesional component of the coefficient of friction is presented in Fig. 8.36, based on Eqs. 8.18 and 8.19.

Based on the assumption that multiple asperities of two rough surfaces in contact have a conical shape, the two-body deformation component of friction can be determined as (Bhushan 2013a, b)

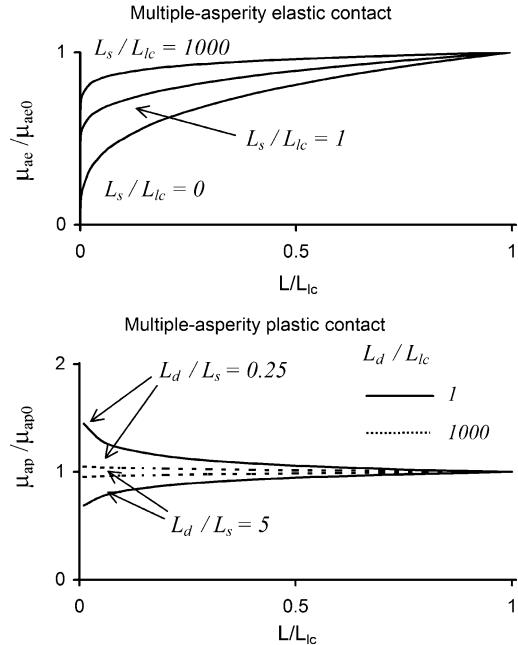


Fig. 8.36 Normalized results for the adhesional component of coefficient of friction, as a function of L/L_{lc} for multiple asperity contact. Data are presented for $m = 0.5$, $n = 0.2$. For multiple-asperity plastic contact, data are presented for two values of L_d/L_{lc} (Bhushan and Nosonovsky 2004a)

$$\mu_{ds} = \frac{2 \tan \theta_r}{\pi} \quad (8.20)$$

where θ_r is the roughness angle (or attack angle) of a conical asperity. Mechanical properties affect the real area of contact and shear strength, and these cancel out in Eq. 8.16 (Bhushan and Nosonovsky 2004a). Based on a statistical analysis of a random Gaussian surface (Bhushan and Nosonovsky 2004a)

$$\mu_{ds} = \frac{2\sigma_0}{\pi\beta_0^*} \left(\frac{L}{L_{lc}} \right)^{n-m} = \mu_{ds0} \left(\frac{L}{L_{lc}} \right)^{n-m}, L < L_{lc} \quad (8.21)$$

where μ_{ds0} is the value of the coefficient of summits deformation component of the coefficient of friction at macroscale ($L \geq L_{lc}$), and σ_0 and β_0^* are macroscale values of standard deviation of surface heights and correlation length, respectively, for a Gaussian surface. The scale

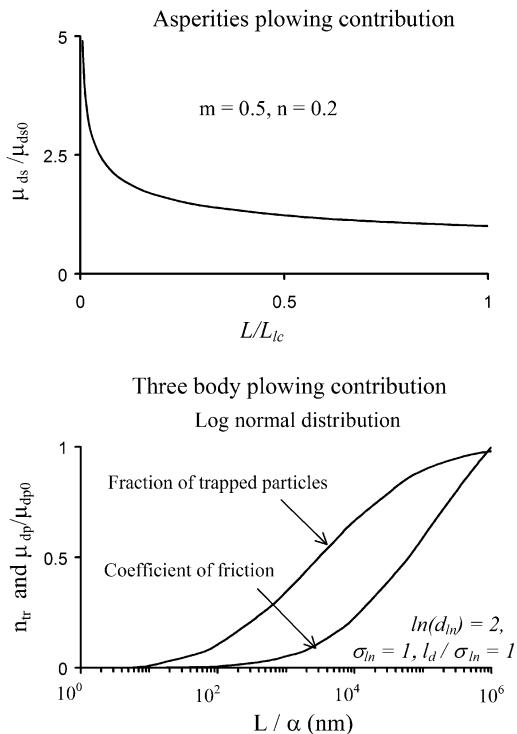


Fig. 8.37 Normalized results for the two-body deformation component of the coefficient of friction, and the number of trapped particles divided by the total number of particles and three-body deformation component of the coefficient of friction, normalized by the macroscale value for log-normal distribution of debris size, where α is the probability of a particle in the border zone to leave the contact region. Various constants given in the figure correspond to log normal distribution (Bhushan and Nosonovsky 2004a)

dependence for the two-body deformation component of the coefficient of friction is presented in Fig. 8.37 (top curve) for $m = 0.5$, $n = 0.2$, based on Eq. 8.21. The coefficient of friction increases with decreasing scale, according to Eq. 8.21. This effect is a consequence of increasing average slope or roughness angle.

For the three-body deformation, it is assumed that wear and contaminant particles at the borders of the contact region are likely to leave the contact region, while the particles in the center are likely to stay (Fig. 8.38). The plowing three-body deformation is plastic and, assuming that particles are harder than the bodies, the shear strength τ_{dp} is equal to the shear yield strength of

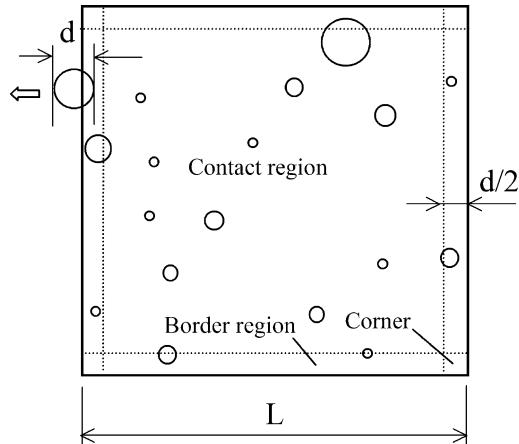


Fig. 8.38 Schematics of debris at the contact zone and at its border region. A particle of diameter d in the border region of $d/2$ is likely to leave the contact zone (Bhushan and Nosonovsky 2004a)

the softer body τ_Y , the three-body deformation component of the coefficient of friction is given by (Bhushan and Nosonovsky 2004b)

$$\mu_{dp} = \mu_{dp0} n_{tr} \frac{\bar{d}^2}{\bar{d}_0^2} \frac{\sqrt{1 + 2\ell_d \sqrt{d}}}{\sqrt{1 + 2\ell_d \sqrt{d_0}}} \quad (8.22)$$

where \bar{d} is the mean particle diameter, \bar{d}_0 is the macroscale value of mean particle diameter, n_{tr} is the number of trapped particles divided by total number of particles, and μ_{dp0} is the macroscale ($L \rightarrow \infty$, $n_{tr} \rightarrow 1$) value of the third-body deformation component of the coefficient of friction. Scale dependence of μ_{dp} is shown in Fig. 8.37 (bottom curve) based on Eq. 8.22. Based on the scale effect predictions presented in Figs. 8.36 and 8.37, trends in the experimental results in Table 8.3 can be explained.

Scale dependence of meniscus effects in friction, wear, and interface temperature can be analyzed in a similar way (Bhushan and Nosonovsky 2004b).

To demonstrate the load dependence of friction at the nano/microscale, coefficient of friction as a function of normal load is presented in Fig. 8.39. The coefficient of friction was measured by Bhushan and Kulkarni (1996) for Si_3N_4 tip versus Si, SiO_2 , and natural diamond using an AFM. They reported that for low loads, the

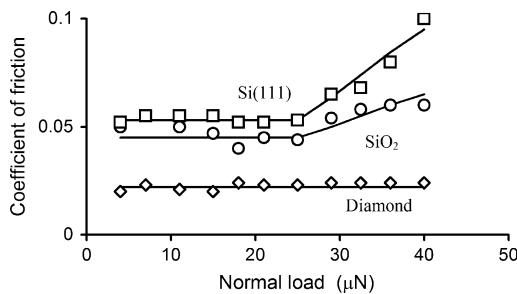


Fig. 8.39 Coefficient of friction as a function of normal load and for Si(111), SiO_2 coating and natural diamond. Inflections in the curves for silicon and SiO_2 correspond to the contact stresses equal to the hardnesses of these materials (Bhushan and Kulkarni 1996)

coefficient of friction is independent of load and increases with increasing load after a certain load. It is noted that the critical value of loads for Si and SiO_2 corresponds to stresses equal to their hardness values, which suggests that the transition to plasticity plays a role in this effect. The friction values at higher loads for Si and SiO_2 approach that of macroscale values.

8.4 Micro/Nanoscale Wear and Scratching, Local Deformation, and Nanofabrication/ Nanomachining

8.4.1 Nanoscale Wear

Bhushan and Ruan (1994) conducted nanoscale wear tests on polymeric magnetic tapes using conventional silicon nitride tips at two different loads of 10 and 100 nN (Fig. 8.40). For a low normal load of 10 nN, measurements were made twice. There was no discernible difference between consecutive measurements for this load. However, as the load was increased from 10 nN to 100 nN, topographical changes were observed during subsequent scanning at a normal load of 10 nN; material was pushed in the sliding direction of the AFM tip relative to the sample. The material movement is believed to occur as a result of plastic deformation of the tape surface.

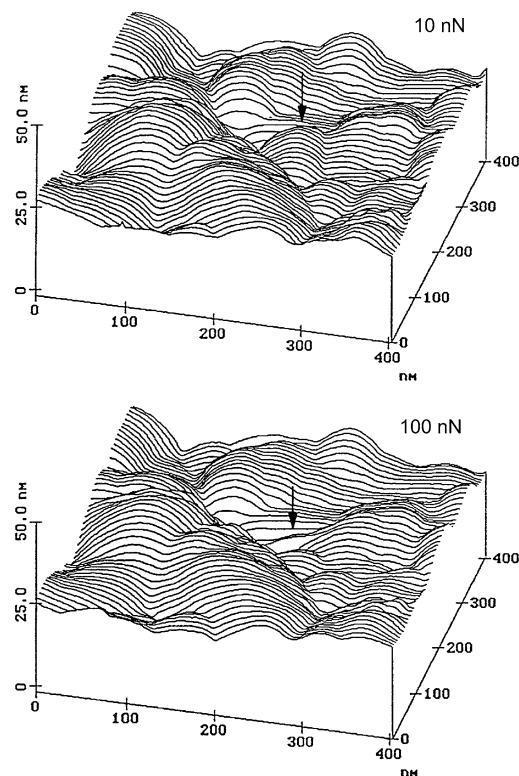


Fig. 8.40 Surface roughness maps of a polymeric magnetic tape at the applied normal load of 10 and 100 nN. Location of the change in surface topography as a result of nanowear is indicated by arrows (Bhushan and Ruan 1994)

Thus, deformation and movement of the soft materials on a nanoscale can be observed.

8.4.2 Microscale Scratching

The AFM can be used to investigate how surface materials can be moved or removed on micro- to nanoscales, for example, in scratching and wear (Bhushan 1999a, 2011) (where these things are undesirable) and nanofabrication/nanomachining (where they are desirable). Figure 8.41a shows microscratches made on Si(111) at various loads and a scanning velocity of 2 $\mu\text{m}/\text{s}$ after 10 cycles (Bhushan et al. 1994). As expected, the scratch depth increases linearly with load. Such microscratching measurements can be used to study

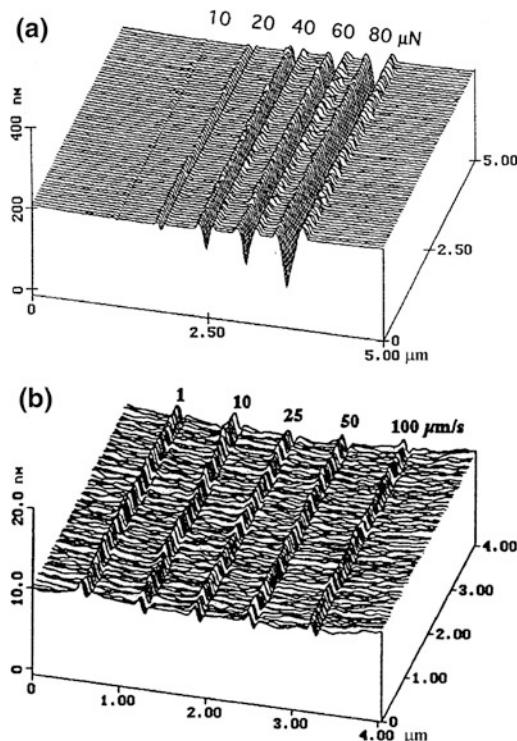


Fig. 8.41 Surface plots of **a** Si(111) scratched for ten cycles at various loads and a scanning velocity of 2 $\mu\text{m}/\text{s}$ (Bhushan et al. 1994). Note that x and y axes are in μm and z axis is in nm, and **b** Si(100) scratched in one unidirectional scan cycle at a normal force of 80 μN and different scanning velocities (Bhushan and Sundararajan 1998)

failure mechanisms on the microscale and to evaluate the mechanical integrity (scratch resistance) of ultra-thin films at low loads.

To study the effect of scanning velocity, unidirectional scratches 5 μm in length were generated at scanning velocities ranging from 1 to 100 $\mu\text{m}/\text{s}$ at various normal loads ranging from 40 to 140 μN (Bhushan and Sundararajan 1998). There is no effect of scanning velocity obtained at a given normal load. For representative scratch profiles at 80 μN , see Fig. 8.41b. This may be because of a small effect of frictional heating with the change in scanning velocity used here. Furthermore, for a small change in interface temperature, there is a large underlying volume to dissipate the heat generated during scratching.

Scratching can be performed under ramped loading to determine the scratch resistance of materials and coatings (Sundararajan and Bhushan 2001). The coefficient of friction is measured during scratching, and the load at which the coefficient of friction increases rapidly is known as the “critical load,” which is a measure of scratch resistance. In addition, post-scratch imaging can be performed *in situ* with the AFM in tapping mode to study failure mechanisms. Figure 8.42 shows data from a scratch test on Si(100) with a scratch length of 25 μm and a scratching velocity of 0.5 $\mu\text{m}/\text{s}$. At the beginning of the scratch, the coefficient of friction is 0.04, which indicates a typical value for silicon. At about 35 μN (indicated by the arrow in the figure), there is a sharp increase in the coefficient of friction, which indicates the critical load. Beyond the critical load, the coefficient of friction continues to increase steadily. In the post-scratch image, we note that at the critical load, a clear groove starts to form. This implies that Si(100) was damaged by plowing at the critical load, associated with the plastic flow of the material. At and after the critical load, small and uniform debris is observed, and the amount of debris increases with increasing normal load. Sundararajan and Bhushan (2001) have also used this technique to measure the scratch resistance of diamondlike carbon coatings ranging in thickness from 3.5 to 20 nm.

8.4.3 Microscale Wear

By scanning the sample in two dimensions with the AFM, wear scars are generated on the surface. Figure 8.43 shows the effect of normal load on wear depth on Si(100). We note that wear depth is very small below 20 μN of normal load (Koinkar and Bhushan 1997c; Zhao and Bhushan 1998). A normal load of 20 μN corresponds to contact stresses comparable to the hardness of silicon. Primarily, elastic deformation at loads below 20 μN is responsible for low wear (Bhushan and Kulkarni 1996).

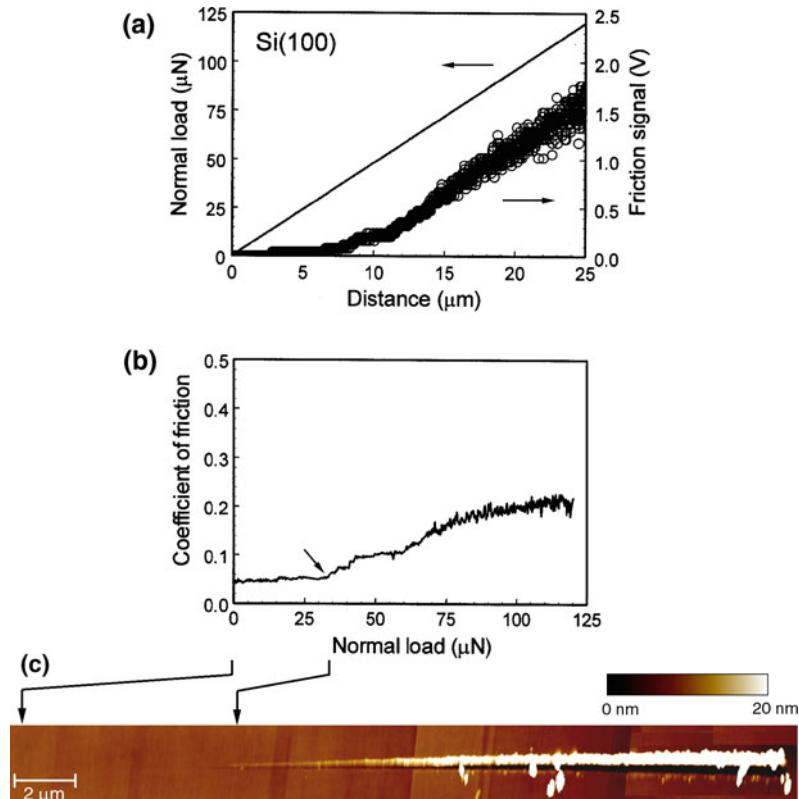


Fig. 8.42 **a** Applied normal load and friction signal measured during the microscratch experiment on Si(100) as a function of scratch distance, **b** friction data plotted in the form of coefficient of friction as a function of normal load, and **c** AFM surface height image of scratch obtained in tapping mode (Sundararajan and Bhushan 2001)

A typical wear mark of the size $2 \mu\text{m} \times 2 \mu\text{m}$ generated at a normal load of $40 \mu\text{N}$ for one scan cycle and imaged using AFM with scan size of $4 \mu\text{m} \times 4 \mu\text{m}$ at 300nN load is shown in Fig. 8.44a. The inverted map of wear marks

shown in Fig. 8.44b indicates the uniform material removal at the bottom of the wear mark (Koinkar and Bhushan 1997c). An AFM image of the wear mark shows debris at the edges, probably swiped during AFM scanning. This indicates that the debris is loose (not sticky) and can be removed during the AFM scanning.

Next the mechanism of material removal on the microscale in AFM wear experiments is examined (Koinkar and Bhushan 1997c; Bhushan and Sundararajan 1998; Zhao and Bhushan 1998). Figure 8.45 shows a secondary electron image of the wear mark and associated wear particles. The specimen used for the scanning electron microscope (SEM) was not scanned with the AFM after initial wear, in order to retain wear debris in the wear region. Wear debris is clearly observed. In the SEM micrographs, the wear debris appear to be agglomerated because of the high surface

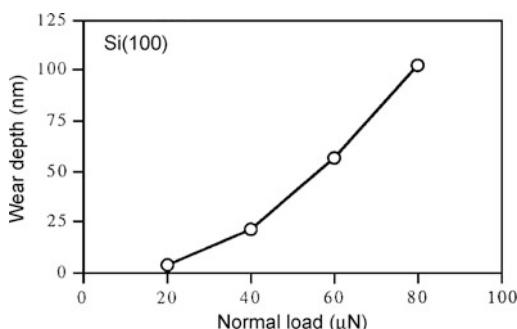


Fig. 8.43 Wear depth as a function of normal load for Si(100) after one cycle (Zhao and Bhushan 1998)

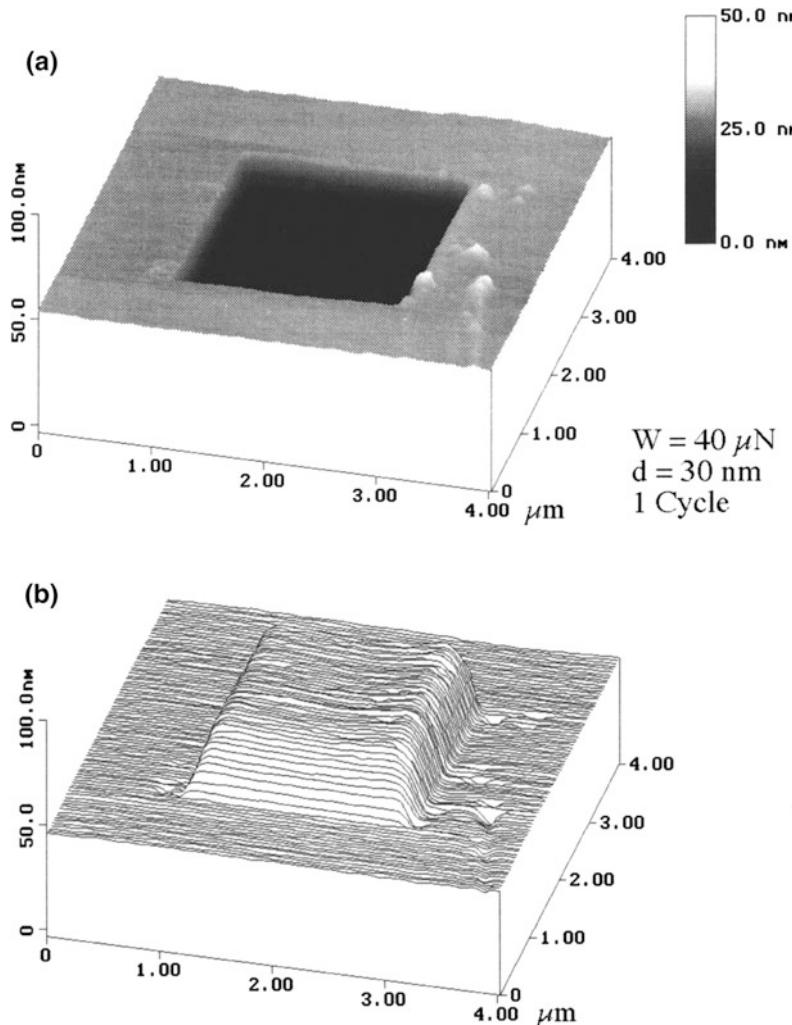


Fig. 8.44 **a** Typical gray scale and **b** inverted AFM images of wear mark created using a diamond tip at a normal load of $40 \mu\text{N}$ and one scan cycle on Si(100) surface (Koinkar and Bhushan 1997c)

energy of the fine particles. Particles appear to be a mixture of rounded and so-called cutting type (feather-like or ribbon-like material). Zhao and Bhushan (1998) reported an increase in the number and size of cutting type particles with the normal load. The presence of cutting type particles indicates that the material is removed primarily by plastic deformation.

To better understand the material removal mechanisms, Zhao and Bhushan (1998) used transmission electron microscopy (TEM). The TEM micrograph of the worn region and

associated diffraction pattern are shown in Fig. 8.46a, b. The bend contours are observed to pass through the wear mark in the micrograph. The bend contours around and inside the wear mark are indicative of a strain field, which in the absence of applied stresses can be interpreted as plastic deformation and/or elastic residual stresses. Often, localized plastic deformation during loading would lead to residual stresses during unloading; therefore, bend contours reflect a mix of elastic and plastic strains. The wear debris is observed outside the wear mark. The enlarged

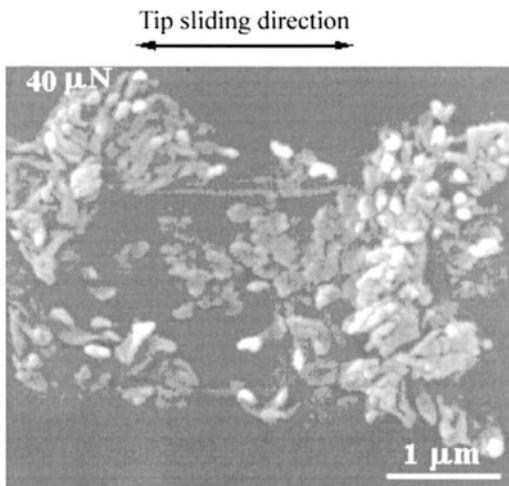


Fig. 8.45 Secondary electron image of wear mark and debris for Si(100) produced at a normal load of 40 μN and one scan cycle (Zhao and Bhushan 1998)

view of the wear debris in Fig. 8.46c shows that much of the debris is ribbon like, indicating that material is removed by a cutting process via plastic deformation, which is consistent with the SEM observations. The diffraction pattern from inside the wear mark was reported to be similar to that of virgin silicon, indicating no evidence of any phase transformation (amorphization) during wear. A selected area diffraction pattern of the wear debris shows some diffuse rings, which indicates the existence of amorphous material in the wear debris, confirmed as silicon oxide products from chemical analysis. It is known that plastic deformation occurs by generation and propagation of dislocations. No dislocation activity or cracking was observed at 40 μN . However, dislocation arrays could be observed at 80 μN . Figure 8.47 shows the TEM micrographs of the worn region at 80 μN ; for better observation of the worn surface, wear debris was moved out of the wear mark by using AFM with a large area scan at 300 nN after the wear test. The existence of dislocation arrays confirms that material removal occurs by plastic deformation. This corroborates the observations made in

scratch tests at ramped load in the previous section. It is concluded that the material on the microscale at high loads is removed by plastic deformation with a small contribution from elastic fracture (Zhao and Bhushan 1998).

To understand wear mechanisms, evolution of wear can be studied using AFM. Figure 8.48 shows the evolution of wear marks of a DLC-coated disk sample. The data illustrate how the microwear profile for a load of 20 μN develops as a function of the number of scanning cycles (Bhushan et al. 1994). Wear is not uniform, but is initiated at the nanoscratches. Surface defects (with high surface energy) present at the nanoscratches act as initiation sites for wear. Coating deposition also may not be uniform on and near nanoscratches which may lead to coating delamination. Thus, scratch-free surfaces will be relatively resistant to wear.

Wear precursors (precursors to measurable wear) can be studied by making surface potential measurements (DeVecchio and Bhushan 1998; Bhushan and Goldade 2000a, b). The contact potential difference, or simply the surface potential between two surfaces depends on a variety of parameters such as electronic work function, adsorption, and oxide layers. The surface potential map of an interface gives a measure of changes in the work function which is sensitive to both physical and chemical conditions of the surfaces including structural and chemical changes. Before material is actually removed in a wear process, the surface experiences stresses that result in surface and subsurface changes of structure and/or chemistry. These can cause changes in the measured potential of a surface. An AFM tip allows the mapping of the surface potential with nanoscale resolution. Surface height and change in surface potential maps of a polished single-crystal aluminum (100) sample, abraded using a diamond tip at loads of 1 and 9 μN , are shown in Fig. 8.49a. It is evident that both abraded regions show a large potential contrast (~ 0.17 V), with respect to the non-abraded area. The black region in the lower right-hand part of the topography scan shows a

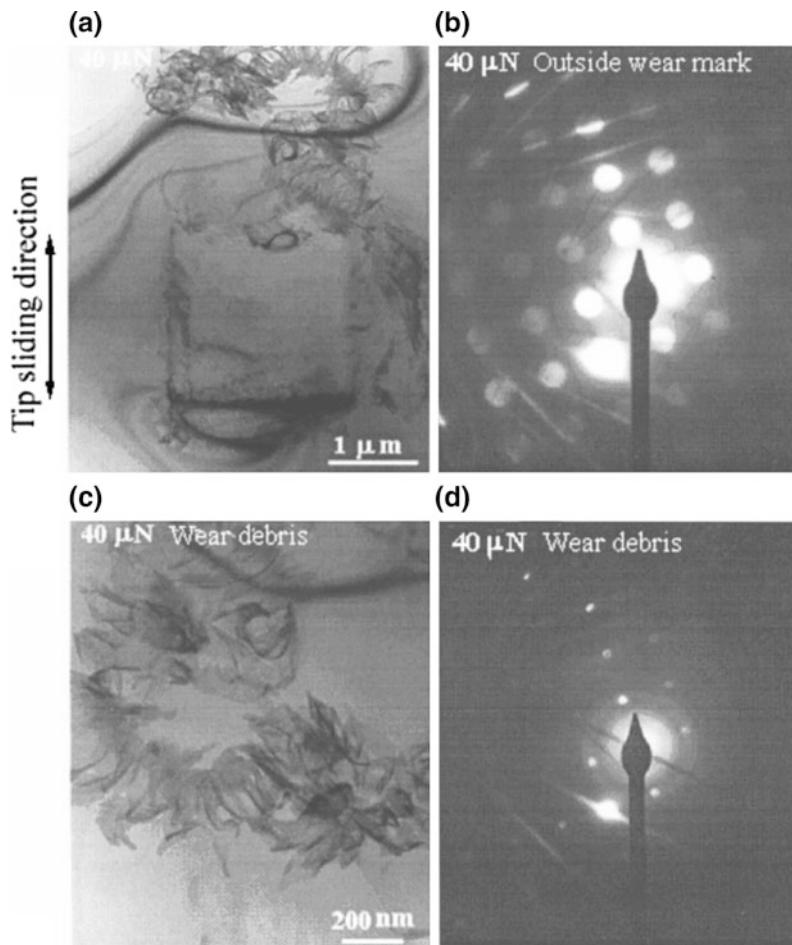


Fig. 8.46 Bright field TEM micrographs (*left*) and diffraction patterns (*right*) of wear mark (**a**, **b**) and wear debris (**c**, **d**) in Si(100) produced at a normal load of 40 μN and one scan cycle. Bend contours around and inside wear mark are observed (Zhao and Bhushan 1998)

step that was created during the polishing phase. There is no potential contrast between the high region and the low region of the sample, indicating that the technique is independent of surface height. Figure 8.49b shows a close up scan of the upper (low load) wear region in Fig. 8.49a. Notice that while there is no detectable change in the surface topography, there is nonetheless, a large change in the potential of the surface in the worn region. Indeed, the wear mark of Fig. 8.49b might not be visible at all in the topography map were it not for the noted absence of wear debris generated

nearby and then swept off during the low load scan. Thus, even in the case of zero wear (no measurable deformation of the surface using AFM), there can be a significant change in the surface potential inside the wear mark which is useful for the study of wear precursors. It is believed that the removal of the thin contaminant layer including the natural oxide layer gives rise to the initial change in surface potential. The structural changes, which precede the generation of wear debris and/or measurable wear scars, occur under ultra-low loads in the top few nanometers of

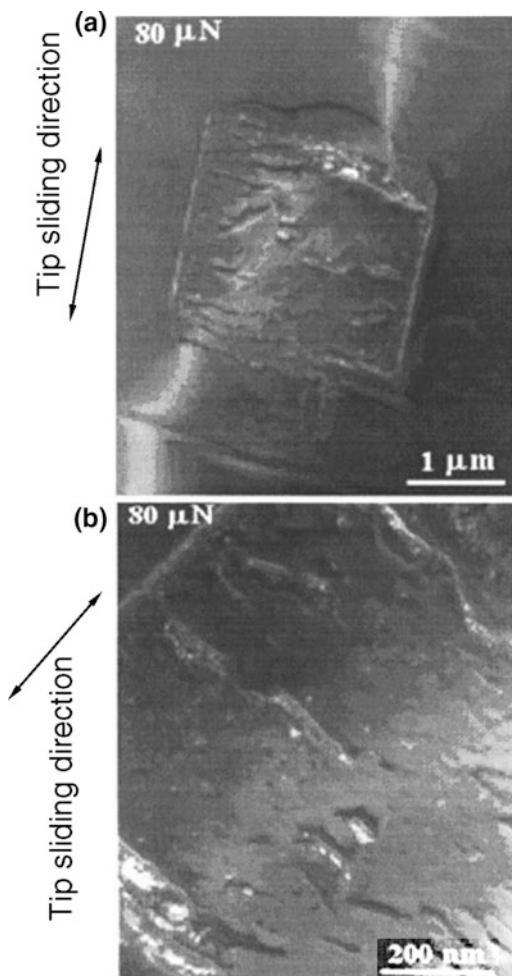


Fig. 8.47 **a** Bright field and **b** weak beam TEM micrographs of wear mark in Si(100) produced at a normal load of $80 \mu\text{N}$ and one scan cycle showing bend contours and dislocations (Zhao and Bhushan 1998)

the sample, and are primarily responsible for the subsequent changes in surface potential.

Surface potential measurements using the nano-Kelvin probe technique has been used to study wear and degradation for various applications such as for human skin (Tang et al. 2010), human hair Lodge et al. (2007); Seshadri and Bhushan 2008c), Li-ion batteries (Nagpure et al. 2011), and MEMS switches (Zaghoul et al. 2011).

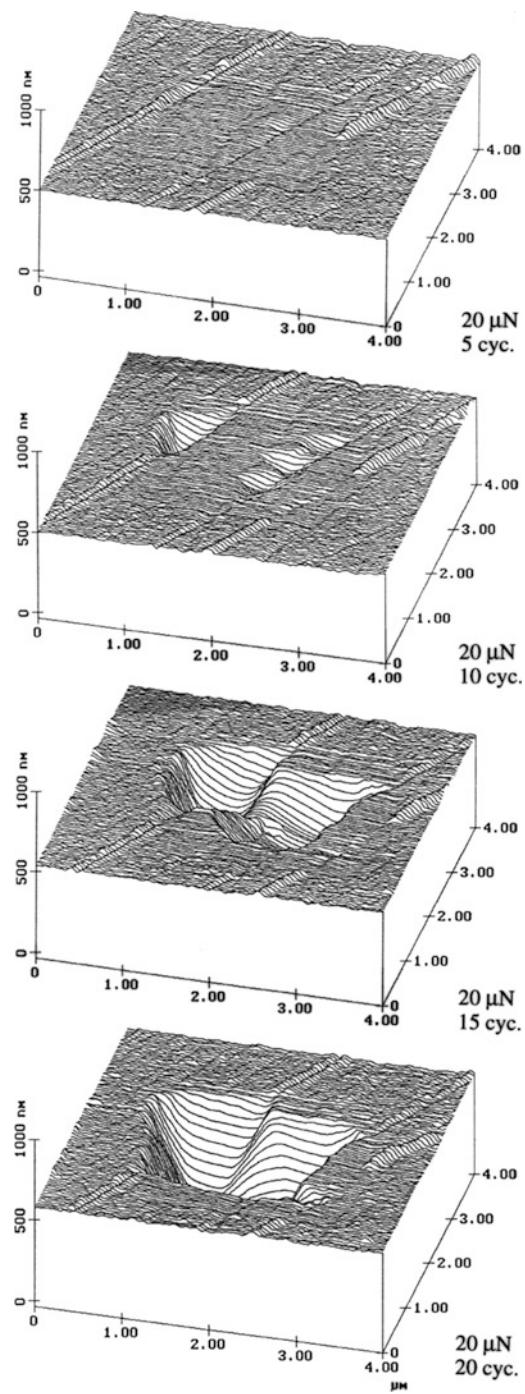


Fig. 8.48 Surface plots of diamond-like carbon-coated thin-film disk showing the worn region; the normal load and number of test cycles are indicated (Bhushan et al. 1994)

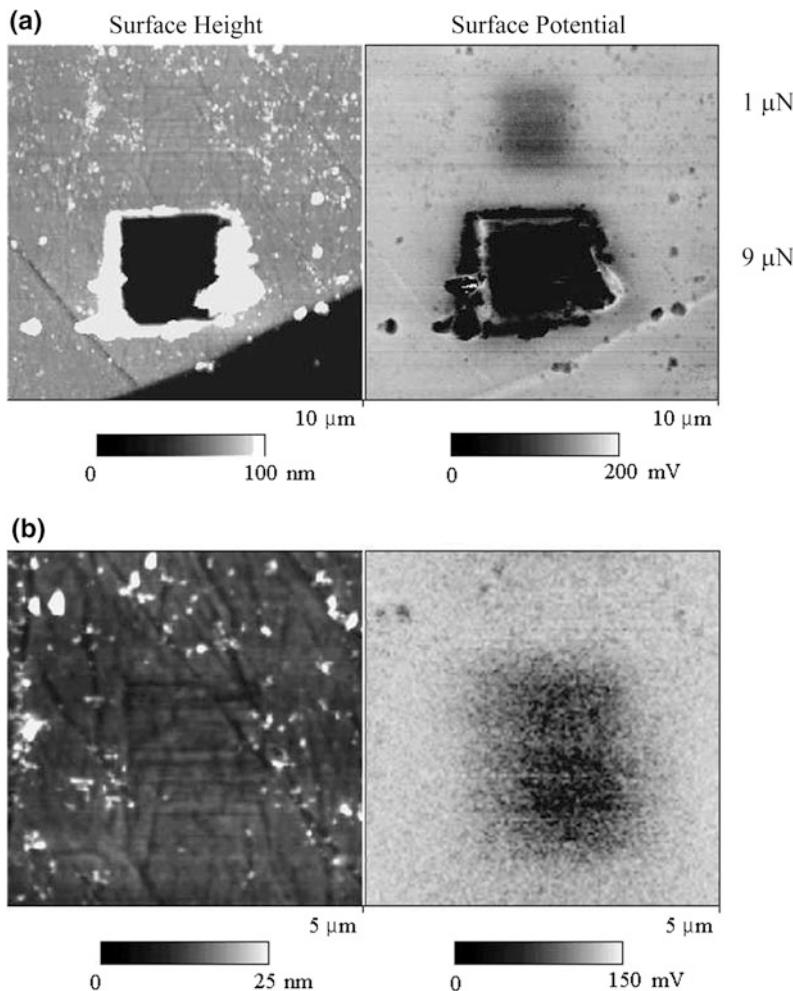


Fig. 8.49 **a** Surface height and change in surface potential maps of wear regions generated at 1 μN (*top*) and 9 μN (*bottom*) on a single crystal aluminum sample showing bright contrast in the surface potential map on the worn regions. **b** Close up of *upper* (*low load*) wear region (DeVecchio and Bhushan 1998)

8.4.4 In Situ Characterization of Local Deformation

In situ surface characterization of local deformation of materials and thin films is carried out using a tensile stage inside an AFM (Bobji and Bhushan 2001a, b; Tambe and Bhushan 2004b; Seshadri and Bhushan 2008a, b). As an example, failure mechanisms of coated polymeric thin films under tensile load were studied by Bobji and Bhushan (2001a, b). The specimens were strained at a rate of $4 \times 10^{-3}\%$ per second, and AFM images were captured at different strains up

to about 10% to monitor generation and propagation of cracks and deformation bands. Bobji and Bhushan (2001a, b) studied three magnetic tapes of thickness ranging from 7 to 8.5 μm . One of these was with acicular-shaped metal particle (MP) coating and the other two with metal-evaporated (ME) coating and with and without a thin diamondlike carbon (DLC) overcoat both on a polymeric substrate and all with particulate back coating (Bhushan 1996). They also studied the polyethylene terephthalate (PET) substrate with 6 μm thickness. They reported that cracking of the coatings started at

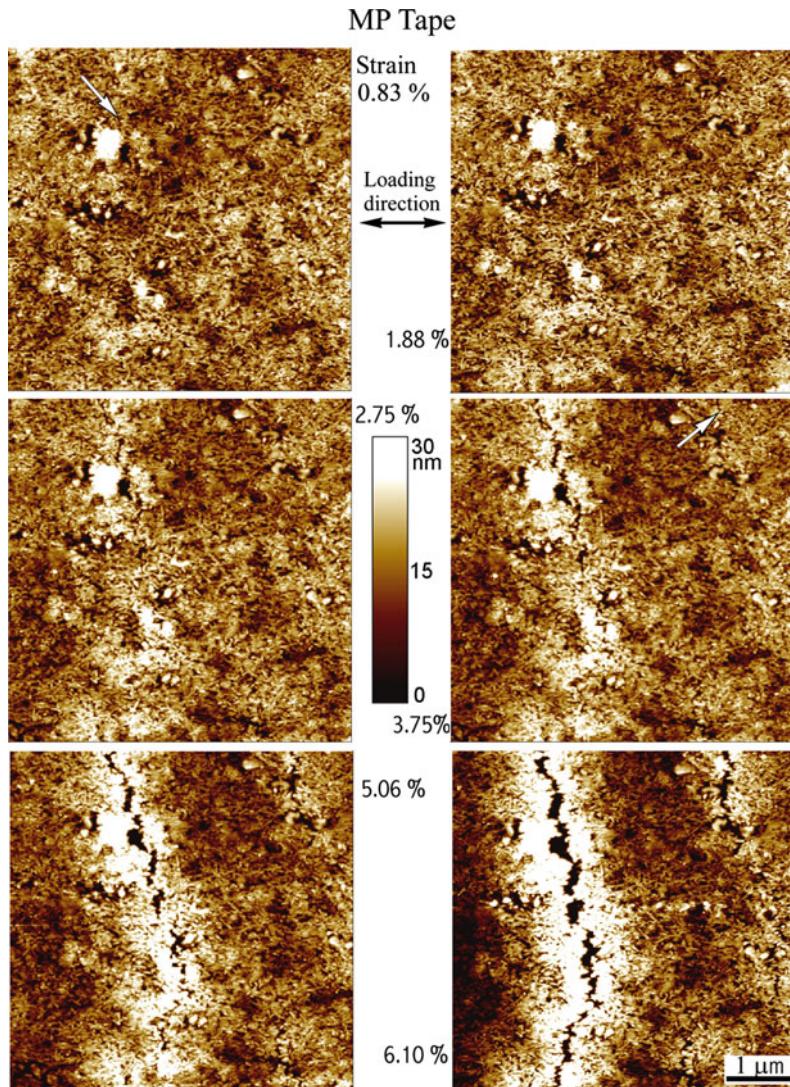


Fig. 8.50 Topographical images of the MP magnetic tape at different strains (Bobji and Bhushan 2001a)

about 1% strain for all tapes much before the substrate starts to yield at about 2% strain. Figure 8.50 shows the topographical images of the MP tape at different strains. At 0.83% strain, a crack can be seen, originating at the marked point. As the tape is further stretched along the direction, as shown in Fig. 8.50, the crack propagates along the shorter boundary of the ellipsoidal particle. However, the general direction of the crack propagation remains perpendicular to the direction of the stretching. The

length, width, and depth of the cracks increase with strain, and at the same time newer cracks keep nucleating and propagating with reduced crack spacing. At 3.75% strain, another crack can be seen nucleating. This crack continues to grow parallel to the first one. When the tape is unloaded after stretching up to a strain of about 2%, i.e. within the elastic limit of the substrate, the cracks close perfectly, and it is impossible to determine the difference from the unstrained tape.

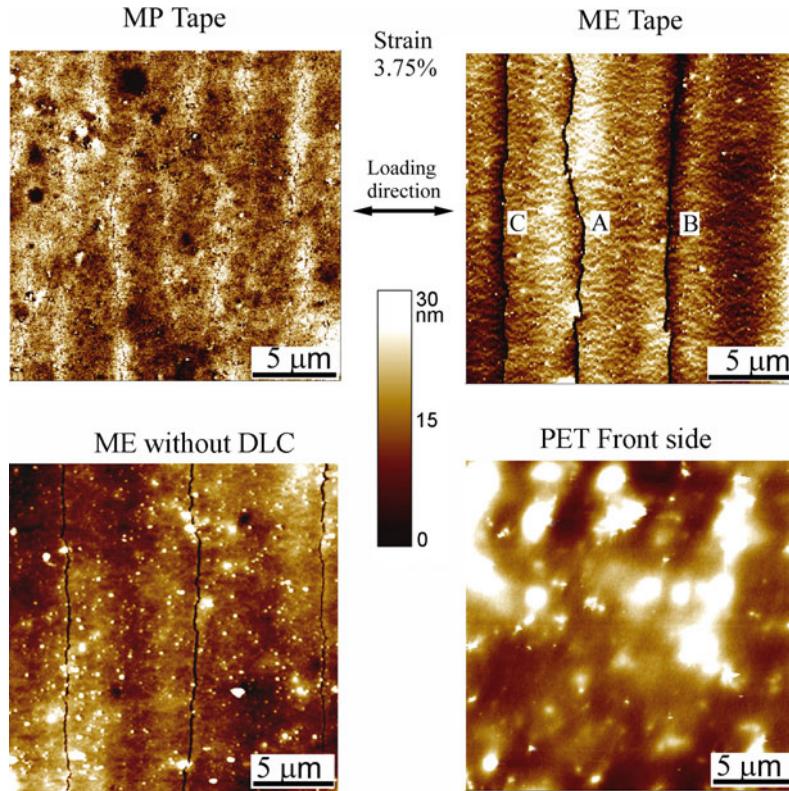


Fig. 8.51 Comparison of crack morphologies at 3.75% strain in three magnetic tapes and PET substrate. Cracks B and C, nucleated at higher strains, are more linear than crack A (Bobji and Bhushan 2001b)

Figure 8.51 shows topographical images of the three magnetic tapes and the PET substrate after being strained to 3.75%, which is well beyond the elastic limit of the substrate. MP tape develops short and numerous cracks perpendicular to the direction of loading. In tapes with metallic coating, the cracks extend throughout the tape width. In ME tape with DLC coating, there is a bulge in the coating around the primary cracks that are initiated when the substrate is still elastic, like crack A in the figure. The white band on the right-hand side of the figure is the bulge of another crack. The secondary cracks like B and C are generated at higher strains and are straighter compared to the primary cracks. In ME tape which has a Co-O film on PET substrate, with a thickness ratio of 0.03, both with and without DLC coating, no difference is observed in the rate of growth between primary and secondary cracks. The failure is cohesive with no bulging of

the coating. This seems to suggest that the DLC coating has residual stresses that relax when the coating cracks, causing delamination. Since the stresses are already relaxed, the secondary crack does not result in delamination. The presence of the residual stress is confirmed by the fact that a free-standing ME tape curls up (in a cylindrical form with its axis perpendicular to the tape length) with a radius of curvature of about 6 mm, and the ME tape without the DLC does not curl. The magnetic coating side of PET substrate is much smoother at smaller scan lengths. However, in 20 μm scans it has a lot of bulging out, which appears as white spots in the figure. These spots change shape even while scanning the samples in tapping mode at very low contact forces.

The variation of average crack width and average crack spacing with strain is plotted in Fig. 8.52. The crack width is measured at a spot

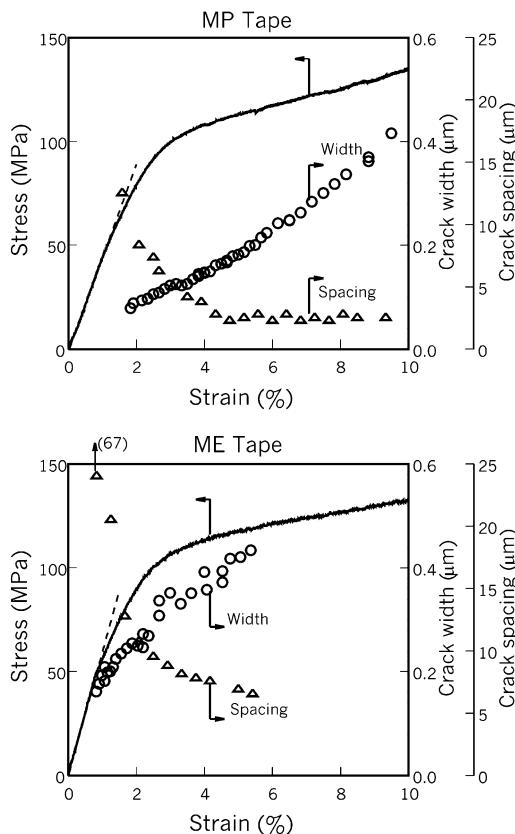


Fig. 8.52 Variation of stress, crack width, and crack spacing with strain in two magnetic tapes (Bobji and Bhushan 2001a)

along a given crack over a distance of 1 μm in the 5 μm scan image at different strains. The crack spacing is obtained by averaging the inter-crack distance measured in five separate 50 μm scans at each strain. It can be seen that the cracks nucleate at a strain of about 0.7–1.0%, well within the elastic limit of the substrate. There is a definite change in the slope of the load-displacement curve at the strain where cracks nucleate, and the slope after that is closer to the slope of the elastic portion of the substrate. This would mean that most of the load is supported by the substrate once the coating fails by cracking.

Fatigue experiments can be performed by applying a cyclic stress amplitude with a certain mean stress (Tambe and Bhushan 2004b). Fatigue life was determined by the first occurrence of cracks. Experiments were performed at various

constant mean stresses and with a range of cyclic stress amplitudes for each mean stress value for various magnetic tapes. Number of cycles to failure were plotted as a function of stress state to obtain a so-called S-N (stress-life) diagram. As the stress is decreased, there is a stress value for which no failure occurs. This stress is termed as the endurance limit or simply fatigue limit. Figure 8.53 shows the S-N curve for an ME tape and an ME tape without DLC. For the ME tape, the endurance limit is seen to go down with a decrease in mean stress. This is consistent with the literature, and it is because for lower mean stress the corresponding stress amplitude is relatively high and this causes failure. The endurance limit is found to be almost the same for all three mean stresses. In the case of ME tape without DLC as well, the critical number of cycles is found to be in the same range.

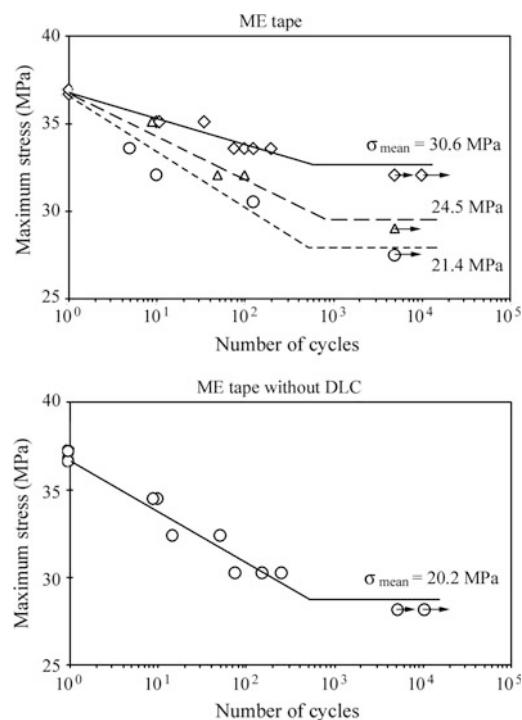


Fig. 8.53 S-N curve for two magnetic tapes with maximum stress plotted on the ordinate and number of cycles to failure on the abscissa. The data points marked with arrows indicate tests for which no failure (cracking) was observed in the scan area, even after a large number of cycles (10,000) (Tambe and Bhushan 2004b)

In situ surface characterization of unstretched and stretched films has been used to measure Poisson's ratio of polymeric thin films by Bhushan et al. (2003). Uniaxial tension is applied by the tensile stage. Surface height profiles obtained from the AFM images of unstretched and stretched samples are used to monitor the changes in displacements of the polymer films in the longitudinal and lateral directions simultaneously.

8.4.5 Nanofabrication/ Nanomachining

An AFM can be used for nanofabrication/nanomachining by extending the microscale scratching operation (Bhushan 1995, 1999a; Bhushan et al. 1994, 1995a). Figure 8.54 shows two examples of nanofabrication. The patterns were created on a single-crystal silicon (100) wafer by scratching the sample surface with a diamond tip at specified locations and scratching angles. Each line is inscribed manually at a normal load of 15 μN and a writing speed of 0.5 $\mu\text{m/s}$. The separation between lines is about 50 nm, and the variation in line width is due to the tip asymmetry. Nanofabrication parameters—normal load, scanning speed, and tip geometry—can be controlled precisely to control the depth and length of the devices.

Nanofabrication using mechanical scratching has several advantages over other techniques. Better control over the applied normal load, scan size, and scanning speed can be used for the nanofabrication of devices. Using the technique, nanofabrication can be performed on any engineering surface. Chemical etching or reactions is not required, and this dry nanofabrication process can be employed where the use of chemicals and electric field is prohibited. One disadvantage of this technique is the formation of debris during scratching. At light loads, debris formation is not a problem compared to high-load scratching. However, debris can be removed easily from the scan area at light loads during scanning.

8.5 Indentation

Mechanical properties on the relevant scales are needed for the analysis of friction and wear mechanisms. Mechanical properties, such as hardness and Young's modulus of elasticity, can be determined on the micro- to picoscales using the AFM (Bhushan and Ruan 1994; Bhushan et al. 1994; Bhushan and Koinkar 1994a, b) and a depth-sensing indentation system used in conjunction with an AFM (Bhushan et al. 1996; Kulkarni and Bhushan 1996a, b, 1997).

8.5.1 Picoindentation

Indentability on the scale of subnanometers of soft samples can be studied in the force calibration mode (Fig. 8.6) by monitoring the slope of cantilever deflection as a function of the sample traveling distance after the tip is engaged and the sample is pushed against the tip. For a rigid sample, the cantilever deflection equals the sample traveling distance, but the former quantity is smaller if the tip indents the sample. In an example for a polymeric magnetic tape shown in Fig. 8.55, the line in the left portion of the figure is curved with a slope of less than 1 shortly after the sample touches the tip, which suggests that the tip has indented the sample (Bhushan and Ruan 1994). Later, the slope is equal to 1, suggesting that the tip no longer indents the sample. This observation indicates that the tape surface is soft locally (polymer rich) but hard (as a result of magnetic particles) underneath. Since the curves in extending and retracting modes are identical, the indentation is elastic up to a maximum load of about 22 nN used in the measurements.

Detection of the transfer of material on a nanoscale is possible with the AFM. Indentation of C_{60} -rich fullerene films with an AFM tip has been shown (Ruan and Bhushan 1993) to result in the transfer of fullerene molecules to the AFM tip, as indicated by discontinuities in the cantilever deflection as a function of the sample traveling distance in subsequent indentation studies.

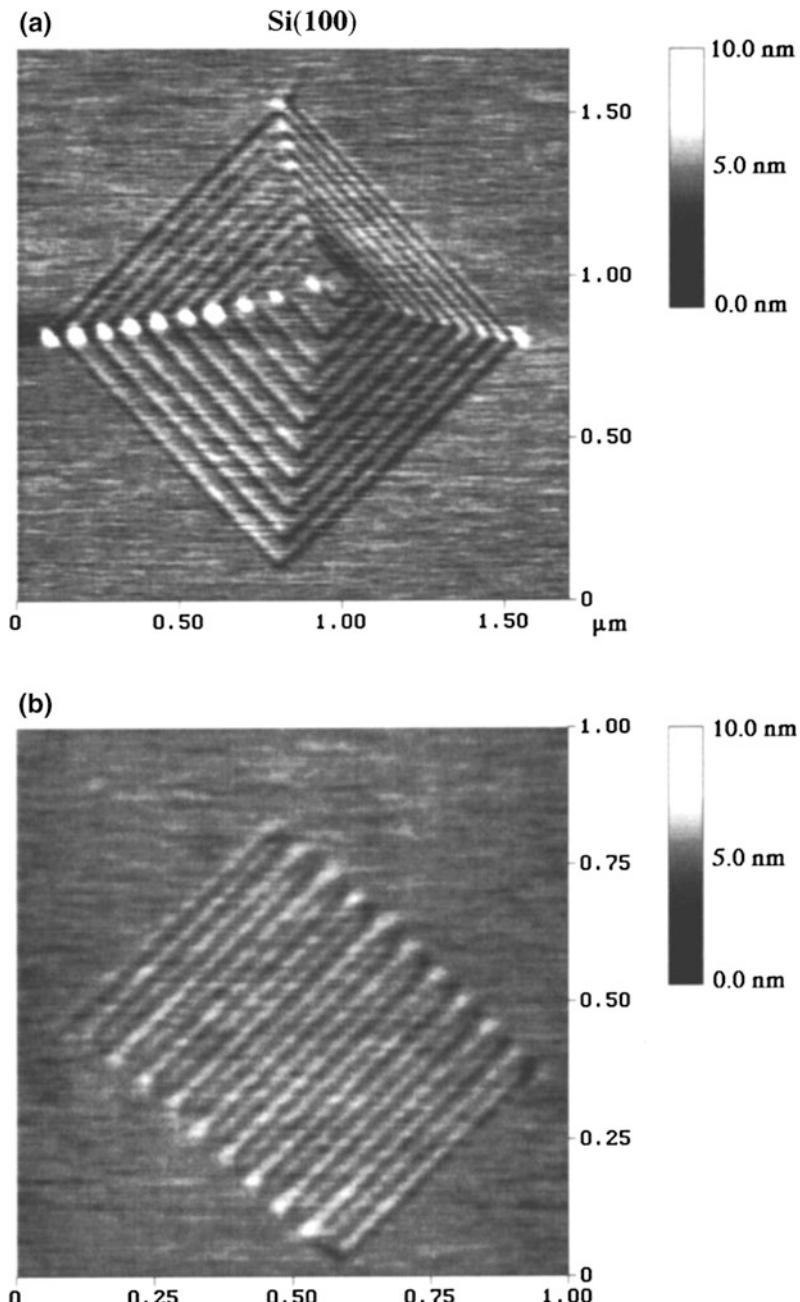


Fig. 8.54 **a** Trim and **b** spiral patterns generated by scratching a Si(100) surface using a diamond tip at a normal load of 15 μN and writing speed of 0.5 $\mu\text{m}/\text{s}$ (Bhushan 1999a, b)

8.5.2 Nanoindentation

The indentation hardness of surface films with an indentation depth of as small as about 1 nm can be measured using an AFM (Bhushan and

Koinkar 1994b; Bhushan et al. 1995a, 1996). Figure 8.56 shows the gray scale plots of indentation marks made on Si(111) at normal loads of 60, 65, 70 and 100 μN . Triangular indents can be clearly observed with very shallow depths. At a

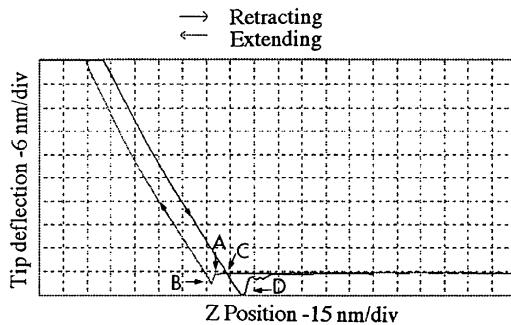


Fig. 8.55 Tip deflection (*normal load*) as a function of the Z (separation distance) curve for a polymeric magnetic tape (Bhushan and Ruan 1994)

normal load of 60 μN , indents are observed, and the depth of penetration is about 1 nm. As the normal load is increased, the indents become clearer, and indentation depth increases. For the case of hardness measurements at shallow depths on the same order as variations in surface roughness, it is desirable to subtract the original (unindented) map from the indent map for an accurate measurement of the indentation size and depth (Bhushan et al. 1994).

To make accurate measurements of hardness at shallow depths, a depth-sensing nano/picoindentation system (Fig. 8.9) is used (Bhushan et al. 1996). Figure 8.57 shows the load-displacement curves at different peak loads for Si(100). Loading/unloading curves often exhibit sharp discontinuities, particularly at high loads. Discontinuities, also referred to as pop-ins, occurring during the initial loading part of the curve mark a sharp transition from pure elastic loading to a plastic deformation of the specimen surface, thus corresponding to an initial yield point. The sharp discontinuities in unloading part of the curves are believed to be due to the formation of lateral cracks which form at the base of the median crack, which results in the surface of the specimen being thrust upward. Load-displacement data at residual depths as low as about 1 nm can be obtained. The indentation hardness of surface films has been measured for

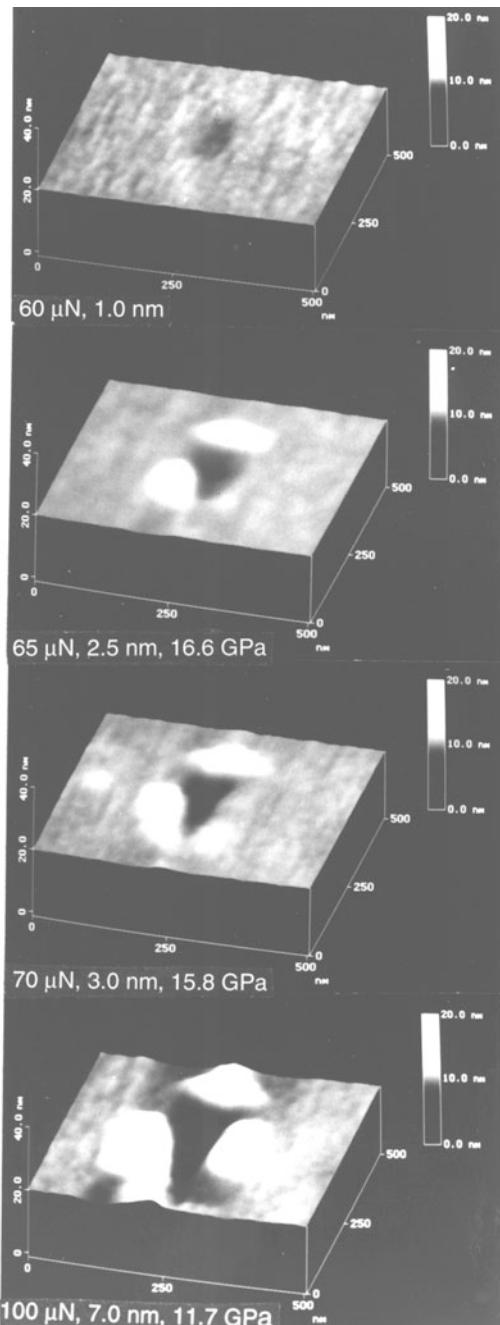


Fig. 8.56 Gray scale plots of indentation marks on the Si (111) sample at various indentation loads. Loads, indentation depths and hardness values are listed in the figure (Bhushan and Koinkar 1994b)

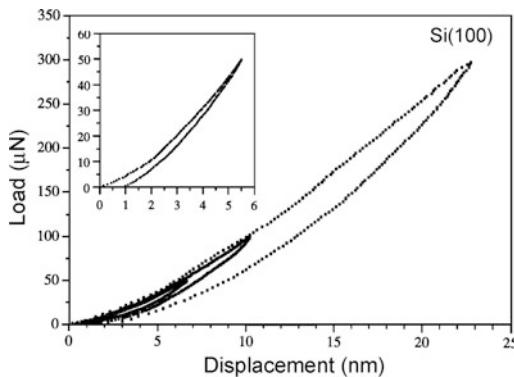


Fig. 8.57 Load-displacement curves at various peak loads for Si(100); inset shows the magnified curve for peak load of 50 μN (Bhushan et al. 1996)

various materials at a range of loads including Si(100) up to a peak load of 500 μN and Al(100) up to a peak load of 2000 μN by Bhushan et al. (1996) and Kulkarni and Bhushan (1996a, b, 1997). The hardnesses of single-crystal silicon and single-crystal aluminum at shallow depths on the order of few nm (on a nanoscale) are found to be higher than at depths on the order of few hundred nm (on a microscale), Fig. 8.58. Microhardness has also been reported to be

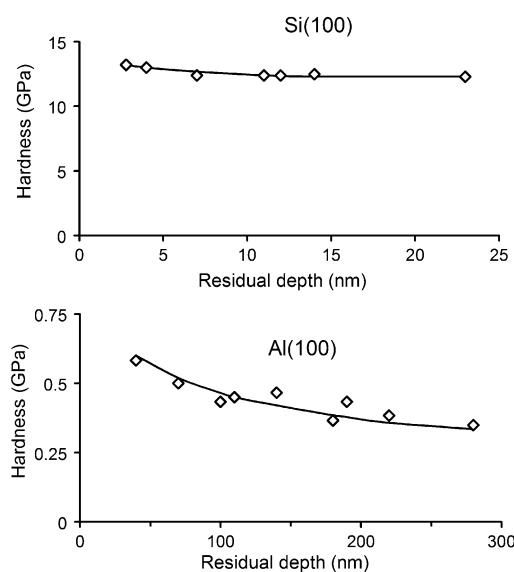


Fig. 8.58 Indentation hardness as a function of residual indentation depth for Si(100) (Bhushan et al. 1996), and Al(100) (Kulkarni and Bhushan 1996a)

higher than that on the millimeter scale by several investigators. The data reported to date show that hardness exhibits scale (size) effect.

During loading, the generation and propagation of dislocations are responsible for plastic deformation. A strain gradient plasticity theory has been developed for micro/nanoscale deformations, and is based on randomly created statistically stored and geometrically necessary dislocations (Fleck et al. 1994; Nix and Gao 1998). Large strain gradients inherent in small indentations lead to the accumulation of geometrically necessary dislocations, located in a certain sub-surface volume, for strain compatibility reasons. The large strain gradients in small indentations require these dislocations to account for the large slope at the indented surface. These dislocations become obstacles to other dislocations that cause enhanced hardening. These are a function of strain gradient, whereas statistically, stored dislocations are a function of strain. Based on this theory, scale dependent hardness is given as

$$H = H_0 \sqrt{1 + \ell_d/a} \quad (8.23)$$

where H_0 is the hardness in the absence of strain gradient or macrohardness, ℓ_d is the material-specific characteristic length parameter, and a is the contact radius. In addition to the role of strain gradient plasticity theory, an increase in hardness with a decrease in indentation depth can possibly be rationalized on the basis that as the volume of deformed material decreases, there is a lower probability of encountering material defects.

Bhushan and Koinkar (1994a) used AFM measurements to show that ion implantation of silicon surfaces increases their hardness and thus their wear resistance. Formation of surface alloy films with improved mechanical properties by ion implantation is of growing technological importance as a means of improving the mechanical properties of materials. Hardness of 20 nm thick DLC films have been measured by Kulkarni and Bhushan (1997).

To measure variations in mechanical properties of layered materials and nonhomogeneous composites, a so-called continuous stiffness technique is used (Li and Bhushan 2002). The

technique offers a direct measurement of dynamic contact stiffness during the loading portion of an indentation test. This allows the measurement of contact stiffness, elastic modulus, hardness, and fatigue properties. For example, the technique has been used in layered and nonhomogenous magnetic tapes.

The creep and strain-rate effects (viscoelastic effects) of ceramics can be studied using a depth-sensing indentation system. Bhushan et al. (1996) and Kulkarni and Bhushan (1996a, b,

1997) have reported that ceramics (single-crystal silicon and diamondlike carbon) exhibit significant plasticity and creep on a nanoscale. Figure 8.59a shows the load-displacement curves for single-crystal silicon at various peak loads held at 180 s. To demonstrate the creep effects, the load-displacement curves for a 500 μ N peak load held at 0 and 30 s are also shown as an inset. Note that significant creep occurs at room temperature. Nanoindenter experiments conducted by Li et al. (1991) exhibited significant

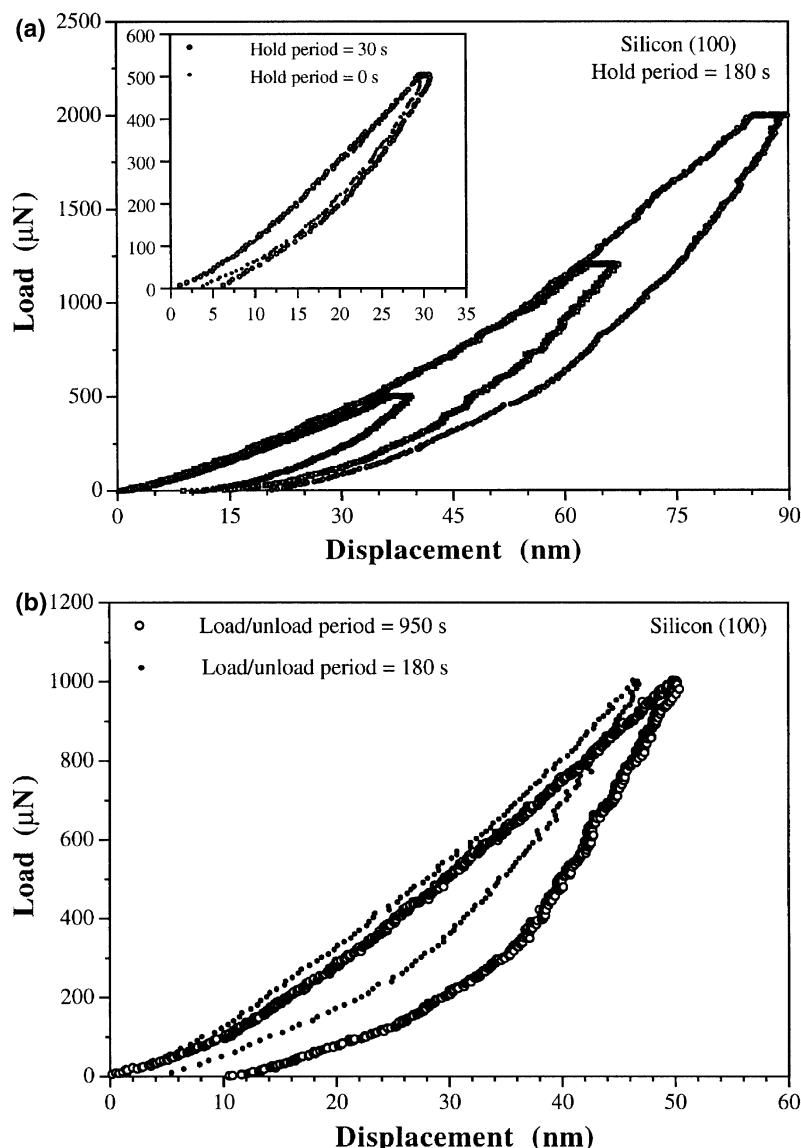


Fig. 8.59 a Creep behavior and b strain-rate sensitivity of Si(100) (Bhushan et al. 1996)

creep only at high temperatures (greater than or equal to 0.25 times the melting point of silicon). The mechanism of dislocation glide plasticity is believed to dominate the indentation creep process on the macroscale. To study the strain-rate sensitivity of silicon, data at two different (constant) rates of loading are presented in Fig. 8.59 b. Note that a change in the loading rate by a factor of about five results in a significant change in the load-displacement data. The viscoelastic effects observed here for silicon at ambient temperature could arise from the size effects mentioned earlier. Most likely, creep and strain rate experiments are being conducted on the hydrated films present on the silicon surface in ambient environment, and these films are expected to be viscoelastic.

8.5.3 Localized Surface Elasticity and Viscoelasticity Mapping

The Young's modulus of elasticity can be calculated from the slope of the indentation curve during unloading. However, these measurements provide a single-point measurement. By using the force modulation technique, it is possible to obtain localized elasticity maps of soft and

compliant materials of near surface regions with nanoscale lateral resolution. This technique has been successfully used for polymeric magnetic tapes, which consist of magnetic and nonmagnetic ceramic particles in a polymeric matrix. Elasticity maps of a tape can be used to identify the relative distribution of the hard magnetic and nonmagnetic ceramic particles on the tape surface, which has an effect on friction and stiction at the head-tape interface (Bhushan 1996). Figure 8.60 shows the surface height and elasticity maps on a polymeric magnetic tape (DeVecchio and Bhushan 1997). The elasticity image reveals sharp variations in the surface elasticity due to the composite nature of the film. As can be clearly seen, regions of high elasticity do not always correspond to high or low topography. Based on a Hertzian elastic-contact analysis, the static indentation depth of these samples during the force modulation scan is estimated to be about 1 nm. We conclude that the contrast seen is influenced most strongly by material properties in the top few nanometers, independent of the composite structure beneath the surface layer.

By using phase contrast microscopy, it is possible to obtain phase contrast maps or the contrast in viscoelastic properties of near surface regions with nanoscale lateral resolution. This technique has been successfully used for

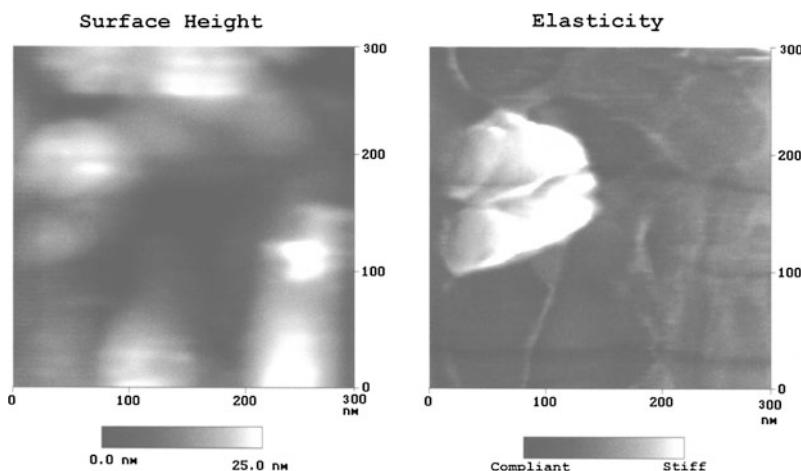


Fig. 8.60 Surface height and elasticity maps on a polymeric magnetic tape ($\sigma = 6.7$ nm and P-V = 32 nm; σ and P-V refer to standard deviation of surface heights and peak-to-valley distance, respectively). The gray scale on the elasticity map is arbitrary (DeVecchio and Bhushan 1997)

polymeric films and magnetic tapes which consist of ceramic particles in a polymeric matrix (Scott and Bhushan 2003; Bhushan and Qi 2003; Kasai et al. 2004; Chen and Bhushan 2005).

Figure 8.61 shows typical surface height, TR amplitude, and TR phase angle images for a MP tape using TR mode II, described earlier. TR amplitude image provides contrast in lateral stiffness, and TR phase angle image provides contrast in viscoelastic properties. In TR amplitude and phase angle images, the distribution of magnetic particles can be clearly seen which have better contrast than that in TR surface height image. MP tape samples show granular structure with elliptical shaped magnetic particle aggregates (50–100 nm in diameter). Studies by Scott and Bhushan (2003), Bhushan and Qi (2003), and Kasai et al. (2004) have indicated that the phase shift can be related to the energy dissipation through the viscoelastic deformation process between the tip and the sample. Recent theoretical analysis has established a quantitative correlation between the lateral surface properties (stiffness and viscoelasticity) of materials and amplitude/phase angle shift in TR measurements (Song and Bhushan 2005). The contrast in the TR amplitude and phase angle images is due to the in-plane (lateral) heterogeneity of the surface. Based on the TR amplitude and phase angle images, the lateral surface properties (lateral stiffness and viscoelasticity) mapping of materials can be obtained.

8.6 Boundary Lubrication

8.6.1 Perfluoropolyether Lubricants

The classical approach to lubrication uses freely supported multimolecular layers of liquid lubricants (Bowden and Tabor 1950; Bhushan 1996, 1999a, 2001a). The liquid lubricants are sometimes chemically bonded to improve their wear resistance (Bhushan 1996, 1999a, 2011, 2013a, b). Partially chemically-bonded, molecularly-thick perfluoropolyether (PFPE) films are used for the lubrication of magnetic storage media because of their thermal stability and extremely low vapor pressure (Bhushan 1996). Chemically-bonded lubricants are considered potential candidate lubricants for MEMS/NEMS. Molecularly-thick PFPEs are well suited because of the following properties: low surface tension and a low contact angle which allows easy spreading on surfaces and provide hydrophobic properties; chemical and thermal stability which minimizes degradation under use; low vapor pressure which provides low out-gassing; high adhesion to substrate via organic functional bonds; and good lubricity which reduces contact surface wear.

For boundary lubrication studies, friction, adhesion, and durability experiments have been performed on virgin Si (100) surfaces and silicon surfaces lubricated with various PFPE lubricants (Koinkar and Bhushan 1996a, b; Liu and Bhushan 2003a; Tao and Bhushan 2005a;

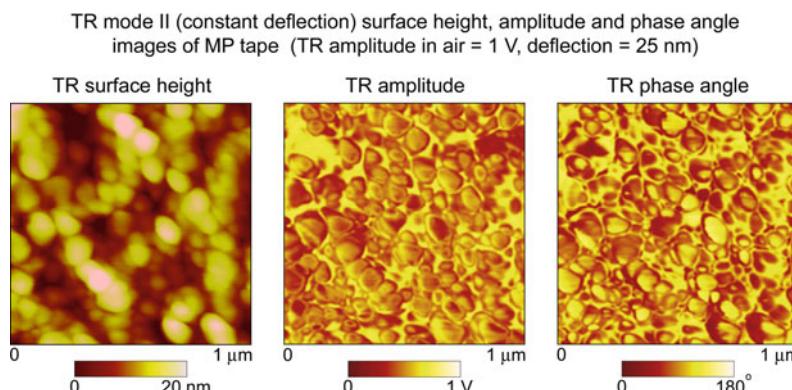


Fig. 8.61 Images of an MP tape obtained with TR mode II (constant deflection). TR mode II amplitude and phase angle images have the largest contrast among tapping, TR mode I and TR mode II techniques (Chen and Bhushan 2005)

Bhushan et al. 2007; Palacio and Bhushan 2007a, b). More recently, there has been interest in selected ionic liquids for lubrication (Bhushan et al. 2008; Palacio and Bhushan 2008, 2009). They possess efficient heat transfer properties. They are also electrically conducting, which is of interest in various MEMS/NEMS applications. Results of the following two PFPE lubricants will be presented here: Z-15 (with $-CF_3$ nonpolar end groups), $CF_3-O-(CF_2-CF_2-O)_m-(CF_2-O)_n-CF_3$ ($m/n \sim 2/3$) and Z-DOL (with $-OH$ polar end groups), $HO-CH_2-CF_2-O-(CF_2-CF_2-O)_m-(CF_2-O)_n-CF_2-CH_2-OH$ ($m/n \sim 2/3$). Z-DOL film was thermally bonded at 150 °C for 30 min, and the unbonded fraction was removed by a solvent (fully bonded) (Bhushan 1996). The thicknesses of Z-15 and Z-DOL films were 2.8 and 2.3 nm, respectively. The lubricant chain diameters of these molecules is about 0.6 nm, and molecularly thick films generally lie flat on surfaces with high coverage.

The adhesive forces of Si(100), Z-15, and Z-DOL (fully bonded) measured by force calibration plot and friction force versus normal load plot are summarized in Fig. 8.62 (Liu and Bhushan 2003a). The data obtained by these two methods are in good agreement. Figure 8.62 shows that the presence of mobile Z-15 lubricant film increases the adhesive force as compared to that of Si(100) by meniscus formation. Whereas, the presence of the solid like phase of the Z-DOL (fully bonded) film reduces the adhesive force as compared that of Si(100) because of the absence of mobile liquid. The schematic (bottom) in Fig. 8.62 shows the relative size and sources of the meniscus. It is well known that the native oxide layer (SiO_2) on the top of Si(100) wafer exhibits hydrophilic properties, and some water molecules can be adsorbed on this surface. The condensed water will form a meniscus as the tip approaches the sample surface. The larger adhesive force in Z-15 is not only caused by the Z-15 meniscus alone, the non-polarized Z-15 liquid does not have good wettability and strong bonding with Si(100). Consequently, in the ambient environment, the condensed water molecules from the environment will permeate through the liquid Z-15 lubricant film and compete with the

lubricant molecules present on the substrate. The interaction of the liquid lubricant with the substrate is weakened, and a boundary layer of the liquid lubricant forms puddles (Koinkar and Bhushan 1996a, b). This dewetting allows water molecules to be adsorbed on the Si(100) surface along with Z-15 molecules, and both of them can form meniscus while the tip approaches the surface. Thus the dewetting of liquid Z-15 film results in a higher adhesive force and poorer lubrication performance. In addition, the Z-15 film is soft compared to the solid Si(100) surface, and penetration of the tip in the film occurs while pushing the tip down. This results in the large area of the tip being wetted by the liquid to form the meniscus at the tip-liquid (mixture of Z-15 and water) interface. It should also be noted that Z-15 has a higher viscosity compared to water, therefore the Z-15 film provides a higher resistance to motion and the coefficient of friction. In the case of Z-DOL (fully bonded) film, both of the active groups of Z-DOL molecules are mostly bonded on Si(100) substrate, thus the Z-DOL (fully bonded) film has low free surface energy and cannot be displaced readily by the water molecules or can readily adsorb the water molecules.

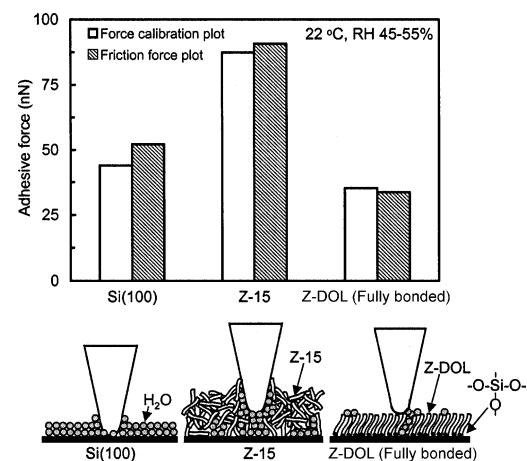


Fig. 8.62 Summary of the adhesive forces of Si(100) and Z-15 and Z-DOL (fully bonded) films measured by force calibration plots and friction force versus normal load plots in ambient air. The schematic (bottom) showing the effect of meniscus, formed between AFM tip and the surface sample, on the adhesive and friction forces (Liu and Bhushan 2003a)

Thus, the use of Z-DOL (fully bonded) can reduce the adhesive force.

To study the velocity effect on friction and adhesion, the variation of friction force, the adhesive force, and the coefficient of friction of Si(100), Z-15 and Z-DOL(fully bonded) as a function of velocity are summarized in Fig. 8.63 (Liu and Bhushan 2003a). It indicates that for a silicon wafer, the friction force decreases logarithmically with increasing velocity. For Z-15, the friction force decreases with increasing velocity up to 10 $\mu\text{m/s}$, after which it remains almost constant. The velocity has a very small effect on the friction force of Z-DOL (fully bonded); it reduced slightly only at very high velocity. Figure 8.63 also indicates that the adhesive force of Si(100) is increased when the

velocity is higher than 10 $\mu\text{m/s}$. The adhesive force of Z-15 is reduced dramatically with a velocity increase up to 20 $\mu\text{m/s}$, after which it is reduced slightly, and the adhesive force of Z-DOL (fully bonded) is also decreased at high velocity. In the testing range of velocity, only the coefficient of friction of Si(100) decreases with velocity, but the coefficients of friction of Z-15 and Z-DOL (fully bonded) almost remain constant. This implies that the friction mechanisms of Z-15 and Z-DOL (fully bonded) do not change with the variation of velocity.

The mechanisms of the effect of velocity on adhesion and friction are explained based on schematics shown in Fig. 8.63 (right) (Liu and Bhushan 2003a). For Si(100), a tribochemical reaction plays a major role. Although, at high

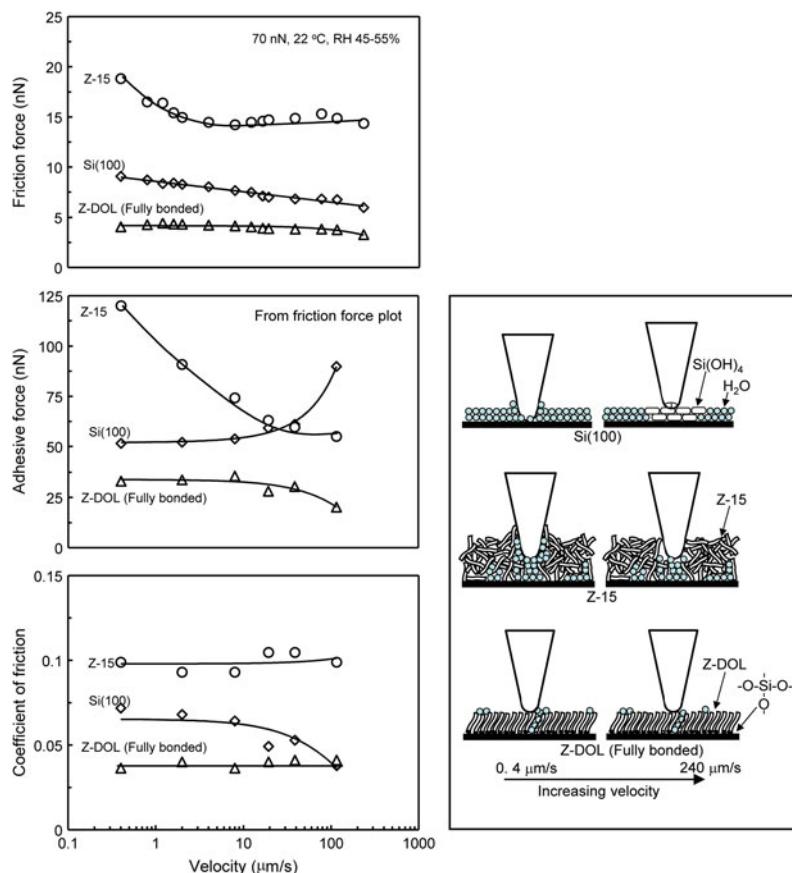


Fig. 8.63 The influence of velocity on the friction force, adhesive force and coefficient of friction of Si(100) and Z-15 and Z-DOL (fully bonded) films at 70 nN, in ambient air. The schematic (right) shows the change of surface composition (by tribochemical reaction) and formation of meniscus while increasing the velocity (Liu and Bhushan 2003a)

velocity, the meniscus is broken and does not have enough time to rebuild, the contact stresses and high velocity lead to tribochemical reactions of Si(100) wafer (which has native oxide (SiO_2)) and Si_3N_4 tip with water molecules and they form $\text{Si}(\text{OH})_4$. The $\text{Si}(\text{OH})_4$ is removed and continuously replenished during sliding. The $\text{Si}(\text{OH})_4$ layer between the tip and the Si(100) surface is known to be of low shear strength and causes a decrease in friction force and coefficient of friction (Bhushan 2013a, b). The chemical bonds of Si–OH between the tip and Si(100) surface induce large adhesive force. For Z-15 film, at high velocity the meniscus formed by the condensed water and the Z-15 molecules is broken and does not have enough time to rebuild, therefore, the adhesive force and consequently friction force is reduced. The friction mechanisms for Z-15 film still shear the same viscous liquid even at high velocity range, thus the coefficient of friction of

the Z-15 does not change with velocity. For Z-DOL (fully bonded) film, the surface can adsorb a few water molecules in an ambient condition, and at high velocity these molecules are displaced, which is responsible for a slight decrease in friction force and adhesive force. Koinkar and Bhushan (1996a, b) have suggested that in the case of samples with mobile films, such as condensed water and Z-15 films, alignment of the liquid molecules (shear thinning) is responsible for the drop in friction force with an increase in scanning velocity. This could be another reason for the decrease in friction force for Si(100) and Z-15 film with velocity in this study.

To study the relative humidity effect on friction and adhesion, the variation of friction force, adhesive force, and the coefficient of friction of Si(100), Z-15, and Z-DOL (fully bonded) as a function of relative humidity are shown in Fig. 8.64 (Liu and Bhushan 2003a). It shows that

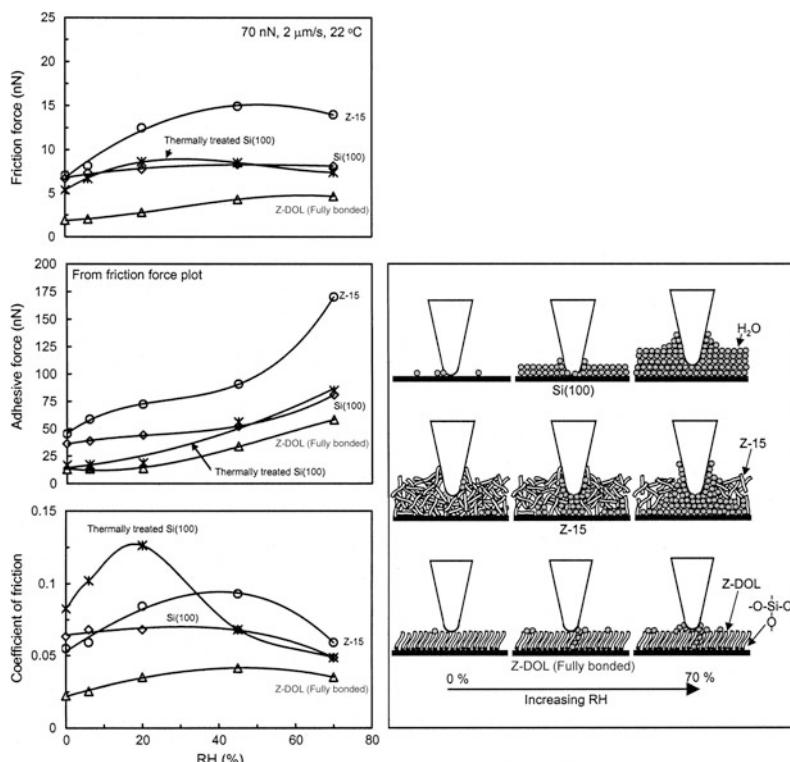


Fig. 8.64 The influence of relative humidity on the friction force, adhesive force, and coefficient of friction of Si(100) and Z-15 and Z-DOL (fully bonded) films at 70 nN, 2 $\mu\text{m}/\text{s}$, and in 22 $^{\circ}\text{C}$ air. Schematic (right) shows the change of meniscus while increasing the relative humidity. In this figure, the thermally treated Si(100) represents the Si(100) wafer that was baked at 150 $^{\circ}\text{C}$ for 1 h in an oven (in order to remove the adsorbed water) just before it was placed in the 0% RH chamber (Liu and Bhushan 2003a)

for Si(100) and Z-15 film, the friction force increases with a relative humidity increase up to 45%, and then it shows a slight decrease with a further increase in the relative humidity. Z-DOL (fully bonded) has a smaller friction force than Si(100) and Z-15 in the whole testing range, and its friction force shows a relative apparent increase when the relative humidity is higher than 45%. For Si(100), Z-15 and Z-DOL (fully bonded), their adhesive forces increase with relative humidity, and their coefficients of friction increase with relative humidity up to 45%, after which they decrease with further increase in the relative humidity. It is also observed that the humidity effect on Si(100) really depends on the history of the Si(100) sample. As the surface of Si(100) wafer readily adsorbs water in air, without any pre-treatment the Si(100) used in our study almost reaches its saturate stage of adsorbed water, and is responsible for less effect during increasing relative humidity. However, once the Si(100) wafer was thermally treated by baking at 150°C for 1 h, a greater effect was observed.

The schematic (right) in Fig. 8.64 shows that for Si(100), because of its high free surface energy, it can adsorb more water molecules during an increase in relative humidity (Liu and Bhushan 2003a). As discussed earlier, for Z-15 film in the humid environment, the condensed water from the humid environment competes with the lubricant film present on the sample surface, and the interaction of the liquid lubricant film with the silicon substrate is weakened and a boundary layer of the liquid lubricant forms puddles. This dewetting allows water molecules to be adsorbed on the Si(100) substrate mixed with Z-15 molecules (Koinkar and Bhushan 1996a, b). Obviously, more water molecules can be adsorbed on Z-15 surface while increasing relative humidity. The more adsorbed water molecules in the case of Si(100), along with lubricant molecules in Z-15 film case, form a bigger water meniscus which leads to an increase of friction force, adhesive force, and coefficient

of friction of Si(100) and Z-15 with humidity, but at very high humidity of 70%, large quantities of adsorbed water can form a continuous water layer that separates the tip and sample surface and acts as a kind of lubricant, which causes a decrease in the friction force and the coefficient of friction. For Z-DOL (fully bonded) film, because of their hydrophobic surface properties, water molecules can be adsorbed at humidity higher than 45%, and cause an increase in the adhesive force and friction force.

To study the temperature effect on friction and adhesion, the variation of friction force, adhesive force, and coefficient of friction of Si(100), Z-15 and Z-DOL (fully bonded) as a function of temperature are summarized in Fig. 8.65 (Liu and Bhushan 2003a). It shows that the increasing temperature causes a decrease of friction force, adhesive force, and coefficient of friction of Si(100), Z-15 and Z-DOL (fully bonded). The schematic (right) in Fig. 8.65 indicates that at high temperature, desorption of water leads to the decrease of the friction force, the adhesive forces and the coefficient of friction for all of the samples. For Z-15 film, the reduction of viscosity at high temperature also has a contribution to the decrease of friction force and coefficient of friction. In the case of Z-DOL (fully bonded) film, molecules are easier to be oriented at high temperature, which may be partly responsible for the low friction force and coefficient of friction.

As a brief summary, the influence of velocity, relative humidity, and temperature on the friction force of mobile Z-15 film is presented in Fig. 8.66 (Liu and Bhushan 2003a). The changing trends are also addressed in this figure.

To study the durability of lubricant films at the nanoscale, the friction of Si(100), Z-15, and Z-DOL (fully bonded) as a function of the number of scanning cycles are shown in Fig. 8.67 (Liu and Bhushan 2003a). As observed earlier, the friction force of Z-15 is higher than that of Si(100) with the lowest values for Z-DOL (fully bonded). During cycling, the friction force and the coefficient of friction of Si(100) show a

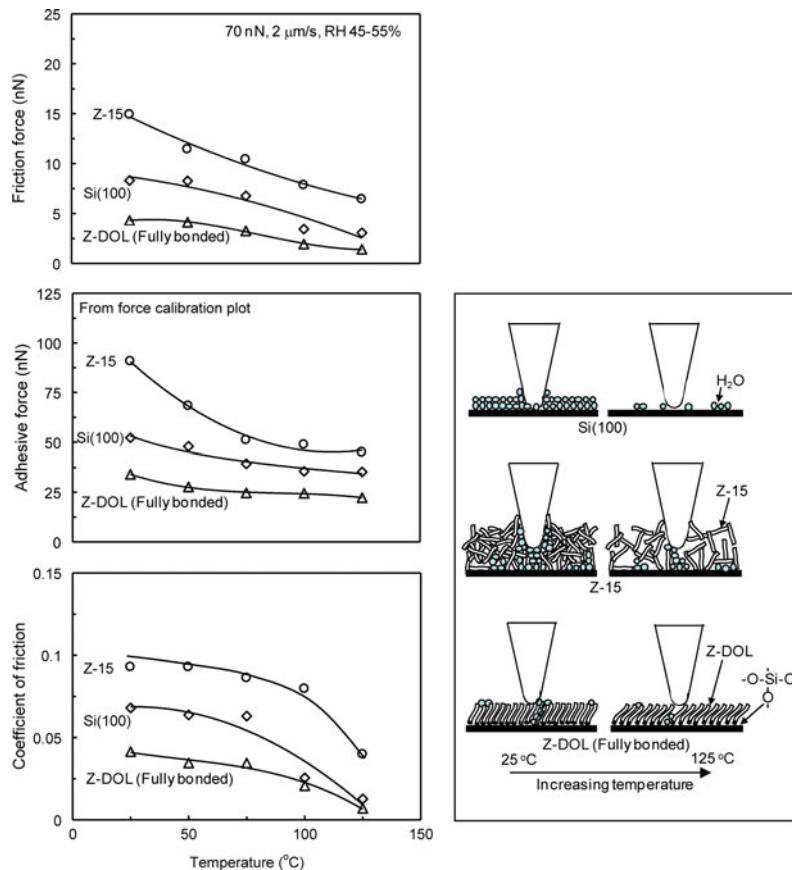


Fig. 8.65 The influence of temperature on the friction force, adhesive force, and coefficient of friction of Si(100) and Z-15 and Z-DOL (fully bonded) films at 70 nN, at 2 $\mu\text{m}/\text{s}$, and in RH 40–50% air. The schematic (right) shows that at high temperature, desorption of water decreases the adhesive forces. And the reduced viscosity of Z-15 leads to the decrease of coefficient of friction. High temperature facilitates orientation of molecules in Z-DOL (fully bonded) film which results in lower coefficient of friction (Liu and Bhushan 2003a)

slight decrease during the initial few cycles then remain constant. This is related to the removal of the native oxide. In the case of Z-15 film, the friction force and coefficient of friction show an increase during the initial few cycles and then approach higher stable values. This is believed to be caused by the attachment of the Z-15 molecules to the tip. After several scans, the molecular interaction reaches equilibrium, and after that the friction force and coefficient of friction remain constant. In the case of Z-DOL (fully bonded) film, the friction force and coefficient of friction start out low and remain low during the entire test for 100 cycles. It suggests that Z-DOL (fully bonded) molecules do not get attached or displaced as readily as Z-15.

8.6.2 Self-Assembled Monolayers

For the lubrication of MEMS/NEMS, another effective approach involves the deposition of organized and dense molecular layers of long-chain molecules. Two common methods to produce monolayers and thin films are the Langmuir-Blodgett (L-B) deposition and self-assembled monolayers (SAMs) by chemical grafting of molecules. LB films are physically bonded to the substrate by weak van der Waals attraction, while SAMs are chemically bonded via covalent bonds to the substrate. Because of the choice of chain length and terminal linking group that SAMs offer, they hold great promise for boundary lubrication of MEMS/NEMS. A number of studies have been

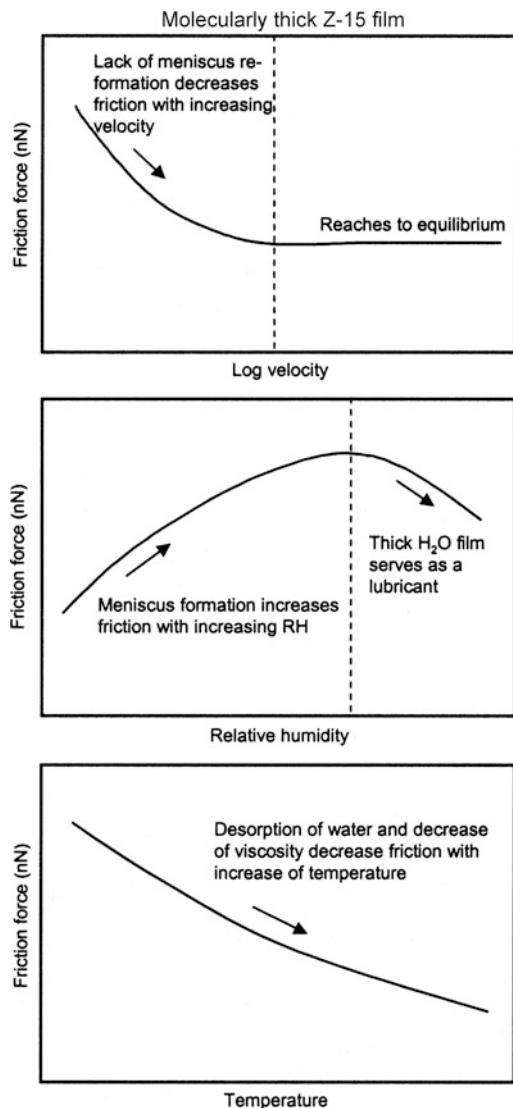


Fig. 8.66 Schematic shows the change of friction force of molecularly thick Z-15 films with log velocity, relative humidity, and temperature. The changing trends are also addressed in this figure (Liu and Bhushan 2003a)

conducted to study the tribological properties of various SAMs deposited on Si, Al, and Cu substrates (Bhushan et al. 1995b, 2005, 2006, 2007; Bhushan and Liu 2001; Liu et al. 2001; Liu and Bhushan 2002; Kasai et al. 2005; Lee et al. 2005; Tambe and Bhushan 2005h; Tao and Bhushan 2005b; Hoque et al. 2006a, b, 2007a, b, 2008, 2009; DeRose et al. 2008).

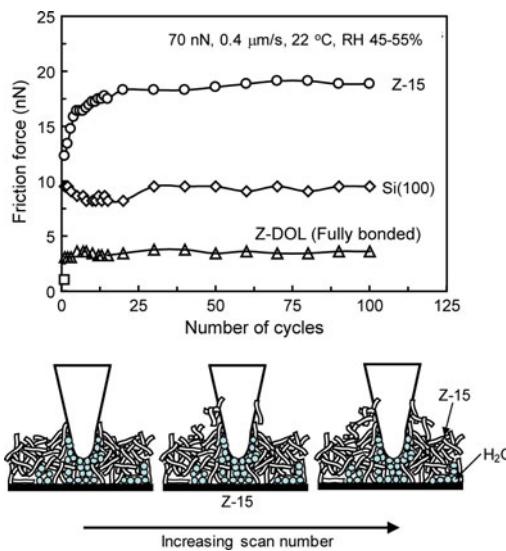


Fig. 8.67 Friction force versus number of sliding cycles for Si(100) and Z-15 and Z-DOL (fully bonded) films at 70 nN, 0.8 μ m/s, and in ambient air. Schematic (bottom) shows that some liquid Z-15 molecules can be attached onto the tip. The molecular interaction between the attached molecules onto the tip with the Z-15 molecules in the film results in an increase of the friction force with multi scanning (Liu and Bhushan 2003a)

Bhushan and Liu (2001) studied the effect of film compliance on adhesion and friction. They used hexadecane thiol (HDT), 1,1,biphenyl-4-thiol (BPT), and crosslinked BPT (BPTC) solvent deposited on Au(111) substrate, Fig. 8.68a. The average values and standard deviation of the adhesive force and coefficient of friction are presented in Fig. 8.68b. Based on the data, the adhesive force and coefficient of friction of SAMs are less than the corresponding substrates. Among various films, HDT exhibits the lowest values. Based on the stiffness measurements of various SAMs, HDT was most compliant, followed by BPT and BPTC. Based on friction and stiffness measurements, SAMs with high-compliance long carbon chains exhibit low friction; chain compliance is desirable for low friction. The friction mechanism of SAMs is explained by a so-called “molecular spring” model (Fig. 8.69). According to this model, the chemically adsorbed self-assembled molecules on a substrate are like assembled molecular springs anchored to the

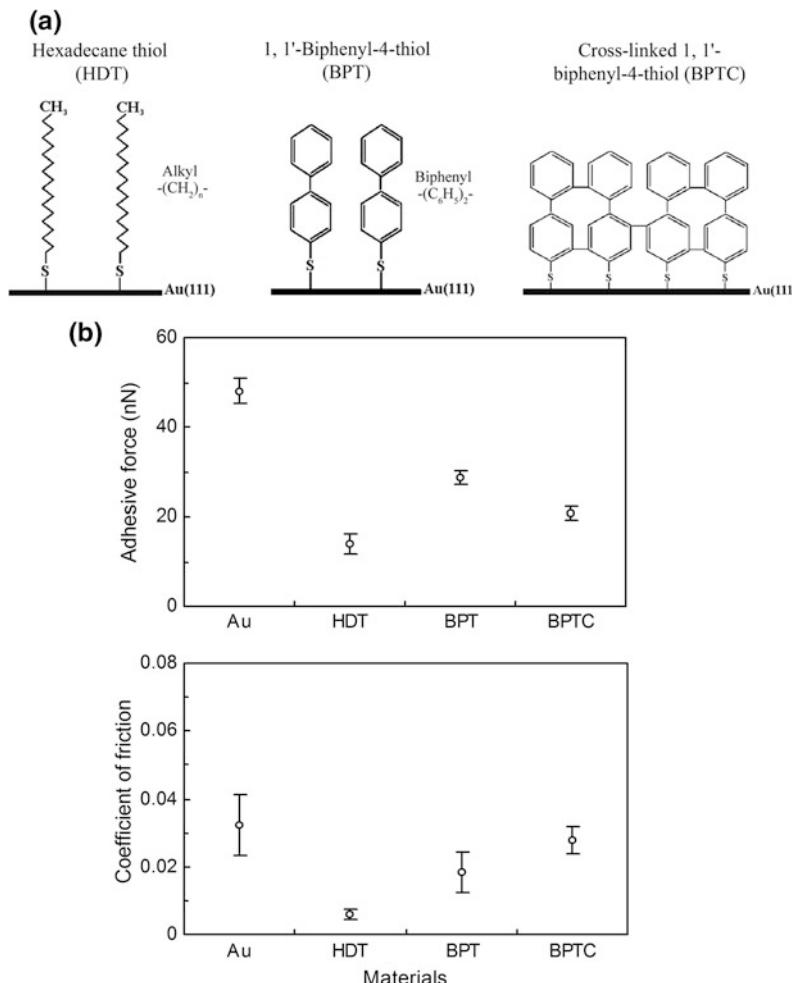


Fig. 8.68 a Schematics of structures of hexadecane thiol and biphenyl thiol SAMs on $\text{Au}(111)$ substrates, and b adhesive force and coefficient of friction of $\text{Au}(111)$ substrate and various SAMs (Bhushan and Liu 2001)

substrate. An asperity sliding on the surface of SAMs is like a tip sliding on the top of “molecular springs or brush.” The molecular spring assembly has compliant features and can experience orientation and compression under load. 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 orientation is determined by the spring constant of a single molecule as well as by the interaction between the neighboring molecules, which can be reflected by the packing density or the packing energy. It should be noted that the ori-

entation can lead to conformational defects along the molecular chains, which lead to energy dissipation.

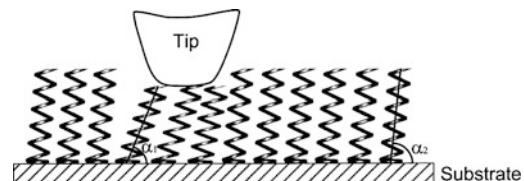


Fig. 8.69 Molecular spring model of SAMs. In this figure, $\alpha_1 < \alpha_2$, which is caused by the further orientation under the normal load applied by an asperity tip (Bhushan and Liu 2001)

An elegant way to demonstrate the influence of molecular stiffness on friction is to investigate SAMs with different structures on the same wafer. For this purpose, a micropatterned SAM was prepared. First the biphenyldimethylchlorosilane (BDCS) was deposited on silicon by a typical self-assembly method (Liu and Bhushan 2002). Then the film was partially crosslinked using mask technique by low energy electron irradiation. Finally the micropatterned BDCS films were realized, which had the as-deposited and crosslinked coating regions on the same wafer. The local stiffness properties of this micropatterned sample were investigated by force modulation AFM technique (DeVecchio and Bhushan 1997). The variation in the deflection amplitude provides a measure of the relative local stiffness of the surface. Surface height, stiffness, and friction images of the micropatterned biphenyldimethylchlorosilane (BDCS) specimen are obtained and presented in Fig. 8.70 (Liu and Bhushan 2002). The circular areas correspond to the as-deposited film, and the

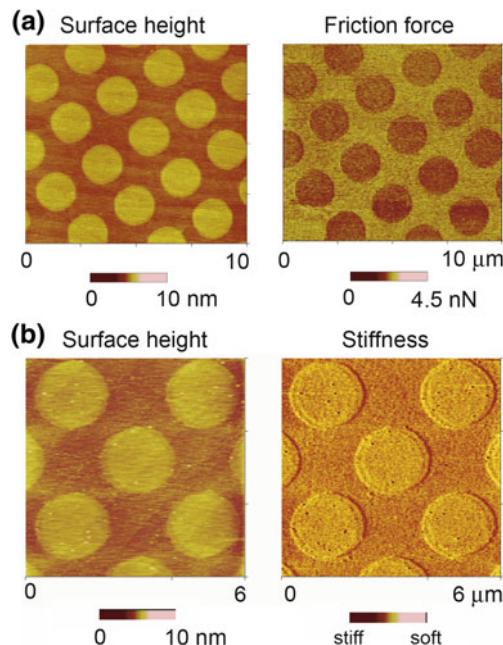


Fig. 8.70 **a** AFM Grayscale surface height and stiffness images, and **b** AFM grayscale surface height and friction force images of micropatterned BDCS (Liu and Bhushan 2002)

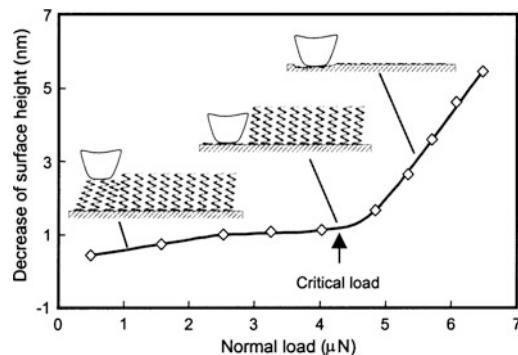


Fig. 8.71 Illustration of the wear mechanism of SAMs with increasing normal load (Liu and Bhushan 2002)

remaining area to the crosslinked film. Figure 8.70a indicates that crosslinking caused by the low energy electron irradiation leads to about 0.5 nm decrease of the surface height of BDCS films. The corresponding stiffness images indicate that the crosslinked area has a higher stiffness than the as-deposited area. Figure 8.70b indicates that the as-deposited area (higher surface height area) has a lower friction force. Obviously, these data of the micropatterned sample prove that the local stiffness of SAMs has an influence on their friction performance. Higher stiffness leads to larger friction force. These results provide strong proof of the suggested molecular spring model.

The SAMs with high-compliance long carbon chains also exhibit the best wear resistance (Bhushan and Liu 2001; Liu and Bhushan 2002). In wear experiments, the wear depth as a function of normal load curves shows a critical normal load, at which the film wears rapidly. A representative curve is shown in Fig. 8.71. Below the critical normal load, SAMs undergo orientation; at the critical load SAMs wear away from the substrate due to the relatively weak interface bond strengths, while above the critical normal load severe wear takes place on the substrate.

8.6.3 Liquid Film Thickness Measurements

Liquid film thickness mapping of ultra-thin films (on the order of couple of 2 nm) can be obtained

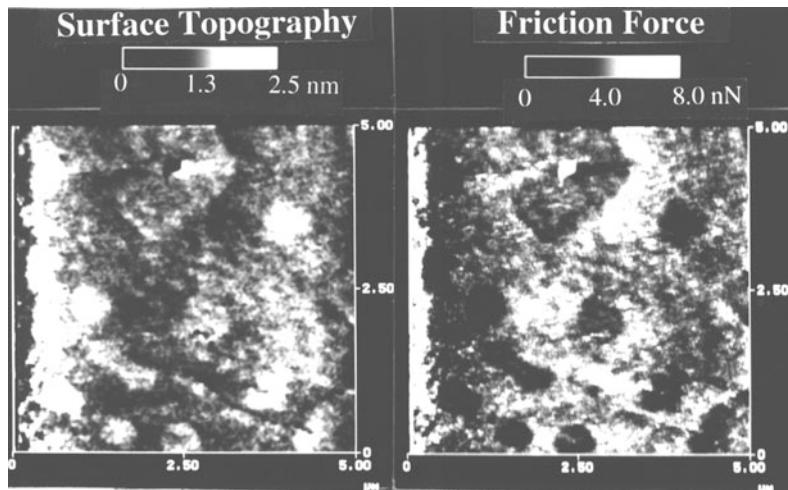


Fig. 8.72 Gray scale plots of the surface topography and friction force obtained simultaneously for unbonded Demnum type perfluoropolyether lubricant film on silicon (Koinkar and Bhushan 1996a)

using friction force microscopy (Koinkar and Bhushan 1996a) and adhesive force mapping (Bhushan and Dandavate 2000). Figure 8.72 shows the gray scale plots of the surface topography and friction force obtained simultaneously for unbonded Demnum S-100 type PFPE lubricant film on silicon. Demnum type PFPE lubricant (Demnum, Daikin, Japan) chains have $-\text{CF}_2-\text{CH}_2-\text{OH}$, a reactive end group on one end, whereas Z-DOL chains have the hydroxyl groups on both ends, as described earlier. The friction force plot shows well distinguished low and high friction regions roughly corresponding to high and low regions in surface topography (thick and thin lubricant regions). A uniformly lubricated sample does not show such a variation in the friction. Friction force imaging can thus be used to measure the lubricant uniformity on the sample surface, which cannot be identified by surface topography alone. Figure 8.73 shows the gray scale plots of the adhesive force distribution for silicon samples coated uniformly and nonuniformly with Z-DOL type PFPE lubricant. It can be clearly seen that there exists a region which has adhesive force distinctly different from the other region for the nonuniformly coated sample.

This implies that the liquid film thickness is nonuniform, giving rise to a difference in the meniscus forces.

Quantitative measurements of liquid film thickness of thin lubricant films (on the order of few nm) with nanometer lateral resolution can be made with the AFM (Bhushan and Blackman 1991; Bhushan 1999a, 2011; Chen and Bhushan 2005). The liquid film thickness is obtained by measuring the force on the tip as it approaches, contacts, and pushes through the liquid film and ultimately contacts the substrate. The distance between the sharp snap-in (owing to the formation of a liquid meniscus and van der Waals forces between the film and the tip) at the liquid surface and the hard repulsion at the substrate surface is a measure of the liquid film thickness. Figure 8.74 shows a plot of forces between tip and virgin and treated hair with hair conditioner. The hair sample was first brought into contact with the tip and then pulled away at a velocity of 400 nm/s. The zero tip-sample separation is defined to be the position where the force on the tip is zero, and the tip is not in contact with the sample. As the tip approaches the sample, a negative force exists which indicates an attractive

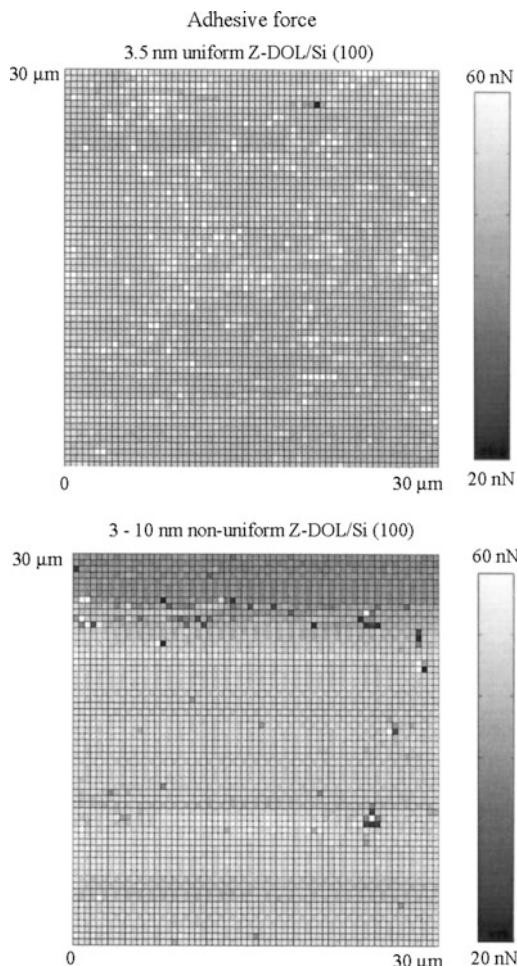


Fig. 8.73 Gray-scale plots of the adhesive force distribution of a uniformly-coated, 3.5-nm thick unbonded Z-DOL film on silicon and 3- to 10-nm thick unbonded Z-DOL film on silicon that was deliberately coated nonuniformly by vibrating the sample during the coating process (Bhushan and Dandavate 2000)

force. The treated hair surface shows a much longer range of interaction with the tip compared to the very short range of interaction between virgin hair surfaces and the tip. Typically, the tip suddenly snaps into contact with the conditioner layer at a finite separation H (about 30 nm), which is proportional to conditioner thickness h . As the tip contacts the substrate, the tip travels with the sample. When the sample is withdrawn, the forces on the tip slowly decrease to zero once the meniscus of liquid is drawn out from the hair surface. It should be noted that the distance H

between the sharp snap-in at the liquid surface and the hard wall contact with the substrate is not the real conditioner thickness h . Due to interaction of the liquid with the tip at some spacing distance, H tends to be thicker than the actual film thickness, but can still provide an estimate of the actual film thickness and upper limit of thickness.

8.7 Closure

For most solid-solid interfaces of technological relevance, contact occurs at multiple asperities. A sharp AFM/FFM tip sliding on a surface simulates just one such contact. However, asperities come in all shapes and sizes. The effect of the radius of a single asperity (tip) on the friction/adhesion performance can be studied using tips of different radii. AFM/FFM are used to study various tribological phenomena including surface roughness, adhesion, friction, scratching, wear, indentation, detection of material transfer, and boundary lubrication. Measurement of the atomic-scale friction of a freshly-cleaved, highly-oriented, pyrolytic graphite exhibits the same periodicity as that of the corresponding topography. However, the peaks in friction and those in the corresponding topography are displaced relative to each other. Variations in atomic-scale friction and the observed displacement can be explained by the variation in interatomic forces in the normal and lateral directions. The relevant friction mechanism is atomic-scale stick-slip. Local variations in microscale friction occur and are found to correspond to the local slopes, suggesting that a ratchet mechanism and collision effects are responsible for this variation. Directionality in the friction is observed on both micro- and macroscales which results from the surface roughness and surface preparation. Anisotropy in surface roughness accentuates this effect. Friction contrast in conventional frictional measurements is based on interactions dependent upon interfacial material properties superimposed by roughness-induced lateral forces. To obtain roughness-independent friction, lateral or torsional modulation techniques can be used. These techniques also allow

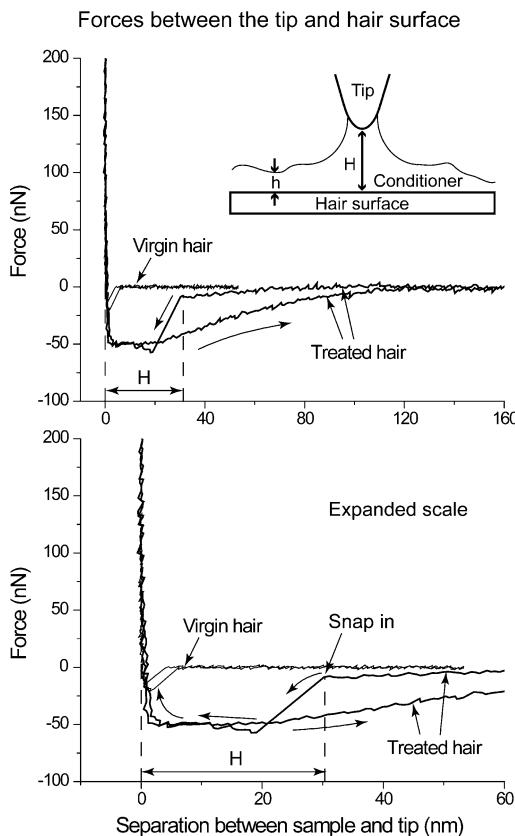


Fig. 8.74 Forces between tip and hair surface as a function of tip sample separation for virgin hair and the conditioner treated hair. A schematic of measurement for localized conditioner thickness is shown in the inset at the top. The expanded scale view of force curve at small separation is shown at the bottom (Chen and Bhushan 2005)

measurements over a small region. AFM/FFM experiments are generally conducted at relative velocities up to about 200 $\mu\text{m/s}$. High velocity experiments can be performed by either mounting a sample on a shear wave transducer driven at very high frequencies or mounting a sample on a high velocity piezo stage. By using these techniques, friction and wear experiments can be performed at a range of sliding velocities as well as normal loads, and the data have been used to develop nanoscale friction and wear maps. Relevant friction mechanisms are different for different ranges of sliding velocities and normal loads.

Adhesion and friction in a wet environment depend on the tip radius, surface roughness, and

relative humidity. Superhydrophobic surfaces can be designed by roughness optimization.

Nanoscale friction is generally found to be smaller than microscale friction. There are several factors responsible for the differences including wear and contaminant particles, the transition from elasticity to plasticity, scale-dependent roughness and mechanical properties, and meniscus effects. Nanoscale friction values increase with an increase in the normal load above a certain critical load (pressure), approaching the macro-scale friction. The critical contact pressure corresponds to the hardness of the softer of the two contacting materials.

The wear rate on the microscale for single-crystal silicon is negligible below 20 μN , and is much higher and remains approximately constant at higher loads. Elastic deformation at low loads is responsible for negligible wear. Most of the wear debris is loose. SEM and TEM studies of the wear region suggest that the material on the microscale is removed by plastic deformation with a small contribution from elastic fracture; this observation corroborates with the scratch data. Evolution of wear has also been studied using AFM. Wear is found to be initiated at nanoscratches. For a sliding interface requiring near-zero friction and wear, contact stresses should be below the hardness of the softer material to minimize plastic deformation and surfaces should be free of nanoscratches. Further, wear precursors can be detected at early stages of wear by using surface potential measurements. It is found that even in the case of zero wear (no measurable deformation of the surface using AFM), there can be a significant change in the surface potential inside the wear mark, which is useful for the study of wear precursors. Detection of material transfer on a nanoscale is possible with AFM.

In situ surface characterization of the local deformation of materials and thin coatings can be carried out using a tensile stage inside an AFM. An AFM can also be used for nanofabrication/nanomachining.

A modified AFM can be used to obtain load-displacement curves and for the measurement of nanoindentation hardness and Young's

modulus of elasticity, with a depth of indentation as low as 1 nm. A continuous stiffness technique can be used to measure variations in mechanical properties of layered materials and nonhomogeneous composites. Mechanical properties are scale dependent. For example, hardness of ceramics on nanoscales is found to be higher than that on the microscale. Ceramics exhibit significant plasticity and creep on a nanoscale. By using the force modulation technique, localized surface elasticity maps of composite materials with penetration depth as low as 1 nm can be obtained. By using phase contrast microscopy in tapping or torsional mode, it is possible to obtain phase contrast maps or the contrast in viscoelastic properties of near-surface regions. Scratching and indentation on the nanoscale are powerful ways to screen for adhesion and resistance to deformation of ultra-thin films.

Boundary lubrication studies and measurement of lubricant-film thickness with a lateral resolution on the nanoscale can be conducted using AFM. Chemically-bonded lubricant films and self-assembled monolayers are superior in friction and wear resistance. For chemically bonded lubricant films, the adsorption of water, formation of meniscus and its changes during sliding, and surface properties play an important role on the adhesion, friction, and durability of these films. Sliding velocity, relative humidity, and temperature affect adhesion and friction. For SAMs, their friction mechanism is explained by a so-called “molecular spring” model. Films with high-compliance, long carbon chains exhibit low friction and wear. Also, perfluoroalkylsilane SAMs on Si appear to be more hydrophobic with lower adhesion than alkylsilane SAMs on Si.

Investigations of adhesion, friction, wear, scratching, and indentation on the nanoscale using an AFM can provide insight into the failure mechanisms of materials. Coefficients of friction, wear rates, and mechanical properties such as hardness have been found to be different on the nanoscale than on the macroscale; generally, the coefficients of friction and wear rates on micro- and nanoscales are smaller, whereas the hardness is greater. Therefore, nanotribological studies may help define the regimes for ultra-low friction

and near-zero wear. These studies also provide insight into the atomic origins of adhesion, friction, wear, and lubrication mechanisms.

References

- Amelio S, Goldade AV, Rabe U, Scherer V, Bhushan B, Arnold W (2001) Measurements of elastic properties of ultra-thin diamond-like carbon coatings using atomic force acoustic microscopy. *Thin Solid Films* 392:75–84
- Anczykowski B, Kruger D, Babcock KL, Fuchs H (1996) Basic properties of dynamic force microscopy with the scanning force microscope in experiment and simulation. *Ultramicroscopy* 66:251–259
- Anonymous (1959) The industrial graphite engineering handbook, National Carbon Company, New York
- Anonymous (2002) Properties of silicon, EMIS data reviews series no. 4. INSPEC, Institution of Electrical Engineers, London. See also Anonymous, MEMS Materials Database, <http://www.memsnet.org/material/>
- Avila A, Bhushan B (2010) Electrical measurement techniques in atomic force microscopy, (invited). *Crit Rev Solid State Mater Sci* 35:38–51
- Bhushan B (1995) Micro/Nanotribology and its applications to magnetic storage devices and MEMS. *Tribol Int* 28:85–95
- Bhushan B (1996) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer-Verlag, New York
- Bhushan B (1997) Micro/Nanotribology and its applications, vol E330. Kluwer Academic Pub, Dordrecht, Netherlands
- Bhushan B (1998) Tribology issues and opportunities in MEMS. Kluwer Academic Pub, Dordrecht, Netherlands
- Bhushan B (1999a) Handbook of Micro/Nanotribology, 2nd edn. CRC Press, Boca Raton, Florida
- Bhushan B (1999b) Nanoscale Tribophysics and Tribomechanics. *Wear* 225–229:465–492
- Bhushan B (1999c) Wear and mechanical characterisation on micro- to picoscales using AFM. *Int Mat Rev* 44:105–117
- Bhushan B (1999d) Chemical, mechanical and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments. *Diam Relat Mater* 8:1985–2015
- Bhushan B (2001a) Modern Tribology handbook, vol 1: principles of Tribology. CRC Press, Boca Raton, Florida
- Bhushan B (2001b) Fundamentals of tribology and bridging the gap between the Macro- and Micro/Nanoscales, NATO Science series II–vol 10, Kluwer Academic Pub, Dordrecht, Netherlands
- Bhushan B (2001c) Nano- to microscale wear and mechanical characterization studies using scanning probe microscopy. *Wear* 251:1105–1123

- Bhushan B (2003) Adhesion and stiction: mechanisms, measurement techniques, and methods for reduction. *J Vac Sci Technol B* 21:2262–2296
- Bhushan B (2005) Nanotribology and nanomechanics. *Wear* 259:1507–1531
- Bhushan B (2008) Nanotribology, nanomechanics and nanomaterials characterization. *Phil Trans R Soc A* 366:1351–1381
- Bhushan B (2011) Nanotribology and nanomechanics I & II—an introduction, 3rd edn. Springer-Verlag, Heidelberg, Germany
- Bhushan B (2013a) Principles and applications of Tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Introduction to Tribology, 2nd edn. Wiley, New York
- Bhushan B (2016) Encyclopedia of nanotechnology, 2nd edn. Springer International, Switzerland
- Bhushan B, Blackman GS (1991) Atomic force microscopy of magnetic rigid disks and sliders and its applications to Tribology, ASME. *J Tribol* 113:452–458
- Bhushan B, Dandavate C (2000) Thin-film friction and adhesion studies using atomic force Microscopy. *J Appl Phys* 87:1201–1210
- Bhushan B, Goldade AV (2000a) Measurements and analysis of surface potential change during wear of single crystal silicon (100) at ultralow loads using kelvin probe Microscopy. *Appl Surf Sci* 157:373–381
- Bhushan B, Goldade AV (2000b) Kelvin probe microscopy measurements of surface potential change under wear at low loads. *Wear* 244:104–117
- Bhushan B, Gupta BK (1991) Handbook of Tribology: Materials, Coatings and Surface Treatments, McGraw-Hill, New York (reprinted Krieger, Malabar Florida, 1997)
- Bhushan B, Kasai T (2004) A surface topography-independent friction measurement technique using torsional resonance mode in an AFM. *Nanotechnology* 15:923–935
- Bhushan B, Koinkar VN (1994a) Tribological studies of silicon for magnetic recording applications. *J Appl Phys* 75:5741–5746
- Bhushan B, Koinkar VN (1994b) Nanoindentation hardness measurements using atomic force microscopy. *Appl Phys Lett* 64:1653–1655
- Bhushan B, Kulkarni AV (1996) Effect of normal load on microscale friction measurements. *Thin Solid Films* 278:49–56, 293, 333
- Bhushan B, Li X (2003) Nanomechanical characterisation of solid surfaces and thin films (invited). *Intern Mat Rev* 48:125–164
- Bhushan B, Liu H (2001) Nanotribological properties and mechanisms of alkylthiol and biphenyl thiol self-assembled monolayers studied by AFM. *Phys. Rev. B* 63:245412–1 to 245412-11
- Bhushan B, Nosonovsky M (2003) Scale effects in friction using strain gradient plasticity and dislocation-assisted sliding (Microslip). *Acta Mater* 51:4331–4345
- Bhushan B, Nosonovsky M (2004a) Comprehensive model for scale effects in friction due to adhesion and two- and three-body deformation (plowing). *Acta Mater* 52:2461–2474
- Bhushan B, Nosonovsky M (2004b) Scale effects in dry and wet friction, wear, and interface temperature. *Nanotechnology* 15:749–761
- Bhushan B, Qi J (2003) Phase contrast imaging of nanocomposites and molecularly-thick lubricant films in magnetic media. *Nanotechnology* 14:886–895
- Bhushan B, Ruan J (1994) Atomic-scale friction measurements using friction force microscopy: part II—application to magnetic media, ASME. *J. Trib.* 116:389–396
- Bhushan B, Sundararajan S (1998) Micro/nanoscale friction and wear mechanisms of thin films using atomic force and friction force microscopy. *Acta Mater* 46:3793–3804
- Bhushan B, Venkatesan S (1993) Mechanical and tribological properties of silicon for micromechanical applications: a review. *Adv Info Storage Syst* 5:211–239
- Bhushan B, Ruan J, Gupta BK (1993) A Scanning Tunnelling Microscopy Study of Fullerene Films. *J Phys D Appl Phys* 26:1319–1322
- Bhushan B, Koinkar VN, Ruan J (1994) Microtribology of Magnetic Media. *Proc Inst Mech. Eng Part J J Eng Tribol* 208:17–29
- Bhushan B, Israelachvili JN, Landman U (1995a) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Kulkarni AV, Koinkar VN, Boehm M, Odoni L, Martelet C, Belin M (1995b) Microtribological characterization of self-assembled and langmuir-blodgett monolayers by atomic and friction force microscopy. *Langmuir* 11:3189–3198
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996) Nano/picoindentation measurement using a capacitance transducer system in atomic force microscopy. *Philos Mag* 74:1117–1128
- Bhushan B, Mokashi PS, Ma T (2003) A new technique to measure poisson's ratio of ultrathin polymeric films using atomic force microscopy. *Rev Sci Instrum* 74:1043–1047
- Bhushan B, Kasai T, Nguyen CV, Meyyappan M (2004a) Multiwalled carbon nanotube AFM Probes for surface characterization of micro/nanostructures. *Microsys Technol* 10:633–639
- Bhushan B, Liu H, Hsu SM (2004b) Adhesion and friction studies of silicon and hydrophobic and low friction films and investigation of scale effects ASME. *J Tribol* 126:583–590
- Bhushan B, Kasai T, Kulik G, Barbieri L, Hoffmann P (2005) AFM study of perfluorosilane and alkylsilane self-assembled monolayers for anti-stiction in MEMS/NEMS. *Ultramicroscopy* 105:176–188
- Bhushan B, Hansford D, Lee KK (2006) Surface modification of silicon and polydimethylsiloxane surfaces with vapor-phase-deposited ultrathin fluorosilane films for biomedical nanodevices. *J Vac Sci Technol A* 24:1197–1202
- Bhushan B, Cichomski M, Tao Z, Tran NT, Ethen T, Merton C, Jewett RE (2007) Nanotribological

- characterization and lubricant degradation studies of metal-film magnetic tapes using novel lubricants, ASME. J Tribol 129:621–627
- Bhushan B, Palacio M, Kinzig B (2008) AFM-based nanotribological and electrical characterization of ultrathin wear-resistant ionic liquid films. *J Colloid Inter Sci* 317:275–287
- Binnig G, Quate CF, Gerber Ch (1986) Atomic force microscopy. *Phys Rev Lett* 56:930–933
- Binnig G, Gerber Ch, Stoll E, Albrecht TR, Quate CF (1987) Atomic resolution with atomic force microscope. *Europhys Lett* 3:1281–1286
- Bobji MS, Bhushan B (2001a) Atomic force microscopic study of the micro-cracking of magnetic thin films under tension. *Scripta Mater* 44:37–42
- Bobji MS, Bhushan B (2001b) In-situ microscopic surface characterization studies of polymeric thin films during tensile deformation using atomic force microscopy. *J Mater Res* 16:844–855
- Bowden FP, Tabor D (1950) The friction and lubrication of solids, part 1. Clarendon Press, Oxford, U.K
- Chen N, Bhushan B (2005) Morphological, nanomechanical and cellular structural characterization of human hair and conditioner distribution using torsional resonance mode in an AFM. *J Microsc* 220:96–112
- DeRose JA, Hoque E, Bhushan B, Mathieu HJ (2008) Characterization of perfluorodecanoate self-assembled monolayers on aluminum and comparison of stability with phosphonate and siloxy self-assembled monolayers. *Surf Sci* 602:1360–1367
- DeVecchio D, Bhushan B (1997) Localized surface elasticity measurements using an atomic force microscope. *Rev Sci Instrum* 68:4498–4505
- DeVecchio D, Bhushan B (1998) Use of a nanoscale kelvin probe for detecting wear precursors. *Rev Sci Instrum* 69:3618–3624
- Field JE (ed) (1992) The properties of natural and synthetic diamond. Academic Press, London
- Fleck NA, Muller GM, Ashby MF, Hutchinson JW (1994) Strain gradient plasticity: theory and experiment. *Acta Metall Mater* 42:475–487
- Frisbie CD, Rozsnyai LF, Noy A, Wrighton MS, Lieber CM (1994) Functional group imaging by chemical force microscopy. *Science* 265:2071–2074
- Fusco C, Fasolino A (2005) Velocity dependence of atomic-scale friction: a comparative study of the one- and two-dimensional tomlinson model. *Phys Rev B* 71:045413
- Garcia R, Tamayo J, Calleja M, Garcia F (1998) Phase contrast in tapping-mode scanning force microscopy. *Appl Phys A* 66:S309–S312
- Gnecco E, Bennewitz R, Gyalog T, Loppacher Ch, Bammerlin M, Meyer E, Guntherodt H-J (2000) Velocity dependence of atomic friction. *Phys Rev Lett* 84:1172–1175
- Guntherodt HJ, Anselmetti D, Meyer E (1995) Forces in scanning probe methods, vol E286. Kluwer Academic Pub, Dordrecht, Netherlands
- Helman JS, Baltensperger W, Holyst JA (1994) Simple-model for dry friction. *Phys Rev B* 49:3831–3838
- Hoque E, DeRose JA, Hoffmann P, Mathieu HJ, Bhushan B, Cichomski M (2006a) Phosphonate self-assembled monolayers on aluminum surfaces. *J Chem Phys* 124:174710
- Hoque E, DeRose JA, Kulik G, Hoffmann P, Mathieu HJ, Bhushan B (2006b) Alkylphosphonate modified aluminum oxide surfaces. *J Phys Chem B* 110:10855–10861
- Hoque E, DeRose JA, Hoffmann P, Bhushan B, Mathieu HJ (2007a) Alkylperfluorosilane self-assembled monolayers on aluminum: a comparison with alkylphosphonate self-assembled monolayers. *J Phys Chem C* 111:3956–3962
- Hoque E, DeRose JA, Hoffmann P, Bhushan B, Mathieu HJ (2007b) Chemical stability of nonwetting, low adhesion self-assembled monolayer films formed by perfluoroalkylsilazation of copper. *J Chem Phys* 126:114706
- Hoque E, DeRose JA, Bhushan B, Mathieu HJ (2008) Self-assembled monolayers on aluminum and copper oxide surfaces: surface and interface characteristics, nanotribological properties, and chemical stability. In: Bhushan B, Fuchs H, Tomitori M (eds) Applied scanning probe methods vol IX—characterization. Springer, Heidelberg, pp 235–281
- Hoque E, DeRose JA, Bhushan B, Hipps KW (2009) low adhesion, non-wetting phosphonate self-assembled monolayer films formed on copper oxide surfaces. *Ultramicroscopy* 109:1015–1022
- Kasai T, Bhushan B, Huang L, Su C (2004) Topography and phase imaging using the torsional resonance mode. *Nanotechnology* 15:731–742
- Kasai T, Bhushan B, Kulik G, Barbieri L, Hoffmann P (2005) Nanotribological study of perfluorosilane sams for anti-stiction and low wear. *J Vac Sci Technol B* 23:995–1003
- Koinkar VN, Bhushan B (1996a) Micro/nanoscale studies of boundary layers of liquid lubricants for magnetic disks. *J Appl Phys* 79:8071–8075
- Koinkar VN, Bhushan B (1996b) Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J Vac Sci Technol A* 14:2378–2391
- Koinkar VN, Bhushan B (1996c) Microtribological studies of Al_2O_3 -TiC, polycrystalline and single-crystal Mn-Zn Ferrite and SiC head slider materials. *Wear* 202:110–122
- Koinkar VN, Bhushan B (1997a) Microtribological properties of hard amorphous carbon protective coatings for thin film magnetic disks and heads. *Proc Inst Mech Eng Part J: J Eng Tribol* 211:365–372
- Koinkar VN, Bhushan B (1997b) Effect of scan size and surface roughness on microscale friction measurements. *J Appl Phys* 81:2472–2479
- Koinkar VN, Bhushan B (1997c) Scanning and transmission electron microscopies of single-crystal silicon microworn/machined using atomic force microscopy. *J Mater Res* 12:3219–3224
- Krotil HU, Stifter T, Waschippy H, Weishaupt K, Hild S, Marti O (1999) Pulse force mode: a new method for

- the investigation of surface properties. *Surf Interface Anal* 27:336–340
- Kulkarni AV, Bhushan B (1996a) Nanoscale mechanical property measurements using modified atomic force microscopy. *Thin Solid Films* 290–291:206–210
- Kulkarni AV, Bhushan B (1996b) Nano/picoindentation measurements on single-crystal aluminum using modified atomic force microscopy. *Mater Lett* 29:221–227
- Kulkarni AV, Bhushan B (1997) Nanoindentation measurement of amorphous carbon coatings. *J Mater Res* 12:2707–2714
- Lee DT, Pelz JP, Bhushan B (2002) Instrumentation for direct, low frequency scanning capacitance microscopy, and analysis of position dependent stray capacitance. *Rev Sci Instr* 73:3523–3533
- Lee KK, Bhushan B, Hansford D (2005) Nanotribological characterization of perfluoropolymer thin films for biomedical micro/nanoelectromechanical systems applications. *J Vac Sci Technol A* 23:804–810
- Li WB, Henshall JL, Hooper RM, Easterling KE (1991) The Mechanism of Indentation Creep. *Acta Metall Mater* 39:3099–3110
- Li X, Bhushan B (2002) A review of nanoindentation continuous stiffness measurement technique and its applications. *Mater Charact* 48:11–36
- Lim SC, Ashby MF (1987) Wear mechanism maps. *Acta Metall* 35:1–24
- Lim SC, Ashby MF, Brunton JH (1987) Wear-rate transitions and their relationship to wear mechanisms. *Acta Metall* 35:1343–1348
- Liu H, Bhushan B (2002) Investigation of nanotribological properties of self-assembled monolayers with Alkyl and Biphenyl spacer chains. *Ultramicroscopy* 91:185–202
- Liu H, Bhushan B (2003a) Nanotribological characterization of molecularly-thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* 97:321–340
- Liu H, Bhushan B (2003b) Adhesion and friction studies of microelectromechanical systems/nanoelectromechanical systems materials using a novel microtriboapparatus. *J Vac Sci Technol A* 21:1528–1538
- Liu H, Bhushan B, Eck W, Staedtler V (2001) Investigation of the adhesion, friction, and wear properties of biphenyl thiol self-assembled monolayers by atomic force microscopy. *J Vac Sci Technol A* 19:1234–1240
- Lodge RA, Bhushan B (2007) Effect of physical wear and triboelectric interaction on surface charges measured by kelvin probe microscopy. *J Colloid Interf Sci* 310:321–330
- Maivald P, Butt HJ, Gould SAC, Prater CB, Drake B, Gurley JA, Elings VB, Hansma PK (1991) Using force modulation to image surface elasticities with the atomic force microscope. *Nanotechnology* 2:103–106
- Marti O, Krotli H-U (2001) Dynamic friction measurement with the scanning force microscope, fundamentals of tribology and bridging the gap between the macro- and micro/nanoscales. Kluwer Academic Publishers, Dordrecht, pp 121–135
- Mate CM, McClelland GM, Erlandsson R, Chiang S (1987) Atomic-scale friction of a tungsten tip on a graphite surface. *Phys Rev Lett* 59:1942–1945
- Meyer E, Overney R, Luthi R, Brodbeck D, Howald L, Frommer J, Guntherodt HJ, Wolter O, Fujihira M, Takano T, Gotoh Y (1992) Friction force microscopy of mixed langmuir-blodgett films. *Thin Solid Films* 220:132–137
- Nagpure SC, Bhushan B, Babu SS (2011) Surface potential measurement of aged li-ion batteries using kelvin probe microscopy. *J Power Sources* 196:1508–1512
- Nix WD, Gao H (1998) Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J Mech Phys Solids* 46:411–425
- Nosonovsky M, Bhushan B (2005) Scale effects in dry friction during multiple-asperity contact, ASME. *J Tribol* 127:37–46
- Palacio M, Bhushan B (2007a) Surface potential and resistance measurements for detecting wear of chemically-bonded and unbonded molecularly-thick perfluoropolyether lubricant films using atomic force microscopy. *J Colloid Interf Sci* 315:261–269
- Palacio M, Bhushan B (2007b) Wear detection of candidate MEMS/NEMS lubricant films using atomic force microscopy-based surface potential measurements. *Scripta Mater* 57:821–824
- Palacio M, Bhushan B (2008) Ultrathin wear-resistant ionic liquid films for novel MEMS/NEMS applications. *Adv Mater* 20:1194–1198
- Palacio M, Bhushan B (2009) Molecularly thick dicationic liquid films for nanolubrication. *J Vac Sci Technol, A* 27:986–995
- Palacio M, Bhushan B (2010) Normal and lateral force calibration techniques for afm cantilevers. *Crit Rev Solid State Mater Sci* 35:73–104; 36:261
- Persson BNJ, Tosatti E (1996) PHYSICS OF SLIDING FRICTION, vol E311. Kluwer Academic Pub, Dordrecht
- Rabe U, Janser K, Arnold W (1996) vibrations of free and surface-coupled atomic force microscope cantilevers: theory and experiment. *Rev Sci Instrum* 67:3281–3293
- Reinstaedtler M, Rabe U, Scherer V, Hartmann U, Goldade A, Bhushan B, Arnold W (2003) on the nanoscale measurement of friction using atomic-force microscope cantilever torsional resonances. *Appl Phys Lett* 82:2604–2606
- Reinstaedtler M, Rabe U, Goldade A, Bhushan B, Arnold W (2005a) investigating ultra-thin lubricant layers using resonant friction force microscopy. *Tribol Inter* 38:533–541
- Reinstaedtler M, Kasai T, Rabe U, Bhushan B, Arnold W (2005b) Imaging and measurement of elasticity and friction using the TR mode. *J Phys D Appl Phys* 38: R269–R282
- Ruan J, Bhushan B (1993) Nanoindentation studies of fullerene films using atomic force microscopy. *J Mater Res* 8:3019–3022
- Ruan J, Bhushan B (1994a) Atomic-scale friction measurements using friction force microscopy: part I—

- general principles and new measurement techniques, ASME. J Tribol 116:378–388
- Ruan J, Bhushan B (1994b) Atomic-scale and microscale friction of graphite and diamond using friction force microscopy. *J Appl Phys* 76:5022–5035
- Ruan J, Bhushan B (1994c) Frictional behavior of highly oriented pyrolytic graphite. *J Appl Phys* 76:8117–8120
- Scherer V, Bhushan B, Rabe U, Arnold W (1997) Local elasticity and lubrication measurements using atomic force and friction force microscopy at ultrasonic frequencies. *IEEE Trans Magn* 33:4077–4079
- Scherer V, Arnold W, Bhushan B (1998) Active friction control using ultrasonic vibration. In: Bhushan B (ed) *Tribology issues and opportunities in MEMS*. Dordrecht, Kluwer Academic Pub, pp 463–469
- Scherer V, Arnold W, Bhushan B (1999) Lateral force microscopy using acoustic friction force microscopy. *Surf Interface Anal* 27:578–587
- Schwarz UD, Zwoerner O, Koester P, Wiesendanger R (1997) Friction force spectroscopy in the low-load regime with well-defined tips. In: Bhushan B (ed) *Micro/Nanotribology and its applications*. Kluwer Academic, Dordrecht, pp 233–238
- Scott WW, Bhushan B (2003) Use of phase imaging in atomic force microscopy for measurement of viscoelastic contrast in polymer nanocomposites and molecularly-thick lubricant films. *Ultramicroscopy* 97:151–169
- Seshadri IP, Bhushan B (2008a) In-situ tensile deformation characterization of human hair with atomic force microscopy. *Acta Mater* 56:774–781
- Seshadri IP, Bhushan B (2008b) Effect of ethnicity and treatments on in situ tensile response and morphological changes of human hair characterized by atomic force microscopy. *Acta Mater* 56:3585–3597
- Seshadri IP, Bhushan B (2008c) Effect of rubbing load on nanoscale charging characteristics of human hair characterized by AFM based kelvin probe. *J Colloid Interf Sci* 325:580–587
- Singer IL, Pollock HM (1992) Fundamentals of friction: macroscopic and microscopic processes, vol E220. Kluwer Academic Pub, Dordrecht
- Song Y, Bhushan B (2005) Quantitative extraction of in-plane surface properties using torsional resonance mode in atomic force microscopy. *J Appl Phys* 87:83533
- Stifter T, Marti O, Bhushan B (2000) Theoretical investigation of the distance dependence of capillary and van der waals forces in scanning probe microscopy. *Phys Rev B* 62:13667–13673
- Sundararajan S, Bhushan B (2000) Topography-induced contributions to friction forces measured using an atomic force/friction force microscope. *J Appl Phys* 88:4825–4831
- Sundararajan S, Bhushan B (2001) Development of a continuous microscratch technique in an atomic force microscope and its application to study scratch resistance of ultra-thin hard amorphous carbon coatings. *J Mater Res* 16:75–84
- Tamayo J, Garcia R (1996) Deformation, contact time, and phase contrast in tapping mode scanning force microscopy. *Langmuir* 12:4430–4435
- Tambe NS, Bhushan B (2004a) Scale dependence of micro/nano-friction and adhesion of MEMS/NEMS materials, coatings and lubricants. *Nanotechnology* 15:1561–1570
- Tambe NS, Bhushan B (2004b) In situ study of nano-cracking of multilayered magnetic tapes under monotonic and fatigue loading using an AFM. *Ultramicroscopy* 100:359–373
- Tambe NS, Bhushan B (2005a) A new atomic force microscopy based technique for studying nanoscale friction at high sliding velocities. *J Phys D Appl Phys* 38:764–773
- Tambe NS, Bhushan B (2005b) Friction model for the velocity dependence of nanoscale friction. *Nanotechnology* 16:2309–2324
- Tambe NS, Bhushan B (2005c) Durability studies of Micro/Nanoelectromechanical system materials, coatings, and lubricants at high sliding velocities (up to 10 mm/s) using a modified atomic force microscope. *J Vac Sci Technol, A* 23:830–835
- Tambe NS, Bhushan B (2005d) Nanoscale friction-induced phase transformation of diamond-like carbon. *Scripta Materialia* 52:751–755
- Tambe NS, Bhushan B (2005e) Identifying Materials with Low Friction and Adhesion for Nanotechnology Applications. *Appl. Phys. Lett* 86:061906. *Nature Mater. Nanozone*, 17 Feb 2005
- Tambe NS, Bhushan B (2005f) Nanoscale friction mapping. *Appl Phys Lett* 86:193102–1 to 3
- Tambe NS, Bhushan B (2005g) Nanowear mapping: a novel atomic force microscopy based approach for studying nanoscale wear at high sliding velocities. *Tribol Lett* 20:83–90
- Tambe NS, Bhushan B (2005h) Nanotribological characterization of self assembled monolayers deposited on silicon and aluminum substrates. *Nanotechnology* 16:1549–1558
- Tambe NS, Bhushan B (2008) Nanoscale friction and wear maps. *Philos Trans R Soc A* 366:1405–1424
- Tang W, Bhushan B, Ge S (2010) Triboelectrification studies of skin and skin cream using kelvin probe microscopy. *J Vac Sci Technol, A* 28:1018–1028
- Tao Z, Bhushan B (2005a) Bonding, degradation, and environmental effects on novel perfluoropolyether lubricants. *Wear* 259:1352–1361
- Tao Z, Bhushan B (2005b) Degradation mechanisms and environmental effects on perfluoropolyether, self assembled monolayers, and diamondlike carbon films. *Langmuir* 21:2391–2399
- Tao Z, Bhushan B (2006a) Surface modification of AFM silicon probes for adhesion and wear reduction. *Trib Lett* 21:1–16
- Tao Z, Bhushan B (2006b) A new technique for studying nanoscale friction at sliding velocities up to 200 mm/s using atomic force microscope. *Rev Sci Instrum* 77:103705

- Tao Z, Bhushan B (2007) Velocity dependence and rest time effect in nanoscale friction of ultrathin films at high sliding velocities. *J Vac Sci Technol, A* 25:1267–1274
- Tomanek D, Zhong W, Thomas H (1991) Calculation of an atomically modulated friction force in atomic force microscopy. *Europhys Lett* 15:887–892
- Tomlinson GA (1929) A molecular theory of friction. *Phil Mag Ser* 7:905–939
- Yamanaka K, Tomita E (1995) Lateral force modulation atomic force microscope for selective imaging of friction forces. *Jpn J Appl Phys* 34:2879–2882
- Zaghoul U, Bhushan B, Coccetti F, Pons P, Plana R (2011) Kelvin probe force microscopy based characterization techniques applied for electrostatic MEMS and thin dielectric films to investigate the dielectric substrate charging phenomena. *J Vac Sci Technol, A* 29:051101
- Zhao X, Bhushan B (1998) Material removal mechanism of single-crystal silicon on nanoscale and at ultralow loads. *Wear* 223:66–78
- Zworner O, Holscher H, Schwarz UD, Wiesendanger R (1998) The Velocity Dependence of Frictional Forces in Point-Contact Friction. *Appl Phys A: MaterSci Process* 66:S263–S267

Surface Forces and Nanorheology of Molecularly Thin Films

9

Dong Woog Lee, Marina Ruths and Jacob N. Israelachvili

Abstract

In this chapter, we describe the static and dynamic normal forces that occur between surfaces in vacuum or liquids and the different modes of friction that can be observed between: (i) bare surfaces in contact (dry or interfacial friction), (ii) surfaces separated by a thin liquid film (lubricated friction), and (iii) surfaces coated with organic monolayers (boundary friction). Experimental methods suitable for measuring normal surface forces, adhesion and friction (lateral or shear) forces of different magnitude at the molecular level are described. We explain the molecular origin of van der Waals, electrostatic, solvation and polymer-mediated interactions, and basic models for the contact mechanics of adhesive and nonadhesive elastically deforming bodies. The effects of interaction forces, molecular shape, surface structure and roughness on adhesion and friction are discussed. Simple models for the contributions of the adhesion force and external load to interfacial friction are illustrated with experimental data on both unlubricated and lubricated systems, as measured with the surface forces apparatus. We discuss rate-dependent

D.W. Lee (✉)

School of Energy and Chemical Engineering, Ulsan
National Institute of Science and Technology
(UNIST), UNIST-gil 50, Ulsan 689-798, Republic of
Korea
e-mail: dongwoog.lee@unist.ac.kr

M. Ruths

Department of Chemistry, University of
Massachusetts Lowell, 1 University Avenue, Lowell,
MA 01854, USA
e-mail: marina_ruths@uml.edu

J.N. Israelachvili

Department of Chemical Engineering and Materials
Department, University of California, Santa Barbara,
CA 93106, USA
e-mail: jacob@engineering.ucsb.edu

adhesion (adhesion hysteresis) and how this is related to friction. Some examples of the transition from wearless friction to friction with wear are shown. Lubrication in different lubricant thickness regimes is described together with explanations of nanorheological concepts. The occurrence of and transitions between smooth and stick-slip sliding in various types of dry (unlubricated and solid boundary lubricated) and liquid lubricated systems are discussed based on recent experimental results and models for stick-slip involving memory distance and dilatancy.

Keywords

Interfacial interactions • Surface forces • Friction forces • Hysteresis • Liquid films • Polymer films

9.1 Introduction: Types of Surface Forces

In this chapter, we discuss the most important types of surface forces and the relevant equations for the force and friction laws. Several different attractive and repulsive forces operate between surfaces and particles. Some forces occur in vacuum, for example, attractive van der Waals and repulsive hard-core interactions. Other types of forces can arise only when the interacting surfaces are separated by another condensed phase, which is usually a liquid. The most common types of surface forces and their main characteristics are listed in Table 9.1.

In vacuum, the two main long-range interactions are the attractive van der Waals and electrostatic (Coulomb) forces. At smaller surface separations (corresponding to molecular contact at surface separations of $D \approx 0.2$ nm), additional attractive interactions can be found such as covalent or metallic bonding forces. These attractive forces are stabilized by the hard-core repulsion. Together they determine the surface and interfacial energies of planar surfaces, as well as the strengths of materials and adhesive junctions. Adhesion forces are often strong enough to elastically or plastically deform bodies or particles when they come into contact.

In vapors (e.g., atmospheric air containing water and organic molecules), solid surfaces in, or

close to, contact will generally have a surface layer of chemisorbed or physisorbed molecules, or a capillary condensed liquid bridge between them. A surface layer usually causes the adhesion to decrease, but in the case of capillary condensation, the additional Laplace pressure or attractive *capillary* force may make the adhesion between the surfaces stronger than in an inert gas or vacuum.

When totally immersed in a *liquid*, the force between particles or surfaces is completely modified from that in vacuum or air (vapor). The van der Waals attraction is generally reduced, but other forces can now arise that can qualitatively change both the range and even the sign of the interaction. The attractive force in such a system can be either stronger or weaker than in the absence of the intervening liquid. Depending on the different forces that may be operating simultaneously in solution, the overall force law is not generally monotonically attractive even at long range; it can be repulsive, or the force can change sign at some finite surface separation. In such cases, the potential-energy minimum, which determines the adhesion force or energy, does not occur at the true molecular contact between the surfaces, but at some small distance further out.

The forces between surfaces in a liquid medium can be particularly complex at *short range*, i.e., at surface separations below a few nanometers or 4–10 molecular diameters. This is partly because, with increasing confinement, a

Table 9.1 Types of surface forces in vacuum versus in liquid (colloidal forces)

Type of force	Subclasses or alternative names	Main characteristics
<i>Attractive forces</i>		
van der Waals	Debye induced dipole force (v & s) London dispersion force (v & s) Casimir force (v & s)	Ubiquitous, occurs both in vacuum and in liquids
Electrostatic	Ionic bond (v) Coulombic force (v & s) Hydrogen bond (v) Charge-exchange interaction (v & s) Acid-base interaction (s) “Harpooning” interaction (v)	Strong, long-range, arises in polar solvents; requires surface charging or charge-separation mechanism
Ion correlation	van der Waals force of polarizable ions (s)	Requires mobile charges on surfaces in a polar solvent
Quantum mechanical	Covalent bond (v) Metallic bond (v) Exchange interaction (v)	Strong, short-range, responsible for contact binding of crystalline surfaces
Solvation	Oscillatory force (s) Depletion force (s)	Mainly entropic in origin, the oscillatory force alternates between attraction and repulsion
Hydrophobic	Attractive hydration force (s)	Strong, apparently long-range; origin not yet understood
Specific binding	“Lock-and-key” or complementary binding (v & s) Receptor-ligand interaction (s) Antibody-antigen interaction (s)	Subtle combination of different noncovalent forces giving rise to highly specific binding; main recognition mechanism of biological systems
<i>Repulsive forces</i>		
van der Waals	van der Waals disjoining pressure (s)	Arises only between dissimilar bodies interacting in a medium
Electrostatic	Coulombic force (v & s)	Arises only for certain constrained surface charge distributions
Quantum mechanical	Hard-core or steric repulsion (v) Born repulsion (v)	Short-range, stabilizing attractive covalent and ionic binding forces, effectively determine molecular size and shape
Solvation	Oscillatory solvation force (s) Structural force (s) Hydration force (s)	Monotonically repulsive forces, believed to arise when solvent molecules bind strongly to surfaces
Entropic	Osmotic repulsion (s) Double-layer force (s) Thermal fluctuation force (s) Steric polymer repulsion (s) Undulation force (s) Protrusion force (s)	Due to confinement of molecular or ionic species; requires mechanism that keeps trapped species between the surfaces
<i>Dynamic interactions</i>		
Non-equilibrium	Hydrodynamic forces (s) Viscous forces (s) Friction forces (v & s) Lubrication forces (s)	Energy-dissipating forces occurring during relative motion of surfaces or bodies

Note (v) applies only to interactions in vacuum, (s) applies only to interactions in solution (or to surfaces separated by a liquid), and (v & s) applies to interactions occurring both in vacuum and in solution

liquid ceases to behave as a structureless continuum with bulk properties; instead, the size and shape of its molecules begin to determine the overall interaction. In addition, the surfaces themselves can no longer be treated as inert and structureless walls (i.e., mathematically flat) and their physical and chemical properties at the atomic scale must now be taken into account. The force laws will then depend on whether the surfaces are amorphous or crystalline (and whether the lattices of crystalline surfaces are matched or not), rough or smooth, rigid or soft (fluid-like), and hydrophobic or hydrophilic.

It is also important to distinguish between *static* (i.e., equilibrium) interactions and *dynamic* (i.e., nonequilibrium) forces such as viscous and friction forces. For example, certain liquid films confined between two contacting surfaces may take a surprisingly long time to equilibrate, as may the surfaces themselves, so that the short-range and adhesion forces appear to be time-dependent, resulting in aging effects.

9.2 Methods Used to Study Surface Forces

9.2.1 Force Laws

The full force law $F(D)$ between two surfaces, i.e., the force F as a function of surface separation D , can be measured in a number of ways (Israelachvili 2011; Claesson et al. 1996; Craig 1997). The simplest is to move the base of a spring by a known amount ΔD_0 . Figure 9.1 illustrates this method when applied to the interaction of two magnets. However, the method is also applicable at the microscopic or molecular level, and it forms the basis of all direct force-measuring apparatuses such as the surface forces apparatus (SFA; (Israelachvili 2011; Israelachvili and Adams 1978)). If there is a detectable force between the surfaces, this will cause the force-measuring spring to deflect by

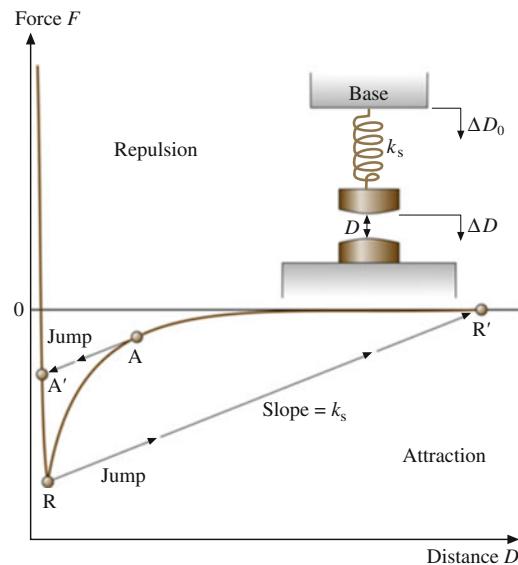


Fig. 9.1 Schematic attractive force law between two macroscopic objects such as two magnets, or between two microscopic objects such as the van der Waals force between a metal tip and a surface. On lowering the base supporting the spring, the latter will expand or contract such that, at any equilibrium separation D , the attractive force balances the elastic spring restoring force. If the gradient of the attractive force dF/dD exceeds the gradient of the spring's restoring force (defined by the spring constant k_s), the upper surface will jump from A into contact at A' (A for advancing). On separating the surfaces by raising the base, the two surfaces will jump apart from R to R' (R for receding). The distance $R-R'$ multiplied by k_s gives the adhesion force, i.e., the value of F at the point R (after (Israelachvili 1995) with permission)

ΔD_s , while the surface separation changes by ΔD . These three displacements are related by

$$\Delta D_s = \Delta D_0 - \Delta D. \quad (9.1)$$

The difference in force, ΔF , between the initial and final separations is given by

$$\Delta F = k_s \Delta D_s, \quad (9.2)$$

where k_s is the spring constant. The equations above provide the basis for measurements of the force difference between any two surface separations. For example, if a force-measuring

apparatus with a known k_s can measure D (and thus ΔD), ΔD_0 , and ΔD_s , the force difference ΔF can be measured between a large initial or reference separation D , where the force is zero ($F = 0$), and another separation $D - \Delta D$. By working one's way in increasing increments of $\Delta D = \Delta D_0 - \Delta D_s$, the full force law $F(D)$ can be constructed over any desired distance regime.

In order to measure an equilibrium force law, it is essential to establish that the two surfaces have stopped moving before the displacements are measured. When displacements are measured while two surfaces are still in relative motion, one also measures a viscous or frictional contribution to the total force. Such dynamic force measurements have enabled the viscosities of liquids near surfaces and in thin films to be accurately determined (Israelachvili 1986; Dhinojwala and Granick 1997). In practice, it is difficult to measure the forces between two perfectly flat surfaces, because of the stringent requirement of perfect alignment for making reliable measurements at distances of a few tenths of a nanometer. It is far easier to measure the forces between curved surfaces, e.g., two spheres, a sphere and a flat surface, or two crossed cylinders. Furthermore, the force $F(D)$ measured between two curved surfaces can be directly related to the energy per unit area $E(D)$ between two flat surfaces at the same separation D by the so-called Derjaguin approximation (Derjaguin 1934a)

$$E(D) = \frac{F(D)}{2\pi R}, \quad (9.3)$$

where R is the radius of the sphere (for a sphere and a flat surface) or the radii of the cylinders (for two crossed cylinders, cf. Table 9.2).

9.2.2 Adhesion Forces

The most direct way to measure the adhesion of two solid surfaces (such as two spheres or a sphere on a flat) is to suspend one of them on a spring and measure the adhesion or pull-off force needed to separate the two bodies, using the deflection of the spring. If k_s is the stiffness of the force-measuring spring and ΔD is the distance the two surfaces

jump apart when they separate, then the adhesion force F_{ad} is given by

$$F_{ad} = F_{max} = k_s \Delta D, \quad (9.4)$$

where we note that, in liquids, the maximum or minimum in the force may occur at some non-zero surface separation. From F_{ad} and a known surface geometry, and assuming that the surfaces were everywhere in molecular contact, one may also calculate the surface or interfacial energy γ . For an elastically deformable sphere of radius R on a flat surface, or for two crossed cylinders of radius R , we have (Israelachvili 2011; Johnson et al. 1971)

$$\gamma = \frac{F_{ad}}{3\pi R}, \quad (9.5)$$

while for two spheres of radii R_1 and R_2

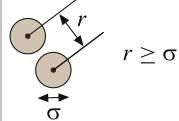
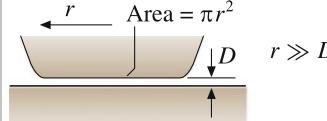
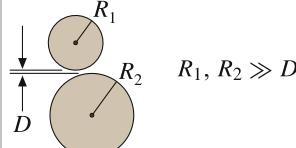
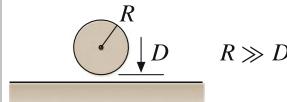
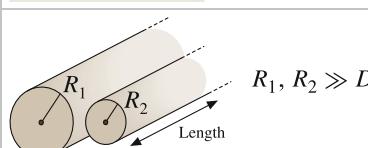
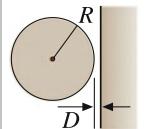
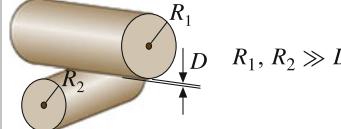
$$\gamma = \frac{F_{ad}}{3\pi} \left(\frac{1}{R_1} + \frac{1}{R_2} \right), \quad (9.6)$$

where γ is in units of $J m^{-2}$ (Sect. 9.5.2).

9.2.3 The SFA

The SFA (Fig. 9.2) is used for measurements of adhesion and force laws between two curved molecularly smooth surfaces immersed in liquids or controlled vapors (Israelachvili 2011; Israelachvili and Adams 1978; Israelachvili and McGuigan 1990). In SFA experiments, ΔD_0 is changed by expanding or contracting a piezoelectric crystal by a known amount or by moving the base of the spring with sensitive motor-driven mechanical stages. The resulting change in surface separation ΔD is measured by multiple beam interferometry with an accuracy of ± 0.1 nm, and the spring deflection ΔD_s can then be obtained according to (9.1). From the shape of the interference fringes one also obtains the radius of the surfaces R and any surface deformation that arises during an interaction (Israelachvili 1973; Heuberger et al. 1997). The resolution in the lateral direction is about 1 μm . The surface separation can be independently controlled to within 0.1 nm,

Table 9.2 Van der Waals interaction energy and force between macroscopic bodies of different geometries (Leckband and Israelachvili 2001)

Geometry of bodies with surfaces D apart ($D \ll R$)	Van der Waals interaction		
	Energy E	Force F	
Two atoms or small molecules		$\frac{-C_{vdW}}{r^6}Z$	$\frac{-6C_{vdW}}{r^7}$
Two flat surfaces (per unit area)		$\frac{-A_H}{12\pi D^2}$	$\frac{-A_H}{6\pi D^3}$
Two spheres or macromolecules of radii R_1 and R_2		$\frac{-A_H}{6D} \left(\frac{R_1 R_2}{R_1 + R_2} \right)$	$\frac{-A_H}{6D^2} \left(\frac{R_1 R_2}{R_1 + R_2} \right)$
Sphere or macromolecule of radius R near a flat surface		$\frac{-A_H R}{6D}$	$\frac{-A_H R}{6D^2}$
Two parallel cylinders or rods of radii R_1 and R_2 (per unit length)		$\frac{-A_H}{12\sqrt{2}D^{3/2}} \left(\frac{R_1 R_2}{R_1 + R_2} \right)^{1/2}$	$\frac{-A_H}{8\sqrt{2}D^{5/2}} \left(\frac{R_1 R_2}{R_1 + R_2} \right)^{1/2}$
Cylinder of radius R near a flat surface (per unit length)		$\frac{-A_H \sqrt{R}}{12\sqrt{2}D^{3/2}}$	$\frac{-A_H \sqrt{R}}{8\sqrt{2}D^{5/2}}$
Two cylinders or filaments of radii R_1 and R_2 crossed at 90°		$\frac{-A_H \sqrt{R_1 R_2}}{6D}$	$\frac{-A_H \sqrt{R_1 R_2}}{6D^2}$

and the force sensitivity is about 10^{-8} N. For a typical surface radius of $R \approx 1$ cm, γ values can be measured to an accuracy of about 10^{-3} mJ m $^{-2}$.

Several different materials have been used to form the surfaces in the SFA, including mica (Pashley 1981, 1982), silica (Horn et al. 1989), sapphire (Horn et al. 1988), and polymer sheets (Merrill et al. 1991). These materials can also be used as supporting substrates in experiments on

the forces between adsorbed or chemically bound polymer layers (Dhinojwala and Granick 1997; Klein 1983; Patel and Tirrell 1989; Watanabe and Tirrell 1993; Ruths et al. 2000), surfactant and lipid monolayers, and bilayers (Helm et al. 1989; Chen et al. 1991; Leckband et al. 1992; Lee et al. 2014, 2015), and metal and metal oxide layers (Coakley and Tabor 1978; Hirz et al. 1992; Steinberg et al. 1993; Vigil et al. 1994; Ruths et al. 2001). The range of liquids and

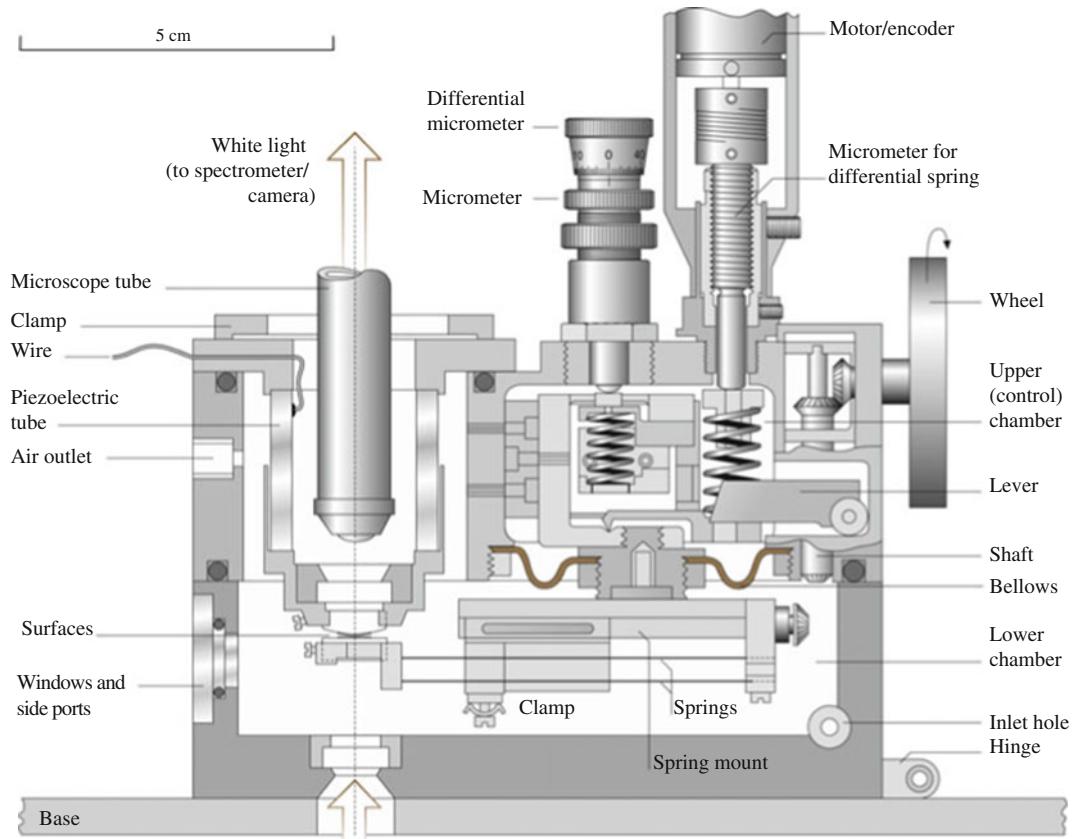


Fig. 9.2 A surface forces apparatus (SFA) where the intermolecular forces between two macroscopic, cylindrical surfaces of local radius R can be directly measured as a function of surface separation over a large distance regime from tenths of a nanometer to micrometers. Local or transient surface deformations can be detected optically. Various attachments for moving one surface laterally with respect to the other have been developed for friction measurements in different regimes of sliding velocity and sliding distance (after (Israelachvili and McGuiggan 1990) with permission)

vapors that can be used is almost endless, and have thus far included aqueous solutions, organic liquids and solvents, polymer melts, various petroleum oils and lubricant liquids, dyes, and liquid crystals.

Friction attachments for the SFA (Homola et al. 1989; Van Alsten and Granick 1990; Klein and Kumacheva 1998; Luengo et al. 1997a; Lee et al. 2013) allow for the two surfaces to be sheared laterally past each other at varying sliding speeds or oscillating frequencies, while simultaneously measuring both the transverse (frictional or shear) force and the normal force (load) between them. The ranges of friction forces and sliding speeds that can be studied with such methods are currently

10^{-7} – 10^{-1} N and 10^{-13} – 10^{-2} m s $^{-1}$, respectively (Kumacheva 1998). The externally applied load L can be varied continuously, and both positive and negative loads can be applied. The distance between the surfaces D , their true molecular contact area, their elastic (or viscoelastic or elastohydrodynamic) deformation, and their lateral motion can all be monitored simultaneously by recording the moving interference fringe pattern. Equipment for dynamic measurements of normal forces has also been developed. Such measurements give information on the viscosity of the medium and the location of the shear or slip planes relative to the surfaces (Israelachvili 1986; Dhinojwala and Granick 1997).

9.3 Normal Forces Between Dry (Unlubricated) Surfaces

9.3.1 Van der Waals Forces in Vacuum and Inert Vapors

Forces between macroscopic bodies (such as colloidal particles) across vacuum arise from interactions between the constituent atoms or molecules of each body across the gap separating them. These intermolecular interactions are electromagnetic forces between permanent or induced dipoles (van der Waals forces), and between ions (electrostatic forces). In certain systems, other types of interactions can also be observed between surfaces in vacuum, e.g., charge exchange interactions, sintering, and cold welding. In this section, we describe the van der Waals forces, which occur between all atoms and molecules and between all macroscopic bodies (Israelachvili 2011).

The interaction between two permanent dipoles with a fixed relative orientation can be attractive or repulsive. For the specific case of two freely rotating permanent dipoles in a liquid or vapor (orientational or Keesom interaction), and for a permanent dipole and an induced dipole in an atom or polar or nonpolar molecule (induction or Debye interaction), the interaction is on average always attractive. The third type of van der Waals interaction, the fluctuation or London dispersion interaction, arises from instantaneous polarization of one nonpolar or polar molecule due to fluctuations in the charge distribution of a neighboring nonpolar or polar molecule (Fig. 9.3). Correlation between these fluctuating induced dipole moments gives an attraction that is present between any two molecules or surfaces across vacuum. At very small separations, the interaction will ultimately be repulsive as the electron clouds of atoms and molecules begin to overlap. The total interaction is thus a combination of a short-range repulsion and a relatively long-range attraction.

Except for in highly polar materials such as water, London dispersion interactions give the largest contribution (70–100%) to the van der Waals attraction. The interaction energy of the

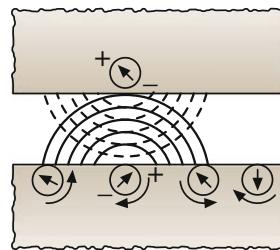


Fig. 9.3 Schematic representation of van der Waals interaction (dipole-induced dipole interaction)

van der Waals force between atoms or molecules depends on the separation r as

$$E(D) = -\frac{C_{vdW}}{r^6}, \quad (9.7)$$

where the constant C_{vdW} depends on the dipole moments and polarizabilities of the molecules. At large separations (>10 nm), the London interaction is weakened by a randomizing effect caused by the rapid fluctuations. That is, the induced temporary dipole moment of one molecule may have changed during the time needed for the transmission of the electromagnetic wave (photon) generated by its fluctuating charge density to another molecule and the return of the photon generated by the induced fluctuation in this second molecule. This phenomenon is called retardation and causes the interaction energy to decay as r^{-7} at large separations (Casimir and Polder 1948).

Dispersion interactions are to a first approximation additive, and their contribution to the interaction energy between two macroscopic bodies (such as colloidal particles) across vacuum can be found by summing the pairwise interactions (Hamaker 1937). The interaction is generally described in terms of the Hamaker constant, A_H . Another approach is to treat the interacting bodies and an intervening medium as continuous phases and determine the strength of the interaction from bulk dielectric properties of the materials (Lifshitz 1956; Dzyaloshinskii et al. 1961). Unlike the pairwise summation, this method takes into account the screening of the interactions between molecules inside the bodies by the molecules closer to the surfaces and the

effects of the intervening medium. For the interaction between material 1 and material 3 across material 2, the nonretarded Hamaker constant given by the Lifshitz theory is approximately (Israelachvili 2011)

$$A_{H,123} = A_{H,v=0} + A_{H,v>0}$$

$$\approx \frac{3}{4} k_B T \left(\frac{\varepsilon_1 - \varepsilon_2}{\varepsilon_1 + \varepsilon_2} \right) \left(\frac{\varepsilon_3 - \varepsilon_2}{\varepsilon_3 + \varepsilon_2} \right) + \frac{3h\nu_e}{8\sqrt{2}} \times \frac{(n_1^2 - n_2^2)(n_3^2 - n_2^2)}{\sqrt{(n_1^2 + n_2^2)} \sqrt{(n_3^2 + n_2^2)} \left(\sqrt{(n_1^2 + n_2^2)} + \sqrt{(n_3^2 + n_2^2)} \right)}, \quad (9.8)$$

where the first term ($v = 0$) represents the permanent dipole and dipole-induced dipole interactions and the second ($v > 0$) the London (dispersion) interaction. ε_i and n_i are the static dielectric constants and refractive indexes of the materials, respectively. ν_e is the frequency of the lowest electron transition (around $3 \times 10^{15} \text{ s}^{-1}$). Either one of the materials 1, 2, or 3 in (9.8) can be vacuum or air ($\varepsilon = n = 1$). A_H is typically $10^{-20}\text{--}10^{-19} \text{ J}$ (the higher values are found for metals) for interactions between solids and liquids across vacuum or air.

The interaction energy between two macroscopic bodies is dependent on the geometry and is always attractive between two bodies of the same material [A_H positive, see (9.8)]. The van der Waals interaction energy and force laws ($F = -dE(D)/dD$) for some common geometries are given in Table 9.2. Because of the retardation effect, the equations in Table 9.2 will lead to an overestimation of the dispersion force at large separations. It is, however, apparent that the interaction energy between macroscopic bodies decays more slowly with separation (i.e., has a longer range) than between two molecules. For inert nonpolar surfaces, e.g., consisting of hydrocarbons or van der Waals solids and liquids, the Lifshitz theory has been found to apply even at molecular contact, where it can be used to predict the surface energies (surface tensions) of such solids and liquids. For example, for hydrocarbon surfaces, $A_H = 5 \times 10^{-20} \text{ J}$. Inserting this value into the equation for two flat surfaces (Table 9.2) and using a cut-off distance of $D_0 \approx 0.165 \text{ nm}$ as an effective separation when

the surfaces are in contact (Israelachvili 2011), we obtain for the surface energy γ (which is defined as half the interaction energy)

$$\gamma = \frac{E}{2} = \frac{A_H}{24\pi D_0^2} \approx 24 \text{ mJ m}^{-2}, \quad (9.9)$$

a value that is typical for hydrocarbon solids and liquids (Fox and Zisman 1952). If the adhesion force is measured between two crossed-cylindrical surfaces of $R = 1 \text{ cm}$ (a geometry equivalent to a sphere with $R = 1 \text{ cm}$ interacting with a flat surface, cf. Table 9.2) using an SFA, we expect the adhesion force to be (Table 9.2) $F = A_H R / (6D_0^2) = 4\pi R \gamma \approx 3.0 \text{ mN}$. Using a spring constant of $k_s = 100 \text{ N m}^{-1}$, such an adhesive force will cause the two surfaces to jump apart by $\Delta D = F/k_s = 30 \mu\text{m}$, which can be accurately measured. (For elastic bodies that deform in adhesive contact, R changes during the interaction and the measured adhesion force is 25% lower, see Sect. 9.5.2). Surface energies of solids can thus be directly measured with the SFA. The measured values are in good agreement with calculated values based on the known surface energies γ of the materials, and for non-polar low-energy solids they are well accounted for by the Lifshitz theory (Israelachvili 2011).

9.4 Normal Forces Between Surfaces in Liquids

9.4.1 Van der Waals Forces in Liquids

The dispersion interaction in a medium will be significantly lower than in vacuum, since the attractive interaction between two solute molecules in a medium (solvent) involves displacement and reorientation of the nearest-neighbor solvent molecules. Even though the surrounding medium may change the dipole moment and polarizability from that in vacuum, the interaction between two identical molecules remains attractive in a binary mixture. The extension of the interactions to the case of two macroscopic bodies is the same as described in Sect. 9.3.1. Typically, the Hamaker constants for interactions in a

medium are an order of magnitude lower than in vacuum. Between macroscopic surfaces in liquids, van der Waals forces become important at distances below 10–15 nm and may at these distances start to dominate interactions of different origin that have been observed at larger separations.

Figure 9.4 shows the measured van der Waals forces between two crossed-cylindrical mica surfaces in water and various salt solutions. Good agreement is obtained between experiment and theory. At larger surface separations, above about 5 nm, the measured forces fall off more rapidly than D^{-2} . This retardation effect (Sect. 9.3.1) is also predicted by the Lifshitz theory and is due to the time needed for propagation of the induced dipole moments over large distances.

From Fig. 9.4 we may conclude that, at separations above about 2 nm, or 8 molecular diameters of water, the *continuum* Lifshitz theory is valid. This would mean that water films as thin as 2 nm may be expected to have bulk-like properties, at least as far as their interaction forces are concerned. Similar results have been obtained with other liquids, where in general continuum properties are manifested, both as regards their interactions and other properties such as viscosity, at a film thickness larger than five or ten molecular diameters. In the absence of a solvent (in vacuum), the agreement of measured van der Waals forces with the continuum Lifshitz theory is generally good at all separations down to molecular contact ($D = D_0$).

Van der Waals interactions in a system of three or more different materials (9.8) can be attractive or repulsive, depending on their dielectric properties. Numerous experimental studies show the attractive van der Waals forces in various systems (Israelachvili 2011), and repulsive van der Waals forces have also been measured directly (Meurk et al. 1997). A practical consequence of the repulsive interaction obtained across a medium with intermediate dielectric properties is that the van der Waals forces will give rise to preferential, nonspecific adsorption of molecules with an intermediate

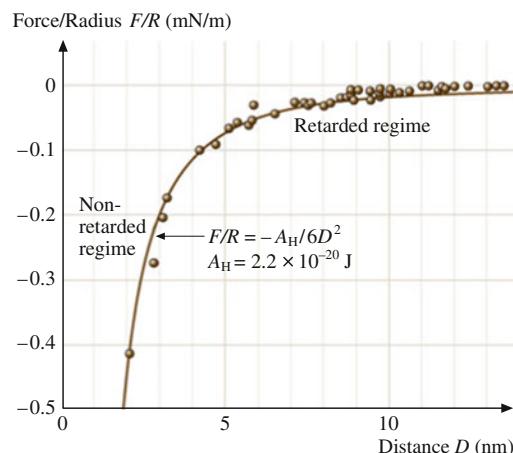


Fig. 9.4 Attractive van der Waals force F between two curved mica surfaces of radius $R \approx 1$ cm measured in water and various aqueous electrolyte solutions. The electrostatic interaction has been subtracted from the total measured force. The measured nonretarded Hamaker constant is $A_H = 2.2 \times 10^{-20}$ J. Retardation effects are apparent at distances larger than 5 nm, as expected theoretically (after Israelachvili 2011), © 1991, with permission from Elsevier Science)

dielectric constant. It is possible to diminish the attractive interaction between dispersed colloidal particles by adsorption of a thin layer of material with dielectric properties close to those of the surrounding medium (matching of refractive index), or by adsorption of a polymer that gives a steric repulsive force that keeps the particles separated at a distance where the magnitude of the van der Waals attraction is negligible. Thermal motion will then keep the particles dispersed.

9.4.2 Electrostatic Forces

Most surfaces in contact with a highly polar liquid (such as water) acquire a surface charge, either by dissociation of ions from the surface into the solution or by preferential adsorption of certain ions from the solution. The surface charge is balanced by a layer of oppositely charged ions (counterions) in the solution at some small distance from the surface (Israelachvili 2011). In dilute solution, this distance is the Debye length

κ^{-1} which is purely a property of the electrolyte solution. The Debye length falls with increasing ionic strength (i.e., with the molar concentration M_i and valency z_i) of the ions in solution

$$\kappa^{-1} = \left(\frac{\varepsilon \varepsilon_0 k_B T}{e^2 N_A \sum_i z_i^2 M_i} \right)^{1/2}, \quad (9.10)$$

where e is the electronic charge. For example, for 1:1 electrolytes at 25 °C, $\kappa^{-1} = 0.304 \text{ nm}/\sqrt{M_{1:1}}$, where M_i is given in M (mol dm⁻³). κ^{-1} is thus about 10 nm in a 1 mM NaCl solution and 0.3 nm in a 1 M solution. In totally pure water at pH 7, where $M_i = 10^{-7} \text{ M}$, κ^{-1} is 960 nm, or about 1 μm. The Debye length also relates the surface charge density σ of a surface to the electrostatic surface potential ψ_0 via the Grahame equation, which for 1:1 electrolytes can be expressed as

$$\sigma = \sqrt{8 \varepsilon \varepsilon_0 k_B T} \sinh \left(\frac{e \psi_0}{2 k_B T} \right) \times \sqrt{M_{1:1} N_A}. \quad (9.11)$$

Since the Debye length is a measure of the thickness of the diffuse atmosphere of counterions near a charged surface, it also determines the range of the electrostatic double-layer interaction between two charged surfaces. The electrostatic double-layer interaction is an entropic effect that arises upon decreasing the thickness of the liquid film containing the dissolved ions. Because of the attractive force between the dissolved ions and opposite charges on the surfaces, the ions stay between the surfaces, but an osmotic repulsion arises as their concentration increases. The long-range electrostatic interaction energy at large separations (weak overlap) between two similarly charged molecules or surfaces is typically repulsive and is roughly an exponentially decaying function of D

$$E(D) \approx + C_{ES} e^{-\kappa D}, \quad (9.12)$$

where C_{ES} is a constant that depends on the geometry of the interacting surfaces, on their surface charge density, and the solution conditions (Table 9.3). We see that the Debye length is

the decay length of the interaction energy between two surfaces (and of the mean potential away from one surface). C_{ES} can be determined by solving the so-called Poisson–Boltzmann equation or by using other theories (Verwey and Overbeek 1948; Sader et al. 1995). The equations in Table 9.3 are expressed in terms of a constant Z defined as

$$Z = 64 \pi \varepsilon \varepsilon_0 \left(\frac{k_B T}{e} \right)^2 \tan h^2 \left(\frac{ze \psi_0}{4 k_B T} \right), \quad (9.13)$$

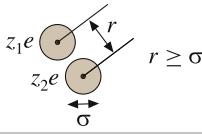
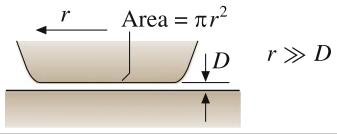
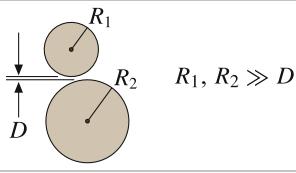
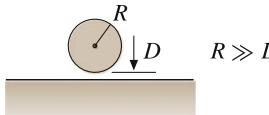
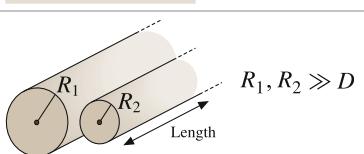
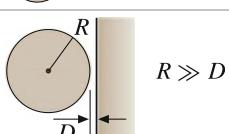
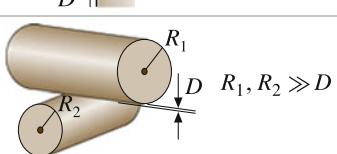
which depends only on the properties of the surfaces.

The above approximate expressions are accurate only for surface separations larger than about one Debye length. At smaller separations one must use numerical solutions of the Poisson–Boltzmann equation to obtain the exact interaction potential, for which there are no simple expressions. In the limit of small D , it can be shown that the interaction energy depends on whether the surfaces remain at constant potential ψ_0 (as assumed in the above equations) or at constant charge σ (when the repulsion exceeds that predicted by the above equations), or somewhere between these limits. In the constant charge limit the total number of counterions in the compressed film does not change as D is decreased, whereas at constant potential, the concentration of counterions is constant. The limiting pressure (or force per unit area) at constant charge is the osmotic pressure of the confined ions

$$F = k_B T \times \text{ion number density} = \frac{2 \sigma k_B T}{zeD}, \quad \text{for } D \ll \kappa^{-1}. \quad (9.14)$$

That is, as $D \rightarrow 0$ the double-layer pressure at constant surface charge becomes infinitely repulsive and independent of the salt concentration (at constant potential the force instead becomes a constant at small D). However, at small separations, the van der Waals attraction (which goes as D^{-2} between two spheres or as D^{-3} between two planar surfaces, see Table 9.2)

Table 9.3 Electrical double-layer interaction energy $E(D)$ and force ($F = -dE/dD$) between macroscopic bodies (Leckband and Israelachvili 2001)

Geometry of bodies with surfaces D apart ($D \ll R$)		Electric double-layer interaction	
		Energy E	Force F
Two ions or small molecules		$\frac{+z_1z_2e^2}{4\pi\epsilon\epsilon_0 r} \frac{e^{-\kappa(r-\sigma)}}{(1+\kappa\sigma)}$	$\frac{+z_1z_2e^2(1-\kappa r)}{4\pi\epsilon\epsilon_0 r} \frac{e^{-\kappa(r-\sigma)}}{(1+\kappa\sigma)}$
Two flat surfaces (per unit area)		$(\kappa/2\pi)Ze^{-\kappa D}$	$\left(\frac{\kappa^2}{2\pi}\right)Ze^{-\kappa D}$
Two spheres or macromolecules of radii R_1 and R_2		$\left(\frac{R_1R_2}{R_1+R_2}\right)Ze^{-\kappa D}$	$\kappa\left(\frac{R_1R_2}{R_1+R_2}\right)Ze^{-\kappa D}$
Sphere or macromolecule of radius R near a flat surface		$RZe^{-\kappa D}$	$\kappa RZe^{-\kappa D}$
Two parallel cylinders or rods of radii R_1 and R_2 (per unit length)		$\frac{\kappa^{1/2}}{\sqrt{2\pi}} \left(\frac{R_1R_2}{R_1+R_2}\right)^{1/2} Z^{-\kappa D}$	$\frac{\kappa^{3/2}}{\sqrt{2\pi}} \left(\frac{R_1R_2}{R_1+R_2}\right)^{1/2} Ze^{-\kappa D}$
Cylinder of radius R near a flat surface (per unit length)		$k^{1/2} \sqrt{\frac{R}{2\pi}} Ze^{-\kappa D}$	$k^{3/2} \sqrt{\frac{R}{2\pi}} Ze^{-\kappa D}$
Two cylinders or filaments of radii R_1 and R_2 crossed at 90°		$\sqrt{R_1R_2} e^{-\kappa D}$	$\kappa\sqrt{R_1R_2} Ze^{-\kappa D}$

wins out over the double-layer repulsion, unless some other short range interaction becomes dominant (Sect. 9.4.4). This is the theoretical prediction that forms the basis of the so-called Derjaguin–Landau–Verwey–Overbeek (DLVO) theory (Verwey and Overbeek 1948; Derjaguin and Landau 1941), illustrated in Fig. 9.5.

Because of the different distance dependence of the van der Waals and electrostatic

interactions, the total force law, as described by the DLVO theory, can show several minima and maxima. Typically, the depth of the outer (secondary) minimum is a few $k_B T$, which is enough to cause reversible flocculation of particles from an aqueous dispersion. If the force barrier between the secondary and primary minimum is lowered, for example, by increasing the electrolyte concentration, particles can be irreversibly

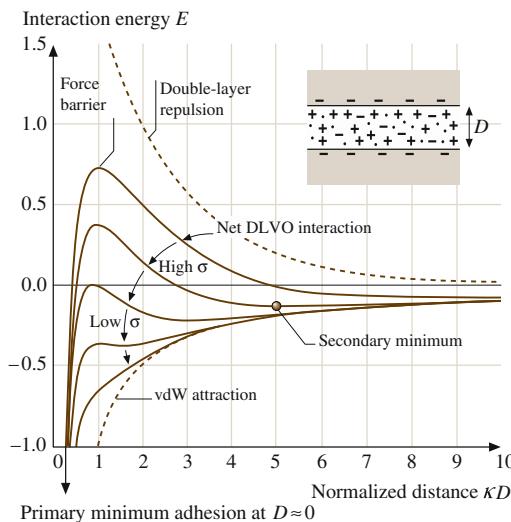


Fig. 9.5 Schematic plots of the DLVO interaction potential energy E between two flat, charged surfaces [or, according to the Derjaguin approximation, (9.3), the force F between two curved surfaces] as a function of the surface separation normalized by the Debye length κ^{-1} . The van der Waals attraction (inverse power-law dependence on D) together with the repulsive electrostatic double-layer force (roughly exponential) at different surface charge σ (or potential, see (9.11)) determine the net interaction potential in aqueous electrolyte solution (after (Leckband and Israelachvili 2001) with permission)

coagulated in the primary minimum. In practice, other forces (described in the following sections) often appear at very small separations, so that the full force law between two surfaces or colloidal particles in solution can be more complex than might be expected from the DLVO theory.

9.4.3 Solvation and Structural Forces

When a liquid is confined within a restricted space, for example, a very thin film between two surfaces, it ceases to behave as a structureless continuum. At small surface separations (below about ten molecular diameters), the van der Waals force between two surfaces or even two solute molecules in a liquid (solvent) is no longer a smoothly varying attraction. Instead, there arises an additional solvation force that generally oscillates between attraction and repulsion with distance, with a periodicity equal to some mean

dimension σ of the liquid molecules (Horn and Israelachvili 1981). Figure 9.6a shows the force law between two smooth mica surfaces across the hydrocarbon liquid tetradecane, whose inert, chainlike molecules have a width of $\sigma \approx 0.4$ nm. The short-range oscillatory force law is related to the density distribution function and potential of mean force characteristic of intermolecular interactions in liquids. These forces arise from the confining effects that the two surfaces have on liquid molecules, forcing them to order into quasi-discrete layers. Such layers are energetically or entropically favored and correspond to the minima in the free energy, whereas fractional layers are disfavored (energy maxima). This effect is quite general and arises in all simple liquids when they are confined between two smooth, rigid surfaces, both flat and curved.

Oscillatory forces are now well understood theoretically, at least for simple liquids, and a number of theoretical studies and computer simulations of various confined liquids (including water) that interact via some form of Lennard-Jones potential have invariably led to an oscillatory solvation force at surface separations below a few molecular diameters (Snook and van Megen 1980; Kjellander and Marcelja 1985; Henderson and Lozadacassou 1986; Gao et al. 1997). In a first approximation, the oscillatory force law may be described by an exponentially decaying cosine function of the form

$$E \approx E_0 \cos\left(\frac{2\pi D}{\sigma}\right) e^{-D/\sigma}, \quad (9.15)$$

where both theory and experiments show that the oscillatory period and the characteristic decay length of the envelope are close to σ .

Once the solvation zones of the two surfaces overlap, the mean liquid density in the gap is no longer the same as in the bulk liquid. Since the van der Waals interaction depends on the optical properties of the liquid, which in turn depends on the density, the van der Waals and the oscillatory solvation forces are not strictly additive. It is more correct to think of the solvation force as the van der Waals force at small separations with the molecular properties and density variations of the

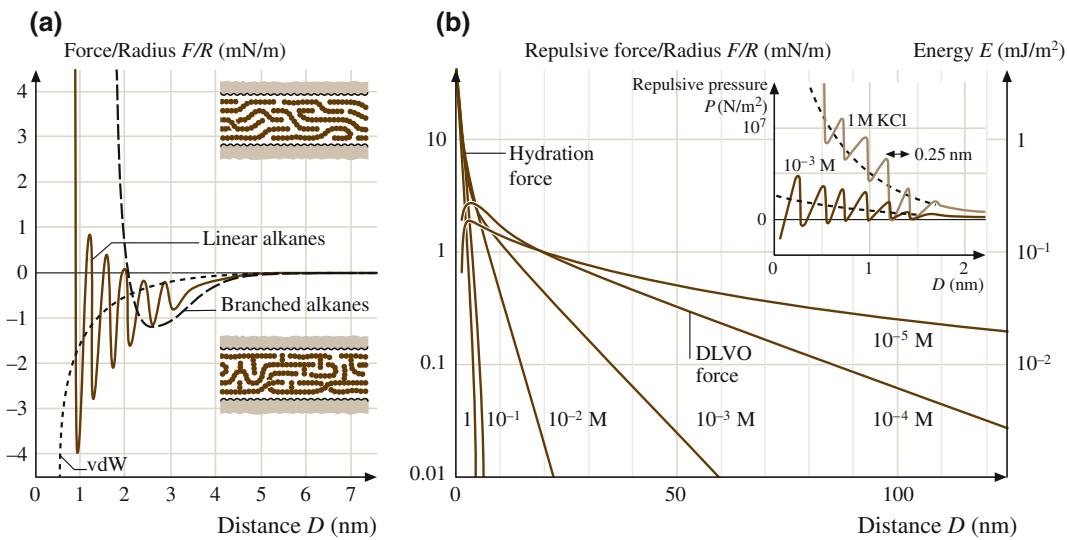


Fig. 9.6 **a** Solid curve forces measured between two mica surfaces across saturated linear chain alkanes such as *n*-tetradecane and *n*-hexadecane (Gee and Israelachvili 1990; Christenson et al. 1987). The 0.4 nm periodicity of the oscillations indicates that the molecules are preferentially oriented parallel to the surfaces, as shown schematically in the upper insert. The theoretical continuum van der Waals attraction is shown as a dotted curve. Dashed curve: smooth, nonoscillatory force law exhibited by irregularly shaped alkanes (such as 2-methyloctadecane) that cannot order into well-defined layers (lower insert) (after (Gee and Israelachvili 1990; Israelachvili et al. 1989)). Similar nonoscillatory forces are also observed between rough surfaces, even when these interact across a saturated linear chain liquid. This is because the irregularly shaped surfaces (rather than the liquid) now prevent the liquid molecules from ordering in the gap. **b** Forces measured between charged mica surfaces in KCl solutions of varying concentrations (Pashley 1981). In dilute solutions (10^{-5} and 10^{-4} M), the measured forces are excellently described by the DLVO theory, based on exact solutions to the nonlinear Poisson–Boltzmann equation for the electrostatic forces and the Lifshitz theory for the van der Waals forces (using a Hamaker constant of $A_H = 2.2 \times 10^{-20}$ J). At higher concentrations, as more hydrated K^+ cations adsorb onto the negatively charged surfaces, an additional hydration force appears superimposed on the DLVO interaction at distances below 3–4 nm. This force has both an oscillatory and a monotonic component. Insert short-range hydration forces between mica surfaces shown as pressure versus distance. The lower and upper curves show surfaces 40 and 95% saturated with K^+ ions. At larger separations, the forces are in good agreement with the DLVO theory (after (Israelachvili 2011), © 1991, with permission from Elsevier Science)

medium taken into account. It is also important to appreciate that solvation forces do not arise simply because liquid molecules tend to structure into semiordered layers at surfaces. They arise because of the disruption or *change* of this ordering during the approach of a second surface. The two effects are related; the greater the tendency toward structuring at an isolated surface the greater the solvation force between two such surfaces, but there is a real distinction between the two phenomena that should be borne in mind.

Oscillatory forces lead to different adhesion values depending on the energy minimum from which two surfaces are being separated. For an

interaction energy described by (9.15), quantized adhesion energies will be E_0 at $D = 0$ (primary minimum), E_0/e at $D = \sigma$, E_0/e^2 at $D = 2\sigma$, etc. E_0 can be thought of as a depletion force (Sect. 9.4.5) that is approximately given by the osmotic limit $E_0 \approx -k_B T/\sigma^2$, which can exceed the contribution to the adhesion energy in contact from the van der Waals forces (at $D_0 \approx 0.15$ –0.20 nm, as discussed in Sect. 9.3.1, keeping in mind that the Lifshitz theory fails to describe the force law at *intermediate distances*).

It is easy to understand how oscillatory forces arise between two flat, plane parallel surfaces. Between two curved surfaces, e.g., two spheres,

one might imagine the molecular ordering and oscillatory forces to be smeared out in the same way that they are smeared out between two randomly rough surfaces (Sect. 9.5.3); however, this is not the case. Ordering can occur as long as the curvature or roughness is itself regular or uniform, i.e., not random. This is due to the Derjaguin approximation (9.3). If the energy between two flat surfaces is given by a decaying oscillatory function (for example, a cosine function as in (9.15)), then the force (and energy) between two curved surfaces will also be an oscillatory function of distance with some phase shift. Likewise, two surfaces with regularly curved regions will also retain their oscillatory force profile, albeit modified, as long as the corrugations are truly regular, i.e., periodic. On the other hand, surface roughness, even on the nanometer scale, can smear out oscillations if the roughness is random and the confined molecules are smaller than the size of the surface asperities (Frink and van Swol 1998; Gao et al. 2000). If an organic liquid contains small amounts of water, the expected oscillatory force can be replaced by a strongly attractive capillary force (Sect. 9.5.1).

9.4.4 Hydration and Hydrophobic Forces

There are many aqueous systems in which the DLVO theory fails and where there is an additional short-range force that is not oscillatory but monotonic. Between hydrophilic surfaces this force is exponentially repulsive and is commonly referred to as the hydration, or structural, force. The origin and nature of this force has long been controversial, especially in the colloidal and biological literature. Repulsive hydration forces are believed to arise from strongly hydrogen-bonding surface groups, such as hydrated ions or hydroxyl ($-OH$) groups, which modify the hydrogen-bonding network of liquid water adjacent to them. Because this network is quite extensive in range (Stanley and Teixeira 1980), the resulting interaction force is also of relatively long range.

Repulsive hydration forces were first extensively studied between clay surfaces (van Olphen 1977). More recently, they have been measured in detail between mica and silica surfaces (Pashley 1981; Pashley 1982; Horn et al. 1989), and even biological membranes (Schrader et al. 2015; Shrestha and Banquy 2016), where they have been found to decay exponentially with decay lengths of about 1 nm. Their effective range is 3–5 nm, which is about twice the range of the oscillatory solvation force in water. Empirically, the hydration repulsion between two hydrophilic surfaces appears to follow the simple equation

$$E = E_0 e^{-D/\lambda_0}, \quad (9.16)$$

where $\lambda_0 \approx 0.6\text{--}1.1$ nm for 1:1 electrolytes and $E_0 = 3\text{--}30$ mJ m⁻² depending on the hydration (hydrophilicity) of the surfaces, higher E_0 values generally being associated with lower λ_0 values.

The interactions between molecularly smooth mica surfaces in dilute electrolyte solutions obey the DLVO theory (Fig. 9.6b). However, at higher salt concentrations, specific to each electrolyte, hydrated cations bind to the negatively charged surfaces and give rise to a repulsive hydration force (Pashley 1981, 1982). This is believed to be due to the energy needed to dehydrate the bound cations, which presumably retain some of their water of hydration on binding. This conclusion was arrived at after noting that the strength and range of the hydration forces increase with the known hydration numbers of the electrolyte cations in the order: $Mg^{2+} > Ca^{2+} > Li^+ \sim Na^+ > K^+ > Cs^+$. Similar trends are observed with other negatively charged colloidal surfaces.

While the hydration force between two mica surfaces is overall repulsive below a distance of 4 nm, it is not always monotonic below about 1.5 nm but exhibits oscillations of mean periodicity of 0.25 ± 0.03 nm, roughly equal to the diameter of the water molecule. This is shown in the insert in Fig. 9.6b, where we may note that the first three minima at $D = 0$, 0.28, and 0.56 nm occur at negative energies, a result that rationalizes observations on certain colloidal systems.

Surface force experiments have shown that hydration forces can be modified or “regulated” by exchanging ions of different hydration on surfaces, an effect that has important practical applications in controlling the stability of colloidal dispersions. It has long been known that colloidal particles can be precipitated (coagulated or flocculated) by increasing the electrolyte concentration, an effect that was traditionally attributed to the reduced screening of the electrostatic double-layer repulsion between the particles due to the reduced Debye length. However, there are many examples where colloids are stabilized at high salt concentrations, not at low concentrations. This effect is now recognized as being due to the increased hydration repulsion experienced by certain surfaces when they bind highly hydrated ions at higher salt concentrations. Hydration regulation of adhesion and interparticle forces is an important practical method for controlling various processes such as clay swelling (Viani et al. 1983), ceramic processing and rheology (Horn 1990), material fracture (Horn 1990), and colloidal particle and bubble coalescence (Lessard and Ziemienski 1971).

Water appears to be unique in having a solvation (hydration) force that exhibits both a monotonic and an oscillatory component. Between hydrophilic surfaces the monotonic component is repulsive (Fig. 9.6b), but between hydrophobic surfaces it is attractive and the final adhesion is much greater than expected from the Lifshitz theory.

A hydrophobic surface is one that is inert to water in the sense that it cannot bind to water molecules via ionic or hydrogen bonds. Hydrocarbons and fluorocarbons are hydrophobic, as is air, and the strongly attractive hydrophobic force has many important manifestations and consequences such as the low solubility or miscibility of water and oil molecules, micellization, protein folding, strong adhesion and rapid coagulation of hydrophobic surfaces, nonwetting of water on hydrophobic surfaces, and hydrophobic particle attachment to rising air bubbles (the basic principle of froth flotation).

In recent years, there has been a steady accumulation of experimental data on the force laws between various hydrophobic surfaces in aqueous solution (Israelachvili and Pashley 1982; Pashley et al. 1985; Claesson et al. 1986; Christenson et al. 1990; Tsao et al. 1993; Rabinovich and Yoon 1994; Craig et al. 1998; Christenson and Claesson 2001; Lin et al. 2005; Meyer et al. 2005a,b; Donaldson et al. 2015). These studies have found that the force law between two macroscopic hydrophobic surfaces is of surprisingly long range, decaying exponentially with a characteristic decay length of 1–2 nm in the separation range of 0–10 nm, and then more gradually further out. The hydrophobic force can be far stronger than the van der Waals attraction, especially between hydrocarbon surfaces in water, for which the Hamaker constant is quite small.

For two surfaces in water the purely hydrophobic interaction energy (ignoring DLVO and oscillatory forces) in the range 0–10 nm is given by

$$E = -2\gamma e^{-D/\lambda_0}, \quad (9.17)$$

where typically $\lambda_0 = 1\text{--}2$ nm, and $\gamma = 10\text{--}50 \text{ mJ m}^{-2}$. The higher value corresponds to the interfacial energy of a pure hydrocarbon–water interface.

At a separation below 10 nm, the hydrophobic force appears to be insensitive or only weakly sensitive to changes in the type and concentration of electrolyte ions in the solution. The absence of a screening effect by ions attests to the nonelectrostatic origin of this interaction. In contrast, some experiments have shown that, at separations greater than 10 nm, the attraction does depend on the intervening electrolyte, and that in dilute solutions, or solutions containing divalent ions, it can continue to exceed the van der Waals attraction out to separations of 80 nm (Christenson et al. 1989, 1990). Recent research suggests that the interactions at very long range might not be a hydrophobic force since they are influenced by the presence of dissolved gas in the solution (Lin et al. 2005; Meyer et al. 2005a,b),

the stability of the hydrophobic surface (Meyer et al. 2005a,b), and, on some types of surfaces, bridging submicroscopic bubbles (Parker et al. 1994; Attard 2003).

The long-range nature of the hydrophobic interaction has a number of important consequences. It accounts for the rapid coagulation of hydrophobic particles in water and may also account for the rapid folding of proteins. It also explains the ease with which water films rupture on hydrophobic surfaces. In this case, the van der Waals force across the water film is repulsive and therefore favors wetting, but this is more than offset by the attractive hydrophobic interaction acting between the two hydrophobic phases across water. Hydrophobic forces are increasingly being implicated in the adhesion and fusion of biological membranes and cells. It is known that both osmotic and electric-field stresses enhance membrane fusion, an effect that may be due to the concomitant increase in the hydrophobic area exposed between two adjacent surfaces (Donaldson et al. 2011).

From the previous discussion we can infer that even though highly related to each other (Donaldson et al. 2015), hydration and hydrophobic forces are not of a simple nature. These interactions are probably the most important, yet the least understood of all the forces in aqueous solutions. The unusual properties of water and the nature of the surfaces (including their homogeneity and stability) appear to be equally important (Donaldson et al. 2011).

How do these exponentially decaying repulsive or attractive forces arise? Theoretical work and computer simulations (Kjellander and Marcelja 1985; Henderson and Lozadacassou 1986) suggest that the solvation forces in water should be purely oscillatory, whereas other theoretical studies (Marcelja et al. 1977; Jonsson and Wennerstrom 1983; Luzar et al. 1987) suggest a monotonically exponential repulsion or attraction, possibly superimposed on an oscillatory force. The latter is consistent with experimental findings, as shown in the inset to Fig. 9.6b, where it appears that the oscillatory force is simply additive with the monotonic hydration

and DLVO forces, suggesting that these arise from essentially different mechanisms.

Like the repulsive hydration force, the origin of the hydrophobic force is still unknown. A Monte Carlo simulation (Luzar et al. 1987) of the interaction between two hydrophobic surfaces across water at separations below 1.5 nm showed a decaying oscillatory force superimposed on a monotonically attractive curve. In more recent computational and experimental work (Leung and Luzar 2000; Mao et al. 2004), it has been suggested that hydrophobic surfaces generate a depleted region of water around them, and that a long-range attractive force due to depletion arises between two such surfaces. Such a difference in density might also cause boundary slip of water at hydrophobic surfaces (Zhu and Granick 2001).

It is questionable whether the hydration or hydrophobic force should be viewed as an ordinary type of solvation or structural force that reflects the packing of water molecules. The energy (or entropy) associated with the hydrogen-bonding network, which extends over a much larger region of space than the molecular correlations, is probably at the root of the long-range interactions of water. The situation in water appears to be governed by much more than the molecular packing effects that dominate the interactions in simpler liquids.

9.4.5 Polymer-Mediated Forces

Polymers or macromolecules are chainlike molecules consisting of many identical segments (monomers or repeating units) held together by covalent bonds. The size of a polymer coil in solution or in the melt is determined by a balance between van der Waals attraction (and hydrogen bonding, if present) between polymer segments, and the entropy of mixing, which causes the polymer coil to expand. In polymer melts above the glass transition temperature, and at certain conditions in solution, the attraction between polymer segments is exactly balanced by the entropy effect. The polymer solution will then

behave virtually ideally, and the density distribution of segments in the coil is Gaussian. This is called the theta (θ) condition, and it occurs at the theta or Flory temperature for a particular combination of polymer and solvent or solvent mixture. At lower temperatures (in a poor or bad solvent), the polymer–polymer interactions dominate over the entropic, and the coil will shrink or precipitate. At higher temperatures (good solvent conditions), the polymer coil will be expanded.

High-molecular-weight polymers form large coils, which significantly affect the properties of a solution even when the total mass of polymer is very low. The radius of the polymer coil is proportional to the segment length a and the number of segments n . At theta conditions, the hydrodynamic radius of the polymer coil (the root-mean-square separation of the ends of one polymer chain) is theoretically given by $R_h = a n^{1/2}$, and the unperturbed radius of gyration (the average root-mean-square distance of a segment from the center of mass of the molecule) is $R_g = a(n/6)^{1/2}$. In a good solvent the perturbed size of the polymer coil, the Flory radius R_F , is proportional to $n^{3/5}$.

Polymers interact with surfaces mainly through van der Waals and electrostatic interactions. The physisorption of polymers containing only one type of segment is reversible and highly dynamic, but the rate of exchange of adsorbed chains with free chains in the solution is low, since the polymer remains bound to the surface as long as one segment along the chain is adsorbed. The adsorption energy per segment is on the order of $k_B T$. In a good solvent, the conformation of a polymer on a surface is very different from the coil conformation in bulk solution. Polymers adsorb in “trains”, separated by “loops” extending into solution and dangling “tails” (the ends of the chain). Compared to adsorption at lower temperatures, good solvent conditions favor more of the polymer chain being in the solvent, where it can attain its optimum conformation. As a result, the extension of the polymer is longer, even though the total amount of adsorbed polymer is lower. In a good solvent, the polymer chains can also be effectively

repelled from a surface, if the loss in conformational entropy close to the surface is not compensated for by a gain in enthalpy from adsorption of segments. In this case, there will be a layer of solution (thickness $\approx R_g$) close to the surfaces that is depleted of polymer.

The interaction forces between two surfaces across a polymer solution will depend on whether the polymer adsorbs onto the surfaces or is repelled from them, and also on whether the interaction occurs at “true” or “restricted” thermodynamic equilibrium. At true or full equilibrium, the polymer between the surfaces can equilibrate (exchange) with polymer in the bulk solution at all surface separations. Some theories (de Gennes 1982; Scheutjens and Fleer 1985) predict that, at full equilibrium, the polymer chains would move from the confined gap into the bulk solution where they could attain entropically more favorable conformations, and that a monotonic attraction at all distances would result from bridging and depletion interactions (which will be discussed below). Other theories suggest that the interaction at small separations would be ultimately repulsive, since some polymer chains would remain in the gap due to their attractive interactions with many sites on the surface (enthalpic)—more sites would be available to the remaining polymer chains if some others desorbed and diffused out from the gap (Evans 1989).

At restricted equilibrium, the polymer is kinetically trapped, and the adsorbed amount is thus constant as the surfaces are brought toward each other, but the chains can still rearrange on the surfaces and in the gap. Experimentally, the true equilibrium situation is very difficult to attain, and most experiments are done at restricted equilibrium conditions. Even the equilibration of conformations assumed in theoretical models for restricted equilibrium conditions can be so slow that this condition is difficult to reach experimentally.

In systems of adsorbing polymer, bridging of chains from one surface to the other can give rise to a long range attraction, since the bridging chains would gain conformational entropy if the surfaces were closer together. In poor solvents, both bridging and intersegment interactions

contribute to an attraction (Patel and Tirrell 1989). However, regardless of solvent and equilibrium conditions, a strong repulsion due to the osmotic interactions is seen at small surface separations in systems of adsorbing polymers at restricted equilibrium.

In systems containing high concentrations of nonadsorbing polymer, the difference in solute concentration in the bulk and between the surfaces at separations smaller than the approximate polymer coil diameter ($2R_g$, i.e., when the polymer has been squeezed out from the gap between the surfaces) may give rise to an attractive osmotic force (the depletion attraction) (Asakura and Oosawa 1958; Joanny et al. 1979; Feigin and Napper 1980; de Gennes 1981). In addition, if the polymer coils become initially compressed as the surfaces approach each other, this can give rise to a repulsion (depletion stabilization) at large separations (Feigin and Napper 1980). For a system of two cylindrical surfaces of radius R , the maximum depletion force F_{dep} is expected to occur when the surfaces are in contact and is given by multiplying the depletion (osmotic) pressure, $P_{\text{dep}} = \rho k_B T$, by the contact area πr^2 , where r is given by the chord theorem: $r^2 = (2R - R_g)R_g \approx 2RR_g$ (Israelachvili 2011)

$$\frac{F_{\text{dep}}}{R} = -2\pi R_g \rho k_B T, \quad (9.18)$$

where ρ is the number density of the polymer in the bulk solution.

If a part of the polymer (typically an end group) is different from the rest of the chain, this part may preferentially adsorb to the surface. End-adsorbed polymer is attached to the surface at only one point, and the extension of the chain is dependent on the grafting density, i.e., the average distance s between adsorbed end groups on the surface (Fig. 9.7). One distinguishes between different regions of increasing overlap of the chains (stretching) called pancake, mushroom, and brush regimes (de Gennes 1987). In the mushroom regime, where the coverage is sufficiently low that there is no overlap between neighboring chains, the thickness of the adsorbed

layer is proportional to $n^{1/2}$ (i.e., to R_g) at theta conditions and to $n^{3/5}$ in a good solvent.

Several models (de Gennes 1980, 1987; Alexander 1977; Milner et al. 1988; Zhulina et al. 1990) have been developed for the extension and interactions between two brushes (strongly stretched grafted chains). They are based on a balance between osmotic pressure within the brush layers (uncompressed and compressed) and the elastic energy of the chains and differ mainly in the assumptions of the segment density profile, which can be a step function or parabolic. At high coverage (in the brush regime), where the chains will avoid overlapping each other, the thickness of the layer is proportional to n .

Experimental work on both monodisperse (Watanabe and Tirrell 1993; Klein et al. 1994) and polydisperse (Ruths et al. 2000) systems at different solvent conditions has confirmed the expected range and magnitude of the repulsive interactions resulting from compression of densely packed grafted layers.

9.5 Adhesion and Capillary Forces

9.5.1 Capillary Forces

When considering the adhesion of two solid surfaces or particles in air or in a liquid, it is easy to overlook or underestimate the important role of capillary forces, i.e., forces arising from the Laplace pressure of curved menisci formed by condensation of a liquid between and around two adhering surfaces (Fig. 9.8).

The adhesion force between a nondeformable spherical particle of radius R and a flat surface in an inert atmosphere (Fig. 9.8a) is

$$F_{\text{ad}} = 4\pi R \gamma_{\text{SV}}. \quad (9.19)$$

But in an atmosphere containing a condensable vapor, the expression above is replaced by

$$F_{\text{ad}} = 4\pi R(\gamma_{\text{LV}} \cos \theta + \gamma_{\text{SL}}), \quad (9.20)$$

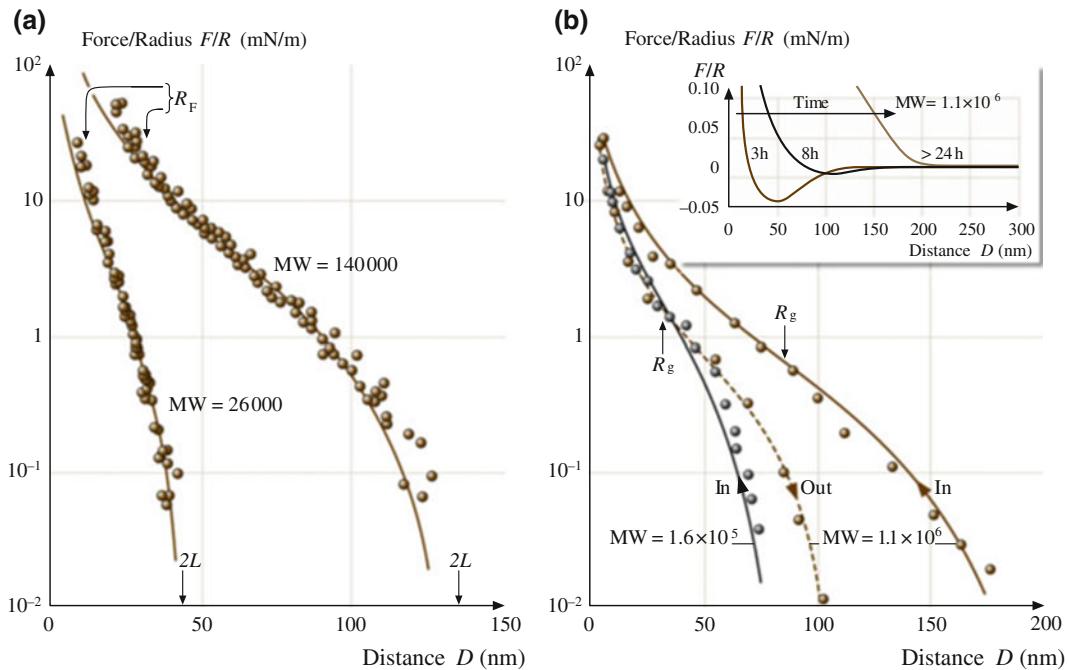


Fig. 9.7 Experimentally determined forces in systems of two interacting polymer layers: **a** polystyrene brush layers grafted via an adsorbing chain-end group onto mica surfaces in toluene (a good solvent for polystyrene). *Left curve:* MW = 26,000 g/mol, $R_F = 12$ nm. *Right curve:* MW = 140,000 g/mol, $R_F = 32$ nm. Both force curves were reversible on approach and separation. The solid curves are theoretical fits using the Alexander-de Gennes theory with the following measured parameters: spacing between attachment sites: $s = 8.5$ nm, brush thickness: $L = 22.5$ and 65 nm, respectively (adapted from (Taunton et al. 1990)). **b** Polyethylene oxide layers physisorbed onto mica from 150 $\mu\text{g}/\text{ml}$ solution in aqueous 0.1 M KNO_3 (a good solvent for polyethylene oxide). *Main figure:* equilibrium forces at full coverage after ~ 16 h adsorption time. *Left curve:* MW = 160,000 g/mol, $R_g = 32$ nm. *Right curve:* MW = 1.1×10^6 g/mol, $R_g = 86$ nm. Note the hysteresis (irreversibility) on approach and separation for this physisorbed layer, in contrast to the absence of hysteresis with grafted chains in case (a). The solid curves are based on a modified form of the Alexander-de Gennes theory. *Inset in (b):* evolution of the forces with the time allowed for the higher MW polymer to adsorb from solution. Note the gradual reduction in the attractive bridging component (adapted from Klein and Luckham 1982, 1984; Luckham and Klein 1990), after (Israelachvili 2011), © 1991, with permission from Elsevier Science)

where the first term is due to the Laplace pressure of the meniscus and the second is due to the direct adhesion of the two contacting solids within the liquid. Note that the above equation does not contain the radius of curvature r of the liquid meniscus (Fig. 9.8b). This is because for smaller r the Laplace pressure γ_{LV}/r increases, but the area over which it acts decreases by the same amount, so the two effects cancel out. Experiments with inert liquids, such as hydrocarbons, condensing between two mica surfaces indicate that (9.20) is valid for values of r as small as 1–2 nm, corresponding to vapor pressures as low as

40% of saturation (Fisher and Israelachvili 1981; Christenson 1988). Capillary condensation also occurs in binary liquid systems, e.g., when water dissolved in hydrocarbon liquids condenses around two contacting hydrophilic surfaces or when a vapor cavity forms in water around two hydrophobic surfaces. In the case of water condensing from vapor or from oil, it also appears that the bulk value of γ_{LV} is applicable for meniscus radii as small as 2 nm.

The capillary condensation of liquids, especially water, from vapor can have additional effects on the physical state of the contact zone. For

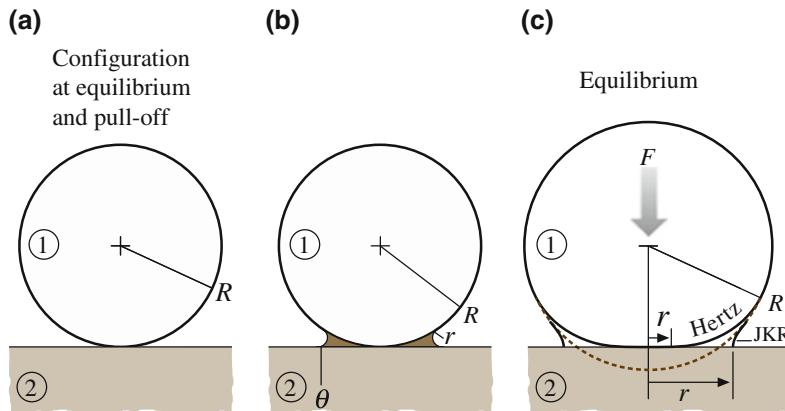


Fig. 9.8 Adhesion and capillary forces: **a** a nondeforming sphere on a rigid, flat surface in an inert atmosphere and **b** in a vapor that can capillary condense around the contact zone. At equilibrium, the concave radius r of the liquid meniscus is given by the Kelvin equation. For a concave meniscus to form, the contact angle θ has to be less than 90° . In the case of hydrophobic surfaces surrounded by water, a vapor cavity can form between the surfaces. As long as the surfaces are perfectly smooth, the contribution of the meniscus to the adhesion force is independent of r (after (Israelachvili 1995) with permission). **c** Elastically deformable sphere on a rigid flat surface in the absence (Hertz) and presence (JKR) of adhesion ((a) and (c) after (Israelachvili 2011), © 1991, with permission from Elsevier Science)

example, if the surfaces contain ions, these will diffuse and build up within the liquid bridge, thereby changing the chemical composition of the contact zone, as well as influencing the adhesion. In the case of surfaces covered with surfactant or polymer molecules (amphiphilic surfaces), the molecules can turn over on exposure to humid air, so that the surface nonpolar groups become replaced by polar groups, which renders the surfaces hydrophilic. When two such surfaces come into contact, water will condense around the contact zone and the adhesion force will also be affected—generally increasing well above the value expected for inert hydrophobic surfaces. It is apparent that adhesion in vapor or a solvent is often largely determined by capillary forces arising from the condensation of liquid that may be present only in very small quantities, e.g., 10–20% of saturation in the vapor, or 20 ppm in the solvent.

9.5.2 Adhesion Mechanics

Two bodies in contact deform as a result of surface forces and/or applied normal forces. For the simplest case of two interacting elastic spheres (a

model that is easily extended to an elastic sphere interacting with an undeformable surface, or vice versa) and in the absence of attractive surface forces, the vertical central displacement (compression) was derived by Hertz (1881) (Fig. 9.8 c). In this model, the displacement and the contact area are equal to zero when no external force (load) is applied, i.e., at the points of contact and of separation. The contact area A increases with normal force or load as $L^{2/3}$.

In systems where attractive surface forces are present between the surfaces, the deformations are more complicated. Modern theories of the adhesion mechanics of two contacting solid surfaces are based on the Johnson–Kendall–Roberts (JKR) theory (Johnson et al. 1971), or on the Derjaguin–Muller–Toporov (DMT) theory (Derjaguin et al. 1975; Muller et al. 1980; Muller et al. 1983). The JKR theory is applicable to easily deformable, large bodies with high surface energy, whereas the DMT theory better describes very small and hard bodies with low surface energy (Tabor 1977). The intermediate regime has also been described (Maugis 1992).

In the JKR theory, two spheres of radii R_1 and R_2 , bulk elastic modulus K , and surface energy γ

will flatten due to attractive surface forces when in contact at no external load. The contact area will increase under an external load L or normal force F , such that at mechanical equilibrium the radius of the contact area r is given by

$$r^3 = \frac{R}{K} \left[F + 6\pi R\gamma + \sqrt{12\pi R\gamma F + (6\pi R\gamma)^2} \right], \quad (9.21)$$

where $R = R_1 R_2 / (R_1 + R_2)$. In the absence of surface energy γ Eq. (9.21) is reduced to the expression for the radius of the contact area in the Hertz model. Another important result of the JKR theory gives the adhesion force or pull-off force

$$F_{ad} = -3\pi R\gamma_s, \quad (9.22)$$

where the surface energy γ_s is defined through $W = 2\gamma_s$, where W is the reversible work of

adhesion. Note that, according to the JKR theory, a finite elastic modulus K , while having an effect on the load-area curve, has no effect on the adhesion force, an interesting and unexpected result that has nevertheless been verified experimentally (Johnson et al. 1971; Horn et al. 1987; Mangipudi et al. 1994). Equations (9.21) and (9.22) provide the framework for analyzing results of adhesion measurements (Fig. 9.9) of contacting solids, known as contact mechanics, and for studying the effects of surface conditions and time on adhesion energy hysteresis (Sect. 9.5.4).

The JKR theory has been extended (Sridhar et al. 1997, 2004) to consider rigid or elastic substrates separated by thin compliant layers of very different elastic moduli, a situation commonly encountered in SFA and AFM experiments. The deformation of the system is then strongly dependent on the ratio of r to the

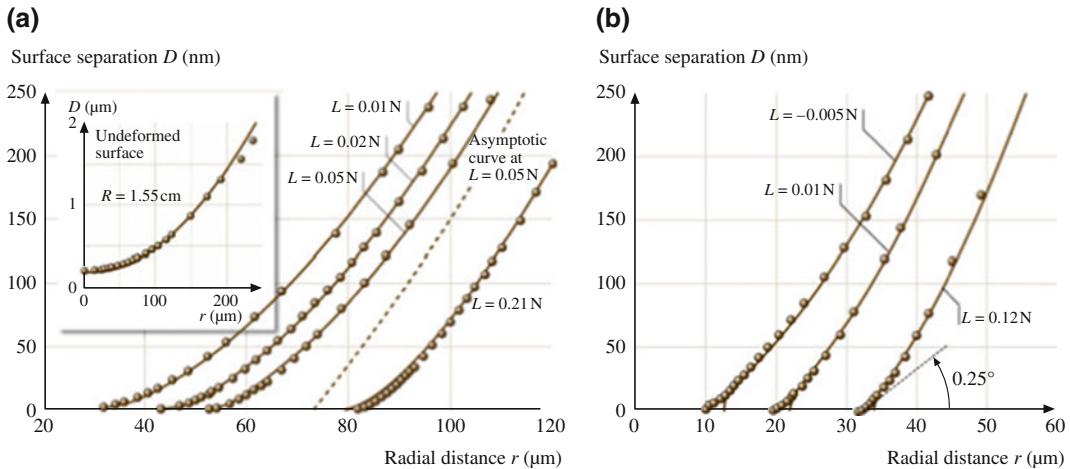


Fig. 9.9 Experimental data on contact mechanics for ideal Hertz and JKR contacts. **a** Measured profiles of surfaces in nonadhesive contact (circles) compared with Hertz profiles (continuous curves). The system was mica surfaces in a concentrated KCl solution in which they do not adhere. When not in contact, the surface shape is accurately described by a sphere of radius $R = 1.55$ cm (insert). The applied loads were 0.01, 0.02, 0.05, and 0.21 N. The last profile was measured in a different region of the surfaces where the local radius of curvature was 1.45 cm. The Hertz profiles correspond to central displacements of $\delta = 66.5, 124, 173$, and 441 nm. The dashed line shows the shape of the undeformed sphere corresponding to the curve at a load of 0.05 N; it fits the experimental points at larger distances (not shown). **b** Surface profiles measured with adhesive contact (mica surfaces adhering in dry nitrogen gas) at applied loads of $-0.005, 0.01$, and 0.12 N. The continuous lines are JKR profiles obtained by adjusting the central displacement in each case to get the best fit to points at larger distances. The values are $\delta = -4.2, 75.6$, and 256 nm. Note that the scales of this figure exaggerate the apparent angle at the junction of the surfaces. This angle, which is insensitive to load, is only about 0.25° (after (Horn et al. 1987), © 1987, with permission from Elsevier Science)

thickness of the confined layer. At small r (low L), the deformation occurs mostly in the thin confined layer, whereas at large r (large L), it occurs mainly in the substrates. Because of the changing distribution of traction across the contact, the adhesion force in a layered system is also modified from that of isotropic systems (9.22) so that it is no longer independent of the elastic moduli.

9.5.3 Effects of Surface Structure, Roughness, and Lattice Mismatch

In a contact between two rough surfaces, the real area of contact varies with the applied load in a different manner than between smooth surfaces (Greenwood and Williamson 1966; Persson 2001). For nonadhering surfaces exhibiting an exponential distribution of *elastically* deforming asperities (spherical caps of equal radius), it has been shown that the contact area for rough surfaces increases approximately linearly with the applied normal force (load) L instead of as $L^{2/3}$ for smooth surfaces (Greenwood and Williamson 1966). It has also been shown that for *plastically* deforming metal microcontacts the real contact area increases with load as $A \propto L$ (Bowden and Tabor 1973; Maugis and Pollock 1984). In systems with attractive surface forces, there is a competition between this attraction and repulsive forces arising from compression of high asperities. As a result, the adhesion in such systems can be very low, especially if the surfaces are not easily deformed (Fuller and Tabor 1975; Persson and Tosatti 2001). The opposite is possible for soft (viscoelastic) surfaces where the real (molecular) contact area might be larger than for two perfectly smooth surfaces (Kim and Russell 2001). The size of the real contact area at a given normal force is also an important issue in studies of nanoscale friction, both of single-asperity contacts (Sect. 9.7) and of contacts between rough surfaces (Sect. 9.9.2).

Adhesion forces may also vary depending on the commensurability of the crystallographic lattices of the interacting surfaces. The adhesion

between two mica surfaces has been studied as a function of the orientation (twist angle) of their surface lattices (McGuiggan and Israelachvili 1990). The forces were measured in air, water, and an aqueous salt solution where oscillatory structural forces were present. In humid air, the adhesion was found to be relatively independent of the twist angle θ due to the adsorption of a 0.4 nm thick amorphous layer of organics and water at the interface. In contrast, in water, sharp adhesion peaks (energy minima) occurred at $\theta = 0, \pm 60, \pm 120$ and 180° , corresponding to the coincidence angles of the surface lattices (Fig. 9.10). As little as $\pm 1^\circ$ away from these peaks, the energy decreased by 50%. In aqueous KCl solution, due to potassium ion adsorption the water between the surfaces becomes ordered, resulting in an oscillatory force profile where the adhesive minima occur at discrete separations of about 0.25 nm, corresponding to integral numbers of water layers. The whole interaction

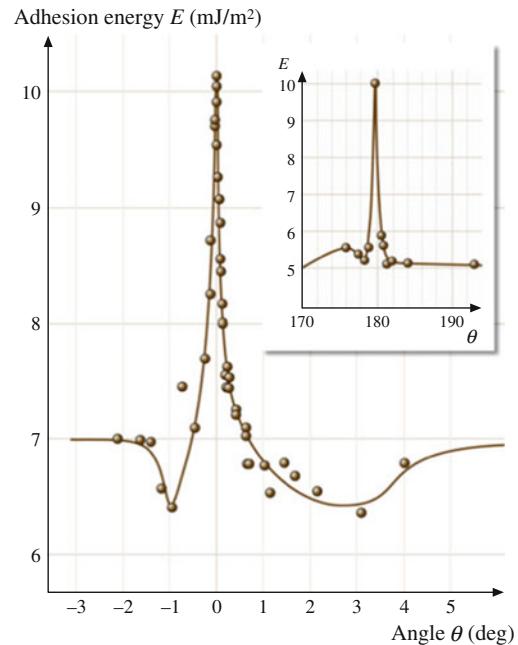


Fig. 9.10 Adhesion energy for two mica surfaces in contact in water (in the primary minimum of an oscillatory force curve) as a function of the mismatch angle θ about $\theta = 0$ and 180° between the mica surface lattices (after (McGuiggan and Israelachvili 1989) with permission)

potential was now found to depend on the orientation of the surface lattices, and the effect extended at least four molecular layers.

It has also been appreciated that the structure of the confining surfaces is just as important as the nature of the liquid for determining the solvation forces (Frink and van Swol 1998; Gao et al. 2000; Rhykerd et al. 1987; Schoen et al. 1989; Thompson and Robbins 1990). Between two surfaces that are completely flat but unstructured, the liquid molecules will order into layers, but there will be no lateral ordering within the layers. In other words, there will be positional ordering normal but not parallel to the surfaces. If the surfaces have a crystalline (periodic) lattice, this may induce ordering parallel to the surfaces, as well, and the oscillatory force then also depends on the structure of the surface lattices. Further, if the two lattices have different dimensions (mismatched or incommensurate lattices), or if the lattices are similar but are not in register relative to each other, the oscillatory force law is further modified (McGuigan and Israelachvili 1990; Ruths and Granick 2000) and the tribological properties of the film are also influenced, as discussed in Sect. 9.9 (Ruths and Granick 2000; Berman 1996).

As shown by the experiments, these effects can alter the magnitude of the adhesive minima found at a given separation within the last one or two nanometers of a thin film by a factor of two. The force barriers (maxima) may also depend on orientation. This could be even more important than the effects on the minima. A high barrier could prevent two surfaces from coming closer together into a much deeper adhesive well. Thus the maxima can effectively contribute to determining not only the final separation of two surfaces, but also their final adhesion. Such considerations should be particularly important for determining the thickness and strength of intergranular spaces in ceramics, the adhesion forces between colloidal particles in concentrated electrolyte solution, and the forces between two surfaces in a crack containing capillary condensed water.

For surfaces that are randomly rough, oscillatory forces become smoothed out and disappear altogether, to be replaced by a purely monotonic solvation force (Frink and van Swol 1998; Gao

et al. 2000; Gee and Israelachvili 1990). This occurs even if the liquid molecules themselves are perfectly capable of ordering into layers. The situation of symmetric liquid molecules confined between rough surfaces is therefore not unlike that of asymmetric molecules between smooth surfaces (Sect. 9.4.3 and Fig. 9.6a). To summarize, for there to be an oscillatory solvation force, the liquid molecules must be able to be correlated over a reasonably long range. This requires that both the liquid molecules and the surfaces have a high degree of order or symmetry. If either is missing, so will the oscillations. Depending on the size of the molecules to be confined, a roughness of only a few tenths of a nanometer is often sufficient to eliminate any oscillatory component of the force law (Ruths et al. 2001; Frink and van Swol 1998).

9.5.4 Nonequilibrium and Rate-Dependent Interactions: Adhesion Hysteresis

Under ideal conditions the adhesion energy is a well-defined thermodynamic quantity. It is normally denoted by E or W (the work of adhesion) or γ (the surface tension, where $W = 2\gamma$) and gives the reversible work done on bringing two surfaces together or the work needed to separate two surfaces from contact. Under ideal, equilibrium conditions these two quantities are the same, but under most realistic conditions they are not; the work needed to separate two surfaces is always greater than that originally gained by bringing them together. An understanding of the molecular mechanisms underlying this phenomenon is essential for understanding many adhesion phenomena, energy dissipation during loading–unloading cycles, contact angle hysteresis, and the molecular mechanisms associated with many frictional processes.

Adhesion hysteresis may be thought of as being due to mechanical effects such as instabilities, or chemical effects such as interdiffusion, interdigitation, molecular reorientations and exchange processes occurring at an interface after contact, as illustrated in Fig. 9.11. Such processes induce roughness and chemical

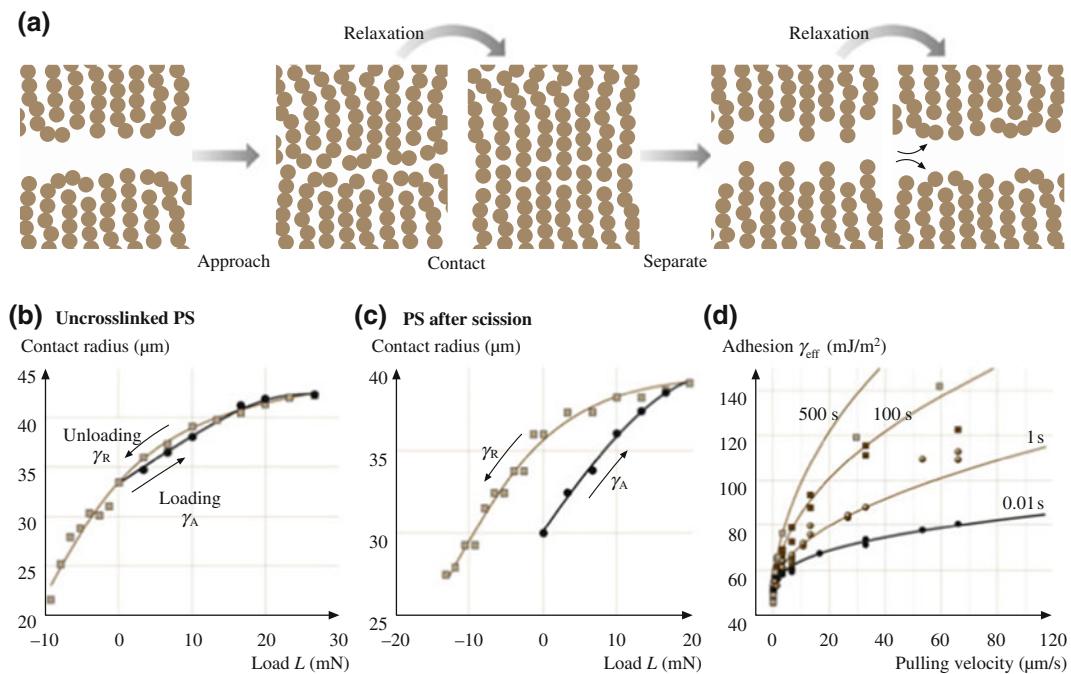


Fig. 9.11 **a** Schematic representation of interpenetrating chains. **b, c** JKR plots (contact radius r as a function of applied load L) showing small adhesion hysteresis for uncrosslinked polystyrene and larger adhesion hysteresis after chain scission at the surfaces after 18 h irradiation with ultraviolet light in an oxygen atmosphere. The adhesion hysteresis continues to increase with the irradiation time. **d** Rate-dependent adhesion of hexadecyl trimethyl ammonium bromide (CTAB) surfactant monolayers. The solid curves (Barthel and Roux 2000) are fits to experimental data on CTAB adhesion after different contact times (Ruths and Granick 1998) using an approximate analytical solution for a JKR model, including crack tip dissipation. Due to the limited range of validity of the approximation, the fits rely on the part of the experimental data with low effective adhesion energy only. From the fits one can determine the thermodynamic adhesion energy, the characteristic dissipation velocity, and the intrinsic dissipation exponent of the model ((a) after (Yoshizawa et al. 1993), © 1993 American Chemical Society, (b, c) after (Maeda et al. 2002), © 2002 American Association for the Advancement of Science, (d) after (Barthel and Roux 2000), © 2000 American Chemical Society)

heterogeneity even though initially (and after separation and reequilibration) both surfaces are perfectly smooth and chemically homogeneous. In general, if the energy change, or work done, on separating two surfaces from adhesive contact is not fully recoverable on bringing the two surfaces back into contact again, the adhesion hysteresis may be expressed as

$$W_R > W_A$$

or

$$\Delta W = (W_R - W_A) > 0, \quad (9.23)$$

where W_R and W_A are the adhesion or surface energies for receding (separating) and advancing (approaching) two solid surfaces, respectively.

Hysteresis effects are also commonly observed in wetting/dewetting phenomena (Miller and Neogi 1985). For example, when a liquid spreads and then retracts from a surface the advancing contact angle θ_A is generally larger than the receding angle θ_R . Since the contact angle θ is related to the liquid-vapor surface tension γ_L and the solid-liquid adhesion energy W by the Dupré equation

$$(1 + \cos \theta) \gamma_L = W, \quad (9.24)$$

we see that wetting hysteresis or contact angle hysteresis ($\theta_A > \theta_R$) actually implies adhesion hysteresis, $W_R > W_A$, as given by (9.23).

Energy-dissipating processes such as adhesion and contact angle hysteresis arise because of practical constraints of the *finite time* of measurements and the *finite elasticity* of materials. This prevents many loading–unloading or approach–separation cycles from being thermodynamically reversible, even though they would be if carried out infinitely slowly. By thermodynamically irreversible one simply means that one cannot go through the approach–separation cycle via a continuous series of equilibrium states, because some of these are connected via spontaneous—and therefore thermodynamically irreversible—instabilities or transitions where energy is liberated and therefore “lost” via heat or phonon release (Israelachvili and Berman 1995). This is an area of much current interest and activity, especially regarding the fundamental molecular origins of adhesion and friction in polymer and surfactant systems, and the relationships between them (Deruelle et al. 1995; Ruths and Granick 1998; Barthel and Roux 2000; Maeda et al. 2002).

9.6 Introduction: Different Modes of Friction and the Limits of Continuum Models

Most frictional processes occur with the sliding surfaces becoming damaged in one form or another (Bowden and Tabor 1971). This may be referred to as normal friction. In the case of brittle materials, the damaged surfaces slide past each other while separated by relatively large, micrometer-sized wear particles. With more ductile surfaces, the damage remains localized to nanometer-sized, plastically deformed asperities. Some features of the friction between damaged surfaces will be described in Sect. 9.7.4.

There are also situations in which sliding can occur between two perfectly smooth, undamaged surfaces. This may be referred to as interfacial sliding or boundary friction and is the focus of the following sections. The term boundary lubrication

is more commonly used to denote the friction of surfaces that contain a thin protective lubricating layer such as a surfactant monolayer, but here we shall use the term more broadly to include any molecularly thin solid, liquid, surfactant, or polymer film.

Experiments have shown that, as a liquid film becomes progressively thinner, its physical properties change, at first quantitatively and then qualitatively (Van Alsten and Granick 1990; Luengo et al. 1997a; Granick 1991; Hu and Granick 1992). The quantitative changes are manifested by an increased viscosity, non-Newtonian flow behavior, and the replacement of normal melting by a glass transition, but the film remains recognizable as a liquid (Fig. 9.12). In tribology, this regime is commonly known as the mixed lubrication regime, where the rheological properties of a film are intermediate between the bulk and boundary

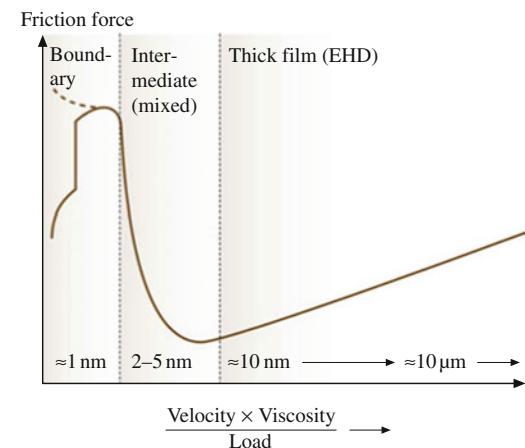


Fig. 9.12 Stribeck curve an empirical curve giving the trend generally observed in the friction forces or friction coefficients as a function of sliding velocity, the bulk viscosity of the lubricating fluid, and the applied load (normal force). The three friction/lubrication regimes are known as the boundary lubrication regime (Sect. 9.7), the intermediate or mixed lubrication regime (Sect. 9.8.2), and thick film or elastohydrodynamic (EHD) lubrication regime (Sect. 9.8.1). The film thicknesses believed to correspond to each of these regimes are also shown. For thick films, the friction force is purely viscous, e.g., Couette flow at low shear rates, but may become complicated at higher shear rates where EHD deformations of surfaces can occur during sliding (after (Israelachvili 1995), with permission)

Table 9.4 The three main tribological regimes characterizing the changing properties of liquids subjected to increasing confinement between two solid surfaces^a

Regime	Conditions for getting into this regime	Static/equilibrium properties ^b	Dynamic properties ^c
Bulk	<ul style="list-style-type: none"> • Thick films (>10 molecular diameters, $\gg R_g$ for polymers) • Low or zero loads • High shear rates 	<p>Bulk (continuum) properties:</p> <ul style="list-style-type: none"> • Bulk liquid density • No long-range order 	<p>Bulk (continuum) properties:</p> <ul style="list-style-type: none"> • Newtonian viscosity • Fast relaxation times • No glass temperature • No yield point • Elastohydrodynamic lubrication
Intermediate mixed	<ul style="list-style-type: none"> • Intermediately thick films (4–10 molecular diameters, $\sim R_g$ for polymers) • Low loads or pressure 	<p>Modified fluid properties include:</p> <ul style="list-style-type: none"> • Modified positional and orientational order^a • Medium- to long-range molecular correlations • Highly entangled states 	<p>Modified rheological properties include:</p> <ul style="list-style-type: none"> • Non-Newtonian flow • Glassy states • Long relaxation times • Mixed lubrication
Boundary	<ul style="list-style-type: none"> • Molecularly thin films (<4 molecular diameters) • High loads or pressure • Low shear rates • Smooth surfaces or asperities 	<p>Onset of nonfluidlike properties:</p> <ul style="list-style-type: none"> • Liquidlike to solidlike phase transitions • Appearance of new liquid-crystalline states • Epitaxially induced long-range ordering 	<p>Onset of tribological properties:</p> <ul style="list-style-type: none"> • No flow until yield point or critical shear stress reached • Solidlike film behavior characterized by defect diffusion, dislocation motion, shear melting • Boundary lubrication

Based on work by Granick (Granick 1991), and others (Hirz et al. 1992; Yoshizawa et al. 1993; Gee et al. 1990) on the dynamic properties of short chain molecules such as alkanes and polymer melts confined between surfaces

^aConfinement can lead to an increased or decreased order in a film, depending both on the surface lattice structure and the geometry of the confining cavity

^bIn each regime both the static and dynamic properties change. The static properties include the film density, the density distribution function, the potential of mean force, and various positional and orientational order parameters

^cDynamic properties include viscosity, viscoelastic constants, and tribological yield points such as the friction coefficient and critical shear stress

properties. One may also refer to it as the intermediate regime (Table 9.4).

For even thinner films, the changes in behavior are more dramatic, resulting in a qualitative change in properties. Thus first-order phase transitions can now occur to solid or liquid-crystalline phases (Klein and Kumacheva 1995, 1998; Thompson and Robbins 1990; Yoshizawa et al. 1993; Gee et al. 1990; Israelachvili et al. 1990a,b; Demirel and Granick

1996), whose properties can no longer be characterized even qualitatively in terms of bulk or continuum liquid properties such as viscosity. These films now exhibit yield points (characteristic of fracture in solids) and their molecular diffusion and relaxation times can be ten orders of magnitude longer than in the bulk liquid or even in films that are just slightly thicker. The three friction regimes are summarized in Table 9.4.

9.7 Relationship Between Adhesion and Friction Between Dry (Unlubricated and Solid Boundary Lubricated) Surfaces

9.7.1 Amontons' Law

Early theories and mechanisms for the dependence of friction on the applied normal force or load L were developed by *da Vinci*, *Amontons*, *Coulomb* and *Euler* (Dowson 1998). For the macroscopic objects investigated, the friction was found to be directly proportional to the load, with no dependence on the contact area. This is described by the so-called Amontons' law

$$F = \mu L, \quad (9.25)$$

where F is the shear or friction force and μ is a constant defined as the coefficient of friction. This friction law is still the principal means of quantitatively describing the friction between surfaces. However, particularly in the case of adhering surfaces, Amontons' law does not adequately describe the friction behavior with load, because of the finite friction force measured at zero and even negative applied loads.

When a lateral force, or shear stress, is applied to two surfaces in adhesive contact, the surfaces initially remain “pinned” to each other until some critical shear force is reached. At this point, the surfaces begin to slide past each other either smoothly or in jerks. The frictional force needed to initiate sliding from rest is known as the static friction force, denoted by F_s , while the force needed to maintain smooth sliding is referred to as the kinetic or dynamic friction force, denoted by F_k . In general, $F_s > F_k$. Two sliding surfaces may also move in regular jerks, known as stick-slip sliding, which is discussed in more detail in Sect. 9.8.3. Such friction forces cannot be described by models used for thick films that are viscous (Sect. 9.8.1) and, therefore, shear as soon as the smallest shear force is applied.

9.7.2 Adhesion Force and Load Contribution to Interfacial Friction

In an adhesive contact, the contact area as a function of load is thus generally well described by the JKR Eq. (9.21). The friction force between two molecularly smooth surfaces sliding in adhesive contact is not simply proportional to the applied load L as might be expected from Amontons' law. There is an additional adhesion contribution that is proportional to the area of contact, A . Thus, in general, the interfacial friction force of dry, unlubricated surfaces sliding smoothly past each other in adhesive contact is given by

$$F = F_k = S_c A + \mu L, \quad (9.26)$$

where S_c is the critical shear stress (assumed to be constant), $A = \pi r^2$ is the contact area of radius r given by (9.21), and μ is the coefficient of friction. For low loads we have

$$\begin{aligned} F &= S_c A = S_c \pi r^2 \\ &= S_c \pi \left[\frac{R}{K} \left(L + 6\pi R \gamma + \sqrt{12\pi R \gamma L + (6\pi R \gamma)^2} \right) \right]^{2/3}, \end{aligned} \quad (9.27)$$

where as for high loads (or high μ), or when γ is very low (Berman et al. 1998; Ruths 2003; Gourdon and Israelachvili 2003; Ruths et al. 2003; Gao et al. 2004), (9.26) reduces to Amontons' law: $F = \mu L$. Depending on whether the friction force in (9.26) is dominated by the first or second term, one may refer to the friction as adhesion-controlled or load-controlled, respectively.

The following friction model (Fig. 9.13a), first proposed by Tabor (1981) and developed further by others (Homola et al. 1989; Sutcliffe et al. 1978; McClelland 1989), has been quite successful at explaining the interfacial and boundary friction of two solid crystalline

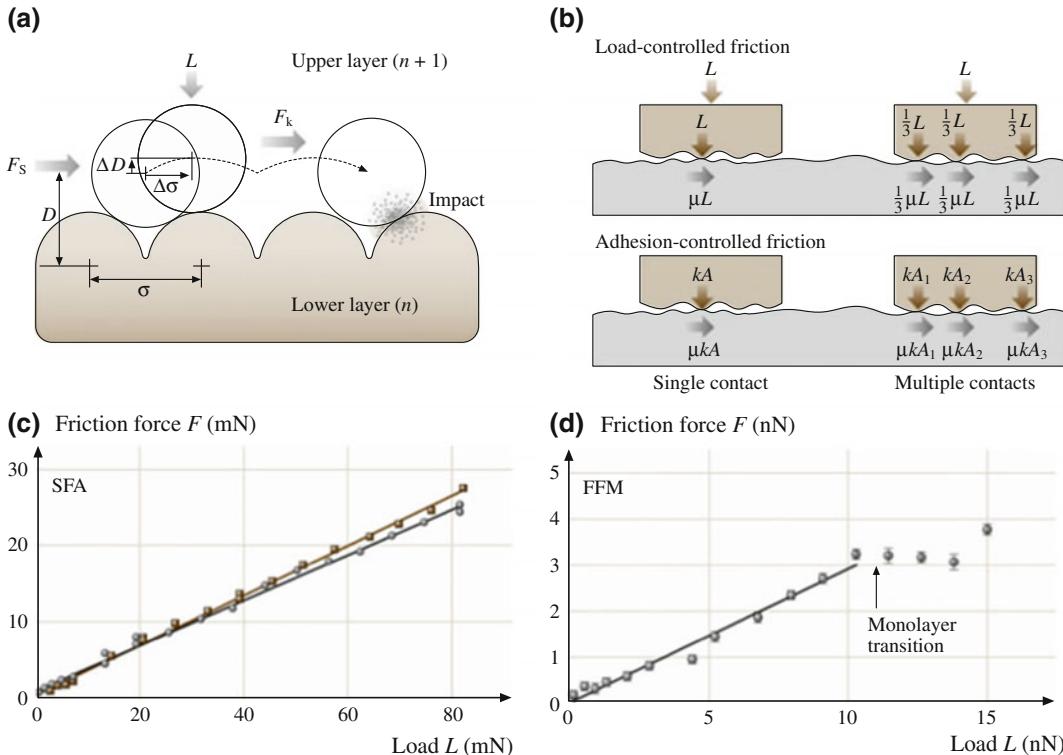


Fig. 9.13 **a** Schematic illustration of how one molecularly smooth surface moves over another when a lateral force F is applied (the *cobblestone model*). As the upper surface moves laterally by some fraction of the lattice dimension $\Delta\sigma$, it must also move up by some fraction of an atomic or molecular dimension ΔD before it can slide across the lower surface. On impact, some fraction ϵ of the kinetic energy is *transmitted* to the lower surface, the rest being *reflected* back to the colliding molecule (*upper surface*) (after (Israelachvili et al. 1994), with permission). **b** Difference in the local distribution of the total applied external load or normal adhesive force between load-controlled nonadhering surfaces and adhesion-controlled surfaces. In the former case, the total friction force F is given either by $F = \mu L$ for one contact point (*left side*) or by $F = 1/3 \mu L + 1/3 \mu L + 1/3 \mu L = \mu L$ for three contact points (*right side*). Thus the load-controlled friction is always proportional to the applied load, independently of the number of contacts and of their geometry. In the case of adhering surfaces, the effective *internal* load is given by kA , where A is the real local contact area, which is proportional to the number of intermolecular bonds being made and broken across each single contact point. The total friction force is now given by $F = \mu kA$ for one contact point (*left side*), and $F = \mu(kA_1 + kA_2 + kA_3) = \mu kA_{\text{tot}}$ for three contact points (*right side*). Thus, for adhesion-controlled friction, the friction is proportional to the real contact area, at least when no additional external load is applied to the system (after (Gao et al. 2004), with permission, © 2004 American Chemical Society). **c, d** Friction force between benzyltrichlorosilane monolayers chemically bound to glass or Si, measured in ethanol ($\gamma < 1 \text{ mJ/m}^2$). **c** SFA measurements where both glass surfaces were covered with a monolayer. Circles and squares show two different experiments: one with $R = 2.6 \text{ cm}$, $v = 0.15 \mu\text{m/s}$, giving $\mu = 0.33 \pm 0.01$; the other with $R = 1.6 \text{ cm}$, $v = 0.5 \mu\text{m/s}$, giving $\mu = 0.30 \pm 0.01$. **d** Friction force microscopy (FFM) measurements of a monolayer functionalized Si tip ($R = 11 \text{ nm}$) sliding on a monolayer-covered glass surface at $v = 0.15 \mu\text{m/s}$, giving $\mu = 0.30 \pm 0.01$. Note the different scales in (c) and (d) (after (Ruths et al. 2003), with permission, © 2003 American Chemical Society)

surfaces sliding past each other in the absence of wear. The surfaces may be unlubricated, or they may be separated by a monolayer or more of

some boundary lubricant or liquid molecules. In this model, the values of the critical shear stress S_c , and the coefficient of friction μ , in (9.26) are

calculated in terms of the energy needed to overcome the attractive intermolecular forces and compressive externally applied load as one surface is raised and then slid across the molecular-sized asperities of the other.

This model is similar to pushing a cart over a road of cobblestones where the cartwheels (which represent the molecules of the upper surface or film) must be made to roll over the cobblestones (representing the molecules of the lower surface) before the cart can move. When at rest, the cartwheels find grooves between the cobblestones where they sit in potential-energy minima, and so the cart is at some stable mechanical equilibrium. A certain lateral force (the push) is required to raise the cartwheels against the force of gravity in order to initiate motion. Motion will continue as long as the cart is pushed, and rapidly stops once it is no longer pushed. Energy is dissipated by the liberation of heat every time a wheel hits the next cobblestone. The cobblestone model is not unlike the Coulomb and interlocking asperity models of friction (Dowson 1998) except that it is being applied at the molecular level and for a situation where the external load is augmented by attractive intermolecular forces.

There are thus two contributions to the force pulling two surfaces together: the externally applied load or pressure, and the (internal) attractive intermolecular forces that determine the adhesion between the two surfaces. Each of these contributions affects the friction force in a different way, which we will discuss in more detail below.

9.7.2.1 Adhesion Force Contribution

When the two surfaces are initially in adhesive contact, the surface molecules will adjust themselves to fit snugly together (Buckley 1977), in an analogous manner to the self-positioning of the cartwheels on the cobblestone road. A small tangential force applied to one surface will therefore not result in the sliding of that surface relative to the other. The attractive van der Waals forces between the surfaces must first be

overcome by having the surfaces separate by a small amount. To initiate motion, let the separation between the two surfaces increase by a small amount ΔD , while the lateral distance moved is $\Delta\sigma$. The energy put into the system by the force F acting over a lateral distance $\Delta\sigma$ is

$$\text{Input energy: } F \times \Delta\sigma. \quad (9.28)$$

This energy may be equated with the change in interfacial or surface energy associated with separating the surfaces by ΔD , i.e., from the equilibrium separation $D = D_0$ to $D = (D_0 + \Delta D)$. Since $\gamma \propto D^{-2}$ for two flat surfaces (Sect. 9.3.1, Table 9.2), the energy cost may be approximated by

Surface energy change \times area:

$$2\gamma A \left[1 - \frac{D_0^2}{(D_0 + \Delta D)^2} \right] \approx 4\gamma A \left(\frac{\Delta D}{D_0} \right), \quad (9.29)$$

where γ is the surface energy, A the contact area, and D_0 the surface separation at equilibrium. During steady state sliding (kinetic friction), not all of this energy will be lost or absorbed by the lattice every time the surface molecules move by one lattice spacing: some fraction will be reflected during each impact of the cartwheel molecules (Israelachvili et al. 1994). Assuming that a fraction ε of the above surface energy is lost every time the surfaces move across the characteristic length $\Delta\sigma$ (Fig. 9.13a), we obtain after equating (9.28) and (9.29)

$$S_c = \frac{F}{A} = \frac{4\gamma\varepsilon\Delta D}{D_0\Delta\sigma}. \quad (9.30)$$

The above model suggests that all interfaces, whether dry or lubricated, dilate just before they shear or slip. This dilation, which provides the crucial extra space for the molecules to slide across each other, is known to occur in macroscopic systems (Cain et al. 2001) and for nanoscopic systems it has been computed (Thompson and Robbins 1990) and also measured experimentally (Dhinojwala et al. 2000).

This model may be extended to lubricated sliding, where a thin liquid film is present between the surfaces. With an increase in the number of liquid layers between the surfaces, D_0 increases while ΔD decreases, hence the friction force decreases. This is precisely what is observed, but with more than one liquid layer between two surfaces the situation becomes too complex to analyze analytically. Furthermore, even in systems as simple as linear alkanes, interdigitation and interdiffusion have been found to contribute strongly to the properties of the system (Qian et al. 2003). Sophisticated modeling based on computer simulations is now required, as discussed in the following section.

9.7.2.2 Relation Between Boundary Friction and Adhesion Energy Hysteresis

While the above equations suggest that there is a direct correlation between friction and adhesion, this is not the case. The correlation is really between friction and adhesion hysteresis, described in Sect. 9.5.4. In the case of friction, this subtle point is hidden in the factor ε , which is a measure of the amount of energy absorbed (dissipated, transferred, or lost) by the lower surface when it is impacted by a molecule from the upper surface. If $\varepsilon = 0$, all the energy is reflected, and there will be no kinetic friction force or any adhesion hysteresis, but the absolute magnitude of the adhesion force or energy will remain finite and unchanged. This is illustrated in Fig. 9.17.

The following simple model shows how adhesion hysteresis and friction may be quantitatively related. Let $\Delta\gamma = \gamma_R - \gamma_A$ be the adhesion energy hysteresis per unit area, as measured during a typical loading—unloading cycle. Now consider the same two surfaces sliding past each other and assume that frictional energy dissipation occurs through the same mechanism as adhesion energy dissipation, and that both occur over the same characteristic molecular length scale σ . Thus, when the two surfaces (of contact area $A = \pi r^2$) move a distance σ , equating the

frictional energy ($F \times \sigma$) to the dissipated adhesion energy ($A \times \Delta\gamma$), we obtain

$$\text{Friction force: } \frac{A \times \Delta\gamma}{\sigma} = \frac{\pi r^2}{\sigma} (\gamma_R - \gamma_A), \quad (9.31)$$

$$\text{or Critical shear stress: } S_c = F/A = \Delta\gamma/\sigma, \quad (9.32)$$

which is the desired expression and has been found to give order-of-magnitude agreement between measured friction forces and adhesion energy hysteresis (Yoshizawa et al. 1993). If we equate Eq. (9.32) with (9.30), since $4\Delta D/(D_0\Delta\sigma) \approx 1/\sigma$, we obtain the intuitive relation

$$\varepsilon = \frac{\Delta\gamma}{\gamma}. \quad (9.33)$$

9.7.2.3 External Load Contribution to Interfacial Friction

When there is no interfacial adhesion, S_c is zero. Thus, the only attractive force that needs to be overcome for sliding to occur is the externally applied load or pressure, as shown in Fig. 9.13b.

For a general semiquantitative analysis, again consider the cobblestone model used to derive (9.30), but now include an additional contribution to the surface-energy change of (9.29) due to the work done against the external load or pressure, $L\Delta D = P_{\text{ext}} A\Delta D$ (this is equivalent to the work done against gravity in the case of a cart being pushed over cobblestones). Thus

$$S_c = \frac{F}{A} = \frac{4\gamma\varepsilon\Delta D}{D_0\Delta\sigma} + \frac{P_{\text{ext}}\varepsilon\Delta D}{\Delta\sigma}, \quad (9.34)$$

which gives the more general relation

$$S_c = \frac{F}{A} = C_1 + C_2 + P_{\text{ext}}, \quad (9.35)$$

where $P_{\text{ext}} = L/A$ and C_1 and C_2 are characteristic of the surfaces and sliding conditions. The constant $C_1 = 4\gamma\varepsilon\Delta D/(D_0\Delta\sigma)$ depends on the

mutual adhesion of the two surfaces, while both C_1 and $C_2 = \varepsilon\Delta D/\Delta\sigma$ depend on the topography or atomic bumpiness of the surface groups (Fig. 9.13a). The smoother the surface groups the smaller the ratio $\Delta D/\Delta\sigma$ and hence the lower the value of C_2 . In addition, both C_1 and C_2 depend on ε (the fraction of energy dissipated per collision), which depends on the relative masses of the shearing molecules, the sliding velocity, the temperature, and the characteristic molecular relaxation processes of the surfaces. This is by far the most difficult parameter to compute, and yet it is the most important since it represents the energy-transfer mechanism in any friction process, and since ε can vary between 1 and 0, it determines whether a particular friction force will be large or close to zero.

Finally, the above equation may also be expressed in terms of the friction force F

$$F = S_c A = C_1 A + C_2 L. \quad (9.36)$$

Equations similar to (9.35) and (9.36) were previously derived by Derjaguin (1934b,

Derjaguin 1988) and by Briscoe and Evans (1982), where the constants C_1 and C_2 were interpreted somewhat differently than in this model. In the absence of any attractive interfacial force, we have $C_1 \approx 0$, and the second term in (9.35) and (9.36) should dominate (Fig. 9.14). Such situations typically arise when surfaces repel each other across the lubricating liquid film. In such cases, the total frictional force should be low and should increase linearly with the external load according to

$$F = C_2 L. \quad (9.37)$$

An example of such lubricated sliding occurs when two mica surfaces slide in water or in salt solution (Fig. 9.18a), where the short-range hydration forces between the surfaces are repulsive. Thus, for sliding in 0.5 M KCl it was found that $C_2 = 0.015$ (Berman et al. 1998). Another case where repulsive surfaces eliminate the adhesive contribution to friction is for polymer chains attached to surfaces at one end and swollen by a good solvent (Klein et al. 1994).

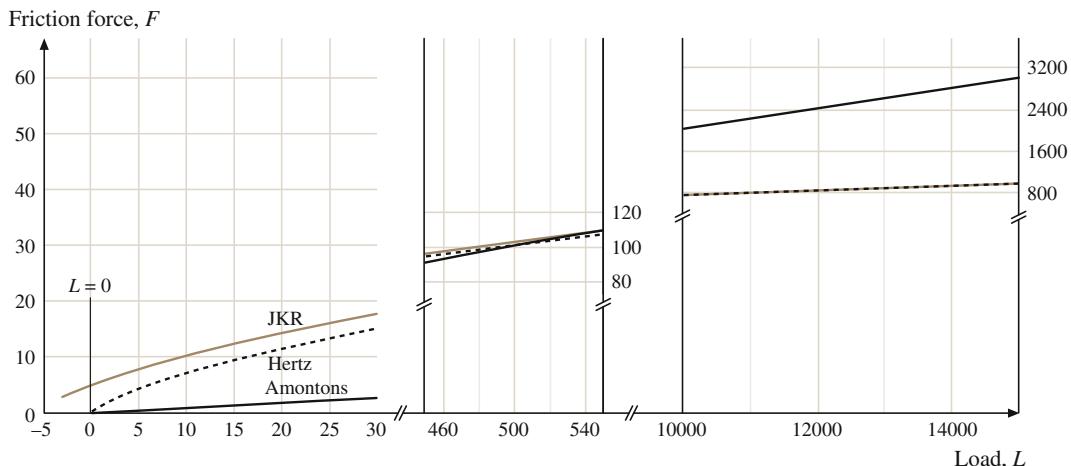


Fig. 9.14 Friction as a function of load for smooth surfaces. At low loads, the friction is dominated by the $C_1 A$ term of (9.36). The adhesion contribution (JKR curve) is most prominent near zero load where the Hertzian and Amontons' contributions to the friction are minimal. As the load increases, the adhesion contribution becomes smaller as the JKR and Hertz curves converge. In this range of loads, the linear $C_2 L$ contribution surpasses the area contribution to the friction. At much higher loads the explicit load dependence of the friction dominates the interactions, and the observed behavior approaches Amontons' law. It is interesting to note that for smooth surfaces the pressure over the contact area does not increase as rapidly as the load. This is because as the load is increased, the surfaces deform to increase the surface area and thus moderate the contact pressure (after (Berman and Israelachvili 1997), with permission of Kluwer Academic Publishers)

For this class of systems, $C_2 < 0.001$ for a finite range of polymer layer compressions (normal loads L). The low friction between the surfaces in this regime is attributed to the entropic repulsion between the opposing brush layers with a minimum of entanglement between the two layers. However, with higher normal loads, the brush layers become compressed and begin to entangle, which results in higher friction (Luckham and Manimaaran 1997). It is important to note that (9.37) has exactly the same form as Amontons' Law

$$F = \mu L, \quad (9.38)$$

where μ is the coefficient of friction.

Figure 9.13c,d shows the kinetic friction force measured with both SFA and FFM (friction force microscopy, using AFM) on a system where both surfaces were covered with a chemically bound benzyltrichlorosilane monolayer (Ruths et al. 2003). When immersed in ethanol, the adhesion in this system is low, and very different contact areas and loads give a linear dependence of F on L with the same friction coefficients, and $F \rightarrow 0$ as $L \rightarrow 0$. In the FFM measurements (Fig. 9.13d), the plateau in the data at higher loads suggest a transition in the monolayers, similar to previous observations on other monolayer systems. The pressure in the contact region in the SFA is much lower than in the FFM, and no transitions in the friction forces or in the thickness of the confined monolayers were observed in the SFA experiments (and no damage to the monolayers or the underlying substrates was observed during the experiments, indicating that the friction was wearless).

At the molecular level a thermodynamic analog of the Coulomb or cobblestone models (Sect. 9.7.2) based on the contact value theorem (Israelachvili 2011; Berman et al. 1998; Berman and Israelachvili 1997) can explain why $F \propto L$ also holds at the microscopic or molecular level. In this analysis we consider the surface molecular groups as being momentarily compressed and decompressed as the surfaces move along. Under irreversible conditions, which always occur when a cycle is completed in a

finite amount of time, the energy lost in the compression–decompression cycle is dissipated as heat. For two nonadhering surfaces, the stabilizing pressure P_i acting locally between any two elemental contact points i of the surfaces may be expressed by the contact value theorem (Israelachvili 2011)

$$P_i = \rho_i k_B T = k_B T / V_i, \quad (9.39)$$

where $\rho_i = V_i^{-1}$ is the local number density (per unit volume) or activity of the interacting entities, be they molecules, atoms, ions or the electron clouds of atoms. This equation is essentially the osmotic or entropic pressure of a gas of confined molecules. As one surface moves across the other, local regions become compressed and decompressed by a volume ΔV_i . The work done per cycle can be written as $\varepsilon P_i \Delta V_i$, where ε ($\varepsilon \leq 1$) is the fraction of energy per cycle lost as heat, as defined earlier. The energy balance shows that, for each compression–decompression cycle, the dissipated energy is related to the friction force by

$$F_i x_i = \varepsilon P_i \Delta V_i, \quad (9.40)$$

where x_i is the lateral distance moved per cycle, which can be the distance between asperities or the distance between surface lattice sites. The pressure at each contact junction can be expressed in terms of the local normal load L_i and local area of contact A_i as $P_i = L_i / A_i$. The volume change over a cycle can thus be expressed as $\Delta V_i = A_i z_i$, where z_i is the vertical distance of confinement. Inserting these into (9.40), we get

$$F_i = \varepsilon L_i (z_i / x_i), \quad (9.41)$$

which is independent of the local contact area A_i . The total friction force is thus

$$\begin{aligned} F &= \sum F_i = \sum \varepsilon L_i (z_i / x_i) \\ &= \varepsilon \langle z_i / x_i \rangle \sum L_i = \mu L, \end{aligned} \quad (9.42)$$

where it is assumed that on average the local values of L_i and P_i are independent of the local slope z_i/x_i . Therefore, the friction coefficient μ is

a function only of the average surface topography and the sliding velocity, but is independent of the local (real) or macroscopic (apparent) contact areas. While this analysis explains non-adhering surfaces, there is still an additional explicit contact area contribution for the case of adhering surfaces, as in (9.36). The distinction between the two cases arises because the initial assumption of the contact value theorem (9.39) is incomplete for adhering systems.

Molecular scale roughness, real contact areas, and changes therein are difficult to quantify in nanoscale friction experiments, which involve deformable substrates in continuous motion under considerable pressure. In contrast, the nominal contact area can be either directly measured (e.g., in the SFA) or estimated from contact mechanics models, although concerns have been raised about the applicability of such models on the nanometer and sub-nanometer scale (Luan and Robbins 2005) where atomic level roughness causes local changes in the pressure. Recent computer simulations (Mo et al. 2009; Mo and Szlufarska 2010; Cheng et al. 2010) have suggested that within the nominal contact area, the size of the real, molecular contact area and its increase with increasing load are adhesion-dependent, but in a different manner than that observed for nominal or macroscopic contacts.

9.7.3 Examples of Experimentally Observed Friction of Dry Surfaces

Numerous model systems, including a variety of both unlubricated and solid- and liquid-lubricated surfaces, have been studied with a surface forces apparatus (SFA) modified for friction experiments (Sect. 9.2.3). The apparatus allows for control of load (normal force) and sliding speed, and simultaneous measurement of surface separation, surface shape, true (molecular) area of contact between smooth surfaces, and friction forces.

Figure 9.15 shows the contact area A and friction force F , both plotted against the applied load L in an experiment in which two

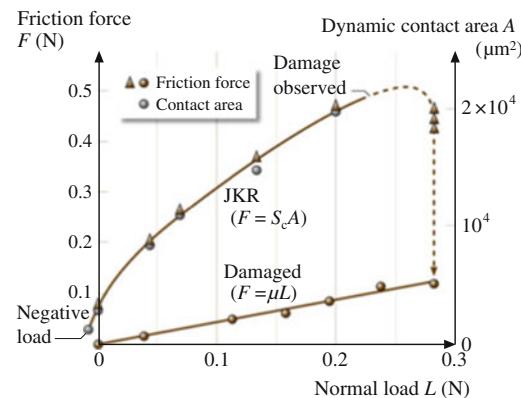


Fig. 9.15 Friction force F and contact area A versus load L for two mica surfaces sliding in adhesive contact in dry air. The contact area is well described by the JKR theory, (9.21), even during sliding, and the friction force is found to be directly proportional to this area, (9.27). The vertical dashed line and arrow show the transition from interfacial to normal friction with the onset of wear (*lower curve*). The sliding velocity is $0.2 \mu\text{m s}^{-1}$ (after (Homola et al. 1989), with permission, © 1989 American Society of Mechanical Engineers)

molecularly smooth surfaces of mica in adhesive contact were slid past each other in an atmosphere of dry nitrogen gas. This is an example of the low-load adhesion-controlled limit, which is excellently described by (9.27). In a number of different experiments, S_c was measured to be $2.5 \times 10^7 \text{ Nm}^{-2}$ and to be independent of the sliding velocity (Homola et al. 1989, 1990). Note that there is a friction force even at negative loads, where the surfaces are still sliding in adhesive contact.

The high friction force of unlubricated sliding can often be reduced by treating the solid surface with a boundary layer of some other solid material that exhibits lower friction, such as a surfactant monolayer, or by ensuring that during sliding a thin liquid film remains between the surfaces (as will be discussed in Sect. 9.8). The effectiveness of a solid boundary lubricant layer on reducing the forces of friction is illustrated in Fig. 9.16. Comparing this with the friction of the unlubricated/untreated surfaces (Fig. 9.15) shows that the critical shear stress has been reduced by a factor of about ten: from 2.5×10^7 to $3.5 \times 10^6 \text{ Nm}^{-2}$. At much higher applied

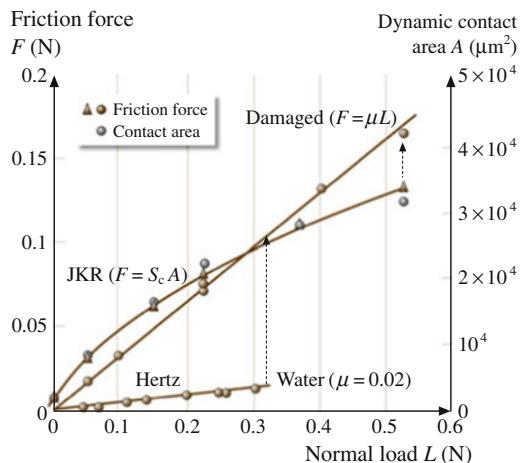


Fig. 9.16 Sliding of mica surfaces, each coated with a 2.5 nm thick monolayer of calcium stearate surfactant, in the absence of damage (obeying JKR-type boundary friction) and in the presence of damage (obeying Amontons-type friction). Note that both for this system and for the bare mica in Figs. 9.15 and 9.18a, the friction force obeys Amontons' law with a friction coefficient of $\mu \approx 0.3$ after damage occurs. At much higher applied loads, the undamaged surfaces also follow Amontons-type sliding, but for a different reason: the dependence on adhesion becomes smaller. Lower line: interfacial sliding with a monolayer of water between the mica surfaces (load-controlled friction, Fig. 9.18a), shown for comparison (after Homola et al. 1990), © 1990, with permission from Elsevier Science)

loads or pressures, the friction force is proportional to the load, rather than the area of contact (Briscoe et al. 1977), as expected from (9.26).

The adhesion and friction of fluorocarbon surfaces (surfactant coated boundary lubricant layers) have been compared to those of hydrocarbon surfaces (Yamada and Israelachvili 1998). It was concluded that well-ordered fluorocarbon surfaces have high friction, in spite of their lower adhesion energy. The low friction coefficient of Teflon (polytetrafluoroethylene, PTFE) must, therefore, be due to some effect other than low adhesion. On a related issue, it has been found that C₆₀ surfaces also exhibited low adhesion but high friction (Luengo et al. 1997b). In both cases the high friction appears to arise from the bulky surface groups—fluorocarbon compared to hydrocarbon groups in the former, large fullerene spheres in the latter. Apparently, the fact that C₆₀

molecules rotate *in their lattice* does not make them a good lubricant: the molecules of the opposing surface must still climb over them in order to slide, and this requires energy that is independent of whether the surface molecules are fixed or freely rotating.

Figure 9.17 illustrates the relationship between adhesion hysteresis and friction for surfactant-coated surfaces under different conditions. This effect, however, is much more general and has been shown to hold for other surfaces as well (Vigil et al. 1994; Maeda et al. 2002; Israelachvili et al. 1994, 1995).

Direct comparisons between absolute adhesion energies and friction forces show little correlation. In some cases, higher adhesion energies for the same system under different conditions correspond to lower friction forces. For example, for hydrophilic silica surfaces it was found that with increasing relative humidity the adhesion energy increases, but the adhesion energy hysteresis measured in a loading–unloading cycle decreases, as does the friction force (Vigil et al. 1994).

A liquid lubricant film (Sect. 9.8.3) is usually much more effective at lowering the friction of two surfaces than a solid boundary lubricant layer. However, to use a liquid lubricant successfully, it must “wet” the surfaces, that is, it should have a high affinity for the surfaces, so that not all the liquid molecules become squeezed out when the surfaces come close together, even under a large compressive load. Another important requirement is that the liquid film remains a liquid under tribological conditions, i.e., that it does not epitaxially solidify between the surfaces.

Effective lubrication usually requires that the lubricant be injected between the surfaces, but in some cases the liquid can be made to condense from the vapor. This is illustrated in Fig. 9.18a for two untreated mica surfaces sliding with a thin layer of water between them. A monomolecular film of water (of thickness 0.25 nm per surface) has reduced S_c from its value for dry surfaces (Fig. 9.15) by a factor of more than 30, which may be compared with the factor of ten attained with the boundary lubricant

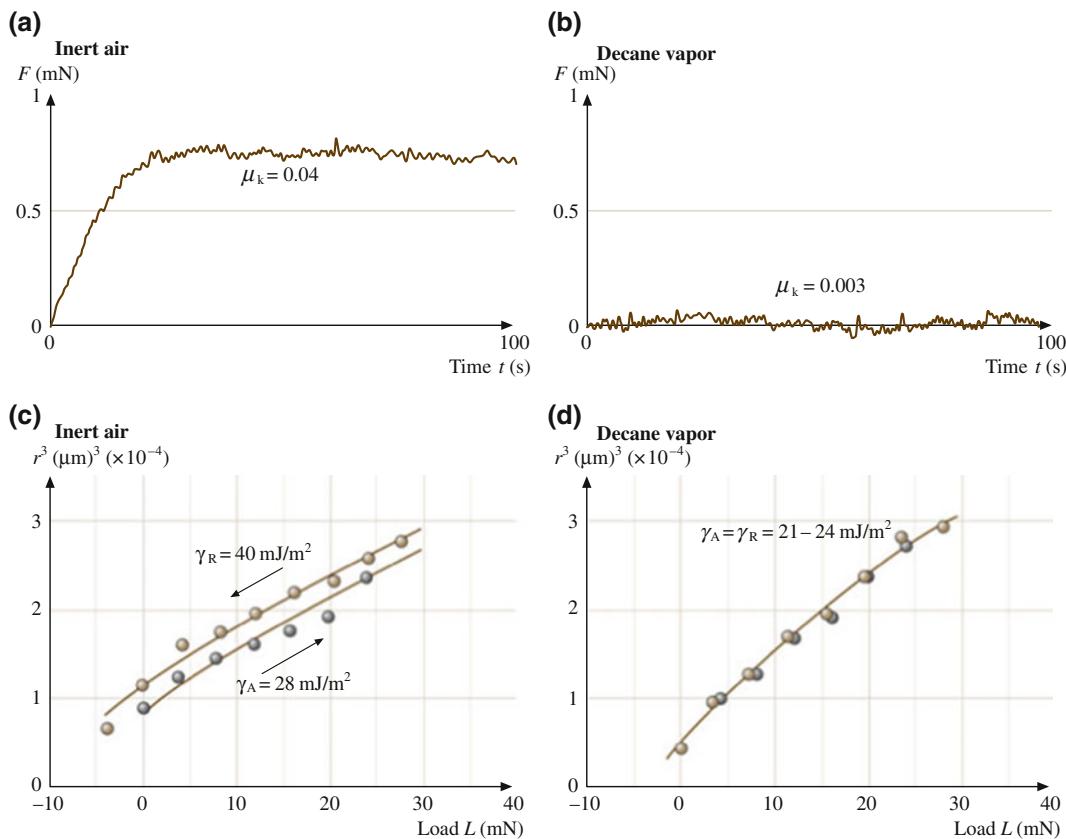


Fig. 9.17 Top: Friction traces for two fluidlike calcium alkylbenzene sulfonate monolayer-coated surfaces at 25 °C showing that the friction force is much higher between dry monolayers (a) than between monolayers whose fluidity has been enhanced by hydrocarbon penetration from vapor (b). Bottom: Contact radius versus load (r^3 versus L) data measured for the same two surfaces as above and fitted to the JKR Eq. (9.21), shown by the solid curves. For dry monolayers (c) the adhesion energy on unloading ($\gamma_R = 40 \text{ mJ m}^{-2}$) is greater than that on loading ($\gamma_A = 28 \text{ mJ m}^{-2}$), which is indicative of an adhesion energy hysteresis of $\Delta\gamma = \gamma_R - \gamma_A = 12 \text{ mJ m}^{-2}$. For monolayers exposed to saturated decane vapor (d) their adhesion hysteresis is zero ($\gamma_A = \gamma_R$), and both the loading and unloading data are well fitted by the thermodynamic value of the surface energy of fluid hydrocarbon chains, $\gamma = 24 \text{ mJ m}^{-2}$ (after (Yoshizawa et al. 1993), with permission, © 1993 American Chemical Society)

layer (of thickness 2.5 nm per surface) in Fig. 9.16. Water appears to have unusual lubricating properties and usually gives wearless friction with no stick-slip (Raviv et al. 2004).

The effectiveness of a water film only 0.25 nm thick to lower the friction force by more than an order of magnitude is attributed to the hydrophilicity of the mica surface (mica is wetted

by water) and to the existence of a strongly repulsive short-range hydration force between such surfaces in aqueous solutions, which effectively removes the adhesion-controlled contribution to the friction force (Ruths 2003). It is also interesting that a 0.25 nm thick water film between two mica surfaces is sufficient to bring the coefficient of friction down to 0.01–0.02, a

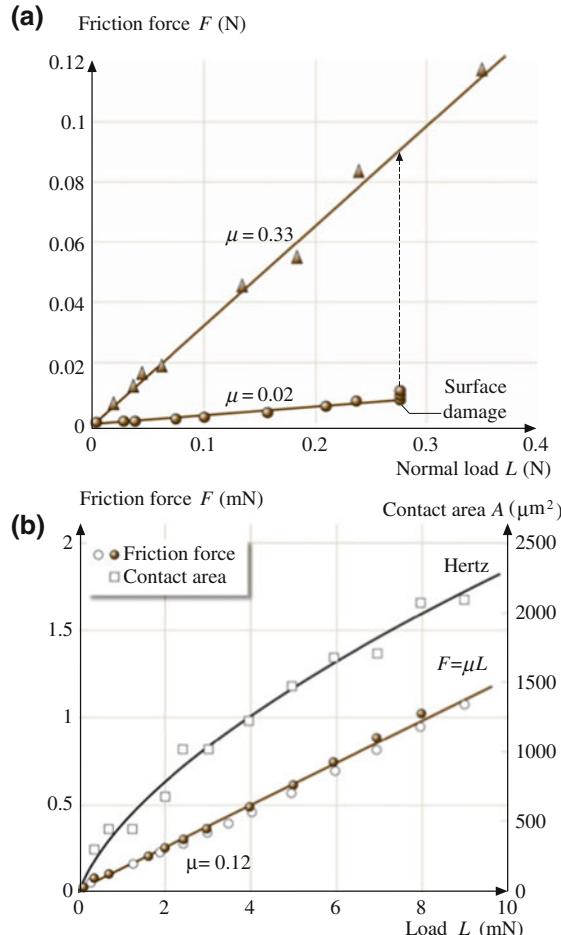


Fig. 9.18 Load-controlled friction. **a** Two mica surfaces sliding past each other while immersed in a 0.01 M KCl salt solution (nonadhesive conditions). The water film is molecularly thin, 0.25–0.5 nm thick, and the interfacial friction force is very low: $S_c \approx 5 \times 10^5 \text{ N m}^{-2}$, $\mu \approx 0.02$ (before damage occurs). After the surfaces have become damaged, the friction coefficient is about 0.3 (after (Homola et al. 1990), © 1990, with permission from Elsevier Science). **b** Steady-state friction force and contact area measured on a confined squalane film between two undamaged mica surfaces at $v = 0.6 \mu\text{m/s}$ in the smooth sliding regime (no stick-slip). Open circles show F obtained on loading (increasing L), solid circles show unloading. Both data sets are straight lines passing through the origin, as shown by the brown line ($\mu = 0.12$). The black curve is a fit of the Hertz equation (Sect. 9.5.2 and Israelachvili 2011) to the A versus L data (open squares) using $K = 10^{10} \text{ N/m}^2$, $R = 2 \text{ cm}$. The thickness D varies monotonically from $D = 2.5$ to 1.7 nm as the load increases from $L = 0$ to 10 mN (adapted from (Gourdon and Israelachvili 2003), © 2003 American Physical Society)

value that corresponds to the unusually low friction of ice. Clearly, a single monolayer of water can be a very good lubricant—much better than most other monomolecular liquid films—for reasons that will be discussed in Sect. 9.9. A linear dependence of F on L has also been observed for mica surfaces separated by certain

hydrocarbon liquids (Gee et al. 1990; Gourdon and Israelachvili 2003). Figure 9.18b shows the kinetic friction forces measured at a high velocity across thin films of squalane, a branched hydrocarbon liquid ($C_{30}H_{62}$), which is a model for lubricating oils. Very low adhesive forces are measured between mica surfaces across this

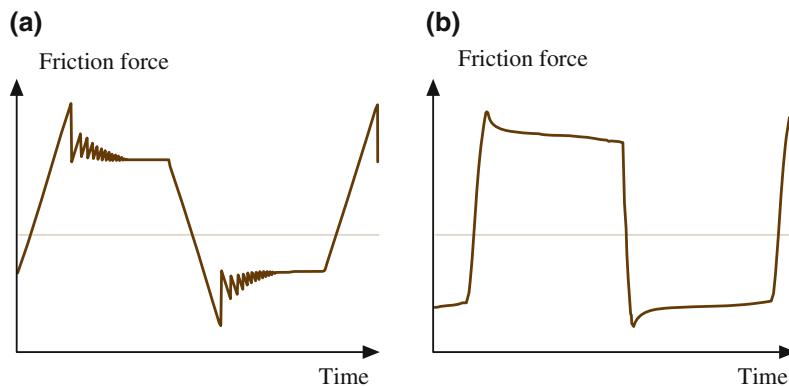


Fig. 9.19 Typical friction traces showing how the friction force varies with the sliding time for two symmetric, glassy polymer films under dry conditions. Qualitative features that are common to both polystyrene and polyvinyl benzyl chloride: **a** Decaying stick-slip motion is observed until smooth sliding is attained if the motion continues for a sufficiently long distance. **b** Smooth sliding observed at sufficiently high speeds. Similar observations have been made (Berthoud et al. 1999) in measurements on polymethyl methacrylate (after (Maeda et al. 2002), with permission, © 2002 American Association for the Advancement of Science)

liquid and the film thickness decreased monotonically with load. The friction force at a given load was found to be velocity-dependent, whereas the contact area was not (Gourdon and Israelachvili 2003).

Dry polymer layers (Fig. 9.19) typically show a high initial static friction (stiction) as sliding commences from rest in adhesive contact. The development of the friction force after a change in sliding direction, a gradual transition from stick-slip to smooth sliding, is shown in Fig. 9.19. As shown in Fig. 9.11b,c, the adhesion hysteresis for polystyrene surfaces can be increased by irradiation to induce scission of chains, and it has been found that the steady-state friction force (kinetic friction) shows a similar increase with irradiation time (Maeda et al. 2002).

9.7.4 Transition from Interfacial to Normal Friction with Wear

Frictional damage can have many causes, such as adhesive tearing at high loads or overheating at high sliding speeds. Once damage occurs, there is a transition from interfacial to normal or load-controlled friction as the surfaces become forced apart by the torn-out asperities (wear particles). For

low loads, the friction changes from obeying $F = S_c A$ to obeying Amontons' law, $F = \mu L$, as shown in Figs. 9.15 and 9.16, and sliding now proceeds smoothly with the surfaces separated by a 10–100 nm forest of wear debris (in this case, mica flakes). The wear particles keep the surfaces apart over an area that is much greater than their size, so that even one submicroscopic particle or asperity can cause a significant reduction in the area of contact and, therefore, in the friction (Homola et al. 1990). For this type of frictional sliding, one can no longer talk of the molecular contact area of the two surfaces, although the macroscopic or apparent area is still a useful parameter.

One remarkable feature of the transition from interfacial to normal friction of brittle surfaces is that, while the strength of interfacial friction, as reflected in the values of S_c , is very dependent on the type of surface and on the liquid film between the surfaces, this is not the case once the transition to normal friction has occurred. At the onset of damage, the material properties of the underlying substrates control the friction. In Figs. 9.15, 9.16 and 9.18a the friction for the damaged surfaces is that of any damaged mica-mica system, $\mu \approx 0.3$, *independent of the initial surface coatings or liquid films between the surfaces*. A similar friction coefficient was found for damaged silica surfaces (Vigil et al. 1994).

Clearly, the mechanism and factors that determine normal friction are quite different from those that govern interfacial friction (Sects. 9.7.1 and 9.7.2). This effect is not general and may only apply to brittle materials. For example, the friction of ductile surfaces is totally different and involves the continuous plastic deformation of contacting surface asperities during sliding, rather than the rolling of two surfaces on hard wear particles (Bowden and Tabor 1971). Furthermore, in the case of ductile surfaces, water and other surface-active components do have an effect on the friction coefficients under normal sliding conditions.

9.8 Liquid Lubricated Surfaces

9.8.1 Viscous Forces and Friction of Thick Films: Continuum Regime

Experimentally, it is usually difficult to unambiguously establish which type of sliding mode is occurring, but an empirical criterion, based on the Stribeck curve (Fig. 9.12), is often used as an indicator. This curve shows how the friction force or the coefficient of friction is expected to vary with sliding speed, depending on which type of friction regime is operating. For thick liquid lubricant films whose behavior can be described by bulk (continuum) properties, the friction forces are essentially the hydrodynamic or viscous drag forces. For example, for two plane parallel surfaces of area A separated by a distance D and moving laterally relative to each other with velocity v , if the intervening liquid is Newtonian, i.e., if its viscosity η is independent of the shear rate, the frictional force experienced by the surfaces is given by

$$F = \frac{\eta Av}{D}, \quad (9.43)$$

where the shear rate $\dot{\gamma}$ is defined by

$$\dot{\gamma} = \frac{v}{D}. \quad (9.44)$$

At higher shear rates, two additional effects often come into play. First, certain properties of liquids may change at high $\dot{\gamma}$ values. In particular, the effective viscosity may become non-Newtonian, one form given by

$$\eta \propto \dot{\gamma}^n, \quad (9.45)$$

where $n = 0$ (i.e., $\eta_{\text{eff}} = \text{constant}$) for Newtonian fluids, $n > 0$ for shear-thickening (dilatant) fluids, and $n < 0$ for shear-thinning (pseudo-plastic) fluids (the latter become less viscous, i.e., flow more easily, with increasing shear rate). An additional effect on η can arise from the higher local stresses (pressures) experienced by the liquid film as $\dot{\gamma}$ increases. Since the viscosity is generally also sensitive to the pressure (usually increasing with P), this effect also acts to increase η_{eff} and thus the friction force.

A second effect that occurs at high shear rates is surface deformation, arising from the large hydrodynamic forces acting on the sliding surfaces. For example, Fig. 9.20 shows how two surfaces deform elastically when the sliding speed increases to a high value. These deformations alter the hydrodynamic friction forces, and this type of friction is often referred to as elastohydrodynamic lubrication (EHD or EHL), as mentioned in Table 9.4.

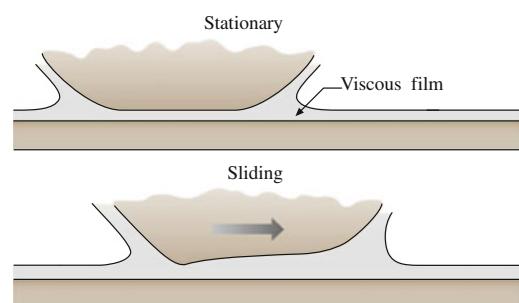


Fig. 9.20 Top: Stationary surfaces (one more deformable and one rigid) separated by a thick liquid film. Bottom: Elastohydrodynamic deformation of the upper surface during sliding (after (Israelachvili 1995), with permission)

How thin can a liquid film be before its dynamic, e.g., viscous flow, behavior ceases to be described by bulk properties and continuum models? Concerning the static properties, we have already seen in Sect. 9.4.3 that films composed of simple liquids display continuum behavior down to thicknesses of 4–10 molecular diameters. Similar effects have been found to apply to the dynamic properties, such as the viscosity, of simple liquids in thin films. Previous studies (Israelachvili 1986; Dhinojwala and Granick 1997; Chan and Horn 1985; Israelachvili and Kott 1989) show that, for simple liquids including linear chain molecules such as alkanes, the viscosity in thin films is the same, within 10%, as the bulk even for films as thin as 10 molecular diameters (or segment widths). This implies that the shear plane is effectively located within one molecular diameter of the solid–liquid interface, and these conclusions were found to remain valid even at the highest shear rates studied (of $\sim 2 \times 10^5 \text{ s}^{-1}$). These results appear to be independent of the existence of electrostatic double-layer or hydration forces. For the case of the simple liquid toluene confined between surfaces with adsorbed layers of C_{60} molecules, this type of viscosity measurement has shown that the traditional no-slip assumption for flow at a solid interface does not always hold (Campbell et al. 1996). The C_{60} layer at the mica–toluene interface results in a full-slip boundary, which dramatically lowers the viscous drag or effective viscosity for regular Couette or Poiseuille flow.

With polymeric liquids (polymer melts) such as polydimethylsiloxanes (PDMS) and polybutadienes (PBD), or with polystyrene (PS) adsorbed onto surfaces from solution, the far-field viscosity is again equal to the bulk value, but with the nonslip plane (hydrodynamic layer thickness) being located at $D = 1\text{--}2R_g$ away from each surface (Israelachvili 1986; Luengo et al. 1997), or at $D = L$ or less for polymer brush layers of thickness L per surface (Dhinojwala and Granick 1997; Klein et al. 1993). In contrast, the same technique was used to show that, for non-adsorbing polymers in solution, there is actually a depletion layer of nearly pure solvent that exists at the surfaces that affects the confined

solution flow properties (Kuhl et al. 1998). These effects are observed from near contact to surface separations in excess of 200 nm.

Further experiments with surfaces closer than a few molecular diameters ($D < 2\text{--}4 \text{ nm}$ for simple liquids, or $D < 2\text{--}4R_g$ for polymer fluids) indicate that large deviations occur for thinner films, described below. One important conclusion from these studies is, therefore, that the dynamic properties of simple liquids, including water, near an isolated surface are similar to those of the bulk *liquid already within the first layer of molecules adjacent to the surface*, only changes when another surface approaches the first due to the constraining effects on the liquid molecules.

9.8.2 Friction of Intermediate Thickness Films

For liquid films in the thickness range between 4 and 10 molecular diameters, the properties can be significantly different from those of bulk films. Still, the fluids do not undergo a phase transition into a solid or liquid-crystalline phase. Films of simple liquids have been found to become non-Newtonian in the 2.5–5 nm regime (about 10 molecular diameters, Fig. 9.21) (Granick 1991) while polymer melts are known to become non-Newtonian at much larger film thicknesses, depending on their molecular weight (Luengo et al. 1997).

Based on studies of interaction forces and friction of small quasi-spherical liquid molecules such as cyclohexane between molecularly smooth mica surfaces (Klein and Kumacheva 1995, 1998), it has been concluded that surface epitaxial effects can cause the liquid film to solidify already at six molecular diameters, resulting in a sudden (discontinuous) jump to high friction at low shear rates. Such dynamic first-order transitions, however, may depend on the shear rate.

A generalized friction map (Fig. 9.21c, d) has been proposed (Luengo et al. 1996) that illustrates the changes in η_{eff} from bulk Newtonian behavior ($n = 0$, $\eta_{\text{eff}} = \eta_{\text{bulk}}$) through the transition regime where n reaches a minimum of -1

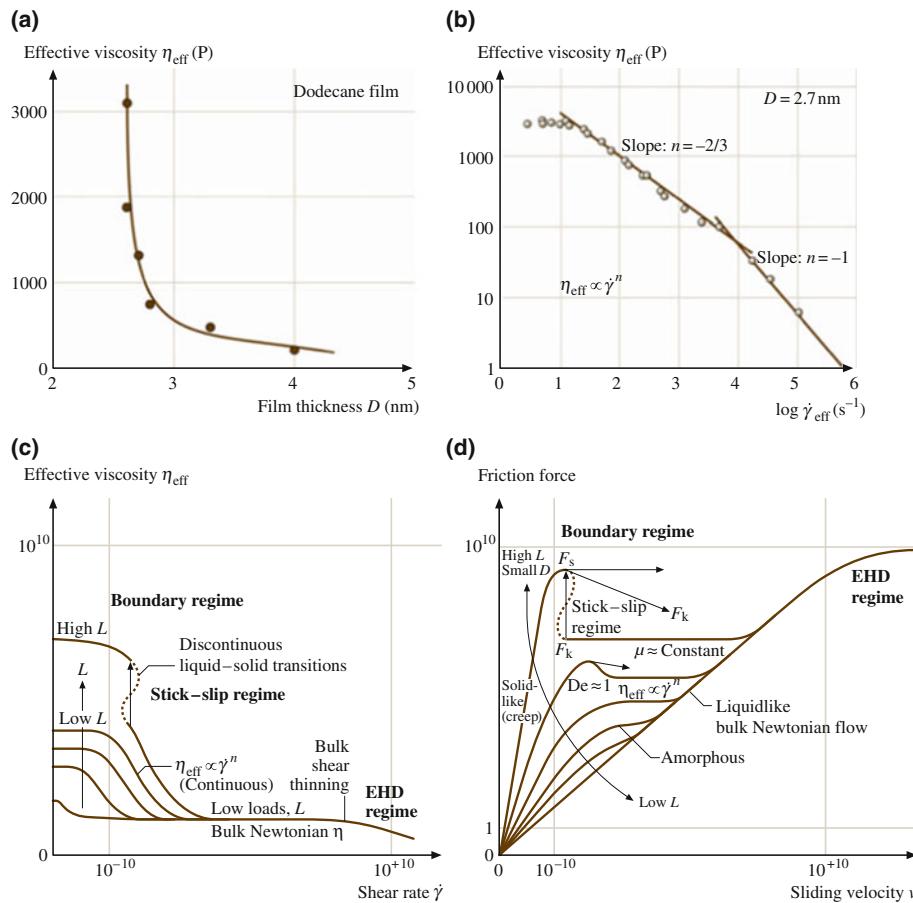


Fig. 9.21 Typical rheological behavior of liquid films in the mixed lubrication regime. **a** Increase in effective viscosity of dodecane film between two mica surfaces with decreasing film thickness. At distances larger than 4–5 nm, the effective viscosity η_{eff} approaches the bulk value η_{bulk} and does not depend on the shear rate γ (after (Granick 1991), © 1991 American Association for the Advancement of Science.). **b** Non-Newtonian variation of η_{eff} with shear rate of a 2.7 nm thick dodecane film at a net normal pressure of 0.12 MPa and at 28 °C. The effective viscosity decays as a power law, as in (9.45). In this example, $n = 0$ at the lowest γ and changes to $n = -2/3$ and -1 at higher γ . For films of bulk thickness, dodecane is a low-viscosity Newtonian fluid ($n = 0$). **c** Proposed general friction map of effective viscosity η_{eff} (arbitrary units) as a function of effective shear rate $\dot{\gamma}$ (arbitrary units) on logarithmic scales. Three main classes of behavior emerge: (i) Thick films elastohydrodynamic sliding. At $L = 0$, approximating bulk conditions, η_{eff} is independent of shear rate except when shear thinning might occur at sufficiently large $\dot{\gamma}$. (ii) Boundary layer films, intermediate regime. A Newtonian regime is again observed ($\eta_{\text{eff}} = \text{constant}$, $n = 0$ in (9.45)) at low loads and low shear rates, but η_{eff} is much higher than the bulk value. As the shear rate $\dot{\gamma}$ increases beyond $\dot{\gamma}_{\min}$, the effective viscosity starts to drop with a power-law dependence on the shear rate (b), with n in the range $-1/2$ to -1 most commonly observed. As the shear rate $\dot{\gamma}$ increases still more, beyond $\dot{\gamma}_{\max}$, a second Newtonian plateau is encountered. (iii) Boundary layer films, high load. The η_{eff} continues to grow with load and to be Newtonian provided that the shear rate is sufficiently low. Transition to sliding at high velocity is discontinuous ($n < -1$) and usually of the stick-slip variety. **d** Proposed friction map of friction force as a function of sliding velocity in various tribological regimes. With increasing load, Newtonian flow in the elastohydrodynamic (EHD) regimes crosses into the boundary regime of lubrication. Note that even EHD lubrication changes, at the highest velocities, to limiting shear stress response. At the highest loads (L) and smallest film thickness (D), the friction force goes through a maximum (the static friction F_s) followed by a regime where the friction coefficient (μ) is roughly constant with increasing velocity (meaning that the kinetic friction, F_k , is roughly constant). Non-Newtonian shear thinning is observed at somewhat smaller load and larger film thickness; the friction force passes through a maximum at the point where $De = 1$. De —the Deborah number—is the point at which the applied shear rate exceeds the natural relaxation time of the boundary layer film. The velocity axis from 10^{-10} to 10^{10} (arbitrary units) indicates a large span (Panels (b–d) after (Luengo et al. 1996), © 1996, with permission from Elsevier Science)

with decreasing shear rate to the solid-like creep regime at very low γ where n returns to 0. A number of results from experimental, theoretical, and computer simulation work have shown values of n from $-1/2$ to -1 for this transition regime for a variety of systems and assumptions (Granick 1991; Hu and Granick 1992; Thompson et al. 1992; Subbotin et al. 1995a,b).

The intermediate regime appears to extend over a narrow range of film thickness, from about 4 to 10 molecular diameters or polymer radii of gyration. Thinner films begin to adopt boundary or interfacial friction properties (described below, see also Table 9.5). Note that the intermediate regime is actually a very narrow one when defined in terms of film thickness, for example, varying from about $D = 2$ to 4 nm for hexadecane films (Granick 1991).

A fluid's effective viscosity η_{eff} in the intermediate regime is usually higher than in the bulk, but η_{eff} usually decreases with increasing sliding velocity, v (known as shear thinning). When two surfaces slide in the intermediate regime, the motion tends to thicken the film (dilatancy). This sends the system into the bulk EHD regime where, as indicated by (9.43), the friction force now increases with velocity. This initial

decrease, followed by an increase, in the frictional forces of many lubricant systems is the basis for the empirical Stribeck curve of Fig. 9.12. In the transition from bulk to boundary behavior there is first a quantitative change in the material properties (viscosity and elasticity), which can be continuous, to discontinuous qualitative changes that result in yield stresses and non-liquid-like behavior.

9.8.3 Boundary Lubrication of Molecularly Thin Films: Nanorheology

When a liquid is confined between two surfaces or within any narrow space whose dimensions are less than 4–10 molecular diameters, both the static (equilibrium) and dynamic properties of the liquid, such as its compressibility and viscosity, can no longer be described even qualitatively in terms of the bulk properties. The molecules confined within such molecularly thin films become ordered into layers (out-of-plane ordering), and within each layer they can also have lateral order (in-plane ordering). Such films may be thought of as behaving more like a liquid

Table 9.5 Effect of molecular shape and short-range forces on tribological properties for molecularly thin liquid films between two shearing mica surfaces at 20 °C

Liquid	Short-range force	Type of friction	Friction coefficient	Bulk liquid viscosity (cP)
<i>Organic (water-free)</i>				
Cyclohexane	Oscillatory	Quantized stick-slip	$\gg 1$	0.6
OMCTS ^a	Oscillatory	Quantized stick-slip	$\gg 1$	2.3
Octane	Oscillatory	Quantized stick-slip	1.5	0.5
Tetradecane	Oscillatory \leftrightarrow smooth	Stick-slip \leftrightarrow smooth	1.0	2.3
Octadecane (branched)	Oscillatory \leftrightarrow smooth	Stick-slip \leftrightarrow smooth	0.3	5.5
PDMS ^a ($M = 3700 \text{ g mol}^{-1}$, melt)	Oscillatory \leftrightarrow smooth	Smooth	0.4	50
PBD ^a ($M = 3500 \text{ g mol}^{-1}$, branched)	Smooth	Smooth	0.03	800
<i>Water</i>				
Water (KCl solution)	Smooth	Smooth	0.01–0.03	1.0

^aOMCTS Octamethylcyclotetrasiloxane, PDMS Polydimethylsiloxane, PBD Polybutadiene

crystal or a solid than a liquid and the structuring of molecules in thin films and the oscillatory forces it gives rise to (Sect. 9.4.3) are now reasonably well understood, both experimentally and theoretically, at least for simple liquids.

Work has also recently been done on the dynamic, e.g., viscous or shear, forces associated with molecularly thin films. Both experiments (Hirz et al. 1992; Klein and Kumacheva 1995, 1998; Ruths and Granick 2000; Gee et al. 1990; Demirel and Granick 1996; Israelachvili et al. 1988) and theory (Schoen et al. 1989; Thompson and Robbins 1990; Thompson et al. 1992; Schoen et al. 1995) indicate that, even when two surfaces are in steady-state sliding, they still prefer to remain in one of their stable potential-energy minima, i.e., a sheared film of liquid can retain its basic layered structure. Thus even during motion the film does not become totally liquid-like. Indeed, if there is some in-plane ordering within a film, it will exhibit a yield point before it begins to flow. Such films can therefore sustain a finite shear stress, in addition to a finite normal stress. The value of the yield stress depends on the number of layers comprising the film and represents another quantized property of molecularly thin films (Schoen et al. 1989).

The dynamic properties of a liquid film undergoing shear are very complex. Depending on whether the film is more liquid-like or solid-like, the motion will be smooth or of the stick-slip type illustrated schematically in Fig. 9.22. During sliding, transitions can occur between n layers and $(n - 1)$ or $(n + 1)$ layers. The details of the motion depend critically on the externally applied load, the temperature, the sliding velocity, the twist angle between the two surface lattices, and the sliding direction relative to the lattices.

9.8.3.1 Smooth and Stick-Slip Sliding

Recent advances in friction-measuring techniques have enabled the interfacial friction of molecularly thin films to be measured with great accuracy. Some of these advances have involved the surface forces apparatus technique (Hirz et al. 1992; Homola et al. 1989; Van Alsten and

Granick 1990; Klein and Kumacheva 1995, 1998; Luengo et al. 1997; Hu and Granick 1992; Gee et al. 1990; Demirel and Granick 1996; Gourdon and Israelachvili 2003; Ruths et al. 2003; Dhinojwala et al. 2000; Qian et al. 2003; Homola et al. 1990; Raviv et al. 2004; Israelachvili et al. 1988; Heuberger et al. 1998). In addition, computer simulations (Gao et al. 2000, 2004; Schoen et al. 1989; Thompson and Robbins 1990; Schoen et al. 1995; Zaloj et al. 1999) have become sufficiently sophisticated to enable fairly complex tribological systems to be studied. All these advances are necessary if one is to probe such subtle effects as smooth or stick-slip friction, transient and memory effects, and ultralow friction mechanisms at the molecular level.

Stick-slip occurs between a static state with high friction and a low-friction kinetic state, and a transition from this sliding regime to smooth sliding can be induced by an increase in velocity. Experimental data on various systems showing this behavior are shown below. Recent studies on adhesive systems have revealed the possibility of other dynamic responses such as inverted stick-slip between two kinetic states of higher and lower friction and with a transition from smooth sliding to stick-slip with increasing velocity (Drummond et al. 2002). Similar friction responses have also been seen in computer simulations (Filippov et al. 2002).

With the added insights provided by computer simulations, a number of distinct molecular processes have been identified during smooth and stick-slip sliding in model systems for the more familiar static-to-kinetic stick-slip and transition from stick-slip to smooth sliding. These are shown schematically in Fig. 9.23 for the case of spherical liquid molecules between two solid crystalline surfaces. The following regimes may be identified:

Surfaces at rest (Fig. 9.23a): even with no externally applied load, solvent–surface epitaxial interactions can cause the liquid molecules in the film to attain a solid-like state. Thus at rest the surfaces are stuck to each other through the film.

Sticking regime (frozen, solid-like film) (Fig. 9.23b): a progressively increasing lateral

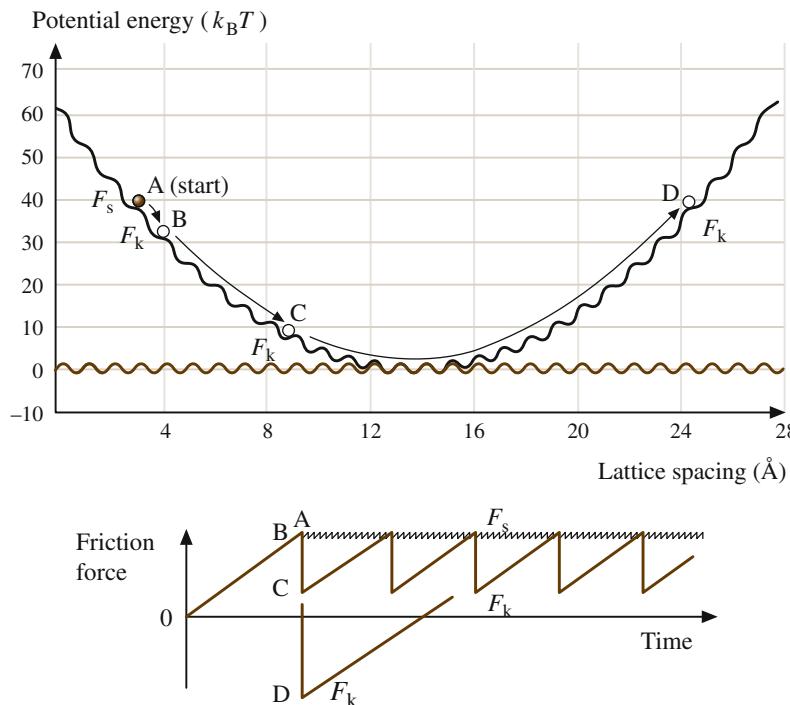


Fig. 9.22 Simple schematic illustration of the most common molecular mechanism leading from smooth to stick-slip sliding in terms of the efficiency of the energy transfer from mechanical to kinetic to phonons. The potential energy of the corrugated surface lattice is shown by the horizontal sine wave. Let the depth of each minimum be ε which is typically $>k_B T$. At equilibrium, a molecule will sit at one of these minima. When the molecule is connected to a horizontal spring, a smooth parabolic curve must be added to the horizontal curve. If this spring is now pushed or pulled laterally at a constant velocity v , the sine curve will move like a wave along the parabola carrying the molecule up with it (towards point A). When the point of inflection at A is reached the molecule will drop and acquire a kinetic energy greater than ε even before it reaches the next lattice site. This energy can be *released* at the next lattice site (i.e., on the first collision), in which case the processes will now be repeated—each time the molecule reaches point A it will fall to point B. This type of motion will give rise to periodic changes in temperature at the interface, as predicted by computer simulations (Xie et al. 2002). The stick-slip here will have a magnitude of the lattice dimension and, the measured macro- and microscopic friction forces will be smooth and independent of v . If the energy dissipation (or *transfer*) mechanism is less than 100% efficient on each collision, the molecule will move further before it stops. In this case the stick-slip amplitude can be large (point C), and the kinetic friction F_k can even be negative in the case of an overshoot (point D) (after (Gao et al. 2004), with permission, © 2004 American Chemical Society)

shear stress is applied. The solid-like film responds elastically with a small lateral displacement and a small increase or dilatancy in film thickness (less than a lattice spacing or molecular dimension σ). In this regime the film retains its frozen, solid-like state: all the strains are elastic and reversible, and the surfaces remain effectively stuck to each other. However, slow creep may occur over long time periods.

Slipping and sliding regimes (molten, liquid-like film) (Fig. 9.23c–c''): when the applied shear

stress or force has reached a certain critical value, the static friction force F_s the film suddenly melts (known as shear melting) or rearranges to allow for wall slip or slip within the film to occur at which point the two surfaces begin to slip rapidly past each other. If the applied stress is kept at a high value, the upper surface will continue to slide indefinitely.

Refreezing regime (resolidification of film) (Fig. 9.23d): In many practical cases, the rapid slip of the upper surface relieves some of the

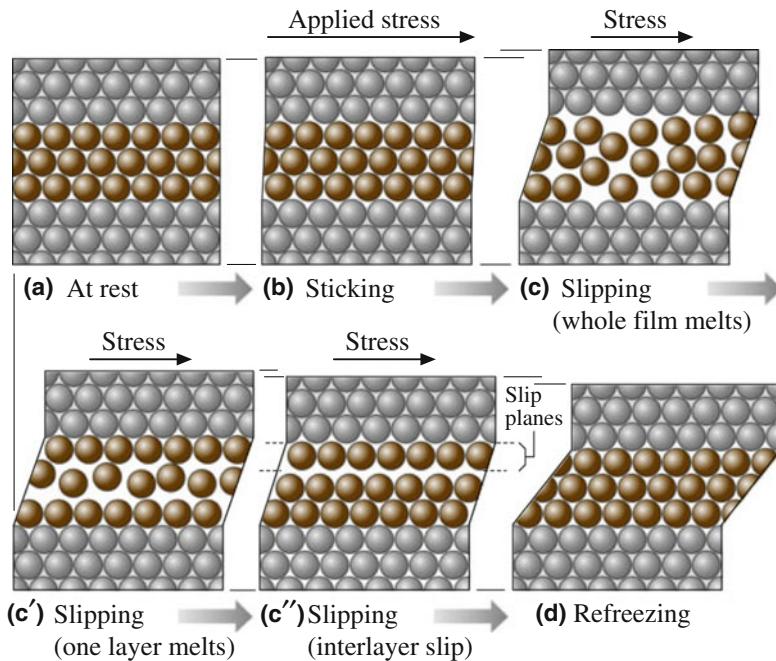


Fig. 9.23 Idealized schematic illustration of molecular rearrangements occurring in a molecularly thin film of spherical or simple chain molecules between two solid surfaces during shear. Depending on the system, a number of different molecular configurations within the film are possible during slipping and sliding, shown here as stages (c): total disorder as the whole film melts; (c'): partial disorder; and (c''): order persists even during sliding with slip occurring at a single slip plane either within the film or at the walls. A dilation is predicted in the direction normal to the surfaces (after Israelachvili et al. 1990), with permission)

applied force, which eventually falls below another critical value, the kinetic friction force F_k , at which point the film resolidifies and the whole stick-slip cycle is repeated. On the other hand, if the slip rate is smaller than the rate at which the external stress is applied, the surfaces will continue to slide smoothly in the kinetic state and there will be no more stick-slip.

Experiments on the friction of two molecularly smooth mica surfaces separated by three molecular layers of the liquid octamethylcyclotetrasiloxane (OMCTS, Fig. 9.24) show how the friction increases to higher values in a quantized way when the number of layers falls from $n = 3$ to $n = 2$ and then to $n = 1$. Computer simulations for simple spherical molecules (Thompson and Robbins 1990) further indicate that during slip the film thickness is roughly 15% higher than at rest (i.e., the film density falls), and that the order parameter within the film drops

from 0.85 to about 0.25. Such dilatancy has been investigated both experimentally (Dhinojwala et al. 2000) and in further computer simulations (Zaloj et al. 1999). The changes in thickness and in the order parameter are consistent with a disorganized state for the whole film during the slip (Schoen et al. 1995), as illustrated schematically in Fig. 9.23c. At this stage, we can only speculate on other possible configurations of molecules in the slipping and sliding regimes. This probably depends on the shapes of the molecules (e.g., whether they are spherical or linear or branched), on the atomic structure of the surfaces, on the sliding velocity, etc. (Drummond and Israelachvili 2001). Figure 9.23 shows three possible sliding modes wherein the shearing film either totally melts, or where the molecules retain their layered structure and where slip occurs between two or more layers. Other sliding modes, for example, involving the movement of

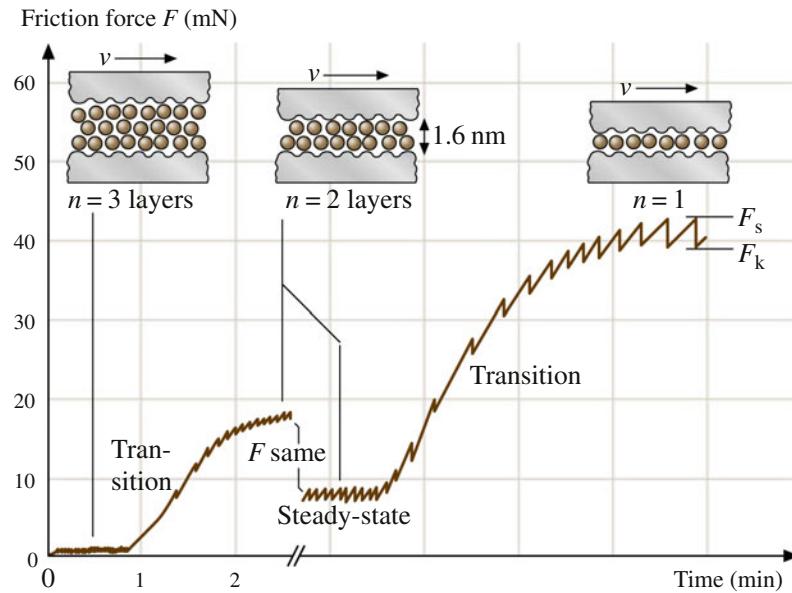


Fig. 9.24 Measured change in friction during interlayer transitions of the silicone liquid octamethylcyclotetrasiloxane (OMCTS), an inert liquid whose quasi-spherical molecules have a diameter of 0.8 nm. In this system, the shear stress $S_c = F/A$ was found to be constant as long as the number of layers, n , remained constant. Qualitatively similar results have been obtained with other quasi-spherical molecules such as cyclohexane (Israelachvili et al. 1988). The shear stresses are only weakly dependent on the sliding velocity v . However, for sliding velocities above some critical value, v_c , the stick-slip disappears and sliding proceeds smoothly at the kinetic value (after (Gee et al. 1990), with permission)

dislocations or disclinations are also possible, and it is unlikely that one single mechanism applies in all cases.

Both friction and adhesion hysteresis vary nonlinearly with temperature, often peaking at some particular temperature T_0 . The temperature dependence of these forces can, therefore, be represented on a friction phase diagram such as the one shown in Fig. 9.25. Experiments have shown that T_0 , and the whole bell-shaped curve, are shifted along the temperature axis (as well as in the vertical direction) in a systematic way when the load, sliding velocity, etc., are varied. These shifts also appear to be highly correlated with one another.

Such effects are also commonly observed in other energy-dissipating phenomena such as polymer viscoelasticity (Ferry 1980). A possible molecular process underlying the energy

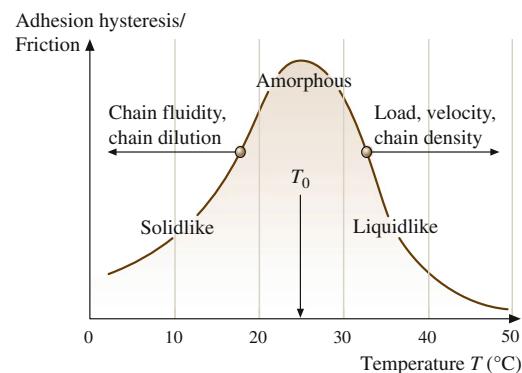


Fig. 9.25 Schematic friction phase diagram representing the trends observed in the boundary friction of a variety of different surfactant monolayers. The characteristic *bell-shaped curve* also correlates with the adhesion energy hysteresis of the monolayers. The arrows indicate the direction in which the whole curve is dragged when the load, velocity, etc., is increased (after (Israelachvili et al. 1994), with permission)

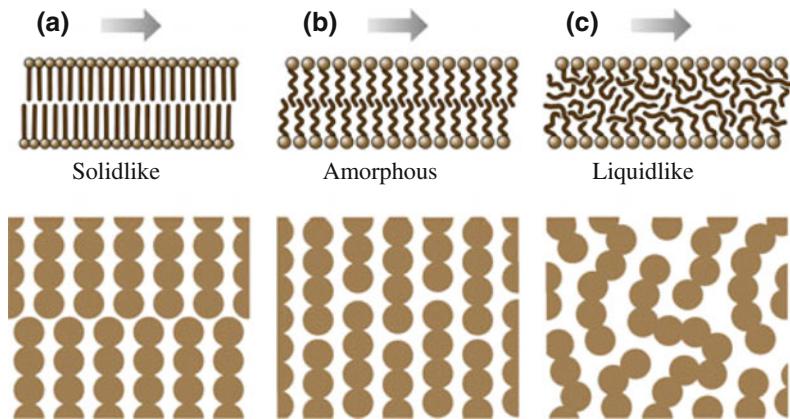


Fig. 9.26 Different dynamic phase states of boundary monolayers during adhesive contact and/or frictional sliding. Solid-like (a) and liquid-like monolayers (c) exhibit low adhesion hysteresis and friction. Increasing the temperature generally shifts a system from the left to the right. Changing the load, sliding velocity, or other experimental conditions can also change the dynamic phase state of surface layers, as shown in Fig. 9.25 (after (Israelachvili et al. 1994), with permission)

dissipation of chain molecules during boundary-layer sliding is illustrated in Fig. 9.26, which shows the three main dynamic phase states of boundary monolayers.

9.8.3.2 Abrupt Versus Continuous Transitions Between Smooth and Stick-Slip Sliding

An understanding of stick-slip is of great practical importance in tribology (Rabinowicz 1995), since these spikes are the major cause of damage and wear of moving parts. Stick-slip motion is a very common phenomenon and is also the cause of sound generation (the sound of a violin string, a squeaking door, or the chatter of machinery), sensory perception (taste texture and feel), earthquakes, granular flow, nonuniform fluid flow such as the spouting flow of polymeric liquids, etc. In the previous section, the stick-slip motion arising from freezing–melting transitions in thin interfacial films was described. There are other mechanisms that can give rise to stick-slip friction, which will now be considered.

9.8.3.3 Rough Surfaces or Surface Topology Model

Rapid slips can occur whenever an asperity on one surface goes over the top of an asperity on

the other surface. The extent of the slip will depend on asperity heights and slopes, on the speed of sliding, and on the elastic compliance of the surfaces and the moving stage. As in all cases of stick-slip motion, the driving velocity v may be constant, but the resulting motion at the surfaces v_0 will display large slips. This type of stick-slip (Rabinowicz 1995) will not be of much concern here since it is essentially a noise-type fluctuation, resulting from surface imperfections rather than from the intrinsic interaction between two surfaces.

9.8.3.4 Distance-Dependent or Creep Model

Another theory of stick-slip, observed in solid-on-solid sliding, is one that involves a characteristic distance, but also a characteristic time τ_s this being the characteristic time required for two asperities to increase their adhesion strength after coming into contact. Originally proposed by Rabinowicz (Rabinowicz 1995, 1958), this model suggests that two rough macroscopic surfaces adhere through their microscopic asperities of a characteristic length. During shearing, each surface must first creep this distance (the size of the contacting junctions) after which the surfaces continue to slide, but with a lower (kinetic) friction force than the

original (static) value. The reason for the decrease in the friction force is that even though, on average, new asperity junctions should form as rapidly as the old ones break, the time-dependent adhesion and friction of the new ones will be lower than the old ones.

The friction force, therefore, remains high during the creep stage of the slip. However, once the surfaces have moved the characteristic distance, the friction rapidly drops to the kinetic value. For any system where the kinetic friction is less than the static force (or one that has a negative slope over some part of its curve of F_0 versus v_0) will exhibit regular stick-slip sliding motion for certain values of K , m , and driving velocity v .

This type of friction has been observed in a variety of dry (unlubricated) systems such as paper-on-paper (Heslot et al. 1994) and steel-on-steel (Rabinowicz 1958; Heymann et al. 1955). This model is also used extensively in geologic systems to analyze rock-on-rock sliding (Dieterich 1978). While originally described for adhering macroscopic asperity junctions, the distance-dependent model may also apply to molecularly smooth surfaces. For example, for polymer lubricant films, the characteristic length would now be the chain-chain entanglement length, which could be much larger in a confined geometry than in the bulk.

9.8.3.5 Velocity-Dependent Friction Model

In contrast to the two friction models mentioned above, which apply mainly to unlubricated, solid-on-solid contacts, the stick-slip of surfaces with thin liquid films between them is better described by other mechanisms. The velocity-dependent friction model is the most studied mechanism of stick-slip and, until recently, was considered to be the only cause of intrinsic stick-slip. If a friction force decreases with increasing sliding velocity, as occurs with boundary films exhibiting shear thinning, the force (F_s) needed to initiate motion will be higher than the force (F_k) needed to maintain motion.

A decreasing intrinsic friction force F_0 with sliding velocity v_0 results in the sliding surface or

stage moving in a periodic fashion, where during each cycle rapid acceleration is followed by rapid deceleration. As long as the drive continues to move at a fixed velocity v , the surfaces will continue to move in a periodic fashion punctuated by abrupt stops and starts whose frequency and amplitude depend not only on the function $F_0(v_0)$, but also on the stiffness K and mass m of the moving stage, and on the starting conditions at $t = 0$.

More precisely, the motion of the sliding surface or stage can be determined by solving the following differential equation:

$$\begin{aligned} m\ddot{x} &= (F_0 - F) = F_0 - (x_0 - x)K \\ \text{or } m\ddot{x} + (x_0 - x)K - F_0 &= 0, \end{aligned} \quad (9.46)$$

where $F_0 = F_0(x_0, v_0, t)$ is the intrinsic or real friction force at the shearing surfaces, F is the force on the spring (the externally applied or measured force), and $(F_0 - F)$ is the force on the stage. To solve (9.46) fully, one must also know the initial (starting) conditions at $t = 0$, and the driving or steady-state conditions at finite t . Commonly, the driving condition is: $x = 0$ for $t < 0$ and $x = vt$ for $t > 0$, where $v = \text{constant}$. In other systems, the appropriate driving condition may be $F = \text{constant}$.

Various, mainly phenomenological, forms for $F_0 = F_0(x_0, v_0, t)$ have been proposed to explain various kinds of stick-slip phenomena. These models generally assume a particular functional form for the friction as a function of velocity only, $F_0 = F_0(v_0)$, and they may also contain a number of mechanically coupled elements comprising the stage (Tomlinson 1929; Carlson and Langer 1989). One version is a two-state model characterized by two friction forces, F_s and F_k , which is a simplified version of the phase transitions model described below. More complicated versions can have a rich $F-v$ spectrum (Persson 1994). Unless the experimental data is very detailed and extensive, these models cannot generally distinguish between different types of mechanisms. Neither do they address the basic question of the origin of the friction force, since this is assumed to begin with.

Experimental data has been used to calculate the friction force as a function of velocity within an individual stick-slip cycle (Nasuno et al. 1997). For a macroscopic granular material confined between solid surfaces, the data shows a velocity-weakening friction force during the first half of the slip. However, the data also shows a hysteresis loop in the friction–velocity plot, with a different behavior in the deceleration half of the slip phase. Similar results were observed for a 1–2 nm liquid lubricant film between mica surfaces (Drummond and Israelachvili 2001). These results indicate that a purely velocity-dependent friction law is insufficient to describe such systems, and an additional element such as the state of the confined material must be considered.

9.8.3.6 Phase Transitions Model

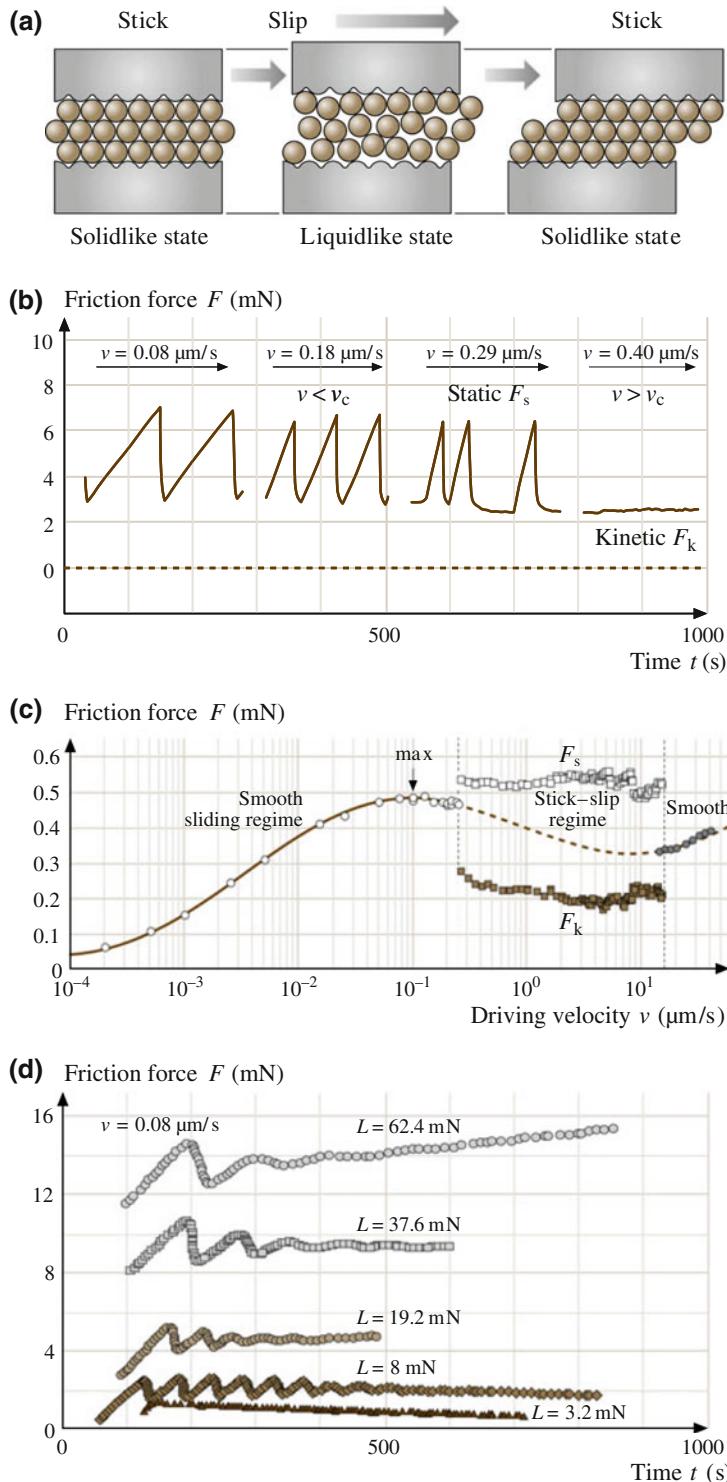
In molecular dynamics computer simulations it has been found that thin interfacial films undergo first-order phase transitions between solid-like and liquid-like states during sliding (Thompson and Robbins 1990; Robbins and Thompson 1991), as illustrated in Fig. 9.27. It has been suggested that this is responsible for the observed stick-slip behavior of simple isotropic liquids between two solid crystalline surfaces. With this interpretation, stick-slip is seen to arise because of the abrupt change in the flow properties of a film at a transition (Israelachvili et al. 1990a,b; Thompson et al. 1992), rather than the gradual or continuous change, as occurs in the previous example. Other computer simulations indicate that it is the stick-slip that induces a disorder (“shear melting”) in the film, not the other way around (Bordarier et al. 1998).

An interpretation of the well-known phenomenon of decreasing coefficient of friction with increasing sliding velocity has been proposed based on computer simulation (Thompson and Robbins 1990). This postulates that it is not the friction that changes with sliding speed v , but rather the time various parts of the system spend in the sticking and sliding modes. In other words, at any instant during sliding, the friction at any local region is always F_s or F_k , corresponding to the static or kinetic values. The measured frictional force, however, is the sum of all these

discrete values averaged over the whole contact area. Since as v increases each local region spends more time in the sliding regime (F_k) and less in the sticking regime (F_s), the overall friction coefficient falls. One may note that this interpretation reverses the traditional way that stick-slip has been explained. Rather than invoking a decreasing friction with velocity to explain stick-slip, it is now the more fundamental stick-slip phenomenon that is producing the apparent decrease in the friction force with increasing sliding velocity. This approach has been studied analytically (Carlson and Batista 1996), with a comprehensive rate- and state-dependent friction force law. This model includes an analytic description of the freezing–melting transitions of a film, resulting in a friction force that is a function of sliding velocity in a natural way. This model predicts a full range of stick-slip behavior observed experimentally.

An example of the rate- and state-dependent model is observed when shearing thin films of OMCTS between mica surfaces (Berman et al. 1996a,b). In this case, the static friction between the surfaces is dependent on the time that the surfaces are at rest with respect to each other, while the intrinsic kinetic friction $F_{k,0}$ is relatively constant over the range of velocities. At slow driving velocities, the system responds with stick-slip sliding with the surfaces reaching maximum static friction before each slip event, and the amplitude of the stick-slip, $F_s - F_k$, is relatively constant. As the driving velocity increases, the static friction decreases as the time at relative rest becomes shorter with respect to the characteristic time of the lubricant film. As the static friction decreases with increasing drive velocity, it eventually equals the intrinsic kinetic friction $F_{k,0}$, which defines the critical velocity v_c , above which the surfaces slide smoothly without the jerky stick-slip motion.

The above classifications of stick-slip are not exclusive, and molecular mechanisms of real systems may exhibit aspects of different models simultaneously. They do, however, provide a convenient classification of existing models and indicate which experimental parameters should be varied to test the different models.



◀ **Fig. 9.27** **a** Phase transitions model of stick-slip where a thin liquid film alternately freezes and melts as it shears, shown here for 22 spherical molecules confined between two solid crystalline surfaces. In contrast to the velocity-dependent friction model, the intrinsic friction force is assumed to change abruptly (at the transitions), rather than smooth or continuously. The resulting stick-slip is also different, for example, the peaks are sharper and the stick-slip disappears above some critical velocity v_c . Note that, while the slip displacement is here shown to be only two lattice spacings, in most practical situations it is much larger, and that freezing and melting transitions at surfaces or in thin films may not be the same as freezing or melting transitions between the bulk solid and liquid phases. **b** Exact reproduction of a chart-recorder trace of the friction force for hexadecane between two untreated mica surfaces at increasing sliding velocity v , plotted as a function of time. In general, with increasing sliding speed, the stick-slip spikes increase in frequency and decrease in magnitude. As the critical sliding velocity v_c is approached, the spikes become erratic, eventually disappearing altogether at $v = v_c$. At higher sliding velocities the sliding continues smoothly in the kinetic state. Such friction traces are fairly typical for simple liquid lubricants and dry boundary lubricant systems (Fig. 9.28a) and may be referred to as the *conventional* type of static-kinetic friction (in contrast to panel (c)). Experimental conditions: contact area $A = 4 \times 10^{-9} \text{ m}^2$, load $L = 10 \text{ mN}$, film thickness $D = 0.4\text{--}0.8 \text{ nm}$, $v = 0.08\text{--}0.4 \mu\text{m s}^{-1}$, $v_c \approx 0.3 \mu\text{m s}^{-1}$, atmosphere: dry N_2 gas, $T = 18^\circ\text{C}$ ((a, b) after (Yoshizawa and Israelachvili 1993) with permission, © 1993 American Chemical Society). **c** Transition from smooth sliding to *inverted* stick-slip and to a second smooth-sliding regime with increasing driving velocity during shear of two adsorbed surfactant monolayers in aqueous solution at a load of $L = 4.5 \text{ mN}$ and $T = 20^\circ\text{C}$. The smooth sliding (open circles) to inverted stick-slip (squares) transition occurs at $v_c \sim 0.3 \mu\text{m/s}$. Prior to the transition, the kinetic stress levels off at after a logarithmic dependence on velocity. The quasi-smooth regime persists up to the transition at v_c . At high driving velocities (filled circles), a new transition to a smooth sliding regime is observed between 14 and 17 $\mu\text{m/s}$ (after (Drummond et al. 2002) with permission). **d** Friction response of a thin squalane (a branched hydrocarbon) film at different loads and a constant sliding velocity $v = 0.08 \mu\text{m s}^{-1}$, slightly above the critical velocity for this system at low loads. Initially, with increasing load, the stick-slip amplitude and the mean friction force decrease with sliding time or sliding distance. However, at high loads or pressures, the mean friction force increases with time, and the stick-slip takes on a more symmetrical, sinusoidal shape. At all loads investigated, the stick-slip component gradually decayed as the friction proceeded towards smooth sliding (after (Gourdon and Israelachvili 2003) with permission, © 2003 American Physical Society)

9.8.3.7 Critical Velocity for Stick-Slip

For any given set of conditions, stick-slip disappears above some critical sliding velocity v_c (Fig. 9.28), above which the motion continues smoothly in the liquid-like or kinetic state (Yoshizawa et al. 1993; Gourdon and Israelachvili 2003; Drummond et al. 2002; Drummond and Israelachvili 2001; Yoshizawa and Israelachvili 1993). The critical velocity is well described by two simple equations. Both are based on the phase transition model, and both include some parameter associated with the inertia of the measuring instrument. The first equation is based on both experiments and simple theoretical modeling (Yoshizawa and Israelachvili 1993)

$$v_c = \frac{(F_s - F_k)}{5K\tau_0}, \quad (9.47)$$

where τ_0 is the characteristic nucleation time or freezing time of the film. For example, inserting

the following typically measured values for a $\sim 1 \text{ nm}$ thick hexadecane film between mica: $(F_s - F_k) \approx 5 \text{ mN}$, spring constant $K \approx 500 \text{ Nm}^{-1}$, and nucleation time (Yoshizawa and Israelachvili 1993) $\tau_0 \approx 5 \text{ s}$, we obtain $v_c \approx 0.4 \mu\text{m s}^{-1}$, which is close to typically measured values (Fig. 9.27b).

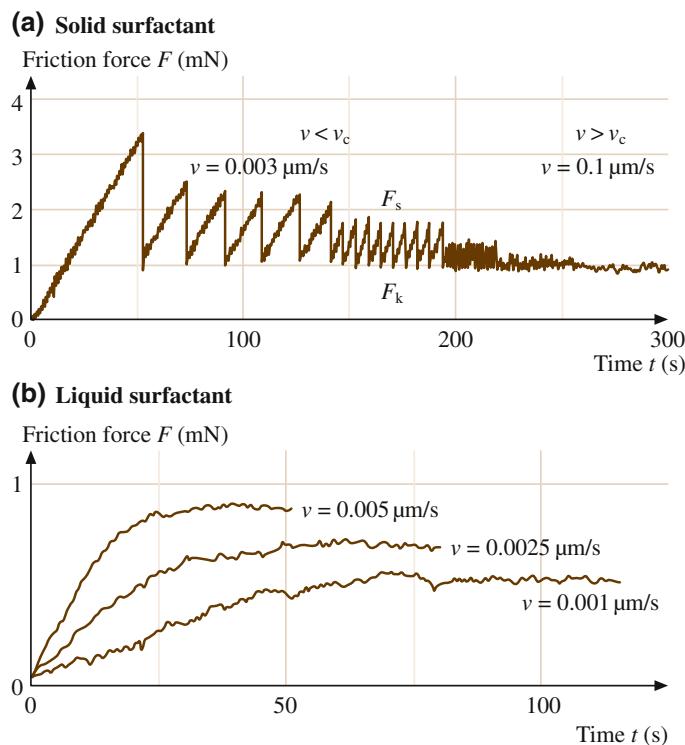
The second equation is based on computer simulations (Robbins and Thompson 1991):

$$v_c \approx 0.1 \sqrt{\frac{F_s \sigma}{m}}, \quad (9.48)$$

where σ is a molecular dimension and m is the mass of the stage. Again, inserting typical experimental values into this equation, viz., $m = 20 \text{ g}$, $\sigma \approx 0.5 \text{ nm}$, and $(F_s - F_k) \approx 5 \text{ mN}$ as before, we obtain $v_c \approx 0.3 \mu\text{m s}^{-1}$, which is also close to measured values.

Stick-slip also disappears above some critical temperature T_c , which is not the same as the melting temperature of the bulk fluid (Gourdon

Fig. 9.28 **a** Exact reproduction of chart-recorder trace for the friction of closely packed surfactant monolayers ($\text{l}-\alpha$ -dimirystoylphosphatidylethanolamine, DMPE) on mica (dry boundary friction) showing qualitatively similar behavior to that obtained with a liquid hexadecane film (Fig. 9.27b). In this case, $L = 0$, $v_c \approx 0.1 \mu\text{m s}^{-1}$, atmosphere: dry N_2 gas, $T = 25^\circ\text{C}$. **b** Sliding typical of liquid-like monolayers, here shown for calcium alkylbenzene sulfonate in dry N_2 gas at $T = 25^\circ\text{C}$ and $L = 0$ (after (Yoshizawa et al. 1993), with permission, © 1993 American Chemical Society)



and Israelachvili 2003). Certain correlations have been found between v_c and T_c and between various other tribological parameters that appear to be consistent with the principle of time–temperature superposition (Sect. 9.8.3), similar to that occurring in viscoelastic polymer fluids (Ferry 1980; McLaren and Tabor 1963; Gorsch 1963).

Work on the coupling between the mechanical resonances of the sliding system and molecular scale relaxations (Heuberger et al. 1998; Zaloj et al. 1999) has resulted in a better understanding of a phenomenon previously noted in various engineering applications: the vibration of one of the sliding surfaces perpendicularly to the sliding direction can lead to a significant reduction of the friction. At certain oscillation amplitudes and a frequency higher than the molecular-scale relaxation frequency, stick–slip friction can be eliminated and replaced by an ultralow kinetic-friction state.

9.9 Effects of Nanoscale Texture on Friction

The above scenario is already quite complicated, and yet this is the situation for the simplest type of experimental system. The factors that appear to determine the critical velocity v_c depend on the type of liquid between the surfaces, as well as on the surface lattice structure.

9.9.1 Role of the Shape of Confined Molecules

Small spherical molecules such as cyclohexane and OMCTS have been found to have very high v_c , which indicates that these molecules can rearrange relatively quickly in thin films. Chain molecules and especially branched-chain

molecules have been found to have much lower v_c , which is to be expected, and such liquids tend to slide smoothly or with erratic stick-slip (Drummond and Israelachvili 2001), rather than in a stick-slip fashion (Table 9.5). With highly asymmetric molecules, such as multiply branched isoparaffins and polymer melts, no regular spikes or stick-slip behavior occurs at any speed, since these molecules can never order themselves sufficiently to solidify. Examples of such liquids are some perfluoropolyethers and polydimethylsiloxanes (PDMS).

Computer simulations (Gao et al. 1997, 2000, 2004) of the structure, interaction forces, and tribological behavior of chain molecules between two shearing surfaces indicate that both linear and singly or doubly branched-chain molecules order between two flat surfaces by aligning into discrete layers parallel to the surfaces. However, in the case of the weakly branched molecules, the expected oscillatory forces do not appear because of a complex cancelation of entropic and enthalpic contributions to the interaction free energy, which results in a monotonically smooth interaction, exhibiting a weak energy minimum rather than the oscillatory force profile that is characteristic of linear molecules. During sliding, however, these molecules can be induced to further align, which can result in a transition from smooth to stick-slip sliding.

Table 9.5 shows the trends observed with some organic and polymeric liquid between smooth mica surfaces. Also listed are the bulk viscosities of the liquids. From the data of Table 9.5 it appears that there is a direct correlation between the shapes of molecules and their coefficient of friction or effectiveness as lubricants (at least at low shear rates). Small spherical or chain molecules have high friction with stick-slip, because they can pack into ordered solid-like layers. In contrast, longer chained and irregularly shaped molecules remain in an entangled, disordered, fluid-like state even in very thin films, and these give low friction and smoother sliding. It is probably for this reason that irregularly shaped branched chain molecules are usually better

lubricants. It is interesting to note that the friction coefficient generally decreases as the bulk viscosity of the liquids increases. This unexpected trend occurs because the factors that are conducive to low friction are generally conducive to high viscosity. Thus molecules with side groups such as branched alkanes and polymer melts usually have higher bulk viscosities than their linear homologues for obvious reasons. However, in thin films the linear molecules have higher shear stresses, because of their ability to become ordered. The only exception to the above correlations is water, which has been found to exhibit both low viscosity and low friction (Fig. 9.18a, and Sect. 9.7.3). In addition, the presence of water can drastically lower the friction and eliminate the stick-slip of hydrocarbon liquids when the sliding surfaces are hydrophilic.

If an “effective” viscosity, η_{eff} , were to be calculated for the liquids of Table 9.5, the values would be many orders of magnitude higher than those of the bulk liquids. This can be demonstrated by the following simple calculation based on the usual equation for Couette flow (9.43):

$$\eta_{\text{eff}} = F_k D / A v, \quad (9.49)$$

where F_k is the kinetic friction force, D is the film thickness, A the contact area, and v the sliding velocity. Using typical values for experiments with hexadecane (Yoshizawa and Israelachvili 1993): $F_k = 5 \text{ mN}$, $D = 1 \text{ nm}$, $A = 3 \times 10^{-9} \text{ m}^2$, and $v = 1 \mu\text{m s}^{-1}$, one gets $\eta_{\text{eff}} \approx 2000 \text{ Ns m}^{-2}$, or $20,000 \text{ P}$, which is ≈ 106 times higher than the bulk viscosity, η_{bulk} , of the liquid. It is instructive to consider that this very high effective viscosity nevertheless still produces a low friction force or friction coefficient μ of about 0.25. It is interesting to speculate that, if a 1 nm film were to exhibit bulk viscous behavior, the friction coefficient under the same sliding conditions would be as low as 0.000001. While such a low value has never been reported for any tribological system, one may consider it a theoretical lower limit that could conceivably be attained under certain experimental conditions.

9.9.2 Effects of Surface Structure

Various studies (Van Alsten and Granick 1990; Granick 1991; Hu and Granick 1992; Ruths 2003; Gourdon and Israelachvili 2003; Ruths et al. 2003) have shown that confinement and load generally increase the effective viscosity and/or relaxation times of molecules, suggestive of an increased glassiness or solid-like behavior (Figs. 9.29 and 9.30). This is in marked contrast to studies of liquids in small confining capillaries where the opposite effects have been observed (Warnock et al. 1986; Awschalom and Warnock 1987). The reason for this is probably because the two modes of confinement are different. In the former case (confinement of molecules between two structured solid surfaces), there is generally little opposition to any lateral or vertical displacement of the two surface lattices relative to each other. This means that the two lattices can shift in the x - y - z planes (Fig. 9.29a) to accommodate the trapped molecules in the most crystallographically commensurate or epitaxial way, which would favor an ordered, solid-like state. In contrast, the walls of capillaries are rigid and cannot easily move or adjust to accommodate the confined molecules (Fig. 9.29b), which will therefore be forced into a more disordered, liquid-like state (unless the capillary wall geometry and lattice are exactly commensurate with the liquid molecules, as occurs in certain zeolites (Awschalom and Warnock 1987)).

Experiments have demonstrated the effects of surface lattice mismatch on the friction between surfaces (Ruths and Granick 2000; Berman 1996; Hirano et al. 1991). Similar to the effects of lattice mismatch on adhesion (Fig. 9.10), the static friction of a confined liquid film is maximum when the lattices of the confining surfaces are aligned. For OMCTS confined between mica surfaces (Berman 1996) the static friction was found to vary by more than a factor of 4, while for bare mica surfaces the variation was by a factor of 3.5 (Hirano et al. 1991). In contrast to the sharp variations in adhesion energy over small twist angles, the variations in friction as a function of twist angle were much broader both in magnitude and angular spread. Similar variations in friction as a function of twist or misfit angles have also been observed in computer simulations (Gyalog and Thomas 1997).

Robbins and coworkers (He et al. 1999) computed the friction forces of two clean crystalline surfaces as a function of the angle between their surface lattices. They found that, for all nonzero angles (finite twist angles), the friction forces fell to zero due to incommensurability effects. They further found that submonolayer amounts of organic or other impurities trapped between two incommensurate surfaces can generate a finite friction force. They therefore concluded that any finite friction force measured between incommensurate surfaces is probably due to such “third-body” effects.

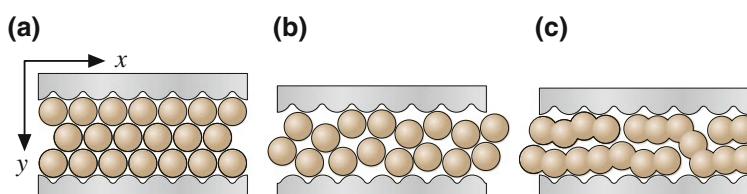


Fig. 9.29 Schematic view of interfacial film composed of spherical molecules under a compressive pressure between two solid crystalline surfaces. (a) If the two surface lattices are free to move in the x - y - z directions, so as to attain the lowest energy state, they could equilibrate at values of x , y , and z , which induce the trapped molecules to become epitaxially ordered into a solidlike film. (b) Similar view of trapped molecules between two solid surfaces that are not free to adjust their positions, for example, as occurs in capillary pores or in brittle cracks. (c) Similar to (a), but with chain molecules replacing the spherical molecules in the gap. These may not be able to order as easily as do spherical molecules even if x , y , and z can adjust, resulting in a situation that is more akin to (b) (after (Yoshizawa and Israelachvili 1993), with permission, ©1993 American Chemical Society)

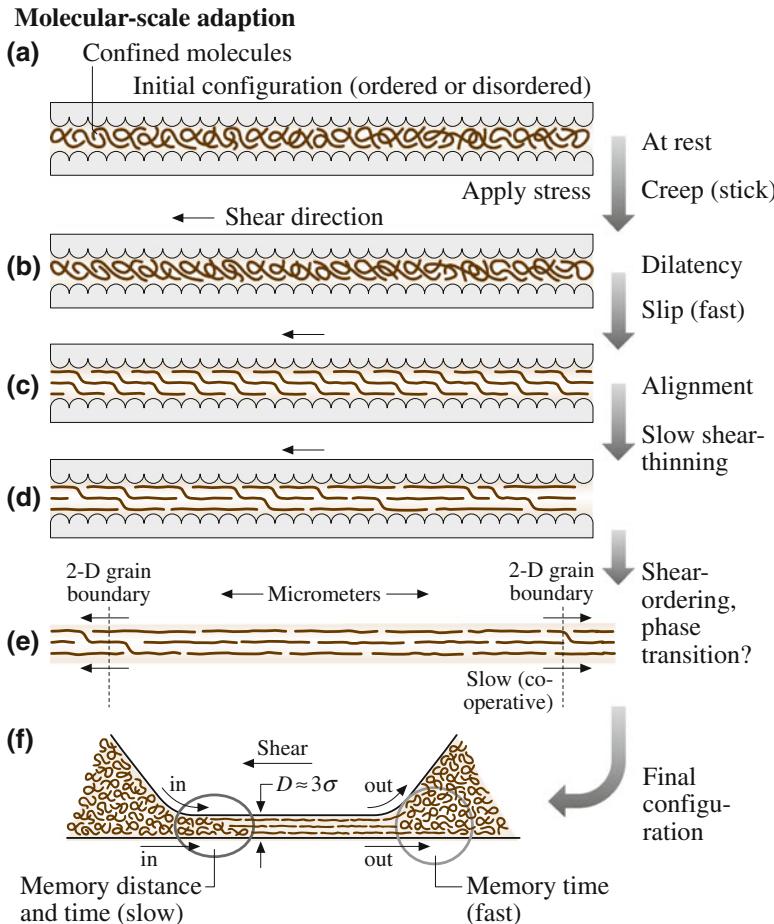


Fig. 9.30 Schematic representation of the film under shear. (a) The lubricant molecules are just confined, but not oriented in any particular direction. Because of the need to shear, the film dilates (b). The molecules disentangle (c) and get oriented in a certain direction related to the shear direction (d). (e) Slowly evolving domains grow inside the contact region. These macroscopic domains are responsible for the long relaxation times. (f) At the steady-state, a continuous gradient of confinement time and molecular order is established in the contact region, which is different for molecules adsorbed on the upper and lower surfaces. Molecules entering into the contact are not oriented or ordered. The required sliding distance to modify their state defines a characteristic distance. Molecules leaving the contact region need some (short) characteristic time to regain their bulk, unconfined configuration (after (Drummond and Israelachvili 2000), with permission, © 2000 American Chemical Society)

The reason why surface texture (lattice structure, roughness, granularity, topography, etc.) has a larger effect on the lateral (shear or friction) forces between two surfaces than on their normal (adhesion) forces is because friction is proportional to the adhesion hysteresis (Sect. 9.7.2), which can be low even when the adhesion force is high. It is also important to recognize that a

system might be defined by more than one length scale. Some systems have well-defined dimensions or size (e.g., a perfect lattice, monodisperse nanoparticles), while others have different lateral and vertical dimensions and macroscopic curvature (Persson and Tosatti 2001). Furthermore, the morphology or texture of many systems, such as asperities that are randomly distributed over a

surface, affects adhesion and tribological properties (Persson 2001; Persson and Tosatti 2001; Gao et al. 2004; Hyun et al. 2004; Israelachvili et al. 2005).

With rough surfaces, i.e., those that have *random* protrusions rather than being periodically structured, we expect a smearing out of the correlated intermolecular interactions that are involved in film freezing and melting (and in phase transitions in general). This should effectively eliminate the highly regular stick-slip and may also affect the location of the slipping planes (Gao et al. 2000, 2004; Berthoud et al. 1999). The stick-slip friction of real surfaces, which are generally rough, may, therefore, be quite different from those of perfectly smooth surfaces composed of the same material. We should note, however, that even between rough surfaces, most of the contacts occur between the tips of microscopic asperities, which may be smooth over their microscopic contact area (Thomas 1999).

References

- Alexander S (1977) Adsorption of chain molecules with a polar head a-scaling description. *J Phys (Paris)* 38:983–987
- Asakura S, Oosawa F (1958) Interaction between particles suspended in solutions of macromolecules. *J Polym Sci* 33:183–192
- Attard P (2003) Nanobubbles and the hydrophobic attraction. *Adv Colloid Interface Sci* 104:75–91
- Awschalom DD, Warnock J (1987) Supercooled liquids and solids in porous-glass. *Phys Rev B* 35:6779–6785
- Barthel E, Roux S (2000) Velocity-dependent adherence: an analytical approach for the JKR and DMT models. *Langmuir* 16:8134–8138
- Berman AD (1996) Dynamics of molecules at surfaces. University of California, Santa Barbara
- Berman A, Israelachvili J (1997) Control and minimization of friction via surface modification. In: Bhushan B (ed) Micro/nanotribology and Its applications. Kluwer Academic Publ, Dordrecht, pp 317–329
- Berman AD, Ducker WA, Israelachvili JN (1996a) Origin and characterization of different stick-slip friction mechanisms. *Langmuir* 12:4559–4563
- Berman AD, Ducker WA, Israelachvili JN (1996b) Experimental and theoretical investigations of stick-slip friction mechanisms. In: Persson BNJ, Tosatti E (eds) Physics of sliding friction. Kluwer Academic Publ, Dordrecht, pp 51–67
- Berman A, Drummond C, Israelachvili J (1998) Amontons' law at the molecular level. *Tribol Lett* 4:95–101
- Berthoud P, Baumberger T, G'Sell C, Hiver JM (1999) Physical analysis of the state- and rate-dependent friction law: static friction. *Phys Rev B* 59:14313–14327
- Bordarier P, Schoen M, Fuchs AH (1998) Stick-slip phase transitions in confined solidlike films from an equilibrium perspective. *Phys Rev E* 57:1621–1635
- Bowden FP, Tabor D (1971) The Friction and Lubrication of Solids. Clarendon, London
- Bowden FP, Tabor D (1973) An Introduction to tribology. Anchor/Doubleday, Garden City
- Briscoe BJ, Evans DCB (1982) The shear properties of Langmuir-Blodgett layers. *Proc R Soc London Ser A-Math Phys Eng Sci* 380:389–407
- Briscoe BJ, Evans DCB, Tabor D (1977) Influence of contact pressure and saponification on sliding behavior of stearic-acid monolayers. *J Colloid Interface Sci* 61:9–13
- Buckley DH (1977) Metal-to-metal interface and its effect on adhesion and friction. *J Colloid Interface Sci* 58:36–53
- Cain RG, Page NW, Biggs S (2001) Microscopic and macroscopic aspects of stick-slip motion in granular shear. *Phys Rev E* 64:8
- Campbell SE, Luengo G, Srđanov VI, Wudl F, Israelachvili JN (1996) Very low viscosity at the solid-liquid interface induced by adsorbed C-60 monolayers. *Nature* 382:520–522
- Carlson JM, Batista AA (1996) Constitutive relation for the friction between lubricated surfaces. *Phys Rev E* 53:4153–4165
- Carlson JM, Langer JS (1989) Mechanical model of an earthquake fault. *Phys Rev A* 40:6470–6484
- Casimir HBG, Polder D (1948) The influence of retardation on the London-van der Waals forces. *Phys Rev* 73:360–372
- Chan DYC, Horn RG (1985) The drainage of thin liquid-films between solid-surfaces. *J Chem Phys* 83:5311–5324
- Chen YL, Helm CA, Israelachvili JN (1991) Molecular mechanisms associated with adhesion and contact-angle hysteresis of monolayer surfaces. *J Phys Chem* 95:10736–10747
- Cheng SF, Luan BQ, Robbins MO (2010) Contact and friction of nanoasperities: effects of adsorbed monolayers. *Phys Rev E* 81:016102
- Christenson HK (1988) Adhesion between surfaces in undersaturated vapors - A reexamination of the influence of meniscus curvature and surface forces. *J Colloid Interface Sci* 121:170–178
- Christenson HK, Claesson PM (2001) Direct measurements of the force between hydrophobic surfaces in water. *Adv Colloid Interface Sci* 91:391–436
- Christenson HK, Gruen DWR, Horn RG, Israelachvili JN (1987) Structuring in liquid alkanes between solid-surfaces—force measurements and mean-field theory. *J Chem Phys* 87:1834–1841
- Christenson HK, Claesson PM, Berg J, Herder PC (1989) Forces between fluorocarbon surfactant monolayers—

- salt effects on the hydrophobic interaction. *J Phys Chem* 93:1472–1478
- Christenson HK, Fang JF, Ninham BW, Parker JL (1990) Effect of divalent electrolyte on the hydrophobic attraction. *J Phys Chem* 94:8004–8006
- Claesson PM, Blom CE, Herder PC, Ninham BW (1986) Interactions between water-stable hydrophobic Langmuir-Blodgett monolayers on mica. *J Colloid Interface Sci* 114:234–242
- Claesson PM, Ederth T, Bergeron V, Rutland MW (1996) Techniques for measuring surface forces. *Adv Colloid Interface Sci* 67:119–183
- Coakley CJ, Tabor D (1978) Direct measurement of van der Waals forces between solids in air. *J Phys D Appl Phys* 11:L77–L82
- Craig VSJ (1997) An historical review of surface force measurement techniques. *Colloid Surf A-Physicochem Eng Asp* 129:75–93
- Craig VSJ, Ninham BW, Pashley RM (1998) Study of the long-range hydrophobic attraction in concentrated salt solutions and its implications for electrostatic models. *Langmuir* 14:3326–3332
- de Gennes PG (1980) Conformations of polymers attached to an interface. *Macromolecules* 13:1069–1075
- de Gennes PG (1981) Polymer solutions near an interface. 1. Adsorption and depletion layers. *Macromolecules* 14:1637–1644
- de Gennes PG (1982) Polymers at an interface. 2. Interaction between two plates carrying adsorbed polymer layers. *Macromolecules* 15:492–500
- de Gennes PG (1987) Polymers at an interface—a simplified view. *Adv Colloid Interface Sci* 27:189–209
- Demirel AL, Granick S (1996) Glasslike transition of a confined simple fluid. *Phys Rev Lett* 77:2261–2264
- Derjaguin BV (1934a) Untersuchungen über die Reibung und Adhäsion, IV. Theorie des Anhaftens kleiner Teilchen, *Kolloid-Zeitschrift* 69:155–164
- Derjaguin BV (1934b) Molekulärtheorie der äußeren Reibung. *Z Phys* 88:661–675
- Derjaguin BV (1988) Mechanical-properties of the boundary lubrication layer. *Wear* 128:19–27
- Derjaguin B, Landau L (1941) Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes. *Acta Physicochimica URSS* 14:633–662
- Derjaguin BV, Muller VM, Toporov YP (1975) Effect of contact deformations on adhesion of particles. *J Colloid Interface Sci* 53:314–326
- Deruelle M, Tirrell M, Marciano Y, Hervet H, Léger L (1995) Adhesion energy between polymer networks and solid surfaces modified by polymer attachment. *Faraday Discuss* 98:55–65
- Dhinojwala A, Granick S (1997) Surface forces in the tapping mode: solvent permeability and hydrodynamic thickness of adsorbed polymer brushes. *Macromolecules* 30:1079–1085
- Dhinojwala A, Bae SC, Granick S (2000) Shear-induced dilation of confined liquid films. *Tribol Lett* 9: 55–62
- Dieterich JH (1978) Time-dependent friction and mechanics of stick-slip. *Pure appl Geophys* 116:790–806
- Donaldson SH, Lee CT, Chmelka BF, Israelachvili JN (2011) General hydrophobic interaction potential for surfactant/lipid bilayers from direct force measurements between light-modulated bilayers. *Proc Natl Acad Sci USA* 108:15699–15704
- Donaldson SH, Royne A, Kristiansen K, Rapp MV, Das S, Gebbie MA, Lee DW, Stock P, Valtiner M, Israelachvili J (2015) Developing a general interaction potential for hydrophobic and hydrophilic interactions. *Langmuir* 31:2051–2064
- Dowson D (1998) History of tribology, 2nd edn. Professional Engineering Publishing, London
- Drummond C, Israelachvili J (2000) Dynamic behavior of confined branched hydrocarbon lubricant fluids under shear. *Macromolecules* 33:4910–4920
- Drummond C, Israelachvili J (2001) Dynamic phase transitions in confined lubricant fluids under shear. *Phys Rev E* 63:11
- Drummond C, Elezgaray J, Richetti P (2002) Behavior of adhesive boundary lubricated surfaces under shear: a new dynamic transition. *Europhys Lett* 58:503–509
- Dzyaloshinskii IE, Lifshitz EM, Pitaevskii LP (1961) The general theory of van der Waals forces. *Adv Phys* 10:165–209
- Evans EA (1989) Force between surfaces that confine a polymer-solution: derivation from self-consistent field-theories. *Macromolecules* 22:2277–2286
- Feigin RI, Napper DH (1980) Stabilization of colloids by free polymer. *J Colloid Interface Sci* 74:567–571
- Ferry JD (1980) Viscoelastic properties of polymers, 3rd edn. Wiley, New York
- Filippov AE, Klafter J, Urbakh M (2002) Inverted stick-slip friction: what is the mechanism? *J Chem Phys* 116:6871–6874
- Fisher LR, Israelachvili JN (1981) Direct measurement of the effect of meniscus forces on adhesion—study of the applicability of macroscopic thermodynamics to microscopic liquid interfaces. *Colloids Surf* 3:303–319
- Fox HW, Zisman WA (1952) The spreading of liquids on low-energy surfaces. 3. Hydrocarbon surfaces. *J Colloid Sci* 7:428–442
- Frink LJD, van Swol F (1998) Solvation forces between rough surfaces. *J Chem Phys* 108:5588–5598
- Fuller KNG, Tabor D (1975) Effect of surface-roughness on adhesion of elastic solids. *Proc R Soc London Ser A-Math Phys Eng Sci* 345:327–342
- Gao JP, Luedtke WD, Landman U (1997) Layering transitions and dynamics of confined liquid films. *Phys Rev Lett* 79:705–708
- Gao JP, Luedtke WD, Landman U (2000) Structures, solvation forces and shear of molecular films in a rough nano-confinement. *Tribol Lett* 9:3–13
- Gao JP, Luedtke WD, Gourdon D, Ruths M, Israelachvili JN, Landman U (2004) Frictional forces

- and Amontons' law: from the molecular to the macroscopic scale. *J Phys Chem B* 108:3410–3425
- Gee ML, Israelachvili JN (1990) Interactions of surfactant monolayers across hydrocarbon liquids. *J Chem Soc-Faraday Trans* 86:4049–4058
- Gee ML, McGuiggan PM, Israelachvili JN, Homola AM (1990) Liquid to solid-like transitions of molecularly thin-films under shear. *J Chem Phys* 93:1895–1906
- Gourdon D, Israelachvili JN (2003) Transitions between smooth and complex stick-slip sliding of surfaces. *Phys Rev E* 68:10
- Granick S (1991) Motions and relaxations of confined liquids. *Science* 253:1374–1379
- Greenwood JA, Williamson JBP (1966) Contact of nominally flat surfaces. *Proc Royal Soc London Ser A-Math Phys Sci* 295:300–319
- Grosch KA (1963) Relation between friction and visco-elastic properties of rubber. *Nature* 197:858–859
- Gyalog T, Thomas H (1997) Friction between atomically flat surfaces. *Europhys Lett* 37:195–200
- Hamaker HC (1937) The London-van der Waals attraction between spherical particles. *Physica* 4:1058–1072
- He G, Muser MH, Robbins MO (1999) Adsorbed layers and the origin of static friction. *Science* 284:1650–1652
- Helm CA, Israelachvili JN, McGuiggan PM (1989) Molecular mechanisms and forces involved in the adhesion and fusion of amphiphilic bilayers. *Science* 246:919–922
- Henderson D, Lozadacassou M (1986) A simple theory for the force between spheres immersed in a fluid. *J Colloid Interface Sci* 114:180–183
- Hertz H (1881) Über die Berührung fester elastischer Körper. *J Reine Angew Math* 92:156–171
- Heslot F, Baumberger T, Perrin B, Caroli B, Caroli C (1994) Creep, stick-slip, and dry-friction dynamics: experiments and a heuristic model. *Phys Rev E* 49:4973–4988
- Heuberger M, Luengo G, Israelachvili J (1997) Topographic information from multiple beam interferometry in the surface forces apparatus. *Langmuir* 13:3839–3848
- Heuberger M, Drummond C, Israelachvili J (1998) Coupling of normal and transverse motions during frictional sliding. *J Phys Chem B* 102:5038–5041
- Heymann F, Rabinowicz E, Rightmire B (1955) Friction apparatus for very low-speed sliding studies. *Rev Sci Instrum* 26:56–58
- Hirano M, Shinjo K, Kaneko R, Murata Y (1991) Anisotropy of frictional forces in muscovite mica. *Phys Rev Lett* 67:2642–2645
- Hirz SJ, Homola AM, Hadzioannou G, Frank CW (1992) Effect of substrate on shearing properties of ultrathin polymer-films. *Langmuir* 8:328–333
- Homola AM, Israelachvili JN, Gee ML, McGuiggan PM (1989) Measurements of and relation between the adhesion and friction of 2 surfaces separated by molecularly thin liquid-films. *J Tribol Trans ASME* 111:675–682
- Homola AM, Israelachvili JN, McGuiggan PM, Gee ML (1990) Fundamental experimental studies in tribology—the transition from interfacial friction of undamaged molecularly smooth surfaces to normal friction with wear. *Wear* 136:65–83
- Horn RG (1990) Surface forces and their action in ceramic materials. *J Am Ceram Soc* 73:1117–1135
- Horn RG, Israelachvili JN (1981) Direct measurement of structural forces between two surfaces in a non-polar liquid. *J Chem Phys* 75:1400–1411
- Horn RG, Israelachvili JN, Pribac F (1987) Measurement of the deformation and adhesion of solids in contact. *J Colloid Interface Sci* 115:480–492
- Horn RG, Clarke DR, Clarkson MT (1988) Direct measurement of surface forces between sapphire crystals in aqueous solutions. *J Mater Res* 3:413–416
- Horn RG, Smith DT, Haller W (1989) Surface forces and viscosity of water measured between silica sheets. *Chem Phys Lett* 162:404–408
- Hu HW, Granick S (1992) Viscoelastic dynamics of confined polymer melts. *Science* 258:1339–1342
- Hyun S, Pei L, Molinari JF, Robbins MO (2004) Finite-element analysis of contact between elastic self-affine surfaces. *Phys Rev E* 70:12
- Israelachvili JN (1973) Thin-film studies using multiple-beam interferometry. *J Colloid Interface Sci* 44:259–272
- Israelachvili JN (1986) Measurements of the viscosity of thin fluid films between two surfaces with and without adsorbed polymers. *Colloid Polym Sci* 264:1060–1065
- Israelachvili JN (1995) Surface forces and microrheology of molecularly thin liquid films. In: Bhushan B (ed) *Handbook of micro/nanotribology*. CRC, Boca Raton, pp 319–467
- Israelachvili JN (2011) *Intermolecular and surface forces*, 3rd edn. Academic Press, London
- Israelachvili JN, Adams GE (1978) Measurements of forces between two mica surfaces in aqueous electrolyte solutions in the range 0–100 nm. *J Chem Soc Faraday Trans 1* 74:975–1001
- Israelachvili J, Berman A (1995) Irreversibility, energy-dissipation, and time effects in intermolecular and surface interactions. *Isr J Chem* 35:85–91
- Israelachvili JN, Kott SJ (1989) Shear properties and structure of simple liquids in molecularly thin-films—the transition from bulk (continuum) to molecular behavior with decreasing film thickness. *J Colloid Interface Sci* 129:461–467
- Israelachvili JN, McGuiggan PM (1990) Adhesion and short-range forces between surfaces. Part 1: new apparatus for surface force measurements. *J Mater Res* 5:2223–2231
- Israelachvili J, Pashley R (1982) The hydrophobic interaction is long-range, decaying exponentially with distance. *Nature* 300:341–342
- Israelachvili JN, McGuiggan PM, Homola AM (1988) Dynamic properties of molecularly thin liquid-films. *Science* 240:189–191

- Israelachvili JN, Kott SJ, Gee ML, Witten TA (1989) Forces between mica surfaces across hydrocarbon liquids—effects of branching and polydispersity. *Macromolecules* 22:4247–4253
- Israelachvili J, Gee M, McGuiggan P, Thompson P, Robbins M (1990) Fall meeting, materials research society, Boston. In: Drake JM, Klafter J, Kopelman R (eds). MRS, Pittsburgh, pp 3–6
- Israelachvili J, McGuiggan P, Gee M, Homola A, Robbins M, Thompson P (1990) Liquid dynamics in molecularly thin-films. *J Phys Cond Matter* 2:SA89–SA98
- Israelachvili JN, Chen YL, Yoshizawa H (1994) Relationship between adhesion and friction forces. *J Adhes Sci Technol* 8:1231–1249
- Israelachvili J, Chen Y-L, Yoshizawa H (1995) Relationship between adhesion and friction forces. In: Rimai DS, DeMejo LP, Mittal KL (eds) Fundamentals of adhesion and interfaces. Utrecht, VSP, pp 261–279
- Israelachvili J, Maeda N, Rosenberg KJ, Akbulut M (2005) Effects of sub-angstrom (pico-scale) structure of surfaces on adhesion, friction, and bulk mechanical properties. *J Mater Res* 20:1952–1972
- Joanny JF, Leibler L, Degennes PG (1979) Effects of polymer-solutions on colloid stability. *J Polym Sci B-Polym Phys* 17:1073–1084
- Johnson KL, Kendall K, Roberts AD (1971) Surface energy and the contact of elastic solids. *Proc Royal Soc A* 324:301–313
- Jonsson B, Wennerstrom H (1983) Image-charge forces in phospholipid-bilayer systems. *J Chem Soc Faraday Trans II* 79:19–35
- Kim HC, Russell TP (2001) Contact of elastic solids with rough surfaces. *J Polym Sci B-Polym Phys* 39:1848–1854
- Kjellander R, Marcelja S (1985) Perturbation of hydrogen-bonding in water near polar surfaces. *Chem Phys Lett* 120:393–396
- Klein J (1983) Forces between mica surfaces bearing adsorbed macromolecules in liquid-media. *J Chem Soc Faraday Trans I* 79:99–118
- Klein J, Kumacheva E (1995) Confinement-induced phase-transitions in simple liquids. *Science* 269:816–819
- Klein J, Kumacheva E (1998) Simple liquids confined to molecularly thin layers. I. Confinement-induced liquid-to-solid phase transitions. *J Chem Phys* 108:6996–7009
- Klein J, Luckham P (1982) Forces between two adsorbed polyethylene oxide layers immersed in a good aqueous solvent. *Nature* 300:429–431
- Klein J, Luckham PF (1984) Long-range attractive forces between two mica surfaces in an aqueous polymer-solution. *Nature* 308:836–837
- Klein J, Kamiyama Y, Yoshizawa H, Israelachvili JN, Fredrickson GH, Pincus P, Fetters LJ (1993) Lubrication forces between surfaces bearing polymer brushes. *Macromolecules* 26:5552–5560
- Klein J, Kumacheva E, Mahalu D, Perahia D, Fetters LJ (1994) Reduction of frictional forces between solid-surface bearing polymer brushes. *Nature* 370:634–636
- Kuhl TL, Berman AD, Hui SW, Israelachvili JN (1998) Part 1. Direct measurement of depletion attraction and thin film viscosity between lipid bilayers in aqueous polyethylene glycol solutions. *Macromolecules* 31:8250–8257
- Kumacheva E (1998) Interfacial friction measurement in Surface Force Apparatus. *Prog Surf Sci* 58:75–120
- Leckband D, Israelachvili J (2001) Intermolecular forces in biology. *Q Rev Biophys* 34:105–267
- Leckband DE, Israelachvili JN, Schmitt FJ, Knoll W (1992) Long-range attraction and molecular-rearrangements in receptor-ligand interactions. *Science* 255:1419–1421
- Lee DW, Banquy X, Israelachvili JN (2013) Stick-slip friction and wear of articular joints. *Proc Natl Acad Sci USA* 110:E567–E574
- Lee DW, Banquy X, Kristiansen K, Kaufman Y, Boggs JM, Israelachvili JN (2014) Lipid domains control myelin basic protein adsorption and membrane interactions between model myelin lipid bilayers. *Proc Natl Acad Sci USA* 111:E768–E775
- Lee DW, Kristiansen K, Donaldson SH, Cadirov N, Banquy X, Israelachvili JN (2015) Real-time inter-membrane force measurements and imaging of lipid domain morphology during hemifusion. *Nat Commun* 6:7238
- Lessard RR, Zieminska SA (1971) Bubble coalescence and gas transfer in aqueous electrolytic solutions. *Ind Eng Chem Fundam* 10:260–269
- Leung K, Luzar A (2000) Dynamics of capillary evaporation. II. Free energy barriers. *J Chem Phys* 113:5845–5852
- Lifshitz EM (1956) The theory of molecular attractive forces between solids. *Soviet Physics JETP* 2:73–83
- Lin Q, Meyer EE, Tadmor M, Israelachvili JN, Kuhl TL (2005) Measurement of the long- and short-range hydrophobic attraction between surfactant-coated surfaces. *Langmuir* 21:251–255
- Luan BQ, Robbins MO (2005) The breakdown of continuum models for mechanical contacts. *Nature* 435:929–932
- Luckham PF, Klein J (1990) Forces between mica surfaces bearing adsorbed homopolymers in good solvents—the effect of bridging and dangling tails. *J Chem Soc Faraday Trans* 86:1363–1368
- Luckham PF, Manimaaran S (1997) Investigating adsorbed polymer layer behaviour using dynamic surface force apparatus—a review. *Adv Colloid Interface Sci* 73:1–46
- Luengo G, Israelachvili J, Dhinojwala A, Granick S (1996) Generalized effects in confined fluids: new friction map for boundary lubrication. *Wear* 200:328–335. Erratum: Wear 205:246 (1997)
- Luengo G, Schmitt FJ, Hill R, Israelachvili J (1997a) Thin film rheology and tribology of confined polymer

- melts: Contrasts with bulk properties. *Macromolecules* 30:2482–2494
- Luengo G, Campbell SE, Srdanov VI, Wudl F, Israelachvili JN (1997b) Measurement of the adhesion and friction of smooth C-60 surfaces. *Chem Mat* 9:1166–1171
- Luzar A, Bratko D, Blum L (1987) Monte-carlo simulation of hydrophobic interaction. *J Chem Phys* 86:2955–2959
- Maeda N, Chen NH, Tirrell M, Israelachvili JN (2002) Adhesion and friction mechanisms of polymer-on-polymer surfaces. *Science* 297:379–382
- Mangipudi V, Tirrell M, Pocius AV (1994) Direct measurement of molecular-level adhesion between poly(ethylene-terephthalate) and polyethylene films—determination of surface and interfacial energies. *J Adhes Sci Technol* 8:1251–1270
- Mao M, Zhang JH, Yoon RH, Ducker WA (2004) Is there a thin film of air at the interface between water and smooth hydrophobic solids? *Langmuir* 20:1843–1849. Erratum: *Langmuir* 20:4310
- Marcelja S, Mitchell DJ, Ninham BW, Sculley MJ (1977) Role of solvent structure in solution theory. *J Chem Soc Faraday Trans II* 73:630–648
- Maugis D (1992) Adhesion of spheres—the JKR-DMT transition using a Dugdale model. *J Colloid Interface Sci* 150:243–269
- Maugis D, Pollock HM (1984) Surface forces, deformation and adherence at metal microcontacts. *Acta Metall* 32:1323–1334
- McClelland GM (1989) Friction at weakly interacting interfaces. In: Grunze M, Kreuzer HJ (eds) *Adhesion and friction*. Springer, Berlin, Heidelberg
- McGuigan PM, Israelachvili JN (1989) Measurements of the effect of angular lattice mismatch on the adhesion energy between two mica surfaces in water. In: Larson BC, Ruhle M, Seidman DN (eds) *Characterization of the structure and chemistry of defects in materials*. Materials Research Soc, Pittsburgh, pp 349–360
- McGuigan PM, Israelachvili JN (1990) Adhesion and short-range forces between surfaces. 2. Effects of surface lattice mismatch. *J Mater Res* 5:2232–2243
- McLaren KG, Tabor D (1963) Visco-elastic properties and friction of solids—friction of polymers—fluence of speed and temperature. *Nature* 197:856–858
- Merrill WW, Pocius AV, Thakker BV, Tirrell M (1991) Direct measurement of molecular-level adhesion forces between biaxially oriented solid polymer films. *Langmuir* 7:1975–1980
- Meurk A, Luckham PF, Bergstrom L (1997) Direct measurement of repulsive and attractive van der Waals forces between inorganic materials. *Langmuir* 13:3896–3899
- Meyer EE, Lin Q, Israelachvili JN (2005a) Effects of dissolved gas on the hydrophobic attraction between surfactant-coated surfaces. *Langmuir* 21:256–259
- Meyer EE, Lin Q, Hassenkam T, Oroudjev E, Israelachvili JN (2005b) Origin of the long-range attraction between surfactant-coated surfaces. *Proc Natl Acad Sci US A* 102:6839–6842
- Miller CA, Neogi P (1985) *Interfacial phenomena: equilibrium and dynamic effects*. Dekker, New York
- Milner ST, Witten TA, Cates ME (1988) Theory of the grafted polymer brush. *Macromolecules* 21:2610–2619
- Mo YF, Szlufarska I (2010) Roughness picture of friction in dry nanoscale contacts. *Phys Rev B* 81:17
- Mo YF, Turner KT, Szlufarska I (2009) Friction laws at the nanoscale. *Nature* 457:1116–1119
- Muller VM, Yushchenko VS, Derjaguin BV (1980) On the influence of molecular forces on the deformation of an elastic sphere and its sticking to a rigid plane. *J Colloid Interface Sci* 77:91–101
- Muller VM, Derjaguin BV, Toporov YP (1983) On 2 methods of calculation of the force of sticking of an elastic sphere to a rigid plane. *Colloids Surf* 7:251–259
- Nasuno S, Kudrolli A, Gollub JP (1997) Friction in granular layers: hysteresis and precursors. *Phys Rev Lett* 79:949–952
- Parker JL, Claesson PM, Attard P (1994) Bubbles, cavities, and the long-ranged attraction between hydrophobic surfaces. *J Phys Chem* 98:8468–8480
- Pashley RM (1981) DLVO and hydration forces between mica surfaces in Li^+ , Na^+ , K^+ , and Cs^+ electrolyte solutions—a correlation of double-layer and hydration forces with surface cation-exchange properties. *J Colloid Interface Sci* 83:531–546
- Pashley RM (1982) Hydration forces between mica surfaces in electrolyte solutions. *Adv Colloid Interface Sci* 16:57–62
- Pashley RM, McGuigan PM, Ninham BW, Evans DF (1985) Attractive forces between uncharged hydrophobic surfaces—direct measurements in aqueous-solution. *Science* 229:1088–1089
- Patel SS, Tirrell M (1989) Measurement of forces between surfaces in polymer fluids. *Annu Rev Phys Chem* 40:597–635
- Persson BNJ (1994) Theory of friction—the role of elasticity in boundary lubrication. *Phys Rev B* 50:4771–4786
- Persson BNJ (2001) Elastoplastic contact between randomly rough surfaces. *Phys Rev Lett* 87:1–4
- Persson BNJ, Tosatti E (2001) The effect of surface roughness on the adhesion of elastic solids. *J Chem Phys* 115:5597–5610
- Qian LM, Luengo G, Perez E (2003) Thermally activated lubrication with alkanes: the effect of chain length. *Europhys Lett* 61:268–274
- Rabinovich YI, Yoon RH (1994) Use of atomic-force microscope for the measurements of hydrophobic forces between silanated silica plate and glass sphere. *Langmuir* 10:1903–1909
- Rabinowicz E (1958) The intrinsic variables affecting the stick-slip process. *Proc Phys Soc London* 71:668–675
- Rabinowicz E (1995) *Friction and wear of materials*. Wiley, New York

- Raviv U, Perkin S, Laurat P, Klein J (2004) Fluidity of water confined down to subnanometer films. *Langmuir* 20:5322–5332
- Rhykerd CL, Schoen M, Diestler DJ, Cushman JH (1987) Epitaxy in simple classical fluids in micropores and near-solid surfaces. *Nature* 330:461–463
- Robbins MO, Thompson PA (1991) Critical velocity of stick-slip motion. *Science* 253:916
- Ruths M (2003) Boundary friction of aromatic self-assembled monolayers: comparison of systems with one or both sliding surfaces covered with a thiol monolayer. *Langmuir* 19:6788–6795
- Ruths M, Granick S (1998) Rate-dependent adhesion between polymer and surfactant monolayers on elastic substrates. *Langmuir* 14:1804–1814
- Ruths M, Granick S (2000) Influence of alignment of crystalline confining surfaces on static forces and shear in a liquid crystal, 4'-*n*-pentyl-4-cyanobiphenyl. *Langmuir* 16:8368–8376
- Ruths M, Johannsmann D, Ruhe J, Knoll W (2000) Repulsive forces and relaxation on compression of entangled, polydisperse polystyrene brushes. *Macromolecules* 33:3860–3870
- Ruths M, Heuberger M, Scheumann V, Hu JJ, Knoll W (2001) Confinement-induced film thickness transitions in liquid crystals between two alkanethiol monolayers on gold. *Langmuir* 17:6213–6219
- Ruths M, Alcantar NA, Israelachvili JN (2003) Boundary friction of aromatic silane self-assembled monolayers measured with the surface forces apparatus and friction force microscopy. *J Phys Chem B* 107:11149–11157
- Sader JE, Carnie SL, Chan DYC (1995) Accurate analytic formulas for the double-layer interaction between spheres. *J Colloid Interface Sci* 171:46–54
- Scheutjens J, Fleer GJ (1985) Interaction between two adsorbed polymer layers. *Macromolecules* 18:1882–1900
- Schoen M, Rhykerd CL, Diestler DJ, Cushman JH (1989) Shear forces in molecularly thin-films. *Science* 245:1223–1225
- Schoen M, Hess S, Diestler DJ (1995) Rheological properties of confined thin-films. *Phys Rev E* 52:2587–2602
- Schrader AM, Donaldson SH, Song J, Cheng CY, Lee DW, Han S, Israelachvili JN (2015) Correlating steric hydration forces with water dynamics through surface force and diffusion NMR measurements in a lipid-DMSO-H₂O system. *Proc Natl Acad Sci USA* 112:10708–10713
- Shrestha BR, Banquy X (2016) Hydration forces at solid and fluid biointerfaces. *Biointerphases* 11:11
- Snook IK, van Megen W (1980) Solvation forces in simple dense fluids 1. *J Chem Phys* 72:2907–2913
- Sridhar I, Johnson KL, Fleck NA (1997) Adhesion mechanics of the surface force apparatus. *J Phys D-App Phys* 30:1710–1719
- Sridhar I, Zheng ZW, Johnson KL (2004) A detailed analysis of adhesion mechanics between a compliant elastic coating and a spherical probe. *J Phys D-App Phys* 37:2886–2895
- Stanley HE, Teixeira J (1980) Interpretation of the unusual behavior of H₂O and D₂O at low-temperatures—tests of a percolation model. *J Chem Phys* 73:3404–3422
- Steinberg S, Ducker W, Vigil G, Hyukjin C, Frank C, Tseng MZ, Clarke DR, Israelachvili JN (1993) Van der Waals epitaxial-growth of alpha-alumina nanocrystals on mica. *Science* 260:656–659
- Subbotin A, Semenov A, Manias E, Hadzioannou G, Tenbrinke G (1995a) Rheology of confined polymer melts under shear-flow—strong adsorption limit. *Macromolecules* 28:1511–1515
- Subbotin A, Semenov A, Hadzioannou G, Tenbrinke G (1995b) Rheology of confined polymer melts under shear-flow—weak adsorption limit. *Macromolecules* 28:3901–3903
- Sutcliffe MJ, Taylor SR, Cameron A (1978) Molecular asperity theory of boundary friction. *Wear* 51:181–192
- Tabor D (1977) Surface forces and surface interactions. *J Colloid Interface Sci* 58:2–13
- Tabor D (1981) The role of surface and intermolecular forces in thin film lubrication. *Tribol Ser* 7:651–682
- Taunton HJ, Toprakcioglu C, Fetters LJ, Klein J (1990) Interactions between surfaces bearing end-adsorbed chains in a good solvent. *Macromolecules* 23: 571–580
- Thomas TR (1999) Rough surfaces, 2nd edn. Imperial College Press, London
- Thompson PA, Robbins MO (1990) Origin of stick-slip motion in boundary lubrication. *Science* 250: 792–794
- Thompson PA, Grest GS, Robbins MO (1992) Phase-transitions and universal dynamics in confined films. *Phys Rev Lett* 68:3448–3451
- Tomlinson GA (1929) A molecular theory of friction. *Phil Mag* 7:905–939
- Tsao YH, Evans DF, Wennerstrom H (1993) Long-range attractive force between hydrophobic surfaces observed by atomic-force microscopy. *Science* 262:547–550
- Van Alsten J, Granick S (1990) Shear rheology in a confined geometry—polysiloxane melts. *Macromolecules* 23:4856–4862
- van Olphen H (1977) An introduction to clay colloid chemistry, 2nd edn. Wiley, New York
- Verwey EJW, Overbeek JTG (1948) Theory of the Stability of lyophobic Colloids. Elsevier, Amsterdam
- Viani BE, Low PF, Roth CB (1983) Direct measurement of the relation between interlayer force and interlayer distance in the swelling of montmorillonite. *J Colloid Interface Sci* 96:229–244

- Vigil G, Xu ZH, Steinberg S, Israelachvili J (1994) Interactions of silica surfaces. *J Colloid Interface Sci* 165:367–385
- Warnock J, Awschalom DD, Shafer MW (1986) Orientational behavior of molecular liquids in restricted geometries. *Phys Rev B* 34:475–478
- Watanabe H, Tirrell M (1993) Measurement of forces in symmetric and asymmetric interactions between diblock copolymer layers adsorbed on mica. *Macromolecules* 26:6455–6466
- Xie HW, Song KY, Mann DJ, Hase WL (2002) Temperature gradients and frictional energy dissipation in the sliding of hydroxylated alpha-alumina surfaces. *Phys Chem Chem Phys* 4:5377–5385
- Yamada S, Israelachvili J (1998) Friction and adhesion hysteresis of fluorocarbon surfactant monolayer-coated surfaces measured with the surface forces apparatus. *J Phys Chem B* 102:234–244
- Yoshizawa H, Israelachvili J (1993) Fundamental mechanisms of interfacial friction. 2. Stick-slip friction of spherical and chain molecules. *J Phys Chem* 97:11300–11313
- Yoshizawa H, Chen YL, Israelachvili J (1993) Fundamental mechanisms of interfacial friction. 1. relation between adhesion and friction. *J Phys Chem* 97:4128–4140
- Zaloj V, Urbakh M, Klafter J (1999) Modifying friction by manipulating normal response to lateral motion. *Phys Rev Lett* 82:4823–4826
- Zhu YX, Granick S (2001) Rate-dependent slip of Newtonian liquid at smooth surfaces. *Phys Rev Lett* 87:096105
- Zhulina EB, Borisov OV, Priamitsyn VA (1990) Theory of steric stabilization of colloid dispersions by grafted polymers. *J Colloid Interface Sci* 137:495–511

Enrico Gnecco, Rémy Pawlak, Marcin Kisiel, Thilo Glatzel
and Ernst Meyer

Abstract

Friction has long been the subject of research: the empirical da Vinci-Amontons friction laws have been common knowledge for centuries. Macroscopic experiments performed by the school of *Bowden* and *Tabor* (Bowden and Tabor 1950) revealed that macroscopic friction can be related to the collective action of small asperities. Over the last 25 years, experiments performed with the atomic force microscope have provided new insights into the physics of single asperities sliding over surfaces. This development, together with the results from complementary experiments using surface force apparatus and the quartz microbalance, have led to the new field of *nanotribology*. At the same time, increasing computing power has permitted the simulation of processes that occur during sliding contact involving several hundreds of atoms. It has become clear that atomic processes cannot be neglected when interpreting nanotribology experiments. Even on well-defined surfaces, experiments have revealed that atomic structure is directly linked to friction force. This chapter will describe friction force microscopy experiments that reveal, more or less directly, atomic processes during sliding contact. We will begin by introducing friction force microscopy, including the calibration of cantilever force sensors and special aspects of the ultrahigh vacuum environment. The empirical Prandtl-Tomlinson model often used to describe atomic stick-slip results is therefore presented in detail. We review experimental results regarding atomic friction, including thermal

E. Gnecco (✉)

Otto Schott Institute of Materials Research (OSIM),
Friedrich Schiller University Jena, 07743 Jena,
Germany
e-mail: enrico.gnecco@uni-jena.de

R. Pawlak · M. Kisiel · T. Glatzel · E. Meyer
Department of Physics, University of Basel,
Klingelbergstrasse 82, 4056 Basel, Switzerland
e-mail: remy.pawlak@unibas.ch

M. Kisiel

e-mail: marcin.kisiel@unibas.ch

T. Glatzel

e-mail: thilo.glatzel@unibas.ch

E. Meyer

e-mail: ernst.meyer@unibas.ch

activation, velocity dependence and temperature dependence. The geometry of the contact is crucial to the interpretation of experimental results, such as the calculation of the lateral contact stiffness. The onset of wear on the atomic scale has recently been studied experimentally and it is described here. The chapter ends with a discussion of recent experiments aimed to detect the dissipative forces acting when a sharp tip is moved parallel and very close to a solid surface without being in contact with it, or when small entities such as single polymer chains, graphene nanoribbons or large organic molecules are manipulated.

Keywords

Friction force microscopy • Prandt-Tomlinson model • Atomic-scale friction • Non-contact friction • Single-molecule friction • Stick-slip motion • Superlubricity • Phononic friction • Joule dissipation

10.1 Friction Force Microscopy in Selected Environments

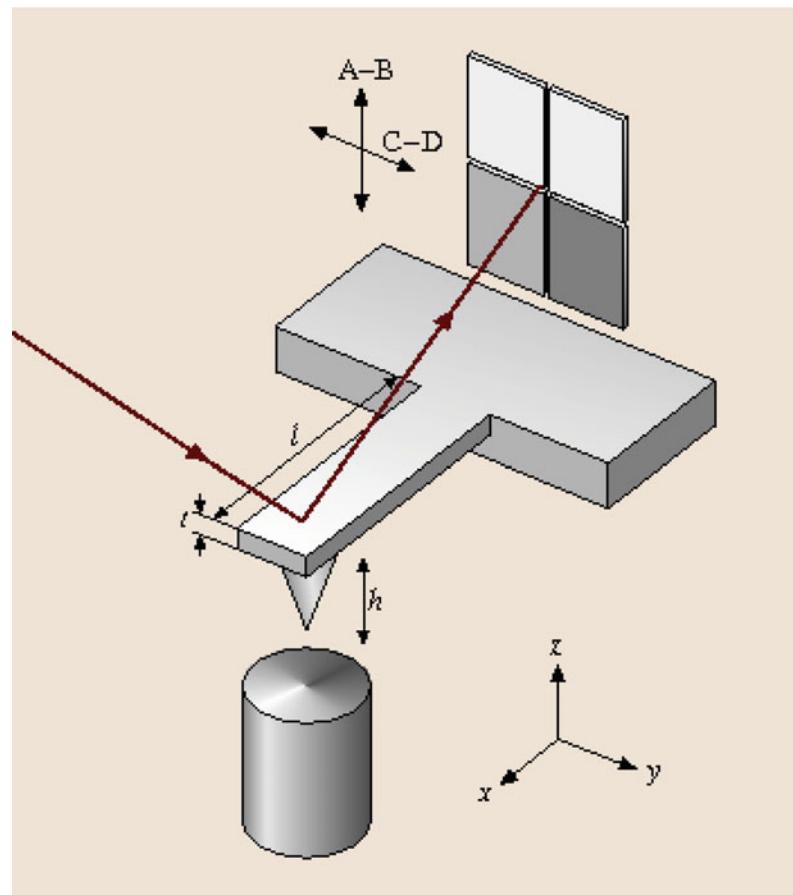
The *friction force microscope* (FFM, also called the *lateral force microscope*, LFM) is used to record the torsion deformation of an elastic cantilever driving a sharp tip on a solid surface and to investigate dissipative processes occurring in a sliding contact down to the atomic scale (Fig. 10.1). The relative motion of tip and surface is realized by a *scanner* made by piezoelectric elements, which moves the surface below the tip with a certain rate. The scanner can be also extended or retracted to vary the normal force F_N applied to the surface. This force is responsible for the deflection of the flexible cantilever supporting the tip. If the normal force F_N increases while scanning due to local topographic variations, the scanner is retracted by a feedback loop. On the other hand, if F_N decreases, the surface is brought closer to the tip by extending the scanner. In this way, the surface topography can be determined line-by-line from the vertical displacement of the scanner. Accurate control of such movement is made possible by a light beam reflected from the rear of the lever into a photodetector. When the cantilever bends, the light spot on the detector moves up or down and causes the photocurrent to

vary, when in turn triggers a corresponding change in the normal force F_N applied.

The relative sliding of tip and surface is usually also accompanied by *friction*. A lateral force F_L , which acts in the opposite direction to the scan velocity v hinders the motion of the tip. This force causes torsion in the cantilever, which can be observed along with the topography if the photodetector can detect not only the normal deflection but also the lateral movement of the lever while scanning. In practice this is achieved using four-quadrant photodetectors, as shown in Fig. 10.1. We should note that friction forces also cause lateral bending of the cantilever, but this effect is negligible if the thickness of the lever is much less than the width.

The FFM was first used by Mate et al. in 1987 to study the friction associated with atomic features (Mate et al. 1987). In their experiment, Mate used a tungsten wire and a slightly different technique (nonfiber interferometry) to detect lateral forces. Optical beam deflection was introduced later by Marti et al. (1990) and Meyer and Amer (1990). Other methods of measuring the forces between tip and surface include capacitance detection (Neubauer et al. 1990), dual fiber interferometry (McClelland and Glosli 1992) and piezolevers (Linnemann et al. 1996). In the first method, two plates close to the

Fig. 10.1 Schematic diagram of a beam-deflection friction force microscope



cantilever reveal the capacitance while scanning. The second technique uses two optical fibers to detect the cantilever deflection along two orthogonal directions aligned 45° with respect to the surface normal. Finally, in the third method, cantilevers with two Wheatstone bridges at their bases reveal normal and lateral forces, which are respectively proportional to the sum and the difference of both bridge signals.

10.1.1 Force Calibration

Force calibration is relatively simple if rectangular cantilevers are used. Due to possible discrepancies with the geometric values provided by manufacturers, one should use optical and electron microscopes to determine the width, thickness and length of the cantilever (w , t , l), the tip

height h and the position of the tip with respect to the cantilever. The thickness of the cantilever can also be determined from the resonance frequency of the lever f_0 using the relation (Nonnenmacher et al. 1991)

$$t = \frac{2\sqrt{12}\pi}{1.875^2} \sqrt{\frac{\rho}{E}} f_0 l^2. \quad (10.1.1)$$

where ρ is the density of the cantilever and E is its Young's modulus. The normal spring constant (c_N) and the lateral spring constant (c_L) of the lever are given by

$$c_N = \frac{Ewt^3}{4l^3}, \quad c_L = \frac{Gwt^3}{3h^2l}, \quad (10.1.2)$$

where G is the shear modulus. Figure 10.2 shows some SEM images of rectangular silicon

cantilevers used for FFM. In the case of silicon, $\rho = 2.33 \times 10^3 \text{ kg/m}^3$, $E = 1.69 \times 10^{11} \text{ N/m}^2$ and $G = 0.5 \times 10^{11} \text{ N/m}^2$. Thus, for the cantilever shown in Fig. 10.2, $c_N = 1.9 \text{ N/m}$ and $c_L = 675 \text{ N/m}$.

The next step consists of measuring the sensitivity of the photodetector S_z (nm/V). For beam-deflection FFMs, the sensitivity S_z can be determined by force versus distance curves measured on hard surfaces (e.g. Al_2O_3), where elastic deformations are negligible and the vertical movement of the scanner equals the deflection of the cantilever. A typical dependence of the vertical signal recorded by the four-quadrant detector V_N in ambient conditions on the tip-surface distance is sketched in Fig. 10.3. When the tip is approached towards the surface, no signal is revealed until the tip jumps into contact at $z = z_1$. Further extension and successive retraction of the scanner result in elastic behavior until the tip jumps out of contact again at a distance $z_2 > z_1$ (when $V_N = V_{\text{off}}$). The slope of the elastic part of the curve gives the required sensitivity S_z .

The normal and lateral forces depend on the voltage V_N and the horizontal signals V_L as

$$F_N = c_N S_Z V_N, \quad F_L = \frac{3}{2} c_L \frac{h}{l} S_Z V_L. \quad (10.1.3)$$

It is assumed here that the light beam is positioned precisely above the probing tip, which may not be always the case in practice.

The normal spring constant c_N can also be calibrated using other methods. Cleveland et al.

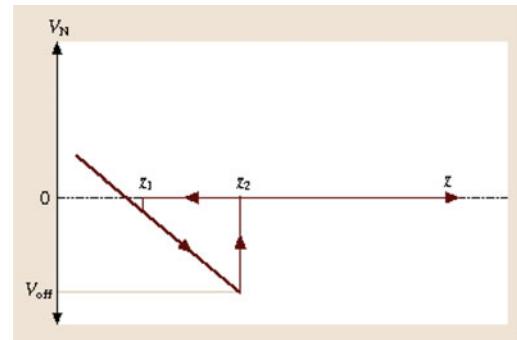


Fig. 10.3 Sketch of a typical force versus distance curve

(1993) attached tungsten spheres to the tip, which changes the resonance frequency f_0 according to the formula

$$f_0 = \frac{1}{2\pi} \sqrt{\frac{c_N}{M + m^*}} \quad (10.1.4)$$

M is the mass of the added object, and m^* is an effective mass of the cantilever, which depends on its geometry. The spring constant can be extrapolated from the frequency shifts corresponding to the different masses attached. As an alternative, Hutter et al. observed that the spring constant c_N can be related to the area of the power spectrum of the thermal fluctuations of the cantilever P (Hutter and Bechhoefer 1993). The correct relation is $c_N = 4k_B T/(3P)$, where $k_B \approx 1.38 \times 10^{-23} \text{ J/K}$ is Boltzmann's constant and T is the temperature (Butt and Jaschke 1995). Cantilevers with different shapes require finite element analysis, although analytical formulas

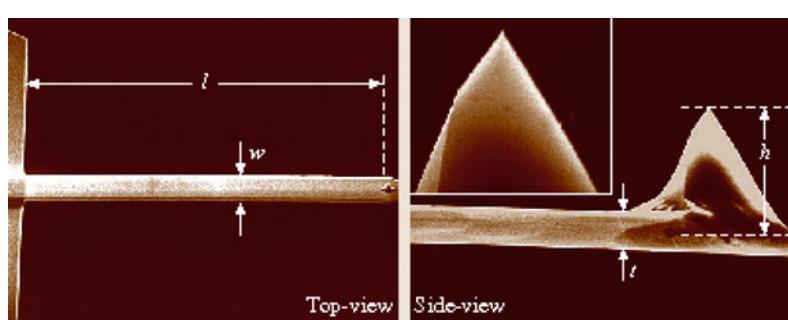


Fig. 10.2 SEM images of a rectangular cantilever. The relevant dimensions are $l = 445 \mu\text{m}$, $w = 43 \mu\text{m}$, $t = 4.5 \mu\text{m}$, $h = 14.75 \mu\text{m}$. Note that h is given by the sum of the tip height and half of the cantilever thickness

can be derived in a few cases. For V-shaped cantilevers, Neumeister et al. derived the following approximation for the lateral spring constant c_L (Neumeister and Ducker 1994)

$$c_L = \frac{\frac{Et^3}{3(1+v)h^2}}{\times \left(\frac{1}{\tan \alpha} \ln \frac{w}{d \sin \alpha} + \frac{L \cos \alpha}{w} - \frac{3 \sin 2\alpha}{8} \right)^{-1}}. \quad (10.1.5)$$

The geometrical quantities L , w , α , d , t and h are defined in Fig. 10.4. The expression for the normal constant is more complex and can be found in the cited reference.

Surfaces with well-defined profiles permit an alternative *in situ* calibration of lateral forces (Ogletree et al. 1996). As an example, Fig. 10.5 shows a commercial grating formed by alternate faces with opposite inclinations with respect to the scan direction. When the tip slides on the inclined planes, the normal force F_N and the lateral force F_L with respect to the surface are different from the two components F_\perp and F_\parallel , which are separated by the photodiode (Fig. 10.6a).

If the Amontons' relation $F_L = \mu F_N$ holds the component F_\parallel can be expressed in terms of F_\perp

$$F_\parallel = \frac{\mu + \tan \theta}{1 - \mu \tan \theta} F_\perp. \quad (10.1.6)$$

The component F_\perp is kept constant by the feedback loop. The sum of and the difference

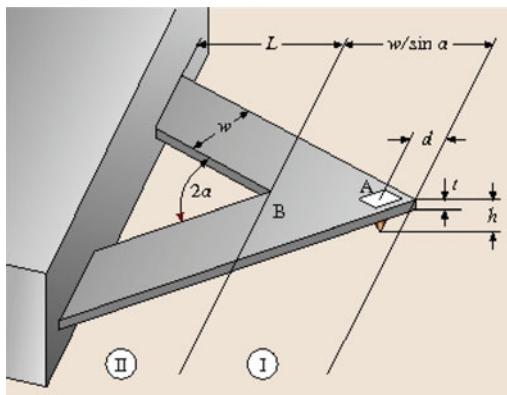


Fig. 10.4 Geometry of a V-shaped cantilever (after Neumeister and Ducker 1994)

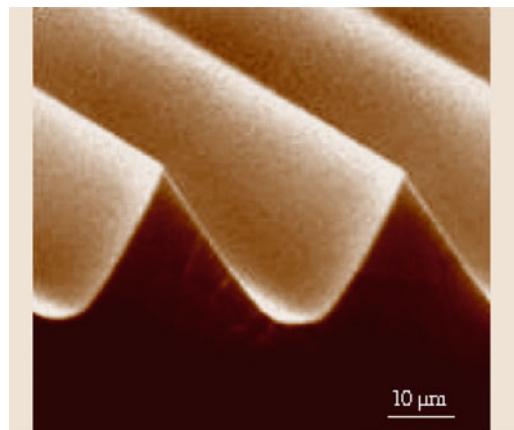


Fig. 10.5 Silicon grating formed by alternated faces angled at $\pm 55^\circ$ from the surface (© Silicon-MDT Ltd., Moscow)

between the F_\parallel values for the two planes (1) and (2) are given by

$$F_+ \equiv F_\parallel^{(1)} + F_\parallel^{(2)} = \frac{2\mu(1 + \tan^2 \theta)}{1 - \mu^2 \tan^2 \theta} F_\perp,$$

$$F_- \equiv F_\parallel^{(1)} - F_\parallel^{(2)} = \frac{2(1 + \mu^2) \tan \theta}{1 - \mu^2 \tan^2 \theta} F_\perp. \quad (10.1.7)$$

The values of F_+ and F_- (in volts) can be measured by scanning the profile back and forth (Fig. 10.6b). If F_+ and F_- are recorded with different values of F_\perp , one can determine the conversion ratio between volts and nanonewtons as well as the coefficient of friction μ .

An accurate error analysis of lateral force calibration was provided by Schwarz et al., who discussed the importance of cantilever oscillations induced by the feedback loop and of geometrical positioning of the laser beam onto the cantilever (Schwarz et al. 1996). Other sources of error (in-plane deformation and cantilever tilt) have been discussed by Sader and Green (2004), Edwards et al. (2008).

An adequate estimation of the radius of curvature of the tip R is also important. This quantity can be directly measured with a scanning electron microscope. If high-aspect structures such as step sites or whiskers are scanned, the resulting images are convolutions of those

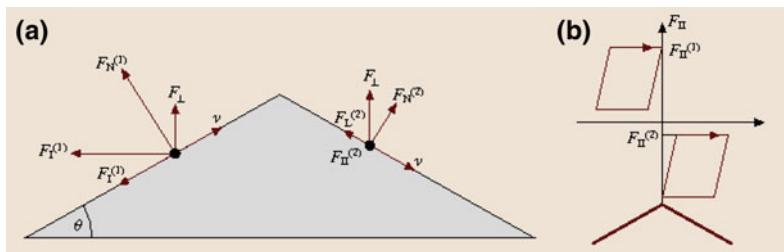


Fig. 10.6 **a** Forces acting on a FFM tip sliding on the grating shown in Fig. 10.5 and **b** friction loops acquired on the two faces

surfaces with the tip shape. A deconvolution algorithm that allows for the extraction of the probe tip's radius of curvature was proposed by Villarrubia (1997).

10.1.2 Friction Force Microscopy in Ultra High Vacuum

Atomic friction studies require well-defined surfaces and—whenever possible—tips. For the surfaces, established methods of surface science performed in ultra-high vacuum (UHV) can be employed. Ionic crystals such as NaCl have become standard materials for friction force microscopy on the atomic scale. Atomically clean and flat surfaces can be prepared by cleavage in UHV. The crystal has to be heated to ≈ 150 °C for 1 h in order to remove charge after the cleavage process. Metal surfaces can be cleaned and flattened by cycles of sputtering with argon ions and annealing. Even surfaces prepared in air or liquids, such as self-assembled molecular monolayers, can be transferred into the vacuum and studied after careful heating procedures that remove water layers.

Tip preparation in UHV is more difficult. Most force sensors for friction studies have silicon nitride or pure silicon tips. Tips can be cleaned and oxide layers removed by sputtering with argon ions. However, the sharpness of the tip is normally reduced by sputtering. As an alternative, tips can be etched in fluoric acid directly before transfer to the UHV. The significance of tip preparation is limited by the fact that the chemical and geometrical structure of the tip

can undergo significant changes when sliding over the surface.

Using FFM in UHV conditions requires some additional effort. First of all, only materials with low vapor pressures can be used, which excludes most plastics and lubricants. Beam-deflection FFMs employ either a light source in the vacuum chamber or an optical fiber guiding the light into the chamber. The positioning of the light beam on the cantilever and the reflected beam on the position-sensitive detector is achieved by motorized mirrors or by moving the light source or detector. Furthermore, a motorized sample approach must be realized.

The quality of the force sensor's electrical signal can seriously deteriorate when it is transferred out of the vacuum chamber. Low noise and high bandwidth can be preserved using a preamplifier in the vacuum. Again, the choice of materials for printing and devices is limited by the need for low vapor pressure. Stronger heating of the electrical circuitry in vacuum, therefore, may be needed.

A typical AFM used in UHV is shown in Fig. 10.7. The housing (1) contains the light source and a set of lenses that focus the light onto the cantilever. Alternatively, the light can be guided via an optical fiber into the vacuum. By using light emitting diodes with low coherency it is possible to avoid interference effects often found in instruments that use a laser as the light source. A plane mirror fixed on the spherical rotor of a first stepping motor (2) can be rotated around vertical and horizontal axes in order to guide the light beam onto the rear of the cantilever, which is mounted on a removable carrier

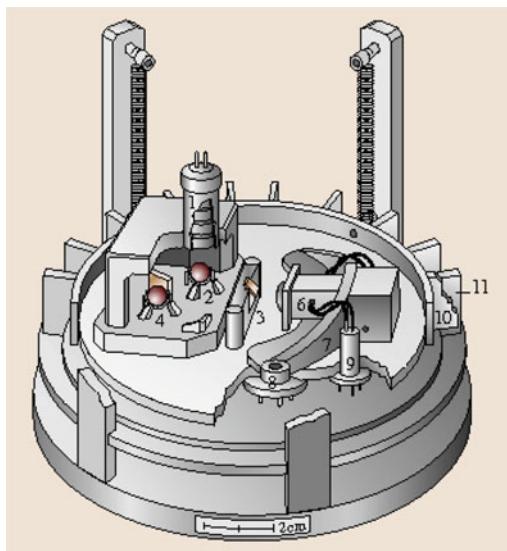


Fig. 10.7 Schematic view of the UHV-FFM realized at the University of Basel. 1 light source, 2, 4 mirrors, 3 cantilever holders, 5 photodetector, 6 scanner, 7 slider, 8 driving piezo, 9 fixed post, 10, 11 eddy current damping (after Howald et al. 1993)

plate (3). The light is reflected off the cantilever toward a second motorized mirror (4) that guides the beam to the center of the quadrant photodiode (5), where the light is then converted into four photocurrents. Four preamplifiers in close vicinity to the photodiode allow low-noise measurements with a bandwidth of 3 MHz.

The two motors with spherical rotors, used to realign the light path after the cantilever has been exchanged, work as *inertial stepping motors*: the sphere rests on three piezoelectric legs that can be moved in small amounts tangentially to the sphere. Each step of the motor consists of the slow forward motion of two legs followed by an abrupt jump backwards. During the slow forward motion, the sphere follows the legs due to friction, whereas it cannot follow the sudden jump due to its inertia. A series of these tiny steps rotates the sphere macroscopically.

The sample, which is also placed on an exchangeable carrier plate, is mounted at the end of a tube scanner (6), which can move the sample in three dimensions over several micrometers. The whole scanning head (7) is the slider of a third inertial stepping motor for coarse

positioning of the sample. It rests with its flat and polished bottom on three supports. Two of them are symmetrically placed piezoelectric legs (8), whereas the third central support is passive. The slider (7) can be moved in two dimensions and rotated about a vertical axis by several millimeters (rotation is achieved by antiparallel operation of the two legs). The slider is held down by two magnets, close to the active supports, and its travel is limited by two fixed posts (9) that also serve as cable attachments. The whole platform is suspended by four springs. A ring of radial copper lamellae (10), floating between a ring of permanent magnets (11) on the base flange, acts to efficiently damp eddy currents.

10.1.3 Energy Dissipation in Noncontact Atomic Force Microscopy

Historically, the measurement of energy dissipation induced by tip–sample interaction has been the domain of FFM. In a typical stick-slip experiment, the energy dissipated in a single atomic slip event is of the order of 1 eV and the lateral resolution is limited by the contact area, which is usually formed by several atoms. These limitations can be overcome using noncontact (NC) AFM. In this case, the tip oscillates with constant amplitude A of typically 1–20 nm at the eigenfrequency f of the cantilever, which shifts by Δf due to interaction forces between tip and sample. In NC-AFM, power dissipation P_0 caused by internal friction of the freely oscillating cantilever is observed. This dissipation is proportional to the eigenfrequency f_0 and to the square of the amplitude A and is inversely proportional to the known Q value of the cantilever. When the tip-sample distance is reduced, the tip interacts with the sample and therefore additional damping of the oscillation is encountered. The extra dissipation P_{ts} caused by the tip–sample interaction can be calculated from the excitation signal A_{exc} (Gotsmann et al. 1999). The observed energy losses per oscillation cycle (in the order of 100 meV) (Loppacher et al. 2000) are comparable to the 1 eV energy loss in contact slip process

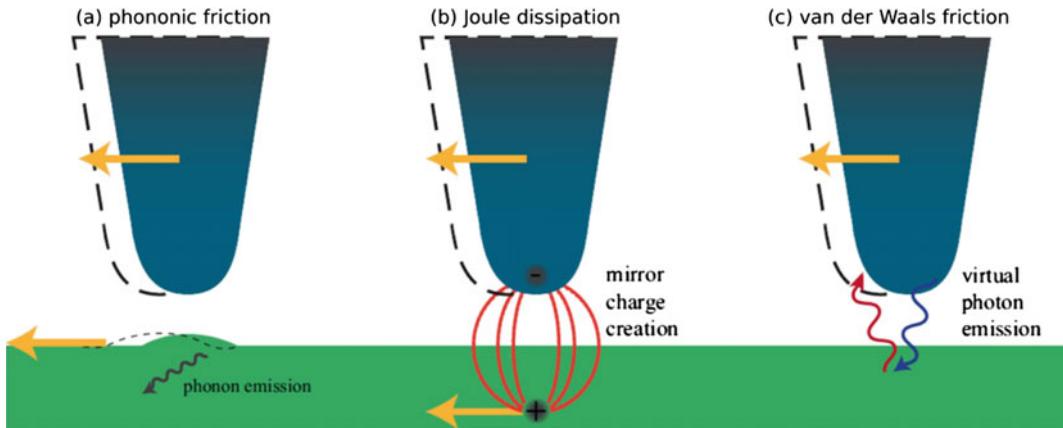


Fig. 10.8 **a** Phononic friction occurs when due to attractive tip-sample interaction the surface is deformed following the tip movement and, upon the strain energy is eventually dissipated into phonons, **b** Joule dissipation takes place when non-compensated charges beneath the surface move through a resistive medium. The tip oscillation causes the displacement current and Joule heating, and **c** the van der Waals friction is initiated by the emission of a virtual photon towards the surface. The reflected photon is adsorbed back in the tip. The friction rises because the two photons are Doppler shifted in the moving reference frame of the oscillating tip (after Gotsmann 2011)

observing that in that a few atoms are involved in the contact. The origins of dissipation in NC-AFM are manifold: phonon excitation, Joule dissipation, and van der Waals friction, as schematized in Fig. 10.8.

A recent improvement in force measurements on the nanoscale using NC-AFM came from the introduction of q-plus sensors and vertically suspended ultrasensitive cantilevers as scanning probes (Fig. 10.9). Both configurations, namely *tuning fork AFM* and *pendulum-AFM* hold two opposite positions in terms of force sensitivity and amplitude stability. Tuning fork AFM is in fact a combination of STM and AFM and is dedicated to perform high resolution imaging. The measurements are typically conducted at a temperature $T = 5$ K and in UHV. STM characterizations are done in constant-current mode with a bias voltage applied to the tungsten tip whereas AFM is performed in non-contact mode using a phase lock loop (PLL). The spring constant k of the q-plus sensor is large, typically of about 2 kN/m. Although not extremely force sensitive the probe provides outstanding atomic-scale resolution due to small amplitudes of oscillation. Amplitude stability and high stiffness of the q-plus sensor also allow to

perform quantitative measurements of molecular friction. This possibility relies on the fact that the variation Δf of the natural contact resonance f_0 of the cantilever are related to the *normal* force F_N on the tip by Albrecht et al. (1991)

$$\frac{\Delta f}{f_0} = -\frac{1}{2k} \frac{dF_N}{dz}$$

(provided that the oscillation amplitude is small compared to the characteristic decay length of F_N).

The pendulum UHV-AFM is a unique design dedicated to the measurement of extremely small forces. In this geometry the cantilever is hovering perpendicularly to the sample surface, shearing a nanometer vacuum gap between tip and sample. In this way, the so-called *non-contact friction* can be measured. In pendulum AFM very soft cantilevers with spring constants k in the order of 10^{-5} to 10^{-3} N/m are used, which avoids snapping into contact with the sample. In contrast to contact friction measurements, where the slider is usually moved at constant velocity, in the pendulum AFM the tip oscillates above the sample surface with a constant frequency. Pendulum AFM operates at cryogenic temperatures and

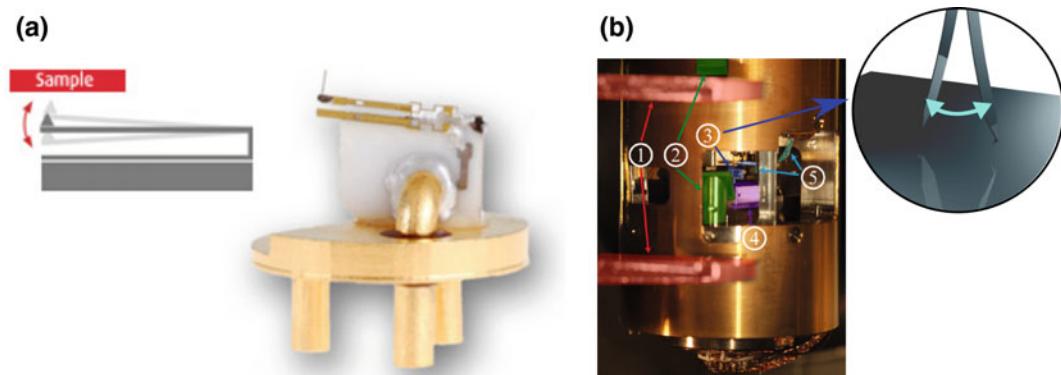


Fig. 10.9 **a** Schematic drawing of a tuning fork sensor in q-plus configuration and **b** pendulum AFM configuration where 1 is the microscope holder, 2 are the mirrors of the beam deflection detection system, 3 is the cantilever suspended in pendulum geometry, 4 is the sample holder and 5 is the STM line

typical quality factors are in the order of $Q \approx 10^5\text{--}10^6$. Such high Q together with extremely small k values (both at the limit for silicon cantilevers) imply that the minimal detectable energy dissipation is few orders of magnitude lower as compared to standard AFM configuration.

10.1.4 Friction Force Microscopy in Water

A robust alternative to UHV is also measuring friction with the tip completely embedded in water or aqueous solution. Here, capillary condensation effects are completely removed and lattice resolution is easily obtained in standard FFM maps. This possibility is achieved using dedicated liquid cells, and is particularly interesting for studying dissolution and growth of mineral surfaces, but also the molecular arrangement of self-assembled monolayers when UHV-AFM is not available. A detailed molecular dynamics (MD) investigation of atomic-scale stick-slip on a (hydrophobic) graphite surface in water, considering collisions of single H_2O molecules with the probing tip, has been recently presented by Vilhena et al. (2016). The main conclusion is that, for small forces in the nN regime, friction in water is almost indistinguishable from friction in UHV except for a slightly

enhanced background noise due to molecular collisions. Although this result cannot be generalized to hydrophilic materials, it is reasonable to expect that tip-water collisions do not play a pivotal role also in that case.

10.2 The Prandtl-Tomlinson Model

The contrast obtained in FFM images of crystal surfaces on the atomic scale is ultimately determined by the *stick-slip* motion of the probing tip. This process was described by Prandtl well before modern scanning probe techniques were introduced (Prandtl 1928) and, for historical reasons, it is usually known as Prandtl-Tomlinson (PT) mechanism. In combination with the Kramer's theory for thermally activated processes, the PT model also explains the temperature and velocity dependence of atomic scale-friction.

10.2.1 One-Dimensional Prandtl-Tomlinson Model

In the PT model, the motion of the tip is determined by both the repulsive interaction with the atomic lattice of the surface and the elastic deformations of the driving support and the contact region. The shape of the tip–surface

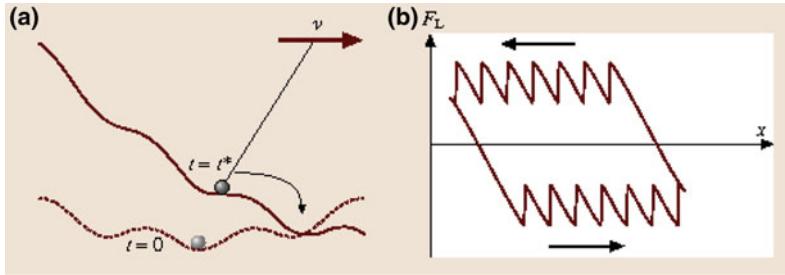


Fig. 10.10 **a** Energy profile experienced by the FFM tip (black circle) at $t = 0$ (dotted line) and $t = t^*$ (continuous line) and **b** friction loop obtained by scanning back and forth in the 1-D PT model. The effective spring constant k_{eff} is the slope of the sticking part of the loop (if $\eta \gg 1$)

potential $U(r)$ depends on several factors, such as the chemical composition of the materials in contact and the atomic arrangement on the crystal surface and the tip apex. For the sake of simplicity, we will start the analysis in the one-dimensional case considering a simple sinusoidal profile with a periodicity a and an amplitude U_0 increasing with the applied load. If the cantilever support moves with a constant velocity v along the x -direction, as usually occurs in FFM measurements, the total energy of the system is

$$U_{\text{tot}}(x, t) = -U_0 \cos \frac{2\pi x}{a} + \frac{1}{2} k_{\text{eff}}(vt - x)^2 \quad (10.2.1)$$

where k_{eff} is an effective lateral stiffness (Sect. 10.5.3). Figure 10.10a shows the energy profile $U_{\text{tot}}(x, t)$ at two different times t . When $t = 0$, the tip is pinned in the absolute minimum of U_{tot} . This minimum becomes less and less pronounced with time, and, at a critical time $t = t^*$, such position becomes unstable and the tip suddenly slips into the next minimum of the potential profile. Correspondingly, the lateral force $F_L = k(vt - x)$ takes the value (Socoliu et al. 2004)

$$F_s = \frac{k_{\text{eff}}a}{2\pi} \sqrt{\eta^2 - 1} \quad (10.2.2)$$

where $\eta = 4\pi^2 U_0 / (k_{\text{eff}}a^2)$. In this way, Eq. (10.2.2) defines the *static friction* force

acting on the tip, and stick-slip is observed only if $\eta > 1$, i.e. when the system is not too stiff or the tip-surface interaction is not too weak.

The lower curve in Fig. 10.10b shows the variation of the lateral force F_L when the cantilever is moved to the right. If, at a certain point, the direction of motion is suddenly inverted, the force follows the profile corresponding to the upper part of the curve. The area of the *friction loop* obtained by scanning forth and back gives the total energy dissipated in the process. On the other hand, when $\eta < 1$, the tip slides continuously on the surface and the lateral force oscillates between negative and positive values. Elastic instabilities are suppressed, which leads to disappearance of lateral force hysteresis (so-called *superlubricity*: see also Sect. 10.2.3).

The average *kinetic friction* can be directly estimated from the energy drop ΔU accompanying the slip as $F_k = \Delta U/a$. It increases linearly with the parameter η when the last one is high enough and it approaches the superlubric transition as $(\eta - 1)^2$ (Gnecco et al. 2012).

Note that all previous considerations are valid in the so-called *quasi-static* limit, assuming that the tip, once unpinned from the surface, immediately rebinds to it. This approximation is acceptable if the damping coefficient γ describing dissipative processes due to tip-substrate coupling is high enough. If this is not the case “long” jumps of several lattice constants are possible and, if γ is very low, a chaotic regime of motion is established (Gnecco et al. 2012).

10.2.2 Two-Dimensional Prandtl-Tomlinson Model

In two dimensions, Eq. (10.2.1) is replaced by

$$U_{tot}(\mathbf{r}, t) = U(r) + \frac{k_{\text{eff}}}{2}(\mathbf{v}t - \mathbf{r})^2 \quad (10.2.3)$$

where $\mathbf{r} \equiv (x, y)$ gives the tip position and the velocity vector \mathbf{v} is arbitrarily oriented on the surface (note that $\mathbf{v} \neq d\mathbf{r}/dt$). Figure 10.11a shows the total energy corresponding to a periodic potential of the form

$$U(x, y, t) = -\frac{U_0}{2} \left(\cos \frac{2\pi x}{a} + \cos \frac{2\pi y}{a} \right) + U_1 \cos \frac{2\pi x}{a} \cos \frac{2\pi y}{a} \quad (10.2.4)$$

In general, the equilibrium condition becomes

$$\nabla U_{tot}(\mathbf{r}, t) = \nabla U(\mathbf{r}) + k_{\text{eff}}(\mathbf{r} - \mathbf{v}t) = 0 \quad (10.2.5)$$

and the stability is described by the Hessian matrix

$$H = \begin{pmatrix} \frac{\partial^2 U}{\partial x^2} + k_{\text{eff}} & \frac{\partial^2 U}{\partial x \partial y} \\ \frac{\partial^2 U}{\partial y \partial x} & \frac{\partial^2 U}{\partial y^2} + k_{\text{eff}} \end{pmatrix}. \quad (10.2.6)$$

When both eigenvalues $\lambda_{1,2}$ of the Hessian are positive, the position of the tip is stable. Figure 10.11b shows these regions for a potential of the form (10.2.4). The tip follows the cantilever adiabatically as long as it remains in the (++)-region. When the tip reaches the border of such

region, it slips into a neighboring (++)-region and, as a result of several repetitions of this process, the tip is expected to follow a zigzag pathway. Analytical expressions for the static friction F_s and the kinetic friction F_k can be derived on simple square (Steiner et al. 2010) and hexagonal (Gnecco 2010) potentials.

10.2.3 Friction Between Atomically Flat Surfaces

Thus far we have implicitly assumed that the tip is ideally terminated by only one atom. It is also instructive to consider the case of a periodic surface sliding on another periodic surface. In the *Frenkel-Kontorova-Tomlinson* (FKT) model, the atoms of one surface are harmonically coupled with their nearest neighbors. In this context, the role of *commensurability* is essential. In 1-D, Weiss and Elmer predicted that friction should decrease with decreasing commensurability, the minimum friction being reached when the ratio between the two lattice constants is equal to the *golden mean* (Weiss and Elmer 1996)

$$\bar{z} = 1 + \frac{1}{1 + \frac{1}{1 + \dots}} = \frac{\sqrt{5} - 1}{2}. \quad (10.2.7)$$

In 2-D, Gyalog and Thomas studied the case of two square lattices with $a_1 = a_2$, and a misalignment given by an angle θ (Fig. 10.12 a). When the sliding direction changes, friction also varies from a minimum value (corresponding to the sliding angle $\varphi = \theta/2$) to a maximum

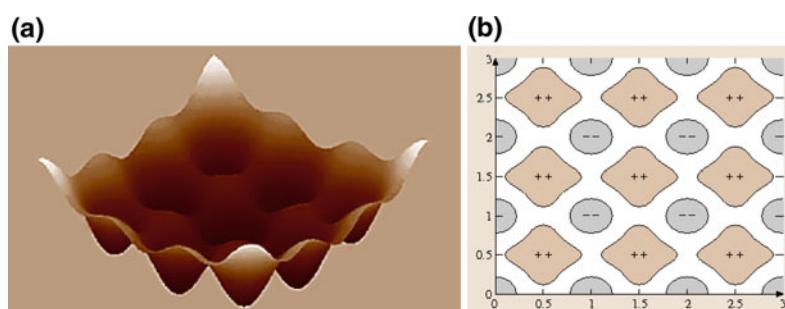


Fig. 10.11 a Energy landscape experienced by the FFM tip in 2-D and b regions on the tip plane labeled according to the signs of the eigenvalues of the Hessian matrix (after Gyalog et al. 1995)

value (which is reached when $\varphi = \theta/2 + \pi/4$; see Fig. 10.12b). The misfit angle θ is related to the commensurability. Since the misfit angles that give rise to commensurate structure form a dense subset, the dependence of friction on θ is expected to be discontinuous as indeed confirmed by additional numerical simulations.

In general, if the contact is incommensurate, the lateral forces of single atoms tend to cancel out if the load decreases or the stiffness increases and a transition from stick-slip to negligible friction is expected (Müser and Robbins 2000). This situation may be called *superlubricity* or, as suggested by Müser, *structural lubricity* (Müser 2004).

10.3 Friction Experiments on the Atomic Scale

In the first lattice resolved FFM measurements, Mate et al. scanned a graphite surface in ambient conditions with normal forces up to few tens of μN (Mate et al. 1987). According to the

continuum mechanics, these values correspond to contact diameters of about 100 nm suggesting that a graphite flake may have detached from the surface and adhered to the tip. A rather linear load dependence of friction was also observed with a small friction coefficient $\mu = 0.01$. More precise and reproducible results are obtained in UHV, where the surface cleanliness is better controlled than in ambient conditions. An atomic-scale friction map, acquired with the setup described in Sect. 10.1.2, is shown in Fig. 10.13. Here the lateral force F_L recorded on a sodium chloride surface periodically increases up to values $F_s \approx 0.4$ nN, when a slip event occurs and the tip promptly rebinds to the surface. The process is repeated several times, and the resulting FFM map perfectly reproduces the symmetry of the surface lattice. Using Eq. (10.2.2) the values $U_0 \approx 0.22$ eV and $k \approx 1$ N/m are estimated for the amplitude of the interaction potential and the effective lateral stiffness of the system. Note that the duration of the slip event, which is related to damping

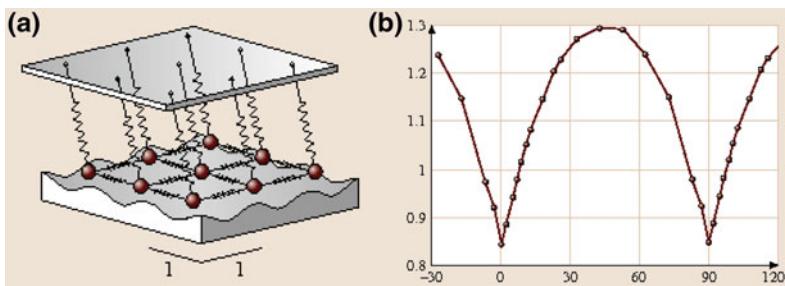
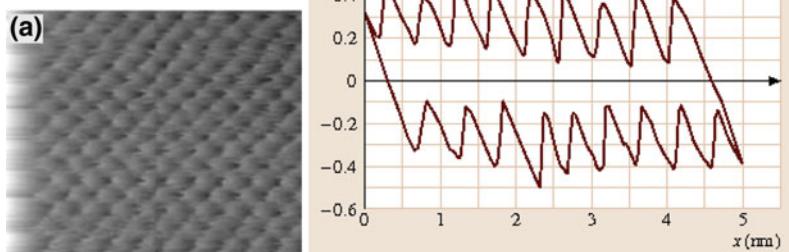


Fig. 10.12 **a** The FKT model in 2-D and **b** Friction as a function of the sliding angle φ in the 2-D FKT model (after Gyalog and Thomas 1997)

Fig. 10.13 **a** FFM map of a sodium chloride in UHV and **b** corresponding friction loop (after Gnecco et al. 2000)



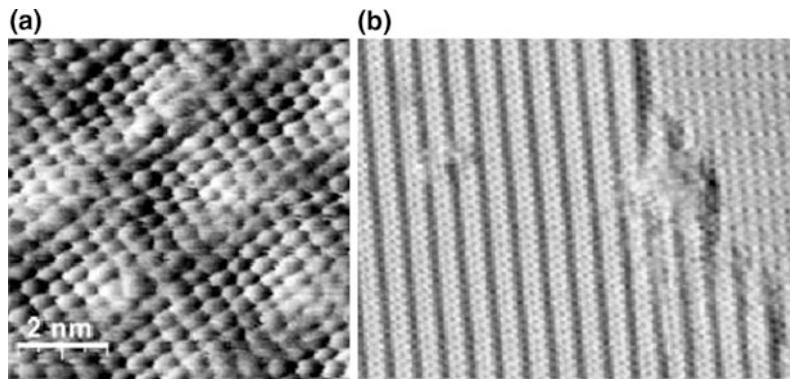


Fig. 10.14 FFM maps **a** of a moiré superstructure formed by potassium bromide on a sodium chloride surface (after Maier et al. 2008) and **b** of two terraces of a reconstructed germanium surface rotated by 90° with respect to each other (after Steiner et al. 2011)

coefficient γ of the tip oscillations, remains beyond the time resolution of state-of-the-art FFMs (Maier et al. 2005). However, this parameter may be indirectly estimated from statistical analysis of the slip length at different loading, as suggested by Roth et al. (2010). By applying subnanonewton loads to a NaCl surface, Socoliu et al. were able to observe the transition from stick-slip to continuous sliding discussed in Sect. 10.2.1 (Socoliu et al. 2004). In other experiments, the transition was induced “dynamically” by exciting flexural or torsional resonance modes of the cantilever in contact with the surface (Socoliu et al. 2006; Roth et al. 2014).

Note that no individual defects are seen in Fig. 10.13. This is expected if the contact between tip and surface is formed by several atoms. The availability of ultrasharp tips with curvature radius below 2 nm has recently allowed to overcome this limitation. High resolution FFM imaging of the moiré pattern formed by an ultrathin KBr film epitaxially grown on NaCl is demonstrated in Fig. 10.14a. This pattern results from the lattice mismatch between the two surfaces, with a substrate versus adsorbate lattice constant ratio of about 6:7. It is not observed if the load is close to the pull-off value. Hexagonal moiré patterns have been also reported by

Filleter et al. using FFM on a graphene film on SiC(0001) (Filleter and Bennewitz 2010). Reconstruction domains on semiconducting surfaces (Goryl et al. 2012), and friction variations across surface step edges of sodium chloride and germanium (Fig. 10.14b) have been also observed with “quasi-atomic” resolution using FFM.

The anisotropy of friction and the zigzag walk of the tip predicted by the 2D PT model (Sect. 10.2.2) was first observed by Fujisawa et al. (1995) and Kawakatsu and Saito (1996) on mica and on MoS₂ using original 2D FFM with two laser beams and two quadrant photodetectors. With another dedicated “Tribolever” Dienwiebel et al. were able to track friction forces with arbitrary scan orientations (Dienwiebel et al. 2004). In this way, stick-slip and a significant energy dissipation were only observed at angles of rotation corresponding to a commensurate contact between a graphite surface and graphene flakes picked up by the tip. More recently, friction anisotropy has been also demonstrated by Fessler et al. on organic crystal surfaces (Fessler et al. 2011). The chemical structure of the used benzylammonium terminated crystalline surface is presented in Fig. 10.15a. The FFM maps in UHV revealed a significant contrast between different molecular orientations yielding

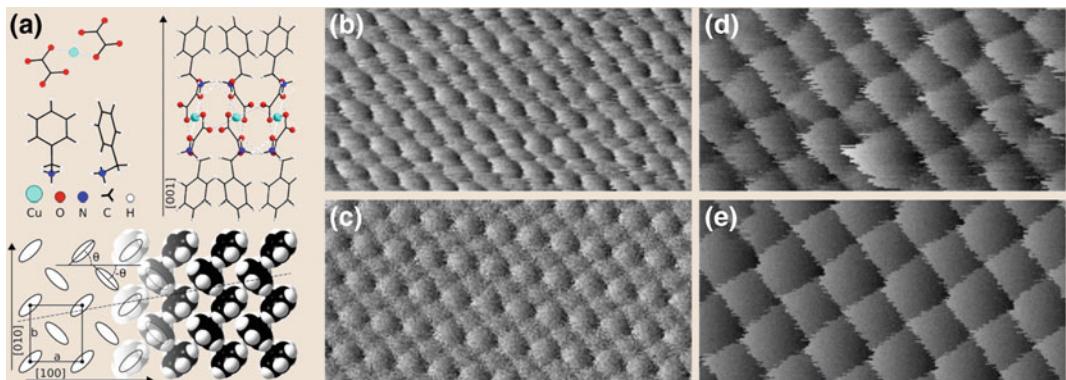


Fig. 10.15 **a** The $[\text{Cu}(\text{C}_2\text{O}_4)_2]^{2-}$ complex builds up an extended two-dimensional structure surrounded by the organic BA cations. The organic crystal is built up by multiple stacks of the illustrated layers and the [001] surface shows the different orientations of the BA molecules. The lattice constants are $a = 7.1$ and $b = 8.0 \text{ \AA}$. **b** and **d** show FFM measurements at different loads (0 and 1.1 nN, respectively). Simulations with the modified potential was done with $\eta = 1$, $c = 0.35$ (**c**) and with $\eta = 3$, $c = 0.1$ (**e**) (after Fessler et al. 2011)

molecular rows which differ from their neighboring ones in dependence of the applied load (Fig. 10.15b). These maps could be reproduced introducing the empirical potential

$$U(x, y) = -U_0 \cos \frac{2\pi x}{a} \left(\cos \frac{2\pi y}{b} + c \right)$$

in the 2D PT model, where a and b are the lattice constants and the parameter c gives rise to two different energy barriers $\Delta E^+ = U_0(1+c)$ and $\Delta E^- = U_0(1-c)$. Two examples of measurements and corresponding simulations are shown in Fig. 10.15b–e. The good agreement between them is demonstrated by the alternate contrast of a row with almost smooth sliding and the neighboring ones with considerably larger friction corresponding to the energy barriers ΔE^- and ΔE^+ respectively.

The potential of liquid environments for high-resolution FFM is also well established. Figure 10.16 shows a friction map of a dolomite cleavage surface and of a thin film of Cu phthalocyanine (PC) molecules previously deposited on it. On the dolomite surface two pinning sites per unit cell are resolved, corresponding to protruding carbonate groups. The

contrast of each group varies with the scan direction, which can be also explained by the 2D PT model in 2D (Pina et al. 2012). A comparison with the structure of PC crystals, as obtained by X-ray diffraction, clearly reveals that the molecules are stacked perpendicular to the substrate. This information could not be obtained in ambient conditions, where, due to capillary condensation, lattice resolution was not achieved.

10.4 Thermal Effects on Atomic Friction

Although the PT model gives a good interpretation of the basic mechanism of the atomic stick-slip discussed in Sect. 10.3, it cannot explain some important features observed in the friction maps. For example, the peaks in the sawtooth profile in Fig. 10.13b have different heights, which is in contrast to the result in Fig. 10.10b. Another effect is observed if the scan velocity v is varied: the mean friction force increases with the logarithm of v (Gnecco et al. 2000). This effect cannot be understood within the mechanism discussed in Sect. 10.2 without further assumptions.

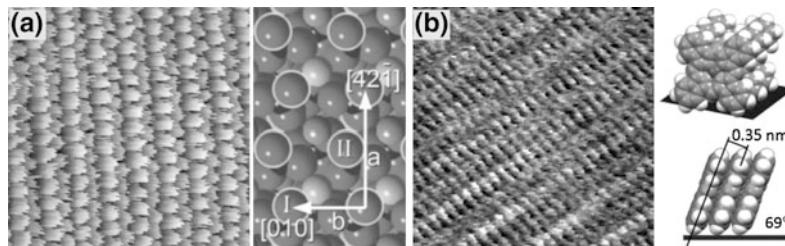


Fig. 10.16 FFM maps of **a** a dolomite cleavage surface and **b** of a Cu phtalocyanine ultrathin film grown on it (after Nita et al. 2013)

10.4.1 The Prandtl-Tomlinson Model at Finite Temperature

Let us focus again on the energy profile discussed in Sect. 10.2.1. For the sake of simplicity, we will suppose that $\eta \gg 1$. At a given time $t < t^*$, the tip jump is prevented by the energy barrier $\Delta E = U(x_{\max}, t) - U(x_{\min}, t)$, where x_{\max} corresponds to the first maximum observed in the energy profile and x_{\min} is the actual position of the tip (Fig. 10.17a). The quantity ΔE decreases with time or, equivalently, with the frictional force F_L until it vanishes when $F_L \approx F^*$ (Fig. 10.17b).

At finite temperature T , the lateral force required to induce a jump is lower than F^* . To estimate the most probable value of F_L at this point, we first consider the probability p that the tip does *not* jump. The probability p changes with time t according to the master equation

$$\frac{dp(t)}{dt} = -f_0 \exp\left(-\frac{\Delta E(t)}{k_B T}\right)p(t), \quad (10.4.1)$$

where f_0 is a characteristic attempt frequency in the order of few tens of kHz (for typical FFM measurements on the atomic scale). Note that the probability of a reverse slip is neglected, since in this case the energy barrier that must be overcome is much higher than ΔE . Assuming a linear dependence of ΔE on the lateral force, a logarithmic dependence of friction on the scan velocity is obtained:

$$F_L(v) = F^* - \frac{k_B T}{\lambda} \ln \frac{v_C}{v} \quad (10.4.2)$$

where λ is in the order of the lattice constant.

However, the linear approximation for ΔE does not hold when the slip occurs very close to the critical point $x = x^*$, which is the case at

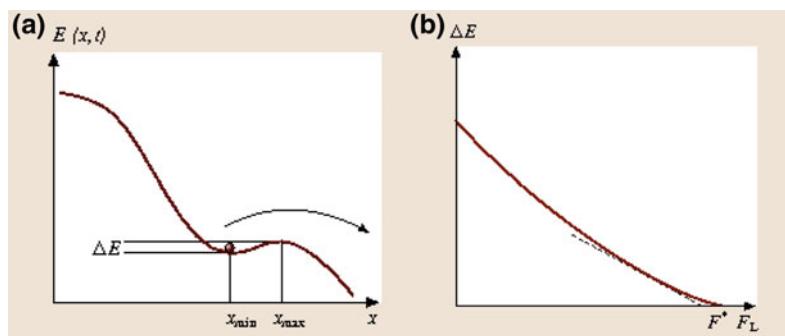


Fig. 10.17 **a** Energy barrier that hinders the tip jump in the PT model and **b** energy barrier ΔE as a function of the lateral force F_L . The dashed line close to the critical value corresponds to the linear approximation (10.4.2)

high velocities. In this case thermal activation does not play any role and friction (due to stick-slip motion) becomes constant, in agreement with the classical Coulomb's law of friction. We should also mention that a better approximation for the energy barrier is $\Delta E \propto (F^* - F_L)^{3/2}$ (Sang et al. 2001). The corresponding relations for the velocity dependence of friction are more complex and difficult to distinguish from the experimental results. Furthermore, the logarithmic law (10.4.1) becomes also invalid at very low values of v . In this case reverse slip cannot be excluded, and, as a result, static friction is found to be proportional to v (Krylov et al. 2005). At room temperature such "thermolubric" regime would be observed for unrealistic velocity values well below 1 pm/s (but the critical value increases rapidly with temperature). The thermally activated PT model has been also discussed in 2D (Fusco and Fasolino 2005).

10.4.2 Velocity and Temperature Dependence of Atomic-Scale Friction

The first velocity dependent atomic-scale FFM measurements in UHV were performed by Bennewitz et al. (1999) and Gnecco et al. surfaces (Gnecco et al. 2000) on copper and sodium chloride; in both cases a logarithmic increase of friction with the sliding velocity was observed up to $v < 1 \mu\text{m/s}$ in perfect agreement with Eq. (10.4.2). It is worth noting that a similar logarithmic dependence was also reported by Bouhacina et al. in polymer layers grafted on silicon oxide and interpreted as the result of stress-aided thermally activation of viscous flow (Bouhacina et al. 1997). We should also emphasize that the so-called *velocity strengthening* of friction can be masked by *velocity weakening* in FFM measurements performed in a humid environment. Here the nucleation of capillary bridges can be assisted by thermal activation and lead to friction forces

logarithmically increasing with the sliding velocity (Riedo et al. 2002).

Thus far we have used thermal activation to explain the velocity dependence of friction. The same mechanism also predicts that friction should change with temperature. The master Eq. (10.4.1) shows indeed that the probability of a tip slip is reduced at low temperatures T until it vanishes when $T = 0$. In this case thermal activation is excluded, and the lateral force F_L is equal to F^* , independent of the scanning velocity v .

Zhao et al. (2007) observed that friction on graphite decreases as $1/T$ over a wide temperature range (140–750 K), supporting the hypothesis of thermal activation of the stick-slip process. However, it was only recently that group of Schirmeisen reported atomic-scale FFM measurements in UHV at different temperatures (Janssen et al. 2010). When silicon, SiC, ionic crystals and graphite surfaces were cooled down from room temperature to cryogenic conditions, a good agreement with the thermally activated PT model was found down to a peak or a plateau, appearing between 50 and 200 K. Below these values, the friction was found to decrease with temperature, which the authors attributed to the competition between thermally activated rupture and formation of chemical bonds (Barel et al. 2010).

10.5 Friction on the Nanometer Scale

Friction is ultimately related to the real shape of the contact between the sliding surfaces. On the macroscopic scale, the contact between two bodies is studied within the context of continuum mechanics, which is based on the elasticity theory developed by Hertz in the nineteenth century. Various FFM experiments have shown that continuum mechanics is still valid down to contact areas just a few nanometers in size. Only when contact is reduced to few atoms does the continuum frame become unsuitable, and other approaches like molecular dynamics become necessary.

10.5.1 Continuum Mechanics of Single Asperities

The lateral force F_L between two surfaces in reciprocal motion depends on the size of the real area of contact, A , which can be a few orders of magnitude smaller than the apparent area of contact. The simplest assumption is that friction is proportional to A ; the proportionality factor is called the *shear strength* τ [Bow50].

$$F_L = \tau A \quad (10.5.1)$$

For plastic deformation, the asperities are compressed until the pressure p equals a certain yield value p^* . The resulting contact area is thus $A = F_N/p^*$, and the well-known Amonton's law is obtained: $F_L = \mu F_N$, where $\mu = \tau/p^*$ is the *coefficient of friction*. The same idea can be extended to contacts formed by many asperities, and it leads again to Amonton's law. The simplicity of this analysis explains why most friction processes were related to plastic deformation for a long time. Such a mechanism, however, should provoke quick disruption of surfaces, which is not observed in practice.

Elastic deformation can be easily studied in the case of a sphere of radius R pressed against a flat surface. In this case, the contact area is

$$A(F_N) = \pi \left(\frac{R}{K} \right)^{2/3} F_N^{2/3} \quad (10.5.2)$$

where $K = 3E^*/4$ and E^* is an effective Young's modulus, related to the Young's moduli (E_1 and E_2) and the Poisson numbers (v_1 and v_2) of sphere and plane, by the following relation (Landau and Lifshitz 1998)

$$\frac{1}{E^*} = \frac{1 - v_1^2}{E_1} + \frac{1 - v_2^2}{E_2}. \quad (10.5.3)$$

The result $A \propto F_N^{2/3}$ contrasts with Amonton's law. However, a linear relation between F_L and F_N can be obtained for contacts formed from several asperities in particular cases. For example, the area of contact between a flat surface and a set of asperities with an exponential height

distribution and the same radius of curvature R depends linearly on the normal force F_N (Greenwood and Williamson 1966). The same conclusion holds approximately even for a Gaussian height distribution. However, the hypothesis that the radii of curvature are the same for all asperities is not realistic. A general model was recently proposed by Persson, who analytically derived the proportionality between contact area and load for a large variety of elasto-plastic contacts formed by surfaces with arbitrary roughness (Persson 2001). However, the discussion of this theory is not straightforward and goes beyond the purposes of this section.

Further effects are observed if adhesive forces between the asperities are taken into account. If the range of action of these forces is smaller than the elastic deformation, (10.5.2) is extended to the Johnson–Kendall–Roberts (JKR) relation

$$A(F_N) = \pi \left(\frac{R}{K} \right)^{2/3} \times (F_N + 3\pi\gamma R + \sqrt{6\pi\gamma RF_N + (3\pi\gamma R)^2})^{2/3}, \quad (10.5.4)$$

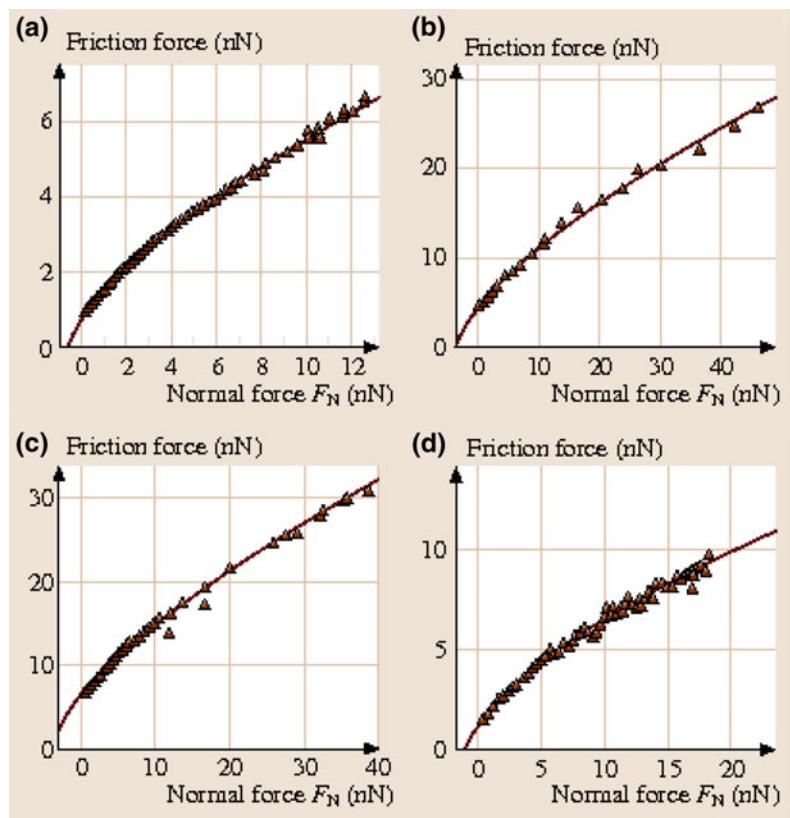
where γ is the surface tension of the sphere and plane (Johnson et al. 1971). The real contact area at zero load is finite and the sphere can be detached only by pulling it away with a certain force. This is also true in the opposite case, in which the range of action of adhesive forces is larger than the elastic deformation. In this case, the relation between contact area and load takes the simple form

$$A(F_N) = \pi \left(\frac{R}{K} \right)^{2/3} (F_N - F_{\text{off}})^{2/3}, \quad (10.5.5)$$

where F_{off} is the negative load required to break the contact. The “Hertz-plus-offset” relation (10.5.5) can be derived from the Derjaguin–Muller–Toporov (DMT) model (Derjaguin et al. 1975). To discriminate between the JKR or DMT models, Tabor introduced a nondimensional parameter

$$\phi = \left(\frac{9R\gamma^2}{4K^2 z_0^3} \right)^{1/3}, \quad (10.5.6)$$

Fig. 10.18 Friction versus load curve on amorphous carbon in argon atmosphere. Curves (a, b) refer to tips with different radii of curvature (after Schwarz et al. 1997)



where z_0 is the equilibrium distance during contact. The JKR model can be applied if $\Phi > 5$; the DMT model holds when $\Phi < 0.1$ (Tabor 1977). For intermediate values of Φ , the Maugis–Dugdale model (Maugis 1992) can reasonably explain experimental results.

10.5.2 Dependence of Friction on Load

The FFM tip represents a single asperity sliding on a surface. The previous discussion suggests a nonlinear dependence of friction on the applied load, provided that continuum mechanics is applicable. Schwarz et al. observed the Hertz-plus-offset relation (10.5.5) on graphite, diamond, amorphous carbon and C_{60} in an argon atmosphere (Fig. 10.18). In their measurements, they used well-defined spherical tips with radii of curvature of tens of nanometers, obtained by contaminating silicon tips with

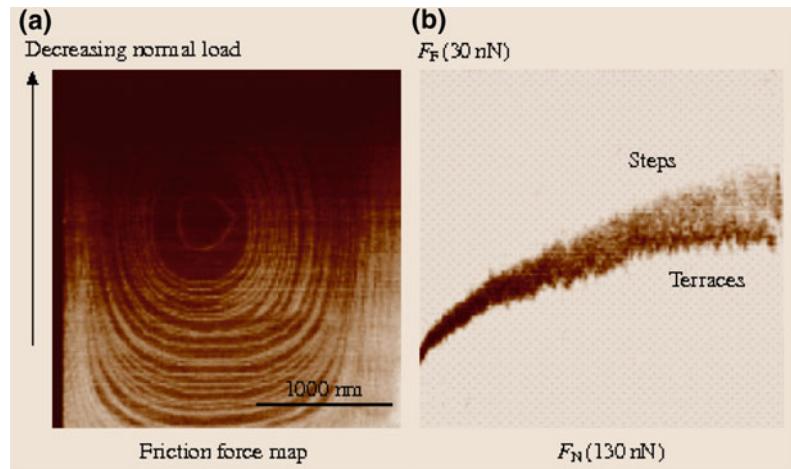
amorphous carbon in a transmission electron microscope.

A good agreement with JKR theory have been reported in earlier UHV-FFM measurements on ionic crystals, mica and metals (Meyer et al. 1996; Carpick et al. 1996; Polaczyk et al. 1998). In order to correlate the lateral and normal forces with improved statistics, Meyer et al. applied an original 2-D histogram technique (Fig. 10.19). Carpick et al. extended the JKR relation (10.5.4) to include nonspherical tips (Carpick et al. 1996). In the case of an axisymmetric tip profile $z \propto r^{2n}$ ($n > 1$), it can be proven analytically that the increase in the friction becomes less pronounced with increasing n .

10.5.3 Estimation of the Contact Area

In contrast to other tribological instruments, such as the surface force apparatus (Israelachvili and

Fig. 10.19 **a** Friction force map on NaCl(100). The load is decreased from 140 to 0 nN (jump-off point) during imaging and **b** 2-D histogram of (a) (after Meyer et al. 1996)



Tabor (1972), the contact area cannot be measured directly by FFM. Indirect methods are provided by contact stiffness measurements. The contact between the FFM tip and the sample can be modeled by a series of two springs (Fig. 10.20). The effective constant k_{eff}^z of the series is given by

$$\frac{1}{k_{\text{eff}}^z} = \frac{1}{k_{\text{contact}}^z} + \frac{1}{c_N}, \quad (10.5.7)$$

where c_N is the normal spring constant of the cantilever and k_{contact}^z is the normal stiffness of the contact. This quantity is related to the radius of the contact area (a) by the simple relation (Johnson 1985)

$$k_{\text{contact}}^z = 2aE^* \quad (10.5.8)$$

where E^* is the effective Young's modulus introduced previously. However, typical values of k_{contact}^z are 1–2 orders of magnitude larger than c_N and a direct application of (10.5.7) is difficult. Alternatively, one can measure the normal resonance frequency f_N of the cantilever contacting the surface while scanning. Steiner et al. reported atomically resolved measurements of f_N on a sodium chloride surface (Steiner et al. 2009). Corresponding variations of the lateral force were simultaneously acquired. The normal

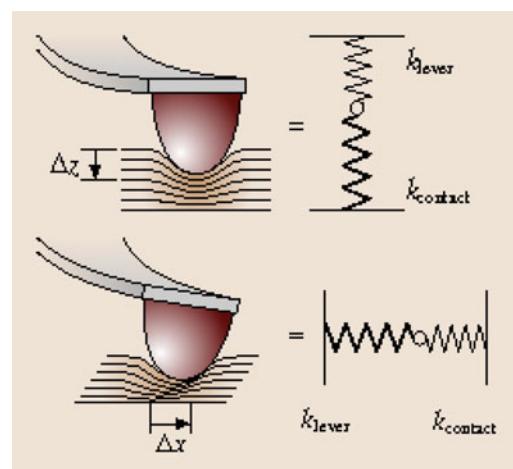


Fig. 10.20 Sketch of normal and lateral stiffness of the contact between tip and surface (after Carpick et al. 1997)

stiffness k_N could be estimated from continuum mechanics relations, which were modified to account for the inclination of the cantilever with respect to the sample surface.

The *lateral* contact stiffness of the contact between a sphere and a flat surface is (Johnson 1985)

$$k_{\text{contact}}^x = 8aG^*, \quad (10.5.9)$$

where the effective shear stress G^* is defined by

$$\frac{1}{G^*} = \frac{2 - v_1^2}{G_1} + \frac{2 - v_2^2}{G_2}, \quad (10.5.10)$$

G_1, G_2 are the shear moduli of the sphere and the plane, respectively. The contact between the FFM tip and the sample can again be modeled by a series of springs (Fig. 10.20). The effective constant k_{eff}^x of the series is given by

$$\frac{1}{k_{\text{eff}}^x} = \frac{1}{k_{\text{contact}}^x} + \frac{1}{k_{\text{tip}}^x} + \frac{1}{c_L}, \quad (10.5.11)$$

where c_L is the lateral spring constant of the cantilever and k_{contact}^x is the lateral stiffness of the contact. As suggested by Lantz et al., (10.5.11) also includes the lateral stiffness of the tip k_{tip}^x which can be comparable to the lateral spring constant (Lantz et al. 1997). The effective spring constant k_{eff}^x is simply given by the slope dF_L / dx of the friction loop (Sect. 10.2.1). Once k_{contact}^x is determined, the contact radius a is easily estimated by (10.5.9).

The lateral stiffness method was applied to contacts between silicon nitride and muscovite mica in air and between NbSe₂ and graphite in UHV. The dependences of both the spring constant k_{eff}^x and the lateral force F_L on the load F_N were explained within the same models (JKR and Maugis–Dugdale, respectively), which confirms that friction is proportional to the contact area for

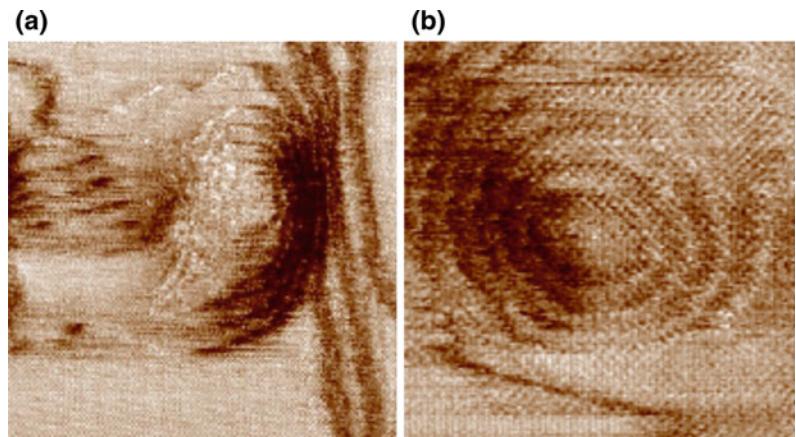
the range of loads applied (up to $F_N = 40$ nN in both experiments). Enachescu et al. estimated the contact area by measuring the contact conductance on diamond as a function of the applied load (Enachescu et al. 1998, 1999). Their experimental data were fitted with the DMT model, which was also used to explain the dependence of friction on load. Since the contact conductance is proportional to the contact area, the validity of the hypothesis (10.5.1) was confirmed again.

10.6 Wear on the Atomic Scale

If the normal force F_N applied by the FFM exceeds a critical value, which depends on the tip shape and on the material under investigation, the surface topography is permanently modified by abrasive wear and related processes. On the nanoscale wear may result in patterns with well-defined shapes. Here we will focus on few experiments which demonstrated the unique ability of the FFM to both scratch and image surfaces down to the atomic scale.

In Fig. 10.21a, a small mound piled up at the end of a groove on KBr(100) is shown at different magnifications. The groove (800 nm long) was produced by repeated scratching with a normal force F_N of about 20 nN. When the groove was formed (in about 20 min) a rather continuous increase of friction with time was

Fig. 10.21 **a** Lateral force image acquired at the end of a groove scratched 256 times with a normal force $F_N = 21$ nN. Frame sizes: 39 nm and **b** lateral force image of a pit produced with $F_N = 22.8$ nN. Frame size: 20 nm (after Gnecco et al. 2002)



observed, which is well fitted by the exponential law

$$F_L = F_0 + F_1 (1 - e^{-t/\tau}). \quad (10.6.1)$$

If $t \gg \tau$ a saturation of the contact area (proportional to friction) is expected, corresponding to the limit value at which the applied load can be perfectly balanced by the tip. In spite of the apparent continuity of the process on long time scales, the friction loops acquired while scraping (not shown) suggest that the atoms (or ion pairs) are removed one by one. Note that this kind of “atom-by-atom attrition” processes has been recently observed in TEM images of nanotips getting worn when scraped on harder surfaces (Jacobs and Carpick 2013). No fracture or plastic deformation was found when Si volumes up to few tens of nm³ were progressively removed from the tip.

Another example is given in Fig. 10.21b, where a pit produced by repeated low-frequency tip oscillations on a 5 nm long line is shown. From a comparison of the number of missing atoms (extracted from the figure) and the energy dissipated during scanning (extracted from the friction loops) it could be estimated that only 30% of the energy dissipation went into rupture of ionic bonds. Furthermore, the damage was observed to increase with the applied load, but did not vary significantly with the scan velocity.

In another study on KBr films on Cu(100) Filleter et al. (2008) reported significant wear at intrinsic step edges of the films, where atomic coordination is lower. In contrast, low friction and no wear were observed across metal steps covered by KBr, which indicates a stabilizing effect of the alkali halide coating on the metal surface.

A different kind of wear is observed on layered materials. Kopta and Salmeron (2000) removed layers from a mica surface by scraping it with large normal forces of few hundreds of nN. Fourier-filtered images acquired on very small areas revealed the different periodicities of the underlying layers, which reflect the complex structure of the substrate. To interpret their results, Kopta et al. assumed that wear is initiated

by atomic defects. When the defects accumulate beyond a critical concentration, they grow to form the scars observed by AFM. This process can be once again thermally activated. The number of defects created in the contact area A (F_N) is

$$N_{\text{def}}(F_N) = t_{\text{res}} n_0 A(F_N) f_0 \exp\left(-\frac{\Delta E}{k_B T}\right), \quad (10.6.2)$$

where t_{res} is the residence time of the tip, n_0 is the surface density of atoms, and f_0 is the frequency of attempts to overcome the energy barrier ΔE to break a Si–O bond, which depends on the applied load. When the defect density reaches a critical value, a hole is nucleated.

10.7 Non-contact Friction

As discussed in Sect. 10.1.3 forces and dissipation can be measured by recording Δf and A_{exc} simultaneously while scanning a surface by NC-AFM. Earlier experiments showed true atomic contrast in topography in the dissipation signal A_{exc} (Guggisberg et al. 2000; Loppacher et al. 2000; Bennewitz et al. 2000, 2001). Figure 10.22a shows the border of an ultrathin NaCl island grown on Cu(111). The strongly enhanced contrast of the step edges and kink sites could be attributed to a slower decay of the electric field and to easier relaxation of the positions of the ions at these locations. The dissipation image shown in Fig. 10.22b was recorded at the same time. Depending on the structure of the tip apex, the topography and damping signal were found to be in phase, partially shifted or even in antiphase. This impressively confirmed by the tip change observed in the experimental images in Fig. 10.22b. The dissipation contrast is seriously enhanced, while the topography contrast remains almost unchanged.

More directly related to friction measurements, where the tip is sliding in contact with the sample, are NC-AFM experiments in the pendulum geometry (Sect. 10.1.3). In the first measurements of this kind Stowe et al. (Stowe et al.

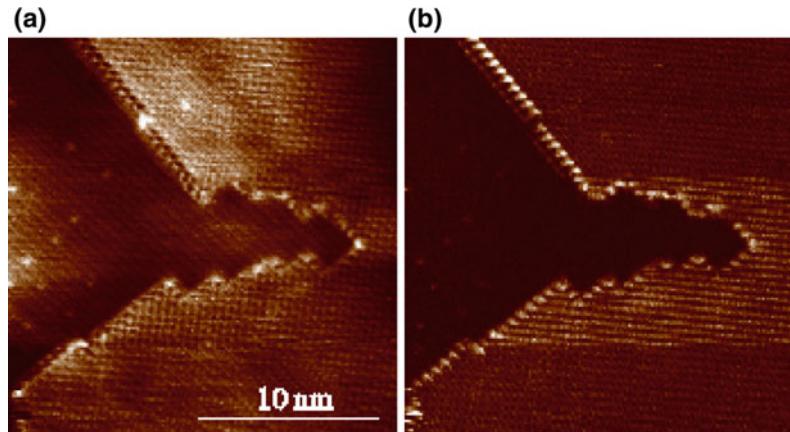


Fig. 10.22 **a** Topography and **b** damping images of a NaCl island on Cu(111). The tip changed after 1/4 of the scan, thereby enhancing the contrast in A_{exc} . After 2/3 of the scan, the contrast in the lower part of the image was restored, indicating a reversible tip change (after Bennewitz et al. 2000)

1999) could image the dopant density for *n*- and *p*-type silicon samples with 150 nm spatial resolution. The damping turned out to scale as U_{ts}^2 with the tip-sample voltage as expected for electric field Joule dissipation (Denk and Pohl 1991). Stipe et al. (2001) measured the noncontact friction between a Au(111) surface and a gold-coated tip using the same setup. They observed the same U_{ts}^2 dependence of the bias voltage and a distance dependence that follows the power law $1/d^n$, where n is between 1.3 and 3. Note that a considerable electric field was present even with zero external bias voltage, which is frequently the case in this kind of measurements.

In order to distinguish phononic and electronic contribution to friction Kisiel et al. (2011) oscillated a pendulum cantilever over a 150 nm thick niobium film. The measurement was performed at two different temperatures, thus allowing to investigate both the superconducting state of Nb, where the phononic channel of energy dissipation prevails, as well as the normal metal state where an electronic contribution to friction is superimposed. The dependence of the damping on the temperature is shown in Fig. 10.23a. During the measurement the tip-sample distance was kept constant. The non-contact friction coefficient rose by a factor of three when the critical temperature $T_c = 9.2$ K was approached from below and

leveled off in the normal state. The variation of the friction coefficient in the vicinity of T_c is smooth as predicted by the BCS (Bardeen-Cooper-Schrieffer) theory of superconductivity (Bardeen et al. 1957). In the superconducting phase, the electrons lose their identity and form Cooper pairs. Therefore, the energy is lost only in the emission of phonons. Close to the transition temperature the gradually growing normal electron population causes acoustic wave attenuation rises due to electron-phonon scattering, which ultimately results in dissipation. Further measurements of voltage V and distance d dependent friction support the hypothesis of separation of phononic and electronic channels.

The idea of replacing a passive substrate with an “active” material undergoing a phase transition for purpose of friction study was further accomplished on a SrTiO₃ sample. According to Benassi et al. (2011) the friction coefficient is peaking when a sliding object undergoes a phase transition. In this context SrTiO₃ is a perfect candidate, because of the presence of a structural second order phase transition at $T = 110$ K. A recent pendulum AFM experiment (Kisiel et al. 2015) have shown that the critical fluctuations in a bulk crystal may affect the dissipation of mechanical probes even if completely external to the crystal surface. NC-AFM dissipation bears clear evidence of a most classic bulk transition,

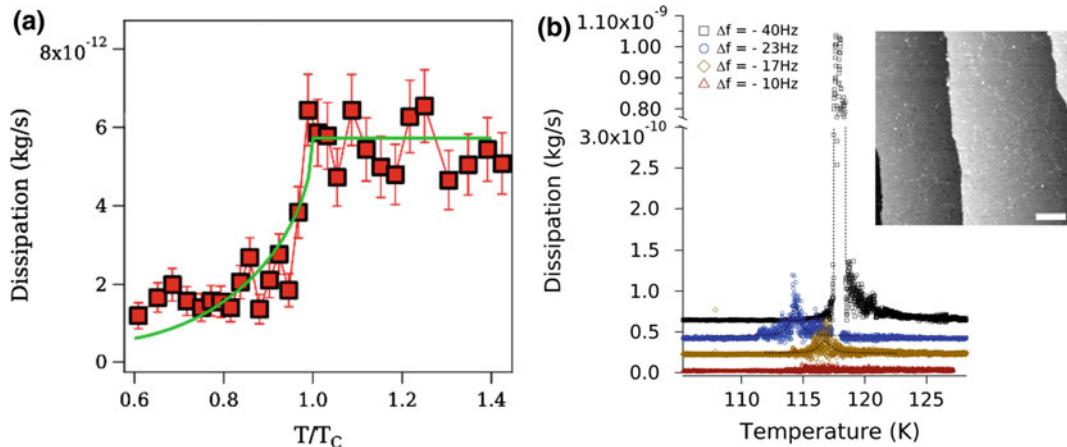


Fig. 10.23 **a** Temperature dependence of non-contact friction across the critical temperature of a Nb sample. The green line is a BCS theory fit to the measured data (red squares) (after Kisiel et al. 2011) and **b** experimental AFM dissipation as a function of temperature and tip—SrTiO₃ sample distance. The black color corresponds to the closest approach and the sharp peak is located at the critical temperature of SrTiO₃ in the bulk region under the tip. Inset shown a low temperature ($T = 5\text{ K}$) STM image of SrTiO₃ surface (after Kisiel et al. 2015)

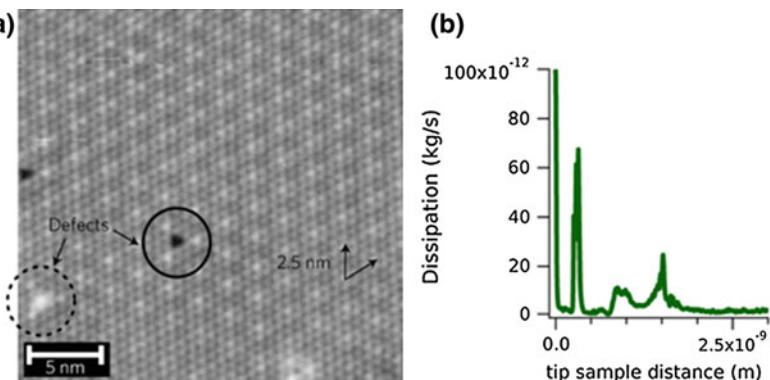
that to the antiferrodistortive state of SrTiO₃, as it is shown in Fig. 10.23b. It also showed that the friction coefficient is peaking at temperatures 110–130 K. In this range the system becomes very “soft” owing to proliferation of domain walls, defects and disorder in the distortive part of the crystal structure. Moreover the frictional response close to the transition was governed by critical exponents characteristic for the 3D Ising model.

Pendulum AFM experiments have also addressed the coupling of mechanical oscillator to *charge-density-waves* (CDWs) on a NbSe₂ sample, an elusively subtle long-range property of matter (Langer et al. 2013). Figure 6.3a shows an atomically resolved surface topography with the additional CDW induced moiré pattern, which is incommensurate with underlying lattice. The distance dependence of the friction coefficient $\Gamma(d)$ shows striking multiple dissipation peaks at few nm distance from the surface (Fig. 10.24b). The temperature evolution of the peaks suggested a strong connection with CDWs. In order to understand the origin of the dissipation spikes a theoretical model was proposed, where a CDW is considered to be an elastic medium, whose phase is perturbed locally by the attractive potential of the cantilever tip

(Pellegrini et al. 2014). This is shown in Fig. 10.24c. 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 a critical value the central peak disappears and an extra 2π phase is pumped locally below the cantilever tip (green line). The energies for the two discussed phase deformations ($N = 0$ and $N = 1$) are shown in red and green solid lines and a crossover energy is marked by a circle. Oscillations close to a crossover point will cause pumping in and out an extra 2π phase and thus be accompanied by a hysteresis cycle in the tip mechanics. The hysteresis explains the giant size of dissipation peaks and approaching the tip further down, might produce another phase slip explaining multiple dissipation peaks.

Lateral oscillations can be also excited in normal AFM geometry when exciting the torsional oscillation of commercial, rectangular AFM cantilevers. Also in this case the tip is oscillating approximately parallel to the surface. This mode, although much less force sensitive than the pendulum AFM, allowed to measure lateral forces acting on the tip at step edges and near impurities quantitatively (Pfeiffer et al. 2002). Enhanced energy dissipation was also

Fig. 10.24 **a** STM image of a NbSe₂ surface at T = 6 K and **b** energy dissipation as a function of tip-sample distance. Three giant dissipation peaks are visible (after Langer et al. 2013)



observed at the impurities. When the tip was moved further toward the sample, the contact formation transformed the nearly free torsional oscillation of the cantilever into a different mode, with the tip-sample contact acting as a hinge. This caused a rapid increase in the power required to maintain a constant tip oscillation amplitude and a positive frequency shift.

10.8 Single Molecule Friction

The experiments described so far focused on the friction forces between a sharp tip and a crystal surface either in contact or in close proximity to each other. In this Section we will address the problem of the frictional response of a nano-object as small as a single molecule sliding on a surface. Such experiments require to intentionally pick up the molecule previously deposited on a solid surface onto the apex of an AFM tip and to manipulate it laterally or vertically. This can be also done using quartz tuning-forks as force sensor (QPlus) operated in NC mode. The high stiffness (~ 2 kN/m) of the sensor allows to achieve sub-angstrom oscillation amplitudes. Furthermore, low temperature conditions are mandatory to ensure the stability of the molecule-surface junction while performing the friction experiments.

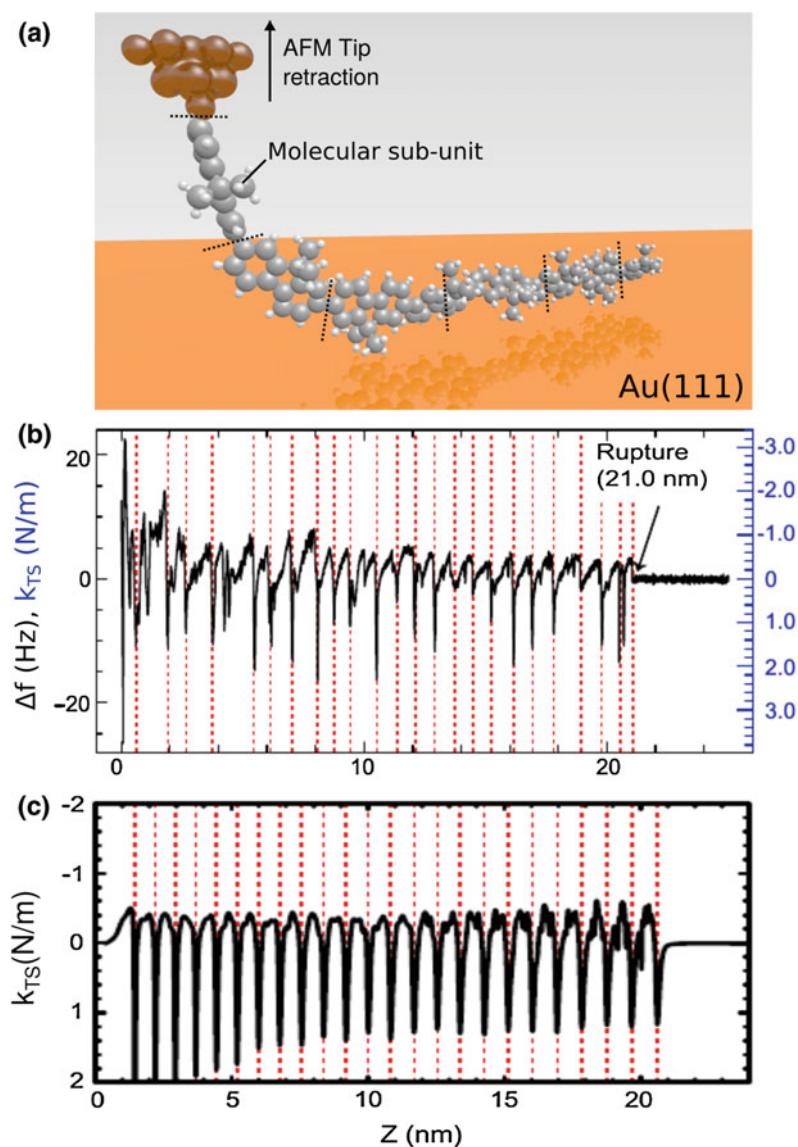
Ternes *et al.* first succeeded in measuring “*the force to move an adatom on a surface*” by means of force spectroscopy (Ternes *et al.* 2008; Emmrich *et al.* 2015). First attempts to measure the force accompanying the lifting of a PTCDA

molecule on Au(111) were conducted by the group of Temirov (Fournier *et al.* 2011; Wagner *et al.* 2014). Systematic vertical manipulations demonstrated reproducible features of the force curves which were related to the detachment of the molecule, i.e. adhesion, as well as to the mechanics of the molecular junction.

To investigate the mechanical degrees of freedom of a molecular structure during vertical displacements, Kawai *et al.* performed pulling manipulation experiments of long polymeric chains (Kawai *et al.* 2014). The chains were synthesized on a Au(111) surface by Ullmann-coupling of several dibromoterfluorene (DBTF) molecules previously adsorbed on the surface. A single chain, having a length of few nm and lying flat on the surface, is bound to the tip of an AFM as shown in Fig. 10.25a. Tip retraction results in vertical manipulation of the molecular chain. During such lifting, force-distance $\Delta F(z)$ curves (Sect. 10.1.3) were acquired (Fig. 10.25b). The effective stiffness, calculated as the derivative of the curve, is indicated on the right scale. A periodic variation is observed, with each period terminated by an abrupt jump. The number of periods (24) exactly corresponds to the number of the fluorene units forming the chain, as previously estimated by STM, and the periodicity of about 0.9 nm corresponds to the length of a single unit. The chain thus detached step-wise from the surface during pulling.

The experimental results were interpreted within an *edge-driven Frenkel-Kontorova* (FK) model. The polymer chain was schematized as a series of particles connected by springs of

Fig. 10.25 **a** On-surface Ullmann coupling of DBTF molecules leads to long polymeric chains on Au(111). The end of the chain can be attached to the AFM tip and then retracted from the surface while acquiring interactions forces, **b** experimental $\Delta f(z)$ spectra obtained during the lifting of the macromolecule (f_0 is the free resonance frequency of the tuning fork and k , its spring constant). The detachment is marked by the arrow, and **c** calculated dependence of the normal force gradient k_{TS} versus tip-surface distance z according the extended Frenkel-Kontorova model (after Kawai et al. 2014)

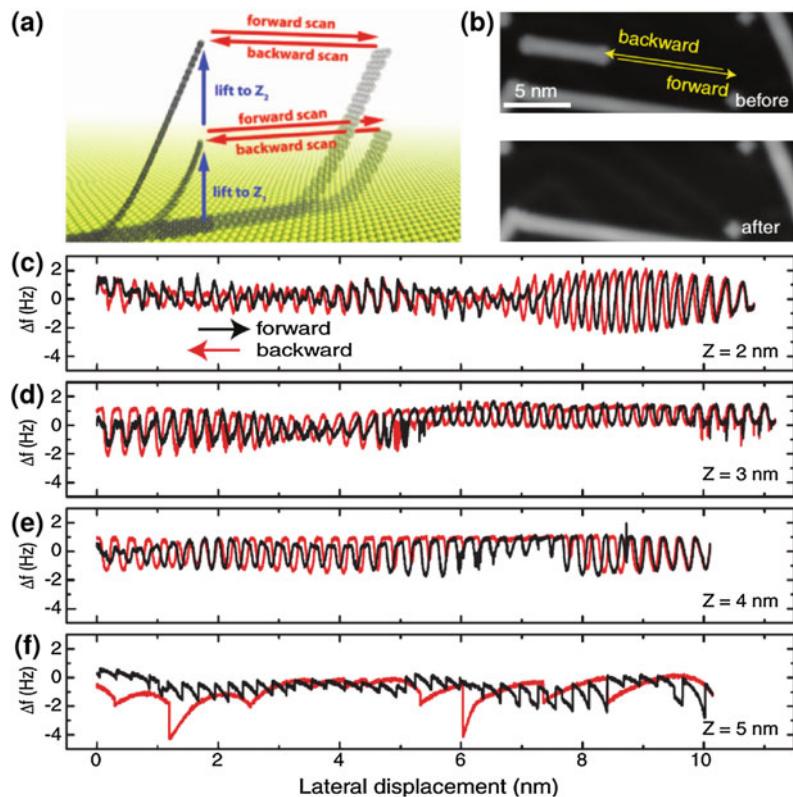


stiffness k and equilibrium length b . Each particle interacts with the substrate via a periodic potential reproducing the symmetry of the Au(111) surface and expressed by a formula introduced by Steele to describe gas-atoms solid interaction (Steele 1973). As shown in Fig. 10.25c, the main features of the experimental $\Delta f(z)$ curves were satisfactorily reproduced in such a way. Simulations also pointed out the fact that the effective stiffness of the pulled segments must be large compared to the stiffness of the chemical bond

formed between tip and chain. Remarkably, they also allowed to estimate the normal force required to detach one unit, which cannot be directly measured in the experiment. This force (in the order of 0.25 nN) is nearly independent of the length of the chain as well as of the unit position in the chain.

In another investigation Kawai et al. were able to measure the friction force acting on graphene nanoribbons (GNRs) sliding on a Au(111) surface using a similar approach (Kawai et al.

Fig. 10.26 Lateral manipulation of a GNR on a Au(111) surface. **a** Schematic drawing of the manipulation protocol at different tip-sample separations, **b** STM topographies before and after removal of a GNR moved on the Au(111) surface in the direction of the yellow arrows. **c–f** Frequency shifts accompanying the lateral motion of another GNR at different distances from the surface ($z = 2, 3, 4$ and 5 nm) (after Kawai et al. 2016)

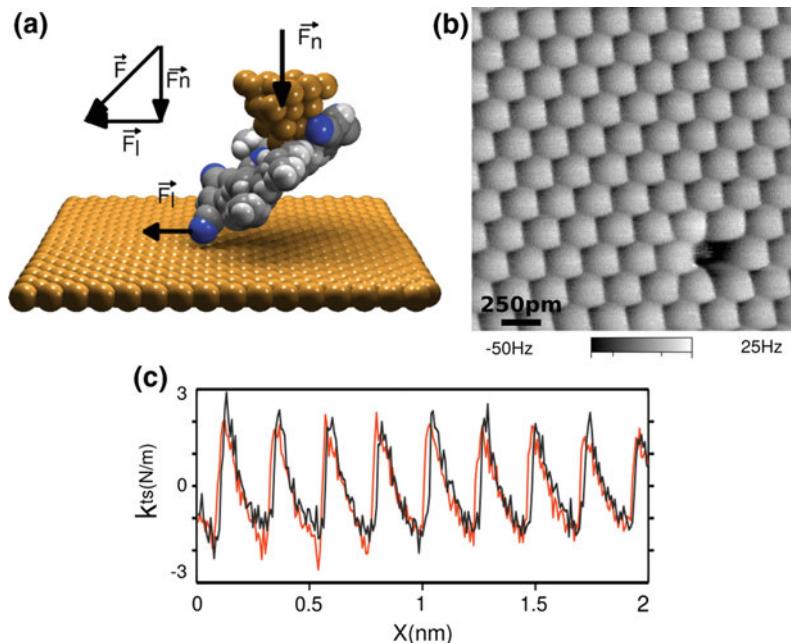


2016). The GNRs were again produced by Ullman-coupling, and first manipulated without attaching them to the tip. The GNRs started to move in correspondence to exceptionally low friction force values of few 100 pN. The static friction per unit length was found to decrease with increasing GNR length, consistently with a strong superlubric behavior of graphene when forming incommensurate contact. As a next step, the ends of selected GNRs were anchored to the tip and dragged back and forth in a controlled way (Fig. 10.26a). The corresponding variations of $\Delta f(x)$ are shown in Fig. 10.26c, d for increasing values of the tip-sample separation z . The frequency shift oscillates with a periodicity of 0.28 nm, except when the GNR was driven backward with $z = 5 \text{ nm}$. The amplitude of the Δf oscillations is not constant along x but is modulated on distances of few nm. Molecular dynamics (MD) simulations were also employed to reproduce the manipulation process. In this

way the herringbone reconstruction of Au(111) was found to be responsible not only for the modulation of $\Delta f(x)$, but also for a significant reduction of commensurability between GNR and substrate. As a consequence, the GNR-Au(111) contact is almost superlubric, with both static and kinetic friction force values in the range of 100 pN.

The frictional response of a single organic molecule sliding on a metal surface was recently investigated by Pawlak et al. (2016). In this experiment, a porphyrin molecule attached to the AFM tip *via* its core as shown in Fig. 10.27a. The surface-molecule junction was formed by a single carbonitrile end-group of the porphyrin structure resulting in a single-point contact with a Cu(111) surface. While dragging the molecule, the junction was robust enough to maintain the contact although mechanical deformation of the molecular spring was expected. A friction $\Delta f(x, y)$ map of the Cu(111) surface obtained in such

Fig. 10.27 Single molecule sliding on a metal substrate:
a a porphyrin is attached to an AFM tip by vertical manipulation and dragged over a Cu(111) surface,
b constant-height $\Delta f(x, y)$ map of the Cu(111) surface revealing a typical stick-slip pattern and **c** cross section showing the saw tooth variation (after Pawlak et al. 2016)



conditions is shown in Fig. 10.27b. The atomic-scale contrast resembles the frictional patterns obtained by conventional FFM on Cu (111) (Bennewitz et al. 1999). A dark depression corresponding to a Cu vacancy in the bottom right part of the image was also observed. In the line profiles of the tip-sample stiffness $k_{TS}(x)$ extracted from the Δf data (Fig. 6.3c) a clear saw-tooth modulation is detected resembling standard atomic-scale friction pattern in FFM. Here, the hysteresis of the friction loop between forward and backward traces is small suggesting superlubric sliding. To verify this assumption, a generalized PT model including the Steele potential was used to simulate the friction process (Steele 1973). The intrinsic parameters characterizing the relevant stretching and bending rigidities of the molecule structure were extracted from first-principles calculations and used to parameterize the simulations. The calculations demonstrated that the variation of the normal force gradient during sliding is coupled to the friction force acting on the CN end group via the molecular entity thus supporting the experimental hypothesis.

10.9 Conclusion

Over the last 25 years, the establishment of friction force microscopy has led to exceptional development in the field of nanotribology. The sliding motion of single-asperities with only few atoms forming the contact can be precisely investigated, and the friction forces accompanying the motion can be studied as a function of time, but also of load, velocity and temperature. Stick-slip motion on the atomic scale is properly described by the Prandtl-Tomlinson model. Extensions of this model predict superlubric transitions and thermally activated slip, both of which have been experimentally verified. Friction force microscopy can also be applied to investigate initial stages of abrasive wear. On the other side, non-contact friction has recently become an active topic of research thanks to instrumental refinements such as pendulum AFM and tuning fork sensors. Distinguishing phononic and electronic contributions to non-contact friction is nowadays possible, as well as estimating the friction forces acting on polymer chains,

graphene ribbons and even single molecules dragged or lifted off a solid surface.

References

- Albrecht TR, Grüter P, Horne D, Rugar D (1991) Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity. *J Appl Phys* 69:668
- Bardeen J, Cooper LN, Schrieffer JR (1957) Theory of Superconductivity. *Phys Rev* 108:1175
- Barel I, Urbakh M, Jansen L, Schirmeisen A (2010) Multibond dynamics of nanoscale friction: the role of temperature. *Phys Rev Lett* 104:066104
- Benassi A, Vanossi A, Santoro GE, Tosatti E (2011) Sliding over a phase transition. *Phys Rev Lett* 106:256102
- Bennewitz R, Gyalog T, Guggisberg M, Bammerlin M, Meyer E, Güntherodt H-J (1999) Atomic-scale stick-slip processes on Cu(111). *Phys Rev B* 60:R11301–R11304
- Bennewitz R, Foster AS, Kantorovich LN, Bammerlin M, Loppacher C, Schär S, Guggisberg M, Meyer E, Shluger AL (2000) Atomically resolved edges and kinks of NaCl islands on Cu(111): experiment and theory. *Phys Rev B* 62:2074–2084
- Bennewitz R, Schär S, Barwich V, Pfeiffer O, Meyer E, Krok F, Such B, Kolodzej J, Szymonski M (2001) Atomic-resolution images of radiation damage in KBr. *Surf Sci* 474:197–202
- Bouhacina T, Aimé JP, Gauthier S, Michel D, Heroguez V (1997) Tribological behavior of a polymer grafted on silanized silica probed with a nanotip. *Phys Rev B* 56:7694–7703
- Bowden FP, Tabor FP (1950) The friction and lubrication of solids. Oxford Univ. Press, Oxford
- Butt HJ, Jaschke M (1995) Calculation of thermal noise in atomic-force microscopy. *Nanotechnology* 6:1–7
- Carpick RW, Agrait N, Ogletree DF, Salmeron M (1996) Measurement of interfacial shear (friction) with an ultrahigh vacuum atomic force microscope. *J Vac Sci Technol B* 14:1289–1295
- Carpick RW, Ogletree DF, Salmeron M (1997) Lateral stiffness: a new nanomechanical measurement for the determination of shear strengths with friction force microscopy. *Appl Phys Lett* 70:1548–1550
- Cleveland J, Manne S, Bocek D, Hansma PK (1993) A nondestructive method for determining the spring constant of cantilevers for scanning force microscopy. *Rev Sci Instrum* 64:403–405
- Denk W, Pohl DW (1991) Local electrical dissipation imaged by scanning force microscopy. *Appl Phys Lett* 59:2171–2173
- Derjaguin BV, Muller VM, Toporov YP (1975) Effect of contact deformations on adhesion of particles. *J Colloid Interface Sci* 53:314–326
- Dienwiebel M, Verhoeven G, Pradeep N, Frenken J, Heimberg J, Zandbergen H (2004) Superlubricity of graphite. *Phys Rev Lett* 92:126101
- Edwards SA, Ducker WA, Sader JE (2008) Influence of atomic force microscope cantilever tilt and induced torque on force measurements. *J Appl Phys* 103:064513
- Emmrich M, Schneiderbauer F, Huber AJ, Weymouth N, Okabayashi FJ (2015) Giessibl Force field analysis suggests a flowerling of diffusion barriers in atomic manipulation due to presence of STM tip. *Phys Rev Lett* 114:146101
- Enachescu M, van den Oetelaar RJA, Carpick RW, Ogletree DF, Flipse CFJ, Salmeron M (1998) Atomic force microscopy study of an ideally hard contact: the diamond(111)/tungsten carbide interface. *Phys Rev Lett* 81:1877–1880
- Enachescu M, van den Oetelaar RJA, Carpick RW, Ogletree DF, Flipse CFJ, Salmeron M (1999) Observation of proportionality between friction and contact area at the nanometer scale. *Tribol Lett* 7:73–78
- Fessler G, Zimmermann I, Glatzel T, Gnecco E, Steiner P, Roth R, Keene TD, Liu SX, Decurtins S, Meyer E (2011) Orientation dependent molecular friction on organic layer compound crystals. *Appl Phys Lett* 98:083119
- Filleter T, Bennewitz R (2010) Structural and frictional properties of graphene lms on SiC(0001) studied by atomic force microscopy. *Phys Rev B* 81:155412
- Filleter T, Paul W, Bennewitz R (2008) Atomic structure and friction of ultrathin films of KBr on Cu(100). *Phys Rev B* 77:035430
- Fournier N, Wagner C, Weiss C, Temirov R, Tautz FS (2011) Force-controlled lifting of molecular wires. *Phys Rev B* 84:035435
- Fujisawa S, Kishi E, Sugawara Y, Morita S (1995) Atomic-scale friction observed with a two-dimensional frictional-force microscope. *Phys Rev B* 51:7849–7857
- Fusco C, Fasolino A (2005) Velocity dependence of atomic-scale friction: a comparative study of the one- and two-dimensional Tomlinson model. *Phys. Rev. B* 71:45413
- Gnecco E (2010) Quasi-isotropy of static friction on hexagonal surface lattices. *Europhys Lett* 91:66008
- Gnecco E, Bennewitz R, Gyalog T, Loppacher C, Bammerlin M, Meyer E, Güntherodt H-J (2000) Velocity dependence of atomic friction. *Phys Rev Lett* 84:1172–1175
- Gnecco E, Bennewitz R, Meyer E (2002) Abrasive wear on the atomic scale. *Phys Rev Lett* 88:215501
- Gnecco E, Roth R, Baratoff A (2012) Analytical expressions for the kinetic friction in the Prandtl-Tomlinson model. *Phys Rev B* 86:035443
- Goryl M, Budzioch J, Krok F, Wojtaszek M, Kolmer M, Walczak L, Konior J, Gnecco E, Szymonski M (2012) Probing atomic-scale friction on reconstructed surfaces of single-crystal semiconductors. *Phys Rev B* 85:085308

- Gotsmann B (2011) Tribology: sliding on vacuum. *Nat Mater* 10:87–88
- Gotsmann B, Seidel C, Anczykowski B, Fuchs H (1999) Conservative and dissipative tip–sample interaction forces probed with dynamic AFM. *Phys Rev B* 60:11051–11061
- Greenwood JA, Williamson JBP (1966) Contact of nominally flat surfaces. *Proc R Soc Lond A* 295:300
- Guggisberg M, Bammerlin M, Baratoff A, Lüthi R, Loppacher C, Battiston FM, Lü J, Bennewitz R, Meyer E, Güntherodt HJ (2000) Dynamic force microscopy across steps on the Si(111)-(7 × 7) surface. *Surf Sci* 461:255–265
- Gyalog T, Thomas H (1997) Friction between atomically flat surfaces. *Europhys Lett* 37:195–200
- Gyalog T, Bammerlin M, Lüthi R, Meyer E, Thomas H (1995) Mechanism of atomic friction. *Europhys Lett* 31:269–274
- Howald L, Meyer E, Lüthi R, Haefke H, Overney R, Rudin H, Güntherodt H-J (1993) Multifunctional probe microscope for facile operation in ultrahigh vacuum. *Appl Phys Lett* 63:117–119
- Hutter JL, Bechhoefer J (1993) Calibration of atomic-force microscope tips. *Rev Sci Instrum* 64:1868–1873
- Israelachvili JN, Tabor D (1972) Measurement of van der Waals dispersion forces in range 1.5 to 130 nm. *Proc R Soc Lond A* 331:19
- Jacobs TDB, Carpick R (2013) Nanoscale wear as a stress-assisted chemical reaction. *Nat Mater* 8:108
- Jansen L, Hölscher H, Fuchs H, Schirmeisen A (2010) Temperature dependence of atomic-scale stick-slip friction. *Phys Rev Lett* 104:256101
- Johnson KL (1985) Contact mechanics. Cambridge Univ. Press, Cambridge
- Johnson KL, Kendall K, Roberts AD (1971) Surface energy and contact of elastic solids. *Proc R Soc Lond A* 324:301
- Kawai S, Koch M, Gnecco E, Sadeghi A, Pawlak R, Glatzel T, Schwarz J, Goedecker S, Hecht S, Baratoff A, Meyer E (2014) Quantifying the atomic-level mechanics of single long physisorbed molecular chains. *Proc Natl Acad Sci USA* 111:3968–3972
- Kawai S, Benassi A, Gnecco E, Söde H, Pawlak R, Feng X, Müllen K, Passerone D, Pignedoli CA, Ruffieux P, Fasel R, Meyer E (2016) Superlubricity of graphene nanoribbons on gold surfaces. *Science* 351:957
- Kawakatsu H, Saito T (1996) Scanning force microscopy with two optical levers for detection of deformations of the cantilever. *J Vac Sci Technol, B* 14:872–876
- Kisiel M, Gnecco E, Gysin U, Marot L, Rast S, Meyer E (2011) Suppression of electronic friction on Nb films in the superconducting state. *Nat Mater* 10:119
- Kisiel M, Pellegrini F, Santoro GE, Samadashvili M, Pawlak R, Benassi A, Gysin U, Buzio R, Gerbi A, Meyer E, Tosatti E (2015) Noncontact atomic force microscope dissipation reveals a Central peak of SrTiO₃ structural phase transition. *Phys Rev Lett* 115:046101
- Kopta S, Salmeron M (2000) The atomic scale origin of wear on mica and its contribution to friction. *J Chem Phys* 113:8249–8252
- Krylov SY, Jinesh KB, Valk H, Dienwiebel M, Frenken JWM (2005) Thermally induced suppression of friction at the atomic scale. *Phys Rev E* 71:65101
- Landau LD, Lifshitz EM (1998) Introduction to theoretical physics, vol 7. Nauka, Moscow
- Langer M, Kisiel M, Pawlak R, Pellegrini F, Santoro GE, Buzio R, Gerbi A, Balakrishnan G, Baratoff A, Tosatti E, Meyer E (2013) Giant frictional dissipation peaks and charge-density-wave slips at the NbSe₂ surface. *Nat Mater* 13:173
- Lantz MA, O’Shea SJ, Welland ME, Johnson KL (1997) Atomic-force-microscope study of contact area and friction on NbSe₂. *Phys Rev B* 55:10776–10785
- Linnemann R, Gotszalk T, Rangelow IW, Dumania P, Oesterschulze E (1996) Atomic force microscopy and lateral force microscopy using piezoresistive cantilevers. *J Vac Sci Technol, B* 14:856–860
- Loppacher C, Bennewitz R, Pfeiffer O, Guggisberg M, Bammerlin M, Schär S, Barwich V, Baratoff A, Meyer E (2000) Experimental aspects of dissipation force microscopy. *Phys Rev B* 62:13674–13679
- Maier S, Sang Y, Filleter T, Grant M, Bennewitz R, Gnecco E, Meyer E (2005) Fluctuations and jump dynamics in atomic friction experiments. *Phys Rev B* 72:245418
- Maier S, Gnecco E, Baratoff A, Meyer E (2008) Atomic-scale friction modulated by a buried interface: combined atomic and friction force microscopy experiments. *Phys Rev B* 78:045432
- Marti O, Colchero J, Mlynek J (1990) Combined scanning force and friction microscopy of mica. *Nanotechnology* 1:141–144
- Mate CM, McClelland GM, Erlandsson R, Chiang S (1987) Atomic-scale friction of a tungsten tip on a graphite surface. *Phys Rev Lett* 59:1942–1945
- Maugis D (1992) Adhesion of spheres: the JKR-DMT transition using a Dugdale model. *J Colloid Interface Sci* 150:243–269
- McClelland GM, Glosli JN (1992) Friction at the atomic scale. In: Singer L, Pollock HM (eds) NATO ASI Series E, vol 220. Kluwer, Dordrecht, pp 405–425
- Meyer G, Amer N (1990) Simultaneous measurement of lateral and normal forces with an optical-beam-deflection atomic force microscope. *Appl Phys Lett* 57:2089–2091
- Meyer E, Lüthi R, Howald L, Bammerlin M, Guggisberg M, Güntherodt H-J (1996) Site-specific friction force spectroscopy. *J Vac Sci Technol, B* 14:1285–1288
- Müser MH (2004) Structural lubricity: role of dimension and symmetry. *Europhys Lett* 66:97
- Müser MH, Robbins MO (2000) Conditions for static friction between flat crystalline surfaces. *Phys Rev B* 61:2335

- Neubauer G, Cohen SR, McClelland GM, Horn DE, Mate CM (1990) Force microscopy with a bidirectional capacitance sensor. *Rev Sci Instrum* 61:2296–2308
- Neumeister JM, Ducker WA (1994) Lateral, normal, and longitudinal spring constants of atomic-force microscopy cantilevers. *Rev Sci Instrum* 65:2527–2531
- Nita P, Casado S, Dietzel D, Schirmeisen A, Gnecco E (2013) Spinning and translational motion of Sb nanoislands manipulated on MoS₂. *Nanotechnol* 24:325302
- Nonnenmacher M, Greschner J, Wolter O, Kassing R (1991) Scanning force microscopy with micromachined silicon sensors. *J Vac Sci Technol, B* 9:1358–1362
- Ogletree DF, Carpick RW, Salmeron M (1996) Calibration of frictional forces in atomic force microscopy. *Rev Sci Instrum* 67:3298–3306
- Pawlak R, Ouyang W, Filippov AE, Kalikhman-Razvozov L, Kawai S, Glatzel T, Gnecco E, Baratoff A, Zheng Q-S, Hod O, Urbakh M, Meyer E (2016) Single molecule tribology: force microscopy manipulation of a porphyrin derivative on a copper surface. *ACS Nano* 10:713–722
- Pellegrini F, Santoro GE, Tosatti E (2014) Charge-density-wave surface phase slips and noncontact nanofriction. *Phys Rev B* 89:245416
- Persson BNJ (2001) Elastoplastic contact between randomly rough surfaces. *Phys Rev Lett* 87:116101
- Pfeiffer O, Bennewitz R, Baratoff A, Meyer E, Grüter P (2002) Lateral-force measurements in dynamic force microscopy. *Phys Rev B* 65:161403
- Pina CM, Miranda R, Gnecco E (2012) Anisotropic surface coupling while sliding on dolomite and calcite crystals. *Phys Rev B* 85:073402
- Polaczyk C, Schneider T, Schöfer J, Santner E (1998) Microtribological behavior of Au(001) studied by AFM/FFM. *Surf Sci* 402:454–458
- Prandtl L (1928) A conceptual model for the kinetic theory of solids (in German). *Z Angew Math Mech* 8:85
- Riedo E, Lévy F, Brune H (2002) Kinetics of capillary condensation in nanoscopic sliding friction. *Phys Rev Lett* 88:185505
- Roth R, Glatzel T, Steiner P, Gnecco E, Baratoff A, Meyer E (2010) Multiple slips in atomic-scale friction: an indicator for the lateral contact damping. *Tribol Lett* 39:63
- Roth R, Fajardo OY, Mazo JJ, Meyer E, Gnecco E (2014) Lateral vibration effects in atomic-scale friction. *Appl Phys Lett* 104:083103
- Sader JE, Green CP (2004) In-plane deformation of cantilever plates with applications to lateral force microscopy. *Rev Sci Instrum* 75:878–883
- Sang Y, Dubé M, Grant M (2001) Thermal effects on atomic friction. *Phys Rev Lett* 87:174301
- Schwarz UD, Köster P, Wiesendanger R (1996) Quantitative analysis of lateral force microscopy experiments. *Rev Sci Instrum* 67:2560–2567
- Schwarz UD, Zwörner O, Köster P, Wiesendanger R (1997) Quantitative analysis of the frictional properties of solid materials at low loads. *Phys Rev B* 56:6987–6996
- Socoliuc A, Bennewitz R, Gnecco E, Meyer E (2004) Transition from stick-slip to continuous sliding in atomic friction: Entering a new regime of ultralow friction. *Phys Rev Lett* 92:134301
- Socoliuc A, Gnecco E, Maier S, Pfeiffer O, Baratoff A, Bennewitz R, Meyer E (2006) Atomic-scale control of friction by actuation of nanometer-sized contacts. *Science* 313:207–210
- Steele WA (1973) The physical interaction of gases with crystalline solids: I. Gas-solid energies and properties of isolated adsorbed atoms. *Surf Sci* 36:317
- Steiner P, Roth R, Gnecco E, Glatzel T, Baratoff A, Meyer E (2009) Modulation of contact resonance frequency accompanying atomic-scale stickslip in friction force microscopy. *Nanotechnology* 20:495701
- Steiner P, Roth R, Gnecco E, Baratoff A, Meyer E (2010) Angular dependence of static and kinetic friction on alkali halide surfaces. *Phys Rev B* 82:205417
- Steiner P, Gnecco E, Krok F, Budzioch J, Walczak L, Konior J, Szymonski M, Meyer E (2011) Atomic-scale friction on stepped surfaces of ionic crystals. *Phys Rev Lett* 106:186104
- Stipe BC, Mamin HJ, Stowe TD, Kenny TW, Rugar D (2001) Noncontact friction and force fluctuations between closely spaced bodies. *Phys Rev Lett* 87:96801
- Stowe TD, Kenny TW, Thomson J, Rugar D (1999) Silicon dopant imaging by dissipation force microscopy. *Appl Phys Lett* 75:2785–2787
- Tabor D (1977) Surface forces and surface interactions. *J Colloid Interface Sci* 58:2–13
- Ternes M, Lutz CP, Hirjibehedin CF, Giessibl FJ, Heinrich AJ (2008) The force needed to move an atom. *Science* 319:1066–1069
- Vilhena G, Pimentel C, Pedraz P, Luo F, Serena PA, Pina CM, Gnecco E, Perez R (2016) Atomic-scale sliding friction on grapheme in water. *ACS Nano* 10:4288–4293
- Villarrubia JS (1997) Algorithms for scanned probe microscope image simulation, surface reconstruction, and tip estimation. *J Res Natl Inst Stand Technol* 102:425–454
- Wagner C, Fournier N, Tautz FS, Temirov R (2014) The role of surface corrugation and tip oscillation in single-molecule manipulation with non-contact atomic force microscope. *Beilstein J Nanotechnol* 5:202–209
- Weiss M, Elmer FJ (1996) Dry friction in the Frenkel-Kontorova-Tomlinson model: Static properties. *Phys Rev B* 53:7539–7549
- Zhao X, Hamilton M, Sawyer WG, Perry SS (2007) Thermally activated friction. *Tribol Lett* 27:113–117

Bharat Bhushan and Michael Nosonovsky

Abstract

A model that explains scale effects in mechanical properties and tribology is presented. Mechanical properties are scale dependent based on the strain gradient plasticity and the effect of dislocation-assisted sliding. Both single asperity and multiple asperity contacts are considered. The relevant scaling length is the nominal contact length, which is the contact diameter for a single-asperity contact, and the scan length for multiple-asperity contacts. For multiple asperity contacts, based on an empirical power-rule for scale dependence of roughness, contact parameters are calculated. The effect of load on the contact parameters and the coefficient of friction is also considered. During sliding, adhesion and two- and three-body deformation, as well as ratchet mechanism, contribute to the dry friction force. These components of the friction force depend on the relevant real areas of contact (dependent on roughness and mechanical properties), average asperity slope, number of trapped particles, and shear strength during sliding. Scale dependence of the components of the coefficient of friction is studied. A scale dependent transition index, which is responsible for transition from predominantly elastic adhesion to plastic

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

B. Bhushan · M. Nosonovsky
Department of Mechanical Engineering, University
of Wisconsin at Milwaukee, EMS Building, Room
EMS981, 53211-0413 Milwaukee, WI, USA
e-mail: nosonovs@uwm.edu

deformation has been proposed. Scale dependence of the wet friction, wear, and interface temperature has been also analyzed. The proposed model is used to explain the trends in the experimental data for various materials at nanoscale and microscale, which indicate that nanoscale values of coefficient of friction are lower than the microscale values due to an increase of the three-body deformation and transition from elastic adhesive contact to plastic deformation.

Keywords

Scale effects • Length scale • Mechanical properties • Tribology • Friction • Wear • Lubrication

Nomenclature

$a, \bar{a}, \bar{a}_0, a_{max}, \bar{a}_{max}$

Contact radius, mean contact radius, macroscale value of mean contact radius, maximum contact radius, mean value of maximum contact radius

$A_a, A_r, A_{ra}, A_{re}, A_{re0}, A_{rp}, A_{rp0}, A_{ds}, A_{dp}$

Apparent area of contact, real area of contact, real area of contact during adhesion, real area of elastic contact, macroscale value of real area of elastic contact, real area of plastic contact, macroscale value of real area of plastic contact, real area of contact during asperity summit deformation, area of contact with particles

b

Burgers vector

c

Constant, specified by crystal structure

C_0

Constant required for normalization of $p(d)$

$d, d_e, d_n, d_{ln}, \bar{d}, \bar{d}_0$

Particle diameter, minimum for exponential distribution, mean for normal distribution, exponential of mean of $\ln(d)$ for log-normal distribution, mean trapped particles diameter, macroscale value of mean trapped particles diameter

Interface zone thickness

D, E_1, E_2, E^*

Elastic moduli of contacting bodies, effective elastic modulus

$F, F_a, F_d, F_{ae}, F_{ap}, F_{a\bar{p}}, F_{ds}, F_{dp}, F_m, F_{m0}$

Friction force, friction force due to adhesion, friction force due to deformation, friction force during elastic adhesional contact, plastic adhesional contact, summit deformation, particles deformation, meniscus force for wet contact, macroscale value of meniscus force

G	Elastic shear modulus
h	Indentation depth
h_f	Liquid film thickness
H, H_0	Hardness, hardness in absence of strain gradient
k, k_0	Wear coefficient, macroscale value of wear coefficient
l_s, l_d	Material-specific characteristic length parameters
$L, L_{lwb}, L_{lc}, L_s, L_d$	Length of the nominal contact zone, long wavelength limit for roughness parameters, long wavelength limit for contact parameters, length parameters related to l_s and l_d
L_p	Peclet number
m, n	Indices of exponents for scale-dependence of σ and β^*
n_{tr}	Number of trapped particles divided by the total number of particles
p_a, p_{ac}	Apparent pressure, critical apparent pressure
$p(d), p_{tr}(d)$	Probability density function for particle size distribution, probability density function for trapped particle size distribution
$P(d)$	Cumulative probability distribution for particle size
$R, R_p, \overline{R}_p, \overline{R}_{p0}$	Effective radius of summit tips, radius of summit tip, mean radius of summit tips, macroscale value of the mean radius of summit tips
$R(\tau)$	Autocorrelation function
s	Spacing between slip steps on the indentation surface
s_d	Separation distance between reference planes of two surfaces in contact
N, N_0	Total number of contacts, macroscale value of total number of contacts
T, T_0	Maximum flash temperature rise, macroscale value of temperature rise
x	Sliding distance
v	Volume of worn material
V	Sliding velocity
W	Normal load
z, z_{min}, z_{max}	Random variable, minimum and maximum value of z
α	Probability for a particle in the border zone to leave the contact region

β^*, β_0^*	Correlation length, macroscale value of correlation length
γ	Surface tension
Γ	Gamma function
ε	Strain
η	Density of particles per apparent area of contact
η_{int}, η_{cr}	Density of dislocation lines per interface area, critical density of dislocation lines per interface area
κ	Curvature
κ_t	Thermal diffusivity
θ	Contact angle between the liquid and surface
θ_i	Indentation angle
θ_r	Roughness angle
$\mu, \mu_a, \mu_{ae}, \mu_{ae0}, \mu_{ap}, \mu_{ap0}, \mu_d, \mu_{ds}, \mu_{ds0}, \mu_{dp}, \mu_{dp0}, \mu_r, \mu_{r0}, \mu_{re}, \mu_{re0}, \mu_{rp}, \mu_{rp0}, \mu_{wet}$	Coefficient of friction, coefficient of adhesional friction, coefficient of adhesional elastic friction, macroscale value of coefficient of adhesional elastic friction, coefficient of adhesional plastic friction, macroscale value of coefficient of adhesional plastic friction, coefficient of deformation friction, coefficient of summits deformation friction, macroscale value of coefficient of summits deformation friction, coefficient of particles deformation friction, macroscale value of coefficient of particles deformation friction, ratchet component of the coefficient of friction, macroscale value of ratchet component of the coefficient of friction, ratchet component of the coefficient of elastic friction, macroscale value of ratchet component of the coefficient of elastic friction, ratchet component of the coefficient of plastic friction, macroscale value of ratchet component of the coefficient plastic of friction, and coefficient of wet friction
v_1, v_2	Poisson's ratios of contacting bodies
ρc_p	Volumetric specific heat

$\sigma, \sigma_0, \sigma_e, \sigma_n, \sigma_{\ln}$

Standard deviation of rough surface profile height, macroscale value of standard deviation of rough surface profile height, standard deviation for the exponential distribution, standard deviation for the normal distributions, standard deviation for $\ln(d)$ of the log normal distribution

ρ, ρ_G, ρ_S

Total density of dislocation lines per volume, density of GND per volume, density of SSD per volume

ϕ, ϕ_0

Transition index, macroscale value of transition index

τ, τ_0

Spatial parameter, value at which the autocorrelation function decays

$\tau_a, \tau_{a0}, \tau_Y, \tau_{Y0}, \tau_{ds}, \tau_{ds0}, \tau_{dp}, \tau_{dp0}, \tau_p$

Adhesional shear strength during sliding, macroscale value of adhesional shear strength, shear yield strength, shear yield strength in absence of strain gradient, shear strength during summits deformation, macroscale value of shear strength during summits deformation, shear strength during particles deformation, macroscale value of shear strength during particles deformation, Peierls stress

11.1 Introduction

Microscale and nanoscale measurements of tribological properties, which became possible due to the development of the Surface Force Apparatus (SFA), Atomic Force Microscope (AFM), and Friction Force Microscope (FFM), demonstrate scale dependence of adhesion, friction, and wear as well as mechanical properties including hardness (Bhushan et al. 1995; Bhushan 1999a, 1999c, 2013, 2017). Advances of micro/nanoelectromechanical systems (MEMS/NEMS) technology in the early 2000s make understanding of scale effects in adhesion, friction, and wear especially important, since

surface to volume ratio grows with miniaturization and surface phenomena dominate. Dimensions of MEMS/NEMS devices range from about 1 mm to a few nm.

Experimental studies of scale dependence of tribological phenomena have been conducted recently. AFM experiments provide data on the nanoscale (Ruan and Bhushan 1994; Bhushan and Kulkarni 1996; Carpick et al. 1996; Schwarz et al. 1997; Bhushan and Sundararajan 1998; Bhushan and Dandavate 2000) whereas micro-tribo apparatus (Liu and Bhushan 2003; Bhushan et al. 2004) and SFA (Homola et al. 1990) provide data on the microscale. Experimental data indicate that wear mechanisms and wear rates are different at macro- and micro/nanoscales

(Koinkar and Bhushan 1997; Zhao and Bhushan 1998). During sliding, the effect of operating conditions such as load and velocity on friction and wear are frequently manifestations of the effect of temperature rise on the variable under study. The overall interface temperature rise is a cumulative result of numerous flash temperature rises at individual asperity contacts. Since contact size is scale dependent, it is expected that temperature rise at each contact also will be scale dependent.

Friction is a complex phenomenon that involves asperity interactions related to adhesion and deformation (plowing), Fig. 11.1. Adhesion and plastic deformation imply energy dissipation, which is responsible for friction. A contact between two bodies takes place on high asperities, and the real area of contact (A_r) is a small fraction of the apparent area of contact (Bhushan 2013). During contact of two asperities, a lateral force may be required for asperities of a given slope to climb against each other. This mechanism is known as ratchet mechanism, and it also contributes to the friction. Wear and contaminant particles present at the interface, referred as the “third body”, also contribute to friction, Fig. 11.2a. In addition, during contact, even at low humidity, a meniscus is formed. Generally, any liquid that wets or has a small contact angle on surfaces will condense from vapor into cracks and pores on surfaces as bulk liquid and in the form of annular-shaped capillary condensate in the contact zone. Figure 11.2b shows a random rough surface in contact with a smooth surface with a continuous liquid film on the smooth surface. The presence of the liquid film of the condensate or pre-existing film of the liquid can significantly increase the adhesion between the solid bodies (Bhushan 2013). The effect of a meniscus is scale-dependent.

A quantitative theory of scale effects in friction should consider scale effect on physical properties relevant to these contributions. However, conventional theories of contact and friction lack characteristic length parameters, which would be responsible for scale effects. The linear elasticity and conventional plasticity theories are scale-invariant and do not include any material

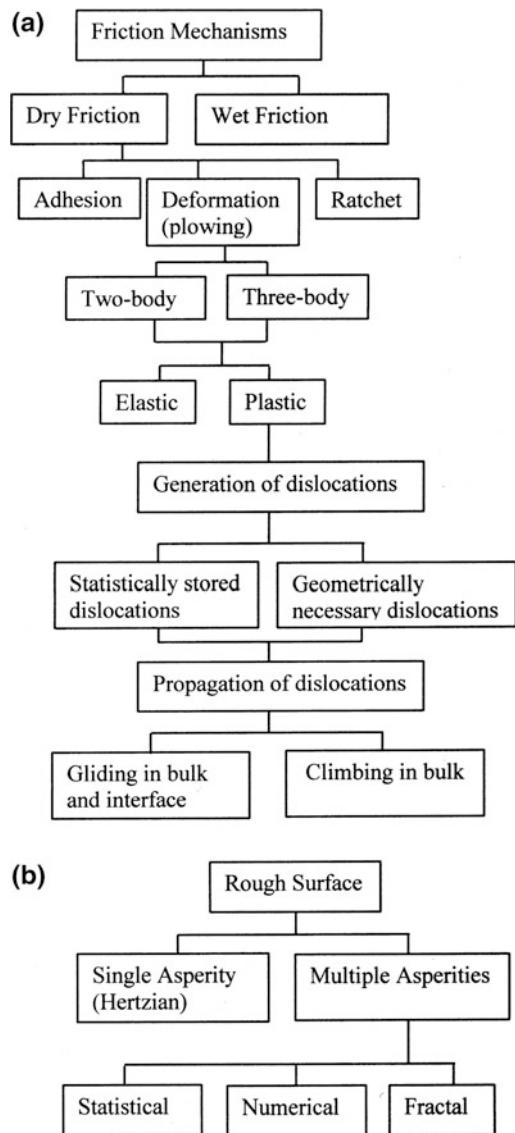
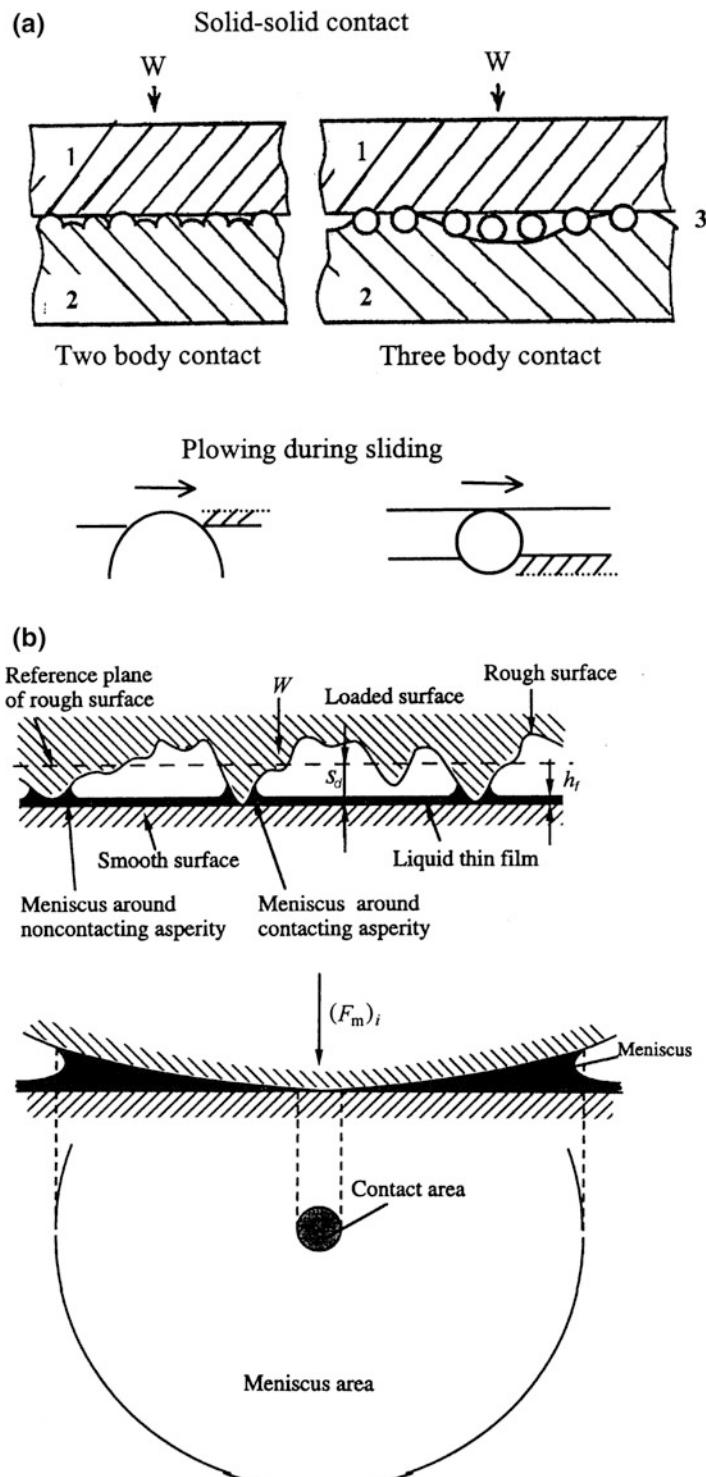


Fig. 11.1 **a** A block diagram showing friction mechanisms and generation and propagation of dislocations during sliding, **b** a block diagram of rough contact models

length scales. A strain gradient plasticity theory has been developed for microscale deformations by Fleck et al. (1994), Nix and Gao (1998), and Hutchinson (2000). Their theory predicts a dependence of mechanical properties on the strain gradient, which is scale dependent: the smaller is the size of the deformed region, the greater is the gradient of plastic strain, yield strength, and hardness.

Fig. 11.2 Schematics of a two-bodies and three-bodies during dry contact of rough surfaces, **b** formation of menisci during wet contact



A comprehensive model of scale effect in friction including adhesion, two- and three-body deformations, and the ratchet mechanism has recently been proposed by Bhushan and Nosonovsky (2003, 2004a, b) and Nosonovsky and Bhushan (2005). The model for adhesive friction during single and multiple asperity contact was developed by Bhushan and Nosonovsky (2003), and is based on the strain gradient plasticity and dislocation assisted sliding (gliding dislocations at the interface or microslip). The model for the two-body and three-body deformation was proposed by Bhushan and Nosonovsky (2004a), and for the ratchet mechanism by Nosonovsky and Bhushan (2005). The model has been extended for wet contacts, wear, and interface temperature by Bhushan and Nosonovsky (2004b). The detailed model is presented in this chapter.

The chapter is organized as follows. In the next section, scale effects in mechanical properties are considered, including yield strength and hardness, based on the strain gradient plasticity, and shear strength at the interface, based on the dislocation assisted sliding (microslip). In the third section, scale effects in surface roughness and contact parameters are considered, including the real area of contact, number of contacts, and mean size of contact. Load dependence of contact parameters is also studied in this section. In the fourth section, scale effects in friction are considered, including adhesion, two- and three-body deformation, ratchet mechanism, meniscus analysis, total value of the coefficient of friction, and comparison with the experimental data. In the fifth and sixth sections, scale effects in wear and interface temperature are analyzed, respectively.

11.2 Scale Effect in Mechanical Properties

In this section, scale-dependence of hardness and shear strength at the interface is considered. A strain gradient plasticity theory has been developed for microscale deformations, by Fleck et al. (1994), Nix and Gao (1998), Hutchinson

(2000), and others, which is based on statistically stored and geometrically necessary dislocations (to be described later). Their theory predicts dependence of mechanical properties on the strain gradient, which is scale dependent: the smaller the size of the deformed region, the greater is the gradient of plastic strain, yield strength, and hardness. Gao et al. (1999) and Huang et al. (2000) proposed a mechanism-based strain gradient (MSG) plasticity theory, which is based on a multiscale framework linking the microscale (10–100 nm) notion of statistically stored and geometrically necessary dislocations to the mesoscale (1–10 μm) notion of plastic strain and strain gradient. Bazant (2002) analyzed scale effects based on the MSG plasticity theory in the limit of small scale, and found that corresponding nominal stresses in geometrically similar structures of different sizes depend on the size according to a power exponent law.

It is believed also, that relative motion of two contacting bodies during sliding takes place due to dislocation-assisted sliding (microslip), which results in scale-dependent shear strength at the interface (Bhushan and Nosonovsky 2003). Scale effects in mechanical properties (yield strength, hardness, and shear strength at the interface) based on the strain gradient plasticity and dislocation-assisted sliding models are considered in this section.

11.2.1 Yield Strength and Hardness

Plastic deformation occurs during asperity contacts because a small real area of contact results in high contact stresses, which are often beyond the limits of the elasticity. As stated earlier, during loading, generation and propagation of dislocations is responsible for plastic deformation. Because dislocation motion is irreversible, plastic deformation provides a mechanism for energy dissipation during friction. The strain gradient plasticity theories (Fleck et al. 1994; Nix and Gao 1998; Hutchinson 2000) consider two types of dislocations: randomly created Statistically Stored Dislocation (SSD) and Geometrically Necessary Dislocations (GND). The GND

are required for strain compatibility reasons. Randomly created SSD during shear and GND during bending are presented in Fig. 11.3a. The density of the GND (total length of dislocation lines per volume) during bending is proportional to the curvature κ and to the strain gradient

$$\rho_G = \frac{\kappa}{b} = \frac{1}{v} \frac{\partial \varepsilon}{\partial z} \propto \nabla \varepsilon \quad (11.1)$$

where ε is strain and $\nabla \varepsilon$ is the strain gradient.

The GND during indentation, Fig. 11.3b, are located in a certain sub-surface volume (Bhushan and Nosonovsky 2003). The large strain gradients in small indentations require GND to account for the large slope at the indented surface. SSD, not shown here, also would be created and would contribute to deformation resistance, and are function of strain rather than strain gradient. According to Nix and Gao (1998), we assume that indentation is accommodated by circular loops of GND with Burgers vector normal to the plane of the surface. If we think of the individual dislocation loops being spaced equally along the surface of the indentation, then the surface slope

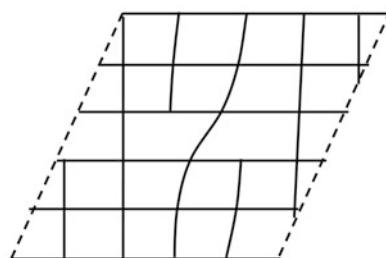
$$\tan \theta_i = \frac{h}{a} = \frac{b}{s} \quad (11.2)$$

where θ_i is the angle between the surface of the conical indenter and the plane of the surface, a is the contact radius, h is the indentation depth, b is the Burgers vector, and s is the spacing between individual slip steps on the indentation surface, Fig. 11.3b. Nix and Gao (1998) further reported that for geometrical (strain compatibility) considerations, the density of the GND is

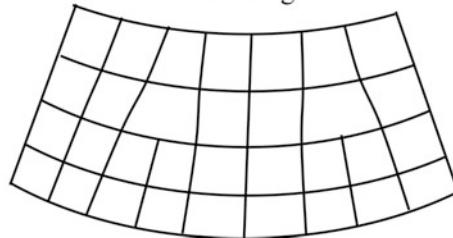
$$\rho_G = \frac{3}{2bh} \tan^2 \theta_i = \frac{3}{2b} \left(\frac{\tan \theta_i}{a} \right) = \frac{3}{2b} \nabla \varepsilon \quad (11.3)$$

Thus, ρ_G is proportional to strain gradient (scale dependent), whereas the SSD density ρ_S is dependent upon the average strain in the indentation, which is related to the slope of the

(a) Statistically stored dislocations during shear



Geometrically necessary dislocations during bending



(b)

Geometrically necessary dislocations during indentation

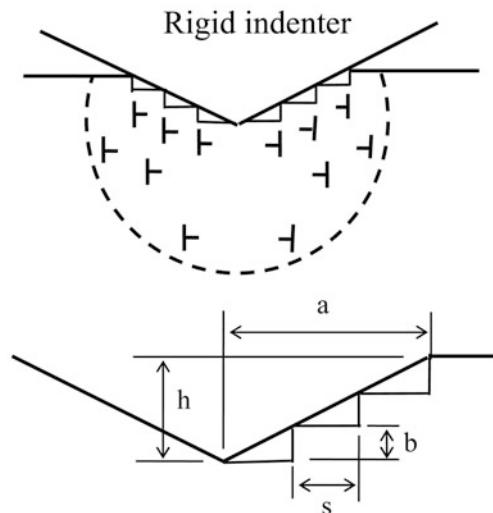


Fig. 11.3 a Illustration of statistically stored dislocations during shear and geometrically necessary dislocations during bending. b Geometrically necessary dislocations during indentation

indenter ($\tan \theta_i$). Based on experimental observations, ρ_s is approximately proportional to strain (Fleck et al. 1994).

According to the Taylor model of plasticity (Friedel 1964), dislocations are emitted from Frank-Read sources. Due to interaction with each other, the dislocations may become stuck in what is called the Taylor network. However, when externally applied stress reaches the order of Peierls stress for the dislocations, they start to move and plastic yield is initiated. The magnitude of the Peierls stress τ_p is proportional to the dislocation's Burgers vector b divided by a distance between dislocation lines s (Friedel 1964; Weertman and Weertman 1966)

$$\tau_p = Gb/(2\pi s) \quad (11.4)$$

where G is the elastic shear modulus. An approximate relation of the shear yield strength τ_Y to the dislocations density at a moment when yield is initiated is given by (Friedel 1964)

$$\tau_{Y0} = cGb/s = cGb\sqrt{\rho} \quad (11.5)$$

where c is a constant on the order of unity, specified by the crystal structure and ρ is the total length of dislocation lines per volume, which is a complicated function of strain ε and strain gradient ($\nabla \varepsilon$)

$$\rho = \rho_s(\varepsilon) + \rho_G(\nabla \varepsilon) \quad (11.6)$$

The shear yield strength τ_Y can be written now as a function of SSD and GND densities (Friedel 1964),

$$\tau_Y = cGb\sqrt{\rho_s + \rho} = \tau_{Y0}\sqrt{1 + (\rho_G/\rho_s)} \quad (11.7)$$

where

$$\tau_{Y0} = cGb\sqrt{\rho_s} \quad (11.8)$$

is the shear yield strength value in the limit of a small ρ_G/ρ_s ratio (large scale) that would arise from the SSD, in the absence of GND. Note that the ratio of the two densities is defined by the

problem geometry and is scale dependent. Based on the relationships for ρ_G (Eq. 11.3) and ρ_s , the ratio ρ_G/ρ_s is inversely proportional to a and Eq. 11.7 reduces to

$$\tau_Y = \tau_{Y0}\sqrt{1 + (l_d/a)} \quad (11.9)$$

where l_d is a plastic deformation length that characterizes depth dependence on shear yield strength. According to Hutchinson (2000), this length is physically related to an average distance a dislocation travels, which was experimentally determined to be between 0.2 and 5 μm for copper and nickel. Note that l_d is a function of the material and the asperity geometry and is dependent on SSD.

Using von Mises yield criterion, hardness $H = 3\sqrt{3}\tau_Y$. From Eq. 11.9 the hardness is also scale-dependent (Nix and Gao 1998)

$$H = H_0\sqrt{1 + (l_d/a)} \quad (11.10)$$

where H_0 is hardness in the absence of strain gradient. Equation 11.9 provides scale-dependence of the resistance force to deformation in a general case of plastic deformation (Bhushan and Nosonovsky 2003).

Scale-dependence of yield strength and hardness has been well-established experimentally. Bhushan and Koinkar (1994) and Bhushan et al. (1996) measured hardness of single-crystal silicon (100) up to a peak load of 500 μN . Kulkarni and Bhushan (1996) measured hardness of single crystal aluminum (100) up to 2000 μN and Nix and Gao (1998) presented data for single crystal copper using a three-sided pyramidal (Berkovich) diamond tip. Hardness on the nanoscale is found to be higher than on the microscale, Fig. 11.4. Similar results have been reported in other tests, including indentation tests for other materials (Gane and Cox 1970; Stelmashenko et al. 1993; McElhaney et al. 1998), torsion and tension experiments on copper wires (Fleck et al. 1994; Hutchinson 2000), and bending experiments on silicon and silica beams (Sundararajan and Bhushan 2002).

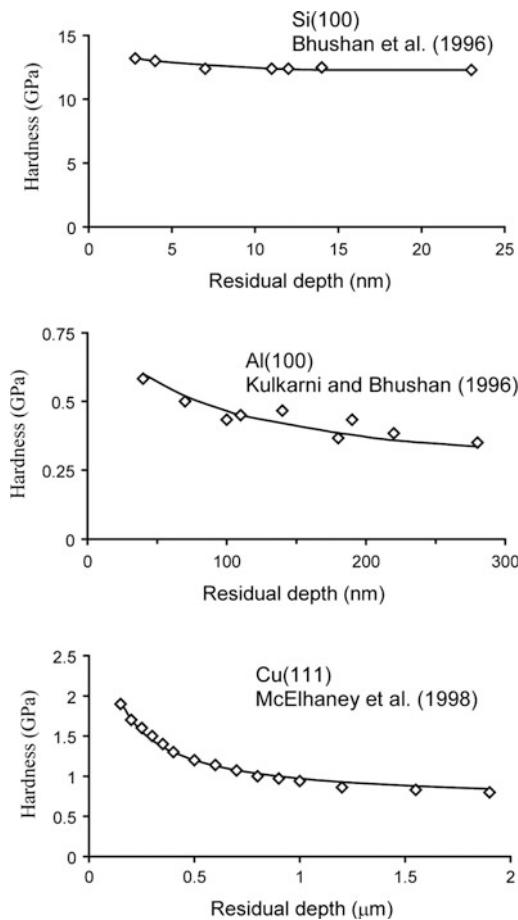


Fig. 11.4 Indentation hardness as a function of residual indentation depth for Si(100), 1996, Al(100), Cu(111)

11.2.2 Shear Strength at the Interface

Mechanism of slip involves motion of large number of dislocations, which is responsible for plastic deformation during sliding. Dislocations are generated and stored in the body and propagate under load. There are two modes of possible line (or edge) dislocation motion: gliding, when dislocation moves in the direction of its Burgers vector b by a unit step of its magnitude, and climbing, when dislocation moves in a direction perpendicular to its Burgers vector, Fig. 11.5a. Motion of dislocations can take place in the bulk of the body or at the interface. Due to periodicity of the lattice, a gliding dislocation experiences a periodic force known as the Peierls force

(Weertman and Weertman 1966). The Peierls force is responsible for keeping the dislocation at a central position between symmetric lattice lines, and it opposes dislocation's gliding, Fig. 11.5b (Bhushan and Nosonovsky 2003). Therefore, an external force should be applied to overcome Peierls force resistance against the dislocation's motion. Weertman (1963) showed that a dislocation or a group of dislocations can glide uniformly along an interface between two bodies of different elastic properties. In the continuum elasticity formulation, this motion is equivalent to a propagating interface slip pulse. However, the physical nature of this deformation is plastic because dislocation motion is irreversible. Local plastic deformation can occur at the interface due to the concentration of dislocations even in predominantly elastic contacts. Gliding of a dislocation along the interface results in a relative displacement of the bodies for a distance equal to the Burgers vector of the dislocation, whereas a propagating set of dislocations effectively results in dislocation-assisted sliding, or microslip, Fig. 11.6.

Several types of microslip are known in the tribology literature (Bhushan 2013). Dislocation-assisted sliding is one type of microslip, which propagates along the interface. The conventional mechanism of sliding is considered to be concurrent slip with simultaneous breaking of all adhesive bonds. Based on Johnson (1997) and Bhushan and Nosonovsky (2003), for contact sizes on the order of few nm to few μm , dislocation-assisted sliding is more energetically profitable than a concurrent slip. Their argument is based on the fact that experimental measurements with the SFA demonstrated that frictional stress for mica is of the same order as Peierls stress, which is required for gliding of dislocations.

Polonsky and Keer (1996) considered the pre-existing dislocation sources and carried out a numerical microcontact simulation based on contact plastic deformation representation in terms of discrete dislocations. They found that when the asperity size decreases and becomes comparable with the characteristic length of materials microstructure (distance between

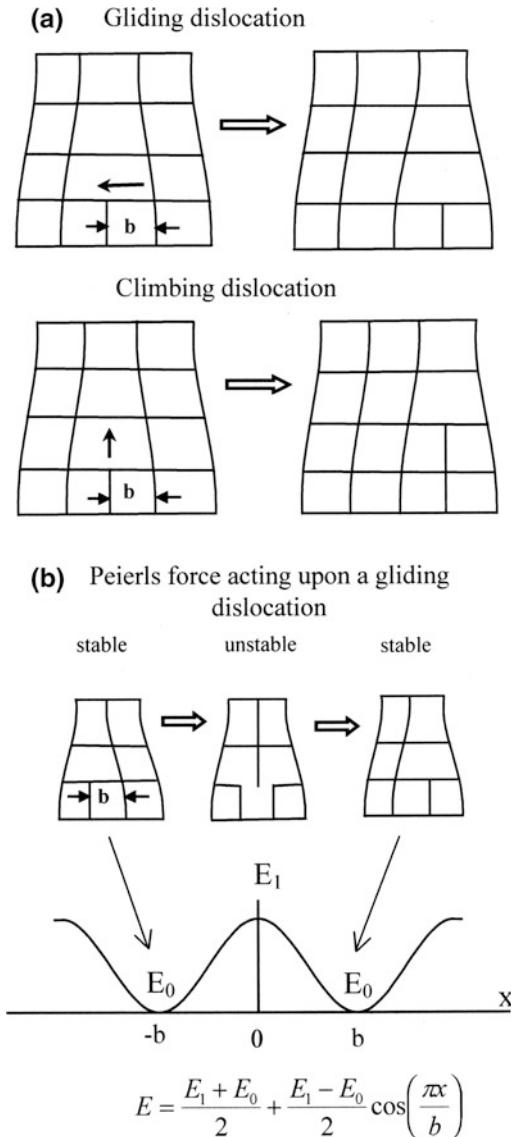


Fig. 11.5 **a** Schematics of gliding and climbing dislocations motion by a unit step of Burgers vector b . **b** Origin of the periodic force acting upon a gliding dislocation (Peierls Force). Gliding dislocation passes locations of high and low potential energy

dislocation sources), resistance to plastic deformation increases. This supports conclusions drawn from strain gradient plasticity. Deshpande et al. (2003) conducted discrete plasticity modeling of cracks in single crystals and considered dislocation nucleation from Frank-Read sources distributed randomly in the material. Pre-existing

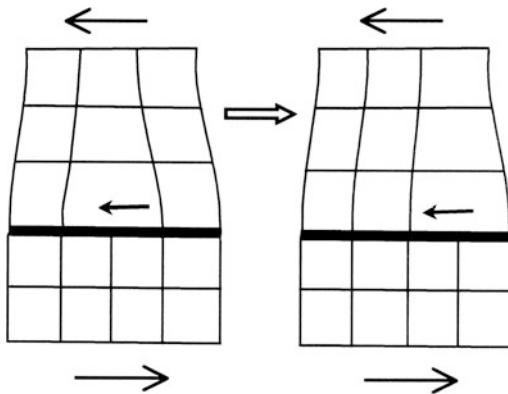


Fig. 11.6 Schematic showing microslip due to gliding dislocations at the interface (Bhushan and Nosonovsky 2003)

sources of dislocations, considered by all of these authors, are believed to be a more realistic reason for an increase in the number of dislocations during loading, rather than newly nucleated dislocations (Friedel 1964). In general, dislocations are emitted under loads from pre-existing sources and propagate along slip lines, shown in Fig. 11.7. In regions of higher loads, number of emitted dislocations is higher (Bhushan and Nosonovsky 2003). The approach of Deshpande et al. (2003) was limited to numerical analysis of special cases.

Bhushan and Nosonovsky (2003) considered a sliding contact between two bodies. Slip along the contact interface is an important special case of plastic deformation. The local dislocation-assisted microslip can exist even if the contact is predominantly elastic due to dislocations concentrating at the interface. Due to these dislocations, the stress at which yield occurs at the interface is lower than shear yield strength in the bulk. This means that average shear strength at the interface is lower than in the bulk.

An assumption that all dislocations produced by externally applied forces are distributed randomly throughout the volume would result in a vanishing small probability for a dislocation to be exactly at the interface. However, many traveling (gliding and climbing) dislocations will be stuck at the interface as soon as they reach it. As a result of this, a certain number of

dislocations will be located at the interface. In order to account for a finite dislocation density at the interface, Bhushan and Nosonovsky (2003) assumed, that the interface zone has a finite thickness D . Dislocations within the interface zone may reach the contact surface due to climbing and contribute to the microslip. In the case of a small contact radius a compared to interface zone thickness D , which is scale-dependent, D is approximately equal to a . However, in the case of a large contact radius, the interface zone thickness is approximately equal to the average distance dislocations can climb l_s , Fig. 11.8 (Bhushan and Nosonovsky 2003). The depth of the subsurface volume, from which dislocations have a high chance to reach the interface is limited by l_s and a , respectively, for these two cases. Based on these geometrical considerations, an approximate relation can be written as

$$D = \frac{al_s}{l_s + a} \quad (11.11)$$

The interface density of dislocations (total length of dislocation lines per interface area) is related to the volume density as

$$\eta_{\text{int}} = \rho D = \rho \left(\frac{al_s}{l_s + a} \right) \quad (11.12)$$

During sliding, dislocations must be generated at the interface with a certain critical density $\eta_{\text{int}} = \eta_{\text{cr}}$. The corresponding shear strength during sliding can be written following Eq. 11.9 as

$$\tau_a = \tau_{a0} \sqrt{1 + (l_s + a)} \quad (11.13)$$

where

$$\tau_{a0} = cGb \sqrt{\frac{\eta_{\text{cr}}}{l_s}} \quad (11.14)$$

is the shear strength during sliding in the limit of $a \gg l_s$.

Equation 11.13 gives scale-dependence of the shear strength at the interface and is based on the following assumptions. First, it is assumed that

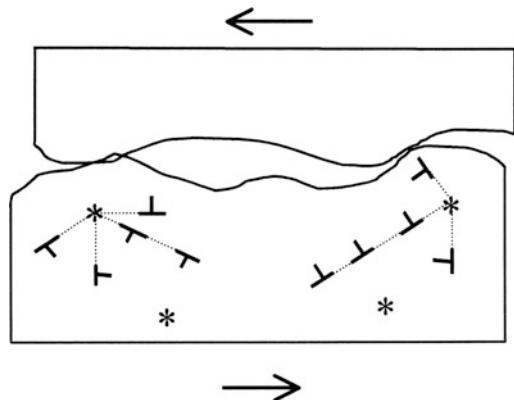


Fig. 11.7 Generation of dislocations from sources (asterisk) during plowing due to plastic deformation

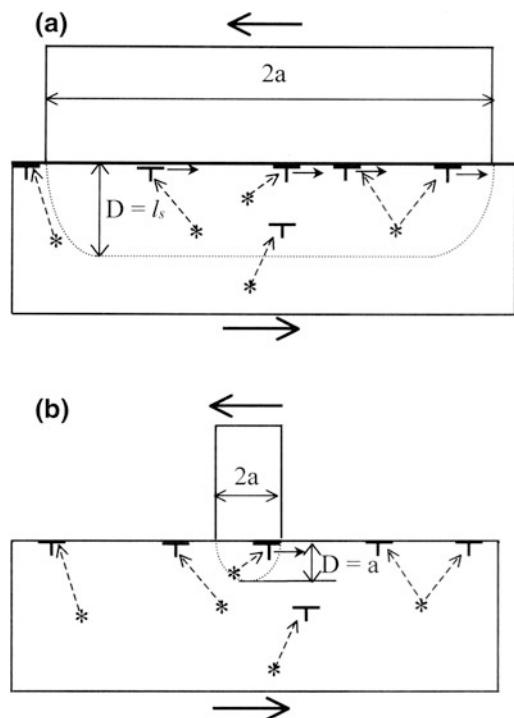


Fig. 11.8 Gliding dislocations at the interface generated from sources (asterisk). Only dislocations generated within the interface zone can reach the interface. **a** For a large contact radius a , thickness of this zone D is approximately equal to an average distance dislocations climb l_s . **b** For small contact radius a , the thickness of the interface zone is approximately equal to a

only dislocations in the interface zone of thickness D , given by Eq. 11.11, contribute into sliding. Second, it is assumed, that a critical

density of dislocations at the interface η_{cr} is required for sliding. Third, the shear strength is equal to the Peierls stress, which is related to the volume density of the dislocations $\rho = \eta/D$ according to Eq. 11.4, with the typical distance between dislocations $s = 1/\sqrt{\rho}$. It is noted that the proposed scaling rule for the dislocation assisted sliding mechanism (Eq. 11.13) has a similar form to that for the yield strength (Eq. 11.9), since both results are consequences of scale-dependent generation and propagation of dislocations under load (Bhushan and Nosonovsky 2003).

11.3 Scale Effect in Surface Roughness and Contact Parameters

During multiple-asperity contact, scale-dependence of surface roughness is a factor that contributes to scale dependence of the real area of contact. Roughness parameters are known to be scale dependent (Bhushan 2013), which, during the contact of two bodies, results in scale-dependence of the real area of contact, number of contacts, and mean contact size. The contact parameters also depend on the normal load, and the load-dependence is similar to the scale-dependence (Nosonovsky and Bhushan 2005). Both effects are analyzed in this section.

11.3.1 Scale-Dependence of Roughness and Contact Parameters

A random rough surface with Gaussian height distribution is characterized by the standard deviation of surface height σ and the correlation length β^* (Bhushan 2013). The correlation length is a measure of how quickly a random event decays and it is equal to the length, over which the autocorrelation function drops to a small fraction of the value at the origin. The correlation length can be considered as a distance, at which two points on a surface have just reached the condition where they can be regarded

as being statistically independent. Thus, σ is a measure of height distribution and β^* is a measure of spatial distribution.

A surface is composed of a large number of length scales of roughness that are superimposed on each other. According to AFM measurements on a glass-ceramic disk surface, both σ and β^* initially increase with the scan size and then approach a constant value at certain scan size, Fig. 11.9. This result suggests that disk roughness has a long wavelength limit, L_{lwl} , which is equal to the scan size at which the roughness values approach a constant value (Bhushan 2013). It can be assumed that σ and β^* depend on the scan size according to an empirical power rule

$$\begin{aligned}\sigma &= \sigma_0 \left(\frac{L}{L_{lwl}} \right)^n & L < L_{lwl} \\ \beta^* &= \beta_0^* \left(\frac{L}{L_{lwl}} \right)^m & L < L_{lwl}\end{aligned}\quad (11.15)$$

where n and m are indices of corresponding exponents and σ_0 and β_0^* are macroscale values (Bhushan and Nosonovsky 2003). Based on the data, presented in Fig. 11.9, it is noted that for a glass-ceramic disk, long-wavelength limit for σ and β^* is about 17 and 23 μm , respectively. The difference is expected to be due to measurement errors. An average value $L_{lwl} = 20 \mu\text{m}$ is taken here for calculations. The values of the indices are found as $m = 0.5$, $n = 0.2$, and the macro-scale values are $\sigma_0 = 5.3 \text{ nm}$, $\beta_0^* = 0.37 \mu\text{m}$ (Nosonovsky and Bhushan 2005).

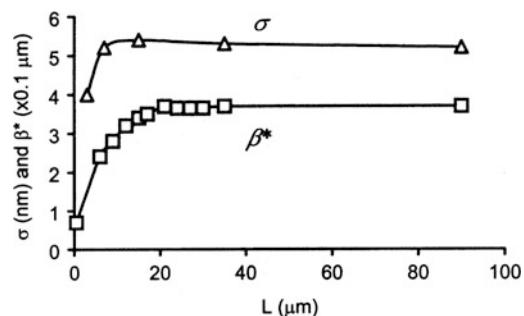


Fig. 11.9 Roughness parameters as a function of scan size for a glass-ceramic disk measured using AFM (Bhushan 2013)

For two random surfaces in contact, the length of the nominal contact size L defines the characteristic length scale of the problem. The contact problem can be simplified by considering a rough surface with composite roughness parameters in contact with a flat surface. The composite roughness parameters σ and β^* can be obtained based on individual values for the two surfaces (Bhushan 2013). For Gaussian surfaces, the contact parameters of interest, to be discussed later, are the real area of contact A_r , number of contacts N , and mean contact radius \bar{a} . The long wavelength limit for scale-dependence of the contact parameters L_{lc} , which is not necessarily equal to that of the roughness, L_{lwl} , will be used for normalization of length parameters. The scale-dependence of the contact parameters exists if $L < L_{lc}$ (Nosonovsky and Bhushan 2005).

The mean of surface height distribution corresponds to the so-called reference plane of the surface. Separation s_d is a distance between reference planes of two surfaces in contact, normalized by σ . For a given s_d and statistical distribution of surface heights, the total real area of contact (A_r), number of contacts (N), and elastic normal load W_e can be found, using statistical analysis of contacts. The real area of contact, number of contacts and elastic normal load are related to the separation distance s_d (Onions and Archard 1973)

$$\begin{aligned} A_r &\propto F_A(s_d) \\ N &\propto \frac{1}{(\beta^*)^2} F_N(s_d) \\ W_e &\propto \frac{E^* \sigma}{\beta^*} F_W(s_d) \end{aligned} \quad (11.16)$$

where $F_A(s_d)$, $F_N(s_d)$, and $F_W(s_d)$, are integral functions defined by Onions and Archard (1973). It should be noted, that A_r and N as a function of d are prescribed by the contact geometry (σ , β^*) and do not depend on whether the contact is elastic or plastic. Based on Onions and Archard data, it is observed that the ratio F_W/F_A is almost constant for moderate $d < 1.4$ and increases slightly for $d > 1.4$. The ratio F_A/F_N decreases rapidly with d and becomes almost constant for $d > 2.0$. For moderate loads, the contact is expected to occur on the upper parts of the asperities ($s_d > 2.0$), and a linear proportionality

of $F_A(s_d)$, $F_N(s_d)$, and $F_W(s_d)$ can be assumed (Bhushan and Nosonovsky 2003).

Based on Eq. 11.16 and the observation that F_W/F_A is almost constant, for moderate loads, A_{re} (the real area of elastic contact), N , and \bar{a} are related to the roughness, based on the parameter L_{lc} , as

$$A_{re} \propto \frac{\beta^*}{\sigma E^*} W = A_{re0} \left(\frac{L}{L_{lc}} \right)^{m-n}, \quad L < L_{lc} \quad (11.17)$$

$$N \propto \frac{W}{\beta^* \sigma E^*} = N_0 \left(\frac{L}{L_{lc}} \right)^{-m-n}, \quad L < L_{lc} \quad (11.18)$$

$$\bar{a} \propto \beta^* = \sqrt{\frac{A_r}{N}} = \bar{a}_0 \left(\frac{L}{L_{lc}} \right)^m, \quad L < L_{lc} \quad (11.19)$$

According to Whitehouse and Archard (1970), the mean radius of summit tips \bar{R}_p is given

$$\bar{R}_p \propto \frac{(\beta^*)^2}{\sigma} = \bar{R}_{p0} \left(\frac{L}{L_{lwl}} \right)^{2m-n}, \quad L < L_{lwl} \quad (11.20)$$

where \bar{a}_0 , N_0 , and \bar{R}_{p0} are macroscale values. E^* is the effective elastic modulus of contacting bodies (Bhushan and Nosonovsky 2004b), which is related to the elastic moduli E_1 , E_2 , and Poisson's ratios v_1 , v_2 , of the two bodies as $1/E^* = (1 - v_1^2)/E_1 + (1 - v_2^2)/E_2$ and which is known to be scale independent, and variables with the subscript "0" are corresponding macroscale values (for $L \geq L_{lc}$).

The dependence of the real area of plastic contact A_{rp} on the load is given by

$$A_{rp} = \frac{W}{H} \quad (11.21)$$

where H is hardness. According to the strain gradient plasticity model (Fleck et al. 1994; Nix and Gao 1998), the yield strength τ_Y is given by Eq. 11.9 and hardness H is given by Eq. 11.10. In the case of plastic contact, the mean contact radius can be determined from Eq. 11.19, which is based on the contact geometry and is independent of load (Bhushan and Nosonovsky 2003). Assuming the contact radius as its mean value from Eq. 11.19 based on elastic analysis,

and combining Eqs. 11.10, 11.19 and 11.21, the real area of plastic contact is given as

$$A_{rp} = \frac{W}{H_0 \sqrt{1 + (l_d/\bar{a})}} = \frac{W}{H_0 \sqrt{1 + (L_d/L)^m}}, \quad L < L_{lc}$$

$$(11.22)$$

where L_d is a characteristic length parameter related to l_d , \bar{a} , and L_{lc} (Bhushan and Nosonovsky 2003)

$$L_d = L_{lc} \left(\frac{l_d}{\bar{a}_0} \right)^{1/m} \quad (11.23)$$

The scale-dependence of A_{re} , N , and \bar{a} is presented in Fig. 11.10 (Nosonovsky and Bhushan 2005).

11.3.2 Dependence of Contact Parameters on Load

The effect of short and long wavelength details of rough surfaces on contact parameters also depends on the normal load. For low loads, the ratio of real to apparent areas of contact A_r/A_a is small, contact spots are small, and long wavelength details are irrelevant. For higher A_r/A_a , long wavelength details become important, whereas small wavelength details of the surface geometry become irrelevant. The effect of increased load is similar to the effect of increased scale length (Nosonovsky and Bhushan 2005).

In the preceding subsections, it was assumed that the roughness parameters are scale-dependent for $L < L_{lw}$, whereas the contact parameters are scale-dependent for $L < L_{lc}$. The upper limit of scale-dependence for the contact parameters, L_{lc} , depends on the normal load, and it is reasonable to assume that L_{lc} is a function of A_r/A_a , and the contact parameters are scale-dependent when A_r/A_a is below a certain critical value. It is convenient to consider the apparent pressure p_a , which is equal to the normal load divided by the apparent area of contact (Nosonovsky and Bhushan 2005).

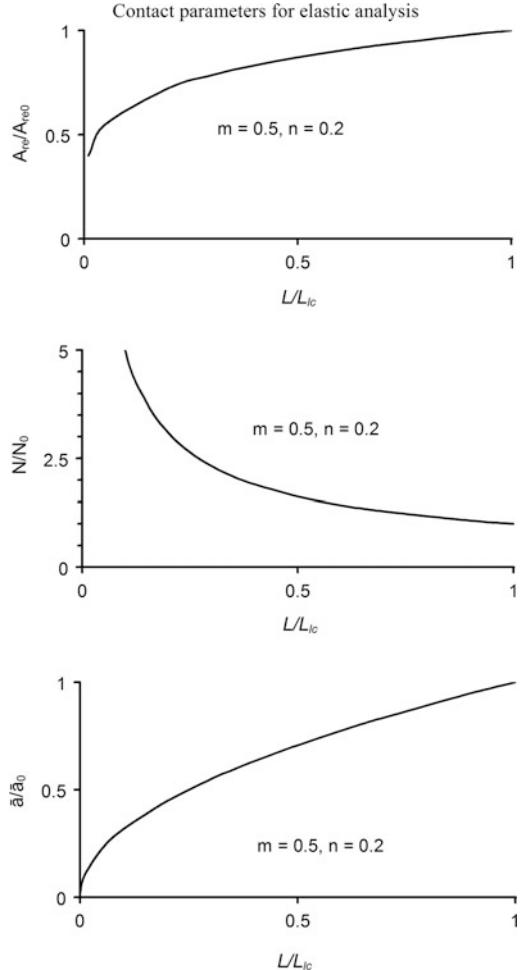


Fig. 11.10 Scale length dependence of normalized contact parameters ($m = 0.5, n = 0.2$) **a** real area of contact, **b** number of contacts, and **c** mean contact radius (Nosonovsky and Bhushan 2005)

For elastic contact, based on Eqs. 11.15 and 11.17, this condition can be written as

$$\frac{A_{re}}{A_a} \propto \frac{\beta^* p_a}{\sigma} = p_a \frac{\beta_0^*}{\sigma_0} \left(\frac{L}{L_{lw}} \right)^{m-n} < p_{ac} \quad (11.24)$$

where p_{ac} is a critical apparent pressure, below which the scale-dependence occurs (Nosonovsky and Bhushan 2005). From Eq. 11.24 one can find

$$L < L_{lw} \left(\frac{\beta_0^* p_a}{\sigma_0 p_{ac}} \right)^{1/(n-m)} \quad (11.25)$$

The right-hand expression in Eq. 11.24 is defined as L_{lc}

$$L_{lc} = L_{lwl} \left(\frac{\beta_0^* p_a}{\sigma_0 p_{ac}} \right)^{1/(n-m)} \quad (11.26)$$

For plastic contact, based on Eq. 11.22

$$\frac{A_{rp}}{A_a} \propto \frac{p_a}{\sqrt{1 + (L_d/L)^m}} < p_{ac} \quad (11.27)$$

In a similar manner to the elastic case, Eq. 11.27 yields (Nosonovsky and Bhushan 2005)

$$L_{lc} = L_d \left[\left(\frac{p_a}{p_{ac}} \right)^2 - 1 \right]^{-1/m} \quad (11.28)$$

Load-dependence of the long wavelength limit for contact parameters, L_{lc} is presented in Fig. 11.11 for an elastic contact based on Eq. 11.28, and for a plastic contact based on Eq. 11.28, for $m = 0.5$, $n = 0.2$ (Nosonovsky and Bhushan 2005). The load (apparent pressure) is normalized by $\beta_0^*/(p_{ac} \sigma_0)$ for the elastic contact and by p_{ac} for the plastic contact. In the case of elastic contact, it is observed, that the long wavelength limit decreases with increasing load. For a case characterized by a given scale length L , an increase of load will result in decrease of L_{lc} and, eventually, the condition $L < L_{lc}$ will be violated. Thus, the contact parameters, including the coefficient of friction, will reach the macro-scale values. A decrease of L_{lc} with increasing load is also observed in the case of plastic contact; the data presented for $p_a/p_{ac} > 1$.

11.4 Scale Effects in Friction

According to the adhesion and deformation model of friction (Bhushan 2013), the coefficient of dry friction μ can be presented as a sum of the adhesion component μ_a and deformation (plowing) component μ_d . The latter, in the presence of particles, is a sum of the asperity summits

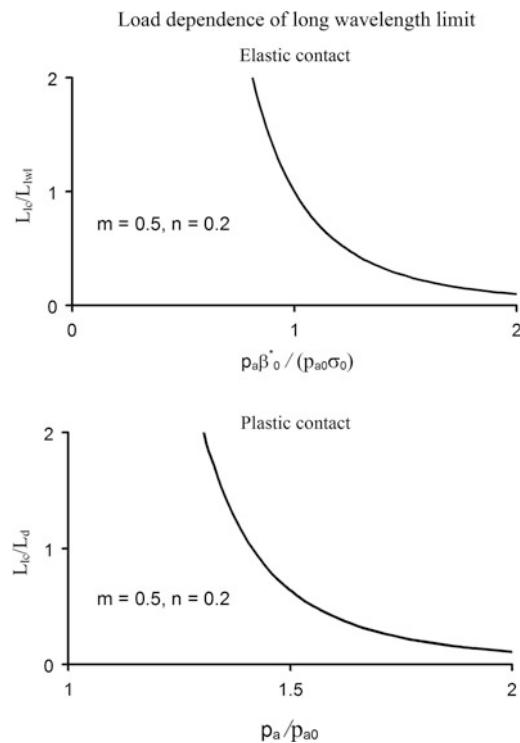


Fig. 11.11 Dependence of the normalized long wavelength limit for contact parameters on load (normalized apparent pressure) for elastic and plastic contacts, for the empirical rule ($m = 0.5$, $n = 0.2$) (Nosonovsky and Bhushan 2005)

deformation component μ_{ds} and particles deformation component μ_{dp} , so that the total coefficient of friction is (Bhushan and Nosonovsky 2004a)

$$\begin{aligned} \mu &= \mu_a + \mu_{ds} + \mu_{dp} = \frac{F_a + F_{ds} + F_{dp}}{W} \\ &= \frac{A_{ra}\tau_a + A_{ds}\tau_{ds} + A_{dp}\tau_{dp}}{W} \end{aligned} \quad (11.29)$$

where W is the normal load, F is the friction force, A_{ra} , A_{ds} , A_{dp} are the real areas of contact during adhesion, two body deformation, and with particles, respectively, and τ is the shear strength. The subscripts a , ds , and dp correspond to adhesion, summit deformation, and particle deformation.

In the presence of a meniscus, the friction force is given by

$$F = \mu(W + F_m) \quad (11.30)$$

where F_m is the meniscus force (Bhushan 2013). The coefficient of friction in the presence of the meniscus force, μ_{wet} , is calculated using only the applied normal load, as normally measured in the experiments (Bhushan and Nosonovsky 2004b)

$$\begin{aligned} \mu_{wet} &= \mu \left(1 + \frac{F_m}{W} \right) \\ &= \frac{A_{ra} + \tau_a + A_{ds}\tau_{ds} + A_{dp}\tau_{dp}}{W} \left(1 + \frac{F_m}{W} \right) \end{aligned} \quad (11.31)$$

Equation 11.31 shows that μ_{wet} is greater than μ , because F_m is not taken into account for calculation of the normal load in the wet contact.

It was shown by Greenwood and Williamson (1966), and by subsequent modifications of their model, that for contacting surfaces with common statistical distributions of asperity heights, the real area of contact is almost linearly proportional to the normal load. This linear-dependence, along with Eq. 11.29, results in linear-dependence of the friction force on the normal load, or coefficient of friction being independent of the normal load. For a review of the numerical analysis of rough surface contacts, see Bhushan (1996, 1998) and Bhushan and Peng (2002). The statistical and numerical theories of contact involve roughness parameters—e.g. the standard deviation of asperity heights and the correlation length (Bhushan 2013). The roughness parameters are scale dependent. In contrast to this, the theory of self-similar (fractal) surfaces solid contact developed by Majumdar and Bhushan (1991) does not include length parameters and is scale-invariant in principle. The shear strength of the contacts in Eq. 11.29 is also scale-dependent. In addition to the adhesional contribution to friction, elastic and plastic deformation on the nano- to macroscale contributes to friction (Bhushan 2013). The deformations are also scale-dependent.

11.4.1 Adhesional Friction

The adhesion component of friction depends on the real area of contact and adhesion shear strength. Here we derive expressions for scale dependence of adhesional friction during single-asperity and multiple-asperity contacts.

11.4.1.1 Single-Asperity Contact

The length scale during single-asperity contact is the nominal contact length, which is equal to the contact diameter $2a$. In the case of predominantly elastic contacts, the real area of contact A_{re} depends on the load according to the Hertz analysis (Johnson 1985)

$$A_{re} = \pi a^2 \quad (11.32)$$

and

$$a = \left(\frac{3WR}{4E^*} \right)^{1/3} \quad (11.33)$$

where R is the effective radius of curvature of summit tips, and E^* is the effective elastic modulus of the two bodies. In the case of predominantly plastic contact, the real area of contact A_{rp} is given by Eq. 11.21, whereas the hardness is given by Eq. 11.10.

Combining Eqs. 11.10, 11.13, 11.29, and 11.32, the adhesional component of the coefficient of friction can be determined for the predominantly elastic contact as

$$\mu_{ae} = \mu_{ae0} \sqrt{1 + (l_s/a)} \quad (11.34)$$

and for the predominantly plastic contact as

$$\mu_{ap} = \mu_{ap0} \sqrt{\frac{1 + (l_s/a)}{1 + (l_d/a)}} \quad (11.35)$$

where μ_{ae0} and μ_{ap0} are corresponding values at the macroscale (Bhushan and Nosonovsky 2003).

The scale-dependence of adhesional friction in single-asperity contact is presented in Fig. 11.12a (Bhushan and Nosonovsky 2003). In

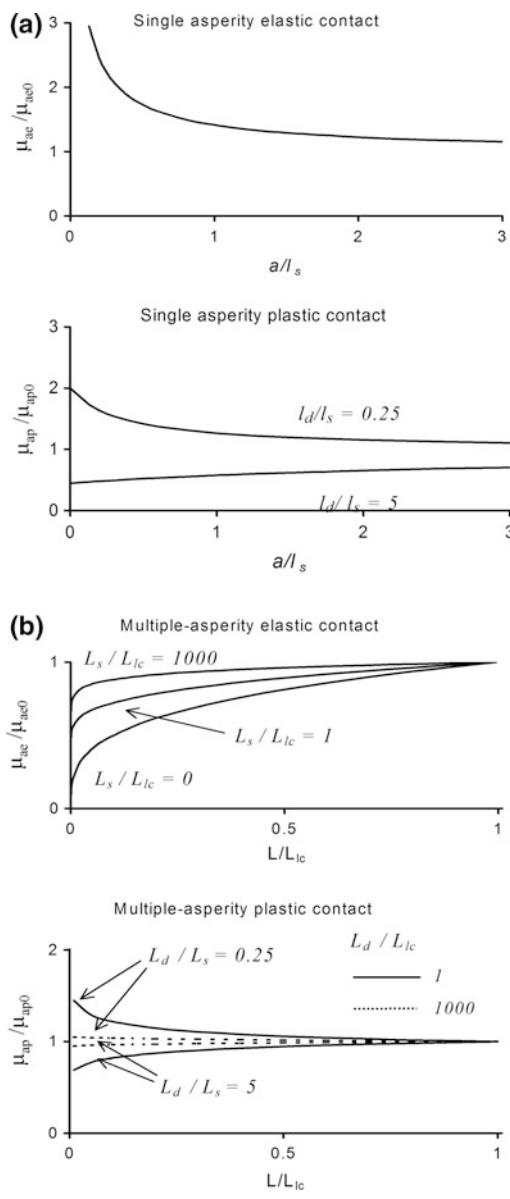


Fig. 11.12 Normalized results for the adhesional component coefficient of friction, as a function of scale [a/l_s for single asperity contact (a) and L/L_{lc} for multi-asperity contact (b)]. In the case of single asperity plastic contact, data are presented for two values of l_d/l_s . In the case of multi-asperity contact, data are presented for $m = 0.5$, $n = 0.2$. For multi-asperity elastic contact, data are presented for three values of L_s/L_{lc} . For multi-asperity plastic contact, data are presented for two values of L_d/L_s (Bhushan and Nosonovsky 2003)

the case of single asperity elastic contact, the coefficient of friction increases with decreasing scale (contact diameter) because of an increase in

the adhesion strength, according to Eq. 11.34. In the case of single asperity plastic contact, the coefficient of friction can increase or decrease with decreasing scale because of an increased hardness or increase in adhesional strength. The competition of these two factors is governed by l_d/l_s , according to Eq. 11.35. There is no direct way to measure l_d and l_s . We will see later from experimental data that the coefficient of friction tends to decrease with decreasing scale. Therefore, it must be assumed that $l_d/l_s > 1$ for the data reported in the paper (Bhushan and Nosonovsky 2003).

11.4.1.2 Multiple-Asperity Contact

The adhesion component of friction depends on the real area of contact and adhesion shear strength. Scale-dependence of the real area of contact was considered in the preceding section. Here we derive expressions for scale-dependence of the shear strength at the interface during adhesional friction. It is suggested by Bhushan and Nosonovsky (2003) that for many materials, dislocation-assisted sliding (microslip) is the main mechanism, which is responsible for the shear strength. They considered dislocation assisted sliding based on the assumption that contributing dislocations are located in a subsurface volume. The thickness of this volume is limited by the distance that dislocations can climb l_s (material parameter) and by the radius of contact a . They showed that τ_a is scale-dependent according to Eq. 11.13. Assuming the contact radii are equal to the mean values given by Eq. 11.19

$$\tau_a = \tau_{a0} \sqrt{1 + (L_s/L)^m}, \quad L < L_{lc} \quad (11.36)$$

where

$$L_s = L_{lc} \left(\frac{l_s}{a_0} \right)^{1/m} \quad (11.37)$$

In the case of absence of microslip (e.g., for an amorphous material), it should be assumed in Eqs. 11.34–11.36, $L_s = l_s = 0$.

Based on Eqs. 11.9, 11.17, 11.24, 11.29, 11.36, and 11.37, the adhesion component of the coefficient of friction in the case of elastic contact

Table 11.1 Scaling factors for the coefficient of adhesional friction (Bhushan and Nosonovsky 2003)

Single asperity elastic contact	Single asperity plastic contact	Multiple-asperity elastic contact	Multiple-asperity plastic contact
$\mu_e = \mu_{e0} \sqrt{1 + (l_s/a)}$	$\mu_e = \mu_{e0} \sqrt{\frac{1 + (l_s/a)}{1 + (l_d/a)}}$	$\mu_e = \mu_{e0} C_E L^{m-n} \sqrt{1 + (L_E/L)^m}$	$\mu_p = \mu_{p0} C_P \sqrt{\frac{1 + (L_s/L)^m}{1 + (L_d/L)^m}}$

μ_{ae} and in the case of plastic contact μ_{ap} , is given as (Bhushan and Nosonovsky 2003)

$$\begin{aligned} \mu_{ae} &= \frac{\tau_a A_{re}}{W} = \frac{\tau_{a0} A_{re0}}{W} \left(\frac{L}{L_{lc}} \right)^{m-n} \sqrt{1 + (L_s/L)^m} \\ &= \frac{\mu_{ae0}}{\sqrt{1 + (l_s/\bar{a}_0)}} \left(\frac{L}{L_{lc}} \right)^{m-n} \sqrt{1 + (L_s/L)^m} \\ &L < L_{lc} \end{aligned} \quad (11.38)$$

$$\begin{aligned} \mu_{ap} &= \frac{\tau_{a0}}{H_0} \sqrt{\frac{1 + (L_s/L)^m}{1 + (L_d/L)^m}} \\ &= \mu_{ap0} \sqrt{\frac{1 + (l_d/\bar{a}_0)}{1 + (l_s/\bar{a}_0)}} \sqrt{\frac{1 + (L_s/L)^m}{1 + (L_d/L)^m}} \\ &L < L_{lc} \end{aligned} \quad (11.39)$$

where μ_{ae0} and μ_{ap0} are values of the coefficient of friction at macroscale ($L \geq L_{lc}$).

The scale-dependence of adhesional friction in multiple-asperity elastic contact for various values of L_s/L_{lc} is presented in Fig. 11.12b, which is based on Eq. 11.38 (Bhushan and Nosonovsky 2003). The change of scale length L affects the coefficient of friction in two different ways: through the change of A_{re} (Eq. 11.17) and τ_a (Eq. 11.36) below L_{lc} . Further, τ_a is controlled by the ratio L_s/L . Based on Eq. 11.36, for a small ratio of L_s/L_{lc} , scale effects on τ_a are insignificant for $L/L_{lc} > 0$. As it is seen from Fig. 11.12b, by comparison of the curve with $L_s/L_{lc} = 0$ (insignificant scale effects on τ_a), $L_s/L_{lc} = 1$ and $L_s/L_{lc} = 1000$ (significant scale effects on τ_a), the results for the normalized coefficient of friction are close. Thus, the main contribution to the scaling effect is due to change of A_{re} . In the case of multiple-asperity plastic contact, the results based on Eq. 11.39 are presented in Fig. 11.12b for $L_d/L_s = 0.25$ and $L_d/L_s = 5$, and $L_d/L_{lc} = 1$ and $L_d/L_{lc} = 1000$. The change of scale affects the coefficient of friction through the change of A_{rp} (Eq. 11.34), which is controlled by L_d , and τ_a

(Eq. 11.36), which is controlled by L_s . It can be observed from Fig. 11.12b that for $L_d > L_s$, the change of A_{rp} prevails over the change of τ_a with decreasing scale, and the coefficient of friction decreases. For $L_d < L_s$, the change of τ_a prevails with decreasing scale, and the coefficient of friction increases (Bhushan and Nosonovsky 2003). Expressions for the coefficient of adhesional friction are presented in Table 11.1.

11.4.2 Two-Body Deformation

Based on the assumption that multiple asperities of two rough surfaces in contact have a conical shape, the two-body deformation component of friction can be determined as

$$\mu_{ds} = \frac{2 \tan \theta_r}{\pi} \quad (11.40)$$

where θ_r is the roughness angle (or attack angle) of a conical asperity (Rabinowicz 1995; Bhushan 2013). Mechanical properties affect the real area of contact and shear strength, and these cancel out in Eq. 11.29.

The roughness angle is scale-dependent and is related to the roughness parameters (Whitehouse and Archard 1970). Based on statistical analysis of a random Gaussian surface,

$$\tan \theta_r \propto \frac{\sigma}{\beta^*} \quad (11.41)$$

From Eq. 11.40 it can be interpreted that stretching the rough surface in the vertical direction (increasing vertical scale parameter σ) increases $\tan \theta_r$, and stretching in the horizontal direction (increasing vertical scale parameter β^*) decreases $\tan \theta_r$.

Using Eqs. 11.40 and 11.41, the scale dependence of the two-body deformation

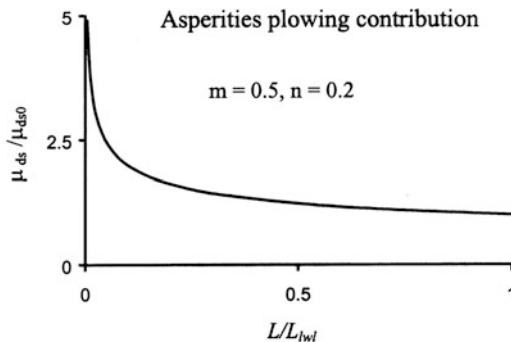


Fig. 11.13 Normalized results for the two-body deformation component of the coefficient of friction (Bhushan and Nosonovsky 2004a)

component of the coefficient of friction is given as (Bhushan and Nosonovsky 2004a)

$$\mu_{ds} = \frac{2\sigma_0}{\pi\beta_0^*} \left(\frac{L}{L_{lc}} \right)^{n-m} = \mu_{ds0} \left(\frac{L}{L_{lc}} \right)^{n-m}, \quad L < L_{lc} \quad (11.42)$$

where μ_{ds0} is the value of the coefficient of summits deformation component of the coefficient of friction at macroscale ($L \geq L_{lc}$).

The scale-dependence for the two-body deformation component of the coefficient of friction is presented in Fig. 11.13 for $m = 0.5$, $n = 0.2$ (Bhushan and Nosonovsky 2004a). The coefficient of friction increases with decreasing scale, according to Eq. 11.42. This effect is a consequence of increasing average slope or roughness angle.

11.4.3 Three-Body Deformation Friction

In this section of the paper, size distribution of particles will be idealized according to the exponential, normal, and log normal density functions, since these distributions are the most common in nature and in industrial applications (see Appendix). The probability for a particle of a given size to be trapped at the interface depends on the size of the contact region. Particles at the edge of the region of contact are likely to leave the contact area, whereas those in the middle are

likely to be trapped. The ratio of the edge region area to the total apparent area of contact increases with decreasing scale size. Therefore, the probability for a particle to be trapped decreases, as well as the three-body component of the coefficient of friction (Bhushan and Nosonovsky 2004a).

Let us consider a square region of contact of two rough surfaces with a length L (relevant scale length), with the density of debris of η particles per unit area (Fig. 11.14) (Bhushan and Nosonovsky 2004a). We assume that the particles have the spherical form and that $p(d)$ is the probability density function of particle size. It is also assumed that, for a given diameter, particles at the border region of the contact zone of the width $d/2$ are likely to leave the contact zone with a certain probability α , whereas particles at the center of the contact region are likely to be trapped. It should be noted that particles in the corners of the contact region can leave in two different directions. Therefore, for them the probability to leave is 2α . The total nominal contact area is equal to L^2 , the area of the border region, without the corners, is equal to $4(L - d)d/2$, and the area of the corners is equal to d^2 .

The probability density of size distribution for the trapped particles $p_{tr}(d)$ can be calculated by

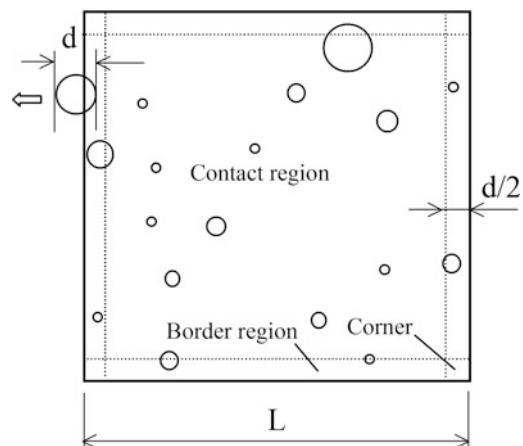


Fig. 11.14 Schematics of debris at the contact zone and at its border region. A particle of diameter d in the border region of $d/2$ is likely to leave the contact zone (Bhushan and Nosonovsky 2004a)

multiplying $p(d)$ by one minus the probability of a particle with diameter d to leave; the latter is equal to the ratio of the border region area multiplied by a corresponding probability of the particle to leave and divided by the total contact area (Bhushan and Nosonovsky 2004a)

$$\begin{aligned} p_{tr}(d) &= p(d) \left(1 - \frac{2\alpha(L-d)d + 2\alpha d^2}{L^2} \right) \\ &= p(d) \left(1 - \frac{2\alpha d}{L} \right), \\ d &< \frac{L}{2\alpha} \end{aligned} \quad (11.43)$$

The ratio of the number of trapped particles to the total number of particles, average radius of a trapped particle \bar{d} , and average square of trapped particles $\overline{d^2}$, as functions of L , can be calculated as

$$\begin{aligned} n_{tr} &= \frac{\int_0^{L/2} p_{tr}(d) dd}{\int_0^\infty p(d) dd} = \frac{\int_0^{L/2} p(d) \left(1 - \frac{2\alpha d}{L} \right) dd}{\int_0^\infty p(d) dd} \\ \bar{d} &= \frac{\int_0^{L/2} dp_{tr}(d) dd}{\int_0^{L/2} p_{tr}(d) dd} \\ \overline{d^2} &= \frac{\int_0^{L/2} d^2 p_{tr}(d) dd}{\int_0^{L/2} p_{tr}(d) dd} \end{aligned} \quad (11.44)$$

Let us assume an exponential distribution of particle sizes (Eq. 11.82) with $d_e = 0$. Substituting Eq. 11.82 into Eq. 11.44 and integrating yields for the ratio of trapped particles (Bhushan and Nosonovsky 2004a)

$$\begin{aligned} n_{tr} &= \frac{\int_0^{L/(2\alpha)} \frac{1}{\sigma_e} \exp\left(-\frac{d}{\sigma_e}\right) \left(1 - \frac{2\alpha d}{L}\right) dd}{\int_0^\infty \frac{1}{\sigma_e} \exp\left(-\frac{d}{\sigma_e}\right) dd} \\ &= \exp\left(-\frac{d}{\sigma_e}\right) \frac{\sigma_e - L/(2\alpha) + d}{L/(2\alpha)} \Big|_0^{L/(2\alpha)} \\ &= \frac{2\alpha\sigma_e}{L} \left(\exp\left(-\frac{L}{2\alpha\sigma_e}\right) - 1 \right) + 1 \end{aligned} \quad (11.45)$$

whereas the mean diameter of the trapped particles is

$$\begin{aligned} \bar{d} &= \frac{\int_0^{L/(2\alpha)} d \exp\left(-\frac{d}{\sigma_e}\right) \left(1 - \frac{2\alpha d}{L}\right) dd}{\int_0^{L/(2\alpha)} \exp\left(-\frac{d}{\sigma_e}\right) \left(1 - \frac{2\alpha d}{L}\right) dd} \\ &= \sigma_e \frac{\exp\left(-\frac{L}{2\alpha\sigma_e}\right) \left(1 + \frac{4\alpha\sigma_e}{L}\right) + 1 - \frac{4\alpha\sigma_e}{L}}{\frac{2\alpha\sigma_e}{L} \left(\exp\left(-\frac{L}{2\alpha\sigma_e}\right) - 1 \right) + 1} \end{aligned} \quad (11.46)$$

and the mean square radius of the trapped particles is

$$\begin{aligned} \overline{d^2} &= \frac{\int_0^{L/(2\alpha)} d^2 \exp\left(-\frac{d}{\sigma_e}\right) \left(1 - \frac{2\alpha d}{L}\right) dd}{\int_0^{L/(2\alpha)} \exp\left(-\frac{d}{\sigma_e}\right) \left(1 - \frac{2\alpha d}{L}\right) dd} \\ &= \sigma_e^2 \frac{\exp\left(-\frac{L}{2\alpha\sigma_e}\right) \left(\frac{L}{2\alpha\sigma_e} + 4 + \frac{12\alpha\sigma_e}{L}\right) + 2 - \frac{12\alpha\sigma_e}{L}}{\frac{2\alpha\sigma_e}{L} \left(\exp\left(-\frac{L}{2\alpha\sigma_e}\right) - 1 \right) + 1} \end{aligned} \quad (11.47)$$

For the normal and log normal distributions, similar calculations can be conducted numerically.

The area supported by particles can be found as the number of trapped particles $\eta L^2 n_{tr}$ multiplied by average particle contact area

$$A_{dp} = \eta L^2 n_{tr} \frac{\pi \overline{d^2}}{4} \quad (11.48)$$

where $\overline{d^2}$ is mean square of particle diameter, η is particle density per apparent area of contact (L^2) and n_{tr} is the number of trapped particles divided by the total number of particles (Bhushan and Nosonovsky 2004a).

The plowing deformation is plastic and, assuming that particles are harder than the bodies, the shear strength τ_{dp} is equal to the shear yield strength of the softer body τ_Y , which is given by Eq. 11.9 with $a = \bar{d}/2$. Combining Eqs. 11.29 with Eqs. 11.9 and 11.48

$$\begin{aligned}\mu_{dp} &= \frac{A_{dp}\tau_{dp}}{W} = \eta \frac{L^2}{W} \frac{\pi \bar{d}^2}{4} n_{tr} \tau_{Y0} \sqrt{1 + 2l_d/\bar{d}} \\ &= \mu_{dp0} n_{tr} \frac{\bar{d}^2}{d_0^2} \frac{\sqrt{1 + 2l_d/\bar{d}}}{\sqrt{1 + 2l_d/d_0}}\end{aligned}\quad (11.49)$$

where \bar{d} is mean particle diameter, d_0 is the macroscale value of mean particle diameter, and μ_{dp0} is macroscale ($L \rightarrow \infty$, $n_{tr} \rightarrow 1$) value of the third-body deformation component of the coefficient of friction given as

$$\mu_{dp0} = \eta \frac{L^2}{W} \frac{\pi \bar{d}_0^2}{4} \tau_{Y0} \sqrt{1 + 2l_d/\bar{d}_0} \quad (11.50)$$

Scale-dependence of the three-body deformation component of the coefficient of friction is presented in Fig. 11.15, based on Eq. 11.49 (Bhushan and Nosonovsky 2004a). The number of trapped particles, as well as the three-body deformation component of the coefficient of friction, are presented as a function of scale size divided by α for the exponential, normal, and log normal distributions. The dependence of μ_d/μ_{d0} is shown as a function of $L/(\alpha\sigma_e)$ for the exponential distribution and normal distribution, for $d_n = d_e = 2\sigma_e$ and $l_d/\sigma_e = 1$, whereas for the log normal distribution the results are presented as a function of L/α , for $\ln(d_{ln}) = 2$, $\sigma_{ln} = 1$, and $l_d/\sigma_{ln} = 1$. This component of the coefficient of friction decreases for all of the three distributions. The results are shown for $l_d/\sigma_{ln} = 1$. However, variation of l_d/σ_{ln} in the range between 0.1 and 10 does not change the shape of the curve significantly. The decrease of the three-body deformation friction force with decreasing scale results from this component being small at the nanoscale.

11.4.4 Ratchet Mechanism

Surface roughness can have an appreciable influence on friction during adhesion. If one of the contacting surfaces has asperities of much

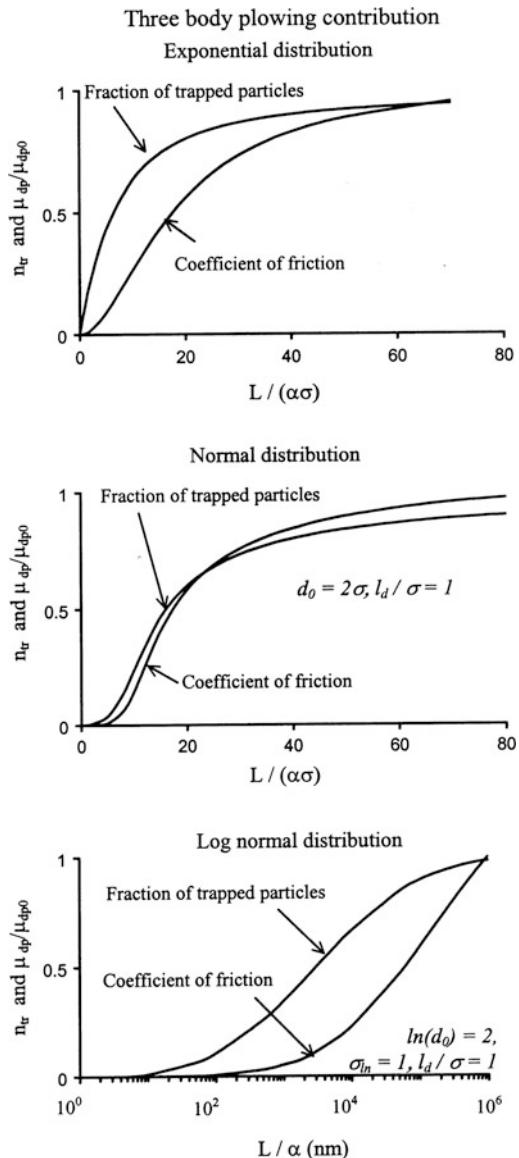
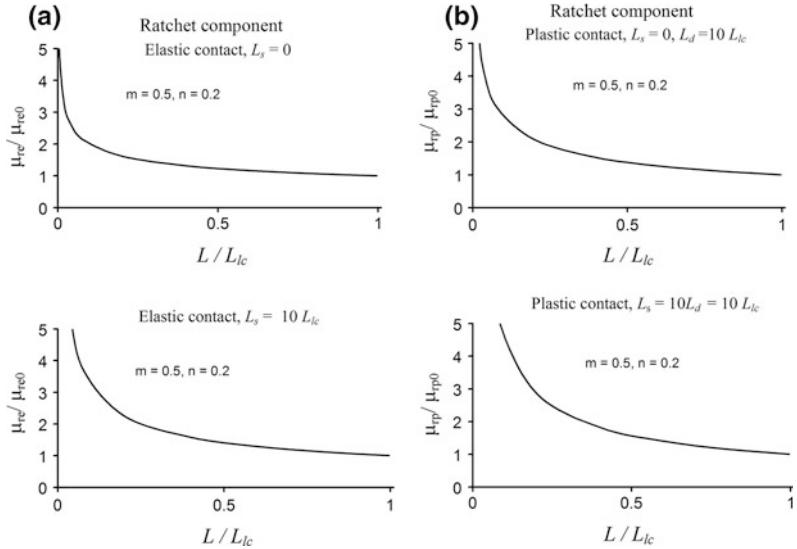


Fig. 11.15 The number of trapped particles divided by the total number of particles and three-body deformation component of the coefficient of friction, normalized by the macroscale value, for three different distributions of debris-size: **a** exponential, **b** normal, **c** and log-normal distributions (Bhushan and Nosonovsky 2004a)

smaller lateral size, such that a small tip slides over an asperity, having the average angle θ_r (so-called ratchet mechanism), the corresponding component of the coefficient of friction is given by

Fig. 11.16 Normalized results for the ratchet component of the coefficient of friction, as a function of scale, for scale independent ($L_s = 0$) and scale dependent ($L_s = 10 L_{lc}$) shear strength ($m = 0.5$, $n = 0.2$). **a** Elastic contact, **b** Plastic contact ($L_d = 10 L_{lc}$) (Nosonovsky and Bhushan 2005)



$$\mu_r = \mu_a \tan^2 \theta_r \quad (11.51)$$

where μ_r is the ratchet mechanism component of friction (Bhushan 2013). Combining Eqs. 11.15, 11.41, 11.38, and 11.39 yields for the scale-dependence of the ratchet component of the coefficient of friction in the case of elastic, μ_{re} , and plastic contact, μ_{rp}

$$\begin{aligned} \mu_{re} &= \mu_{ae} \left(\frac{2\sigma_0}{\pi\beta_0^*} \left(\frac{L}{L_{lc}} \right)^{n-m} \right)^2 \\ &= \frac{\mu_{re0}}{\sqrt{1 + (l_s/\bar{a}_0)}} \left(\frac{L}{L_{lc}} \right)^{n-m} \sqrt{1 + (L_s/L)^m} \\ L &< L_{lc} \end{aligned} \quad (11.52)$$

$$\begin{aligned} \mu_{rp} &= \mu_{ap} \left(\frac{2\sigma_0}{\pi\beta_0^*} \left(\frac{L}{L_{lc}} \right)^{n-m} \right)^2 \\ &= \mu_{rp0} \left(\frac{L}{L_{lc}} \right)^{2(n-m)} \sqrt{\frac{1 + (l_d/\bar{a}_0)}{1 + (l_s/\bar{a}_0)}} \sqrt{\frac{1 + (L_s/L)^m}{1 + (L_d/L)^m}} \\ L &< L_{lc} \end{aligned} \quad (11.53)$$

where μ_{re0} and μ_{rp0} are the macroscale values of the ratchet component of the coefficient of friction for elastic and plastic contact correspondingly (Nosonovsky and Bhushan 2005).

Scale-dependence of the ratchet component of the coefficient of friction, normalized by the macroscale value, is presented in Fig. 11.16 for

scale-independent adhesional shear strength, $\tau_a = const$, ($L_s = 0$) and for scale-dependent τ_a ($L_s = 10 L_d$), based on Eqs. 11.52 and 11.53 (Nosonovsky and Bhushan 2005). The ratchet component during adhesional elastic friction, μ_{re} , is presented in Fig. 11.16a. It is observed that with decreasing scale, μ_{re} increases. The ratchet component during adhesional plastic friction, μ_{rp} , is presented in Fig. 11.16b. It is observed, that, for $L_s = 0$, with decreasing scale, μ_{rp} increases (Nosonovsky and Bhushan 2005).

11.4.5 Meniscus Analysis

During contact, if a liquid is introduced at the point of asperity contact, the surface tension results in a pressure difference across a meniscus surface referred to as capillary pressure or Laplace pressure. The attractive force for a sphere in contact with a planar surface is proportional to the sphere radius R_p . For a sphere close to a surface with separation s , or for a sphere close to a surface with continuous liquid film (Bhushan 2013)

$$F_m \propto R_p \quad (11.54)$$

The case of multiple-asperity contact is shown in Fig. 11.1b. Note, that both contacting and near-contacting asperities wetted by the liquid film contribute to the total meniscus force.

A statistical approach can be used to model the contact. In general, given the interplanar separation s_d , the mean peak radius \bar{R}_p , the thickness of liquid film h_f , the surface tension γ , liquid contact angle between the liquid and surface θ , and the total number of summits in the nominal contact area N

$$F_m = 2\pi\bar{R}_p\gamma(1 + \cos\theta)N \quad (11.55)$$

In Eq. 11.55, γ and θ are material properties that are not expected to depend on scale, whereas \bar{R}_p and N depend on surface topography and are scale-dependent, according to Eqs. 11.18 and 11.20. Therefore, Eq. 11.55 is rewritten as

$$F_m \propto \bar{R}_p N = F_{m0} \left(\frac{L}{L_{lwl}} \right)^{m-2n}, \quad L < L_{lwl} \quad (11.56)$$

where F_{m0} is the macroscale value of the meniscus force ($L \geq L_{lwl}$).

Scale-dependence of the meniscus force is presented in Fig. 11.17, based on Eq. 11.56 for $m = 0.5$, $n = 0.2$ (Bhushan and Nosonovsky 2004b). It may be observed that, depending on the value of D , the meniscus force may increase or decrease with decreasing scale size.

11.4.6 Total Value of Coefficient of Friction and Transition from Elastic to Plastic Regime

During transition from elastic to plastic regime, contribution of each of the three components of the coefficient of friction in Eq. 11.29 changes. In the elastic regime, the dominant contribution is expected to be adhesion involving elastic deformation, and in the plastic regime the dominant contribution is expected to be deformation. Therefore, in order to study transition from elastic to plastic regime, the ratio of deformation to the adhesion component should be considered. The expression for the total value of the

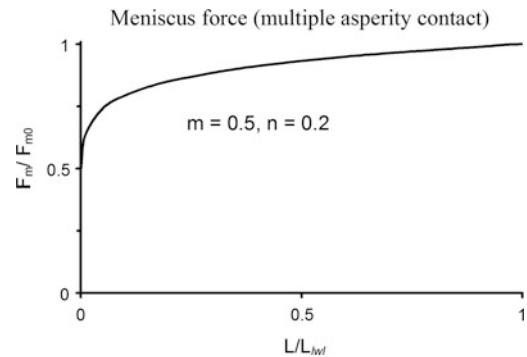


Fig. 11.17 Meniscus force for $m = 0.5$, $n = 0.2$ (Bhushan and Nosonovsky 2004b)

coefficient of friction, which includes meniscus force contribution, based on Eqs. 11.29 and 11.31 can be rewritten as (Bhushan and Nosonovsky 2004a)

$$\mu_{wet} = \mu_a \left(1 + \frac{\mu_{ds}}{\mu_a} + \frac{\mu_{dp}}{\mu_a} \right) \left(1 + \frac{F_m}{W} \right) \quad (11.57)$$

The ratchet mechanism component is ignored here since it is present only in special cases. Results in the preceding subsection provide us with data about the adhesion and two-body and three-body deformation components of the coefficient of friction, normalized by their values at the macroscale. However, that analysis does not provide any information about their relation to each other or about transition from the elastic to plastic regime. In order to analyze the transition from pure adhesion involving elastic deformation to plastic deformation, a transition index ϕ can be considered (Bhushan and Nosonovsky 2004a). The transition index is equal to the ratio of average pressure in the elastic regime (normal load per real area of elastic contact) to hardness or simply the ratio of the real area of plastic contact divided by the real area of elastic contact

$$\phi = \frac{W}{A_{re}H} = \frac{A_{rp}}{A_{re}} \quad (11.58)$$

Using Eqs. 11.17 and 11.22, the scale-dependence of ϕ is

$$\phi = \frac{W}{A_{re}(L/L_{lc})^{m-n} H_0 \sqrt{1 + (L_s/L)^m}} = \phi_0 \frac{\sqrt{1 + (L_s/\bar{a})(L/L_{lc})^{n-m}}}{\sqrt{1 + (L_s/L)^m}} \quad L < L_{lc} \quad (11.59)$$

where ϕ_0 is the macroscale value of the transition index (Bhushan and Nosonovsky 2004a).

With a low value of ϕ close to zero, the contacts are mostly elastic and only adhesion contributes to the coefficient of friction involving elastic deformation. Whereas with increasing ϕ approaching unity, the contacts become predominantly plastic and deformation becomes a dominant contributor. It can be argued that A_{ds}/A_{re} and A_{dp}/A_{re} will also be a direct function of ϕ , and here these will be assumed to have a linear relationship.

Next, the ratio of adhesion and deformation components of the coefficient of friction in terms of ϕ is obtained. In this relationship, τ_{ds} and τ_{dp} are equal to the shear yield strength, which is

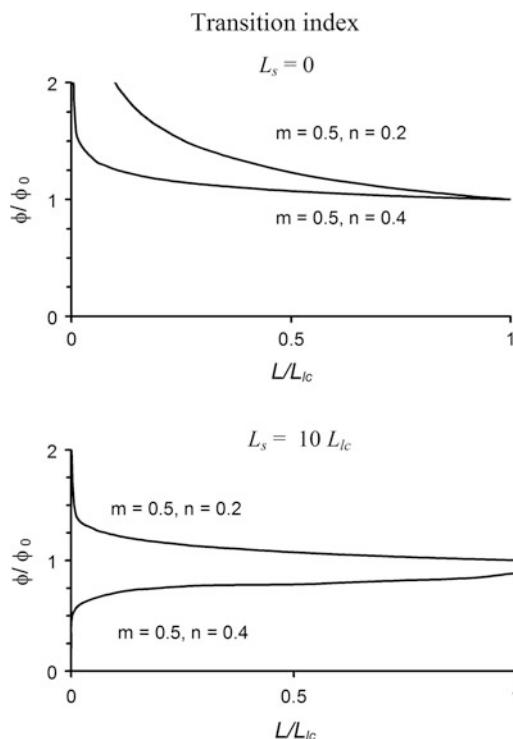


Fig. 11.18 The transition index as a function of scale. Presented for $m = 0.5$, $n = 0.2$ and $m = 0.5$, $n = 0.4$ (Nosonovsky and Bhushan 2005)

proportional to hardness, and can be obtained from Eq. 11.9 using Eq. 11.19 and 11.36

$$\frac{\mu_{ds}}{\mu_{ae}} = \frac{A_{ds}\tau_{ds}}{A_{re}\tau_a} \propto \phi \frac{\tau_{ds}}{\tau_a} = \phi \frac{\tau_{ds0}\sqrt{1 + (L_d/L)^m}}{\tau_{a0}\sqrt{1 + (L_s/L)^m}}, \quad L < L_{lc} \quad (11.60)$$

$$\frac{\mu_{dp}}{\mu_{ae}} = \frac{A_{dp}\tau_{dp}}{A_{re}\tau_a} \propto \phi \frac{\tau_{dp}}{\tau_a} = \phi \frac{\tau_{y0}\sqrt{1 + 2l_d/\bar{d}}}{\tau_{a0}\sqrt{1 + (L_s/L)^m}}, \quad L < L_{lc} \quad (11.61)$$

The sum of the adhesion and deformation components (Bhushan and Nosonovsky 2004a)

$$\mu_{wet} = \mu_{ae} \left[1 + \varphi \left(\frac{\tau_{ds0}\sqrt{1 + (L_d/L)^m}}{\tau_{a0}\sqrt{1 + (L_s/L)^m}} + \frac{\tau_{y0}\sqrt{1 + 2l_d/\bar{d}}}{\tau_{a0}\sqrt{1 + (L_s/L)^m}} \right) \right] \times \left(1 + \frac{F_{m0}}{W} \left(\frac{L}{L_{lwI}} \right)^{m-2n} \right), \quad L < L_{lc} \quad (11.62)$$

Note that ϕ itself is a complicated function of L , according to Eq. 11.59.

Scale-dependence of the transition index, normalized by the macroscale value, is presented in Fig. 11.18, based on Eq. 11.59 (Nosonovsky and Bhushan 2005). It is observed that for $L_s = 0$, the transition index decreases with increasing scale. For $L_s = 10L_{lc}$, the same trend is observed for $m > 2n$, but, in the case $m < 2n$, ϕ decreases. An increase of the transition index means that the ratio of plastic to elastic real areas of contact increases. With decreasing scale, the mean radius of contact decreases, causing hardness enhancement and decrease of the plastic area of contact. Based on this, the model may predict an increase or decrease of the transition index, depending on whether elastic or plastic area decreases faster.

The dependence of the coefficient of friction on ϕ is illustrated in Fig. 11.19, based on Eq. 11.62 (Bhushan and Nosonovsky 2004a). It

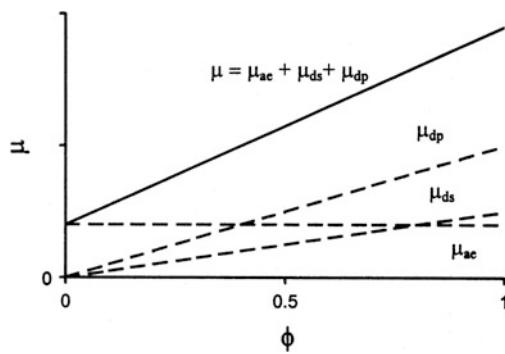


Fig. 11.19 The coefficient of friction (dry contact) as a function of the transition index for given scale length L . With increasing ϕ and onset of plastic deformation, both μ_{ds} and μ_{dp} grow, as a result of this, the total coefficient of friction μ grows as well (Bhushan and Nosonovsky 2004a)

is assumed in the figure that the slope for the dependence of μ_{dp} on ϕ is greater than the slope for the dependence of μ_{ds} on ϕ . For ϕ close to zero, the contact is predominantly elastic, whereas for ϕ approaching unity the contact is predominantly plastic.

11.4.7 Comparison with the Experimental Data

Experimental data on friction at micro- and nanoscale are presented in this subsection and compared with the model. First, a single-asperity predominantly-elastic contact is considered (Bhushan and Nosonovsky 2003). Then, transition to plastic deformation involving multiple asperity contacts is analyzed (Nosonovsky and Bhushan 2005).

11.4.7.1 Single-Asperity Predominantly-Elastic Contact

Nanoscale dependence of friction force upon the normal load was studied for various tip materials. Carpick et al. (1996) studied a Pt-coated AFM tip versus mica in ultra-high vacuum (UHV); Schwarz et al. (1997) studied a Si tip versus diamond and amorphous carbon; and Bhushan and

Kulkarni (1996) studied a Si_3N_4 tip on Si, SiO_2 , and diamond, Fig. 11.20a. Homola et al. (1990) conducted SFA experiments with mica rolls with a single contact zone (before onset of wear), Fig. 11.21b. Contacts relevant in these experiments can be considered as single-asperity, predominantly-elastic in all of these cases. For a single-asperity elastic contact of radius a , the expression for μ is given by Eq. 11.17. For the limit of a small contact radius $a \ll l_s$, Eq. 11.13 combined with the Hertzian dependence of the contact area upon the normal load (Eq. 11.33) yields

$$F_e \approx \pi a^2 \tau_0 \sqrt{l_s/a} \propto a^{3/2} \propto W^{1/2} \quad (11.63)$$

If an adhesive pull-off force W_0 is large, Eq. 11.63 can be modified as

$$F_e = C_0 \sqrt{W + W_0} \quad (11.64)$$

where C_0 is a constant. Friction force increases with square root of the normal load, opposed to the two third exponent in scale-independent analysis.

The results in Fig. 11.20 demonstrate a reasonable agreement of the experimental data with the model (Bhushan and Nosonovsky 2003). The platinum-coated tip versus mica (Carpick et al. 1996) has a relatively high pull-off force and the data fit with $C_0 = 23.7 \text{ (nN)}^{1/2}$ and $W_0 = 170 \text{ nN}$. For the silicon tip versus amorphous carbon and natural diamond, the fit is given by $C_0 = 8.0, 19.3 \text{ (nN)}^{1/2}$ and small W_0 . For the virgin $\text{Si}(111)$, SiO_2 , and natural diamond sliding versus Si_3N_4 tip (Schwarz et al. 1997), the fit is given by $C_0 = 0.40, 0.76, 0.86 \text{ (nN)}^{1/2}$ for $\text{Si}(111)$, SiO_2 , and diamond, respectively and small W_0 . For two mica rolls (Homola et al. 1990), the fit is given by $C_0 = 10 \text{ N}^{1/2}$ and $W_0 = 0.5 \text{ N}$ (Bhushan and Nosonovsky 2003).

AFM experiments provide data on nanoscale, whereas SFA experiments provide data on microscale. Next, we study scale-dependence on the shear strength based on these data. In the AFM measurements by Carpick et al. (1996), the average shear strength during sliding for a

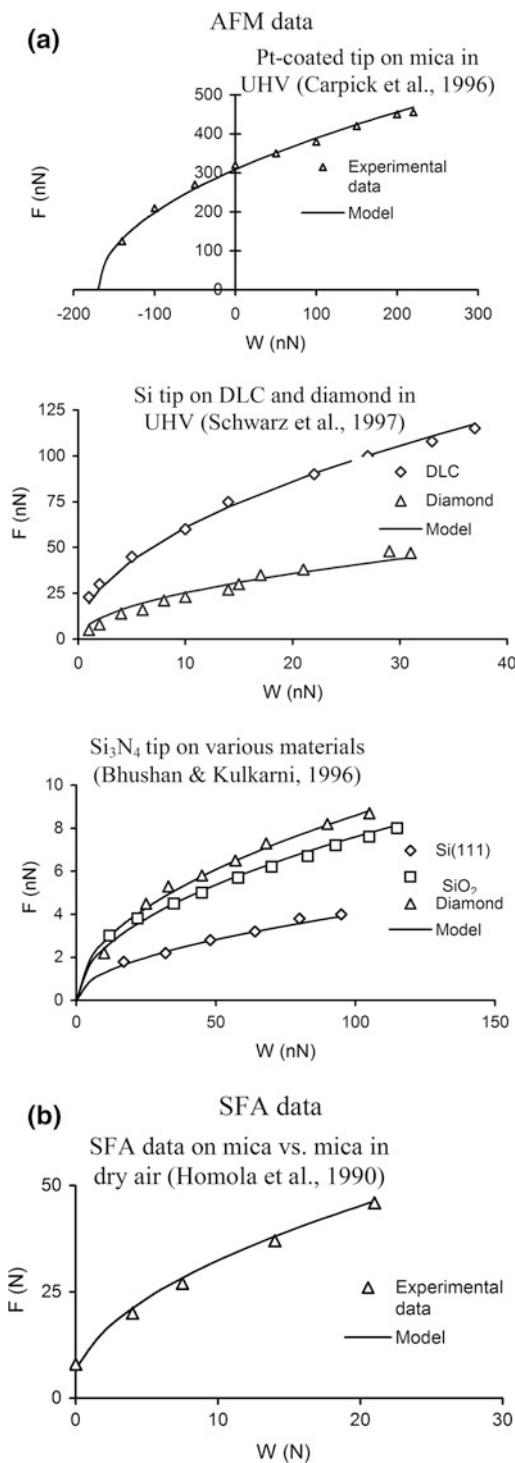


Fig. 11.20 Summary of **a** AFM data and **b** SFA data for friction force as a function of normal load (Bhushan and Nosonovsky 2003)

Pt-mica interface was reported as 0.86 GPa, whereas the pull-off contact radius was reported as 7 nm. In the SFA measurements by Homola et al. (1990), the average shear strength during sliding for a mica-mica interface was reported as 25 MPa, whereas the contact area during high loads was on the order of 10^{-8} m^2 , which corresponds to a contact radius on the order 100 μm . To normalize shear strength, we need shear modulus. The shear modulus for mica is $G_{\text{mica}} = 25.7 \text{ GPa}$ (Clauser 1963) and for Pt is $G_{\text{Pt}} = 63.7 \text{ GPa}$ (Bhushan and Gupta 1997). For a mica-Pt interface, the effective shear modulus is

$$G = 2G_{\text{mica}}G_{\text{Pt}}/(G_{\text{mica}} + G_{\text{Pt}}) = 36.6 \text{ GPa} \quad (11.65)$$

This yields the value of the shear stress normalized by the shear modulus $\tau_a/G = 2.35 \times 10^{-2}$ for the AFM data of Carpick et al. (1996), and 9.73×10^{-4} for the SFA data. These values are presented in Fig. 11.21 together with the values predicted by the model for assumed values of $l_s = 1$ and 10 μm (Bhushan and Nosonovsky 2003). It can be seen that the model (Eq. 11.13) provides an explanation of adhesional shear strength increase with a scale decrease (Bhushan and Nosonovsky 2003).

11.4.7.2 Transition to Predominantly Plastic Deformation Involving Multiple Asperity Contacts

Next, we analyze the effect of transition from predominantly elastic adhesion to predominantly plastic deformation involving multiple-asperity contacts (Nosonovsky and Bhushan 2005). The data on nano- and microscale friction for various materials are presented in Table 11.2 based on Ruan and Bhushan (1994), Liu and Bhushan (2003), and Bhushan et al. (2004) for Si(100), graphite (HOPG), natural diamond, and diamond like carbon (DLC). There are several factors responsible for the differences in the coefficients of friction at micro- and nanoscale. Among them are the contributions from ratchet mechanism, wear and contaminant particles, scale dependent

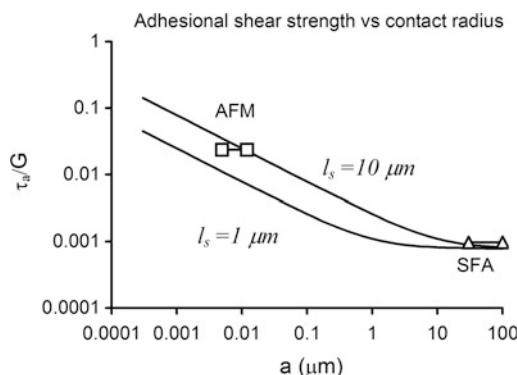


Fig. 11.21 Shear stress as a function of contact radius. Microscale and nanoscale data compared with the model for $l_s = 1 \mu\text{m}$ and $l_s = 10 \mu\text{m}$ (Bhushan and Nosonovsky 2003)

roughness and mechanical properties, transition from elasticity to plasticity, and meniscus effect. The ratchet mechanism and meniscus effect result in an increase of friction with decreasing scale and cannot explain the decrease of friction found in the experiments. The contribution of wear and contaminant particles is more significant at macro/microscale because of the larger number of trapped particles (Fig. 11.15). It can be argued that for the nanoscale AFM experiments the contacts are predominantly elastic, and adhesion is the main contribution to the friction, whereas for the microscale experiments the contacts are predominantly plastic and deformation is an important factor. Therefore, transition from elastic contacts in nanoscale contacts to plastic deformation in microscale contacts is an important effect (Nosonovsky and Bhushan 2005).

According to Eq. 11.29, the friction force depends on the shear strength and a relevant real area of contact. For calculation of contact radii and contact pressures, the elastic modulus, Poisson's ratio, and hardness for various samples, are required and presented in Table 11.2 (Anonymous 1959, 2002; Field 1992; Bhushan and Venkatesan 1993; Bhushan and Gupta 1997; Bhushan 1999b). In the nanoscale AFM experiments a sharp tip was slid against a flat sample. The apparent contact size and mean contact pressures are calculated based on the assumption, that the contacts are single asperity, elastic

contacts (contact pressures are small compared to hardness). Based on the Hertz equation (Johnson 1985), for a spherical asperity of radius R in contact with a flat surface, with an effective elastic modulus E^* , under normal load W , the contact radius a and mean apparent contact pressure p_a are given by

$$a = \left(\frac{3WR}{4E^*} \right)^{1/3} \quad (11.66)$$

$$p_a = \frac{W}{\pi a^2} \quad (11.67)$$

The surface energy effect (Bhushan 2013) was neglected in Eqs. 11.66 and 11.67, because the experimental value of the normal adhesion force was small compared to W (Ruan and Bhushan 1994). The calculated values of a and p_a for the relevant normal load are presented in Table 11.2 (Nosonovsky and Bhushan 2005).

In the microscale experiments, a ball was slid against a nominally flat surface. The contact in this case is multiple-asperity contact due to the roughness, and the contact pressure of the asperity contacts is higher than the apparent pressure. For calculation of a characteristic scale length for the multiple-asperity contacts, which is equal to the apparent length of contact, Eq. 11.66 was also used. Apparent radius and mean apparent contact pressure for microscale contacts at relevant load ranges are also presented in Table 11.2 (Nosonovsky and Bhushan 2005). This analysis provides an upper limit on apparent radius and lower limit on the mean contact pressure.

A quantitative estimate of the effect of the shear strength and the real area of contact on friction is presented in Table 11.3. The friction force at mean load (average of maximum and minimum loads) is shown, based on the experimental data presented in Table 11.2. For microscale data, the real area of contact was estimated based on the assumption that the contacts are plastic and based on Eq. 11.33 for mean loads given in Table 11.2. For nanoscale data, the apparent area of contact was on the

Table 11.2 Micro- and nanoscale values of the coefficient of friction, typical physical properties of specimens, and calculated apparent contact radii and apparent contact pressures at loads used in micro- and nanoscale measurements

Specimen	Coefficient of friction	Elastic modulus (GPa)		Poisson's ratio	Hardness (GPa)	Apparent contact radius at test load for		Mean apparent pressure at test load for
		Microscale	Nanoscale			Microscale (μm) (upper limit)	Nanoscale (nm)	
Si(100) wafer	0.47 ^a	0.06 ^c	130 ^{e,f}	0.28 ^f	9–10 ^{e,f}	0.8–2.2 ^a	1.6–3.4 ^c	0.05–0.13 ^a
Graphite (HOPG)	0.1 ^b	0.006 ^c	9–15 ^g (9)	–(0.25)	0.01 ^j	62 ^b	3.4–7.4 ^c	0.082 ^b
Natural diamond	0.2 ^b	0.05 ^c	1140 ^h	0.07 ^h	80–104 ^{g,h}	21 ^b	1.1–2.5 ^c	0.74 ^b
DLC film	0.19 ^a	0.03 ^d	280 ⁱ	0.25 ⁱ	20–30 ⁱ	0.7–2.0 ^a	1.3–2.9 ^d	0.06–0.16 ^a
For calculation purposes, it is assumed that contacts on micro- and nanoscale are single-asperity elastic contacts (Nosonovsky and Bhushan 2005)								

^a500 μm radius Si(100) ball at 100–2000 μN and 720 μm/s in dry air, Bhushan et al. (2004)

^b3 mm radius Si₃N₄ ball (Elastic modulus 310 GPa, Poisson's ratio 0.22, Bhushan and Gupta 1997) at 1 N and 800 μm/s, Ruan and Bhushan (1994)

^c50 nm radius Si₃N₄ tip at load range from 10 to 100 nN and 0.5 nm/s, in dry air, Ruan and Bhushan (1994)

^d50 nm radius Si₃N₄ tip at load range from 10 to 100 nN in dry air, Bhushan et al. (2004)

^eBhushan and Venkatesan (1993)

^fAnonymous (2002)

^gBhushan and Gupta (1997)

^hField (1992)

ⁱBhushan (1999b)

^jAnonymous (1959)

order of several square nanometers, and it was assumed that the real area of contact is comparable with the mean apparent area of contact, which can be calculated for the mean apparent contact radius, given in Table 11.2. The estimate provides the upper limit of the real area of contact. The lower limit of the shear strength is calculated as friction force divided by the upper limit of the real area of contact, and presented in Table 11.3 (Nosonovsky and Bhushan 2005). Based on the data in Table 11.3, for Si(100), natural diamond, and DLC film, the microscale value of shear strength is about two orders of magnitude higher, than the nanoscale value, which indicates that transition from the adhesion to the deformation mechanism of friction and the third-body effect are responsible for an increase of friction at microscale. For graphite, this effect is less pronounced due to the molecularly smooth structure of the graphite surface (Nosonovsky and Bhushan 2005).

Based on data available in the literature (Bhushan and Kulkarni 1996), load-dependence of friction at nano/microscale as a function of normal load is presented in Fig. 11.22. Coefficient of friction was measured for a Si_3N_4 tip

versus Si, SiO_2 , and natural diamond using an AFM. They reported that for low loads, the coefficient of friction is independent of load and increases with increasing load after a certain load. It is noted that the critical value of loads for Si and SiO_2 corresponds to stresses equal to their hardness values, which suggests that transition to plasticity plays a role in this effect. The friction values at higher loads for Si and SiO_2 approach to that of macroscale values. This result is consistent with predictions of the model for plastic contact (Fig. 11.11), which states that with increasing normal load, the long wavelength limit for the contact parameters decreases. This decrease results in violation of the condition $L < L_{lc}$, and the contact parameters and the coefficient of friction reach the macroscale values, as was discussed earlier. It must be noted that the values of $m = 0.5$ and $n = 0.2$ are taken based on available data for the glass-ceramic disk (Fig. 11.9). These parameters depend on material and on surface preparation, and may be different for Si, SiO_2 , and natural diamond. However, no experimental data on the scale-dependence of roughness parameters for the materials of interest are available.

Table 11.3 Mean friction force, real area of contact, and lower limit of shear strength (Nosonovsky and Bhushan 2005)

Specimen	Friction force at mean load ^a		Upper limit of real area of contact at mean load		Lower limit of mean shear strength (GPa)	
	Microscale (mN)	Nanoscale (nN)	Microscale ^b (μm^2)	Nanoscale ^c (nm^2)	Microscale ^d	Nanoscale ^d
Si(100) wafer	0.49	3.3	0.11	19	4.5	0.17
Graphite (HOPG)	100	0.33	10^5	92	0.001	0.004
Natural diamond	200	2.7	10.9	10	18.4	0.27
DLC film	0.2	1.7	0.042	14	4.8	0.12

^aBased on the data from Table 11.2. Mean load at microscale is 1050 μN for Si(100) and DLC film and 1 N for HOPG and natural diamond, and 55 nN for all samples at nanoscale

^bFor plastic contact, based on hardness values from Table 11.2. Scale-dependent hardness value will be higher at relevant scale, presented values of the real area of contact are an upper estimate

^cUpper limit for the real area is given by the apparent area of contact calculated based on the radius of contact data from Table 11.2

^dLower limit for the mean shear strength is obtained by dividing the friction force by the upper limit of the real area of contact

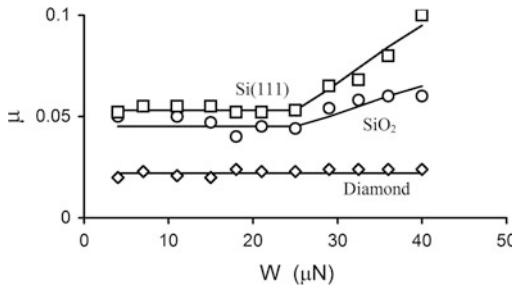


Fig. 11.22 Coefficient of friction as a function of normal load (Bhushan and Kulkarni 1996)

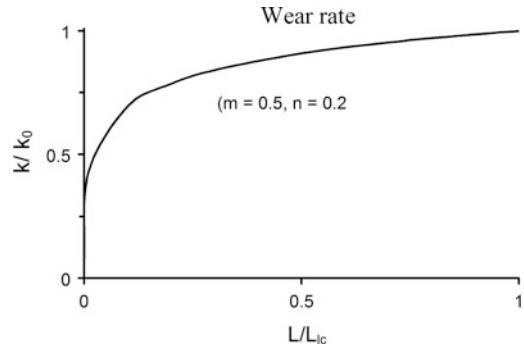


Fig. 11.23 The wear coefficient as a function of scale, presented for $m = 0.5$, $n = 0.2$ (Bhushan and Nosonovsky 2003)

11.5 Scale Effects in Wear

The amount of wear during adhesive or abrasive wear involving plastic deformation is proportional to the load and sliding distance x , divided by hardness (Bhushan 2013)

$$v = k_0 \frac{Wx}{H} \quad (11.67)$$

where v is volume of worn material and k_0 is a nondimensional wear coefficient. Using Eqs. 11.10 and 11.19, the relationships can be obtained for scale-dependence of the coefficient of wear in the case of the fractal surface and power-law-dependence of roughness parameters

$$v = k \frac{Wx}{H_0} \quad (11.68)$$

and

$$k = \frac{k_0}{\sqrt{1 + (l_d/\bar{a})}} = \frac{k_0}{\sqrt{1 + (L_d/L)^m}}, \quad L < L_{lw} \quad (11.69)$$

where k is scale-dependent wear coefficient, and k_0 corresponds to the macroscale limit of the value of k (Bhushan and Nosonovsky 2004b).

Scale-dependence of the wear coefficient is presented in Fig. 11.23 for $m = 0.5$ and $n = 0.2$, based on Eq. 11.69. It is observed that the wear coefficient decreases with decreasing scale; this is due to the fact that the hardness increases with decreasing mean contact size.

11.6 Scale Effect in Interface Temperature

Frictional sliding is a dissipative process, and frictional energy is dissipated as heat over asperity contacts. Therefore, a high amount of heat per unit area is generated during sliding. A contact is formed and destroyed as one asperity passes the other at a given velocity. When an asperity comes into contact with another asperity, the real area of contact starts to grow. When the asperities are directly above each other, the area is at maximum. As they move away from each other, the area starts to get smaller. There are a number of contacts at a given time during sliding. For each individual asperity contact, a flash temperature rise can be calculated. High temperature rise affects mechanical and physical properties of contacting bodies.

For thermal analysis, a dimensionless Peclet number is used

$$L_p = \frac{6Va_{max}}{16\kappa_t} \quad (11.70)$$

where V is sliding velocity, a_{max} is maximum radius of contact for a given contact spot, and κ_t is thermal diffusivity. This parameter indicates whether the sliding is high-speed or low-speed. If $L_p > 10$, the contact falls into the category of high speed; if $L_p < 0.5$, it falls into the category of low speed; if $0.5 \leq L_p \leq 10$, a transition

regime should be considered (Bhushan 2013). For high L_p , there is not enough time for the heat to flow to the sides during the lifetime of the contact and the heat flows only in the direction perpendicular to the sliding surface. Based on the numerical calculations for flash temperature rise of asperity contact for adhesional contact (Bhushan 2013), the following relation holds for the maximum temperature rise T , normalized by the rate at which heat is generated q , divided by the volumetric specific heat ρc_p

$$\begin{aligned} \frac{T\rho c_p V}{q} &= 0.95 \left(\frac{2Va_{\max}}{\kappa_t} \right)^{1/2}, \quad L_p > 10 \\ &= 0.33 \left(\frac{2Va_{\max}}{\kappa_t} \right), \quad L_p < 0.5 \end{aligned} \quad (11.71)$$

The rate at which heat generated per time per unit area depends on the coefficient of friction μ , sliding velocity V , apparent normal pressure p_a , and ratio of the apparent to real areas of contact (A_a/A_r)

$$q = \mu p_a V \frac{A_a}{A_r} \quad (11.72)$$

Based on Eqs. 11.71 and 11.72,

$$\begin{aligned} \frac{T\rho c_p}{p_a} &= 0.95 \frac{A_r}{A_a} \mu \left(\frac{2Va_{\max}}{\kappa_t} \right)^{1/2}, \quad L_p > 10 \\ &= 0.33 \frac{A_r}{A_a} \mu \left(\frac{2Va_{\max}}{\kappa_t} \right), \quad L_p < 0.5 \end{aligned} \quad (11.73)$$

For a multiple asperity contact, mean temperature in terms of average of maximum contact size can be written as

$$\begin{aligned} \frac{\bar{T}\rho c_p}{p_a} &= 0.95 \frac{A_r}{A_a} \mu \left(\frac{2V\bar{a}_{\max}}{\kappa_t} \right)^{1/2}, \quad L_p > 10 \\ &= 0.33 \frac{A_r}{A_a} \mu \left(\frac{2V\bar{a}_{\max}}{\kappa_t} \right), \quad L_p < 0.5 \end{aligned} \quad (11.74)$$

In Eq. 11.74, \bar{a}_{\max} , μ , and A_a/A_r are scale-dependent parameters. During adhesional

contact, the maximum radius \bar{a}_{\max} is proportional to the contact radius \bar{a} , and the scale-dependence for \bar{a}_{\max} is given by Eqs. 11.19, for μ by Eqs. 11.38–11.39, and for A_{re} and A_{rp} by Eqs. 11.17 and 11.21. The scale-dependence of q , involving μ and A_r , and \bar{a}_{\max} in Eq. 11.71 can be considered separately and then combined. For the sake of simplicity, we only consider the scale-dependence of \bar{a}_{\max} . For the empirical rule-dependence of surface roughness parameters and the fractal model, in the case of high and low velocity, Eq. 11.74 yields (Bhushan and Nosonovsky 2004b)

$$\begin{aligned} \frac{\bar{T}\rho c_p V}{q} &= 0.95 \left(\frac{2VC_A L^m}{\kappa} \right)^{1/2}, \quad L < L_{lwl}, L_p > 10 \\ &= 0.33 \left(\frac{2VC_A L^m}{\kappa} \right), \quad L < L_{lwl}, L_p < 0.5 \end{aligned} \quad (11.75)$$

Scale-dependence for the ratio of the flash temperature rise to the amount of heat generated per unit time per unit area for a given sliding velocity as a function of scale is presented in Fig. 11.24, based on Eq. 11.75, for the high-speed and low-speed cases. For the empirical rule-dependence of roughness parameters, the results are shown for $m = 0.5$, $n = 0.2$.

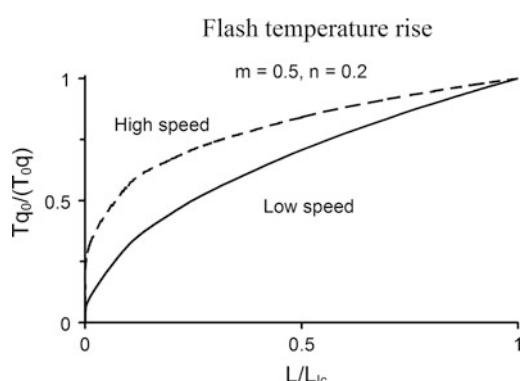


Fig. 11.24 Ratio of the flash temperature rise to the amount of heat generated per unit time per unit area, for a given sliding velocity, as a function of scale. Presented for $m = 0.5$, $n = 0.2$ (Bhushan and Nosonovsky 2004b)

11.7 Closure

This chapter presented a model that explains scale effects in mechanical properties (yield strength, hardness, and shear strength at the interface) and tribology (surface roughness, contact parameters, friction, wear, and interface temperature). Both mechanical properties and roughness parameters are scale-dependent.

According to the strain gradient plasticity theory, the scale-dependence of the so-called geometrically necessary dislocations causes enhanced yield strength and hardness with decreasing scale. The shear strength at the interface is scale-dependent due to the effect of dislocation-assisted sliding. An empirical rule for the scale-dependence of the roughness parameters has been proposed. It was assumed that the standard deviation of surface height and autocorrelation length depend on scale according to a power law when scale is less than the long wavelength limit value.

Both single asperity and multiple asperity contacts were considered. For multiple asperity contacts, based on the empirical power-rule for the scale-dependence of roughness, contact parameters were calculated. The effect of load on the contact parameters was also studied. The effect of increasing load is similar to that of increasing scale because it results in increased relevance of longer wavelength details of roughness of surfaces in contact.

During sliding, adhesion and two- and three-body deformation, as well as ratchet mechanism, contribute to the friction force. These components of the friction force depend on the relevant real areas of contact (dependent on roughness, mechanical properties, and load), average asperity slope, number of trapped particles, and relevant shear strength during sliding. The relevant scaling length is the nominal contact length—contact diameter ($2a$) for a single-asperity contact, only considered in adhesion, and scan length (L) for multiple-asperity contacts, considered in adhesion and deformation.

For the adhesional component of the coefficient of friction, the shear yield strength and hardness increase with decreasing scale. In the

case of elastic contact, the real area of contact is scale-independent for single-asperity contact, and may increase or decrease for multiple-asperity contact depending on roughness parameters. In the case of plastic contact, enhanced hardness results in a decrease in the real area of contact. The adhesional shear strength at the interface may remain constant or increase with decreasing scale due to dislocation-assisted sliding (or microslip). The model predicts that the adhesion component of the coefficient of friction may increase or decrease with scale, depending on the material parameters and roughness. The coefficient of friction during two-body deformation and the ratchet component depend on the average slope of the rough surface. The average slope increases with scale due to scale dependence of the roughness parameters. As a result, the two-body deformation component of the coefficient of friction increases with decreasing scale. The three-body component of the coefficient of friction depends on the concentrations of particles trapped at the interface, which decrease with decreasing scale.

The transition index was proposed, which is responsible for transition from predominantly elastic adhesional friction to plastic deformation. It was found to change with scale, due to scale-dependence of roughness parameters. For the transition index close to zero, the contact is predominantly elastic, and the dominant contribution to friction is adhesion involving elastic deformation. An increase of the transition index leads to an increase in plastic deformation with increasing contribution of the deformation component of friction, resulting in a larger value of the total coefficient of friction.

In the presence of a meniscus force, the measured value of the coefficient of friction is greater than the value of the coefficient of dry friction. The difference is important especially for small loads when the normal load is comparable with the meniscus force. The meniscus force depends on peak radii, and may either increase or decrease with scale depending on the surface parameters.

The wear coefficient and the ratio of the maximum flash temperature rise to the amount of

heat generated per unit time per unit area for a given sliding velocity, as a function of scale, decrease with decreasing scale. This is due to a decrease in the mean contact size.

The proposed model is used to explain the trends in the experimental data for various materials at nanoscale and microscale, which indicate that nanoscale values of the coefficient of friction are lower than the microscale values (Tables 11.2 and 11.3). The two factors responsible for this trend are the increase of the three-body deformation and transition from elastic adhesive contact to plastic deformation. Experimental data show that the coefficient of friction increases with increasing load after a certain load and reaches the macroscale value. This is due to the onset of plastic deformation with increasing load and the effect of load on contact parameters, which affect the coefficient of friction.

Appendix A: Statistics of Particle Size Distribution

A.1 Statistical Models of Particle Size Distribution

Particle size analysis is an important field for different areas of engineering, environmental, and biomedical studies. In general, size distribution of particles depends on how the particles were formed and sorted. Several statistical distributions that govern distribution of random variables including particle size have been suggested (Fig. 11.25) (Bernhardt 1994; Devoro 1995; Everitt 1998; Zwillinger and Kokoska 2000; Wolfram 2003; Bhushan and Nosonovsky 2004a). Statistical distributions commonly used are either the probability density (or frequency) function (PDF), $p(z)$, or cumulative distribution function (CDF), $P(h)$. $P(h)$ associated with random variable $z(x)$, which can take any value between $-\infty$ and $+\infty$ or z_{min} and z_{max} , is defined as the probability of the event $z(x) \leq z'$ and is written as (Bendet and Piersol 1986)

$$P(z) = Prob(z \leq z') \quad (11.76)$$

With $P(-\infty) = 0$ and $P(\infty) = 1$.

The PDF is the slope of the CDF given by its derivative

$$p(z) = \frac{dP(z)}{dz} \quad (11.77)$$

or

$$P(z \leq z') = P(z') = \int_{-\infty}^{z'} p(z) dz \quad (11.78)$$

Furthermore, the total area under the PDF must be unity; that is, it is certain that the value of z at any x must fall somewhere between plus and minus infinity or z_{min} and z_{max} .

The PDF, $p(d)$, in the exponential form is the simplest distribution mathematically

$$p(d) = \frac{1}{\sigma_e} \exp\left(-\frac{d - d_e}{\sigma_e}\right) \quad d \geq d_0 \quad (11.79)$$

where d is particle diameter, σ_e is standard deviation, and d_e is minimum value (for this distribution). For convenience, the density function can be normalized by σ_e in terms of a normalized variable, d^* , equal to $(d - d_e)/\sigma_e$

$$p(d^*) = \exp(-d^*), \quad d^* \geq 0 \quad (11.80)$$

which has zero minimum and unity standard deviation. The CDF, $P(d')$, is given as

$$P(d') = P(d^* \leq d') = 1 - \exp(-d') \quad (11.81)$$

The Gaussian or normal distribution is used to represent data for a wide collection of random physical phenomena in practice such as surface roughness. The PDF and CDF are given as

$$p(d) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{(d - d_n)^2}{2\sigma_n^2}\right), \\ -\infty < d < \infty, -\infty < d_n < \infty, \sigma_n > 0 \quad (11.82)$$

where d_n is the mean value. The integral of $p(d)$ in the interval $-\infty < d < \infty$ is equal to 1. In terms of the normalized variables, Eq. 11.81 reduces to

$$p(d^*) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{d^{*2}}{2}\right) \quad (11.83)$$

and

$$\begin{aligned} P(d') &= P(d^* \leq d') \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d'} \exp\left(-(d^*)^2/2\right) dd^* = \text{erf}(d') \end{aligned} \quad (11.84)$$

where $\text{erf}(d')$ is called the “error function” and its values are listed in most statistical handbooks. The PDF is bell-shaped and the CDF is S-shaped.

For particle size distribution, of interest here, the diameter cannot be less than zero. For this condition, Eq. 11.82 must be modified by using a constant on the right side

$$p(d) = \frac{C_0}{\sqrt{2\pi}\sigma_e} \exp\left(-\frac{(d-d_n)^2}{2\sigma_e^2}\right), \quad (11.85)$$

$$0 \leq d < \infty$$

$$\text{where } C_0 = \left[\frac{1}{\sqrt{2\pi}} \int_{-d_0/\sigma}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt \right]^{-1}$$

The constant is calculated by integrating $p(d)$ in the interval $0 \leq d \leq \infty$ and equating to one

$$\int_0^{\infty} p(d) dd = 1 \quad (11.86)$$

The log-normal distribution is commonly used to describe particle size distribution. A variable d is log-normally distributed if $\ln d$ is normally distributed. Log-normal PDF for variable d , for which $\ln(d)$ has a Gaussian distribution with a mean $\ln(d_{\ln})$ and standard deviation σ_{\ln} , is given as

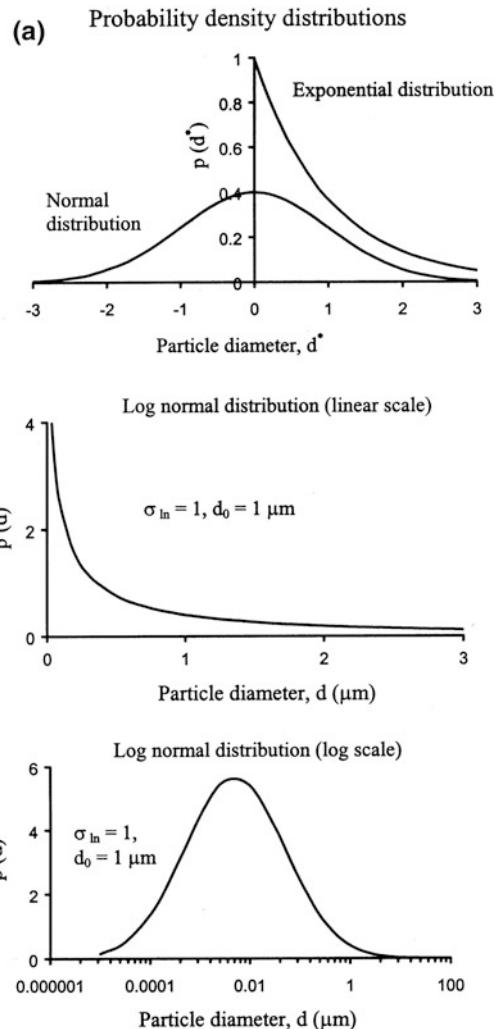


Fig. 11.25 Common statistical distributions of particle size (Bhushan and Nosonovsky 2004a)

$$p(d) = \frac{1}{\sqrt{2\pi}\sigma_{\ln}} \left(\frac{1}{d} \right) \exp\left(-\frac{(\ln(d/d_{\ln}))^2}{2\sigma_{\ln}^2}\right), \quad 0 < d < \infty \quad (11.87)$$

The mean of the log-normal distribution is $\exp(\ln d_{\ln} + \sigma_{\ln}^2/2)$, the standard deviation is $\exp(2\ln d_{\ln} + \sigma_{\ln}^2)(\exp(\sigma_{\ln}^2) - 1)$, the skewness is $(\exp(\sigma_{\ln}^2) + 2)(\exp(\sigma_{\ln}^2) - 1)^{1/2}$, and kurtosis is $\exp(4(\sigma_{\ln}^2)) + 2\exp(3(\sigma_{\ln}^2)) + 3\exp(2(\sigma_{\ln}^2)) - 3$

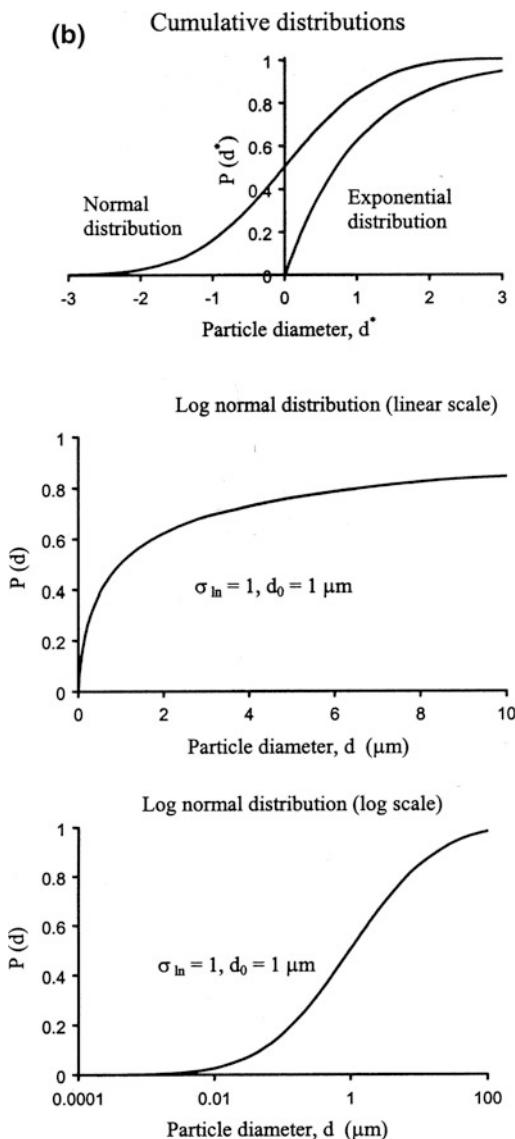


Fig. 11.25 (continued)

(Everitt 1998). The case where $d_{\ln} = 0$ is called the standard log-normal distribution. The density function can be normalized by σ_{\ln} in terms of a normalized variable, $d^* = (\ln d - d_{\ln})/\sigma_{\ln}$

$$p(d^*) = \frac{1}{\sqrt{2\pi}} \left(\frac{1}{d^*} \right) \exp \left(-\frac{(d^*)^2}{2} \right) \quad (11.88)$$

and

$$P(d') = P(d^* \leq d') = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{d'}{\sqrt{2}} \right) \right] \quad (11.89)$$

The log-normal distribution of particle size occurs when the dispersion is attained by comminution (e.g., milling, grinding, or crushing). The size distribution of pulverized silica, granite, calcite, limestone, quartz, soda, ash, alumina, clay, as well as of wear particles, is often governed by the log-normal distribution (Herdan 1960). A size distribution is usually presented either as probability density or frequency $p(d)$, or as cumulative percent (percent of particles greater than given size) $P(d)$, or as cumulative mass versus particle size. All these presentations are interrelated (Herdan 1960).

A.2 Typical Particle Size Distribution Data

Typical experimental data for size distributions of atmospheric particles (dust), sand, and abrasive diamond particles are presented in Fig. 11.26a (Bhushan and Nosonovsky 2004a). It can be seen, that the atmospheric particles (Cadle 1965) follow the normal distribution function. The dune sand is low in heavy mineral content, so the curve is concave downward. Micaceous dune sand is sorted by gravity slide on a sharp mountain slope and appears to follow log-normal distribution, as many distributions of sediments, which are sorted by gravity. Whereas beach sand distribution curve is concave upward due to richness in smaller size component (Herdan 1960). The abrasive diamond particles follow log-normal distribution (Xie and Bhushan 1996).

Size distribution of wear particles has been studied actively since 1970s, when ferrography was introduced (Seifert and Westcott 1972; Scott and Westcott 1977). The data for wear particles is presented in Fig. 11.26b (Bhushan and Nosonovsky 2004a). Xuan et al. (1990) studied the size distribution of submicrometer particles

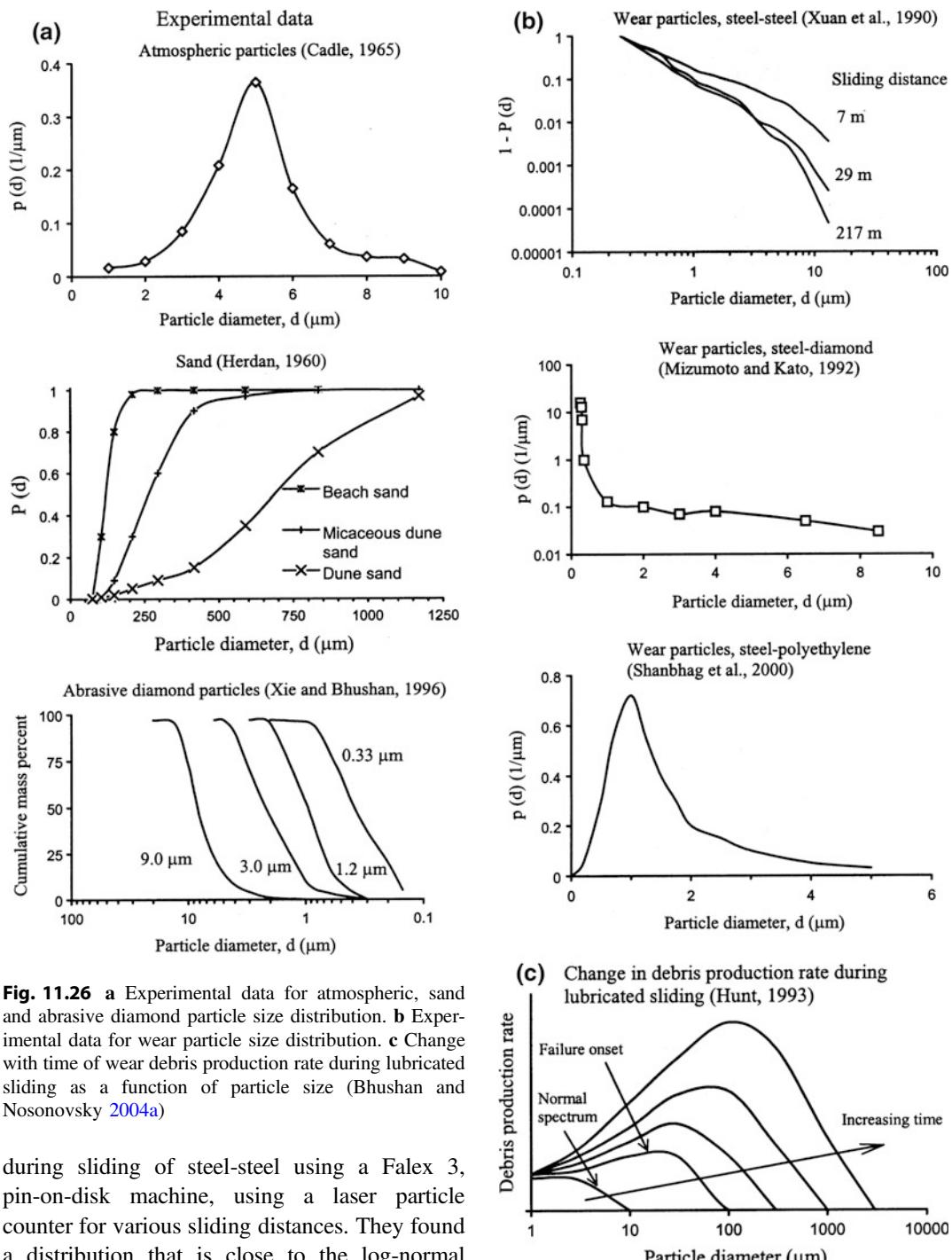


Fig. 11.26 **a** Experimental data for atmospheric, sand and abrasive diamond particle size distribution. **b** Experimental data for wear particle size distribution. **c** Change with time of wear debris production rate during lubricated sliding as a function of particle size (Bhushan and Nosonovsky 2004a)

during sliding of steel-steel using a Falex 3, pin-on-disk machine, using a laser particle counter for various sliding distances. They found a distribution that is close to the log-normal function. Mizumoto and Kato (1992) studied size distribution of particles generated during

Fig. 11.26 (continued)

pin-on-disk test, for diamond, sapphire, silicon carbide, and tungsten carbide pins versus steel disk, using a laser particle counter. They found that the PDF is exponential for particles greater than 1 μm diameter. However, for smaller particles a linear law was assumed. Shanbhag et al. (2000) studied wear particles for ultrahigh molecular weight polyethylene (UHMWPE) versus titanium in biomedical applications (total knee replacement) using a scanning electron microscope. They found that the distribution is close to that of the normal distribution. Numerous data for wear particles are presented by Anderson (1991). Hunt (1993) discusses various techniques of debris measurement and analysis in lubricants. A typical change in wear debris generation rate, which occurs with time, is presented in Fig. 11.26c (Bhushan and Nosonovsky 2004a). Change in the size distribution of wear particles in lubricant indicates an onset of mechanical failure.

References

- Anderson DP (1991) Wear particle atlas, 2nd edn. Spectro Inc. Industrial Tribology Systems, Littleton, Mass
- Anonymous (1959) The industrial graphite engineering handbook. National Carbon Company, New York
- Anonymous (2002) Properties of silicon. EMIS data reviews series No. 4. INSPEC, Institution of Electrical Engineers, London. See also, Anonymous, MEMS Materials Database. <http://www.memsnet.org/material/>
- Bazant ZP (2002) Scaling of dislocation-based strain-gradient plasticity. *J Mech Phys Solids* 50 (2002):435–448
- Bendet JS, Piersol AG (1986) Engineering applications of correlation and spectral analysis, 2nd edn. Wiley, New York
- Bernhardt C (1994) Particle Size analysis. Chapman & Hall, London
- Bhushan B (1996) Contact mechanics of rough surfaces in tribology: single asperity contact. *Appl Mech Rev* 49 (1996):275–298
- Bhushan B (1998) Contact mechanics of rough surfaces in tribology: multiple asperities contact. *Tribol Lett* 4:1–35
- Bhushan B (1999a) Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton, Florida
- Bhushan B (1999b) Chemical, mechanical and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments. *Diamond Relat Mater* 8:1985–2015
- Bhushan B (1999c) Nanoscale tribophysics and tribomechanics. *Wear* 225–229:465–492
- Bhushan B (2013) Introduction to tribology, 2nd edn. Wiley, NY
- Bhushan B (2017) Springer handbook of nanotechnology, 4th edn. Springer International, Switzerland
- Bhushan B, Dandavate C (2000) Thin-film friction and adhesion studies using atomic force microscopy. *J Appl Phys* 87:1201–1210
- Bhushan B, Gupta BK (1997) Handbook of tribology: materials, coatings, and surface treatments. McGraw-Hill, New York, 1991; Krieger, Malabar, Florida, 1997
- Bhushan B, Koinkar AV (1994) Nanoindentation hardness measurements using atomic force microscopy. *Appl Phys Lett* 64:1653–1655
- Bhushan B, Kulkarni AV (1996) Effect of normal load on microscale friction measurements. *Thin Solid Films* 278:49–56; 293:333
- Bhushan B, Nosonovsky M (2003) Scale effects in friction using strain gradient plasticity and dislocation-assisted sliding (microslip). *Acta Mater* 51:4331–4345
- Bhushan B, Nosonovsky M (2004a) Comprehensive model for scale effects in friction due to adhesion and two- and three-body deformation (plowing). *Acta Mater* 52:2461–2474
- Bhushan B, Nosonovsky M (2004b) Scale effects in dry and wet friction, wear, and interface temperature. *Nanotechnology* 15:749–761
- Bhushan B, Peng W (2002) Contact modeling of multilayered rough surfaces. *Appl Mech Rev* 55:435–480
- Bhushan B, Sundararajan S (1998) Micro/nanoscale friction and wear mechanisms of thin films using atomic force and friction force microscopy. *Acta Mater* 46:3793–3804
- Bhushan B, Venkatesan S (1993) Mechanical and tribological properties of silicon for micromechanical applications: a review. *Adv Info Storage Syst* 5:211–239
- Bhushan B, Israelachvili JN, Landman U (1995) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996) Nano/picoindentation measurement using a capacitive transducer system in atomic force microscopy. *Philos Mag* 74:1117–1128
- Bhushan B, Liu H, Hsu SM (2004) Adhesion and friction studies of silicon and hydrophobic and low friction films and investigation of scale effects. *ASME J Trib* 126:583–590
- Cadle RD (1965) Particle size—theory and industrial applications. Reinhold Publishing, NY
- Carpick RW, Agrait N, Ogletree DF, Salmeron M (1996) Measurement of interfacial shear (friction) with an ultrahigh vacuum atomic force microscope. *J Vac Sci Technol B* 14:1289–1295

- Clauser HR (ed) (1963) The encyclopedia of engineering materials and processes. Reinhold Publishing Corp, London
- Deshpande VS, Needleman A, Van der Giessen E (2003) Discrete dislocation plasticity modeling of short cracks in single crystals. *Acta Mater* 51:1–15
- Devoro JL (1995) Probability and statistics for engineering and the sciences. Duxbury Press, New York
- Everitt BS (1998) The Cambridge dictionary of statistics. Cambridge University Press, Cambridge, UK
- Field JE (ed) (1992) The properties of natural and synthetic diamond. Academic Press, London
- Fleck NA, Muller GM, Ashby MF, Hutchinson JW (1994) Strain Gradient plasticity: theory and experiment. *Acta Metall Mater* 42:475–487
- Friedel J (1964) Dislocations. Pergamon Press, New York, Oxford
- Gane N, Cox JM (1970) The micro-hardness of metals at very low loads. *Philos Mag* 22:881–891
- Gao H, Huang Y, Nix WD, Hutchinson JW (1999) Mechanism-based strain-gradient plasticity—I. Theory. *J Mech Phys Solids* 47:1239–1263
- Greenwood JA, Williamson JBP (1966) Contact of nominally flat surfaces. *Proc Royal Soc Lond A* 295:300–319
- Herdan G (1960) Small particle statistics. Butterworth, London
- Homola AW, Israelachvili JN, McGuiggan PM, Gee ML (1990) Fundamental experimental studies in tribology: the transition from “interfacial” friction of undamaged molecularly smooth surfaces to normal friction with wear. *Wear* 136:65–83
- Huang Y, Gao H, Nix WD, Hutchinson JW (2000) Mechanism-based Strain-gradient plasticity—II. Analysis. *J Mech Phys Solids* 48:99–128
- Hunt TM (1993) Handbook of wear debris analysis and particle detection in liquids. Elsevier Applied Science, London
- Hutchinson JW (2000) Plasticity at the micron scale. *Inter J Solids Struc* 37:225–238
- Johnson KL (1985) Contact mechanics. Clarendon Press, Oxford, UK
- Johnson KL (1997) Adhesion and friction between a smooth elastic spherical asperity and a plane surface. *Proc Royal Soc Lond A* 453:163–179
- Koinkar VN, Bhushan B (1997) Scanning and transmission electron microscopies of single-crystal silicon microworn/machined using atomic force microscopy. *J Mater Res* 12:3219–3224
- Kulkarni AV, Bhushan B (1996) Nanoscale mechanical property measurements using modified atomic force microscopy. *Thin Solid Films* 290–291:206–210
- Liu H, Bhushan B (2003) Adhesion and friction studies of microelectromechanical systems/nanoelectromechanical systems materials using a novel microtriboapparatus. *J Vac Sci Technol, A* 21:1528–1538
- Majumdar A, Bhushan B (1991) Fractal model of elastic-plastic contact between rough surfaces. *ASME J Tribol* 113:1–11
- McElhaney KW, Vlassak JJ, Nix WD (1998) Determination of indenter tip geometry and indentation contact area of depth-sensing indentation experiments. *J Mater Res* 13:1300–1306
- Mizumoto M, Kato K (1992) Size distribution and number of wear particles generated by the abrasive sliding of a model asperity in the SEM-tribosystem. In: Dowson D, Taylor CM, Childs THC, Godet M, Dalmau G (eds) Wear particles: from the cradle to the grave. Elsevier, Amsterdam, pp 523–530
- Nix WD, Gao H (1998) Indentation size effects in crystalline materials: a law for strain gradient plasticity. *J Mech Phys Solids* 46:411–425
- Nosonovsky M, Bhushan B (2005) Scale effects in dry friction during multiple-asperity contact. *ASME J Tribol* 127:37–46
- Onions RA, Archard JF (1973) The contact of surfaces having a random structure. *J Phys D Appl Phys* 6:289–304
- Polonsky IA, Keer LM (1996) Scale effects of elastic-plastic behavior of microscopic asperity contact. *ASME J Tribol* 118:335–340
- Rabinowicz E (1995) Friction and wear of materials, 2nd edn. Wiley, NY
- Ruan J, Bhushan B (1994) Atomic-scale friction measurements using friction force microscopy: Part I—general principles and new measurement technique. *ASME J Tribol* 116:378–388
- Schwarz UD, Zwörner O, Köster P, Wiesendanger R (1997) Quantitative analysis of the frictional properties of solid materials at low loads. 1. Carbon compounds. *Phys Rev B* 56:6987–6996
- Scott D, Westcott VC (1977) Predictive maintenance by ferrography. *Wear* 44:173–182
- Seifert WW, Westcott VC (1972) A method for the study of wear particles in lubricating oil. *Wear* 21:27–42
- Shanbhag AS, Bailey HO, Hwang DS, Cha CW, Errol NG, Rubash HE (2000) Quantitative analysis of ultrahigh molecular weight polyethylene (UHMWPE) wear debris associated with total knee replacements. *J Biomed Mater Res* 53:100–110
- Stelmashenko MA, Walls MG, Brown LM, Miman YV (1993) Microindentation on W and Mo oriented single crystal An SEM Study. *Acta Met Mater* 41:2855–2865
- Sundararajan S, Bhushan B (2002) Development of AFM-based techniques to measure mechanical properties of nanoscale structures. *Sens Actuators A* 101:338–351
- Weertman JJ (1963) Dislocations moving uniformly on the interface between isotropic media of different elastic properties. *J Mech Phys Solids* 11:197–204
- Weertman J, Weertman JR (1966) Elementary dislocations theory. The MacMillan Company, NY
- Whitehouse DJ, Archard JF (1970) The properties of random surfaces of significance in their contact. *Proc Roy Soc Lond A* 316:97–121
- Wolfram S (2003) The mathematica book, 5th edn. Wolfram Media, Champaign, IL. Also see <http://mathworld.wolfram.com>

- Xie Y, Bhushan B (1996) Effect of particle size, polishing pad and contact pressure in free abrasive polishing. *Wear* 200:281–295
- Xuan JL, Cheng HS, Miller RJ (1990) Generation of submicrometer particles in dry sliding. *ASME J Tribol* 112:664–691
- Zhao X, Bhushan B (1998) Material removal mechanisms of single-crystal silicon on nanoscale and at ultralow loads. *Wear* 223:66–78
- Zwillinger D, Kokoska S (2000) CRC standard probability and statistics tables and formulas. CRC Press, Boca Raton, Florida

Part IV

Molecularly-Thick Films for Lubrication

Bharat Bhushan

Abstract

One of the best materials to use in applications that require very low wear and reduced friction is diamond, especially in the form of a diamond coating. Unfortunately, true diamond coatings can be deposited only at high temperatures and on selected substrates, and they require surface finishing. However, the mechanical, thermal, and optical properties of hard, amorphous carbon—commonly known as diamondlike carbon or a DLC coating—can be tailored to be similar to those of diamond. It can also be deposited at a wide range of thicknesses using a variety of deposition processes on various substrates at or near room temperature. The coatings reproduce the topography of the substrate, removing the need for surface finishing. Various physical and chemical characterization techniques have been used to study the structure, stoichiometry, and mechanical and tribological properties. The friction and wear properties of some DLC coatings make them very attractive for some tribological applications. The most significant current industrial applications of DLC coatings are in magnetic storage devices, micro/nanoelectromechanical systems, and various other consumer products including razor blades, sunglasses, and optical windows. In this chapter, the state-of-the-art in the chemical, mechanical, and tribological characterization of ultrathin amorphous carbon coatings as thin as 5 nm is presented.

Keywords

Amorphous carbon • Diamondlike carbon • Films • Friction • Wear • Lubrication • Nanotribology

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

12.1 Introduction

Carbon exists in both crystalline and amorphous forms and exhibits both metallic and nonmetallic characteristics (Bhushan and Gupta 1991; Bhushan 2013a, b). Crystalline carbon includes graphite, diamond and a family of fullerenes, Fig. 12.1. The graphite and diamond are infinite periodic network solids with a planar structure, whereas the fullerenes are a molecular form of

pure carbon with a finite network with a non-planar structure. Graphite has a hexagonal, layered structure with weak interlayer bonding forces, and exhibits excellent lubrication properties. The graphite crystal may be visualized as infinite parallel layers of hexagons stacked 0.34 nm apart with a 0.1415-nm interatomic distance between the carbon atoms in the basal plane. The atoms lying in the basal planes are trigonally coordinated and closely packed with

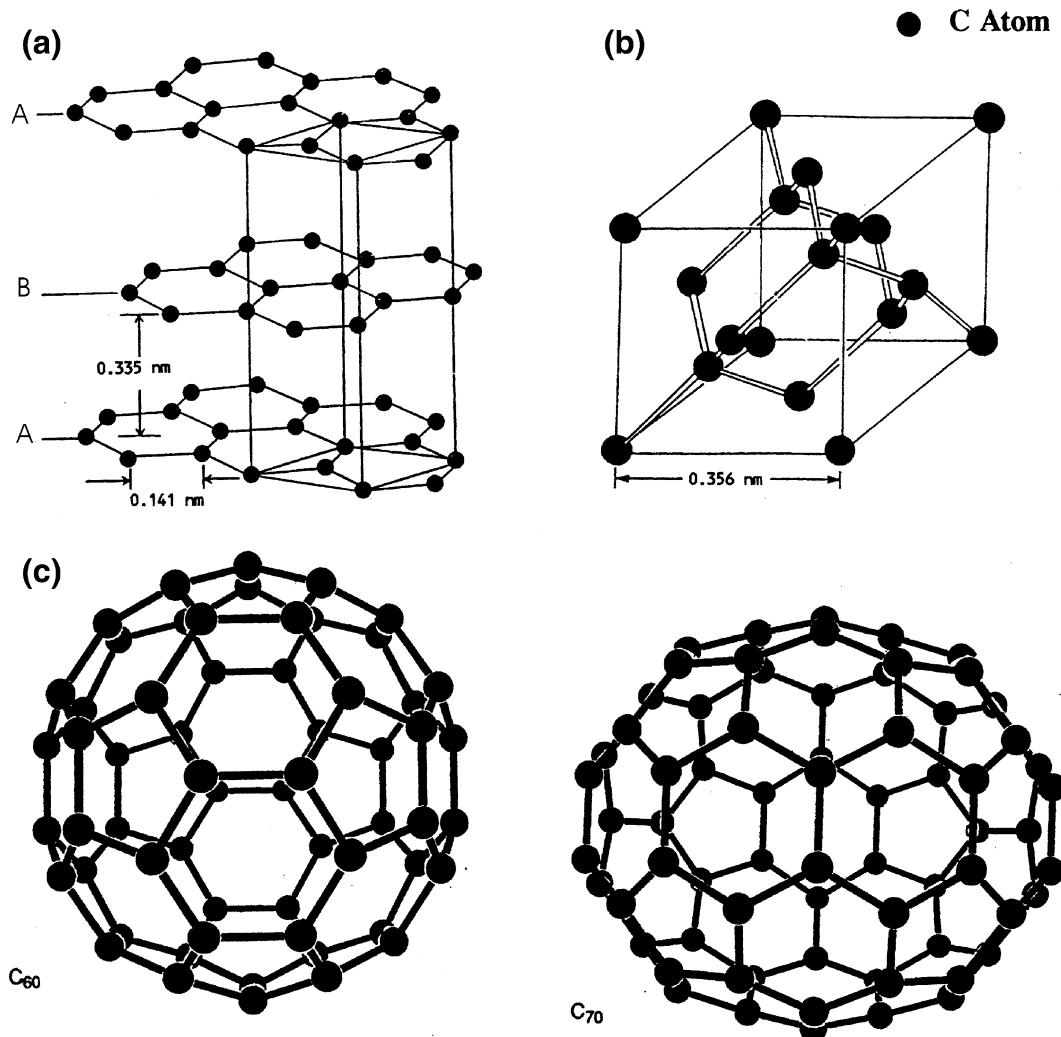


Fig. 12.1 The structure of three known forms of crystalline carbon **a** hexagonal structure of graphite, **b** modified face-centered cubic (fcc) structure, two interpenetrating fcc lattices displaced by one one-quarter of the cube diagonal, of diamond (each atom is bonded to four others that form the corners of the pyramidal structure called tetrahedron), and **c** the structure of two most common forms of fullerenes-soccer ball C₆₀ and rugby ball C₇₀ molecules

strong σ (covalent) bonds to its three carbon neighbors using the hybrid sp^2 orbitals. The fourth electron lies in a p_z orbital lying normal to the σ bonding plane. It forms a weak π bond by overlapping side to side with a p_z orbital of an adjacent atom to which carbon is attached by a σ bond. The layers (basal planes) themselves are relatively far apart and the forces that bond them are weak van der Waals forces. These layers can align themselves parallel to the direction of the relative motion, and slide over one another with relative ease, providing low friction. Strong interatomic bonding and packing in each layer is thought to help reduce wear. Operating environment has a significant influence on the lubrication properties, i.e. low friction and low wear, of graphite. It lubricates better in a humid environment than a dry one which results from adsorption of water vapor and other gases from the environment. This further weakens the interlayer bonding forces, resulting in easy shear and transfer of the crystallite platelets to the mating surface. Thus, transfer plays an important role in controlling friction and wear. Graphite oxidizes at high operating temperatures, and can be used up to about 430 °C.

One of the fullerene molecules is C_{60} , commonly known as a Buckyball. Since the C_{60} molecules are very stable and do not require additional atoms to satisfy chemical bonding requirements, they are expected to have low adhesion to the mating surface and low surface energy. Since C_{60} molecules with a perfect spherical symmetry are weakly bonded to other molecules, C_{60} clusters get detached readily, similar to other layered-lattice structures. They either transfer to the mating surface by mechanical compaction, or are present as a loose wear particle that may roll like a tiny ball bearings in a sliding contact, resulting in low friction and wear. The wear particles are expected to be harder than as-deposited C_{60} molecules because of their phase transformation at high asperity contact pressures present in a sliding interface. The low surface energy, spherical shape of C_{60} molecules, weak intermolecular bonding, and high load bearing capacity offer potential for various mechanical and tribological applications.

The sublimed C_{60} coatings and fullerene particles as an additive to mineral oils and greases, have been reported to be good solid lubricants comparable to graphite and MoS_2 (Bhushan et al. 1993a; Gupta et al. 1994; Gupta and Bhushan 1994).

Diamond crystallizes in a modified face centered cubic (fcc) structure with an interatomic distance of 0.154 nm. The diamond cubic lattice consists of two interpenetrating fcc lattices displaced by one-quarter of the cube diagonal. Each carbon atom is tetrahedrally coordinated, making strong σ (covalent) bonds to its four carbon neighbors using the hybrid sp^3 atomic orbitals. This accounts for its hardness (80–104 GPa) and thermal conductivity (900–2100 W/m K, on the order of five times that of copper), the highest of any known solid, and a high electrical resistivity, optical transmission and large optical bandgap. It is relatively chemically inert, and it exhibits poor adhesion with other solids with consequent low friction and wear. Its high thermal conductivity allows dissipation of frictional heat during sliding and protects the interface. The dangling carbon bonds on the surface react with the environment to form hydrocarbons, which act as good lubrication films. These are some of the reasons for low friction and wear of diamond. Diamond and its coatings find many industrial applications: tribological applications (low friction and wear), optical applications (exceptional optical transmission, high abrasion resistance), and thermal management or heat sink applications (high thermal conductivity). Diamond can be used to high temperatures and it starts to become graphite at about 1000 °C in ambient air and at about 1400 °C in vacuum. Diamond is an attractive material for cutting tools, as an abrasive for grinding wheels and lapping compounds, and other extreme wear applications.

Natural diamond, particularly in large sizes, is very expensive and its coatings, a low cost alternative, are attractive. True diamond coatings are deposited by chemical vapor deposition (CVD) processes at high substrate temperatures (on the order of 800 °C). They adhere best on a silicon substrate and require an interlayer for other substrates. A major roadblock to the

widespread use of true diamond films in tribological, optical, and thermal management applications is surface roughness. Growth of the diamond phase on a nondiamond substrate is initiated by nucleation either at randomly seeded sites or at thermally favored sites due to statistical thermal fluctuation at the substrate surface. Based on growth temperature and pressure conditions, favored crystal orientations dominate the competitive growth process. As a result, the grown films are polycrystalline in nature with relatively large grain size ($>1\text{ }\mu\text{m}$), and terminate in very rough surfaces with RMS roughnesses ranging from few tenth of a micron to tens of microns. Techniques for polishing these films have been developed. It has been reported that the laser polished films exhibit friction and wear properties almost comparable to that of bulk polished diamond (Bhushan et al. 1993b, 1994).

Amorphous carbon has no long range order and the short range order of carbon atoms can have one or more of three bonding configurations — sp^3 (diamond), sp^2 (graphite), or sp^1 (with two electrons forming strong σ bonds and the remaining two electrons left in orthogonal p_y and p_z orbitals to form weak π bonds). Short range order controls the properties of amorphous materials and coatings. Hard amorphous carbon (a-C) coatings are commonly known as diamondlike carbon or DLC (implying high hardness) coatings. They are a class of coatings that are mostly metastable amorphous materials, but include a micro- or nanocrystalline phase. The coatings are random networks of covalently-bonded carbon in hybridized tetragonal (sp^3) and trigonal (sp^2) local coordination, with some of the bonds terminated by hydrogen. These coatings have been deposited successfully by a variety of vacuum deposition techniques on variety of substrates at or near room temperature. These coatings generally reproduce substrate topography and do not require any post finishing. However, these coatings mostly adhere best on silicon substrates. Best adhesion is obtained on substrates that form carbides, e.g. Si, Fe and Ti. Based on depth profile analyses using Auger and XPS of DLC coatings deposited on silicon substrates, it has been reported that a substantial

amount of silicon carbide (on the order of 5–10 nm in thickness) is present at the carbon-silicon interface for the coatings with good adhesion and high hardness (e.g., Sander et al. 1987). For good adhesion of DLC coatings to other substrates, in most cases, an interlayer of silicon is required, except for cathodic arc deposited coatings.

There is significant interest in DLC coatings because of their unique combination of desirable properties. These properties include high hardness and wear resistance, chemical inertness to both acids and alkalis, lack of magnetic response, and an optical band gap ranging from zero to a few eV, depending upon the deposition conditions. These are used in a wide range of applications including tribological, optical, electronic, and biomedical applications (Bhushan and Gupta 1991; Matthews and Eskildsen 1994; Lettington 1998). The high hardness, good friction and wear properties, versatility in deposition and substrates and no requirements of post finishing make them very attractive for tribological applications. The largest industrial application of the family of amorphous carbon coatings are typically deposited by DC/RF magnetron sputtering, plasma-enhanced chemical vapor deposition, or ion beam deposition techniques in magnetic storage devices (Gupta and Bhushan 1995a; Bhushan 1996a, 2000, 2001a; Bhushan et al. 1996; Xu and Bhushan 1998; Scott and Bhushan 2000a, b; Scott et al. 2000). These coatings are employed to protect against wear and corrosion, magnetic coatings on thin-film rigid disks and metal evaporated tapes, and the thin-film head structure of a read/write disk head (Fig. 12.2). To maintain low physical spacing between the magnetic element of a magneto resistive-type (MR) read/write head and the magnetic layer of a media, thicknesses ranging from 3 to 10 nm are employed. Mechanical properties affect friction and wear, and these need to be optimized. In 1998, Gillette introduced Mach 3 razor blades with ultra-thin DLC coatings, which is a large industrial application. DLC coatings are also used in other commercial applications such as cutting tools, glass windows of supermarket laser barcode scanners, other optical windows, and sunglasses. These coatings are also

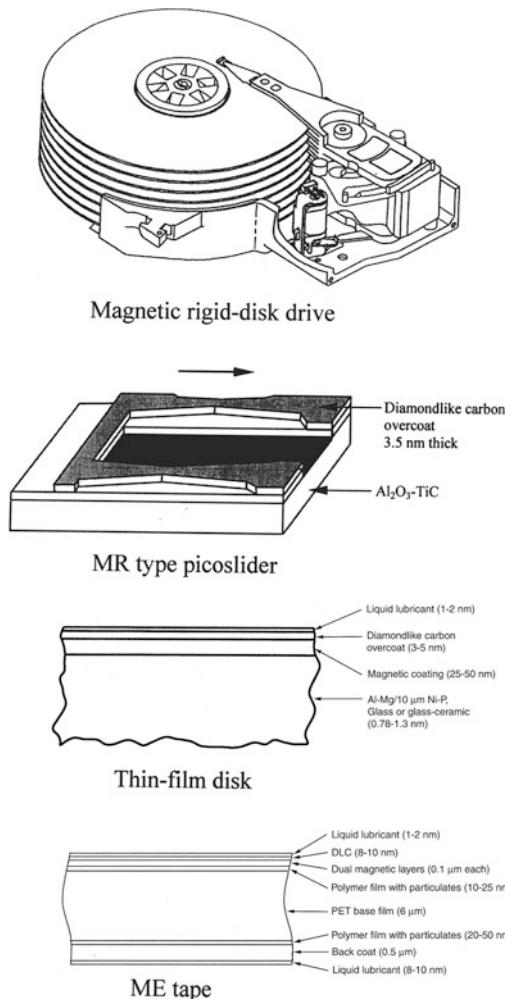


Fig. 12.2 Schematic of a magnetic rigid disk drive and MR type picoslider and cross-sectional schematics of a magnetic thin-film rigid disk and a metal evaporated (ME) tape

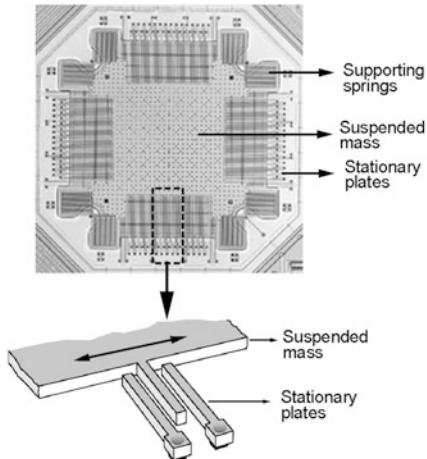
being used in micro/nanoelectromechanical systems (MEMS/NEMS) components, Fig. 12.3 (Bhushan 1996b, 1998, 1999a, 2001b, 2016).

In this chapter, an overview of the chemical, mechanical, and tribological characterization of ultrathin amorphous carbon coatings, commonly referred to as diamondlike carbon, is presented. First, an overview of the most commonly used deposition techniques is presented, followed by typical chemical and mechanical characterization data and typical tribological data both from coupon level testing and functional testing.

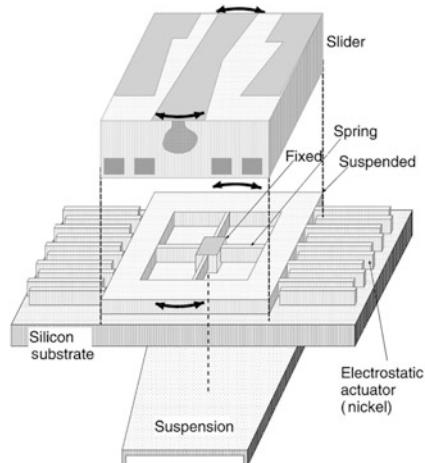
12.2 Description of Commonly Used Deposition Techniques

The first hard amorphous carbon coatings were deposited by a beam of carbon ions produced in an argon plasma on room-temperature substrates as reported by Aisenberg and Chabot (1971). Subsequent confirmation by Spencer et al. (1976) led to explosive growth of this field. Following the first work, several alternative techniques have been developed. The amorphous carbon coatings have been prepared by a variety of deposition techniques and precursors, including evaporation, DC, RF or ion beam sputtering, RF or DC plasma-enhanced chemical vapor deposition (PECVD), electron cyclotron resonance chemical vapor deposition (ECR-CVD), direct ion beam deposition, pulsed laser vaporization and vacuum arc, from a variety of carbon-bearing solids or gaseous source materials (Bhushan and Gupta 1991; Grill and Meyerson 1994). Coatings with both graphitic and diamondlike properties have been produced. Evaporation and ion plating techniques have been used to produce coatings with graphitic properties (low hardness, high electrical conductivity, very low friction, etc.) and all techniques have been used to produce coatings with diamondlike properties.

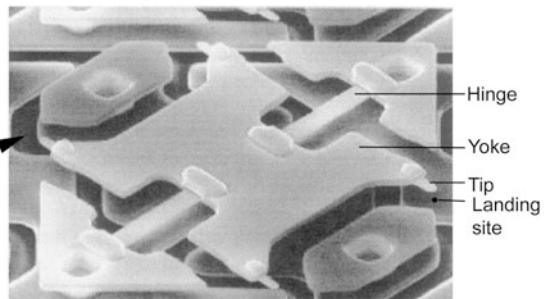
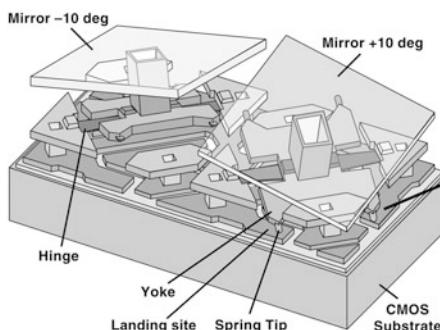
Structure and properties of a coating are dependent upon the deposition technique and the deposition parameters. High-energy surface bombardment has been used to produce harder and denser coatings. It is reported that sp³/sp² fractions are in the decreasing order for cathodic arc deposition, pulsed laser vaporization, direct ion beam deposition, plasma-enhanced chemical vapor deposition, ion beam sputtering, and DC/RF sputtering (Gupta and Bhushan 1995a; Catherine 1991; Cuomo et al. 1991). Common feature to these techniques is that the deposition is energetic, i.e. carbon species arrive with an energy significantly greater than that represented by the substrate temperature. The resultant coatings are amorphous in structure, with hydrogen content up to 50%, and display a high degree of sp³ character. From the results of previous investigations, it has been proposed that deposition of sp³-bonded carbon requires the depositing



Capacitive type silicon accelerometer for automotive sensory applications (Sulouff, 1998)



Electroplated-nickel rotary microactuator for magnetic disk drives (Hirano et al., 2000)



Digital micromirror device for displays (Hornbeck, 1999)

Fig. 12.3 Schematics of a capacitive type silicon accelerometer for automotive sensory applications, digital micrometer devices for high-projection displays, and polysilicon rotary microactuator for magnetic disk drives

species to have kinetic energies on the order of 100 eV or higher, well above those obtained in thermal processes like evaporation (0–0.1 eV). The species must then be quenched into the metastable configuration via rapid energy removal. Excess energy, such as that provided by substrate heating, is detrimental to the achievement of a high sp^3 fraction. In general, a high fraction of the sp^3 -bonded carbon atoms in an amorphous network results in a higher hardness (Angus and Hayman 1988; Angus and Jensen 1988; Green et al. 1989; Cho et al. 1990; Cuomo et al. 1991; Angus 1992; Bhushan et al. 1992; Robertson 1992). The mechanical and

tribological properties of a carbon coating depend on the sp^3/sp^2 -bonded carbon ratio, the amount of hydrogen in the coating, and adhesion of the coating to the substrate, which are influenced by the precursor material, kinetic energy of the carbon species prior to deposition, deposition rate, substrate temperature, substrate biasing and the substrate itself (Savvides and Window 1985; Angus et al. 1986; Robertson 1986; Cho et al. 1990; Rubin et al. 1990; Vandentop et al. 1990; Cuomo et al. 1991, 1992; Bhushan et al. 1992; Pappas et al. 1992; Scheibe and Schultrich 1994; Donnet and Grill 1997; Grill 1997). The kinetic energies and deposition rates involved in selected

Table 12.1 Summary of most commonly used deposition techniques and the kinetic energy of depositing species and deposition rates

Deposition technique	Process	Kinetic energy (eV)	Deposition rate (nm/s)
Filtered cathodic arc (FCA)	Energetic carbon ions produced by a vacuum arc discharge between a graphite cathode and grounded anode	100–2500	0.1–1
Direct ion beam (IB)	Carbon ions produced from methane gas in an ion source and accelerated towards a substrate	50–500	0.1–1
Plasma enhanced chemical vapor deposition (PECVD)	Hydrocarbon species, produced by plasma decomposition of hydrocarbon gases (e.g. acetylene) are accelerated towards a DC-biased substrate	1–30	1–10
Electron cyclotron resonance plasma chemical vapor deposition (ECR-CVD)	Hydrocarbon ions, produced by plasma decomposition of ethylene gas in the presence of a plasma in electron cyclotron resonance condition, are accelerated towards an RF-biased substrate	1–50	1–10
DC/RF sputtering	Sputtering of graphite target by argon ion plasma	1–10	1–10

deposition processes used in the deposition of DLC coatings are compared in Table 12.1 (Catherine 1991; Bhushan and Gupta 1991).

In the studies by Gupta and Bhushan (1995a, b), Li and Bhushan (1999a, b) and Sundararajan and Bhushan (1999), DLC coatings, ranging in thickness typically from 3.5 to 20 nm, were deposited on single-crystal silicon, magnetic Ni-Zn ferrite and Al_2O_3 -TiC substrates (surface roughness \approx 1–3 nm RMS) by filtered cathodic arc (FCA) deposition, (direct) ion beam deposition (IBD), electron cyclotron resonance chemical vapor deposition (ECR-CVD), plasma-enhanced chemical vapor deposition (PECVD), and DC/RF planar, magnetron sputtering (SP) deposition techniques (Bhushan 1999b). In this article, we will limit the presentation of data of coatings deposited by FCA, IBD, ECR-CVD and SP deposition techniques.

12.2.1 Filtered Cathodic Arc Deposition Technique

In the filtered cathodic arc deposition of carbon coating (Cuomo et al. 1991; Aksenov and Strel'Nitskii 1991; McKenzie et al. 1991; Lossy

et al. 1992; Brown et al. 1993; Fallon et al. 1993; Anders et al. 1994a, b, 1995), a vacuum arc plasma source is used to form carbon film. In the FCA technique used by Bhushan and coworkers (e.g., Gupta and Bhushan 1995a), energetic carbon ions are produced by a vacuum arc discharge between a planar graphite cathode and grounded anode, Fig. 12.4a. The cathode is a 6 mm diameter high density graphite disk mounted on a water-cooled copper block. The arc is driven at an arc current of 200 A, arc duration of 5 ms, and arc repetition rate of 1 Hz. The plasma beam is guided by a magnetic field that transports current between the electrodes to form tiny, rapidly moving spots on the cathode surface. The source is coupled to a 90° bent magnetic filter to remove the macroparticles produced concurrently with the plasma in the cathode spots. The ion current density at the substrate is in the range of 10–50 mA/cm². The base pressure is less than 10⁻⁴ Pa. Compared with electron beam evaporation with auxiliary discharge, much higher plasma density is achieved with the aid of powerful arc discharge. In this process, the cathodic material suffers a complicated transition from the solid phase to an expanding, non-equilibrium plasma via liquid and dense, equilibrium non-ideal

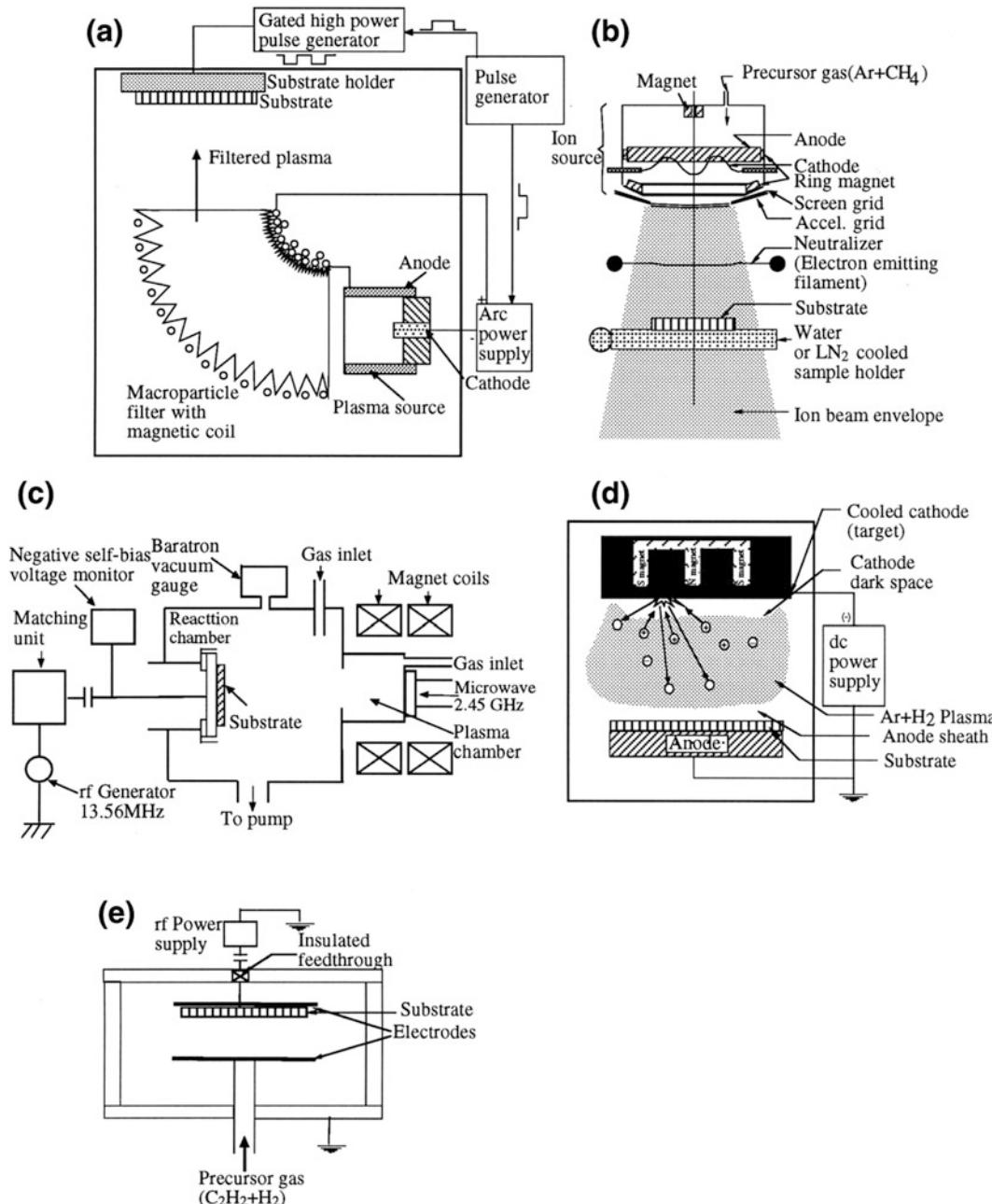


Fig. 12.4 Schematic diagrams of deposition by **a** filtered cathodic arc deposition, **b** ion beam deposition, **c** electron cyclotron resonance chemical vapor deposition (ECR-CVD), **d** DC planar magnetron sputtering, and **e** plasma enhanced chemical vapor deposition (PECVD)

plasma phases (Anders et al. 1994b). The carbon ions in the vacuum arc plasma have a direct kinetic energy of 20–30 eV. The high voltage pulses are applied to the substrate mounted on a

water-cooled sample holder, ions are accelerated through the sheath and arrive at the substrate with an additional energy given by the potential difference between the plasma and the substrate. The

substrate holder is pulse biased to a negative voltage up to -2 kV with a pulse duration of $1\text{ }\mu\text{s}$. The negative biasing of -2 kV corresponds to 2 keV kinetic energy of the carbon ions. The use of a pulse bias instead of a DC bias has advantages of applying a much higher voltage and building a surface potential on a nonconducting film. The energy of the ions is varied during the deposition. For the first 10% of the deposition the substrates are pulsed biased to -2 keV with a pulse duty cycle of 25% , i.e. for 25% of the time the energy is 2 keV, for the remaining 75% time it is 20 eV which is the “natural” energy of carbon ions in a vacuum discharge. For the last 90% of the deposition the pulsed bias voltage is reduced to -200 eV with a pulse bias duty cycle of 25% , i.e. the energy is 200 eV for 25% and 20 eV for 75% of the deposition. The high energy at the beginning leads to a good intermixing and good adhesion of the films whereas the lower energy at later stage leads to hard films. Under the conditions described, the deposition rate at the substrate is about 0.1 nm/s which is slow. As compared with most gaseous plasma, the cathodic arc plasma is nearly fully ionized and the ionized carbon atoms have high kinetic energy of carbon ions help achieving a high fraction of sp^3 -bonded carbon ions which in turn result into a high hardness and higher interfacial adhesion. Cuomo et al. (1992) have reported that based on electron energy loss spectroscopy (EELS) analysis, the sp^3 -bonded carbon fraction of cathodic arc coating is 83% as compared with 38% for the ion beam sputtered carbon. These coatings are reported to be *nonhydrogenated*.

This technique does not require an adhesion underlayer for non-silicon substrates. However, adhesion of the DLC coatings on the electrically-insulating substrate is poor as negative pulse biasing forms an electrical sheath which accelerates depositing ions to the substrate and enhances the adhesion of the coating to the substrate with associated ion implantation. It is

difficult to build potential on an insulating substrate and lack of biasing results in poor adhesion.

12.2.2 Ion Beam Deposition Technique

In the direct ion beam deposition of carbon coating (Swec et al. 1988; Erdemir et al. 1991, 1993; Wei et al. 1993; Erdemir and Donnet 2001), as used by Bhushan and coworkers (e.g., Gupta and Bhushan 1995a) (deposition carried out by Ion Tech. Inc., Fort Collins, Colorado), the carbon coating is deposited from an accelerated carbon ion beam. The sample is precleaned by ion etching. For the case of non-silicon substrates, $2\text{--}3$ nm thick amorphous silicon adhesion layer is deposited by ion beam sputtering using an ion beam of a mixture of methane and argon at 200 V. For the carbon deposition, chamber is pumped to about 10^{-4} Pa and methane gas is fed through the cylindrical ion source and is ionized by energetic electrons produced by a hot-wire filament, Fig. 12.4b. Ionized species then pass through a grid with a bias voltage of about 50 eV where they gain a high acceleration energy and reach a hot-wire filament, emitting thermionic electrons that neutralize the incoming ions. Discharging of ions is important when insulating ceramics are used as substrates. The species are then deposited on a water cooled substrate. Operating conditions are adjusted to give an ion beam with an acceleration energy of about 200 eV and a current density of about 1 mA/cm^2 . At this operating condition, the deposition rate is about 0.1 nm/s which is slow. Incidentally, for deposition of tough and soft coatings, these are deposited at high acceleration energy of about 400 eV at a deposition rate of about 1 nm/s . The ion beam deposited carbon coatings are reported to be hydrogenated ($30\text{--}40$ at.% hydrogen).

12.2.3 Electron Cyclotron Resonance Chemical Vapor Deposition Technique

The lack of electrodes and its ability to create high densities of charged and excited species at low pressures ($\leq 10^{-4}$ Torr) in ECR plasma, make it an attractive processing discharge in the coating depositions (Asmussen 1989). In the ECR-CVD deposition process of carbon coating described by Suzuki and Okada (1995) and used by Li and Bhushan (1999a, b) and Sundararajan and Bhushan (1999), microwave power is generated by a magnetron operating in continuous mode at a frequency of 2.45 GHz, Fig. 12.4c. The plasma chamber functions as a microwave cavity resonator. The magnetic coils arranged around the plasma chamber generate a magnetic field of 875 G necessary for electron cyclotron resonance condition. The substrate is placed on a stage which is capacitively connected to a 13.56 MHz RF generator. The process gas is introduced into the plasma chamber and the hydrocarbon ions generated are accelerated by a negative self-bias voltage, which is generated by applying RF power to the substrate. Both the substrate stage and plasma chamber are water-cooled. The process gas used is 100% ethylene and its flow rate is held constant at 100 sccm. The microwave power is 100–900 W. The RF power is 30–120 W. The pressure during deposition is kept close to the optimum value of 5.5×10^{-3} Torr. Before the deposition, the substrates are cleaned using Ar ions generated in the ECR plasma chamber.

12.2.4 Sputtering Deposition Technique

In DC planar magnetron sputtered carbon coating (Banks and Rutledge 1982; Weissmantel et al. 1982; Dimigen and Hubsch 1983–1984; Savvides and Window 1985; Cho et al. 1990; Rubin et al. 1990; Bhushan et al. 1995a, b; Bhushan 1996a), the carbon coating is deposited by the sputtering of graphite target with Ar ion plasma. In the glow discharge, positive ions from the

plasma strike the target with sufficient energy to dislodge the atoms by momentum transfer which are intercepted by the substrate. First an about 5 nm thick amorphous silicon adhesion layer is deposited by sputtering if the deposition is to be carried out on a non-silicon surface. In the process used by Bhushan and coworkers (e.g., Gupta and Bhushan 1995a), the coating is deposited by sputtering of graphite target with Ar ion plasma, at 300 W power for 200 mm diameter target and pressure of about 0.5 Pa (6 mTorr), Fig. 12.4d. Plasma is generated by applying a DC potential between the substrate and a target. Bhushan et al. (1992) reported that sputtered carbon coating contain about 35 at.% hydrogen. Hydrogen comes from the hydrocarbon contaminants present in the deposition chamber. To produce hydrogenated carbon coating with a larger concentration of hydrogen, deposition is carried out in Ar and hydrogen plasma.

12.2.5 Plasma-Enhanced Chemical Vapor Deposition Technique

In the RF-PECVD deposition of carbon coating as used by Bhushan and coworkers (e.g., Gupta and Bhushan 1995a), carbon coating is deposited by adsorption of most free radicals of hydrocarbon to the substrate and chemical bonding to other atoms on the surface. The hydrocarbon species are produced by the RF plasma decomposition of hydrocarbon precursors such as acetylene (C_2H_2), Fig. 12.4e (Holland and Ojha 1976; Andersson 1981; Bubenzer et al. 1983; Dimigen and Hubsch 1983–1984; Grill et al. 1990a; Grill and Meyerson 1994). Instead of requiring thermal energy in thermal CVD, the energetic electrons in the plasma (at pressures ranging from 1 to 5×10^2 Pa, typically less than 10 Pa) can activate almost any reaction among the gases in the glow discharge at relatively low substrate temperatures ranging from 100 to 600 °C (typically less than 300 °C). To deposit the coating on non-silicon substrates, an about 4 nm thick amorphous silicon adhesion layer, used to improve adhesion, is first deposited under similar

conditions from a gas mixture of 1% silane in argon (Grill et al. 1988). In the process used by Bhushan and coworkers (Gupta and Bhushan 1995a), the plasma is sustained in a parallel-plate geometry by a capacitive discharge at 13.56 MHz, at a surface power density on the order of 100 mW/cm². The deposition is performed at a flow rate on the order of 6 sccm and a pressure on the order of 4 Pa (30 mTorr) on a cathode-mounted substrate maintained at a substrate temperature of 180 °C. The cathode bias held fixed at on the order of -120 V with external DC power supply attached to substrate (powered electrode). The carbon coatings deposited by PECVD usually contain hydrogen up to 50% (Grill et al. 1990b; Bhushan et al. 1992).

12.3 Chemical Characterization and Effect of Deposition Conditions on Chemical Characteristics and Physical Properties

The chemical structure and properties of amorphous carbon coatings are a function of deposition conditions. It is important to understand the relationship of the chemical structure of amorphous carbon coatings to the properties in order to define useful deposition parameters. Amorphous carbon films are metastable phases formed when carbon particles are condensed on a substrate. The prevailing atomic arrangement in the DLC coatings is amorphous or quasi-amorphous with small diamond (sp^3), graphite (sp^2) and other unidentifiable micro- or nanocrystallites. The coatings dependent upon the deposition process and its conditions contain varying amounts of sp^3/sp^2 ratio and hydrogen. The sp^3/sp^2 ratio of DLC coatings ranges typically from 50% to close to 100% with an increase in hardness with sp^3/sp^2 ratio. Most DLC coatings except those produced by filtered cathodic arc contain from a few to about 50 at.% hydrogen. Sometimes hydrogen and nitrogen are deliberately added to produce hydrogenated (a-C:H) and nitrogenated amorphous carbon (a-C:N)

coatings, respectively. Hydrogen helps to stabilize sp^3 sites (most of the carbon atoms attached to hydrogen have a tetrahedral structure); therefore, the sp^3/sp^2 ratio for hydrogenated carbon is higher (Angus and Hayman 1988). Optimum sp^3/sp^2 in random covalent network composed of sp^3 and sp^2 carbon sites (N_{sp^2} and N_{sp^3}) and hydrogen is (Angus and Hayman 1988),

$$\frac{N_{sp^3}}{N_{sp^2}} = \frac{6X_H - 1}{8 - 13X_H} \quad (12.1)$$

where X_H is the atom fraction of hydrogen. The hydrogenated carbon, has a larger optical band gap, higher electrical resistivity (semiconductor) and has a lower optical absorption or high optical transmission. The hydrogenated coatings have a lower density probably because of reduction of cross-linking due to hydrogen incorporation. However, hardness decreases with an increase of the hydrogen even though the proportion of sp^3 sites increases (that is, as the local bonding environment becomes more diamondlike) (Jansen et al. 1985; Kaplan et al. 1985). It is speculated that the high hydrogen content introduces frequent terminations in the otherwise strong three-dimensional network and hydrogen increases the soft polymeric component of the structure more than it enhances the cross-linking sp^3 fraction.

A number of investigations have been performed to identify the microstructure of amorphous carbon films using a variety of techniques such as Raman spectroscopy, EELS, nuclear magnetic resonance, optical measurements, transmission electron microscopy, and X-ray photoelectron spectroscopy (Cho et al. 1990). The structure of diamondlike amorphous carbon is amorphous or quasi-amorphous with small graphitic (sp^2) and tetrahedrally coordinated (sp^3) and other unidentifiable nanocrystallites (typically on the order of couple nm, randomly oriented) (Tsai et al. 1988; Marchon et al. 1989; Cho et al. 1990). These studies indicate that the chemical structure and physical properties of the coatings are quite variable, depending on the deposition techniques and film growth conditions. It is clear that both sp^2 and sp^3 -bonded atomic sites are incorporated in diamondlike amorphous carbon coatings and that

the physical and chemical properties of the coatings depend strongly on their chemical bonding and microstructure. Systematic studies have been conducted to carry out chemical characterization and to investigate how the physical and chemical properties of amorphous carbon coatings vary as a function of deposition parameters (e.g., Cho et al. 1990; Rubin et al. 1990; Bhushan et al. 1992). EELS and Raman spectroscopy are commonly used to characterize the chemical bonding and microstructure. Hydrogen concentration of the coatings is obtained by means of forward recoil spectrometry (FRS). A variety of physical properties of the coatings relevant to tribological performance are measured.

To present the typical data obtained for characterization of typical amorphous carbon coatings and their relationships to physical properties, we present data on several sputtered,

one RF-PECVD amorphous carbon and one microwave-PECVD (MPECVD) diamond coatings (Rubin et al. 1990; Bhushan et al. 1992; Cho et al. 1990). The sputtered coatings were DC magnetron sputtered at a chamber pressure of 10 mTorr, under sputtering power densities of 0.1 and 2.1 W/cm^2 in a pure Ar plasma, labeled as W1 and W2, respectively; and prepared at a power density of 2.1 W/cm^2 with various hydrogen fractions of 0.5, 1, 3, 5, 7, and 10% of Ar/H the gas mixtures labeled as H1, H2, H3, H4, H5, and H6, respectively.

12.3.1 EELS and Raman Spectroscopy

EELS and Raman spectra of four sputtered (W1, W2, H1, and H3) and one PECVD carbon samples were obtained. Figure 12.5 shows the EELS

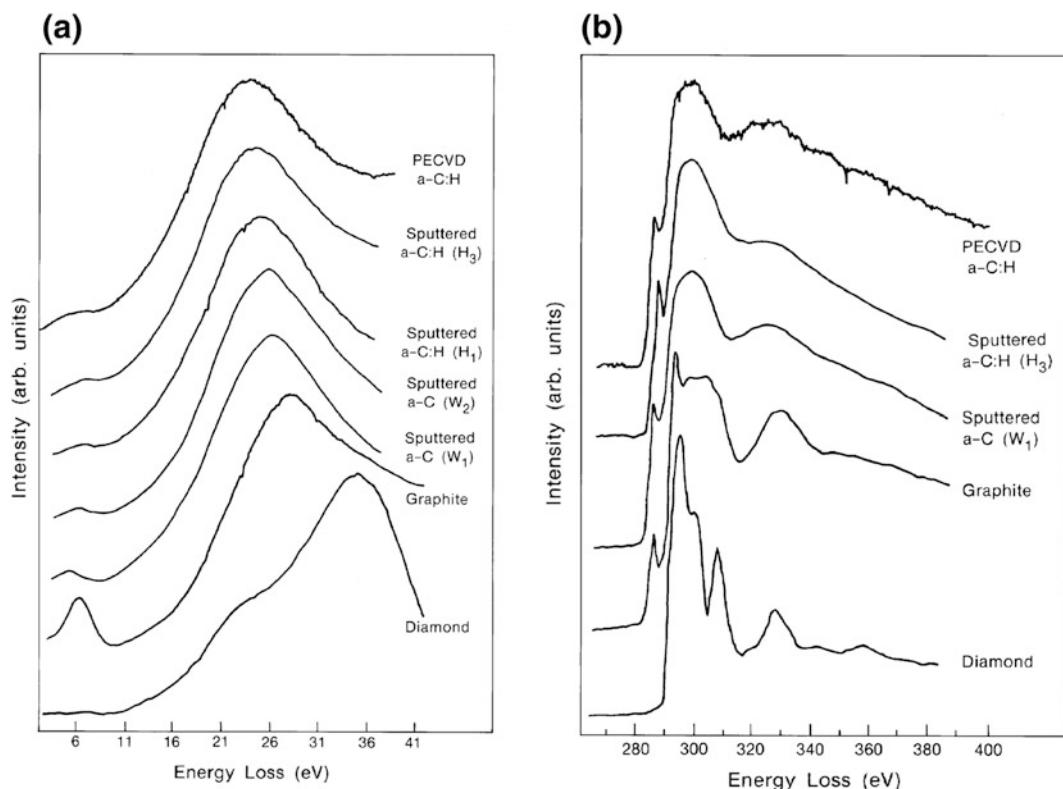


Fig. 12.5 **a** Low energy, and **b** high energy EELS of DLC coatings produced by DC magnetron sputtering and RF-PECVD techniques. Data for bulk diamond and polycrystalline graphite are included for comparisons (Bhushan et al. 1992)

Table 12.2 Experimental results from EELS and Raman spectroscopy (Bhushan et al. 1992)

Sample	EELS peak position		Raman peak position		Raman FWHM ^a		I_D/I_G^d
	π (eV)	$\pi + \sigma$ (eV)	G-band ^b (cm ⁻¹)	D-band ^c (cm ⁻¹)	G-band (cm ⁻¹)	D-band (cm ⁻¹)	
Sputtered α -C coating (W1)	5.0	24.6	1541	1368	105	254	2.0
Sputtered α -C coating (W2)	6.1	24.7	1560	1379	147	394	5.3
Sputtered α -C:H coating (H1)	6.3	23.3	1542	1334	95	187	1.6
Sputtered α -C:H coating (H3)	6.7	22.4	e	e	e	e	e
PECVD α -C:H coating	5.8	24.0	1533	1341	157	427	1.5
Diamond coating	1525 ^f	1333 ^g	...	8 ^g	...
Graphite (for reference)	6.4	27.0	1580	1358	37	47	0.7
Diamond (for reference)	...	37.0	...	1332 ^g	...	2 ^g	...

^aFull width at half maximum width^bPeak associated with sp^2 “graphite” carbon^cPeak associated with sp^2 “disordered” carbon (not sp^3 -bonded carbon)^dIntensity ratio of the D-band to the G-band^eFluorescence^fIncludes D and G band, signal too weak to analyze^gPeak position and width for diamond phonon

spectra of these carbon coatings. EELS spectra for bulk diamond and polycrystalline graphite in an energy range up to 50 eV are also shown in Fig. 12.5. One prominent peak is seen at 35 eV in diamond, while two peaks are seen at 27 eV and 6.5 eV in graphite, which are called $(\pi + \sigma)$ and (π) peaks, respectively. These peaks are produced by the energy loss of transmitted electrons to the plasmon oscillation of the valence electrons. The $\pi + \sigma$ peak in each coating is positioned at a lower energy region than that of graphite. The π peaks in the W series and PECVD samples are also seen at a lower energy region than that of the graphite; however, the π peaks in the H-series are comparable to or higher than that of graphite (see Table 12.2). The plasmon oscillation frequency is proportional to the square root of the corresponding electron density to a first approximation. Therefore, the samples in the H-series most likely have a higher density of π electrons than the other samples.

Amorphous carbon coatings contain (mainly) a mixture of sp^2 - and sp^3 -bonds even though there is some evidence for the presence of sp-bonds as well (Dischler et al. 1983). The PECVD coatings and the H-series coatings in this study have nearly the same mass density as seen in Table 12.4, to be

presented later, but the former have a lower concentration of hydrogen (18.1%) than the H-series (35–39%), as seen in Table 12.3, to be presented later. The relatively low energy position of π peaks of PECVD coatings, compared to those of the H-series, indicates that the PECVD coatings contain a higher fraction of sp^3 -bonds than the sputtered hydrogenated carbon coatings (H-series).

Figure 12.5b shows the EELS spectra associated with the inner-shell (K-shell) ionization. Again, the spectra for diamond and polycrystalline graphite are included for comparisons. Sharp peaks are observed at 285.5 and 292.5 eV in graphite, while no peak is seen at 285.5 eV in diamond. The general features of the K-shell EELS spectra for the sputtered and PECVD carbon samples resemble the features of graphite, but with the higher energy features smeared. The observation of the peak at 285.5 eV in the sputtered and PECVD coatings also indicates the presence of sp^2 -bonded atomic sites in the coatings. All these spectra peak at 292.5 eV, similar to the spectra of graphite, but the peak in graphite is sharper.

Raman spectra from samples W1, W2, H1, and PECVD are shown in Fig. 12.6. Raman

Table 12.3 Experimental results of FRS analysis (Bhusan et al. 1992)

Sample	Ar/H ratio	C (at. % ± 0.5)	H (at. % ± 0.5)	Ar (at. % ± 0.5)	O (at. % ± 0.5)
Sputtered <i>a</i> -C coating (W2)	100/0	90.5	9.3	0.2	...
Sputtered <i>a</i> -C:H coating (H2)	99/1	63.9	35.5	0.6	...
Sputtered <i>a</i> -C:H coating (H3)	97/3	56.1	36.5	...	7.4
Sputtered <i>a</i> -C:H coating (H4)	95/5	53.4	39.4	...	7.2
Sputtered <i>a</i> -C:H coating (H5)	93/7	58.2	35.4	0.2	6.2
Sputtered <i>a</i> -C:H coating (H6)	90/10	57.3	35.5	...	7.2
PECVD <i>a</i> -C:H coating	99.5% CH ₄	81.9	18.1
Diamond coating	H ₂ -1 mol% CH ₄	94.0	6.0

spectra could not be observed in specimens H2 and H3 due to high fluorescence signals. The Raman spectra of single crystal diamond and polycrystalline graphite are also shown for comparison in Fig. 12.6. The results of the spectral fits are summarized in Table 12.2. We will focus on the G-band position, which has been shown to be related to the fraction of sp³-bonded sites. Increasing the power density in the amorphous carbon coatings (W1 and W2) results in a higher G-band frequency, implying a smaller fraction of sp³-bonding in W2 than in W1. This is consistent with higher density of W1. H1 and PEVCD have still lower G-band positions than W1, implying an even higher fraction of sp³-bonding, which is presumably caused by the incorporation of H atoms into the lattice. The high hardness of H3 might be attributed to efficient sp³ cross-linking of small, sp²-ordered domains.

The Raman spectrum of a MPECVD diamond coating is shown in Fig. 12.6. The diamond Raman peak is at 1333 cm⁻¹ with a line width of 7.9 cm⁻¹. There is a small broad peak around 1525 cm⁻¹, which is attributed to a small amount of *a*-C:H. This impurity peak is not intense enough to fit to separate G- and D-bands. The diamond peak frequency is very close to that of natural diamond [1332.5 cm⁻¹, e.g. Fig. 12.6], indicating that the coating is not under stress (Knight and White 1989). The large line width compared to that of natural diamond (2 cm⁻¹) indicates that the microcrystallites likely have a high concentration of defects (Ager et al. 1991).

12.3.2 Hydrogen Concentrations

FRS analysis of six sputtered (W2, H2, H3, H4, H5, and H6) coatings, one PECVD coating, and one diamond coating was performed. Figure 12.7 shows an overlay of the spectra from the six sputtered samples. Similar spectra were obtained from the PECVD and the diamond films. Table 12.3 shows the H and C fractions as well as the amount of impurities (Ar and O) in the films in atomic %. Most apparent is the large fraction of H in the sputtered films. Regardless of how much H₂ is in the Ar sputtering gas, the H content of the coatings is about the same, ~35 at.%. Interestingly, there is still ~10% H present in the coating sputtered in pure Ar (W2). It is interesting to note that Ar is present only in coatings grown under low (<1%) H content in the Ar carrier gas. The presence of O in the coatings, combined with the fact that the coatings were prepared approximately 9 months before the FRS analysis, caused suspicion that they had absorbed water vapor, and that this may be the cause for the H peak in specimen W2.

All samples were annealed for 24 h at 250 °C in a flowing He furnace and then reanalyzed. Surprisingly, the H content of all coatings measured increased slightly, even though the O content decreased, and W2 still had a substantial amount of H₂. This slight increase in H concentration is not understood; however, since H concentration did not decrease with the oxygen as a result of annealing, it suggests that high H concentration is not due to adsorbed water vapor.

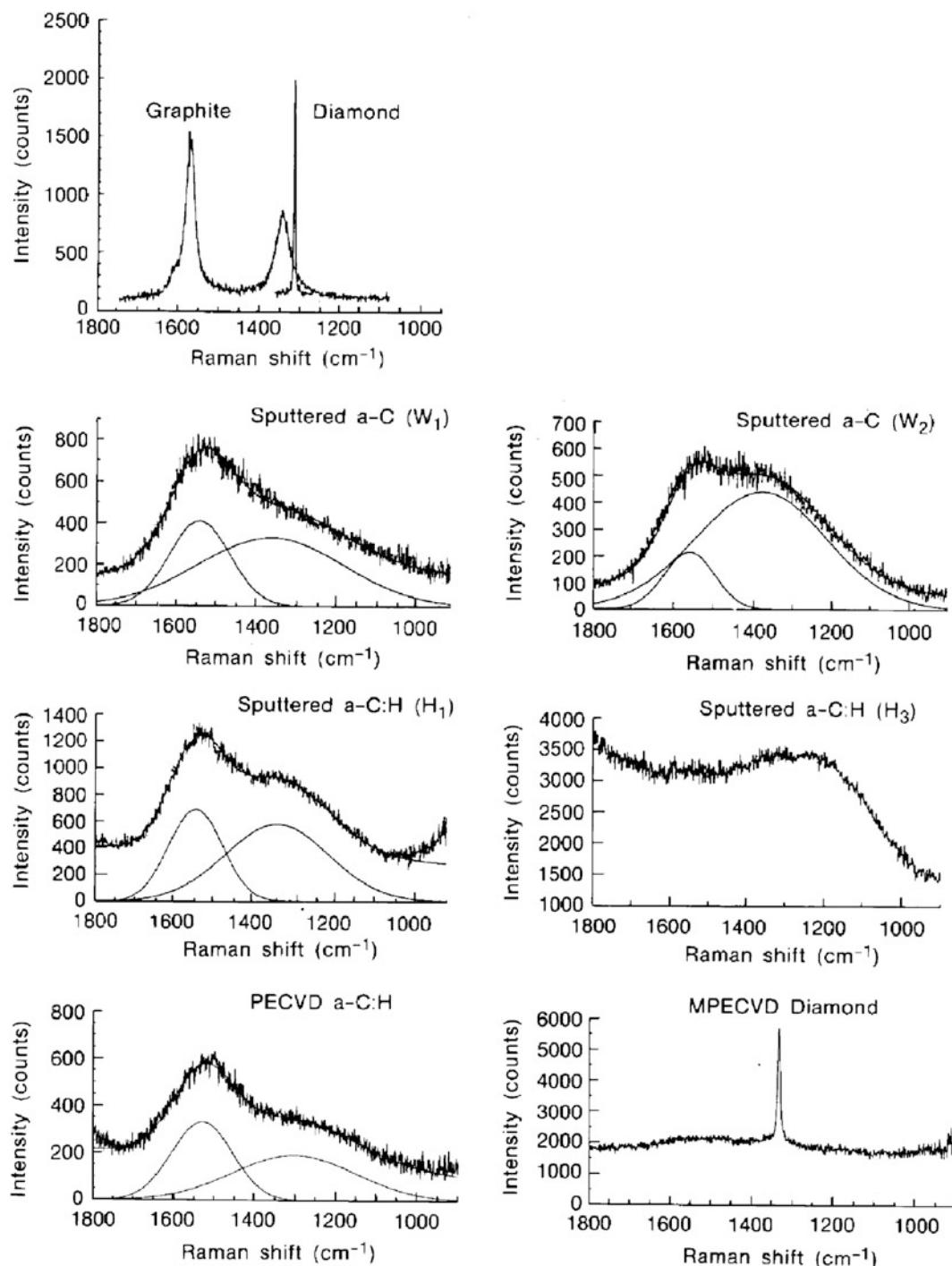


Fig. 12.6 Raman spectra of DLC coatings produced by DC magnetron sputtering and RF-PECVD techniques and a diamond film produced by MPECVD technique. Data for bulk diamond and microcrystalline graphite are included for comparison (Bhushan et al. 1992)

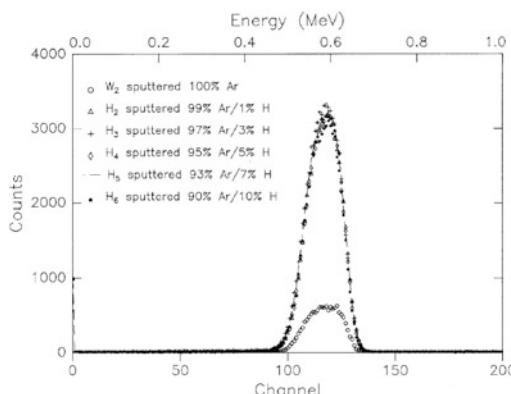


Fig. 12.7 FRS spectra of six DLC coatings produced by DC magnetron sputtering (Bhushan et al. 1992)

The PECVD film has more H ($\sim 18\%$) than the sputtered films initially, but after annealing it has the same fraction as specimen W2, the film sputtered in pure Ar. The diamond film has the smallest amounts of hydrogen, as seen in Table 12.3.

12.3.3 Physical Properties

Physical properties of the four sputtered (W1, W2, H1, and H3) coatings, one PECVD coating, one diamond coating, coatings and bulk diamond and graphite are presented in Table 12.4. The hydrogenated carbon and the diamond coatings have very high resistivity compared to

unhydrogenated carbon coatings. It appears that unhydrogenated carbon coatings have a higher density than the hydrogenated carbon coatings, although both groups are less dense than graphite. The density depends on the deposition technique and the deposition parameters. It appears that unhydrogenated sputtered coatings deposited at low power exhibit the highest density. Nanohardness of hydrogenated carbon is somewhat lower than that of the unhydrogenated carbon coatings. PECVD coatings are significantly harder than sputtered coatings. The nanohardness and modulus of elasticity of the diamond coatings are very high compared to that of DLC coatings even though the hydrogen content is similar. The compressive residual stresses of the PECVD coatings are substantially higher than those of sputtered coatings, which is consistent with the hardness results.

Figure 12.8a shows the effect of hydrogen in the plasma on the residual stresses and the nanohardness for sputtered coatings W2 and H1 to H6. The coatings made with H_2 flow between 0.5 and 1.0% delaminate very quickly, even when only a few tens of nm thick. In pure Ar and at H_2 flows greater than 1%, the coatings appear to be more adhesive. The tendency of some coatings to delaminate can be caused by intrinsic stress in the coating, which is measured by substrate bending. All of the coatings in the figure are in compressive stress. The maximum stress occurs between 0 and

Table 12.4 Experimental results of physical properties (Bhushan et al. 1992)

Sample	Mass density (g/cm ³)	Nano-hardness (GPa)	Elastic modulus (GPa)	Electrical resistivity (Ω·cm)	Compressive residual stress (GPa)
Sputtered α -C coating (W1)	2.1	15	141	1300	0.55
Sputtered α -C:H coating (W2)	1.8	14	136	0.61	0.57
Sputtered α -C:H coating (H1)	...	14	96	...	>2
Sputtered α -C:H coating (H3)	1.7	7	35	>10 ⁶	0.3
PECVD α -C:H coating	1.6–1.8	33–35	~200	>10 ⁶	1.5–3.0
Diamond coating	...	40–75	370–430
Graphite (for reference)	2.267	Soft	9–15	5×10^{-5} ^a , 4×10^{-3} ^b	0
Diamond (for reference)	3.515	70–102	900–1050	10^7 – 10^{20}	0

^aParallel to layer planes

^bPerpendicular to layer planes

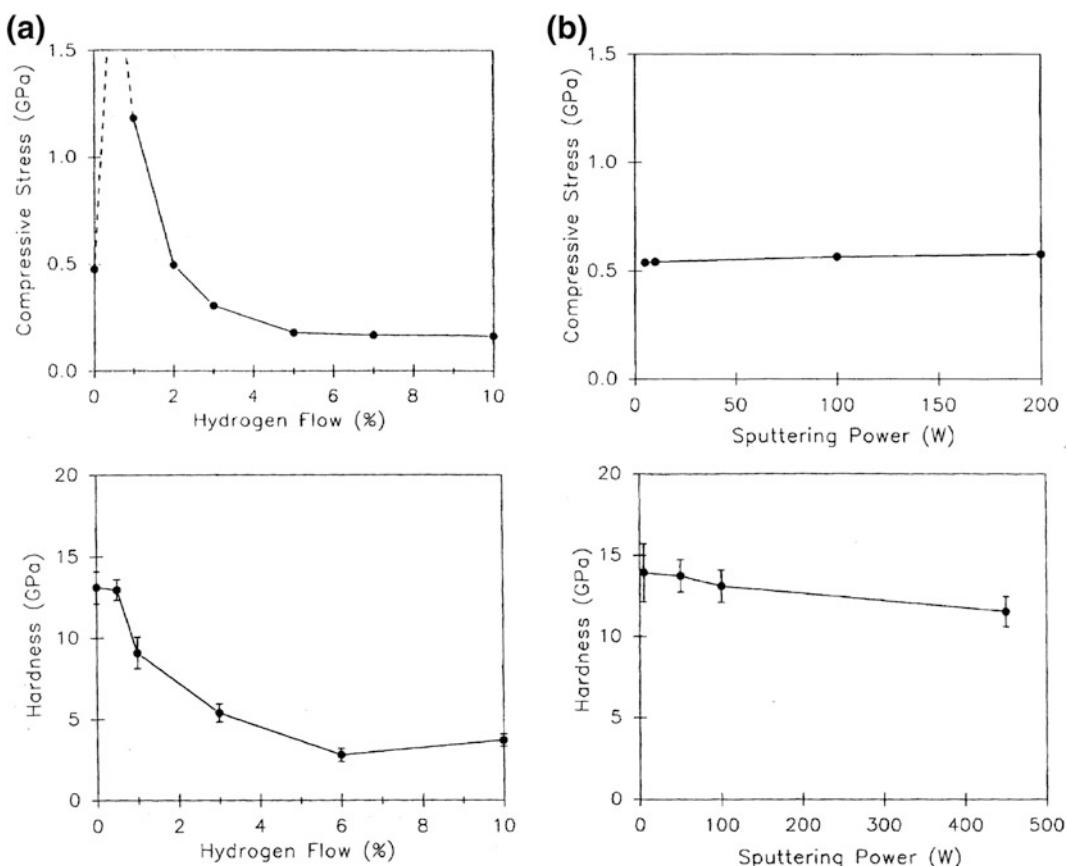


Fig. 12.8 Residual compressive stresses and nanohardness **a** as a function of hydrogen flow rate; sputtering power is 100 W and target diameter is 75 mm (power density = 2.1 W/cm²), and **b** as a function of sputtering power over a 75 mm diameter target with no hydrogen added in the plasma (Rubin et al. 1990)

1% H₂ flow, but the stress cannot be quantified in this range because the coatings instantly delaminate upon exposure to air. At higher hydrogen concentrations the stress gradually diminishes. A generally decreasing trend is observed in the hardness of the coatings as hydrogen content increased. The hardness decreases slightly, going from 0 to 0.5% H₂, and then decreases sharply. These results are probably lower than the true values because of local delamination around the indentation point. This is especially likely for the 0.5 and 1.0% coatings where delamination is visually apparent, but may also be true to a lesser extent for the other coatings. Such an adjustment would bring the hardness profile into closer correlation with the stress profile. Weissmantel et al. (1982) and Scharff et al. (1989) observed a

downturn in hardness for high bias and low pressure of hydrocarbon gas in ion plated carbon coating and therefore presumably low hydrogen content in support of the above contention.

Figure 12.8b shows the effect of sputtering power (with no hydrogen added in the plasma) on the residual stresses and nanohardness for various sputtered coatings. As the power decreases, compressive stress does not seem to change while nanohardness slowly increases. The rate of change becomes more rapid at very low power levels.

Addition of H₂ during sputtering of carbon coatings increases H concentration in the coating. Hydrogen causes the character of the C-C bonds to shift from sp² to sp³, and the rising number of C-H bonds which ultimately relieves stress and produces a softer “polymer-like” material. Low

power deposition, like the presence of hydrogen, appears to stabilize the formation of sp^3 C-C bonds, increasing hardness. These coatings have relieved stress and better adhesion. An increase in temperature during deposition at high power density results in graphitization of the coating material, responsible for a decrease in hardness with an increase in power density. Unfortunately, low power also means impractically low deposition rates.

12.3.4 Summary

Based on the EELS and Raman data, all DLC coatings have both sp^2 - and sp^3 -bondings. sp^2/sp^3 bonding ratio depends on the deposition techniques and parameters. The DLC coatings deposited by sputtering and PECVD, contain significant concentrations of hydrogen while the diamond coating contains only small amounts of hydrogen impurity. Sputtered coatings with no deliberate addition of hydrogen in the plasma contain a significant amount of hydrogen. Regardless of how much hydrogen is in the Ar sputtering gas, the hydrogen content of the coatings increases initially with no further increase.

Hydrogen flow and sputtering power density affect the mechanical properties of these coatings. Maximum compressive residual stress and hardness occur between 0 and 1% hydrogen flow, resulting in rapid delamination. Low sputtering power moderately increases hardness while relieving residual stress.

12.4 Micromechanical and Tribological Characterizations of Coatings Deposited by Various Techniques

12.4.1 Micromechanical Characterization

Common mechanical characterizations include measurement of hardness and elastic modulus, fracture toughness, fatigue life, and scratch and

wear testing. Nanoindentation and atomic force microscopy (AFM) are used for mechanical characterization of ultrathin films.

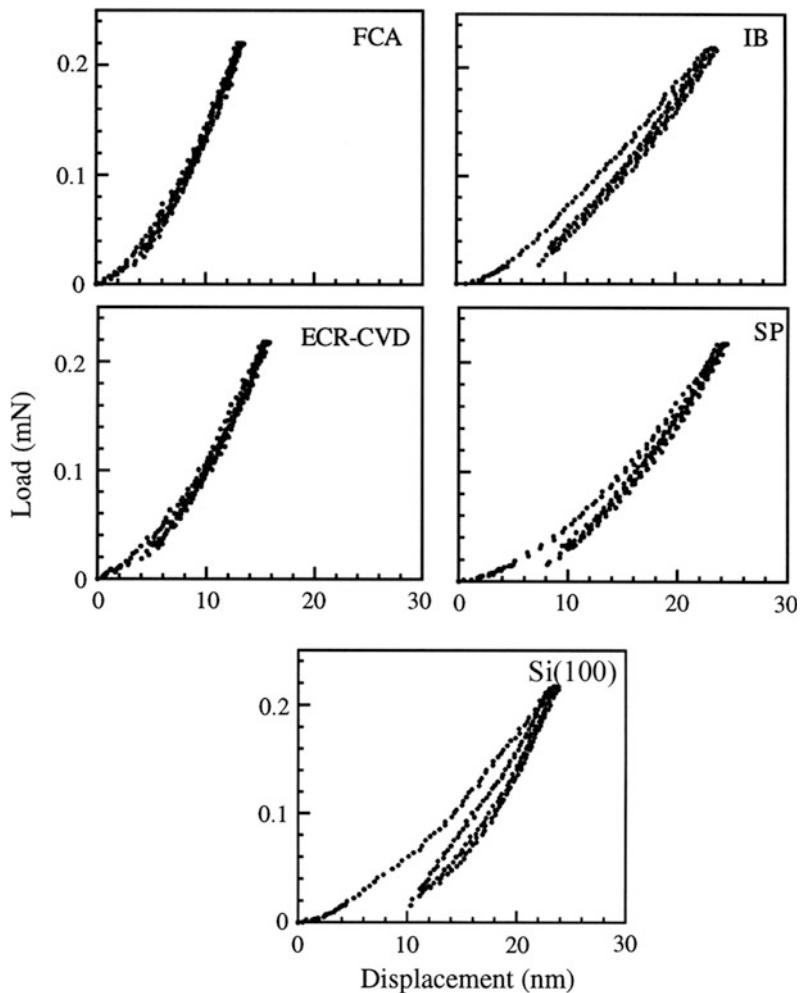
Hardness and elastic modulus are calculated from the load displacement data obtained by nanoindentation at loads ranging typically from 0.2 to 10 mN using a commercially available nanoindenter (Bhushan 1999a; Bhushan and Li 2003). This instrument monitors and records the dynamic load and displacement of the three-sided pyramidal diamond (Berkovich) indenter during indentation. For fracture toughness measurement of ultra-thin films ranging from 100 nm to few μm , a nanoindentation-based technique is used in which through-thickness cracking in the coating is detected from a discontinuity observed in the load-displacement curve and energy released during the cracking is obtained from the curve (Li et al. 1997; Li and Bhushan 1998b, 1999c). Based on the energy released, fracture mechanics analysis is then used to calculate fracture toughness. An indenter with a cube corner tip geometry is preferred because the through-thickness cracking of hard films can be accomplished at lower loads. In fatigue measurement, a conical diamond indenter having a tip radius of about one micron is used and load cycles of a sinusoidal shape are applied (Li and Bhushan 2002a, b). The fatigue behavior of coatings is studied by monitoring the change in contact stiffness which is sensitive to the damage formation.

12.4.1.1 Hardness and Elastic Modulus

For materials which undergo plastic deformation, high hardness and elastic modulus are generally needed for low friction and wear, whereas for brittle materials, high fracture toughness is needed (Bhushan and Gupta 1991; Bhushan 1996b, 2013a, b). DLC coatings used for many applications are hard and brittle and values of hardness and fracture toughness need to be optimized.

Representative load-displacement plots of indentations made at 0.2 mN peak indentation load on 100-nm thick DLC coatings deposited by the four deposition techniques on single-crystal silicon substrate are compared in Fig. 12.9. The indentation depths at the peak load range from

Fig. 12.9 Load versus displacement plots for various 100-nm thick amorphous carbon coatings on single-crystal silicon substrate and bare substrate



about 18 to 26 nm, smaller than that of the coating thickness. Many of the coatings exhibit a discontinuity or pop-in marks in the loading curve, which indicate a sudden penetration of the tip into the sample. A nonuniform penetration of the tip into a thin coating possibly results due to either formation of cracks in the coating, formation of cracks at the coating-substrate interface, or debonding or delamination of the coating from the substrate.

The hardness and elastic modulus values at a peak load of 0.2 mN on the various coatings and single-crystal silicon substrate are summarized in Table 12.5 and Fig. 12.10 (Gupta and Bhushan 1995b; Li and Bhushan 1999b, c, 2002a). Typical values for the peak and residual indentation

depths ranged from 18 to 26 nm and 6 to 12 nm, respectively. The FCA coatings exhibit the highest hardness of 24 GPa and elastic modulus of 280 GPa of various coatings, followed by the ECR-CVD, IB, and SP coatings. Hardness and elastic modulus have been known to vary over a wide range with sp^3 - to sp^2 -bonding ratio, which depends on the kinetic energy of the carbon species and amount of hydrogen (Angus and Hayman 1988; Robertson 1993; Gupta and Bhushan 1994, 1995b; Bull 1995). The high hardness and elastic modulus of the FCA coatings are attributed to the high kinetic energy of carbon species involved in the FCA deposition (Gupta and Bhushan 1995a, b). Anders et al. (1994a) also reported a high hardness, measured

Table 12.5 Hardness, elastic modulus, fracture toughness, fatigue life, critical load during scratch, coefficient of friction during accelerated wear testing and residual stresses of various DLC coatings on single-crystal silicon substrate

Coating	Hardness ^a (Li and Bhushan 1999a) (GPa)	Elastic modulus ^a (Li and Bhushan 1999a) (GPa)	Fracture toughness ^a (Li and Bhushan 1999c) (MPa m ^{1/2})	Fatigue life ^b , N _f (Li and Bhushan 2002a) (×10 ⁴)	Critical load during scratch ^b (Li and Bhushan 1999a) (mN)	Coefficient of friction during accelerated wear testing ^b (Li and Bhushan 1999a)	Compressive residual stress ^c (Gupta and Bhushan 1995b) (GPa)
Cathodic arc carbon coating (a-C)	24	280	11.8	2.0	3.8	0.18	12.5
Ion beam carbon coating (a-C:H)	19	140	4.3	0.8	2.3	0.18	1.5
ECR-CVD carbon coating (a-C:H)	22	180	6.4	1.2	5.3	0.22	0.6
DC sputtered carbon coating (a-C:H)	15	140	2.8	0.2	1.1	0.32	2.0
Bulk graphite (for comparison)	Very soft	9–15	–	–	–	–	–
Diamond (for comparison)	80–104	900–1050	–	–	–	–	–
Si(100) substrate	11	220	0.75	–	0.6	0.55	0.02

^aMeasured on 100-nm thick coatings^bMeasured on 20-nm thick coatings^cMeasured on 400-nm thick coatings^dN_f was obtained at a mean load of 10 μN and a load amplitude of 8 μN

by nanoindentation, of about 45 GPa for cathodic arc carbon coatings. They observed a change in hardness from 25 to 45 GPa with pulsed bias voltage and bias duty cycle. The high hardness of cathodic arc carbon was attributed to the high percentage (more than 50%) of sp³ bonding. Savvides and Bell (1992) reported an increase in hardness from 12 to 30 GPa and in elastic modulus from 62 to 213 GPa with an increase of sp³-to-sp² bonding ratio, from 3 to 6, for a C:H coating deposited by low-energy ion-assisted

unbalanced magnetron sputtering of a graphite target in an Ar-H₂ mixture.

Bhushan et al. (1992) reported hardnesses of about 15 and 35 GPa and elastic moduli of about 140 and 200 GPa, measured by nanoindentation, for a-C:H coatings deposited by DC magnetron sputtering and RF-plasma-enhanced chemical vapor deposition techniques, respectively. The high hardness of RF-PECVD a-C:H coatings is attributed to a higher concentration of sp³ bonding than in a sputtered hydrogenated a-C:H coating. Hydrogen is believed to play a crucial

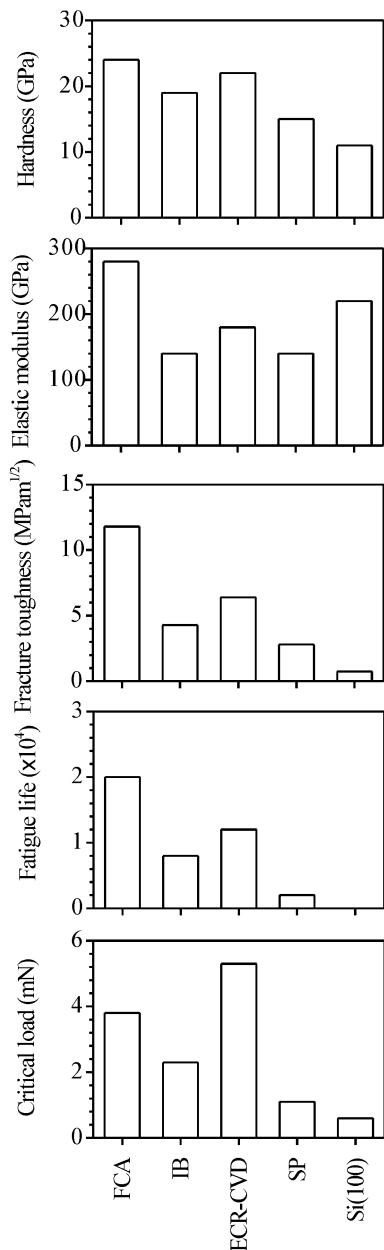


Fig. 12.10 Bar charts summarizing the data of various coatings and single-crystal silicon substrate. Hardness, elastic modulus, and fracture toughness were measured on 100-nm thick coatings, and fatigue life and critical load during scratch were measured on 20-nm thick coatings

role in the bonding configuration of carbon atoms by helping to stabilize tetrahedra-coordination (sp^3 bonding) of carbon species. Jansen et al. (1985) suggested that the incorporation of

hydrogen efficiently passivates the dangling bonds and saturates the graphitic bonding to some extent. However, a large concentration of hydrogen in the plasma in sputter deposition is undesirable. Cho et al. (1990) and Rubin et al. (1990) observed that hardness decreased from 15 to 3 GPa with increased hydrogen content. Bhushan and Doerner (1989) reported a hardness of about 10–20 GPa and an elastic modulus of about 170 GPa, measured by nanoindentation, for 100 nm thick DC-magnetron sputtered a-C:H on the silicon substrate.

Residual stresses measured using a well-known curvature measurement technique are also presented in Table 12.5. The DLC coatings are under significant compressive internal stresses. Very high compressive stresses in FCA coatings are believed to be partly responsible for their high hardness. However, high stresses result in coating delamination and buckling. For this reason, the coatings thicker than about 1 μm have a tendency to delaminate from the substrates.

12.4.1.2 Fracture Toughness

Representative load-displacement curves of indentations on the 400-nm thick cathodic arc carbon coating on silicon at various peak loads are shown in Fig. 12.11. Steps are found in all curves as shown by arrows in Fig. 12.11a. In the 30 mN SEM micrograph, in addition to several radial cracks, ring-like through-thickness cracking is observed with small lips of material overhanging the edge of indentation. The step at about 23 mN in the loading curves of indentations made at 30 and 100 mN peak indentation loads results from the ring-like through-thickness cracking. The step at 175 mN in the loading curve of indentation made at 200 mN peak indentation load is caused by spalling and second ring-like through-thickness cracking.

Based on Li et al. (1997), the fracture process progresses in three stages: (1) first ring-like through-thickness cracks form around the indenter by high stresses in the contact area, (2) delamination and buckling occur around contact area at the coating/substrate interface by high lateral pressure, (3) second ring-like through-thickness

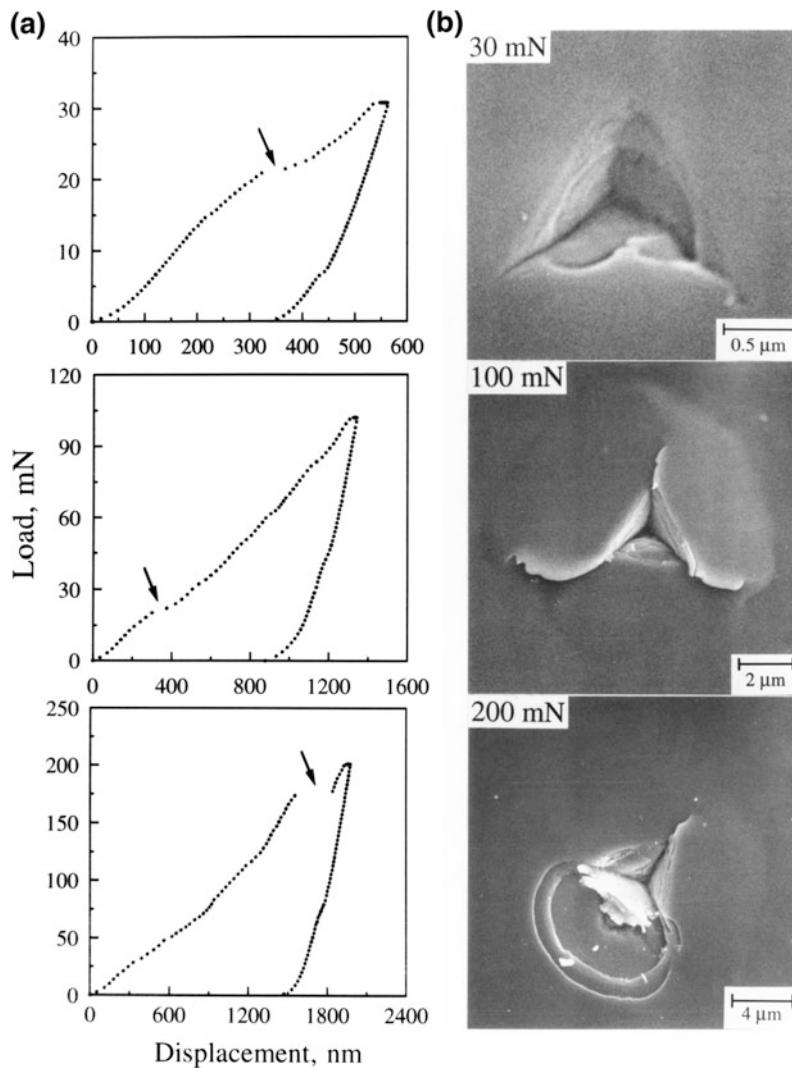


Fig. 12.11 **a** Load-displacement curves of indentations made at 30, 100 and 200 mN peak indentation loads using the cube corner indenter, and **b** the SEM micrographs of indentations on the 400-nm thick cathodic arc carbon coating on silicon. Arrows indicate steps during loading portion of the load-displacement curve (Li et al. 1997)

cracks and spalling are generated by high bending stresses at the edges of the buckled coating, see Fig. 12.12a. In the first stage, if the coating under the indenter is separated from the bulk coating via the first ring-like through-thickness cracking, a corresponding step will be present in the loading curve. If discontinuous cracks form and the coating under the indenter is not separated from the remaining coating, no step appears in the loading curve because the coating still supports the indenter and the indenter cannot suddenly

advance into the material. In the second stage, for the coating used in the present study, the advances of the indenter during the radial cracking, delamination and buckling are not big enough to form steps in the loading curve because the coating around the indenter still supports the indenter, but generate discontinuities which change the slope of the loading curve with increasing indentation loads. In the third stage, the stress concentration at the end of the interfacial crack cannot be relaxed by the propagation of the interfacial crack.

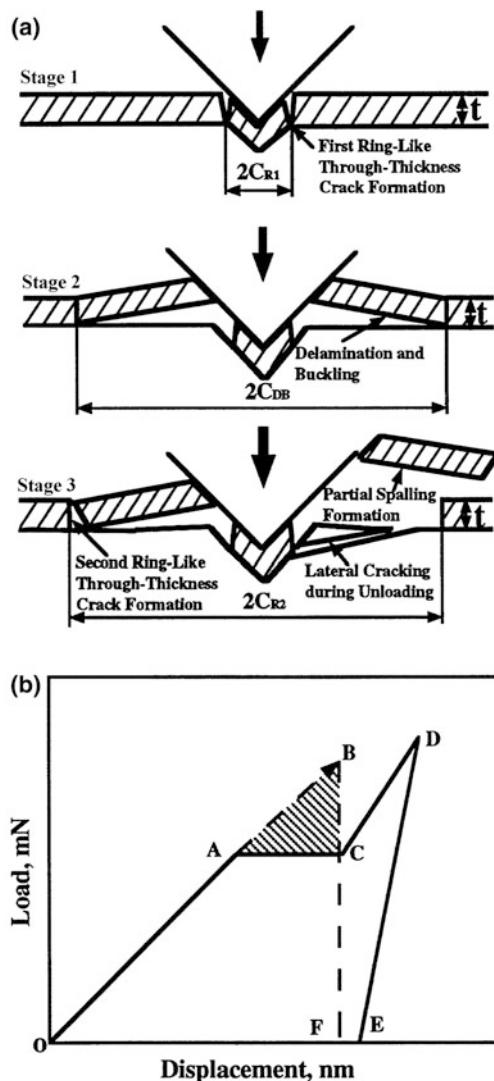


Fig. 12.12 **a** Schematic of various stages in nanoindentation fracture for the coatings/substrate system, and **b** schematic of a load-displacement curve, showing a step during the loading cycle and associated energy release

With an increase in indentation depth, the height of the bulged coating increases. When the height reaches a critical value, the bending stresses caused by the bulged coating around the indenter will result in the second ring-like through-thickness crack formation and spalling at the edge of the buckled coating as shown in Fig. 12.12a, which leads to a step in the loading curve. This is a single event and results in the

separation of the part of the coating around the indenter from the bulk coating via cracking through coatings. The step in the loading curve is totally from the coating cracking and not from the interfacial cracking or the substrate cracking.

The area under the load-displacement curve is the work performed by the indenter during elastic-plastic deformation of the coating/substrate system. The strain energy release in the first/second ring-like cracking and spalling can be calculated from the corresponding steps in the loading curve. Figure 12.12b shows a modeled load-displacement curve. OACD is the loading curve and DE is the unloading curve. Since the first ring-like through-thickness crack should be considered. It should be emphasized that the edge of the buckled coating is far from the indenter, therefore, it does not matter if the indentation depth exceeds the coating thickness or if deformation of the substrate occurs around the indenter when we measure fracture toughness of the coating from the released energy during the second ring-like through-thickness cracking (spalling). Suppose that the second ring-like through-thickness cracking occurs at AC. Now, let us consider the loading curve OAC. If the second ring-like through-thickness crack does not occur, it can be understood that OA will be extended to OB to reach the same displacement as OC. This means that the crack formation changes the loading curve OAB into OAC. For point B, the elastic-plastic energy stored in the coating/substrate system should be OBF. For point C, the elastic-plastic energy stored in the coating/substrate system should be OACF. Therefore, the energy difference before and after the crack generation is the area of ABC, i.e., this energy stored in ABC will be released as strain energy to create the ring-like through-thickness crack. According to the theoretical analysis by Li et al. (1997), the fracture toughness of thin films can be written as

$$K_{Ic} = \left[\left(\frac{E}{(1 - v^2)2\pi C_R} \right) \left(\frac{U}{t} \right) \right]^{1/2} \quad (12.2)$$

where E is the elastic modulus, ν is the Poisson's ratio, $2\pi C_R$ is the crack length in the coating plane, t is the coating thickness, and U is the strain energy difference before and after cracking.

Using Eq. 12.2, the fracture toughness of the coatings are calculated. The loading curve is extrapolated along the tangential direction of the loading curve from the starting point of the step up to reach the same displacement as the step. The area between the extrapolated line and the step is the estimated strain energy difference before and after cracking. C_R is measured from SEM micrographs or AFM images of indentations. Second ring-like crack is where the spalling occurs. For example, for the 400-nm thick cathodic arc carbon coating data presented in Fig. 12.11, U of 7.1 nNm is assessed from the steps in Fig. 12.11a at the peak indentation loads of 200 mN. For C_R of 7.0 μm from Fig. 12.11b, $E = 300 \text{ GPa}$ measured using nanoindenter and an assumed value of 0.25 for ν , fracture toughness values is calculated as 10.9 MPa $\sqrt{\text{m}}$ (Li et al. 1997; Li and Bhushan 1998b). The fracture toughness and related data for various 100-nm thick DLC coatings are presented in Fig. 12.10 and Table 12.5.

12.4.1.3 Nanofatigue

Delayed fracture resulting from extended service is called fatigue (Suresh 1991). Fatigue fracturing progresses through a material via changes within the material at the tip of a crack, where there is a high stress intensity. There are several situations: cyclic fatigue, stress corrosion, and static fatigue. Cyclic fatigue results from cyclic loading of machine components. In a low-flying slider in a magnetic head-disk interface, isolated asperity contacts occur during use and the fatigue failure occurs in the multilayered thin-film structure of the magnetic disk (Bhushan 1996a). In many MEMS components, impact occurs and the failure mode is cyclic fatigue. Asperity contacts can be simulated using a sharp diamond tip in an oscillating contact with the component.

Figure 12.13 shows the schematic of a fatigue test on a coating/substrate system using a continuous stiffness measurement (CSM) technique.

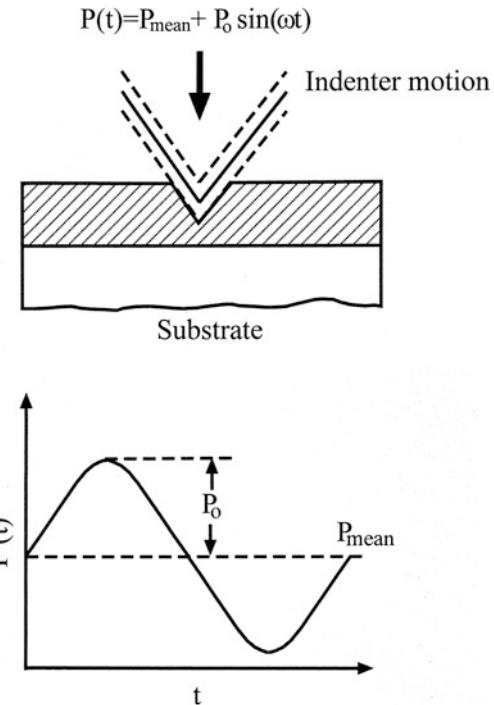


Fig. 12.13 Schematic of a fatigue test on a coating/substrate system using the continuous stiffness measurement technique

Load cycles are applied to the coating, resulting in a cyclic stress; P is the cyclic load, P_{mean} is the mean load, P_0 is the oscillation load amplitude, and ω is the oscillation frequency. The following results can be obtained: (1) endurance limit; i.e., the maximum load below which there is no coating failure for a preset number of cycles; (2) number of cycles at which the coating failure occurs; and (3) changes in contact stiffness measured by using the unloading slope of each cycle which can be used to monitor the propagation of the interfacial cracks during cyclic fatigue process.

Figure 12.14a shows the contact stiffness as a function of the number of cycles for 20-nm thick FCA coatings cyclically deformed by various oscillation load amplitudes with a mean load of 10 μN at a frequency of 45 Hz. At 4 μN load amplitude, no change in contact stiffness was found for all coatings. This indicates that 4 μN load amplitude is not high enough to damage the

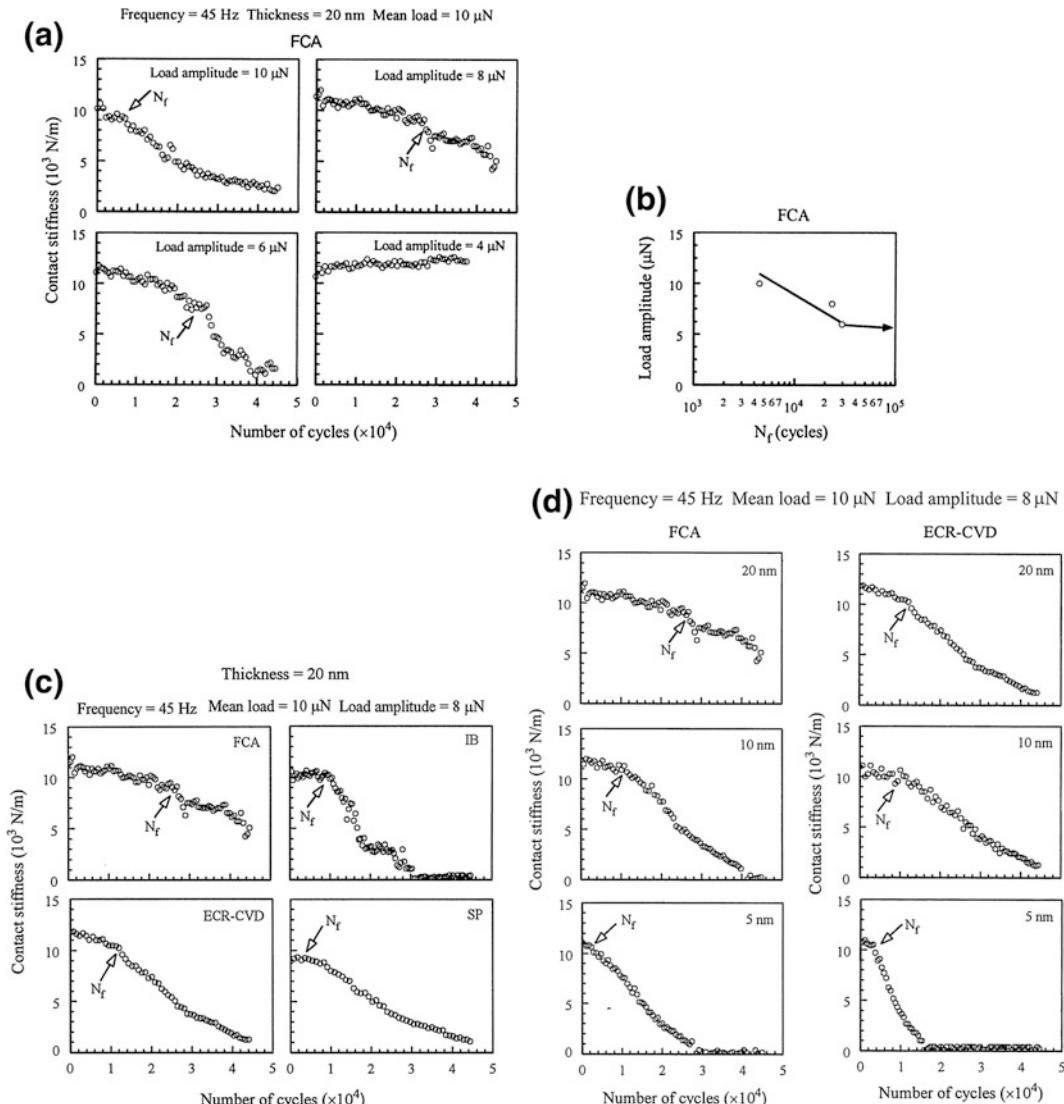


Fig. 12.14 **a** Contact stiffness as a function of the number of cycles for 20-nm thick FCA coatings cyclically deformed by various oscillation load amplitudes with a mean load of $10 \mu\text{N}$ at a frequency of 45 Hz; **b** plot of load amplitude versus N_f ; **c** contact stiffness as a function of the number of cycles for four different 20-nm thick coatings; and **d** contact stiffness as a function of the number of cycles for two coatings of the different thicknesses

coatings. At $6 \mu\text{N}$ load amplitude, an abrupt decrease in contact stiffness was found at a certain number of cycles for each coating, indicating that fatigue damage has occurred. With increasing load amplitude, the number of cycles to failure, N_f , decreases for all coatings. Load amplitude versus N_f , a so-called S-N curve, is plotted in Fig. 12.14b. The critical load

amplitude, below which no fatigue damage occurs (an endurance limit), was identified for each coating. This critical load amplitude together with mean load is of critical importance to the design of head-disk interfaces or MEMS/NEMS device interfaces.

To compare the fatigue lives of different coatings studied, the contact stiffness as a

function of the number of cycles for 20-nm thick FCA, IB, ECR-CVD and SP coatings cyclically deformed by an oscillation load amplitude of 8 μN with a mean load of 10 μN at a frequency of 45 Hz is shown in Fig. 12.14c. FCA coating has the longest N_f , followed by ECR-CVD, IB, and SP coatings. In addition, after the N_f , the contact stiffness of the FCA coating shows a slower decrease than the other coatings. This indicates that after the N_f , the FCA coating had less damage than others. The fatigue behavior of FCA and ECR-CVD coatings of different thicknesses is compared in Fig. 12.14d. For both coatings, N_f decreases with decreasing coating thickness. At 10 nm, FCA and ECR-CVD have almost the same fatigue life. At 5 nm, ECR-CVD coating shows a slightly longer fatigue life than FCA coating. This indicates that even for nanometer thick DLC coatings their microstructure and residual stresses are not uniform across the thickness direction. Thinner coatings are more influenced by interfacial stresses than thicker coating.

Figure 12.15a shows the high magnification SEM images of 20-nm thick FCA coatings before, at, and after N_f . In the SEM images, the net-like structure is the gold film coated on the DLC coating, which should be ignored in analyzing the indentation fatigue damage. Before the N_f , no delamination or buckling was found except the residual indentation mark at magnifications up to 1,200,000 \times using SEM. This suggests that only plastic deformation occurred before the N_f . At the N_f , the coating around the indenter bulged upwards, indicating delamination and buckling. Therefore, it is believed that the decrease in contact stiffness at the N_f results from the delamination and buckling of the coating from the substrate. After the N_f , the buckled coating was broken down around the edge of the buckled area, forming a ring-like crack. The remaining coating overhung at the edge of the buckled area. It is noted that the indentation size increases with increasing number of cycles. This

indicates that deformation, delamination and buckling, and ring-like crack formation occurred over a period of time.

The schematic in Fig. 12.15b shows various stages in the indentation fatigue damage for a coating/substrate system. Based on this study, three stages in the indentation fatigue damage appear to exist: (1) indentation induced compression; (2) delamination and buckling; (3) ring-like crack formation at the edge of the buckled coating. Residual stresses are often induced in coatings by the deposition process. The model shown in Fig. 12.15b considers a coating with a uniform biaxial residual compression σ_r . In the first stage, indentation induces elastic/plastic deformation, exerting an outward acting pressure on the coating around the indenter. Interfacial defects like voids and impurities act as original cracks. These cracks propagate and link up as the indentation compressive stress increases. At this stage, the coating, which is under the indenter and above the interfacial crack (with a crack length of $2a$), still maintains a solid contact with the substrate; the substrate still fully supports the coating. Therefore, this interfacial crack does not lead to an abrupt decrease in contact stiffness, but gives a rise to a slight decrease in contact stiffness as shown in Fig. 12.14. The coating above the interfacial crack is treated as a rigidly clamped disk. We assume that the crack radius, a , is large compared with the coating thickness t . Since the coating thickness ranges from 20 to 5 nm, this assumption is easily satisfied in this study (radius of the delaminated and buckled area, shown in Fig. 12.15a, is on the order of 100 nm). The compressive stress caused by indentation is given as (Marshall and Evans 1984)

$$\sigma_i = \frac{EV_i}{2\pi t a^2(1-v)} \quad (12.3)$$

where v and E are the Poisson's ratio and elastic modulus of the coating, V_i is the indentation

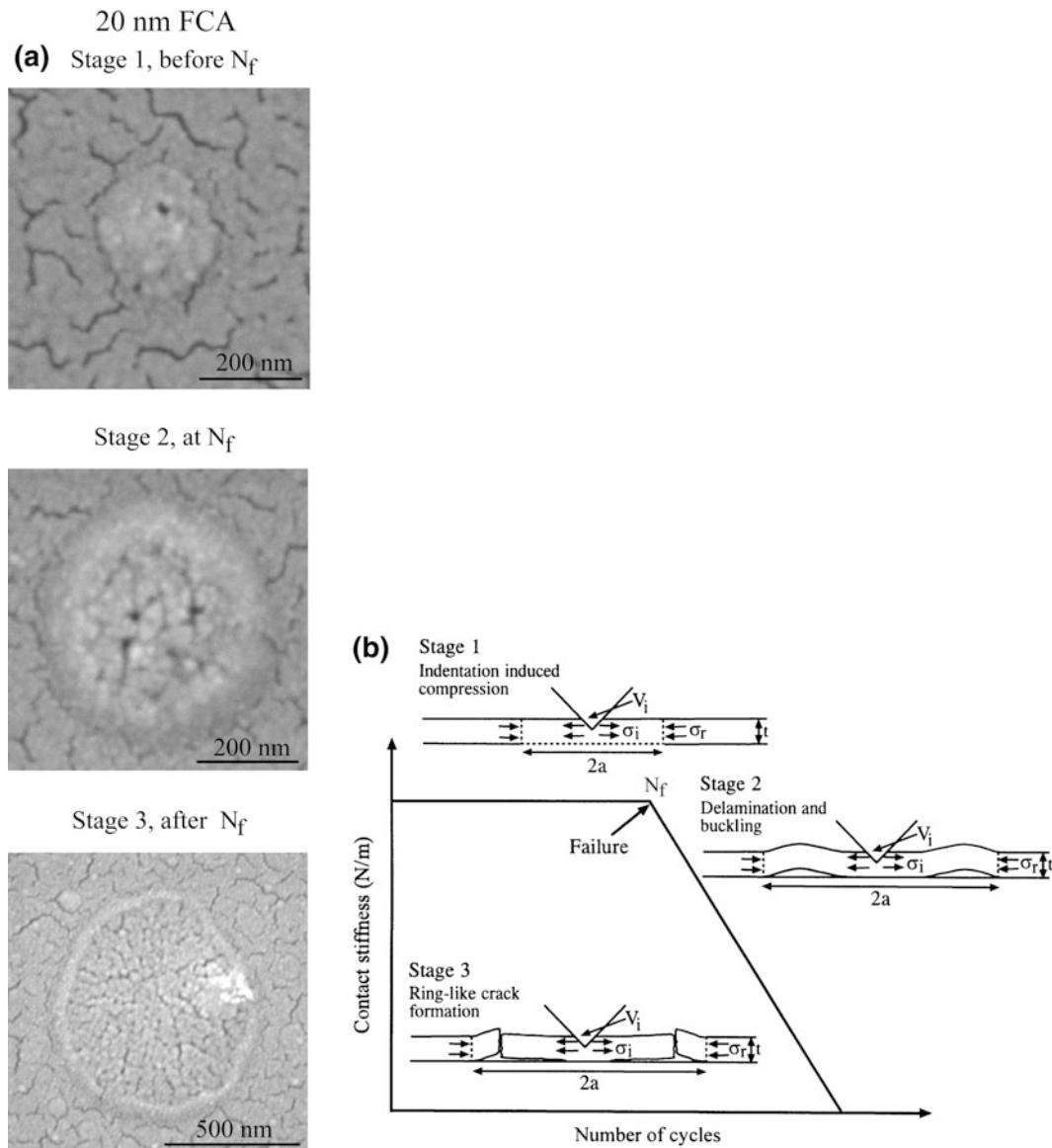


Fig. 12.15 **a** High magnification SEM images of a coatings before, at, and after N_f , and **b** schematic of various stages in the indentation fatigue damage for a coating/substrate system (Li and Bhushan 2002a)

volume, t is the coating thickness, and a is the crack radius. With increasing number of cycles, the indentation volume V_i increases. Therefore, the indentation compressive stress σ_i increases accordingly. In the second stage, buckling occurs during the unloading segment of fatigue testing cycle when the sum of indentation compressive stress σ_i and the residual stress σ_r exceed the critical buckling stress σ_b for the delaminated

circular section as given by (Evans and Hutchinson 1984)

$$\sigma_b = \frac{\mu^2 E}{12(1 - \nu^2)} \left(\frac{t}{a}\right)^2 \quad (12.4)$$

where the constant μ equals 42.67 for a circular clamped plate with a constrained center point and 14.68 when the center is unconstrained. The

buckled coating acts as a cantilever. In this case, the indenter indents a cantilever rather than a coating/substrate system. This ultrathin coating cantilever has much less contact stiffness than the coating/substrate system. Therefore, the contact stiffness shows an abrupt decrease at the N_f . In the third stage, with increasing the number of cycles, the delaminated and buckled size increases, resulting in a further decrease in contact stiffness since the cantilever beam length increases. On the other hand, a high bending stress acts at the edge of the buckled coating. Larger the buckled size, higher is the bending stress. The cyclically bending stress causes fatigue damage at the end of the buckled coating, forming a ring-like crack. The coating under the indenter is separated from the bulk coating (caused by the ring-like crack at the edge of the buckled coating) and the substrate (caused by the delamination and buckling in the second stage). Therefore, the coating under the indenter is not constrained but is free to move with the indenter during fatigue testing. At this point, the sharp nature of the indenter is lost because the coating under the indenter gets stuck on the indenter. The indentation fatigue experiment results in the contact of a relative huge blunt tip with the substrate. This results in a low contact stiffness value.

Compressive residual stresses result in delamination and buckling. A coating with higher adhesion strength and a less compressive residual stress is required for a higher fatigue life. Interfacial defects should be avoided in coating deposition process. We know that the ring-like crack formation occurs in the coating. Formation of fatigue cracks in the coating depends upon the hardness and fracture toughness. Cracks are more difficult to form and propagate in the coating with higher strength and fracture toughness.

It is now accepted that long fatigue life in a coating/substrate almost always involves ‘living with crack,’ that the threshold or limit condition is associated with the non-propagation of exiting cracks or defects, even though these cracks may be undetectable (Suresh 1991). For all coatings studied, at 4 μN , contact stiffness does not change much. This indicates that delamination

and buckling did not occur within the number of cycles tested in this study. This is probably because the indentation induced compressive stress was not high enough to allow the cracks to propagate and link up under the indenter or the sum of indentation compressive stress σ_i and the residual stress σ_r did not exceed the critical buckling stress σ_b .

Figure 12.10 and Table 12.5 summarize the hardness, elastic modulus, fracture toughness, and fatigue life of all coatings studied. A good correlation exists between fatigue life and other mechanical properties. Higher mechanical properties result in a longer fatigue life. The mechanical properties of DLC coatings are controlled by the sp^3 -to- sp^2 ratio. The sp^3 -bonded carbon exhibits the outstanding properties of diamond (Bhushan 1999b). A higher deposition kinetic energy will result in a larger fraction of sp^3 -bonded carbon in an amorphous network. Thus, the higher kinetic energy for the FCA could be responsible for its better carbon structure and higher mechanical properties (Li and Bhushan 1999a, b; Sundararajan and Bhushan 1999, 2001). Higher adhesion strength between the FCA coating and substrate makes the FCA coating more difficult to delaminate from the substrate.

12.4.2 Microscratch and Microwear Studies

For microscratch studies, a conical diamond indenter (e.g. having a tip radius of about one micron and included angle of 60°) is drawn over the sample surface, and the load is increased steadily (typically from 2 to 25 mN), until substantial damage occurs. The coefficient of friction is monitored during scratching. Scratch-induced damage of coating, specifically fracture or delamination, can be monitored by in situ friction force measurements and by optical and SEM imaging of the scratches after tests. A gradual increase in friction is associated with plowing and abrupt increase in friction is associated with fracture or catastrophic failure (Bhushan and Gupta 1995). The load, corresponding to the

abrupt increase in friction or an increase in friction above a certain value (typically $2\times$ of initial value), provides a measure of scratch resistance or adhesive strength of a coating and is called “critical load”. Depth of scratches with increasing scratch length or normal load is measured by using an AFM, typically with an area of $10 \times 10 \mu\text{m}$ (Li and Bhushan 1998a, 1999a; Bhushan 2013a, b).

The microscratch and microwear studies are also conducted using an AFM (Bhushan and Koinkar 1995; Koinkar and Bhushan 1997; Bhushan 1999a; Sundararajan and Bhushan 1999, 2001). A square pyramidal diamond tip (tip radius $\sim 100 \text{ nm}$) or a three-sided pyramidal diamond (Berkovich) tip, with an apex angle of 60° and a tip radius of about 100 nm , mounted on a platinum-coated, rectangular stainless steel cantilever of stiffness of about 40 N/m , is scanned orthogonal to the long axis of the cantilever to generate scratch and wear marks. During the scratch test, the normal load is either kept constant or is increased (typically from 0 to $100 \mu\text{N}$), until damage occurs. Topography images of the scratch are obtained *in situ* with the AFM at a low load. By scanning the sample during scratching, wear experiments can be conducted. Wear at a constant load is monitored as a function of number of cycles. Normal loads ranging from 10 to $80 \mu\text{N}$ are typically used.

12.4.2.1 Microscratch

Scratch test conducted with a sharp diamond tip simulate a sharp asperity contact. In a scratch test, the cracking or delamination of a hard coating is signaled by a sudden increase in the coefficient of friction (Bhushan 1999a). The load associated with this event is called the “critical load.”

Wu (1991), Bhushan et al. (1995a), Gupta and Bhushan (1995a, b), and Li and Bhushan (1998a, 1999a, b) have used a nanoindenter to perform microscratch studies (mechanical durability) of various carbon coatings. The coefficient of friction profiles as a function of increasing normal load and AFM surface height maps of regions over scratches at the respective critical loads (indicated by the arrows in the friction profiles

and AFM images), made on the various coatings of different thicknesses and single-crystal silicon substrate using a conical tip are compared in Figs. 12.16 and 12.17. Bhushan and Koinkar (1995), Koinkar and Bhushan (1997), Bhushan (1999a), and Sundararajan and Bhushan (1999, 2001) have used an AFM to perform microscratch studies. Data for various coatings of different thicknesses and silicon substrate using a Berkovich tip are compared in Figs. 12.18 and 12.19. Critical loads for various coatings tested using nanoindenter and AFM are summarized in Fig. 12.20. The selected data for 20-nm thick coatings obtained using nanoindenter, are also presented in Fig. 12.10 and Table 12.5.

It can be seen that there exists a well-defined critical load for each coating. The AFM images clearly show that below the critical loads the coatings were plowed by the scratch tip, associated with the plastic flow of materials. At and after the critical loads, debris (chips) or buckling were observed on the sides of scratches. Delamination or buckling can be observed around or after the critical loads which suggests that the damage starts from delamination and buckling. For the 3.5 and 5 nm thick FCA coatings, before the critical loads small debris is observed on the sides of scratches. This suggests that the thinner FCA coatings are not so durable. It is obvious that for a given deposition method, the critical loads increase with increasing coating thickness. This indicates that the critical load is determined not only by the adhesion strength to the substrate but also by the coating thickness. We note that the debris generated on the thicker coatings is larger than that generated on the thinner coatings. For a thicker coating, it is more difficult to be broken, the broken coating chips (debris) for a thicker coating are larger than those for the thinner coatings. The difference in the residual stresses of the coatings of different thicknesses could also affect the size of debris. The AFM image shows that the silicon substrate was damaged by plowing, associated with the plastic flow of materials. At and after the critical load, small and uniform debris is observed and the amount of debris increases with increasing normal load.

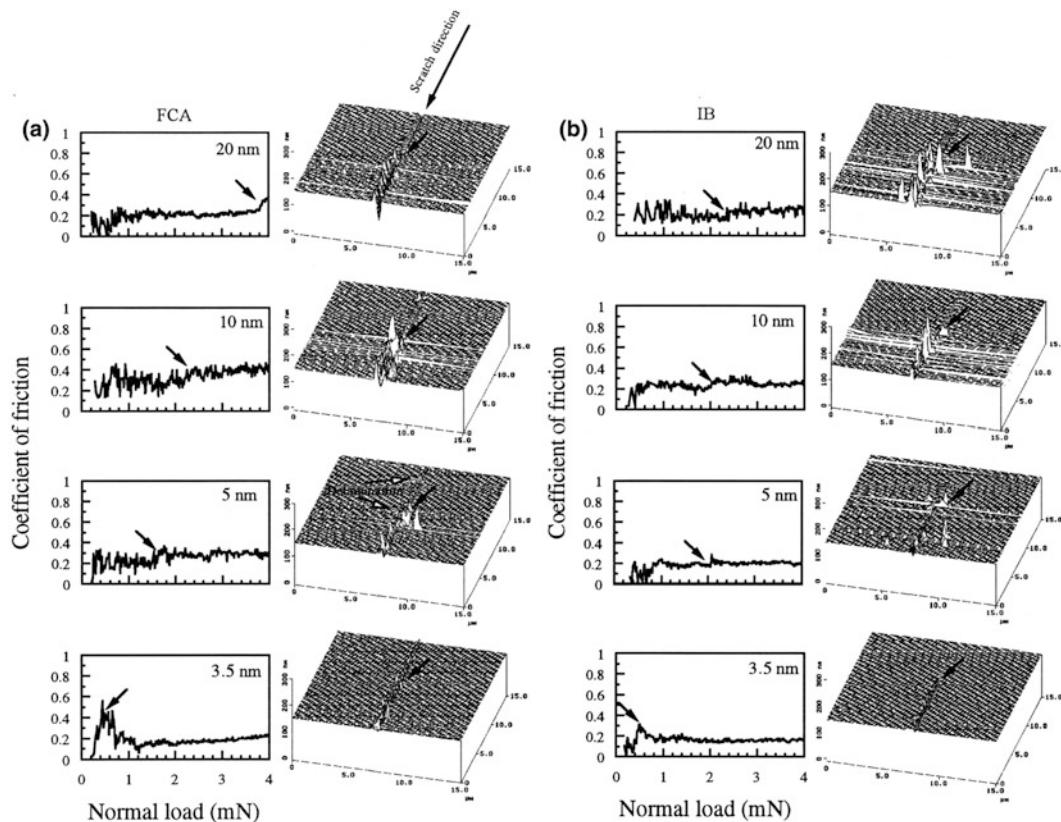


Fig. 12.16 Coefficient of friction profiles as a function of normal load and corresponding AFM surface height maps of regions over scratches at the respective critical loads (indicated by the arrows in the friction profiles and AFM images) for coatings of different thicknesses deposited by various deposition techniques: **a** FCA, **b** IB, **c** ECR-CVD, **d** SP

Since, at the critical load, the damage mechanism appears to be onset of plowing, higher hardness and fracture toughness of a coating will therefore result in higher load required for deformation and hence higher critical load. Figure 12.21 shows critical loads of the various coatings, obtained with AFM tests, as a function of the coating hardness and fracture toughness (from Table 12.5). It can be seen that, in general, higher coating hardness and fracture toughness result in higher critical load. The only exceptions are the FCA coatings at 5- and 3.5-nm coating thickness, which show the lowest critical loads despite their high hardness and fracture toughness. The brittleness of the thinner FCA coatings may be one reason for their low critical loads. The mechanical properties of coatings that are

less than 10-nm thick are unknown. The FCA process may result in coatings with low hardness at such low thickness due to differences in coating stoichiometry and structure as compared to the coatings of higher thickness. Also, at these thicknesses stresses at the coating-substrate interface may affect adhesion and load-carrying capacity of the coatings.

Based on the experimental results, a schematic of scratch damage mechanisms of the DLC coatings used in this study is shown in Fig. 12.22. Below the critical load, if a coating has a good combination of strength and fracture toughness, plowing associated with the plastic flow of materials is responsible for the damage of coating (Fig. 12.22a). Whereas, if a coating has a lower fracture toughness, cracking could occur

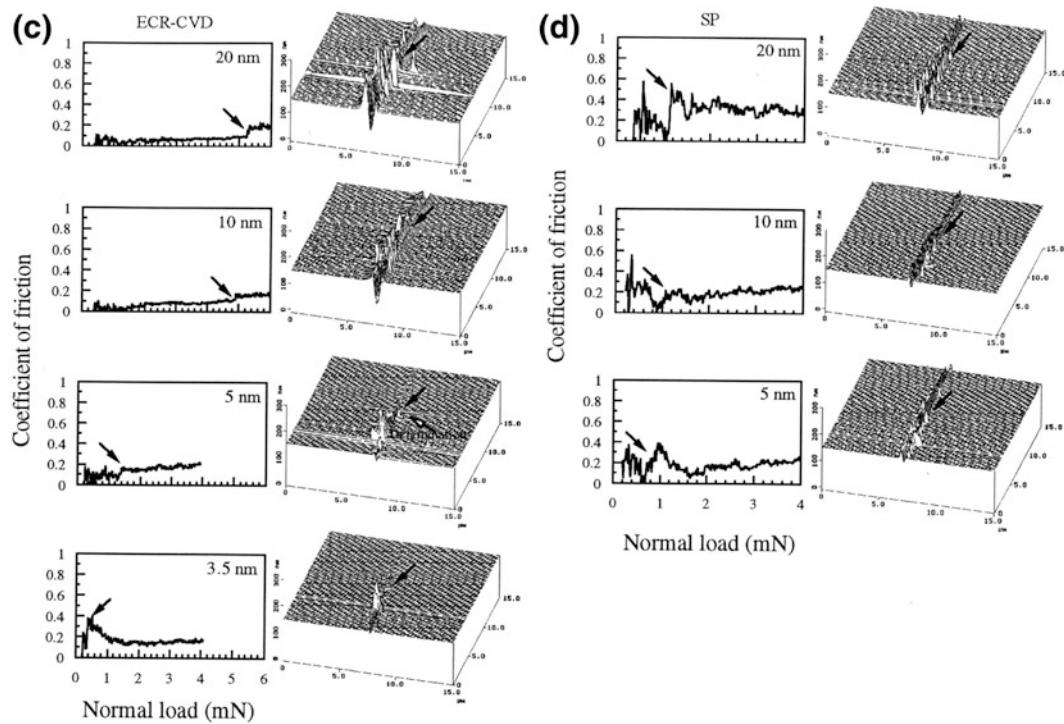


Fig. 12.16 (continued)

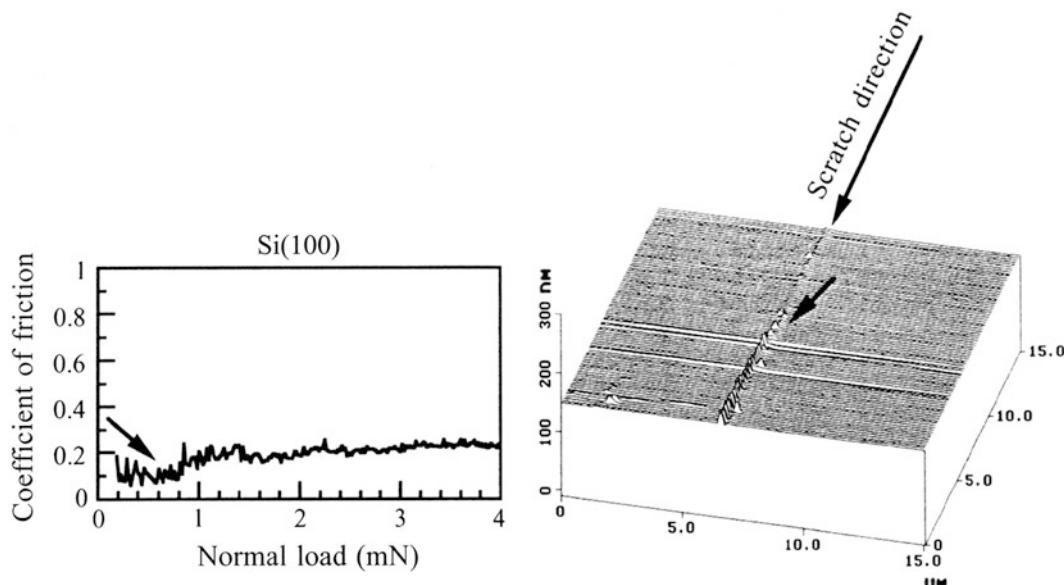


Fig. 12.17 Coefficient of friction profiles as a function of normal load and corresponding AFM surface height maps of regions over scratches at the respective critical loads (indicated by the *arrows* in the friction profiles and AFM images) for Si(100)

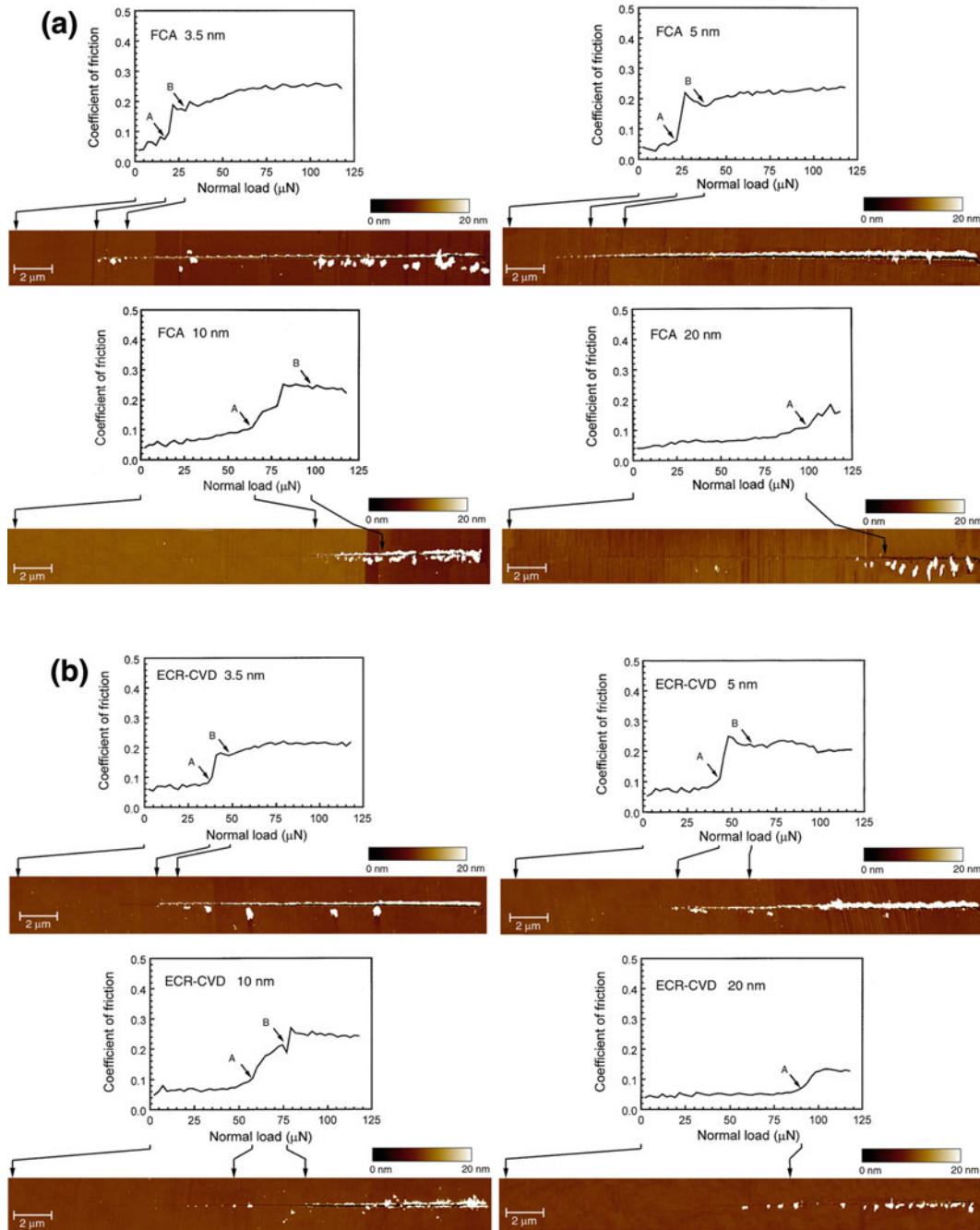


Fig. 12.18 Coefficient of friction profiles during scratch as a function of normal load and corresponding AFM surface height maps for **a** FCA, **b** ECR-CVD, and **c** SP coatings (Sundararajan and Bhushan 2001)

during plowing, associated with the formation of small debris (Fig. 12.22b). When normal load increases up to the critical load, delamination or

buckling will occur at the coating/substrate interface (Fig. 12.22c). Further increase in normal load will result in breakdown of coating via

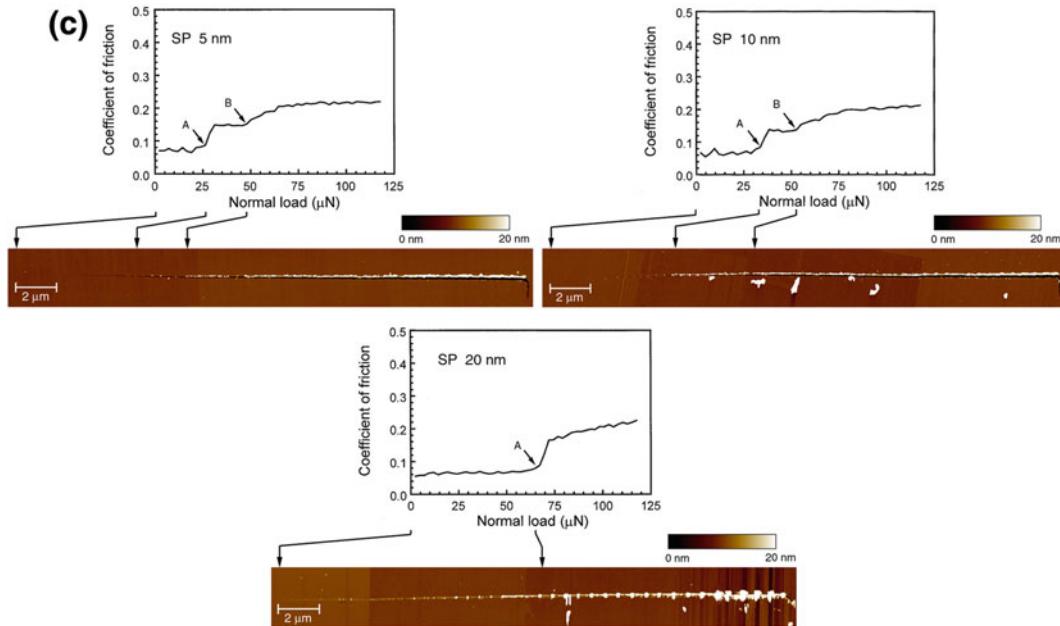
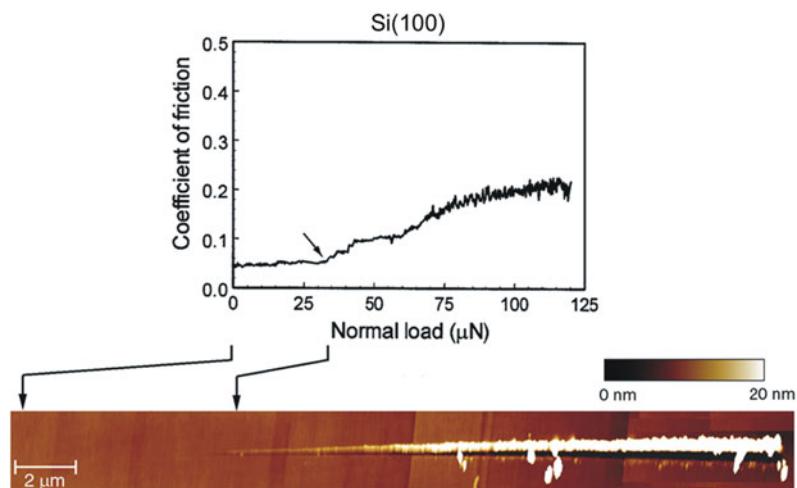


Fig. 12.18 (continued)

Fig. 12.19 Coefficient of friction profiles during scratch as a function of normal load and corresponding AFM surface height maps for Si (100) (Sundararajan and Bhushan 2001)



through coating thickness cracking, as shown in Fig. 12.22d. Therefore, adhesion strength plays a crucial role in the determination of critical load. If a coating has stronger adhesive strength with the substrate, the coating is more difficult to be delaminated, which will result in a higher critical load. The interfacial and residual stresses of a coating could greatly affect the delamination and

buckling (Bhushan and Gupta 1991). The coating with higher interfacial and residual stresses is more easily delaminated and buckled, which will result in a low critical load. It has been reported earlier that the FCA coatings have higher residual stresses than the other coatings (Gupta and Bhushan 1995b). Interfacial stresses play a more important role when a coating gets thinner.

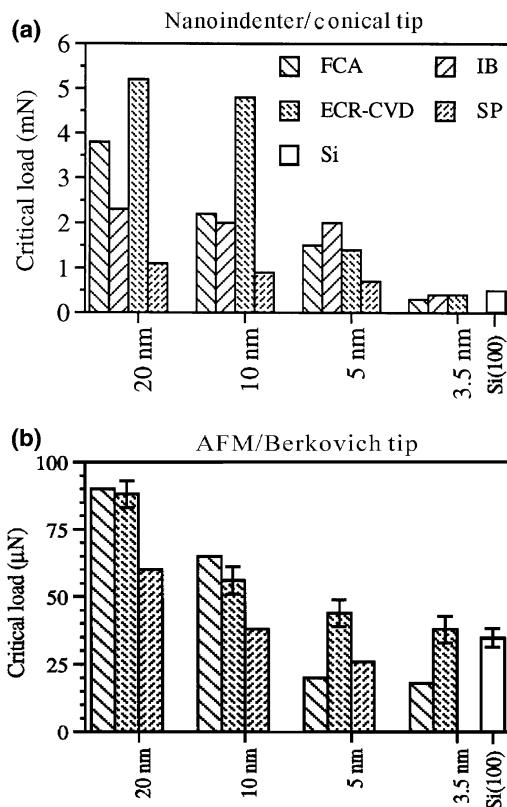


Fig. 12.20 Critical loads estimated from the coefficient of friction profiles from **a** nanoindenter and **b** AFM tests for various coatings of different thicknesses and Si(100) substrate

A large mismatch in elastic modulus between the FCA coatings and the silicon substrate may cause large interfacial stresses. This may be why the thinner FCA coatings show relatively low critical loads compared with the thicker FCA coatings, even though the FCA coatings have higher hardness and elastic modulus. The brittleness of the thinner FCA coatings may be another reason for the lower critical loads. The strength and fracture toughness of a coating also affect the critical load. Higher strength and fracture toughness will make the coating more difficult to be broken after delamination and buckling. The high scratch resistance/adhesion of FCA coatings is attributed to an atomic intermixing at the coating-substrate interface because of high kinetic energy (2 keV) plasma formed during the cathodic arc deposition process (Anders et al.

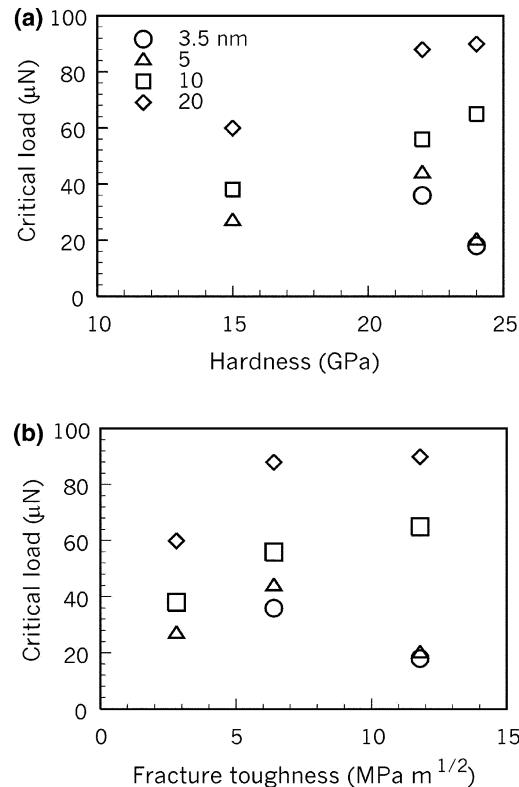
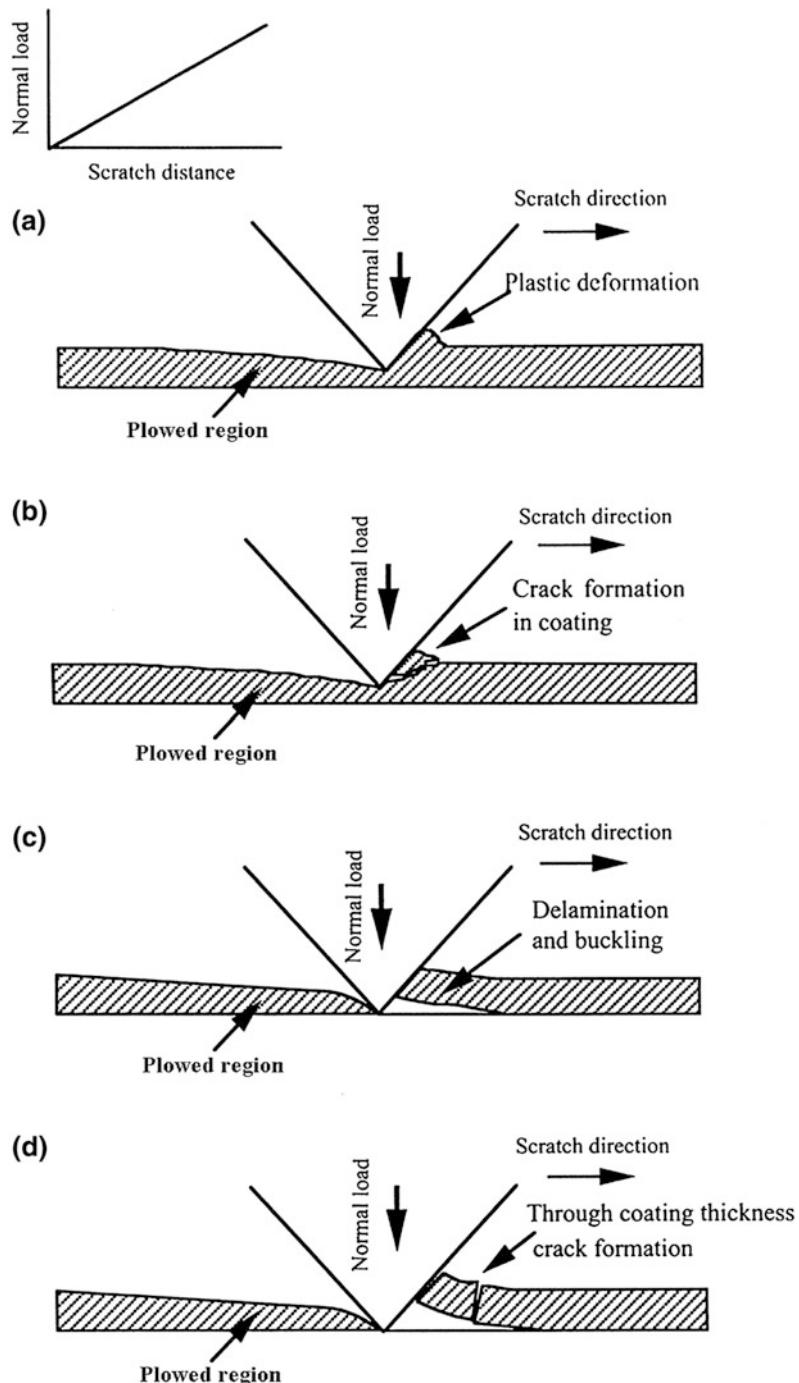


Fig. 12.21 Measured critical loads estimated from the coefficient of friction profiles from AFM tests as a function of **a** coating hardness and **b** fracture toughness. Coating hardness and fracture toughness values were obtained using a nanoindenter on 100-nm thick coatings (Table 12.5)

1994a). The atomic intermixing at the interface provides a graded compositional transition between the coating and the substrate materials. In all other coatings used in this study, the kinetic energy of the plasma was insufficient for atomic intermixing.

Gupta and Bhushan (1995a, b) and Li and Bhushan (1999a, b) measured scratch resistance of DLC coatings deposited on Al₂O₃-TiC, Ni-Zn ferrite and single-crystal silicon substrates. For good adhesion of DLC coating to other substrates, in most cases, an interlayer of silicon is required except for cathodic arc deposited coatings. The best adhesion with cathodic arc carbon coating is obtained on electrically conducting substrates such as Al₂O₃-TiC and silicon as compared to Ni-Zn ferrite.

Fig. 12.22 Schematic of scratch damage mechanisms of the DLC coatings,
a plowing associated with the plastic flow of materials,
b plowing associated with the formation of small debris,
c delamination and buckling at the critical load, and
d breakdown via through coating thickness cracking at and after the critical load (Li and Bhushan 1999a)



12.4.2.2 Microwear

Microwear studies can be conducted using an AFM (Bhushan 1999a). For microwear studies, a three-sided pyramidal single-crystal natural

diamond tip with an apex angle of about 80° and a tip radius of about 100 nm is used at relatively high loads of 1–150 μN . The diamond tip is mounted on a stainless steel cantilever beam with

a normal stiffness of about 30 N/m. The sample is generally scanned in a direction orthogonal to the long axis of the cantilever beam (typically at a rate of 0.5 Hz). The tip is mounted on the beam such that one of its edge is orthogonal to the beam axis. For wear studies, typically an area of $2 \mu\text{m} \times 2 \mu\text{m}$ is scanned for a selected number of cycles.

Microwear studies of various types of DLC coatings have been conducted (Bhushan and Koinkar 1995; Koinkar and Bhushan 1997; Sundararajan and Bhushan 1999). Figure 12.23a shows a wear mark on uncoated Si(100). Wear occurs uniformly and material is removed layer by layer via plowing from the first cycle, resulting in the constant friction force seen during the wear (Fig. 12.24a). Figure 12.23b shows AFM images of the wear marks on all 10 nm

coatings. It is seen that coatings wear nonuniformly. Coating failure is sudden and accompanied by a sudden rise in the friction force (Fig. 12.24b). Figure 12.24 shows the wear depth of Si(100) substrate and various DLC coatings at two different loads. FCA and ECR-CVD, 20-nm thick coatings show excellent wear resistance up to 80 μN , the load that is required for the IB 20 nm coating to fail. In these tests, “failure” of a coating results when the wear depth exceeds the quoted coating thickness. The SP 20 nm coating fails at the much lower load of 35 μN . At 60 μN , the coating hardly provides any protection. Moving onto the 10 nm coatings, ECR-CVD coating requires about 45 cycles at 60 μN to fail as compared to IB and FCA which fail at 45 μN . The FCA coating exhibits slight roughening in the wear track after

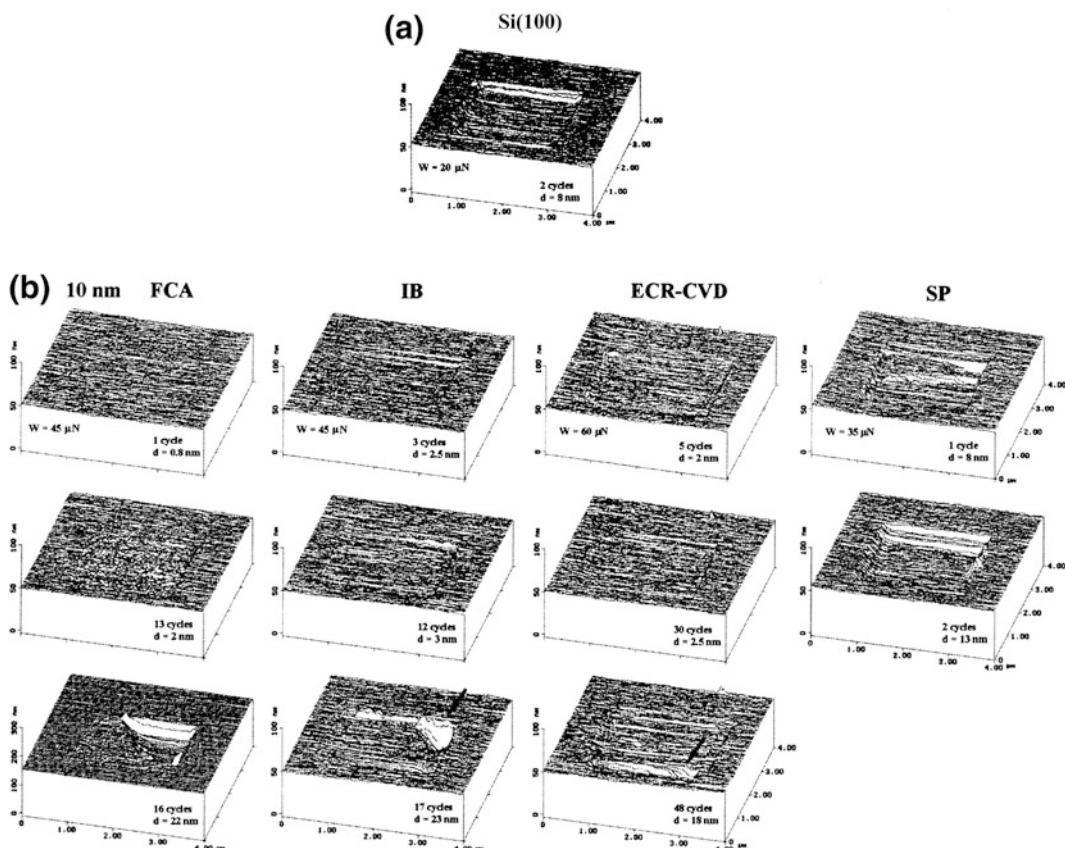


Fig. 12.23 AFM images of wear marks on **a** bare Si(100), and **b** all 10-nm thick DLC coatings (Sundararajan and Bhushan 1999)

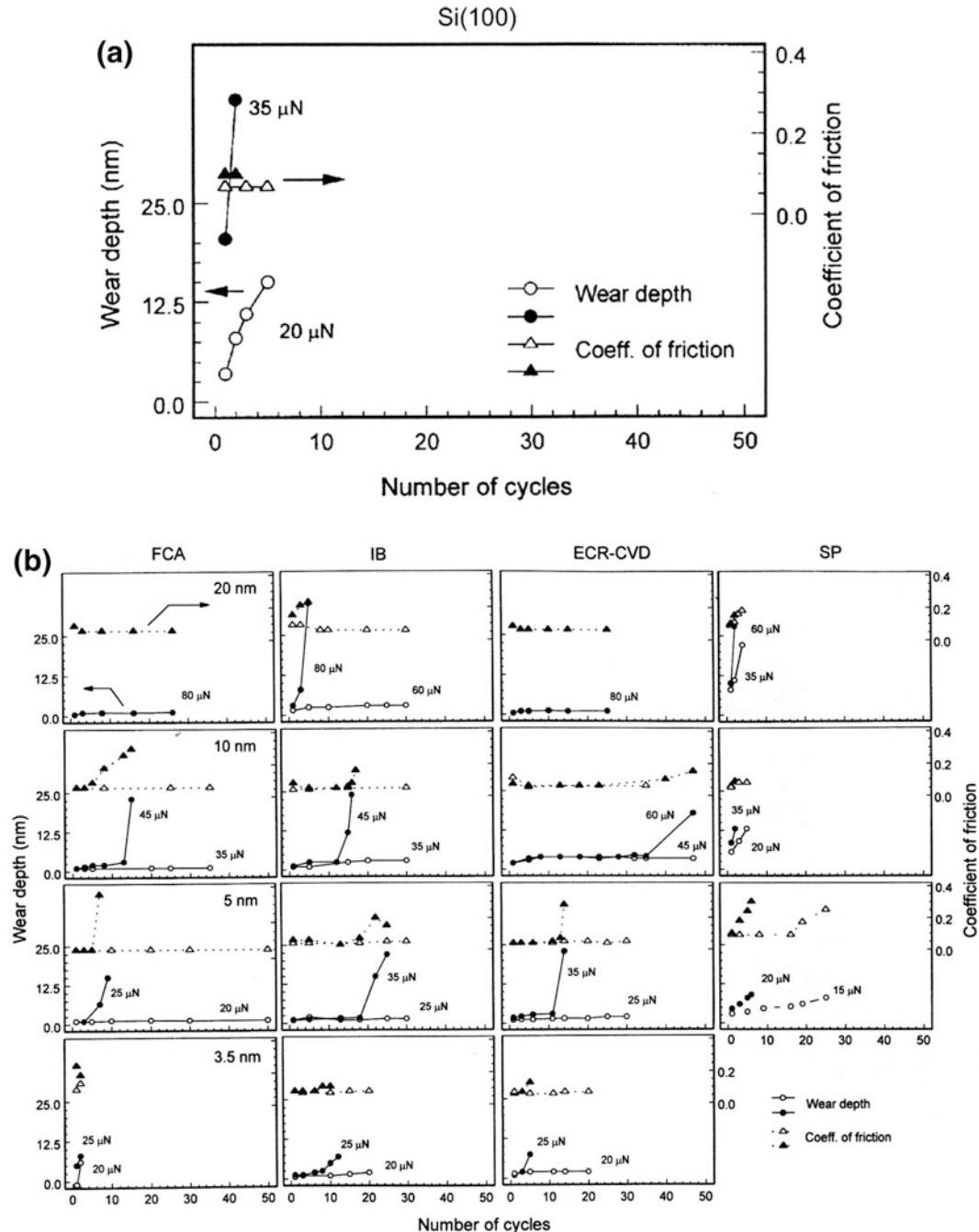


Fig. 12.24 Wear data of **a** bare Si(100) and **b** all DLC coatings. Coating thickness is constant along each row in **(b)**. Both wear depth and coefficient of friction during wear for a given cycle are plotted (Sundararajan and Bhushan 1999)

the first few cycles, which leads to an increase in the friction force. The SP coating continues to exhibit poor resistance, failing at 20 μN . For the

5 nm coatings, the load required to fail the coatings continue to decrease. But IB and ECR-CVD still provide adequate protection as

compared to bare Si(100) in that order, failing at 35 μN compared to FCA at 25 μN and SP at 20 μN . Almost all the 20, 10 and 5 nm coatings provide better wear resistance than bare silicon. At 3.5 nm, FCA coating provides no wear resistance, failing almost instantly at 20 μN . The IB and ECR-CVD coating show good wear resistance at 20 μN compared to bare Si(100). But IB lasts only about 10 cycles and ECR-CVD about 3 cycles at 25 μN .

The wear tests highlight the differences in the coatings more vividly than the scratch tests. At higher thicknesses (20 and 10 nm), the ECR-CVD and FCA coating appear to show the best wear resistance. This is probably due to higher hardness of the coatings (see Table 12.5). At 5 nm, IB coating appears to be the best. FCA coatings show poorer wear resistance with decreasing coating thickness. This suggests that the trends in hardness seen in Table 12.5 no longer hold at low thicknesses. SP coatings showed consistently poor wear resistance at all thicknesses. The IB 3.5 nm coating does provide reasonable wear protection at low loads.

12.4.3 Macroscale Tribological Characterization

So far data on mechanical characterization and microscratch and microwear studies using a nanoindenter and an AFM have been presented. Mechanical properties affect tribological performance of the coatings and microwear studies simulate a single asperity contact which helps in the development of fundamental understanding of the wear process. These studies are useful in screening various candidates as well as to understand the relationships between deposition conditions and properties of various samples. As a next step, macroscale friction and wear tests need to be conducted to measure tribological performance of the coatings.

Macroscale accelerated friction and wear tests to screen a large number of candidates and functional tests on selected candidates have been conducted. An accelerated test is designed to accelerate the wear process such that it does not

change the failure mechanism. The accelerated friction and wear tests are generally conducted using a ball-on-flat tribometer under reciprocating motion (Bhushan et al. 1995a). Typically, a diamond tip with 20 μm tip radius or a sapphire ball with 3 mm diameter and surface finish of about 2 nm RMS is slid against the coated substrates at selected loads. Coefficient of friction is monitored during the tests.

Functional tests are conducted using an actual machine under close to actual operating conditions for which coatings is developed. Generally, the tests are accelerated somewhat to fail the interface in a short time.

12.4.3.1 Accelerated Friction and Wear Tests

Li and Bhushan (1999a) conducted accelerated friction and wear tests on DLC coatings deposited by various deposition techniques, using a ball-on-flat tribometer. Average values of coefficient of friction are presented in Table 12.5. The optical micrographs of wear tracks and debris formed on all samples when slid against a sapphire ball after sliding distance of 5 m are presented in Fig. 12.25. The normal load used for the 20 and 10 nm thick coatings is 200 mN and the normal load used for the 5 and 3.5 nm thick coatings and silicon substrate is 150 mN.

Among the 20 nm thick coatings, the SP coating exhibits a higher coefficient of friction of about 0.3 than the other coatings which have comparable values of coefficient of friction of about 0.2. The optical micrographs show that the SP coating has a larger wear track and more debris than the IB coatings. No wear track and debris were found on the 20 nm thick FCA and ECR-CVD coatings. The optical micrographs of 10-nm thick coatings show that the SP coating was severely damaged showing a large wear track with scratches and lots of debris. The FCA and ECR-CVD coatings show smaller wear tracks and less debris than the IB coatings.

For the 5 nm thick coatings, the wear tracks and debris of the IB and ECR-CVD coatings are comparable. The bad wear resistance of the 5 nm thick FCA coating is in good agreement with the low scratch critical load, which may be due to the

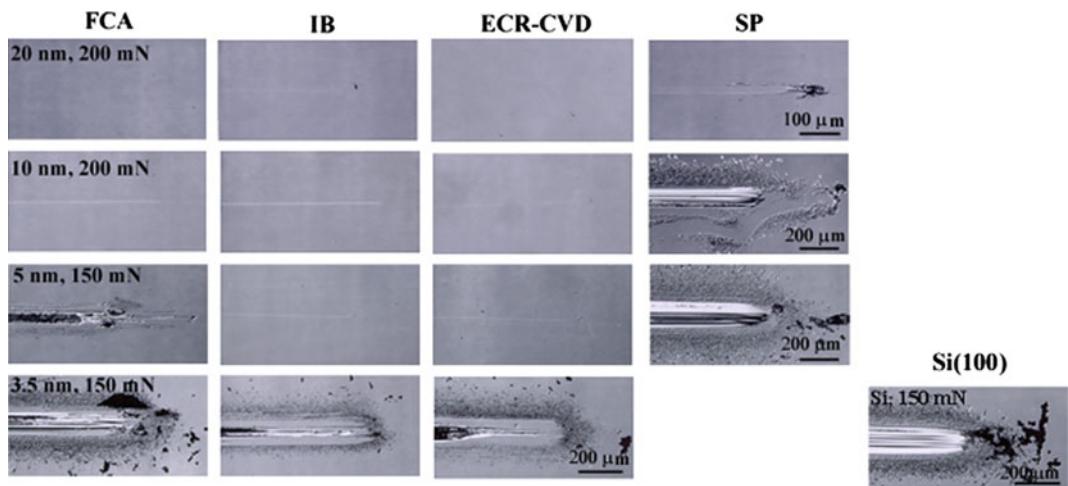


Fig. 12.25 Optical micrographs of wear tracks and debris formed on the various coatings of different thicknesses and silicon substrate when slid against a sapphire ball after sliding distance of 5 m. The end of the wear track is on the right hand side of the image

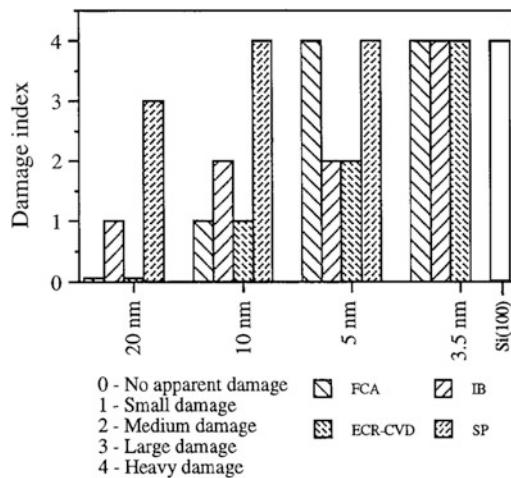


Fig. 12.26 Wear damage index bar chart of the various coatings of different thicknesses and Si(100) substrate based on optical examination of the wear tracks and debris

higher interfacial and residual stresses as well as brittleness of the coating.

At 3.5 nm, all coatings exhibit wear. The FCA coating provides no wear resistance, failing instantly like the silicon substrate. Large block-like debris is observed on the sides of the wear track of the FCA coating. This indicates that large region delamination and buckling

occurred during sliding, resulting in large block-like debris. These block-like debris, in turn, scratched the coating, making the coating damage even more severe. The IB and ECR-CVD coatings are able to provide some protection against wear at 3.5 nm.

In order to better evaluate the wear resistance of various coatings, based on the optical examination of the wear tracks and debris after tests, the wear damage index bar chart of the various coatings of different thicknesses and an uncoated silicon substrate is presented in Fig. 12.26. Among the 20 and 10 nm thick coatings, the SP coatings show the worst damage, followed by FCA/ECR-CVD. At 5 nm, the FCA and SP coatings show the worst damage, followed by IB, and ECR-CVD coatings. All 3.5 nm thick coatings show as the same heavy damage as the uncoated silicon substrate.

The wear damage mechanisms of thick and thin DLC coatings studied are believed to be as illustrated in Fig. 12.27. At the early stages of sliding, deformation zone and Hertzian and wear fatigue cracks formed beneath the surface extend within the coating upon subsequent sliding (Bhushan and Gupta 1991). Formation of fatigue cracks depend upon the hardness and subsequent cycles. These are controlled by the sp^3 -to- sp^2

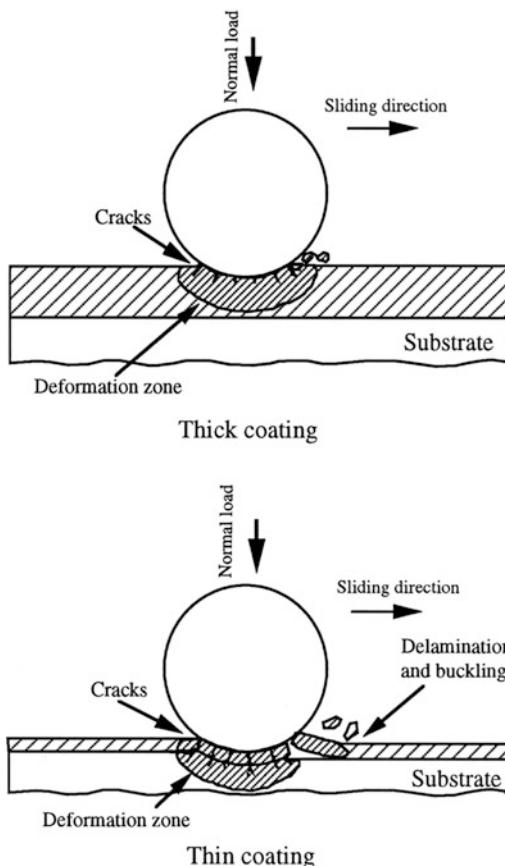


Fig. 12.27 Schematic of wear damage mechanisms of thick and thin DLC coatings (Li and Bhushan 1999a)

ratio. For thicker coating, the cracks generally do not penetrate the coating. While for a thinner coating, the cracks easily propagate down to the interface with the aid of the interfacial stresses, they get diverted along the interface just enough to cause local delamination of the coating. When this happens, the coating experiences excessive plowing. At this point, the coating fails catastrophically, resulting in a sudden rise in the coefficient of friction. All 3.5 nm thick coatings, failed much quicker as compared to the thicker coatings. It appears that these thin coating have very low load-carrying capacity and therefore the substrate undergoes deformation almost immediately. This generates stresses at the interface that weaken the coating adhesion and lead to delamination of the coating. Another reason may be that the thickness

is insufficient to produce a coating comprised of the DLC structure. Instead the bulk may be made up of a matrix characteristic of the interface region where atomic mixing occurs with the substrate and/or any interlayer used. This would also result in poor wear resistance and silicon-like behavior of the coating, especially in the case of FCA coatings, which show the worst performance at 3.5 nm. In contrast to the other coatings, all SP coatings show the worst wear performance at any thicknesses (Fig. 12.25). This may be due to their poor mechanical properties, such as lower hardness and scratch resistance compared to the other coatings.

Comparison of Figs. 12.20 and 12.26 shows a very good correlation between the wear damage and scratch critical loads. Less wear damage corresponds to higher scratch critical load. Based on the data, thicker coatings do show better scratch and wear resistance than thinner coatings. This is probably due to better load-carrying capacity of the thick coatings as compared to the thinner ones. For a given coating thickness, higher hardness and fracture toughness and better adhesion strength are believed to be responsible for the superior wear performance.

12.4.3.2 Effect of Environment

Friction and wear performance of amorphous carbon coatings are found to be strongly dependent on the water vapor content and partial gas pressure in the test environment. The friction data for an amorphous carbon film on the silicon substrate sliding against steel are presented as a function of the partial pressure of water vapor in Fig. 12.28 (Dimigen and Hubsch 1983–1984; Memming et al. 1986; Bhushan 1996a; Bhushan and Gupta 1991; Donnet et al. 1998). Friction increases dramatically above a relatively humidity of about 40%. At high relatively humidity, condensed water vapor forms meniscus bridges at the contacting asperities and the menisci result in an intrinsic attractive force which is responsible for an increase in friction. For completeness, the data for the coefficient of friction of bulk graphitic carbon are also presented in Fig. 12.28. Note that friction decreases

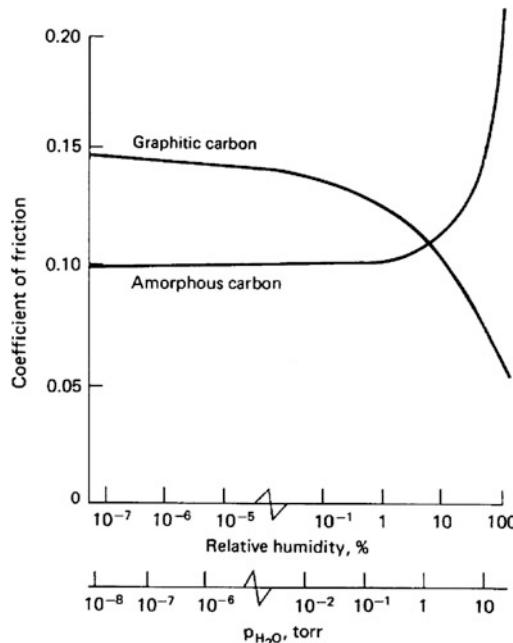


Fig. 12.28 Coefficient of friction as a function of relative humidity and water vapor partial pressure of RF-plasma deposited amorphous carbon coating and bulk graphitic carbon coating sliding against a steel ball

with an increase in the relative humidity (Bowden and Young 1951). Graphitic carbon has a layered lattice crystal structure. Graphite absorbs polar gases (H_2O , O_2 , CO_2 , NH_3 , etc.) at the edges of the crystallites, which weakens the interlayer bonding forces facilitating interlayer slip and results in lower friction (Bhushan and Gupta 1991).

To better study the effect of environment for carbon coated magnetic disks, a number of tests have been conducted in controlled environments. Marchon et al. (1990) conducted tests in alternating environments of oxygen and nitrogen gases, Fig. 12.29. The coefficient of friction increases as soon as oxygen is added to test environment whereas in a nitrogen environment the coefficient of friction reduces slightly. Tribochemical oxidation of the DLC coating in the oxidizing environment is responsible for an increase in the coefficient of friction implying wear. Dugger et al. (1990), Strom et al. (1991), Bhushan and Ruan (1994), and Bhushan et al. (1995b) conducted tests with DLC coated

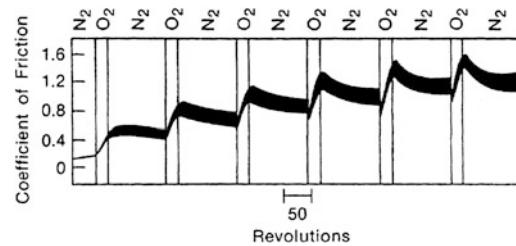


Fig. 12.29 Coefficient of friction as a function of sliding distance for a ceramic slider against a magnetic disk coated with 20-nm thick DC magnetron sputtered DLC coating, measured at a speed of 0.06 m/s and 10 g load. The environment is alternated between oxygen and nitrogen gases (Marchon et al. 1990)

magnetic disks (with about 2-nm thick of perfluoropolyether lubricant film) in contact with Al_2O_3 -TiC sliders, in different gaseous environments including a high vacuum of 2×10^{-7} Torr, Fig. 12.30. The wear lives are shortest in high vacuum and the longest in most atmospheres of nitrogen and argon with the following order (from best to worst): argon or nitrogen, $Ar + H_2O$, ambient, $Ar + O_2$, and $Ar + H_2O$, vacuum. From this sequence of wear performance, we can see that having oxygen and water in an operating environment worsen wear performance of the coatings, but having nothing in it (vacuum) is the worst of all. Indeed, failure

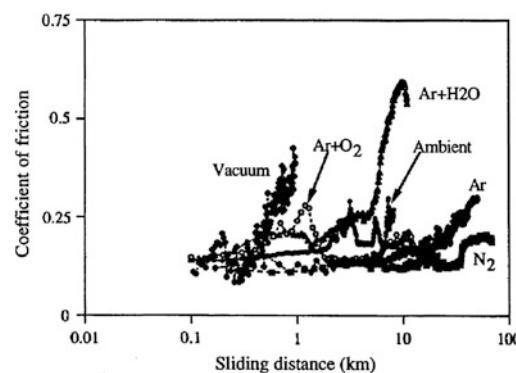


Fig. 12.30 Durability measured by sliding a Al_2O_3 -TiC magnetic slider against a magnetic disk coated with 20-nm thick DC sputtered amorphous carbon coating and 2-nm thick perfluoropolyether film, measured at a speed of 0.75 m/s and 10 g load. Vacuum refers to 2×10^{-7} Torr (Bhushan et al. 1995b)

mechanisms differ in various environments. In high vacuum, intimate contact between disk and slider surfaces results in significant wear. In ambient air, Ar + O₂, and Ar + H₂O, tribocorrosion oxidation of the carbon overcoat is responsible for interface failure. For experiments performed in pure argon and nitrogen, mechanical shearing of the asperities cause the formation of debris which is responsible for the formation of scratch marks on the carbon surface as could be observed with an optical microscope (Bhushan et al. 1995b).

12.4.3.3 Functional Tests

Magnetic thin-film heads made with Al₂O₃-TiC substrate are used in magnetic storage applications (Bhushan 1996a). Multilayered thin-film pole-tip structure present on the head surface wears more rapidly than the Al₂O₃-TiC substrate which is much harder. Pole-tip recession (PTR) is a serious concern in magnetic storage (Bhushan et al. 1996, 1997; Xu and Bhushan 1998; Scott and Bhushan 2000a, b; Scott et al. 2000). Two of the diamondlike carbon coatings superior in mechanical properties—ion beam and cathodic arc carbon—were deposited on the air bearing surfaces of Al₂O₃-TiC head sliders (Bhushan et al. 1996). The functional tests were conducted by running a metal-particle (MP) tape in a computer tape drive. Average PTR as a function of sliding distance data are presented in Fig. 12.31. We note that PTR increases for the uncoated head, whereas for the coated heads there is a slight increase in PTR in early sliding followed by little change. Thus, coatings provide protection.

Micromechanical and accelerated and functional tribological data presented here clearly suggest that there is a good correlation between the scratch resistance and wear resistance measured using accelerated tests and functional tests. Thus, scratch tests can be successfully used to screen coatings for wear applications.

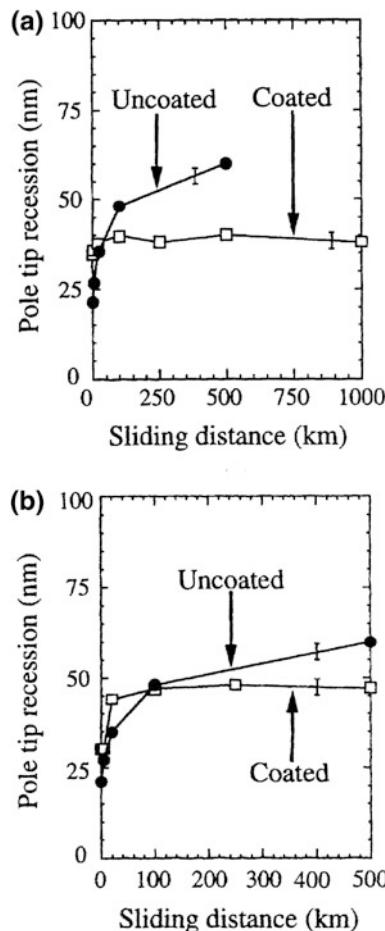


Fig. 12.31 Pole tip recession as a function of sliding distance as measured with an AFM for **a** uncoated and 20-nm thick ion beam carbon coated, and **b** uncoated and 20-nm thick cathodic arc carbon coated Al₂O₃-TiC heads run against MP tapes (Bhushan et al. 1996)

12.4.4 Coating Continuity Analysis

Ultrathin coatings less than 10 nm, may not uniformly coat the sample surface. In other words, coating may be discontinuous and deposited in the form of islands on the micro-scale. A possible reason for poor wear protection and the non-uniform failure of the coatings may

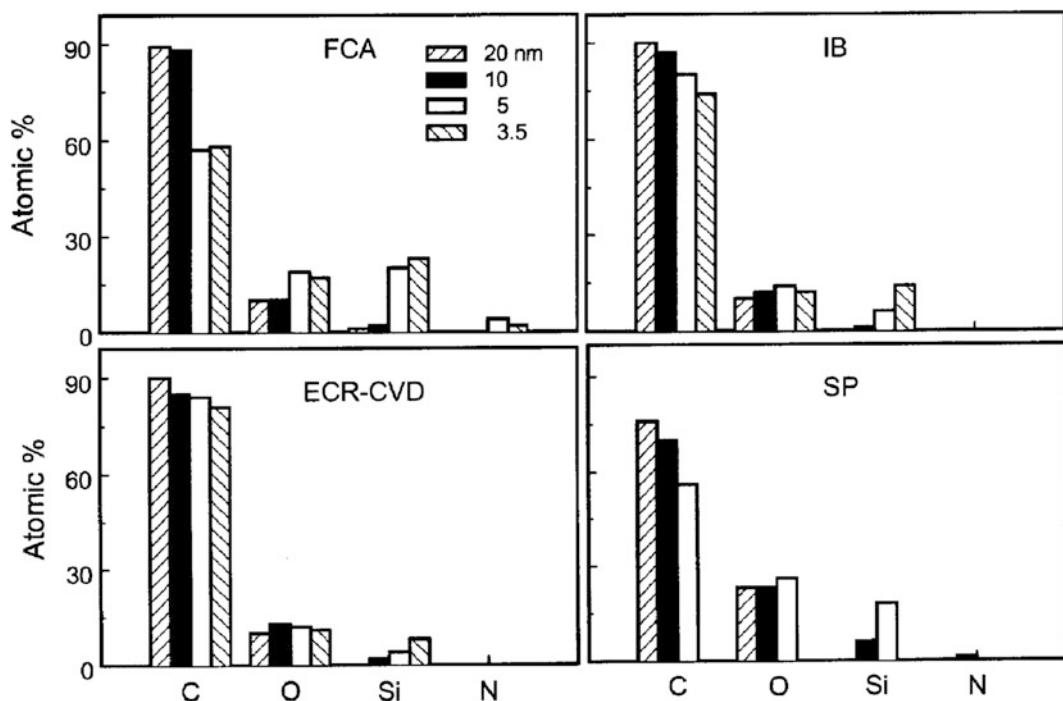


Fig. 12.32 Quantified XPS data for various DLC coatings on Si(100) substrate (Sundararajan and Bhushan 1999). Atomic concentrations are shown

be due to poor coverage of the thin coatings on the substrate. Coating continuity can be studied by using surface analytical techniques such as Auger and/or XPS analyses. Any discontinuity with coating thickness less than sampling depth of the instrument will detect locally the substrate species (Bhushan and Koinkar 1995; Li and Bhushan 1999b; Sundararajan and Bhushan 1999).

The results of XPS analysis on various coatings, deposited on Si(100) substrates, over a 1.3 mm^2 region (single point measurement with spot diameter $1300 \mu\text{m}$) are shown in Fig. 12.32. The sampling depth is about 2–3 nm. The poor SP coatings show much less carbon content (<75% atomic concentration) as does the poor 5 and 3.5 nm FCA coatings (<60%) as compared to the IB and ECR-CVD coatings. Silicon is detected in all 5 nm coatings. From the data it is hard to say if the Si is from the substrate or from exposed regions due to coating discontinuity.

Based on the sampling depth any Si detected in 3.5 nm coatings would likely be from the substrate. The other interesting observation is that all poor coatings (all SP and FCA 5 and 3.5 nm) show almost twice the oxygen content than that of the other coatings. Any oxygen present may be due to leaks in the deposition chamber and is present in the form of silicon oxides.

AES measurements, averaged over a scan area of $900 \mu\text{m}^2$, were conducted on FCA and SP 5 nm coatings at 6 different regions on each sample. Very little silicon was detected on this scale and the detected peaks were characteristic of oxides. The oxygen levels were comparable to that seen for the good coatings by XPS. These contrast to the XPS measurements at a larger scale suggest that the coating possesses discontinuities only at isolated areas and that the 5 nm coatings are generally continuous on the micro-scale. Figure 12.33 shows representative XPS and AES spectra of selected samples.

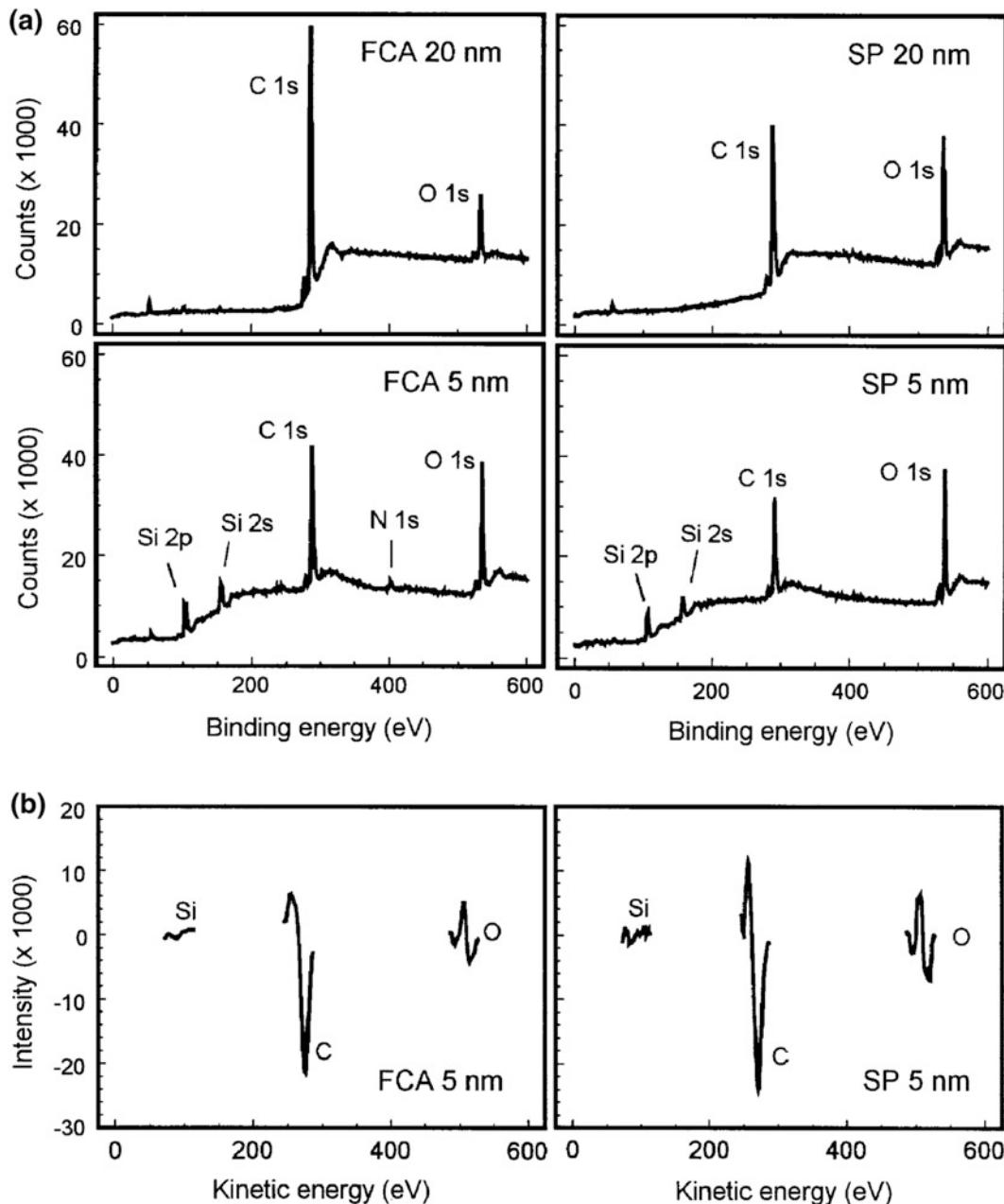


Fig. 12.33 a XPS spectra for FCA and SP coatings of 5 nm and 20 nm coating thicknesses on Si(100) substrate, and b AES spectra for FCA and SP coatings at 5 nm thickness on Si(100) substrate (Sundararajan and Bhushan 1999)

12.5 Closure

Diamond material and its smooth coatings exhibit relatively low wear and friction. Major limitations of true diamond coatings are that they need to be

deposited at high temperatures, can only be deposited on selected substrates, and require surface finishing. Mechanical, thermal, and optical properties of hard amorphous carbon (a-C), commonly known as DLC coatings, can be tailored to be similar to diamond. These can be deposited

with a large range of thicknesses using a variety of deposition processes on variety of substrates at or near room temperature. The coatings reproduce substrate topography, avoiding the need for post finishing. Friction and wear properties of some DLC coatings can be very attractive for tribological applications. The largest industrial application of these coatings is in magnetic storage devices. They are also being used in MEMS/NEMS.

EELS and Raman spectroscopies can be used successfully for chemical characterization of amorphous carbon coatings. The prevailing atomic arrangement in DLC coatings is amorphous or quasi-amorphous with small diamond (sp^3), graphite (sp^2), and other unidentifiable micro- or nanocrystallites. Most DLC coatings, except those produced by filtered cathodic arc, contain from a few to about 50 at. % hydrogen. Sometimes hydrogen deliberately is incorporated in the sputtered and ion plated coatings to tailor their properties.

Amorphous carbon coatings deposited by various techniques exhibit different mechanical and tribological properties. The nanoindenter can be used successfully for measurement of hardness, elastic modulus, fracture toughness, and fatigue life. Microscratch and microwear experiments can be performed using either a nanoindenter or an AFM. Thin coatings deposited by filtered cathodic arc, ion beam and ECR-CVD hold promise for tribological applications. Coatings as thin as 5 nm or even thinner provide wear protection. Microscratch, microwear, and accelerated wear testing, if simulated properly, can be used successfully to screen coating candidates for industrial applications. In the examples shown in this chapter, trends observed in the microscratch, microwear, and accelerated macrofriction wear tests are similar to those found in functional tests.

References

- Ager JW, Veirs DK, Rosenblatt CM (1991) Spatially resolved Raman studies of diamond films grown by chemical vapor deposition. *Phys Rev B* 43:6491–6499
- Aisenberg S, Chabot R (1971) Ion beam deposition of thin films of diamond like carbon. *J Appl Phys* 49:2953–2958
- Aksenen II, Strel'Nitskii VE (1991) Wear resistance of diamond-like carbon coatings. *Surf Coat Technol* 47:252–256
- Anders S, Anders A, Brown IG, Wei B, Komvopoulos K, Ager JW III, Yu KM (1994a) Effect of vacuum arc deposition parameters on the properties of amorphous carbon thin films. *Surf Coat Technol* 68–69:388–393
- Anders S, Anders A, Brown IG, Dickinson MR, MacGill RA (1994b) Metal plasma immersion ion implantation and deposition using arc plasma sources. *J Vac Sci Technol B12:815–820*
- Anders S, Anders A, Brown IG (1995) Transport of vacuum arc plasma through magnetic macroparticle filters. *Plasma Sources Sci 4:1–12*
- Andersson LP (1981) A review of recent work on hard i-C films. *Thin Solid Films* 86:193–200
- Angus JC (1992) Diamond and diamondlike films. *Thin Solid Films* 216:126–133
- Angus JC, Hayman CC (1988) Low pressure metastable growth of diamond and diamondlike phase. *Science* 241:913–921
- Angus JC, Jensen F (1988) Dense diamondlike hydrocarbons as random covalent networks. *J Vac Sci Technol A6:1778–1782*
- Angus JC, Koidl P, Domitz S (1986) Carbon Thin Films. In: Mort J, Jensen F (eds) *Plasma deposited thin films*. CRC Press, Boca Raton, pp 89–127
- Asmussen J (1989) Electron cyclotron resonance microwave discharges for etching and thin-film deposition. *J Vac Sci Technol A7:883–893*
- Banks BA, Rutledge SK (1982) Ion beam sputter deposited diamond like films. *J Vac Sci Technol B1:807–814*
- Bhushan B (1996a) Tribology and mechanics of magnetic storage devices, second edn. Springer, New York
- Bhushan B (1996b) Nanotribology and nanomechanics of MEMS devices. In: Proceedings of ninth annual workshop on micro electro mechanical systems (IEEE, New York), pp 91–98
- Bhushan B (ed) (1998) Tribology issues and opportunities in MEMS. Kluwer Academic Pub, Dordrecht
- Bhushan B (1999a) Handbook of micro/nanotribology, 2nd edn. CRC, Boca Raton
- Bhushan B (1999b) Chemical, mechanical, and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments. *Diam Relat Mater* 8:1985–2015
- Bhushan B (2000) Mechanics and reliability of flexible magnetic media, 2nd edn. Springer, New York
- Bhushan B (2001a) Macro- and microtribology of magnetic storage devices. In: Bhushan B (ed) *Modern tribology handbook*, vol 2—materials, coatings, and industrial applications. CRC Press, Boca Raton, pp 1413–1513
- Bhushan B (2001b) Macro- and microtribology of MEMS materials. In: Bhushan B (ed) *Modern tribology handbook*, vol 2—materials, coatings, and industrial applications. CRC Press, Boca Raton, pp 1515–1548
- Bhushan B (2013a) Principles and applications of tribology, 2nd edn. Wiley, New York

- Bhushan B (2013b) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2016) Encyclopedia of nanotechnology, 2nd edn. Springer International, Switzerland
- Bhushan B, Doerner MF (1989) Role of mechanical properties and surface texture in the real area of contact of magnetic rigid disks. *ASME J Tribol* 111:452–458
- Bhushan B, Gupta BK (1991) Handbook of tribology: materials, coatings, and surface treatments. McGraw Hill, New York. Reprint ed., Krieger, Malabar, Florida, 1997
- Bhushan B, Gupta BK (1995) Micromechanical characterization of Ni-P coated aluminum-magnesium, glass and glass-ceramic substrates and finished magnetic thin-film rigid disks. *Adv Info Storage Syst* 6:193–208
- Bhushan B, Koinkar VN (1995) Microscale mechanical and tribological characterization of hard amorphous coatings as thin as 5 nm for magnetic disks. *Surf Coatings Technol* 76–77:655–669
- Bhushan B, Li X (2003) Nanomechanical characterization of solid surfaces and thin films. *Int Mater Rev* 48:125–164
- Bhushan B, Ruan J (1994) Tribological performance of thin film amorphous carbon overcoats for magnetic recording rigid disks in various environments. *Surf Coatings Technol* 68(69):644–650
- Bhushan B, Kellock AJ, Cho NH, Ager JW III (1992) Characterization of chemical bonding and physical characteristic of diamond-like amorphous carbon and diamond films. *J Mater Res* 7:404–410
- Bhushan B, Gupta BK, VanCleef GW, Capp C, Coe JV (1993a) Fullerene (C₆₀) films for solid lubrication. *Tribol Trans* 36:573–580
- Bhushan B, Subramaniam VV, Malshe A, Gupta BK, Ruan J (1993b) Tribological properties of polished diamond films. *J Appl Phys* 74:4174–4180
- Bhushan B, Gupta BK, Subramaniam VV (1994) Polishing of diamond films. *Diam Films Technol* 4:71–97
- Bhushan B, Gupta BK, Azarian MH (1995a) Nanoindentation, microscratch, friction and wear studies for contact recording applications. *Wear* 181–183:743–758
- Bhushan B, Yang L, Gao C, Suri S, Miller RA, Marchon B (1995b) Friction and wear studies of magnetic thin-film rigid disks with glass-ceramic, glass and aluminum-magnesium substrates. *Wear* 190:44–59
- Bhushan B, Patton ST, Sundaram R, Dey S (1996) Pole tip recession studies of hard carbon-coated thin-film tape heads. *J Appl Phys* 79:5916–5918
- Bhushan B, Theunissen GSAM, Li X (1997) Tribological studies of chromium oxide films for magnetic recording applications. *Thin Solid Films* 311:67–80
- Bowden FP, Young JE (1951) Friction of diamond, graphite and carbon and the influence of surface films. *Proc R Soc Lond* 208:444–455
- Brown IG, Anders A, Anders S, Dickinson MR, Ivanov IC, MacGill RA, Yao XY, Yu KM (1993) Plasma synthesis of metallic and composite thin films with atomically mixed substrate bonding. *Nucl Instrum Meth Phys Res B* 80–B81:1281–1287
- Bubenzier A, Dischler B, Brandt B, Koidl P (1983) R.F. plasma deposited amorphous hydrogenated hard carbon thin films, preparation, properties and applications. *J Appl Phys* 54:4590–4594
- Bull SJ (1995) Tribology of carbon coatings: DLC, diamond and beyond. *Diam Relat Mater* 4:827–836
- Catherine Y (1991) Preparation techniques for diamond-like carbon. In: Clauzing RE, Horton LL, Angus JC, Koidl P (eds) Diamond and diamond-like films and coatings. Plenum, New York, pp 193–227
- Cho NH, Krishnan KM, Veirs DK, Rubin MD, Hopper CB, Bhushan B, Bogy DB (1990) Chemical structure and physical properties of diamond-like amorphous carbon films prepared by magnetron sputtering. *J Mater Res* 5:2543–2554
- Cuomo JJ, Pappas DL, Bruley J, Doyle JP, Seigner KL (1991) Vapor deposition processes for amorphous carbon films with sp³ fractions approaching diamond. *J Appl Phys* 70:1706–1711
- Cuomo JJ, Pappas DL, Lossy R, Doyle JP, Bruley J, Di Bello GW, Krakow W (1992) Energetic carbon deposition at oblique angles. *J Vac Sci Technol A* 10:3414–3418
- Dimigen H, Hubsch H (1983–1984) Applying low-friction wear-resistant thin solid films by physical vapor deposition. *Philips Tech Rev* 41:186–197
- Dischler B, Bubenzier A, Koidl P (1983) Hard carbon coatings with low optical-absorption. *Appl Phys Lett* 42:636–638
- Donnet C, Grill A (1997) Friction control of diamond-like carbon coatings. *Surf Coat Technol* 94–95:456
- Donnet C, Le Mogne T, Ponsonnet L, Belin M, Grill A, Patel V (1998) The respective role of oxygen and water vapor on the tribology of hydrogenated diamond-like carbon coatings. *Tribol Lett* 4:259
- Dugger MT, Chung YW, Bhushan B, Rothschild W (1990) Friction, wear, and interfacial chemistry in thin film magnetic rigid disk files. *ASME J Tribol* 112:238–245
- Erdemir A, Donnet C (2001) Tribology of diamond, diamond-like carbon, and related films. In: Bhushan B (ed) Modern tribology handbook, vol 2—materials, coatings, and industrial applications. CRC Press, Boca Raton, pp 871–908
- Erdemir A, Switala M, Wei R, Wilbur P (1991) A tribological investigation of the graphite-to-diamond-like behavior of amorphous carbon films ion beam deposited on ceramic substrates. *Surf Coat Technol* 50:17–23
- Erdemir A, Nicols FA, Pan XZ, Wei R, Wilbur PJ (1993) Friction and wear performance of ion-beam deposited diamond-like carbon films on steel substrates. *Diam Relat Mat* 3:119–125
- Evans AG, Hutchinson JW (1984) On the mechanics of delamination and spalling in compressed films. *Int J Solids Structures* 20:455–466
- Fallon PJ, Veerasamy VS, Davis CA, Robertson J, Amaralunga GAJ, Milne WI, Koskinen J (1993)

- Properties of filtered-ion-beam-deposited diamond-like carbon as a function of ion energy. *Phys Rev B* 48:4777–4782
- Green DC, McKenzie DR, Lukins PB (1989) The microstructure of carbon thin films. *Mater Sci Forum* 52–53:103–124
- Grill A (1997) Tribological properties of diamondlike carbon and related materials. *Surf Coat Technol* 94–95:507
- Grill A, Meyerson BS (1994) Development and status of diamondlike carbon. In: Spear KE, Dismukes JP (eds) *Synthetic diamond: emerging CVD science and technology*. Wiley, New York, pp 91–141
- Grill A, Meyerson BS, Patel VV (1988) Interface modification for improving the adhesion of a-C: H to metals. *J Mater Res* 3:214
- Grill A, Meyerson BS, Patel VV (1990a) Diamond-like carbon films by RF plasma-assisted chemical vapor deposition from acetylene. *IBM J Res Develop* 34:849–857
- Grill A, Patel VV, Meyerson BS (1990b) Optical and tribological properties of heat-treated diamond-like carbon. *J Mater Res* 5:2531–2537
- Gupta BK, Bhushan B (1994) Fullerene particles as an additive to liquid lubricants and greases for low friction and wear. *Lub Eng* 50:524–528
- Gupta BK, Bhushan B (1995a) Mechanical and tribological properties of hard carbon coatings for magnetic recording heads. *Wear* 190:110–122
- Gupta BK, Bhushan B (1995b) Micromechanical properties of amorphous carbon coatings deposited by different deposition techniques. *Thin Solid Films* 270:391–398
- Gupta BK, Bhushan B, Capp C, Coe JV (1994) Material characterization and effect of purity and ion implantation on the friction and wear of sublimed fullerene films. *J Mater Res* 9:2823–2838
- Holland L, Ojha SM (1976) Deposition of hard and insulating carbonaceous films of an RF target in butane plasma. *Thin Solid Films* 38:L17–L19
- Jansen F, Machonkin M, Kaplan S, Hark S (1985) The effect of hydrogenation on the properties of ion beam sputter deposited amorphous carbon. *J Vac Sci Technol A3*:605–609
- Kaplan S, Jansen F, Machonkin M (1985) Characterization of amorphous carbon-hydrogen films by solid-state nuclear magnetic resonance. *Appl Phys Lett* 47:750–753
- Knight DS, White WB (1989) Characterization of diamond films by Raman spectroscopy. *J Mater Res* 4:385–393
- Koinkar VN, Bhushan B (1997) Microtribological properties of hard amorphous carbon protective coatings for thin-film magnetic disks and heads. *Proc Inst Mech Eng J* 211:365–372
- Lettington AH (1998) Applications of diamond-like carbon thin films. *Carbon* 36:555–560
- Li X, Bhushan B (1998a) Micromechanical and tribological characterization of hard amorphous carbon coatings as thin as 5 nm for magnetic recording heads. *Wear* 220:51–58
- Li X, Bhushan B (1998b) Measurement of fracture toughness of ultra-thin amorphous carbon films. *Thin Solid Films* 315:214–221
- Li X, Bhushan B (1999a) Micro/nanomechanical and tribological characterization of ultra-thin amorphous carbon coatings. *J Mater Res* 14:2328–2337
- Li X, Bhushan B (1999b) Mechanical and tribological studies of ultra-thin hard carbon overcoats for magnetic recording heads. *Z Metallkd* 90:820–830
- Li X, Bhushan B (1999c) Evaluation of fracture toughness of ultra-thin amorphous carbon coatings deposited by different deposition techniques. *Thin Solid Films* 355–356:330–336
- Li X, Bhushan B (2002a) Development of a nanoscale fatigue measurement technique and its application to ultrathin amorphous carbon coatings. *Scripta Mater* 47:473–479
- Li X, Bhushan B (2002b) Nanofatigue studies of ultrathin hard carbon overcoats used in magnetic storage devices. *J Appl Phys* 91:8334–8336
- Li X, Diao D, Bhushan B (1997) Fracture mechanisms of thin amorphous carbon films in nanoindentation. *Acta Mater* 45:4453–4461
- Lossy R, Pappas DL, Roy RA, Cuomo JJ (1992) Filtered arc deposition of amorphous diamond. *Appl Phys Lett* 61:171–173
- Marchon B, Salmeron M, Siekhaus W (1989) Observation of graphitic and amorphous structures on the surface of hard carbon films by scanning tunneling microscopy. *Phys Rev B* 39:12907–12910
- Marchon B, Heiman N, Khan MR (1990) Evidence for tribochemical wear on amorphous carbon thin films. *IEEE Trans Magn* 26:168–170
- Marshall DB, Evans AG (1984) Measurement of adherence of residual stresses in thin films by indentation. I. Mechanics of interface delamination. *J Appl Phys* 51:2632–2638
- Matthews A, Eskildsen SS (1994) Engineering applications for diamond-like carbon. *Diam Relat Mat* 3:902–911
- McKenzie DR, Muller D, Pailthorpe BA, Wang ZH, Kravtchinskaia E, Segal D, Lukins PB, Martin PJ, Amarutunga G, Gaskell PH, Saeed A (1991) Properties of tetrahedral amorphous carbon prepared by vacuum arc deposition. *Diam Relat Mater* 1:51–59
- Memming R, Tolle HJ, Wierenga PE (1986) Properties of polymeric layers of hydrogenated amorphous carbon produced by plasma-activated chemical vapor deposition: tribological and mechanical properties. *Thin Solid Films* 143:31–41
- Pappas DL, Saegner KL, Bruley J, Krakow W, Cuomo JJ (1992) Pulsed laser deposition of diamondlike carbon films. *J Appl Phys* 71:5675–5684
- Robertson J (1986) Amorphous Carbon. *Adv Phys* 35:317–374
- Robertson J (1992) Properties of diamond-like carbon. *Surf Coat Technol* 50:185–203

- Robertson J (1993) Deposition of diamond-like carbon. *Philos Trans R Soc Lond A* 342:277–286
- Rubin M, Hooper CB, Cho NH, Bhushan B (1990) Optical and mechanical properties of DC sputtered carbon films. *J Mater Res* 5:2538–2542
- Sander P, Kaiser U, Altebockwinkel M, Wiedmann L, Benninghoven A, Sah RE, Koidl P (1987) Depth profile analysis of hydrogenated carbon layers on silicon by X-ray photoelectron spectroscopy, auger electron spectroscopy, electron energy-loss spectroscopy, and secondary ion mass spectrometry. *J Vac Sci Technol A5*:1470–1473
- Savvides N, Bell TJ (1992) Microhardness and young's modulus of diamond and diamondlike carbon films. *J Appl Phys* 72:2791–2796
- Savvides N, Window B (1985) Diamondlike amorphous carbon films prepared by magnetron sputtering of graphite. *J Vac Sci Technol A3*:2386–2390
- Scharff W, Hammer K, Stenzel O, Ullman J, Vogel M, Fraueneheim T, Eibisch B, Roth S, Schulze S, Muelling I (1989) Preparation of amorphous i-C films by ion-assisted methods. *Thin Solid Films* 171:157–169
- Scheibe HJ, Schultrich B (1994) DLC film deposition by laser-arc and study of properties. *Thin Solid Films* 246:92–102
- Scott WW, Bhushan B (2000a) Corrosion and wear studies of uncoated and ultra-thin DLC coated magnetic tape-write heads and magnetic tapes. *Wear* 243:31–42
- Scott WW, Bhushan B (2000b) Loose debris and head stain generation and pole tip recession in modern tape drives. *J Info Storage Proc Syst* 2:221–254
- Scott WW, Bhushan B, Lakshminikumaran AV (2000) Ultrathin diamond-like carbon coatings used for reduction of pole tip recession in magnetic tape heads. *J Appl Phys* 87:6182–6184
- Spencer EG, Schmidt PH, Joy DC, Sansalone FJ (1976) Ion beam deposited polycrystalline diamond-like films. *Appl Phys Lett* 29:118–120
- Strom BD, Bogy DB, Bhatia CS, Bhushan B (1991) Tribocorrosion effects of various gases and water vapor on thin film magnetic disks with carbon overcoats. *ASME J Tribol* 113:689–693
- Sundararajan S, Bhushan B (1999) Micro/nanotribology of ultra-thin hard amorphous carbon coatings using atomic force/friction force microscopy. *Wear* 225–229:678–689
- Sundararajan S, Bhushan B (2001) Development of a continuous microscratch technique in an atomic force microscope and its application to study scratch resistance of ultrathin hard amorphous carbon coatings. *J Mater Res* 16:437–445
- Suresh S (1991) Fatigue of materials. Cambridge University Press, Cambridge
- Suzuki J, Okada S (1995) Deposition of diamondlike carbon films using electron cyclotron resonance plasma chemical vapor deposition from ethylene gas. *Jpn J Appl Phys* 34:L1218–L1220
- Swec DM, Mirtich MJ, Banks BA (1988) Ion beam and plasma methods of producing diamondlike carbon films. Report No. NASATM102301, NASA, Cleveland, Ohio
- Tsai HC, Bogy DB, Kundmann MK, Veirs DK, Hilton MR, Mayer ST (1988) Structure and properties of sputtered carbon overcoats on rigid magnetic media disks. *J Vac Sci Technol A6*:2307–2315
- Vandentop GJ, Kawasaki M, Nix RM, Brown IG, Salmeron M, Somorjai GA (1990) Formation of hydrogenated amorphous carbon films of controlled hardness from a methane plasma. *Phys Rev B* 41:3200–3210
- Wei R, Wilbur PJ, Liston MJ (1993) Effects of diamond-like hydrocarbon films on rolling contact fatigue of bearing steels. *Diam Relat Mat* 2:898–903
- Weissmantel C, Bewilogua K, Breuer K, Dietrich D, Ebersbach U, Erler HJ, Rau B, Reisse G (1982) Preparation and properties of hard i-C and i-BN coatings. *Thin Solid Films* 96:31–44
- Wu TW (1991) Microscratch and load relaxation tests for ultra-thin films. *J Mater Res* 6:407–426
- Xu J, Bhushan B (1998) Pole tip recession studies of thin-film rigid disk head sliders II: effects of air bearing surface and pole tip region designs and carbon coating. *Wear* 219:30–41

Self-assembled Monolayers (SAMs) for Nanotribology and Surface Protection

13

Bharat Bhushan

Abstract

In many applications, hydrophobic and lubricating films are needed to minimize adhesion, stiction, friction, and wear. In various applications, surfaces need to be protected from exposure to the operating environment, and hydrophobic films are of interest. The self-assembled monolayers are molecularly thick, well-organized, and chemically bonded to the substrate. Some can be insensitive to environment and can provide lower adhesion, friction, and wear. Ordered molecular assemblies with high hydrophobicity and desirable nanotribological properties can be engineered using chemical grafting of various polymer molecules with suitable functional head groups, spacer chains, and nonpolar surface terminal groups. In this chapter, we present an overview of self-assembled monolayers (SAMs) with hydrophobicity and desirable nanotribological properties. SAMs are produced by various organic precursors. We first present a primer to organic chemistry, followed by an overview of selected SAMs with various substrates, spacer chains, and terminal groups in the molecular chains, and an overview of nanotribological properties of SAMs. The contact angle, adhesion, friction, and wear properties of SAMs having various spacer chains with different surface terminal and head groups (hexadecane thiol, biphenyl thiol, perfluoroalkylsilane, alkylsilane, perfluoroalkylphosphonate, and alkylphosphonate) and on various substrates (Au, Si, and Al) are surveyed. Chemical degradation mechanisms

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: bhushan.2@osu.edu

and environmental effects are presented. Based on the contact angle and nanotribological properties of various SAM films by atomic force microscopy (AFM), perfluoroalkylsilane and perfluorophosphonate SAMs exhibit attractive hydrophobic and nanotribological properties.

Keywords

Self-assembled monolayers • Molecularly-thick films • Organized films • Chemically bonded films • Friction • Wear • Lubrication • Nanotribology

13.1 Introduction

The reliability concerns for engineering devices require the use of hydrophobic and lubricating films to minimize adhesion, stiction, friction, and wear (Bhushan et al. 1995a; Bhushan 1996, 1998a, 1999a, 2007, 2008; Man et al. 1998; Tanner et al. 2000). In various applications, surfaces need to be protected from exposure to the operating environment. Examples include micro/nanoelectromechanical systems (MEMS/NEMS) and BioMEMS/BioNEMS, requiring relative motion, and magnetic storage devices such as magnetic rigid disk and tape drives. In some biomedical applications, such as biosensors and implantable biomedical devices, undesirable protein adsorption, biofouling, and biocompatibility are some of the major issues (Bhushan 2007, 2008, 2016). In micro- and nanofluidic based sensors, the fluid drag in micro- and nanochannels can be reduced by using hydrophobic coatings.

13.1.1 Need for Hydrophobic Surfaces for Nanotribology

The source of the liquid film at the interface can be either a pre-existing film of liquid and/or capillary condensates of water vapor from the environment. If the liquid wets the surface ($0 \leq \theta < 90^\circ$, where θ is the contact angle¹ between

the liquid-vapor interface and the liquid-solid interface for a liquid droplet sitting on a solid surface, Fig. 13.1a), the liquid surface is thereby constrained to lie parallel with the surface (Adamson 1990; Israelachvili 1992; Schrader and Loeb 1992), and the complete liquid surface therefore is concave in shape, Fig. 13.1b. Surface tension results in a pressure difference across any meniscus surface, referred to as capillary pressure or Laplace pressure, and is negative for a concave meniscus (Bhushan 2013a, b). The negative Laplace pressure results in an intrinsic attractive (adhesive) force which depends on the interface roughness (local geometry of interacting asperities and number of asperities), the surface tension, and the contact angle. During normal separation, this intrinsic force needs to be overcome (Cai and Bhushan 2008; Bhushan 2013a, b). During sliding, frictional effects need to be overcome, not only because of external load but also because of intrinsic adhesive force. Measured value of high static friction force contributed largely by liquid mediated adhesion (meniscus contribution) is generally referred to as “stiction.” It becomes a major concern in micro- and nanodevices operating at ultralow loads as the liquid mediated adhesive force may be on the same order as the external load. The effect of liquid-mediated adhesion can be minimized by increasing surface roughness and/or the use of a liquid with low surface tension with a

(Footnote 1 continued)

by aligning a tangent with the drop profile at the point of contact with the solid surface using a telescope equipped with a goniometer eyepiece.

¹The direct measurement of contact angle is most widely made from sessile drops. The angle is generally measured

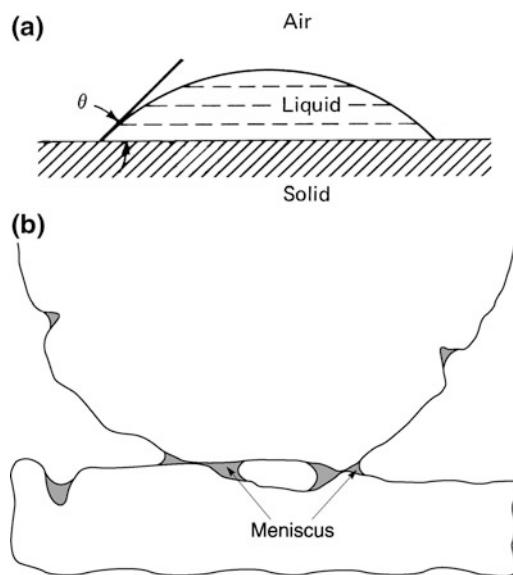


Fig. 13.1 **a** Schematic of a sessile-drop on a solid surface and the definition of contact angle, and **b** formation of meniscus bridges as a result of liquid present at an interface

film thickness on the order of surface roughness, as well as its chemical bonding to the substrate (Bhushan 1998b, 2001, 2003, 2013a, b; Bhushan and Zhao 1999; Bhushan and Peng 2002; Cai and Bhushan 2008). The formation of menisci and/or condensation of water vapor from the environment at the interface can be minimized by the use of hydrophobic (water-fearing) coatings. The surfaces can be made superhydrophobic by the introduction of controlled roughness on the surfaces to take advantage of so-called Lotus effect (Bhushan and Jung 2008; Nosonovsky and Bhushan 2008; Bhushan 2016).

13.1.2 Surface Films for Nanotribology and Surface Protection

Surfaces can be treated or coated with a liquid with relatively low surface tension or certain solid films to make them hydrophobic and/or to control adhesion, stiction, friction, and wear.

The classical approach to lubrication uses freely supported multimolecular layers of liquid

lubricants (Bowden and Tabor 1950; Zisman 1959; Bhushan 1996, 1999a, b, 2013a, b; Koinkar and Bhushan 1996; Bhushan and Zhao 1999; Liu and Bhushan 2003a). Boundary lubricant films are formed by either physisorption, chemisorption, or chemical reaction. The physisorbed films can be either monomolecularly or polymolecularly thick. The chemisorbed films are monomolecular, but stoichiometric films formed by chemical reaction can be multilayers. In general, stability and durability of surface films decrease in the following order: chemically reacted films, chemisorbed films, and physisorbed films. A good boundary lubricant should have a high degree of interaction between its molecules and the sliding surface. As a general rule, liquids will have a more desirable performance when they are polar and thus able to grip on solid surfaces (or be adsorbed). Polar lubricants contain reactive functional end groups. Boundary lubrication properties are also dependent upon the molecular conformation and lubricant spreading. It should be noted that the liquid films with the thickness on the order of a few nm, may be discontinuous and may deposit in an island form with nonuniform thickness with a lateral resolution on the nm scale.

Solid films are also commonly used for controlling hydrophobicity and/or adhesion, stiction, friction, and wear. Hydrophobic films have non-polar surface terminal groups (to be described later) which repel water. These films have low surface energy (15–30 dyn/cm) and a high contact angle ($\theta \geq 90^\circ$) which minimize wetting (e.g., Zisman 1959; Ulman 1991, 1996). Multimolecularly-thick (few tenths of nm) films of conventional solid-lubricants have been studied. Hansma et al. (1992) reported the deposition of multi-molecularly-thick, highly-oriented PTFE films from the melt, vapor phase or from solution by a mechanical deposition technique by dragging the polymer at controlled temperature, pressure and speed against a smooth glass substrate. Scandella et al. (1998) reported that the coefficient of nanoscale friction of MoS₂ platelets on mica, obtained by the exfoliation of lithium intercalated MoS₂ in water, was a factor of 1.4 less than that of mica itself. However, MoS₂ is reactive to water,

and its friction and wear properties degrade with an increase in humidity (Bhushan 2013a, b).

Amorphous diamond-like carbon (DLC) coatings can be produced with extremely high hardness and are commercially used as wear resistant coatings (Bhushan 1999b; Erdemir and Donnet 2001). They are widely used in magnetic storage devices (Bhushan 1996). Doping of the DLC matrix with elements like hydrogen, nitrogen, oxygen, silicon, and fluorine influences their hydrophobicity and tribological properties (Dorfman 1992; Grischke et al. 1995; Bhushan 1999b). Nitrogen and oxygen reduce the contact angle (or increase the surface energy) due to the strong polarity which is formed when these elements are bonded to carbon. On the other hand, silicon and fluorine increase the contact angle ranging from 70 to 100° (or reduce the surface energy to 20–40 dyn/cm), making them hydrophobic (Butter et al. 1997; Grischke et al. 1998). Nanocomposite coatings with a diamond like carbon (a-C:H) network and a glasslike a-Si:O network are generally deposited using a PECVD (plasma-enhanced chemical vapor deposition) technique in which plasma is formed from a siloxane precursor using a hot filament. For a fluorinated DLC, CF₄ as the fluorocarbon source is added to an acetylene plasma. In addition, fluorination of DLC can be achieved by the post deposition treatment of DLC coatings in a CF₄ plasma. Silicon- and fluorine-containing DLC coatings mainly reduce their polarity due to the loss of sp² bonded carbon (due to polarization potential of the involved π electrons) and dangling bonds of the DLC network. As silicon and fluorine are unable to form double bonds, they force carbon into a sp³ bonding state (Grischke et al. 1998). Friction and wear properties of both silicon-containing and fluorinated DLC coatings have been reported to be superior to that of conventional DLC coatings (Donnet et al. 1997; Kester et al. 1999). However, DLC coatings require a line of sight deposition process which prevents deposition on complex geometries. Furthermore, it has been reported that some self-assembled monolayers (SAMs) are superior to DLC coatings in hydrophobicity and tribological performance (Liu and Bhushan 2003b; Bhushan et al. 2004).

Organized and dense molecular-scale layers of, preferably, long-chain organic molecules are known to be superior lubricants on both macro- and micro/nanoscales as compared with freely supported multimolecular layers (Bhushan 1999a). Common techniques to produce molecular scale organized layers are Langmuir-Blodgett (LB) deposition and chemical grafting of organic molecules to realize SAMs (Ulman 1991, 1996). In the LB technique, organic molecules from suitable amphiphilic molecules are first organized at the air-water interface and then physisorbed on a solid surface to form mono- or multi-molecular layers (Zasadzinski et al. 1994). Whereas in the case of SAMs, the functional groups of molecules chemisorb on a solid surface which results in the spontaneous formation of robust, highly ordered, oriented, and dense monolayers (Ulman 1996). In both cases, the organic molecules used have well distinguished amphiphilic properties (a hydrophilic functional head and a hydrophobic aliphatic tail) so that an adsorption of such molecules on an active inorganic substrate leads to their firm attachment to the surface. Direct organization of SAMs on the solid surfaces allows coating in inaccessible areas. The weak adhesion of classical LB films to the substrate surface restricts their lifetimes during sliding, whereas certain SAMs can be very durable. As a result, SAMs are of great interest in tribological applications.

Much research into the application of SAMs has been carried out using the so called soft lithographic technique (Tian et al. 1998; Xia and Whitesides 1998). This is a nonphotolithographic technique. Photolithography is based on a projection-printing system used for projection of an image from a mask to a thin-film photoresist, and its resolution is limited by optical diffraction limits. In soft lithography, an elastomeric stamp or mold is used to generate micropatterns of SAMs by either contact printing (known as microcontact printing or μCP (Kumar and Whitesides 1993)), by embossing (nanoimprint lithography) (Chou et al. 1996), or by replica molding (Xia et al. 1996), which circumvent the diffraction limits of photolithography. The stamps are generally cast from photolithographically generated patterned masters, and the stamp

material is generally polydimethylsiloxane (PDMS). In μ CP, the ink is a SAM precursor to produce nm-thick resists with lines thinner than 100 nm. Soft lithography requires little capital investment. μ CP and embossing techniques may be used to produce microdevices which are substantially cheaper and more flexible in choice of material for construction than the conventional photolithography (e.g., SAMs and non-SAM entities for μ CP and elastomers for embossing).

The largest industrial application for SAMs is in digital micromirror devices (DMD) used in optical projection displays (Hornbeck 2001; Douglass 1998). The chip set of a DMD consists of half a million to more than two million independently controlled reflective aluminum alloy micromirrors of about 12 μ m square. These micromirrors switch forward and backward at a frequency on the order of 5–7 kHz with a rotation of $\pm 12^\circ$ with respect to the horizontal plane; the movement is limited by a mechanical stop. Mechanical contact leads to stiction and wear in contacting surfaces. A SAM of vapor deposited perfluorinated n-alkanoic acid ($C_nF_{2n-1}\cdot O_2H$) (e.g., perfluorodecanoic acid, or PFDA, $CF_3(CF_2)_8\cdot COOH$) is used to coat contacting surfaces to make them hydrophobic in order to minimize meniscus formation. Furthermore, the entire DMD chip set is hermetically sealed in order to prevent particulate contamination and excessive condensation of water at the contacting surfaces. A so-called “getter” strip of PFDA is included inside the hermetically sealed enclosure containing the chip, which acts as a reservoir in order to maintain a PFDA vapor within the package. Degradation mechanisms of SAMs leading to stiction have been studied by Liu and Bhushan (2004a, b). Nanotribological studies of various SAMs on Al substrates have been carried out by Tambe and Bhushan (2005a), Bhushan et al. (2006a), Hoque et al. (2006a, b, 2007a) and DeRose et al. (2008).

There are various other micro/nanodevices which require SAMs for hydrophobicity in order to minimize meniscus formation. Examples include micromotors, microgears, microvalves, microswitches, mirror-based optical switches, and

atomic force microscopy probes (Bhushan 2007, 2008, 2016). Nanotribological studies on Si substrates have been carried out by Bhushan et al. (1995b, 2005, 2006b), Kasai et al. (2005), Lee et al. (2005), Tambe and Bhushan (2005a) and Tao and Bhushan (2005, 2006). SAMs deposition on Cu surfaces are also being explored for corrosion inhibition for micro/nanoelectronics and/or heat-exchange surfaces exploiting dropwise condensation (Hoque et al. 2007b, 2008, 2009).

Other industrial applications for SAMs are in the areas of bio/chemical and optical sensors, devices for use as drug-delivery vehicles, and in the construction of electronic components (Manz and Becker 1998; Cheng and Kricka 2001; Heller and Guttman 2001; van den Berg 2003; Bhushan 2016). Bio/chemical sensors require highly sensitive organic layers with tailored biological properties that can be incorporated into electronic, optical, or electrochemical devices. Self-assembled microscopic vesicles are being developed to ferry potentially lifesaving drugs to cancer patients. By assembling organic, metal, and phosphonate molecules (complexes of phosphorous and oxygen atoms) themselves into conductive materials, these can be produced as self-made sandwiches for use as electronic components. Several applications have been proposed based on silicon, glass, or polymer nanochannels, including cell immunoisolation chambers, DNA separation devices, and biocapsules for drug delivery (Bhushan 2007, 2008).

SAMs are also being considered for protection of the surfaces from exposure to the operating environment. They are being developed to reduce corrosion and oxidation of Cu in heat exchangers (Hoque et al. 2007b, 2008, 2009). They are being developed to minimize undesirable protein adsorption, biofouling and improve biocompatibility in biosensors and implantable biomedical devices (Bhushan et al. 1995b, 2005, 2006a, b; Kasai et al. 2005; Lee et al. 2005; Tambe and Bhushan 2005a; Tao and Bhushan 2005; Tokachichu and Bhushan 2006; Bhushan and Cichomski 2007). These films can also be used to reduce fluid drag in micro/nanochannels.

13.1.3 Scope of the Chapter

An overview of molecularly thick layers of liquid lubricants and conventional solid lubricants can be found in various references, such as Bhushan (1996, 1999a, b, 2001, 2013a, b, 2016), Bhushan and Zhao (1999), and Liu and Bhushan (2003a). In this chapter, we present an overview of SAMs for high hydrophobicity, low adhesion, friction, and wear. SAMs are produced by various organic precursors. We first present a primer to organic chemistry followed by an overview on suitable substrates, spacer chains and end groups in the molecular chains, an overview on contact angle, adhesion, friction, and wear properties of various SAMs, and some concluding remarks.

13.2 A Primer to Organic Chemistry

All organic compounds contain the carbon (C) atom. Carbon, in combination with hydrogen, oxygen, nitrogen, sulfur and phosphorus results in a large number of organic compounds. The atomic number of carbon is 6, and its electron structure is $1s^2 2s^2 2p^2$. Two stable isotopes of carbon, ^{12}C and ^{13}C , exist. With four electrons in its outer shell, carbon forms four covalent bonds with each bond resulting from two atoms sharing a pair of electrons. The number of electron pairs that two atoms share determines whether or not the bond is single or multiple. In a single bond, only one pair of electrons is shared by the atoms. Carbon can also form multiple bonds by sharing two or three pairs of electrons between the atoms. For example, the double bond formed by sharing two electron pairs is stronger than a single bond, and it is shorter than a single bond. An organic compound is classified as saturated if it contains only single bonds and as unsaturated if the molecules possess one or more multiple carbon-carbon bonds.

13.2.1 Electronegativity/Polarity

When two different kinds of atoms share a pair of electrons, a bond is formed in which electrons are

shared unequally: one atom assumes a partial positive charge and the other a negative charge with respect to each other. This difference in charge occurs because the two atoms exert unequal attraction in the pair of shared electrons. The attractive force that an atom of an element has for shared electrons in a molecule or polyatomic ion is known as its electronegativity. Elements differ in their electronegativities. A scale of relative electronegatives, in which the most electronegative element, fluorine, is assigned a value of 4.0, was developed by L. Pauling. Relative electronegativities of the elements in the periodic table can be found in most undergraduate chemistry textbooks (e.g., Hein et al. 1997). Relative electronegativity of the nonmetals is high as compared to that of metals. Relative electronegativity of selected elements of interest with high values is presented in Table 13.1.

The polarity of a bond is determined by the difference in electronegativity values of the atoms forming the bond. If the electronegativities are the same the bond is nonpolar, and the electrons are shared equally. In this type of bond, there is no separation of positive and negative charge between atoms. If the atoms have greatly different electronegativities the bond is very polar. A dipole is a molecule that is electrically asymmetrical, causing it to be oppositely charged at two points. As an example, in hydrogen chloride, HCl, both hydrogen and chlorine need one electron to form stable electron configurations. They share a pair of electrons. Chlorine is more electronegative and therefore has a greater attraction for the shared electrons than does

Table 13.1 Relative electronegativity of selected elements

Element	Relative electronegativity
F	4.0
O	3.5
N	3.0
Cl	3.0
C	2.5
S	2.5
P	2.1
H	2.1

hydrogen. As a result, the pair of electrons is displaced towards the chlorine atom, giving it a partial negative charge and leaving the hydrogen atom with a partial positive charge, (Fig. 13.2). However, the entire molecule, HCl, is electrically neutral. The hydrogen atom with a partial positive charge (exposed proton on one end) can be easily attracted to the negative charge of other molecules and this is responsible for the polarity of the molecule. A partial charge is usually indicated by δ , and the electronic structure of HCl is given as:



Similar to the HCl molecule, HF is polar, and both behave as a small dipole. On the other hand, methane (CH_4), carbon tetrachloride (CCl_4), and carbon dioxide (CO_2) are nonpolar. In CH_4 and CCl_4 , the four C–H and C–Cl polar bonds are identical, and because these bonds emanate from the center to the corners of a tetrahedron in the molecule, the effect of their polarities cancels one another. CO_2 ($\text{O}=\text{C}=\text{O}$) is nonpolar because carbon-oxygen dipoles cancel each other by acting in opposite direction. Water ($\text{H}-\text{O}-\text{H}$) is a polar molecule. If the atoms in water were linear as in CO_2 , two O–H dipoles would cancel each other, and the molecule would be nonpolar. However, water has a bent structure with an

angle of 105° between the two bonds, which is responsible for the water being a polar molecule.

13.2.2 Classification and Structure of Organic Compounds

Table 13.2 presents selected organic compounds grouped into classes.

13.2.2.1 Hydrocarbons

Hydrocarbons are compounds that are composed entirely of carbon and hydrogen atoms bonded to each other by covalent bonds. Saturated hydrocarbons (alkanes) contain single bonds. Unsaturated hydrocarbons that contain carbon-carbon double bonds are called alkenes, and ones with triple bonds are called alkynes. Unsaturated hydrocarbons that contain aromatic rings, e.g., benzene rings, are called aromatic hydrocarbons.

Saturated Hydrocarbons: Alkanes

The alkanes also known as paraffins, are saturated hydrocarbons, straight- or branched-chain hydrocarbons with only single covalent bonds between the carbon atoms. The general molecular formula for alkanes is $\text{C}_n\text{H}_{2n+2}$, where n is the number of carbon atoms in the molecule. Each carbon atom is connected to four other atoms by four single covalent bonds. These

Fig. 13.2 Schematic representation of the formation of a polar HCl molecule

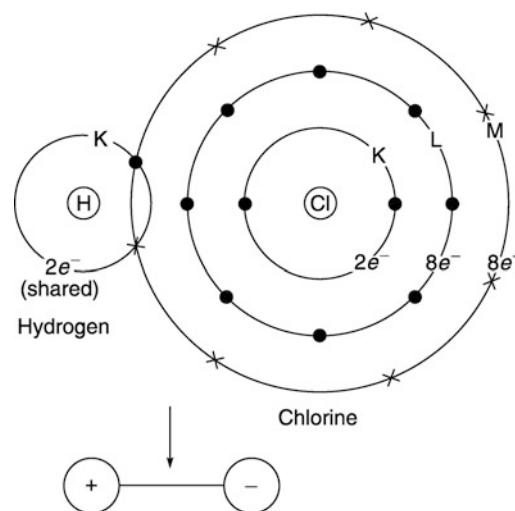
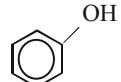


Table 13.2 (a) Names and formulas of selected hydrocarbons, (b) names and formulas of selected alcohols, ethers, phenols and thiols, (c) Names and formulas of selected aldehydes and ketones, (d) names and formulas of selected carboxylic acids and esters, (e) names and formulas of selected organic nitrogen compounds (amides and amines)

Name	Formula
a	
<i>Saturated hydrocarbons</i>	
Straight-chain alkanes	C_nH_{2n+2}
e.g., Methane	CH_4
Ethane	C_2H_6 or CH_3CH_3
Alkyl groups	C_nH_{2n+1}
e.g., Methyl	$-CH_3$
Ethyl	$-CH_2CH_3$
<i>Unsaturated hydrocarbons</i>	
Alkenes	$(CH_2)_n$
e.g., Ethene	C_2H_4 or $CH_2=CH_2$
Propene	C_3H_6 or $CH_3CH=CH_2$
Alkynes	$HC\equiv CH$
e.g., Acetylene	
Aromatic hydrocarbons	
e.g., Benzene	C_6H_5OH or 
b	
Alcohols	$R-OH$
e.g., Methanol	CH_3OH
Ethanol	CH_3CH_2OH
Ethers	$R-O-R'$
e.g., Dimethyl ether	CH_3-O-CH_3
Diethyl ether	$CH_3CH_2-O-CH_2CH_3$
Phenols	C_6H_5OH or 
Thiols	$-SH$
e.g., Methanethiol	CH_3SH
The letters R- and R'-represent an alkyl group. The R-groups in ethers can be the same or different and can be alkyl or aromatic (Ar) groups	
c	
Aldehydes	$RCHO$ or $HCHO$
	$\begin{array}{c} O \\ \\ R-C-H \end{array}$
e.g., Methanal or Formaldehyde	$ArCHO$ or $Ar-C-H$
Ethanal or Acetaldehyde	CH_3CHO
Ketones	$RCOR'$ or $RCOAr$ or $ArCOAr$
	$\begin{array}{c} O \\ \\ R-C-R' \\ O \\ \\ R-C-Ar \\ O \\ \\ Ar-C-Ar \end{array}$

(continued)

Table 13.2 (continued)

Name	Formula
e.g., Butanone or Methyl Ethyl ketone	$\text{CH}_3\text{COCH}_2\text{CH}_3$
The letters R and R' represent an alkyl group and Ar represents an aromatic group	
d Carboxylic acid ^a	$\begin{array}{c} \text{O} \\ \\ \text{RCOOH} \end{array}$ or $\begin{array}{c} \text{O} \\ \\ \text{ArCOOH} \end{array}$ or $\begin{array}{c} \text{Ar}-\text{C}-\text{OH} \\ \\ \text{O} \end{array}$
e.g., Methanoic acid (formic acid) Ethanoic acid (acetic acid) Octadecanoic acid (stearic acid)	HCOOH CH_3COOH $\text{CH}_3(\text{CH}_2)_{16}\text{COOH}$
Esters ^b	$\begin{array}{c} \text{O} \\ \\ \text{R}-\text{C}-\text{O}-\text{R}' \\ \\ \text{RCOOR}' \end{array}$ or acid alcohol
e.g., Methyl propanoate	$\text{CH}_3\text{CH}_2\text{COOCH}_3$
^a The letter R represents an alkyl group and Ar represents an aromatic group	
^b The letter R represents hydrogen, alkyl group or aromatic group and R' represents alkyl group or aromatic group	
e Amides	$\begin{array}{c} \text{O} \\ \\ \text{R}-\text{C}-\text{NH}_2 \\ \\ \text{RCONH}_2 \end{array}$ or $\begin{array}{c} \text{HCONH}_2 \\ \\ \text{CH}_3\text{CONH}_2 \end{array}$
e.g., Methanamide (formamide) Ethanamide (acetamide)	
Amines	$\begin{array}{c} \text{H} \\ \\ \text{R}-\text{N}-\text{H} \\ \\ \text{H} \\ \text{RNH}_2 \end{array}$ or $\begin{array}{c} \text{R}_2\text{NH} \\ \\ \text{R}_3\text{N} \end{array}$
e.g., Methylamine Ethylamine	$\begin{array}{c} \text{CH}_3\text{NH}_2 \\ \\ \text{CH}_3\text{CH}_2\text{NH}_2 \end{array}$
The letter R represents an alkyl group or aromatic group	

bonds are separated by angles of 109.5° (the angle by lines drawn from the center of a regular tetrahedron to its corners). Alkane molecules contain only carbon–carbon and carbon–hydrogen bonds, which are symmetrically directed towards the corners of a tetrahedron. Therefore alkane molecules are essentially nonpolar.

Common alkyl groups have the general formula $\text{C}_n\text{H}_{2n+1}$ (one hydrogen atom less than the corresponding alkane). The missing H atom may be detached from any carbon in the alkane. The name of the group is formed from the name of the corresponding alkane by replacing -ane with -yl ending. Some examples are shown in Table 13.2a.

Unsaturated Hydrocarbons

Unsaturated hydrocarbons consist of three families of compounds that contain fewer hydrogen atoms than the alkane with the corresponding number of carbon atoms, and contain multiple bonds between carbon atoms. These include alkenes (with carbon-carbon double bonds), alkynes (with carbon-carbon triple bonds), and aromatic compounds (with benzene rings which are arranged in a six-membered ring with one hydrogen atom bonded to each carbon atom). Some examples are shown in Table 13.2a.

13.2.2.2 Alcohols, Ethers, Phenols, and Thiols

Organic molecules with certain functional groups are synthesized for desirable properties. Alcohols, ethers and phenols are derived from the structure of water by replacing the hydrogen atoms of water with alkyl groups (R) or aromatic (Ar) rings. For example, phenol is a class of compounds that has a hydroxy group attached to an aromatic ring (benzene ring). Organic compounds that contain the –SH group are analogs of alcohols, and are known as thiols. Some examples are shown in Table 13.2b.

13.2.2.3 Aldehydes and Ketones

Both aldehydes and ketones contain the carbonyl group, >C=O , a carbon-oxygen double bond. Aldehydes have at least one hydrogen atom bonded to the carbonyl group, whereas ketones have only alkyl or aromatic group bonded to the carbonyl group. The general formula for the saturated homologous series of aldehydes and ketones is $\text{C}_n\text{H}_{2n}\text{O}$. Some examples are shown in Table 13.2c.

13.2.2.4 Carboxyl Acids and Esters

The functional group of the carboxylic acids is known as a carboxyl group represented as –COOH. Carboxylic acids can be either aliphatic (RCOOH) or aromatic (ArCOOH). The carboxylic acids with even numbers of carbon atoms, n, ranging from 4 to about 20 are called fatty acids (e.g., n = 10, 12, 14, 16 and 18 are called capric acid, lauric acid, myristic acid, palmitic acid, and stearic acid, respectively).

Esters are alcohol derivates of carboxylic acids. Their general formula is RCOOR' , where R may be a hydrogen, alkyl group, or aromatic group, and R' may be an alkyl group or aromatic group, but not a hydrogen. Esters are found in fats and oils. Some examples are shown in Table 13.2d.

13.2.2.5 Amides and Amines

Amides and amines are organic compounds containing nitrogen. Amides are nitrogen derivatives of carboxylic acids. The carbon atom of a carbonyl group is bonded directly to a nitrogen atom of an $-\text{NH}_2$, $-\text{NHR}$, or $-\text{NR}_2$ group. The characteristic structure of amide is RCONH_2 .

An amine is a substituted ammonia molecule which has a general structure as RNH_2 , R_2NH or R_3N , where R is an alkyl or aromatic group. Some examples are shown in Table 13.2e

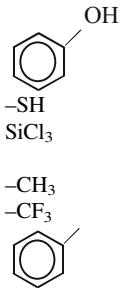
13.2.3 Polar and Nonpolar Groups

Table 13.3a summarizes polar and nonpolar groups commonly used in the construction of hydrophobic and hydrophilic molecules. Table 13.3b lists the relative polarity of selected polar groups (Mohrig et al. 1998). Thiol, silane, carboxylic acid, and alcohol (hydroxyl) groups are the most commonly used polar anchor groups for their attachment to surfaces. Silane anchor groups are commonly used for Si or SiO_2 surfaces as $-\text{Si}-\text{O}-$ bonds are strong. Methyl and trifluoromethyl are commonly used as end groups for hydrophobic film surfaces.

13.3 Self-assembled Monolayers: Substrates, Spacer Chains, and End Groups in the Molecular Chains

SAMs are formed as a result of spontaneous, self-organization of functionalized organic molecules onto the surfaces of appropriate substrates into stable, well defined structures, Fig. 13.3. The final structure is close to or at thermodynamic equilibrium, and as a result it

Table 13.3 (a) Some examples of polar (hydrophilic) and nonpolar (hydrophobic) groups, (b) organic groups listed in the increasing order of polarity

Name	Formula
a	
<i>Polar</i>	
Alcohol (hydroxyl)	-OH
Carboxyl acid	-COOH
Aldehyde	-COH
Ketone	O R-C-R
Ester	-COO-
Carbonyl	>C=O
Ether	R-O-R
Amine	-NH ₂
Amide	O -C-NH ₂
Phenol	
Thiol	-SH
Trichlorosilane	SiCl ₃
<i>Nonpolar</i>	
Methyl	-CH ₃
Trifluoromethyl	-CF ₃
Aryl (benzene ring)	
The letter R represents an alkyl group	
b	
Alkanes	
Alkenes	
Aromatic hydrocarbons	
Ethers	
Trichlorosilanes	
Aldehydes, ketones, esters, carbonyls	
Thiols	
Amines	
Alcohols, phenols	
Amides	
Carboxylic acids	

tends to form spontaneously and rejects defects. SAMs consist of three building groups: a head group that binds strongly to a substrate, a surface terminal (tail or end) group that constitutes the outer surface of the film, and a spacer chain (backbone chain) that connects the head and surface terminal groups. The SAMs are named based on the surface terminal group followed by the spacer chain and the head group (or type of compound formed at the surface). In order to

control hydrophobicity, adhesion, friction, and wear, it should be strongly adherent to the substrate, and the surface terminal group of the organic molecular chain should be nonpolar. For a strong attachment of the organic molecules to the substrate, the head group of the molecular chain should contain a polar terminal group resulting in the exothermic process (energies on the order of tens of kcal/mol), i.e., it results in an apparent pinning of the head group to a specific

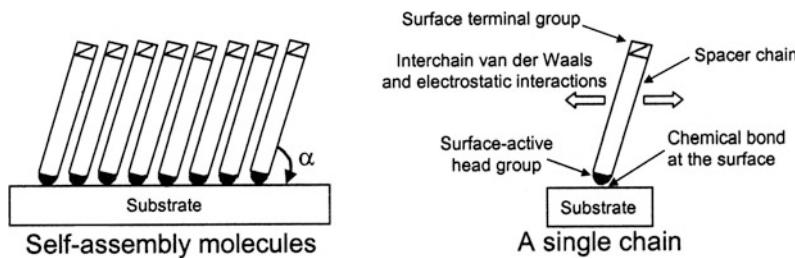


Fig. 13.3 Schematic of a self-assembled monolayer on a surface and the associated forces

site on the surface through a chemical bond. Furthermore, the molecular structure and any crosslinking would have a significant effect on their friction and wear performance. The substrate surface should have a high surface energy (hydrophilic), so that there will be a strong tendency for molecules to adsorb on the surface. The surface should be highly functional with polar groups and dangling bonds (generally unpaired electrons) so that they can react with organic molecules and provide a strong bond. Because of the exothermic head group-substrate interactions, molecules try to occupy every available binding site on the surface, and during this process they generally push together molecules that have already adsorbed. The process results in the formation of ordered molecular assemblies. The interactions between molecular chains are van der Waals or electrostatic types with energies on the order of a few (<10) kcal/mol, exothermic. The molecular chains in SAMs are not perpendicular to the surface; the tilt angle depends on the anchor group as well as

on the substrate and the spacer group. For example, the tilt angle for alkanethiolate on Au is typically about a 30–35° angle with respect to substrate normal.

Table 13.4 lists selected systems which have been used for the formation of SAMs (Xia and Whitesides 1998). The spacer chain of the SAM is mostly an alkyl chain ($(-\text{CH}_2)_n$) or made of a derivatized alkyl group. By attaching different terminal groups at the surface, the film surface can be made to attract or repel water. The commonly used surface terminal group of a hydrophobic film with low surface energy, in the case of a single alkyl chain, is a nonpolar methyl ($-\text{CH}_3$) or trifluoromethyl ($-\text{CF}_3$) group. For a hydrophilic film, the commonly used surface terminal groups are alcohol ($-\text{OH}$) or carboxylic acid ($-\text{COOH}$) groups. The surface active head groups most commonly used are thiol ($-\text{SH}$), silane (e.g., trichlorosilane or $-\text{SiCl}_3$), and carboxyl ($-\text{COOH}$) groups. The substrates most commonly used are gold, silver, platinum, copper, hydroxylated (activated) surfaces of SiO_2 on Si, Al_2O_3 on Al, and glass.

Table 13.4 Selected substrates and precursors which have commonly been used for formation of SAMs

Substrate	Precursor	Binding with substrate
Au	R SH (thiol)	RS-Au
Au	Ar SH (thiol)	ArS-Au
Au	RSSR' (disulfide)	RS-Au
Au	RSR' (sulfide)	RS-Au
Si/SiO ₂ , glass	R SiCl ₃ (trichlorosilane)	Si-O-Si (siloxane)
Si/Si-H	R COOH (carboxyl)	R-Si
Metal oxides (e.g., Al ₂ O ₃ , SnO ₂ , TiO ₂)	RCOOH (carboxyl)	RCOO-...MO _n

R represents alkane ($\text{C}_n\text{H}_{2n+2}$) and Ar represents aromatic hydrocarbon. It consists of various surface active headgroups and mostly with methyl terminal group

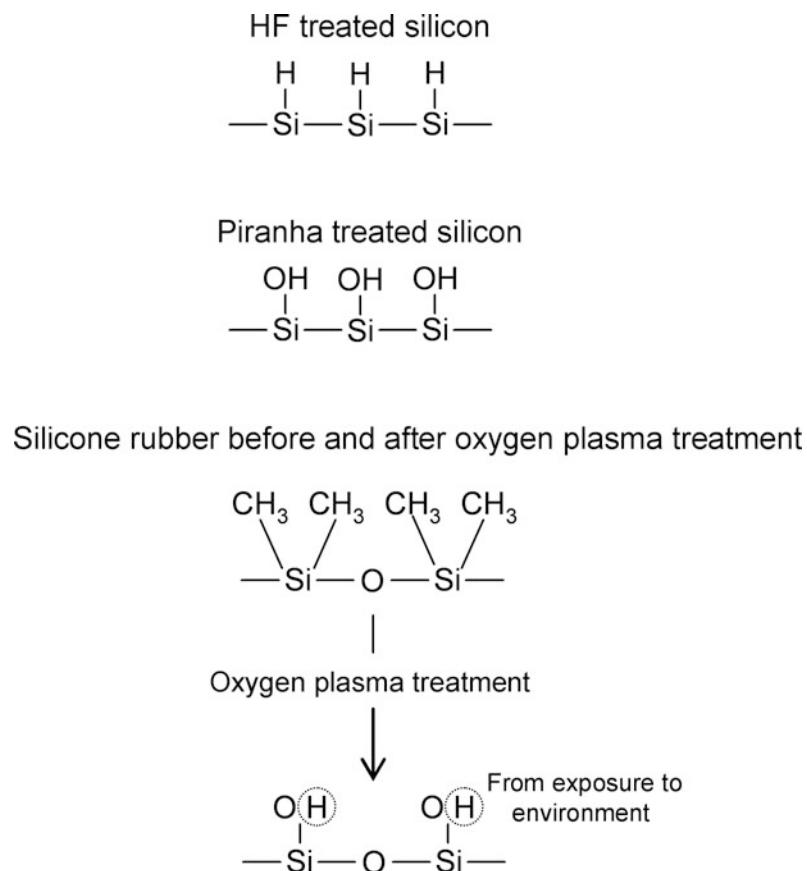


Fig. 13.4 Schematic showing HF treated silica and hydroxylation process occurring on a silica and elastomeric surfaces using Piranha solution and oxygen plasma, respectively

As an example of substrate, epitaxial Au film on glass, mica or single-crystal silicon, produced by e-beam evaporation, is commonly used because it can be deposited on smooth surfaces as a film which is atomically flat and defect free. Bulk single-crystal Si, sputtered Al film and bulk Al sheets, and copper disks with natural oxide layers have been selected because of their use in the construction of MEMS/NEMS (e.g., digital projection displays) and heat exchange applications, respectively. The substrate surface should be clean before deposition. For silicon substrates, a concentrated HF solution (typically 49% HF) is commonly used to remove the oxide layer followed by a rinse with deionized water (Lee et al.

2005; Bhushan et al. 2006b). Hydrogen passivates the surface by saturating the dangling bonds, and it results in a hydrogen-terminated silicon surface with hydrophobic properties. For deposition of multimolecularly-thick polymer films with nonpolar ends, hydrophobic substrates may lead to a coated surface with a high contact angle, and are preferred. For SAM deposition, the substrate should be hydrophilic in order to form strong interfacial bonds with their head groups. Hydroxylation of oxide surfaces is carried out to make them hydrophilic. Silicon and other metals get oxidized and get hydroxylated to some degree when exposed to the environment. Bulk silicon, polysilicon film or SiO_2 film

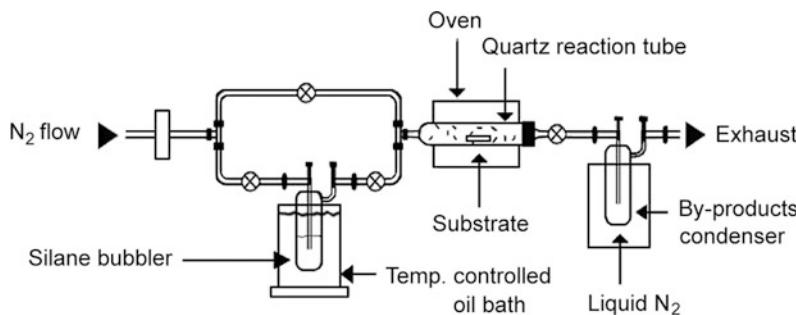


Fig. 13.5 Schematic showing vapor phase deposition system for silane SAMs (Bhushan et al. 2005)

surfaces are commonly treated to produce a hydroxylated silica surface by immersion in Piranha solution (a mixture of typically 3:1 v/v 98% H₂SO₄: 30% H₂O₂) at temperatures of about 90 °C for about 30 min followed by a rinse in deionized water (Liu and Bhushan 2003a; Lee et al. 2005; Bhushan et al. 2006b). Piranha solution also removes any organic and metallic contaminants, whereas HF would not necessarily remove organics. Oxygen plasma is another technique used for hydroxylation for SiO₂ as well as polymer surfaces (Bhushan et al. 2005; Kasai et al. 2005; Tao and Bhushan 2006; Bhushan and Cichomski 2007). For complex silicon geometries or fine structures, such as AFM tips, oxygen plasma may be preferable. Figure 13.4 shows the schematics of surfaces after various surface treatments. Surfaces after piranha or oxygen plasma treatment remain hydrophilic for a few hours to about a day and become hydrophobic when they come in contact with carbon. They can retain hydrophilicity longer in dry nitrogen. To retain hydrophobicity, polymers are generally stored in DI water. Treated surfaces with HF remain hydrophobic for about 2 to 3 h and can retain hydrophobicity longer in dry nitrogen.

For organic molecules to pack together and provide a better ordering, a substrate for given molecules should be selected such that the cross-sectional diameter of the spacer chains of the molecule is equal to or smaller than the distance between the anchor groups attached to the substrate. For the case of alkanethiol film, the advantage of Au substrate over SiO₂ substrate is that it results in better ordering because the cross-sectional diameter

of the alkane molecule is slightly smaller than the distance between sulfur atoms attached to the Au substrate (~0.53 nm). The thickness of the film can be controlled by varying the length of the hydrocarbon chain, and the surface properties of the film can be modified by the terminal group.

SAMs are usually produced by immersing a substrate in a solution containing the precursor (ligand) that is reactive to the substrate surface or by exposing the substrate to the vapor of the reactive chemical precursors (Ulman 1991). A schematic of the vapor deposition system is shown in Fig. 13.5 (Bhushan et al. 2005, 2006b; Bhushan and Cichomski 2007). Samples are placed in the quartz reaction tube. The silane bubbler is used for introducing gas phase silane into the quartz reaction tube placed in an oven at a controlled temperature. An inert gas flow (N₂) is used as a carrier gas. A by-products condenser is used for trapping the by-products and/or non-reacted silanes.

Research on some SAMs has been widely reported. SAMs of long-chain fatty acids C_nH_{2n}₊₁COOH or (CH₃)_{(CH₂)_n}COOH (n = 10, 12, 14 or 16) on glass or alumina substrates have been widely studied since the 1950s (Bowden and Tabor 1950; Zisman 1959; Ulman 1991, 1995, 1996). Probably the most studied SAMs to date are n-alkanethiolate² monolayers CH₃(CH₂)_nS- prepared from the adsorption of alkanethiol – (CH₂)_nSH solution onto a Au film (Jung et al. 1998; Xia and Whitesides 1998; Geyer et al. 1999; Bhushan and Liu 2001; Liu et al. 2001;

²n-alkyl and n-alkane are used interchangeably.

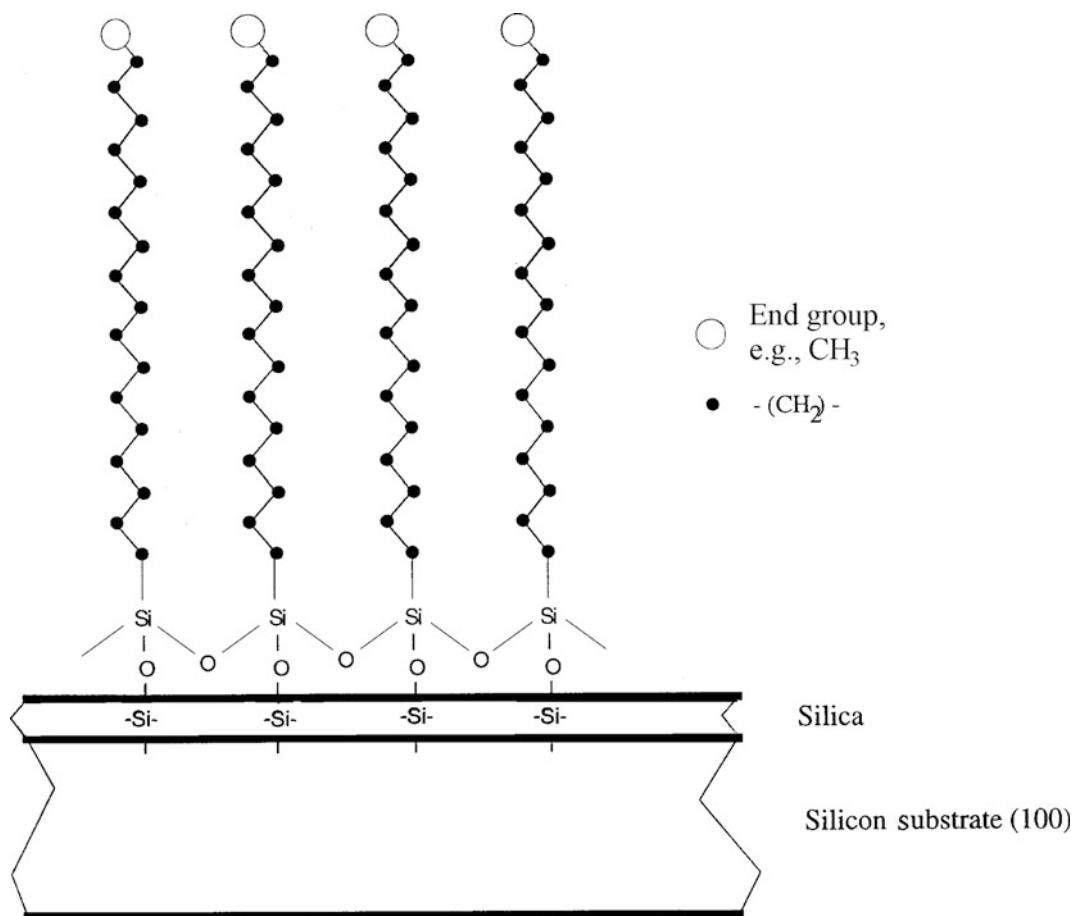


Fig. 13.6 Schematics of a methyl-terminated, n-alkylsiloxane monolayer on Si/SiO₂

Liu and Bhushan 2002a) and n-alkylsiloxane³ monolayers produced by adsorption of n-alkyltrichlorosilane $-(\text{CH}_2)_n\text{SiCl}_3$ onto a hydroxylated Si/SiO₂ substrate with siloxane (Si—O—Si) binding; see Fig. 13.6 (Wasserman et al. 1989; Bhushan et al. 1995b, 2005, 2006b; Kasai et al. 2005; Lee et al. 2005; Tambe and Bhushan 2005a; Tao and Bhushan 2005). Tambe and Bhushan (2005a), Bhushan et al. (2006a), Hoque et al. (2006a, b, 2007a, b, 2008, 2009). DeRose et al. (2008) have produced perfluoro-

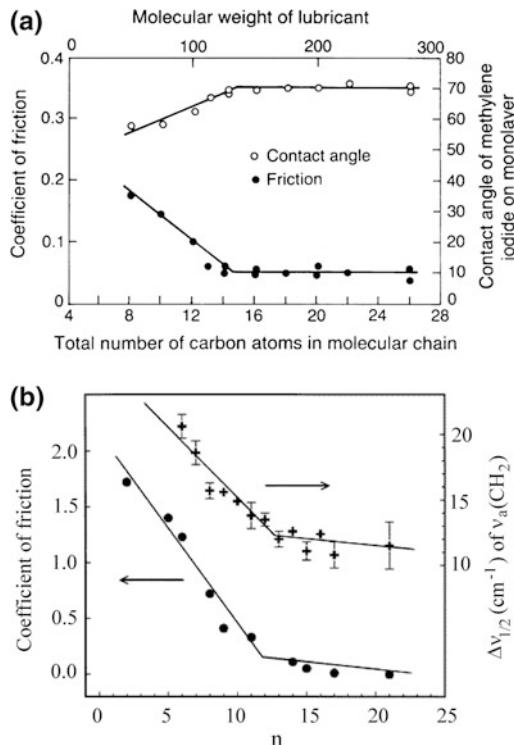
alkylsilane and perfluoroalkylphosphonate on Al and Cu surfaces.

13.4 Contact Angle and Nanotribological Properties of SAMs

The basis for the molecular design and tailoring of SAMs should start from a complete knowledge of the interrelationships between the molecular structure and contact angle and nanotribological properties of SAMs, as well as a deep understanding of the adhesion, friction, and wear mechanisms of SAMs at the molecular level. Friction and wear studies of SAMs have been carried out on macro- and nanoscales. Macroscale

³Siloxane (Si—O—Si) refers to the bond and silane ($\text{Si}_n\text{X}_{2n+2}$ which includes a covalently bonded compounds containing the elements Si and other atoms or groups such as H and Cl to form SiH_4 and SiCl_4 , respectively) refers to the head group of the precursor. These terms are used interchangeably.

Fig. 13.7 **a** Effect of chain length (or molecular weight) on coefficient of macroscale friction of stainless steel sliding on glass lubricated with a monolayer of fatty acid and contact angle of methyl iodide on condensed monolayers of fatty acid on glass (Zisman 1959), **b** effect of chain length of methyl-terminated, n-alkanethiolate over Au film $\text{AuS}(\text{CH}_2)_n\text{CH}_3$ on the coefficient of microscale friction and peak bandwidth at half maximum ($\Delta\nu_{1/2}$) for the bandwidth of the methylene stretching mode [$\nu_a(\text{CH}_2)$] (McDermott et al. 1997)



tests are conducted using a so called pin-on-disk tribotester apparatus in which a ball specimen slides against a lubricated flat specimen (Bhushan 2013a, b). Nanoscale tests are conducted using an atomic force/friction force microscope (AFM/FFM) (Bhushan 1995a, 1999a, 2011, 2013a, b). In the AFM/FFM experiments, a sharp tip of a radius ranging from about 5 to 50 nm slides against a SAM specimen. A Si_3N_4 tip is commonly used for friction studies, and a Si or natural diamond tip is commonly used for scratch, wear, and indentation studies.

In early studies, the effect of chain length of the carbon atoms of fatty acid monolayers on the coefficient of friction and wear on the macroscale was studied by Bowden and Tabor (1950) and Zisman (1959). Zisman (1959) reported that for the monolayers deposited on a glass surface sliding against a stainless steel surface, there is a steady decrease in friction with increasing chain length. At a significantly long chain length, the

coefficient of friction reaches a lower limit, Fig. 13.7a. He further reported that monolayers having a chain length below 12 carbon atoms behave as liquids (poor durability), and those with a chain length of 12–15 carbon atoms behave like a plastic solid (medium durability), whereas those with a chain length above 15 carbon atoms behave like a crystalline solid (high durability). Investigations by Ruhe et al. (1993) indicated that the lifetime of the alkylsilane monolayer coating on a silicon surface increases greatly with an increase in the chain length of the alkyl substituent. DePalma and Tillman (1989) showed that a monolayer of n-octadecyltrichlorosilane ($n\text{-C}_{18}\text{H}_{37}\text{SiCl}_3$, OTS) is an effective lubricant on silicon.

With the development of AFM techniques, researchers have successfully characterized the nanotribological properties of self-assembled monolayers (Bhushan et al. 1995a; Bhushan 1999a, 2008, 2011). Studies by Bhushan et al.

(1995b) showed that C₁₈ alkylsiloxane films exhibit the lowest coefficient of friction and can withstand a much higher normal load during sliding as compared to LB films, soft Au films, and hard SiO₂ coatings. McDermott et al. (1997) studied the effect of the length of alkyl chains on the frictional properties of methyl-terminated n-alkylthiolate CH₃(CH₂)_nS⁻ films chemisorbed on Au(111) using an AFM. They reported that the longer chain monolayers exhibit a markedly lower friction and a reduced propensity to wear than shorter chain monolayers, Fig. 13.7b. These results are in good agreement with the macroscale results by Zisman (1959). They also conducted infrared reflection spectroscopy to measure the bandwidth of the methylene stretching mode [ν_a(CH₂)] which exhibits a qualitative correlation with the packing density of the chains. It was found that the chain structures of monolayers prepared with longer chain lengths are more ordered and more densely packed in comparison to those of monolayers prepared with shorter chain lengths. They further reported that the ability of the longer chain monolayers to retain molecular scale order during shear leads to a lower observed friction. Monolayers having a chain length of more than 12 carbon atoms, preferably 18 or more, are desirable for tribological applications. (Incidentally, monolayers with 18 carbon atoms, octadecanethiol films, have been widely studied.)

Xiao et al. (1996) and Lio et al. (1997) also studied the effect of the length of the alkyl chains on the frictional properties of n-alkanethiolate films on gold and n-alkylsilane films on mica. Friction was found to be particularly high with short chains of less than eight carbon atoms. Thiols and silanes exhibit similar friction forces for the same n when n > 11, while for n < 11, silanes exhibit higher friction, larger than that for thiols by a factor of about 3 for n = 6. The increase in friction was attributed to the large number of dissipative modes in the less ordered chains that occurs when going from a thiol to a silane anchor or when decreasing n. Longer chains (n > 11), stabilized by van der Waals attraction, form more compact and rigid layers and act as better lubricants. Schonherr and

Vancso (1999) also correlated the magnitude of friction with the order of the alkane chains. The disorder of the short chain hydrocarbon disulfide SAMs was found to result in a significant increase in the magnitude of friction.

Tsukruk and Bliznyuk (1998) studied the adhesion and friction between a Si sample and a Si₃N₄ tip, in which both surfaces were modified by -CH₃, -NH₂ and -SO₃H terminated silane based SAMs. Various polymer molecules for the backbone were used. They reported that a very broad maximum adhesive force in a pH range from 4 to 8 with a minimum adhesion at pH > 9 and pH < 3 were observed for all of the studied mating surfaces. This observation can be understood by considering a balance of electrostatic and van der Waals interaction between composite surfaces with multiple isoelectric points. The friction coefficient of NH₂/NH₂ and SO₃H/SO₃H mating SAMs was very high in aqueous solutions. Cappings of NH₂ modified surfaces (3-aminopropyltriethoxysilane) with rigid and soft polymer layers resulted in a significant reduction in adhesion to a level lower than that of untreated surface (Tsukruk et al. 1998). Fujihira et al. (2001) studied the influence of surface terminal groups of SAMs and functional tip on adhesive force. It was found that the adhesive forces measured in air increase in the order of CH₃/CH₃, CH₃/COOH, and COOH/COOH.

Bhushan and Liu (2001), Liu et al. (2001), and Liu and Bhushan (2002a, b) studied adhesion, friction and wear properties of alkylthiol and biphenylthiol SAMs on Au(111) films. They explained the friction mechanisms using a molecular spring model in which local stiffness and intermolecular forces govern the friction properties. They studied the influence of relative humidity, temperature, and velocity on adhesion and friction. They also investigated the wear mechanisms of SAMs by a continuous micro-scratch AFM technique.

Fluorinated carbon (fluorocarbon) molecules are known to have low surface energy and are commonly used for lubrication (Bhushan 2013a, b). Bhushan and Cichomski (2007) deposited fluorosilane SAMs on polydimethylsiloxane (PDMS). To make a hydrophobic PDMS surface

chemically active, PDMS surface was oxygenated using an oxygen plasma, which introduces silanol groups (SiOH). They reported that SAM coated PDMS was more hydrophobic with lower adhesion, friction, and wear. Bhushan et al. (2005, 2006b), Kasai et al. (2005), Lee et al. (2005), Tambe and Bhushan (2005a), and Tao and Bhushan (2006) studied the adhesion, friction, and wear of methyl- and/or perfluoro-terminated alkylsilanes on silicon. They reported that perfluoroalkylsilane SAMs exhibited lower surface energy, higher contact angle, lower adhesive force, and lower wear as compared to that of alkylsilanes. Kasai et al. (2005) also reported the influence of relative humidity, temperature, and velocity on adhesion and friction. Tao and Bhushan (2005) studied degradation mechanisms of alkylsilanes and perfluoroalkylsilane SAMs on Si. They reported that oxygen in the air causes thermal oxidation of SAMs.

Tambe and Bhushan (2005a), Bhushan et al. (2006a), Hoque et al. (2006a, b) and DeRose et al. (2008) studied the nanotribological properties of methyl- and perfluoro-terminated alkylphosphonate, perfluorodecyldimethylchlorosilane, and perfluorodecanoic acid on aluminum, of industrial interest. Hoque et al. (2007a) and DeRose et al. (2008) studied the nanotribological properties of alkylsilanes and perfluoroalkylsilanes on aluminum. Hoque et al. (2007b, 2008, 2009) studied the nanotribological properties of alkylphosphonate and perfluoroalkylsilane SAMs on copper. The authors found that these SAMs on aluminum and copper perform well irrespective of the substrate used. They confirmed the presence of respective films using X-ray photoelectron spectroscopy (XPS).

Hoque et al. (2007b, 2008, 2009) studied the chemical stability of various SAMs deposited on Cu substrates via exposure to various corrosive conditions. DeRose et al. (2008) studied the chemical stability of various SAMs deposited on Al substrates via exposure to corrosive conditions (aqueous nitric acid solutions of a low pH of 1.8 at temperatures ranging from 60 to 80 °C for times ranging from 30 to 70 min). The exposed samples were characterized by XPS and contact

angle measurements. They reported that perfluorodecanoic acid/Al is less stable than perfluorodecylphosphonate/Al and octadecylphosphonate/Al, but more stable than perfluorodecyldimethylchlorosilane/Al, which has implications in DMD applications, discussed earlier. In general, chemical stability data of various SAMs deposited on Cu and Al surfaces to corrosive environments has been reported by these authors. Based on these studies, it was concluded that chemisorption occurs at the interface and is responsible for strong interfacial bonds.

To date, the contact angle and nanotribological properties of alkanethiol, biphenylthiol, alkylsilane, perfluoroalkylsilane, alkylphosphonate and perfluoroalkylphosphane SAMs have been widely studied. In this chapter, we review, in some detail, the nanotribological properties of various SAMs having alkyl and biphenyl spacer chains with different surface terminal groups ($-\text{CH}_3$, $-\text{CF}_3$) and head groups ($-\text{S}-\text{H}$, $-\text{Si}-\text{O}-$, $-\text{OH}$, and $\text{P}-\text{O}-$) which have been investigated by AFM at various operating conditions, Fig. 13.8a, b (Bhushan and Liu 2001; Liu et al. 2001; Liu and Bhushan 2002a, b; Bhushan et al. 2005, 2006a; Kasai et al. 2005; Tambe and Bhushan 2005a; Tao and Bhushan 2005). Hexadecane thiol (HDT), 1,1'-biphenyl-4-thiol (BPT), and crosslinked BPT (BPTC) were deposited on Au(111) films on Si(111) substrates by immersing the substrate in a solution containing the precursor (ligand) that is reactive to the substrate surface. Crosslinked BPTC was produced by irradiation of BPT monolayers with low energy electrons. Perfluoroalkylsilane and alkylsilane SAMs were deposited on Si(100) by exposing the substrate to the vapor of the reactive chemical precursors. Perfluoroalkylphosphonate and alkylphosphonate SAMs were deposited on sputtered Al film on Si substrate as well as bulk Al substrates. Representative data follow.

13.4.1 Measurement Techniques

Experimental techniques used for measurement of the static contact angle, surface energy, adhesion, friction, and wear are described next.

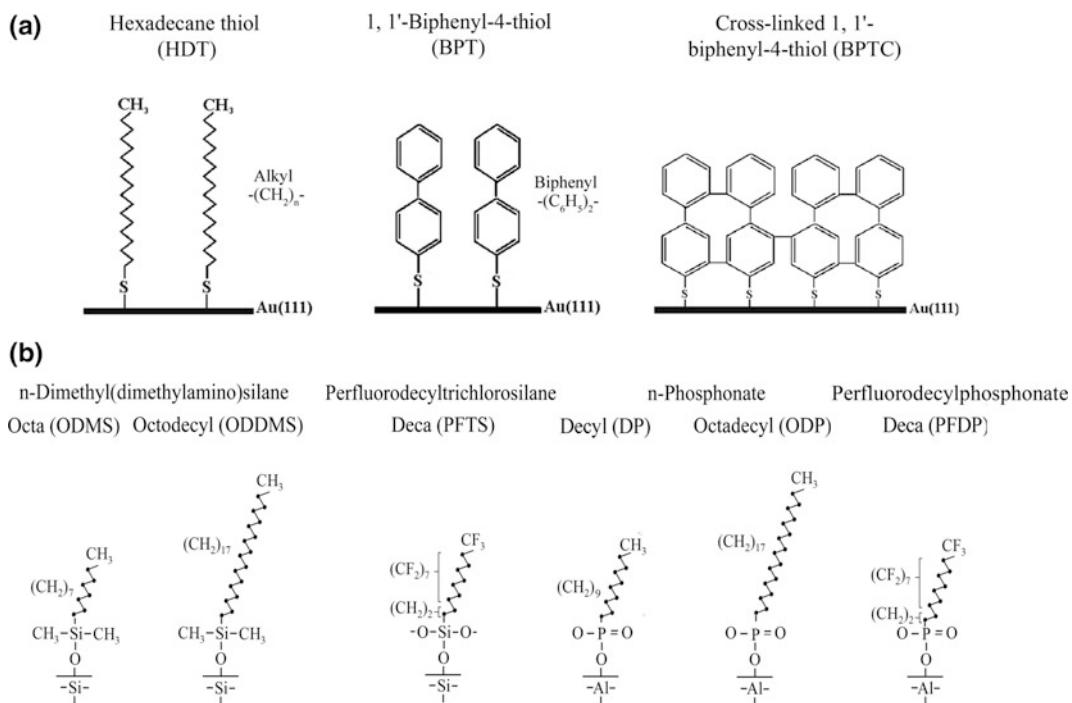


Fig. 13.8 Schematics of the structures of **a** hexadecane and biphenyl thiol SAMs on Au(111) substrates, and **b** perfluoroalkylsilane and alkylsilane SAMs on Si with native oxide substrates, and perfluoroalkylphosphonate and alkylphosphonate SAMs on Al with native oxide

13.4.1.1 Static Contact Angle and Surface Energy Measurements Using DI Water

The static contact angle, a measure of water repellent property, was measured using a Rame-Hart model 100 contact angle goniometer (Mountain Lakes, NJ, USA) (Good and Oss 1992; Adao et al. 1999). Typically, 10 μl droplets of DI water were used for making contact angle measurements. At least two measurements of the contact angle were made and were found to be reproducible within $\pm 2^\circ$. The critical surface tension, a measure of interfacial surface energy, was obtained from the so-called Zisman plot. Contact angles of SAMs with liquids with a range of surface tensions were measured. The cosines of the contact angles were plotted as a function of surface tension of the various n-alkane liquids (hexadecane, dodecane, undecane and decane) used. The plot is a linear line

for low polarizable SAMs. The horizontal intercept of the line passing through $\cos(\text{contact angle}) = 1$ provides the critical surface tension which is a measure of the surface energy of the SAM (Bhushan et al. 2005).

13.4.1.2 AFM Adhesion and Friction Measurements

The adhesion and friction tests were conducted using a commercial AFM system. Square pyramidal Si_3N_4 tips with a 30–50 nm tip radius on gold back-coated triangular Si_3N_4 cantilevers with a typical spring constant of 0.58 N/m were used. Adhesion can be calculated using either force calibration plots or from the negative intercepts on the friction force versus normal load plots. Both methods generally yield similar results. The force calibration plot technique was used in this study. The coefficient of friction was obtained from the slope of the friction force versus normal load plots. Normal loads typically

ranged from 5 to 100 nN. Friction force measurements were generally performed at a scan rate of 1 Hz along the fast scan axis and over a scan size of $2 \mu\text{m} \times 2 \mu\text{m}$. The fast scan axis was perpendicular to the longitudinal direction of the cantilever. The friction force was calibrated by the method described in Bhushan (1999a, 2008, 2011).

13.4.1.3 AFM Wear Measurements

Wear tests were conducted using a diamond tip with a nominal radius of 50 nm and nominal cantilever stiffness of 10 N/m. Wear tests were performed on a $1 \mu\text{m} \times 1 \mu\text{m}$ scan area at the desired normal load and at a scan rate of 1 Hz. After each wear test, a $3 \mu\text{m} \times 3 \mu\text{m}$ area was imaged, and the average wear depth was calculated.

13.4.1.4 Effect of Relative Humidity, Temperature, and Sliding Velocity

The influence of relative humidity on adhesive force, friction force, and wear was studied in an environmentally controlled chamber. Relative humidity was controlled by introducing a mixture of dry and moist air inside the chamber. The temperature was maintained at $22 \pm 1^\circ\text{C}$. The sample was kept in the environmental chamber at desired humidity for at least 2 h prior to the tests so that the system could reach equilibrium condition.

In order to study the effect of temperature on adhesion and friction force, the samples were placed on a thermal stage during the measurements. A glass plate was placed under the thermal stage to prevent the heat from being transported away. The temperature range studied was from 20 to 110°C . The relative humidity was maintained at $50 \pm 5\%$ during the temperature effect measurements.

The effect of sliding velocity on friction force was monitored in ambient conditions using a high velocity piezo stage designed for achieving high relative sliding velocities on a commercial AFM set up (Tambe and Bhushan 2005b). The

traveling distance of the sample, i.e., the scan size, was set at $25 \mu\text{m}$, while the scan frequency was varied between 0.1 Hz ($5 \mu\text{m/s}$) and 100 Hz ($5000 \mu\text{m/s}$).

13.4.1.5 Chemical Degradation and Environmental Studies

The chemical degradation experiments were carried out in a high vacuum tribotest apparatus (Bhushan and Ruan 1994; Bhushan et al. 1995c). The system was equipped with a mass spectrometer so that gaseous emissions from the interface could be monitored in situ during the sliding in high vacuum and other controlled environments. The normal loads and friction forces at the contacting interface were measured using resistive-type strain-gage transducers. For the sliding tests, the coated flat sample was glued on a flat surface at the end of a rotating shaft. The sample was slid against a Si(100) wafer mounted on the flat surface of a slider integrated with a flexible cantilever used in magnetic rigid disk drives. The sliding speed used was 0.3 m/s, and the applied pressure was 50 kPa. The environmental effects were investigated in high vacuum (2×10^{-7} Torr), argon, dry air (less than 2% RH), ambient air (30% RH), and high humidity air (70% RH).

13.4.2 Hexadecane Thiol and Biphenyl Thiol SAMs on Au(111)

Hexadecane thiol on Au(111) film was selected as it is a widely studied film. Biphenyl thiol was selected to study the effect of rigidity on nanotribological performance. Biphenyl thiol film was crosslinked to further increase its stiffness.

13.4.2.1 Surface Roughness, Adhesion, and Friction

Surface height and friction force images of SAMs were recorded simultaneously on an area of $1 \mu\text{m} \times 1 \mu\text{m}$ by an AFM, and adhesive forces were

measured by using the force calibration mode in an AFM (Bhushan and Liu 2001).

For further analysis, presented later in this chapter, the measured roughness, thickness, tilt angles, and spacer chain lengths of Si(111), Au(111) and various SAMs are listed in Table 13.5 (Bhushan and Liu 2001). The roughness of BPT is very close to that of Au(111), but the roughness of BPTC is lower than that of Au(111) and BPT; this is caused by electron irradiation. Table 13.5 indicates that the roughness value of HDT is much higher than its substrate roughness

of Au(111). This is caused by local aggregation of organic compounds on the substrates during SAMs deposition. Table 13.5 also indicates that the thickness of biphenyl thiol SAMs are generally smaller than the alkylthiol, because of shorter spacer chains in biphenyl thiol.

The average values and standard deviation of the adhesive force and coefficient of friction are presented in Fig. 13.9 (Bhushan and Liu 2001). Based on the data, the adhesive force and coefficient of friction of SAMs are less than those of their corresponding substrates. Among various films,

Table 13.5 The R_a roughness, thickness, tilt angles and spacer chain lengths of SAMs

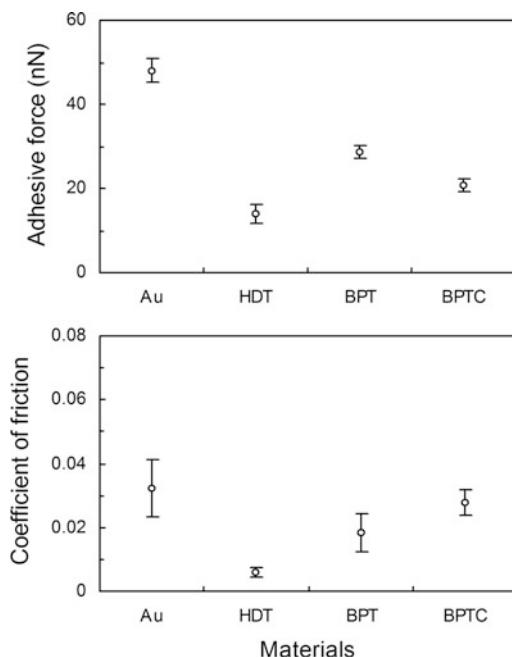
Samples	R_a roughness ^a (nm)	Thickness ^b (nm)	Tilt angle ^c (°)	Spacer length ^d (nm)
Si(111)	0.07			
Au(111)	0.37			
HDT	0.92	1.89	30	1.91
BPT	0.36	1.25	15	0.89
BPTC	0.14	1.14	25	0.89

^aMeasured by an AFM with 1 $\mu\text{m} \times 1 \mu\text{m}$ scan size, using Si_3N_4 tip under 3.3 nN normal load

^bThe thickness and tilt angles of BPT and BPTC are reported by Geyer et al. (1999). The thickness and tilt angles of HDT are reported by Ulman (1996)

^cThe spacer chain lengths of alkylthiols were calculated by the method reported by Miura et al. (1998). The spacer chain lengths of biphenyl thiols were calculated by the data reported by Ratajczak-Sitarz et al. (1987)

Fig. 13.9 Adhesive forces and coefficients of friction of Au(111) and various SAMs



HDT exhibits the lowest values. The ranking of adhesive forces F_a is in the following order: $F_{a-Au} > F_{a-BPT} > F_{a-BPTC} > F_{a-HDT}$, and the ranking of the coefficients of friction μ is in the following order: $\mu_{Au} > \mu_{BPTC} > \mu_{BPT} > \mu_{HDT}$. The ranking of various SAMs for adhesive force and coefficient of friction are similar. It suggests that alkylthiol and biphenyl thiol SAMs can be used as effective molecular lubricants for micro/nanodevices.

In micro/nano scale contact, liquid capillary condensation is one of the sources for the adhesion and friction. In the case of a sphere in contact with a flat surface, the attractive Laplace force caused by water capillary is (Bhushan 2013a, b),

$$F_L = 2\pi R\gamma_{la}(\cos \theta_1 + \cos \theta_2) \quad (13.1)$$

where R is the radius of the sphere, γ_{la} is the surface tension of the liquid against air, and θ_1 and θ_2 are the contact angles between liquid and flat and spherical surfaces, respectively. In an AFM adhesive study, the tip-flat sample contact is just like a sphere in contact with a flat surface, and the liquid is water. Since a single tip was used in the adhesion measurements, $\cos \theta_2$ can be treated as a constant. Therefore,

$$\begin{aligned} F_L &= 2\pi R\gamma_{la}(1 + \cos \theta_1) - 2\pi R\gamma_{la}(1 - \cos \theta_2) \\ &= 2\pi R\gamma_{la}(1 + \cos \theta_1) - C \end{aligned} \quad (13.2)$$

where C is a constant.

Based on the following Young-Dupre equation, the work of adhesion W_a (the work required to pull apart the unit area of the solid-liquid interface) can be written as (Israelachvili 1992)

$$W_a = \gamma_{la}(1 + \cos \theta_1) \quad (13.3)$$

It indicates that W_a is determined by the contact angle of SAMs, i.e. is influenced by the surface chemistry properties (polarization and hydrophobicity) of SAMs. By substituting Eq. (13.3) into Eq. (13.2), F_L can be expressed as

$$F_L = 2\pi RW_a - C \quad (13.4)$$

When the influence of other factors, such as van der Waals force, are very small on the

adhesive force, then adhesive force $F_a \approx F_L$. Thus the adhesive force F_a should be proportional to the work of adhesion W_a .

Contact angle is a measure of the wettability of a solid by a liquid and determines the W_a value (Good and Oss 1992; Adao et al. 1999). The static contact angles of distilled water on Au (111) and SAMs were measured and are summarized in Fig. 13.10a (Bhushan and Liu 2001). For water, $\gamma_{la} = 72.6 \text{ mJ/m}^2$ at 22 °C, and by using Eq. (13.3), the W_a data are obtained and presented in Fig. 13.10b. The W_a can be ranked in the following order: $W_{a-Au} (97.1) > W_{a-BPT} (86.8) > W_{a-BPTC} (82.1) > W_{a-HDT} (61.4)$. Except W_{a-Au} , this order exactly matches the order of adhesion force in Fig. 13.9. The relationship between F_a and W_a is summarized in Fig. 13.11 (Bhushan and Liu 2001). It indicates that the adhesive force F_a (nN) increases with the work of adhesion W_a (mJ/m^2) by the following linear relationship:

$$F_a = 0.57W_a - 22 \quad (13.5)$$

These experimental results agree well with the modeling prediction presented earlier in Eq. (13.4). It proves that on the nanoscale at ambient conditions the adhesive force of SAMs is mainly influenced by the water capillary force. Though neither HDT nor BPT has polar surface groups, the surface terminal of HDT has a symmetrical structure, which causes a smaller electrostatic attractive force and yields a smaller adhesive force than BPT. It is believed that the easy attachment of Au onto the tip should be one of the reasons that causes the large adhesive force, which do not fit the linear relationship described by Eq. (13.5).

13.4.2.2 Stiffness, Molecular Spring Model, and Micropatterned SAMs

Next the friction mechanisms of SAMs were examined. The Monte Carlo simulation of the mechanical relaxation of $\text{CH}_3(\text{CH}_2)_{15}\text{SH}$ self-assembled monolayer performed by Siepmann and McDonald (1993) indicated that SAMs compress

Fig. 13.10 **a** The static contact angle, and **b** work of adhesion of Au(111) and various SAMs. All of the points in this figure represent the mean value of six measurements. The uncertainty associated with the average contact angle is within $\pm 2^\circ$

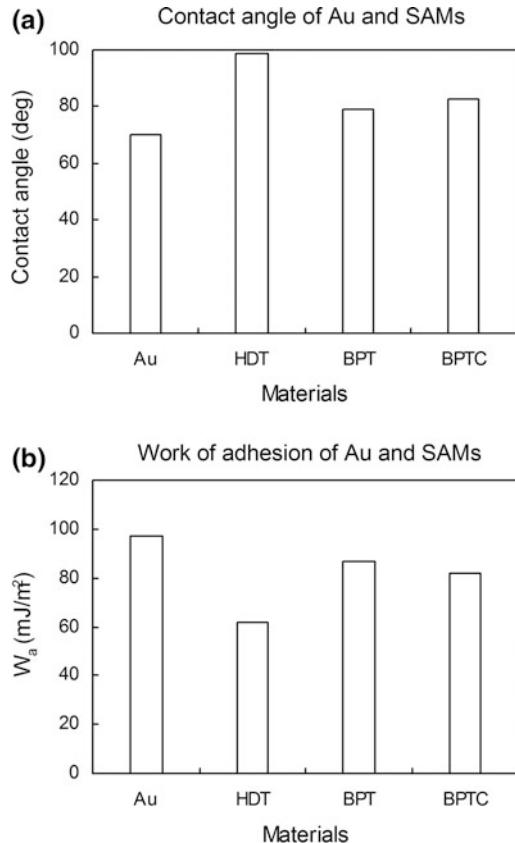
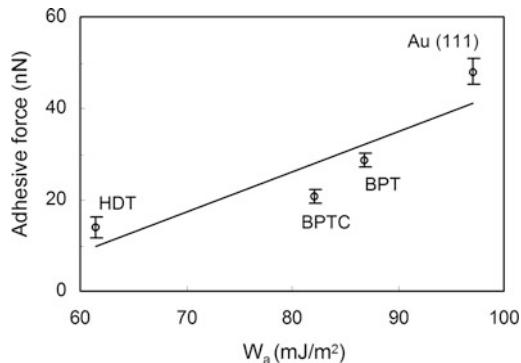


Fig. 13.11 Relationship between the adhesive forces and work of adhesion of different specimen



and respond nearly elastically to microindentation by an AFM tip when the load was below a critical normal load. Compression can lead to major changes in the mean molecular tilt (i.e. orientation), but the original structure is

recovered as the normal load is removed. Garcia-Parajo et al. (1997) also reported the compression and relaxation of octadecyltrichlorosilane (OTS) film during loading and unloading.

To study the difference in the stiffness of various films, the stiffness properties were measured by an AFM in force modulation mode (Bhushan 1999a, 2008, 2011; Liu and Bhushan 2002a). They reported that BPT was stiffer than HDT. Since BPT has a rigid benzene structure, it is more difficult to compress than HDT. Figure 13.12 shows the variation of the displacement with normal load during indentation mode. It clearly indicates that SAMs can be compressed. At a given normal load, long carbon chain structure SAMs such as HDT are easy to be compressed as compared to rigid benzene ring structure SAMs such as BPT, which implies that BPT is more rigid than HDT.

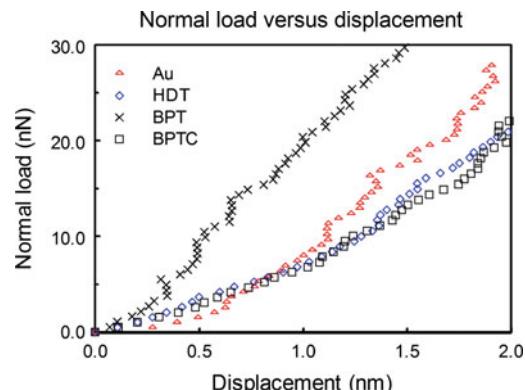
In order to explain the frictional difference of SAMs, based on the friction and stiffness measurements by AFM and the Monte Carlo simulation, a molecular spring model is presented in Fig. 13.13. It is believed that the self-assembled molecules on a substrate are just like assembled molecular springs anchored to the substrates (Bhushan and Liu 2001). An AFM tip sliding on the surface of SAMs is like a tip sliding on the top of “molecular springs or brush.” The molecular spring assembly has compliant features and can experience compression and orientation under normal load. The orientation of the molecular springs or brush 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 leads to energy dissipation. In the study of BPT by AFM, it was found that after the first several scans, the friction force is significantly reduced, but the surface height does not show any apparent change. This suggests that the molecular orientation can be facilitated by initial sliding and is reversible (Liu and Bhushan 2002b).

Based on the stiffness measurements obtained using Fig. 13.12 and the view of molecular structures in Fig. 13.13, biphenyl is a more rigid structure due to the contribution of two rigid benzene rings. Therefore the spring constant of BPT is larger than that of HDT. The hydrogen (H^+) in a biphenyl chain has an electrostatic attractive force with the π electrons in the neighboring benzene ring. Thus the intermolecular force between biphenyl chains is stronger than that for alkyl chains. The larger spring constant of BPT and stronger intermolecular force require a larger external force to allow it to orient, thus causing a higher coefficient of friction. The crosslinking of BPT leads to a larger packing energy for BPTC. Therefore it requires a larger external force to allow BPTC orientation, i.e., the coefficient of BPTC is higher than BPT.

An elegant way to demonstrate the influence of molecular stiffness on friction is to investigate SAMs with different structures on the same wafer.

Fig. 13.12 Normal load versus displacement curves of Au(111) and various SAMs



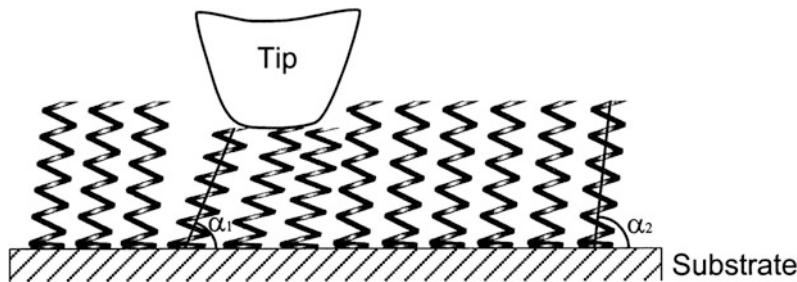


Fig. 13.13 Molecular spring model of SAMs. In this figure, $\alpha_1 < \alpha_2$, which is caused by the orientation under the normal load applied by AFM tip. The orientation of the molecular springs reduces the shearing force at the interface, which in turn reduces the friction force. The molecular spring constant, as well as the inter-molecular forces can determine the magnitude of the coefficients of friction of SAMs. In this figure, the size of the tip and molecular springs do not in the exactly scale (Bhushan and Liu 2001)

For this purpose, a micropatterned SAM was prepared. First the biphenyldimethylchlorosilane (BDCS) was deposited on silicon by a typical self-assembly method (Liu and Bhushan 2002a). Then the film was partially crosslinked using a mask technique by low energy electron irradiation. Finally the micropatterned BDCS films were realized, which had the as-deposited and cross-linked coating regions on the same wafer. The local stiffness properties of these micropatterned samples were investigated by force modulation AFM technique (DeVecchio and Bhushan 1997). The variation in the deflection amplitude provides a measure of the relative local stiffness of the surface. Surface height, stiffness, and friction images of the micropatterned biphenyldimethylchlorosilane (BDCS) specimen are obtained and presented in Fig. 13.14 (Liu and Bhushan 2002a). The circular areas correspond to the as-deposited film and the remaining area to the crosslinked film. Figure 13.14a indicates that crosslinking caused by the low energy electron irradiation leads to about a 0.5 nm decrease of the surface height of BDCS films. The corresponding stiffness images indicate that the crosslinked area has a higher stiffness than the as-deposited area. Figure 13.14b indicates that the as-deposited area (area with higher surface height) has a lower friction force. Obviously, these data of the micropatterned sample prove that the local stiffness of SAMs has an influence on their friction performance. Higher stiffness leads to larger friction force. These results

provide strong proof of the suggested molecular spring model.

In summary, it is found that SAMs exhibit compliance and can experience compression and orientation under normal load. The orientation of SAMs reduces the shear stress at the interface; therefore SAMs can serve as good lubricants. The molecular spring constant (local stiffness), as well as the intermolecular forces, can influence the magnitude of the coefficients of friction of SAMs.

13.4.2.3 Wear and Scratch Resistance

Wear resistance was studied on an area of $1 \mu\text{m} \times 1 \mu\text{m}$. The variation of wear depth with normal loads is presented in Fig. 13.15 (Bhushan and Liu 2001). HDT exhibits the best wear resistance. For all of the tested SAMs, in the wear depth as a function of normal load curves, there appears a critical normal load, which is marked by arrows in Fig. 13.15. When the normal load is smaller than the critical normal load, the monolayer only shows a slight height change in the scan areas. When the normal load is higher than the critical value, the height change of SAM increases dramatically. Relocation and accumulation of BPT molecules have been observed during the initial several scans, which lead to the formation of a larger terrace. Wear studies of a single BPT terrace indicate that the wear life of BPT increase exponentially with terrace size (Liu et al. 2001; Liu and Bhushan 2002a).

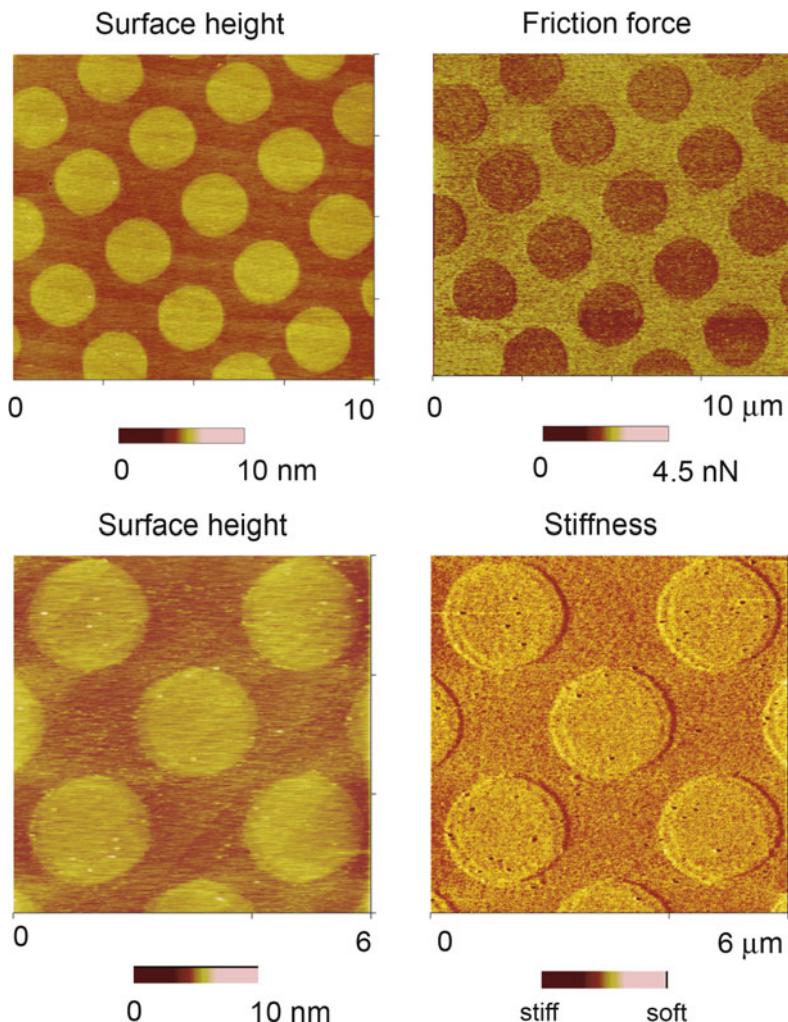
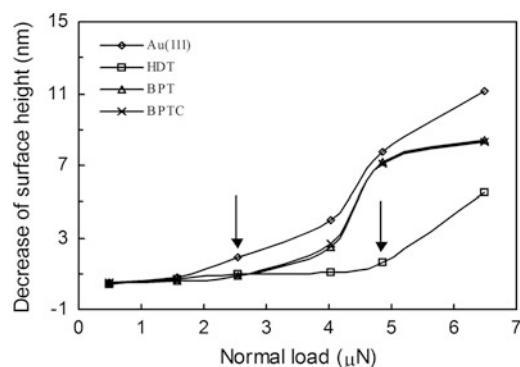


Fig. 13.14 **a** AFM Grayscale surface height and stiffness images, and **b** AFM grayscale surface height and friction force images of micropatterned BDCS (Liu and Bhushan 2002a)

Fig. 13.15 Wear depth as a function of normal load after one scan cycle



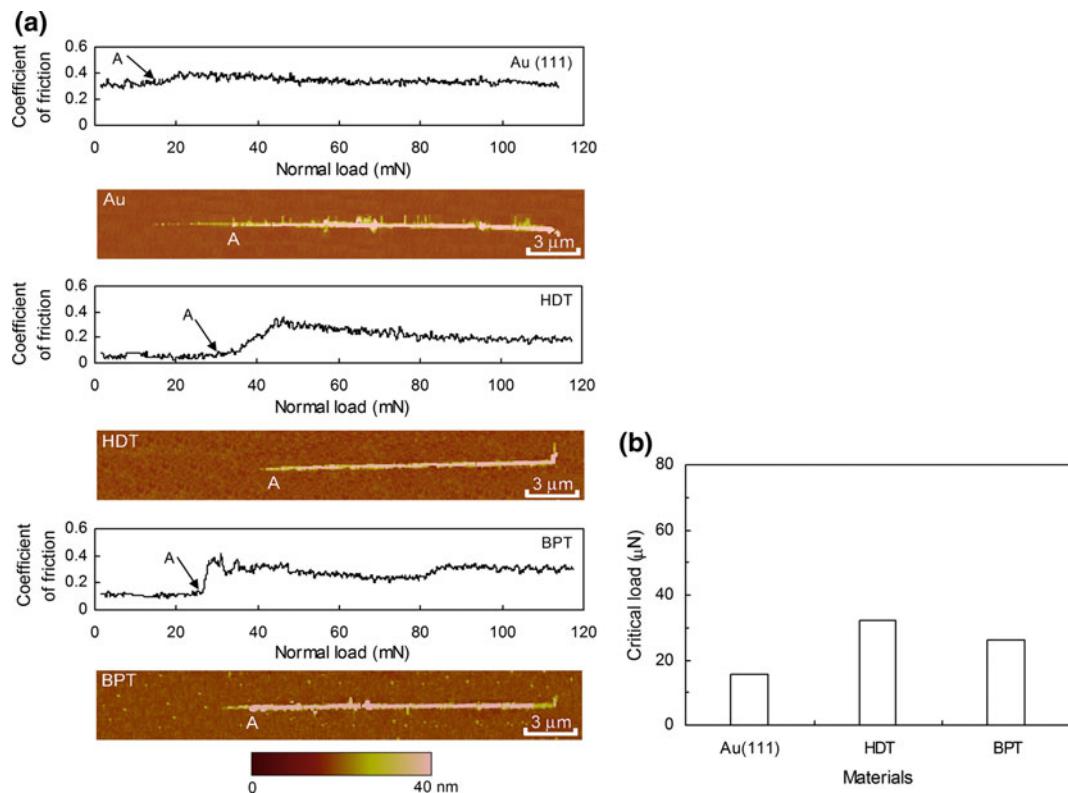


Fig. 13.16 **a** Coefficient of friction profiles during scratch as a function of normal load and corresponding AFM surface height images, and **b** critical loads estimated from the coefficient of friction profile and AFM images for Au(111), HDT/Au(111), and BPT/Au(111) (Liu and Bhushan 2002a)

Scratch resistance of Au(111) and SAMs were studied by a continuous AFM microscratch technique. Figure 13.16a shows coefficient of friction profiles as a function of increasing normal load and corresponding tapping mode AFM surface height images of the scratches captured on Au(111) and SAMs (Liu and Bhushan 2002a). Figure 13.16a indicates that there is an abrupt increase in the coefficient of friction for all of the tested samples. The normal load associated with this event is termed the critical load (indicated by the arrows labeled “A”). At the initial stages of the scratch, all the samples exhibit a low coefficient of friction indicating that the friction force is dominated by the shear component. This is in agreement with analysis of the AFM images, which shows negligible damage on the surfaces prior to the critical load. At the critical load, a clear groove is formed, which is accompanied by the formation

of material pileup at the sides of the scratch. This suggests that the initial damage that occurs at the critical load is due to plowing associated with plastic deformation, and this causes the sharp rise in the coefficient of friction. Beyond the critical load, debris can be seen in addition to material pileup at the sides of the scratch. Figure 13.16b summarizes the critical loads for the various samples obtained in this study. It clearly indicates that all SAMs can increase the critical load of the corresponding substrate.

The mechanisms responsible for a sudden drop in decrease in surface height with an increase in load during wear and scratch test need to be understood. Barrena et al. (1999) observed that the height of self-assembled alkylsilanes decrease in discrete amounts with normal load. This steplike behavior is due to the discrete molecular tilts, which are dictated by the

geometrical requirements of the close packing of molecules. Only certain angles are allowed due to the zigzag arrangement of the carbon atoms. The relative height of the monolayer under pressure can be calculated by the following equation

$$\left(\frac{h}{L}\right) = \left[1 + \left(\frac{na}{d}\right)^2\right]^{-1/2} \quad (13.6)$$

where L is the total length of the molecule, h is the height of the SAMs in the tilt configuration (monolayer thickness), a is the distance between alternate carbon atoms in the molecule, d is the separation of the molecules, and n is the step number. The values of a (0.25 nm) and d (0.47 nm) are used in the calculation for HDT. The calculated and measured relative heights of HDT are listed in Table 13.6. When the normal loads are smaller than the critical values in Fig. 13.15, the measured relative height values of HDT is very close to the calculated values. It means that HDT underwent step tilting below critical normal loads.

The residual SAMs thickness after wear under critical normal load was measured by profiling the worn film using AFM. The results are listed in Table 13.7. For an alkanethiol monolayer, the relationship between the monolayer thickness h and intercept length L_0 can be expressed as (see Fig. 13.17):

$$h = b \cos(\alpha)n + L_0 \quad (13.7)$$

where b is the length of the projection of the C–C bond onto the main chain axis ($b = 0.127$ nm for alkanethiol), n is the chain length defined by $\text{CH}_3(\text{CH}_2)_n\text{SH}$, and α is the tilt angle (Miura et al. 1998). For BPT and BPTC, based on the same principle, and using the bond lengths reported in reference (Ratajczak-Sitarz et al. 1987), the L_0 values are also calculated, Table 13.7. It indicates that the measured residual thickness values of SAMs under critical load are very close to the calculated intercept length L_0 values. It means that under the critical normal load, the Si_3N_4 tip approaches the interface, and SAMs wear severely away from the substrate.

Table 13.6 Calculated $\left[1 + \left(\frac{na}{d}\right)^2\right]^{-1/2}$ and measured $\left(\frac{h}{L}\right)$ relative heights of HDT self-assembled monolayer (Bhushan and Liu 2001)

Steps (n)	Calculated ^a $\left[1 + \left(\frac{na}{d}\right)^2\right]^{-1/2}$	Measured $\left(\frac{h}{L}\right)$
1	0.883	
2	0.685	0.674 ^b
3	0.531	0.532 ^b
4	0.425	0.416 ^b
5	0.352	0.354 ^b
6	0.299	

^aCalculations are based on the assumption that the molecules tilt in discrete steps (n), upon compression with a diamond AFM tip (Barrena et al. 1999)

^bThese measured values correspond to the normal loads of 0.50, 1.57, 2.53, and 4.03 μN , respectively

Table 13.7 Calculated L_0 and measured residual film thickness for SAMs under critical load

	L_0 (nm) ^a	Residual thickness (nm) ^b
HDT	0.24	0.25
BPT	0.39	0.42
BPTC	0.33	0.38

^aCalculated by the equation of $h = b \cos(\alpha)n + L_0$ (Miura et al. 1998)

^bMeasured by AFM using diamond tip under critical normal load. All of the data are the mean value of three tests

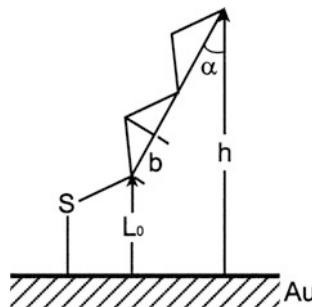


Fig. 13.17 Illustration of the relationship between the components of the equation $h = b \cos(\alpha)n + L_0$ (Bhushan and Liu 2001)

This is due to the interface chemical adsorption bond strength (S–Au) being generally smaller than the other chemical bond strengths in SAMs spacer chains (see Table 13.9, to be introduced later).

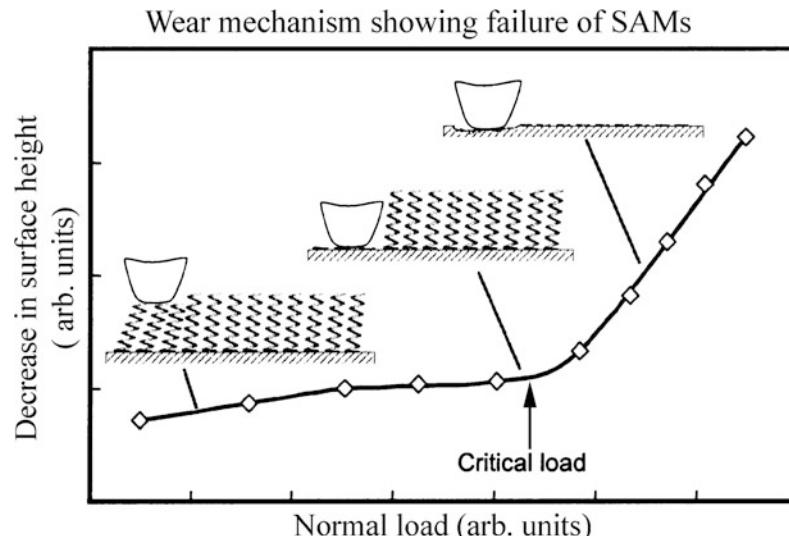
According to the wear and scratch results reported here and the above discussion, the transition of the wear mechanisms of SAMs with increasing normal load is illustrated in Fig. 13.18. Below the critical normal load SAMs undergo step orientation, at the critical load SAMs wear away from the substrate due to the weak interface bond strengths, while above the critical normal load severe wear take place on the substrate. In

order to improve wear resistance, the interface bond must be enhanced; a rigid spacer chain and a hard substrate are also preferred.

13.4.3 Perfluoroalkyl silane and Alkylsilane SAMs on Si(100), and Perfluoroalkylphosphonate and Alkylphosphonate SAMs on Al

Perfluorodecyltricholorosilane (PFTS), $\text{CF}_3-(\text{CF}_2)_7-(\text{CH}_2)_2-\text{SiCl}_3$, n-octyldimethyl (dimethyl amino)silane (ODMS), $\text{CH}_3-(\text{CH}_2)_n-\text{Si}(\text{CH}_3)_2-\text{N}(\text{CH}_3)_2$ ($n = 7$), and n-octadecyldimethyl (dimethylamino)silane ($n = 17$) (ODDMS) vapor deposited on Si(100) substrate and perfluorodecylphosphonate (PFDP) $\text{CF}_3-\text{O}-$
 $(\text{CF}_2)_7-\text{CH}_2-\overset{\text{O}}{\underset{\parallel}{\text{P}}}-\text{OH}$, decylphosphonate
 (DP) , $\text{CH}_3-(\text{CH}_2)_n-\overset{\text{O}}{\underset{\parallel}{\text{P}}}-\text{OH}$ ($n = 9$) and octadecylphosphonate (ODP) ($n = 17$) by liquid deposition on sputtered Al film on Si substrate were selected. Perfluoro SAMs were selected

Fig. 13.18 Illustration of the wear mechanisms of SAMs with increasing normal load (Liu and Bhushan 2002a)



because fluorinated films are known to have low surface energy. Two chain lengths of alkylsilanes (with 8 and 18 carbon atoms) were selected to compare their nanotribological performance with that of the former as well as to study the effect of chain length. Al substrate was selected because of the application of Al micromirrors in digital projection displays. Perfluoroalkylphosphane (with 10 carbon atoms) and alkylphosphonate SAMs (with 10 and 18 carbon atoms) on Al were selected.

13.4.3.1 Static Contact Angle and Surface Free Energy Measurements

Static contact angles and surface energy were measured to get a measure of the hydrophobicity. Figure 13.19 shows a Zisman plot for the SAMs deposited on Si and their surface energy values using various alkane liquids (Bhushan et al. 2005). Critical surface tension, a measure of surface energy values for the two SAMs, are presented in the figure. Zisman analysis for the Si substrate was not available because the alkane liquids used for the measurement instantly spread on these surfaces. For PFTS, significant reduction in the critical surface tension or surface

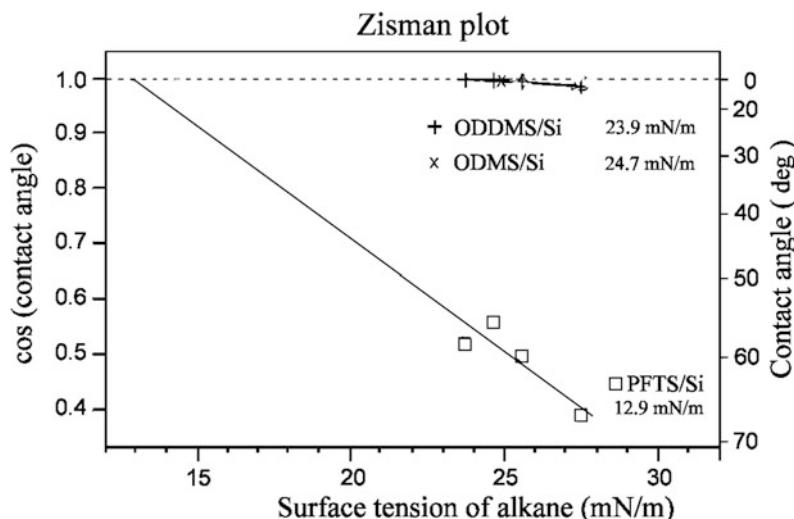


Fig. 13.19 Zisman plot for PFTS/Si, ODMS/Si, and ODDMS/Si used for calculating critical surface tension, a measure of surface energy, which is given by the x-intercept [$\cos(\text{contact angle}) = 1$] of the fitted line to the data

Table 13.8 A summary of RMS roughness, contact angle and film thickness of various SAMs

SAM/substrate	Acronym	RMS roughness (nm)	Contact angle (°)	Film thickness (nm)
Silicon(111) ^a	Si	0.07	48	–
Perfluorodecyltricholorosilane/Si ^a	PFTS/Si	0.09	112	~1.8
n-Octyldimethyl(dimethylamino)silane/Si ^a	ODMS/Si	0.10	99	~1.9
n-Octadecyldimethyl(dimethylamino)silane/Si ^b	ODDMS/Si	0.10	92	~2.1
Aluminum ^b	Al	32	<15	–
Perfluorodecylphosphonate	PFDP	34	137	~1.9
Decylphosphonate/Al ^b	DP/Al	31	129	~1.9
Octadecylphosphonate/Al ^b	ODP/Al	36	130	~2.1

^aKasai et al. (2005)

^bBhushan et al. (2006a)

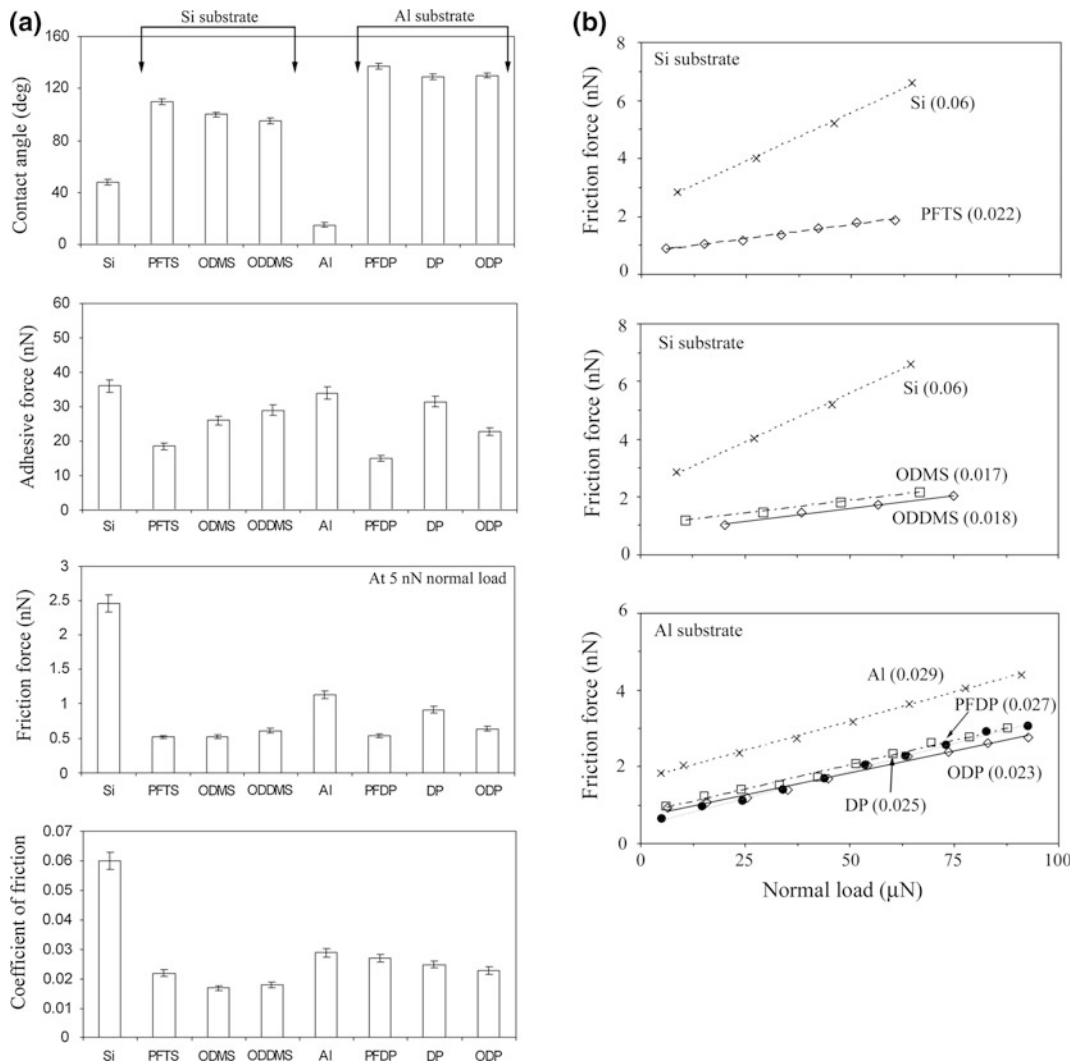


Fig. 13.20 **a** The static contact angle, adhesive force, friction force and coefficient of friction measured using an AFM for various SAMs on Si and Al substrates, and **b** friction force versus normal load plots for various SAMs on Si and Al substrates (Bhushan et al. 2005, 2006a)

energy was observed (12.9 mN/m for PFTS/Si), as compared to ODMS (24.7 mN/m for ODMS/Si) and ODDMS (23.9 mN/m for ODDMS/Si). The surface energy for ODMS and ODDMS was comparable. It suggests that the surface was comparably covered by the SAMs without bare substrate appearing.

For comparison, static contact angles were measured for various SAMs on Si and Al substrates. The measured values are compared among

the samples in Fig. 13.20a (Bhushan et al. 2005, 2006a). A summary of root-mean-square (RMS) roughness measured by using an AFM, and static contact angles and film thickness measured using an ellipsometer are summarized in Table 13.8 (Kasai et al. 2005; Bhushan et al. 2006a). Significant improvement in the water repellent property was observed for perfluorinated SAMs as compared to bare Si and Al substrates. Static contact angles of alkylsilanes and

alkylphosphonates were also higher than corresponding substrates, but lower than corresponding perfluorinated films. The contact angle generally increases with a decrease in surface energy (Eustathopoulos et al. 1999), which is consistent with the data obtained. The contact angles can be influenced by the packing density as well as the sample roughness (Ren et al. 2003). The higher contact angles for the SAMs deposited on Al substrates than those on Si substrate are probably due to this effect. The $-\text{CH}_3$ groups in ODMS, ODDMS, DP, and ODP are non-polar and are known to contribute to the water repellent property. Perfluorinated SAMs exhibited the highest contact angle among the SAMs tested in this study.

13.4.3.2 AFM Adhesion and Friction Measurements Under Ambient Conditions

Figure 13.20a shows the adhesive force, friction force, and the coefficient of friction measured under ambient conditions using an AFM, and Fig. 13.20b shows the friction force versus normal load plots for various SAMs deposited onto the Si and Al substrates (Bhushan et al. 2005, 2006a). Figure 13.21 shows surface height and friction force maps for Si and PFTS, ODMS, and ODDMS on Si (Bhushan et al. 2005).

The bare substrates showed higher adhesive force than the SAMs coatings. ODMS and ODDMS shows an adhesive force comparable to DP and ODP despite their lower water contact angles. These SAMs have the same tail groups, and during AFM measurements the AFM tip interacts only with the tail groups, whereas the contact angles can also be influenced by the head groups in these SAMs. This is probably the reason as to why the adhesive forces for these SAMs are comparable. PFTS and PFDP, which have the highest contact angles, showed the lowest adhesion.

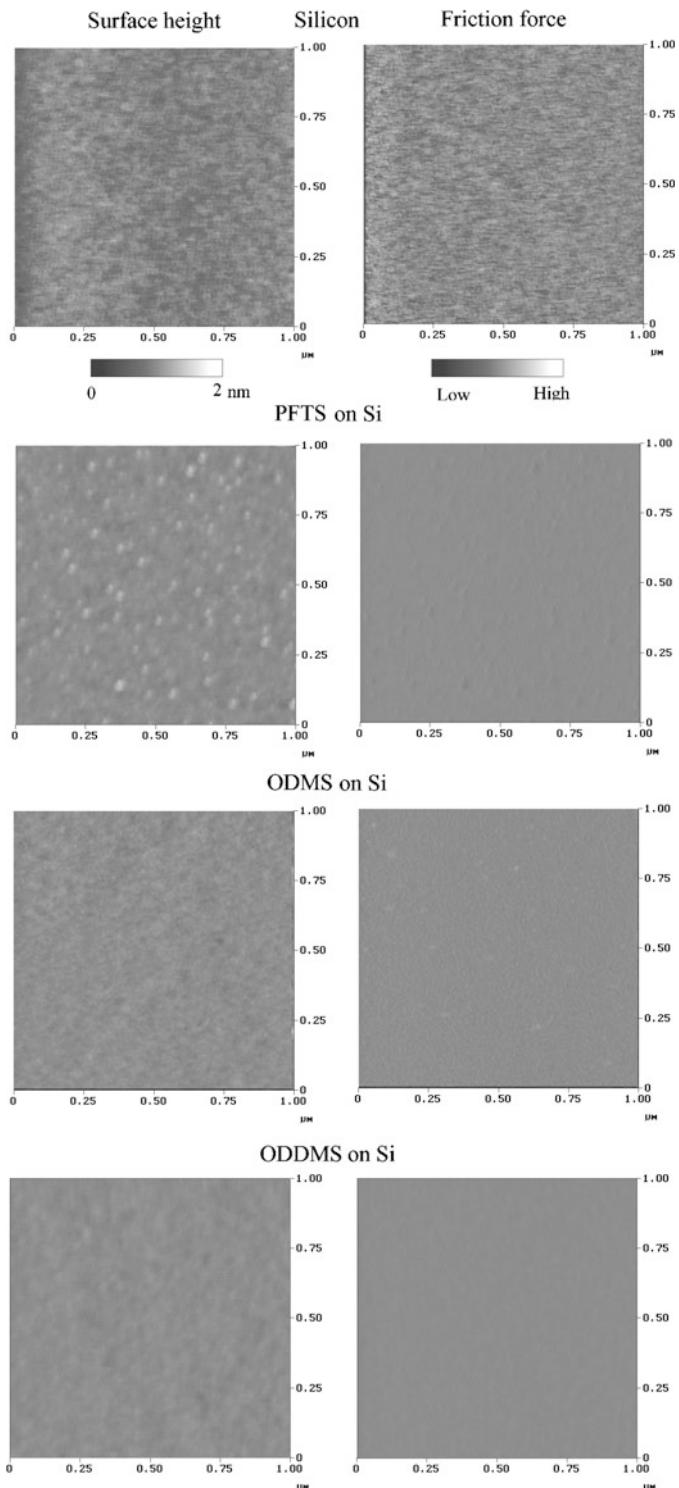
Friction force images of SAMs on Si exhibit more uniform contrast than those of bare Si. The coefficient of friction was higher for the bare substrates as compared to the corresponding SAMs deposited on them. The SAMs deposited on the Si substrate showed lower coefficient of friction than those deposited on the Al substrates.

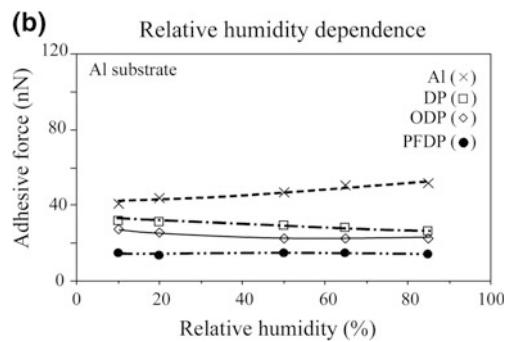
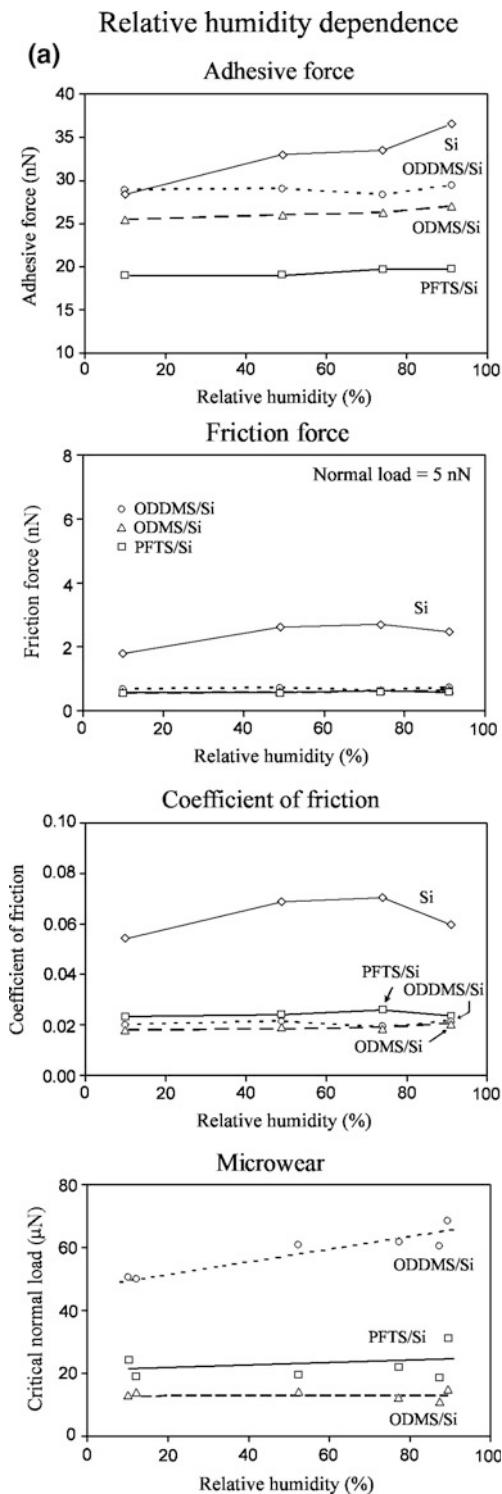
The primary reason for this is believed to be the higher roughness of the Al substrates. The SAMs with fluorocarbon backbone chains were found to have a higher coefficient of friction than those with hydrocarbon backbone chains. This might be attributed to the higher stiffness of the fluorocarbon backbone (Kasai et al. 2005; Bhushan et al. 2006a). For the fluorocarbon backbone chains, it is harder to rotate the backbone structure due to the larger size of the F atoms in comparison to the H atoms (Clark 1999). The C–C bonds of hydrocarbon chains can, on the other hand, rotate more freely. We have presented earlier a molecular spring or brush model to explain why less compliant SAMs show larger friction. The SAMs with a higher spring constant or stiffer backbone structure may need more energy to be elastically deformed during sliding; therefore friction is higher for these SAMs. In terms of the chain length effect, it has been reported that the coefficient of friction for SAM surfaces decreases with the carbon backbone chain length (n) up to 12 carbon atoms ($n \sim 12$) (McDermott et al. 1997). This chain length effect for the coefficient of friction was not obvious in these data.

13.4.3.3 Effect of Relative Humidity, Temperature, and Sliding Velocity on AFM Adhesion and Friction

The effect of relative humidity for various SAMs on adhesion and friction was studied. Adhesive force, friction force at 5 nN of normal load, coefficient of friction, and microwear data are presented in Fig. 13.22 (Kasai et al. 2005; Bhushan et al. 2006a). The result of adhesive force for silicon showed an increase with relative humidity, Fig. 13.22a. This is expected since the surface of silicon is hydrophilic, as shown in Fig. 13.20a. More condensation of water at the tip-sample interface at higher humidity increases the adhesive force due to capillary effect. On the other hand, the adhesive force for the SAMs showed a very weak dependency on the change in humidity. This occurs since the surface of the SAMs is hydrophobic. The adhesive force of

Fig. 13.21 Surface height and friction force maps for a Si, PFTS/Si, ODMS/Si, and ODDMS/Si (Bhushan et al. 2005)



**Fig. 13.22** (continued)

ODMS/Si and ODDMS/Si showed a slight increase from 75 to 90% RH. Such an increase was absent for PFTS/Si, possibly because of the hydrophobicity of PFTS/Si. The Al substrate is hydrophilic and hence shows relative humidity dependence, Fig. 13.22b. The PFDP, DP, and ODP SAMs deposited on Al substrates showed almost no change in adhesive force with humidity. The highly hydrophobic nature of these monolayers means that the contribution of water menisci to the overall adhesive force is negligible at all humidities.

The friction force of silicon showed an increase with relative humidity up to about 75% RH and a slight decrease beyond this point, see Fig. 13.22a. The initial increase can result from the increase in adhesive force. The decrease in friction force at higher humidity could be attributed to the lubricating effect of the water layer. This effect is more pronounced in the coefficient of friction. Since the adhesive force increased and coefficient of friction decreased in this range, those effects cancel each other out and the resulting friction force showed slight changes. On the other hand, the friction force and coefficient of friction of SAMs showed very small changes with relative humidity like that found for adhesive force. This suggests that the adsorbed water layer on the surface maintained a similar thickness throughout the relative humidity range tested. The differences among the SAM types were small within the measurement error, however a closer look at the coefficient of friction for ODMS/Si showed a slight increase from 75 to 90% RH as compared to PFTS/Si, possibly due to the same reason for the adhesive force increment. The inherent hydrophobicity of SAMs

Fig. 13.22 Relative humidity effect on **a** adhesive force, friction force, coefficient of friction and microwear for various SAMs on **(a)** Si substrates (Kasai et al. 2005), and **b** adhesive force for various SAMs on Al substrates (Bhushan et al. 2006a)

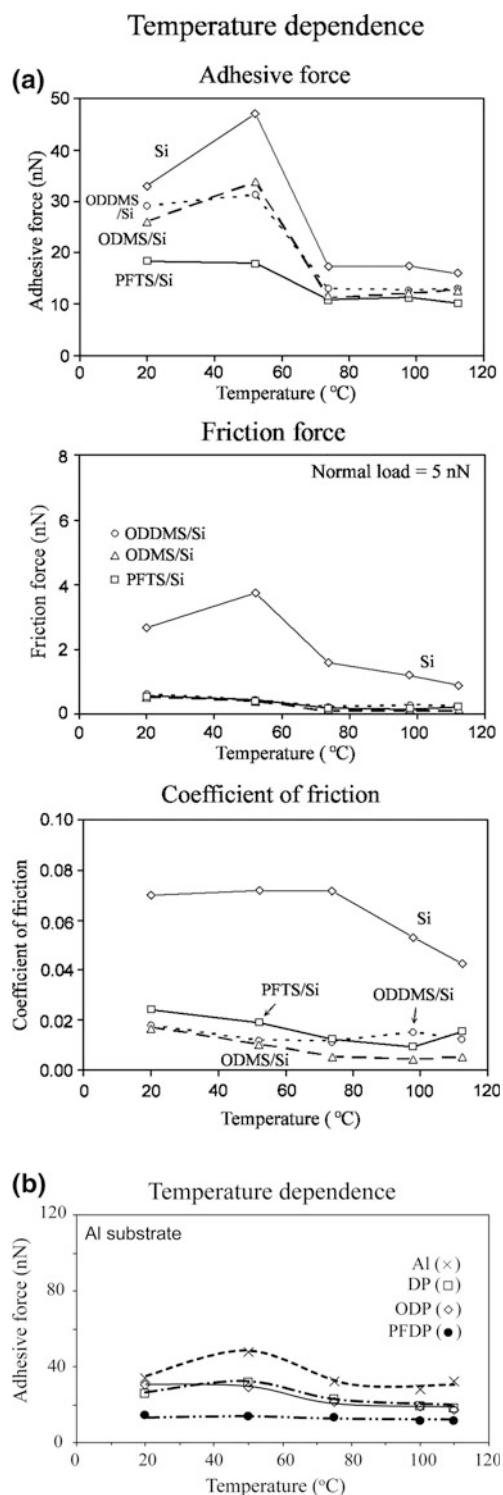


Fig. 13.23 Temperature effect **a** on adhesive force, friction force, and coefficient of friction for various SAMs on Si substrates (Kasai et al. 2005), and **b** on adhesive force for various SAMs on Al substrates (Bhushan et al. 2006a)

means that they did not show much relative humidity dependence.

Figure 13.23a shows the effect of temperature on adhesive force, friction force at 5 nN of normal load, and coefficient of friction for various SAMs on Si substrate (Kasai et al. 2005). Figure 13.23b shows the effect of temperature on the adhesive force for SAMs on Al (Bhushan et al. 2006a). The adhesive force for silicon showed an increase with the temperature, from room temperature (RT) to about 55 °C, followed by a decrease from 55 to 75 °C, and eventually leveled off from 75 to 110 °C. The adhesive force increased for Al substrate up to 50 °C and then decreased to a stable value for higher temperatures. The initial increase of adhesive force for Si and Al substrates at lower temperatures is not well understood. The observed decrease could be attributed to the desorption of water molecules on the surface. After almost full depletion of the water layer, the adhesive force remains constant. The SAMs with hydrocarbon backbones on both Si and Al substrates showed similar behavior as that of the Si and Al substrates, but the initial increase in the adhesive force with temperature was smaller. The SAMs with fluorocarbon backbone chains showed almost no temperature dependence. For the SAMs with hydrocarbon backbone chains, the initial increase in adhesive force is believed to be caused by the melting of the SAM film. The melting point for a linear carbon chain molecule such as $\text{CH}_3(\text{CH}_2)_{14}\text{CH}_2\text{OH}$ is 50 °C (Lide 2004). With an increase in temperature, the SAM film softens, thereby increasing the real area of contact and consequently the adhesive force. Once the temperature is higher than the melting point, the lubrication regime is changed from boundary lubrication in a solid SAM to liquid lubrication in the melted SAM (Liu and Bhushan 2002a).

The friction force for silicon showed an increase with temperature followed by a steady decrease. The friction force is highly affected by the change in adhesion. The decrease in friction can result from the depletion of the water layer. The coefficient of friction for silicon remained constant followed by a decrease starting at about 80 °C. For SAMs, the coefficient of friction exhibited a monotonic decrease with temperature. The decrease in friction and coefficient of

friction for SAMs possibly results from the decrease in stiffness. As introduced before, the spring model suggests a smaller friction for more compliant SAMs (Liu and Bhushan 2002a). The difference among the SAM types was not significant. PFTS could maintain its stiffness more than ODMS and ODDMS when temperature is increased (Callister 1997), however, it was not pronounced in the results.

Figure 13.24a shows the effect of sliding velocity on adhesive force, friction force and coefficient of friction for various SAMs on a Si substrate (Kasai et al. 2005). The adhesive force for silicon remained rather constant at a lower sliding velocity, and then increased rapidly. A similar trend was found for the SAMs. The increase in adhesive force for silicon is believed to be the result of a tribochemical reaction at the tip and sample interface (Liu and Bhushan 2003a) and increase of contact area by mechanical plowing. For the SAMs, the higher adhesive force at the higher velocity can result from viscous drag of the SAM molecules (Tambe and Bhushan 2005c). SAMs can be detached from the surface and attached to an AFM tip. In addition, another reason may be an increase of contact area, which may be caused by more penetration of the AFM tip into the SAMs. The rate of increase was larger for ODMS than PFTS presumably because of the higher stiffness and more densely packed structure of PFTS.

Coefficient of friction showed an increase with sliding velocity and reached a plateau for Si, ODMS/Si, and ODDMS/Si. As the sliding speed is increased, the reorientation of the SAMs needs additional work, which might lead to increase in friction. For PFTS, coefficient of friction decreased at larger sliding velocity forming a peak. The peak structure may result from the viscoelastic property of SAMs (Clear and Nealey 2001).

Figure 13.24b shows the effect of sliding velocity on the friction force for various SAMs on Al substrate (Bhushan et al. 2006a). Friction force increases slowly at lower sliding velocities for the bare Al substrate and SAMs, followed by a rapid increase at higher sliding velocities, except for PFDP. The increase in friction force at high velocities (>1 mm/s) is the result of asperity impacts and a corresponding high frictional

Sliding velocity dependence

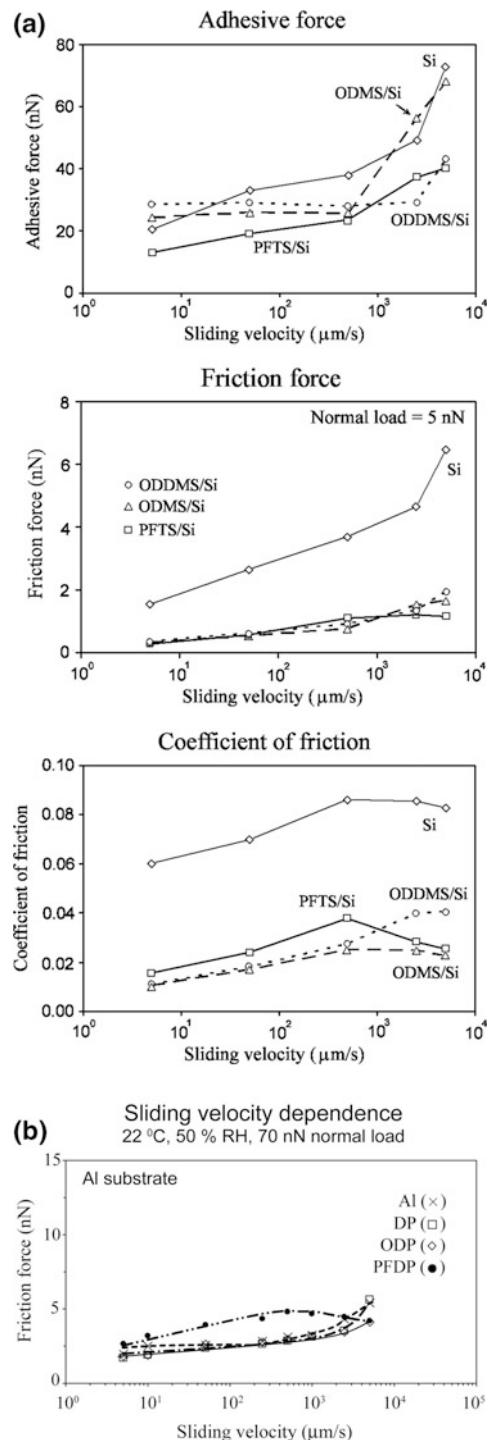


Fig. 13.24 Sliding velocity effect **a** on adhesive force, friction force, and coefficient of friction for various SAMs on Si substrates (Kasai et al. 2005), and **b** on friction force for various SAMs on Al substrates (Bhushan et al. 2006a)

energy dissipation at the sliding interface for Al (Tambe and Bhushan 2005d). Tambe and Bhushan (2005c) extended the “molecular spring” model presented by Bhushan and Liu (2001) to explain this velocity dependent increase in friction force for compliant SAM molecules. Based on this model for the DP and ODP SAMs, the increase in friction force is believed to result from the reorientation of the SAM molecules under the tip load and during tip motion. The reorientation of the SAMs can act as an additional hindrance to tip motion when the AFM tip reverses during scanning and thus result in higher friction. The molecules can consequently get entangled and/or get detached from the substrate and attach to the AFM tip.

13.4.3.4 AFM Wear Measurements

Figure 13.25a shows the relationship between the decrease in surface height as a function of the normal load during wear tests (Kasai et al. 2005; Bhushan et al. 2006a). As shown in the figure, the SAMs exhibit a critical normal load, beyond the point of which the surface height drastically decreases. Figure 13.25a also shows the wear behavior of the Al and Si substrates. Unlike the SAMs, the substrates show a monotonic decrease in surface height with the increasing normal load with wear initiating from the very beginning, i.e., even for low normal loads. Si (Young’s modulus of elasticity, $E = 130$ GPa (Anonymous 1988), hardness, $H = 11$ GPa (Bhushan 1999b) is relatively hard in comparison to Al ($E = 77$ GPa, $H = 0.41$ GPa), and hence the decrease in surface height for Al is much larger than that for Si for similar normal loads.

The critical loads corresponding to the sudden failure of SAMs are shown in Fig. 13.25b. Amongst all the SAMs, ODDMS shows the best performance in the wear tests, and this is believed to be because of the longer chain length effect. Fluorinated SAMs—PFTS and PFDP show a higher critical load as compared to ODMS and DP with similar chain lengths. ODP shows a higher critical load as compared to DP because of its longer chain length. The mechanism of failure of compliant SAMs during wear tests has been presented earlier in Fig. 13.18. It is

believed that the SAMs fail mostly due to shearing of the molecule at the head group, that is, by means of shearing of the molecules off the substrate. Table 13.9 gives the bond strengths for various intermolecular bonds. The weakest bonds are at the interface, and hence failure is expected to be initiated at the interface.

To study the effect of relative humidity, wear tests were performed at various humidities. The bottom of Fig. 13.22a shows critical normal load as a function of relative humidity. The critical normal load showed weak dependency on relative humidity for ODMS/Si and PFTS/Si and was larger for ODMS/Si than PFTS/Si throughout the humidity range. This suggests that water molecules could penetrate into the ODDMS, which might work as lubricant (Tian et al. 1999; Liu and Bhushan 2002a). This effect was absent for PFTS/Si and ODMS/Si.

13.4.4 Chemical Degradation and Environmental Studies

Chemical degradation and environmental studies were carried out for HDT/Au, PFTS/Si, ODMS/Si, and ODDMS/Si films.

13.4.4.1 Chemical Degradation Studies

The coefficient of friction and detected gaseous products for HDT/Au are shown in Fig. 13.26a (Tao and Bhushan 2005). A normal pressure of 50 kPa was applied on HDT films. The coefficient of friction increased after a sliding distance of about 10 m. During sliding, $(CH_2)_{15}S$, C_2H_3 , CH_3 , CH_2 , and H_2 were detected by mass spectrometer. Partial pressure of HS fragments is of interest as it corresponds to the interface bonds, and it is reported here. Increase of $(CH_2)_{15}S$ was much more than that of other species, due to the breaking of the S–Au bond. Partial pressures of C_2H_3 , CH_3 , CH_2 , and H_2 were also found to increase during the sliding. There was no noticeable change in the partial pressure of HS.

HDT film is deposited on a Au(111) layer. The bond strength of S–Au is 184 kJ/mol (see Table 13.9), which is lower than that of C–C

Fig. 13.25 **a** Decrease in surface height as a function of normal load after one scan cycle for various SAMs on Si and Al substrates, and **b** comparison of critical loads for failure during wear tests for various SAMs (Kasai et al. 2005; Bhushan et al. 2006a)

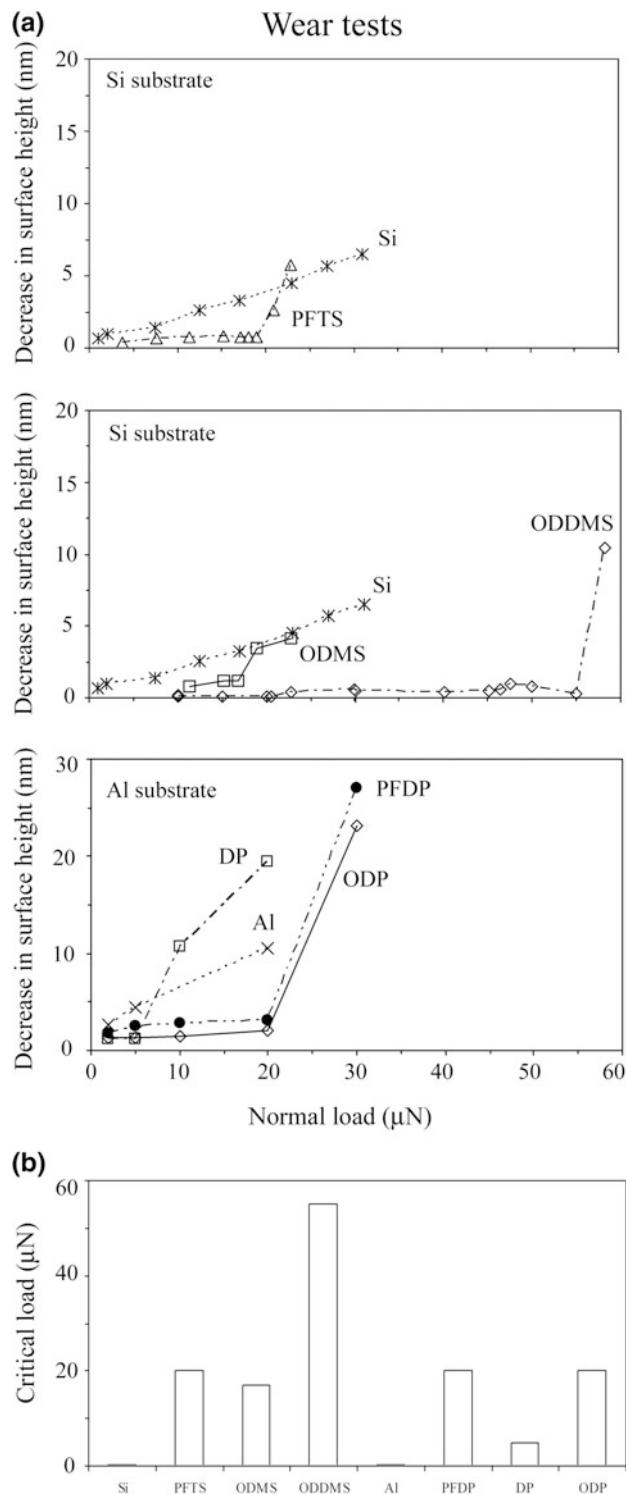


Table 13.9 Typical bond strengths^a in SAMs

SAMs	Bond	Hexadecane thiol (HDT) (kJ/mol)	Biphenyl thiol (BPT) (kJ/mol)	Bond	Perfluorodetyl- trichlorosilane(PFTS) (kJ/mol)	Alkylsilane (ODMS or ODDMS) (kJ/mol)	Perfluoro- decylphosphonate (PFDP) (kJ/mol)	Alkylphosphonate (DP and ODP) (kJ/mol)
Interfacial bonds	S–Au	184 ^b	184 ^b	Si–O	242 ^c	242 ^c		
	S–C	286 ^a		Si–C	800 ^d	800 ^d		
	C ₆ H ₅ –S		362 ^a		414 ^a	414 ^a		
				Al–O		511 ^a		
				P–C		513 ^a		
				P–O		599 ^a		
				C–C				
				CH ₂ –CH ₂	326 ^e		511 ^a	
				CF ₂ –CF ₂	~326 ^f		513 ^a	
				CF ₂ –CH ₂	~326 ^f		599 ^a	
				CF ₃ –CF ₂	~326 ^f			
				CH ₃ –CH ₂	~305 ^a			
						~305 ^a		
							~305 ^a	

^aLide (2004)^bChemical adsorption bond from Lio et al. (1997)^cChemical adsorption bond from Hoshino (1999)^dIn diatomic molecules^eCottrell (1958)^fDue to C–C bond, should be close to that of CH₂–CH₂

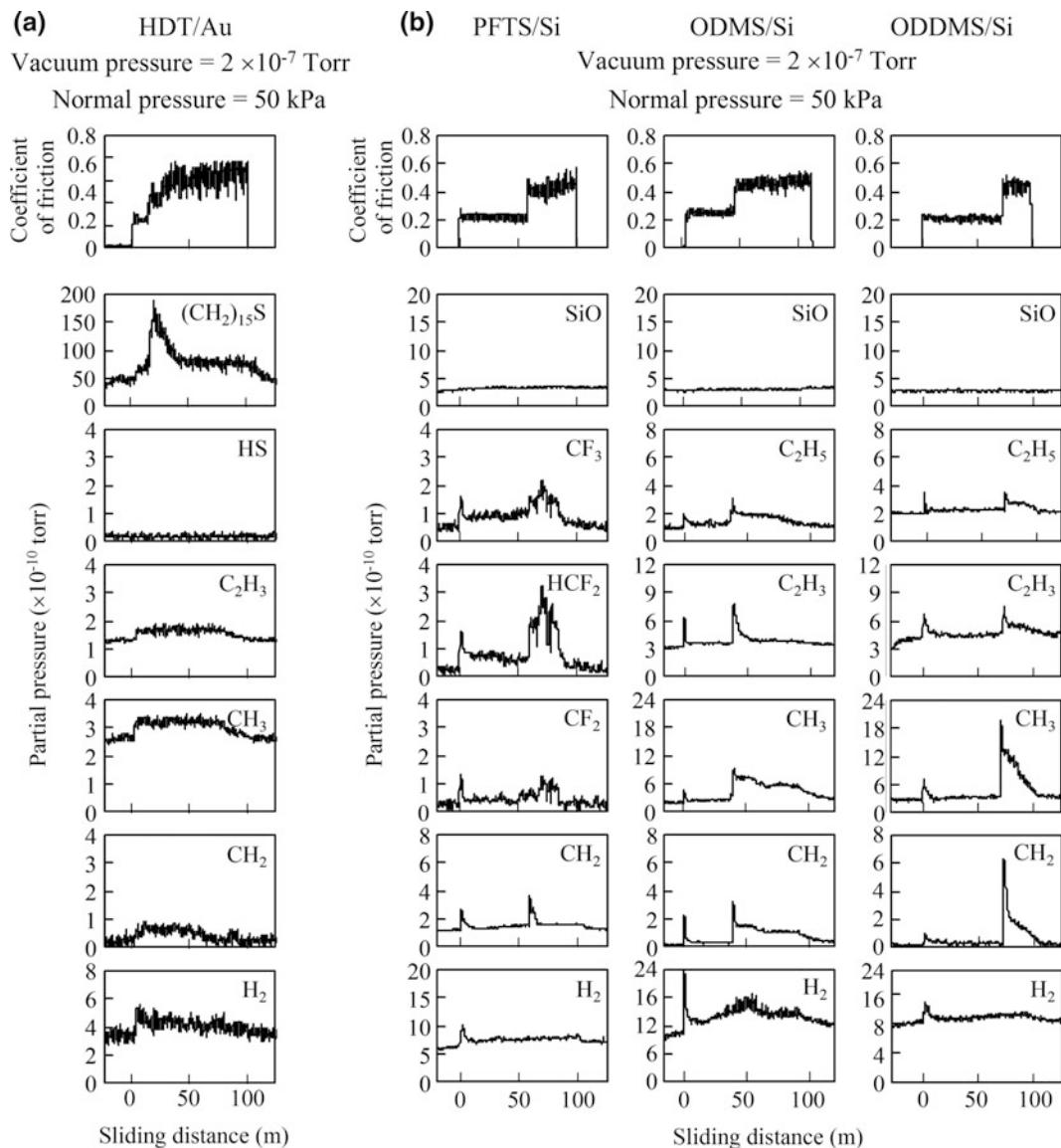


Fig. 13.26 Coefficients of friction and mass spectra data on **a** HDT/Au (1.9 nm), and **b** PFTS/Si (1.8 nm), ODMS/Si (~1.9 nm), and ODDMS/Si (~2.1 nm) in high vacuum (Tao and Bhushan 2005)

bond (425 kJ/mol), C–H bond (422 kJ/mol), and C–S bond (286 kJ/mol) in the alkyl chain. Since the S–Au bond is the weakest bond in the alkanethiol chain, the whole chain should be sheared away from the substrate. Because the upper atomic mass unit (amu) limit of the mass

spectrometer used is 250, we monitored $(\text{CH}_2)_{15}\text{S}$ (amu = 242), which is the chain with CH_3 sheared. The rate of generation of $(\text{CH}_2)_{15}\text{S}$ is much larger than that of other species. This suggests that the mechanical shear of the whole alkanethiol chain is the dominant factor causing

the failure of the HDT film. The cleavage of the S–Au bonds has been reported in literature. Based on the bond strength as well as the above studies, mechanical shearing of the C–C bonds and C–H bonds does not likely happen during sliding. The reaction induced by low-energy electrons, generated by triboelectrical emission during the sliding, could be responsible for the degradation of the alkanethiol chain. Thermal desorption of HDT from Au is another possibility of degradation mechanism of HDT.

The coefficient of friction and generated gaseous products for PFTS/Si, ODMS/Si, and ODDMS/Si are shown in Fig. 13.26b (Tao and Bhushan 2005). The coefficients of friction for PFTS/Si, ODMS/Si, and ODDMS/Si increase sharply after a certain sliding distance, which indicates the degradation of the film. At the same time, gaseous products of CF_3 , HCF_2 , CF_2 , CH_2 , and H_2 were detected for PFTS/Si, and C_2H_5 , C_2H_3 , CH_3 , CH_2 , and H_2 were detected for ODMS/Si, and ODDMS/Si.

PFTS/Si showed lower friction than ODMS/Si in the tests. ODDMS/Si showed lower friction than both PFTS/Si and ODMS/Si. This is because of the chain length effect. It has been reported that the coefficient of friction for SAM surfaces decreases with the carbon backbone chain length (n) when the carbon atoms are less than 12. For chains with more than 12 carbons, the change of carbon atoms will not influence the coefficient of friction in a noticeable extent.

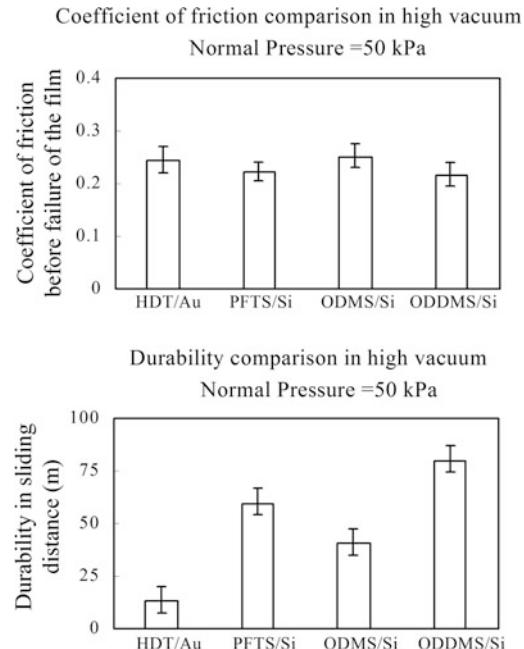
PFTS/Si showed a higher durability than ODMS/Si. For the case of a perfluorinated carbon backbone, it is harder to rotate the backbone structure (due to the different size of F atoms versus H atoms) which implies that this structure is more rigid than a hydrocarbon backbone (Clark 1999). Chambers (1973) has reported that the C–C bond strength increases when hydrogen is replaced with fluorine. This suggests that the rigid perfluorinated carbon backbone may be responsible for the increased durability. The length of the alkyl chain also influences the desorption energy

of alkanes. Based on studies of the adsorption of alkanes on Cu(100), Au(111), Pt(110), and others, the physisorption energy increases with the alkyl chain length (Sexton and Hughes 1984; Dubois et al. 1990; McMaster et al. 1993). Therefore, ODDMS are more durable than ODMS.

During sliding on PFTS films, gaseous products of CF_3 , HCF_2 , CF_2 , CH_2 , and H_2 were detected. From the structure of perfluoroalkylsilane, the only source of H on the molecular chain which would cause a partial pressure increase of H_2 is from the $(\text{CH}_2)_2$, which is located at the bottom of the chain. Since the partial pressure of H_2 increases immediately after sliding and remains high until the end of sliding, it is probably generated by the low-energy electrons coming from a triboelectrical emission. The partial pressure of CH_2 had a sharp peak at the beginning of sliding and at the moment when friction changes. Meanwhile, the partial pressures of CH_3 , HSF_2 , and CF_2 increased significantly when the friction increased. For ODMS and ODDMS, C_2H_5 , C_2H_3 , CH_3 , CH_2 , and H_2 were detected during sliding. The partial pressure of the carbon-related products increase considerably when the friction is increased. The SiO, which is related to the interface bonds, has no noticeable change during sliding.

Perfluoroalkylsilanes and alkylsilanes are attached to the naturally oxidized silicon by Si–O bonds. The Si–O bond strength varies over a large range (see Table 13.9), depending on the forming condition. In the alkylsilane chain, C–Si bond strength (414 kJ/mol) is slightly lower than C–C bond strength. Based on Table 13.9, interfacial bonds (Si–O) are weaker than the C–C bonds in the backbone. It is believed that the cleavage of films occurs at the interface. We have previously reported evidence of the cleavage of the interfacial bonds using an AFM. To explain the hydrogen, C_1 and C_2 hydrocarbon (in the tests for PFTS/Si, ODMS/Si, and ODDMS/Si) or fluorocarbon (in the tests for PFTS/Si) products, Kluth et al. (1998) suggested that the alkylsilane

Fig. 13.27 Coefficient of friction and durability comparison of HDT/Au, PFTS/Si, ODMS/Si, and ODDMS/Si in high vacuum. Error bars represent $\pm 3\sigma$ based on five measurements (Tao and Bhushan 2005)



(perfluoroalkylsilane as well) chains break and create radicals. The radical could remain on the surface and decompose to generate a shorter radical and an alkene. The radical could repeatedly decompose to ever short radicals and alkenes as long as it remains on the surface.

A summary of the coefficients of friction and durability of all films in vacuum is presented in Fig. 13.27 (Tao and Bhushan 2005).

13.4.4.2 Environmental Studies

To study the effect of environment, the friction tests on HDT/Au, PFTS/Si, ODMS/Si, and ODDMS/Si were conducted in high vacuum, argon, dry air (less than 2% RH), air with 30% RH, and air with 70% RH; see Fig. 13.28 (Tao and Bhushan 2005). By comparing the coefficient of friction in different environments, it was found that the friction in argon is the lowest for the SAMs tested. In high vacuum, the intimate contact leads to high friction. In dry air, the friction is higher than in the argon. This shows that oxygen has an apparent effect on the performance of SAMs. Kim et al. (2003) studied the thermal stability of alkylsiloxane SAMs in air. They found that the

alkylsiloxane decomposes at about 200 °C, which is much lower than the decomposition temperature of 470 °C in vacuum reported by Kluth et al. (1998). This difference could be attributed to the oxygen in air. The water contained in the air is found to have a significant influence on the friction of SAMs. A humidity effects study of alkylsilane on mica substrate by Tian et al. (1999) have indicated that the water molecules can penetrate the alkylsilane film, which alters their molecular chain ordering and can also detach the alkylsilane molecules from the substrate.

A summary of the coefficients of friction before the failure of the lubricant films in various environments is presented in Fig. 13.29 (Tao and Bhushan 2005). The data in Fig. 13.29 are average values based on five measurements. To summarize the highlights, friction of tested lubricant films is high in high vacuum because of the intimate contact between the lubricants and the counterpart surface. Friction of tested lubricant films is lower in argon than that in dry air. Friction of SAMs is significantly influenced by water molecules.

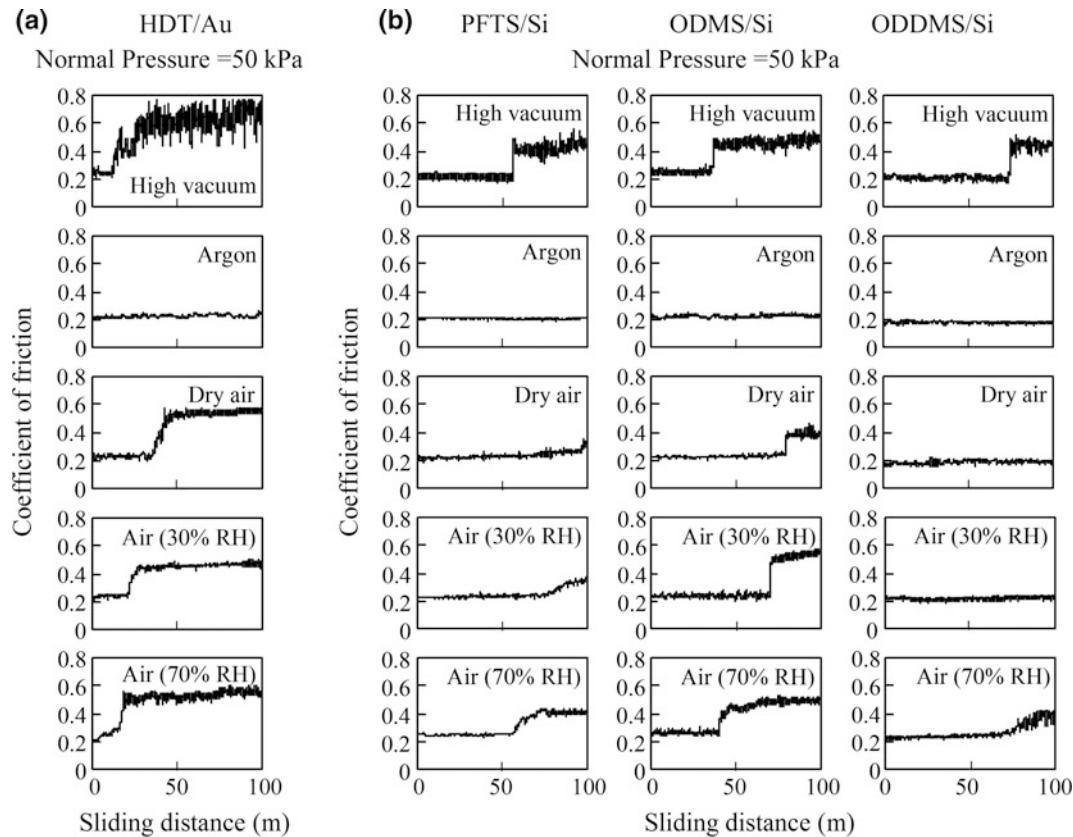
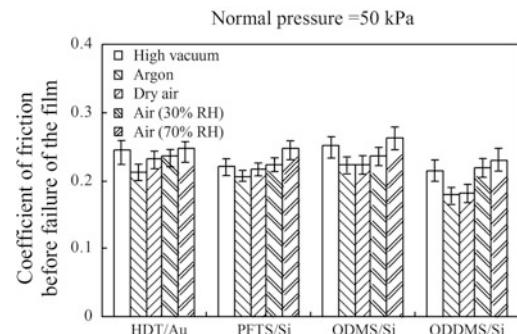


Fig. 13.28 Coefficient of friction data in high vacuum, argon, and air with different humidity levels of **a** HDT/Au (1.9 nm) and **b** PFTS/Si (1.8 nm), ODMS/Si (\sim 1.9 nm), and ODDMS/Si (\sim 2.1 nm) (Tao and Bhushan 2005)

Fig. 13.29 Comparison of coefficient of friction data for HDT/Au, PFTS/Si, ODMS/Si, and ODDMS/Si in high vacuum, argon, and air with different humidity levels. Error bars represent $\pm 3\sigma$ based on five measurements (Tao and Bhushan 2005)

Coefficient of friction comparison in various environments



13.5 Closure

The adhesive force can be significant in an interface with ultrasmooth surfaces and it can be on the same order as the external load if the latter is small,

such as in micro- and nanodevices. High adhesion also leads to stiction, friction, and wear issues in sliding surfaces. In various applications, surfaces need to be protected from exposure to the operating environment, and hydrophobic films are of interest. Films are also needed to provide low

adhesion, stiction, friction, and wear. These films should be molecularly thick, well-organized, chemically bonded to the substrate, and insensitive to environment. Ordered molecular assemblies with high hydrophobicity can be engineered using chemical grafting of various polymer molecules with suitable functional head groups and nonpolar surface terminal groups.

The contact angle, adhesion, friction, and wear properties of SAMs having alkyl, biphenyl, and perfluoroalkyl spacer chains with different surface terminal groups ($-\text{CH}_3$ and $-\text{CF}_3$) and head groups ($-\text{SH}$, $-\text{Si}-\text{O}-$, $-\text{OH}$, and $\text{P}-\text{O}-$) were presented in this chapter. It was found that the adhesive force varies linearly with the work of adhesion value of SAMs, which indicates that capillary condensation of water plays an important role to the adhesion of SAMs on the nanoscale at ambient conditions. SAMs with high-compliance long carbon spacer chains exhibited the lowest adhesive force and friction force. The friction data were explained using a molecular spring model, in which the local stiffness and intermolecular force govern frictional performance. The results of the stiffness and friction characterization of the micropatterned sample with different structures support this model. Perfluoroalkylsilane and perfluoroalkylphosphonate SAMs exhibited lower surface energy, higher contact angle, and lower adhesive force as compared to that of alkylsilane and alkylphosphonate SAMs, respectively. The substrate had little effect. The coefficient of friction of various SAMs were comparable.

The influence of relative humidity on adhesion and friction of SAMs was dominated by the thickness of the adsorbed water layer. At higher humidity, water increases friction through increased adhesion by meniscus effect in the contact zone. With an increase in temperature, in the case of Si(111), the desorption of the adsorbed water layer and reduction of the surface tension of water reduced the adhesive force and friction force. A decrease in adhesion and friction with temperature was found for all films. An increase in adhesive force and friction with sliding velocity was found for all films.

PFTS/Si showed a better wear resistance than ODMS/Si. ODDMS/Si showed a better wear resistance than ODMS/Si due to the chain length effect. Wear behavior of the SAMs was mostly determined by the molecule-substrate bond strengths. Similar trends were observed for films on Al substrates.

The nanotribological characterization studies of SAMs deposited on Si and Al substrates are summarized in Table 13.10 (Tambe and Bhushan 2005a). SAMs deposited on the Si and Al substrates showed low friction and low adhesion, both of which are desirable for MEMS/NEMS applications.

Based on studies in high vacuum (2×10^{-7} Torr), the coefficients of friction of the SAMs are in the following order (from low to high): ODDMS/Si, PFTS/Si, HDT/Au, ODMS/Si. HDT on Au showed lower durability than perfluoroalkylsilane and alkylsilane on Si because of the weak interfacial bond. PFTS/Si had better durability than ODMS/Si. This indicates that

Table 13.10 Summary of nanotribological characterization studies for SAMs on Si and Al substrates

SAMs property			Friction force	Adhesive force	Wear
Substrate	{	Hard	High	Low	Low
		Soft	Low	Low	High
Chemical structure	{	Linear chain molecule	High	Low	High
		Ring molecule	High	High	Low
Backbone	{	Fluorocarbon backbone	Low	Low	Low
		Hydrocarbon backbone	Low	High	High
Chain length	{	Long backbone chain		High	High
		Short backbone chain		Low	Low

fluorination of alkylsilane can improve durability. ODDMS/Si was more durable than ODMS/Si and PFTS/Si because of the chain length effect. The friction of SAMs in high vacuum was higher than in argon because of intimate contact. Based on studies in argon and air with various relative humidities, oxygen can increase friction and decrease durability of SAMs. The water molecule can detach SAMs molecules from the substrate and result in high friction and low durability.

In summary, based on the contact angle, adhesion, friction, and wear measurements of SAM films by AFM, they exhibit attractive hydrophobic and tribological properties. Fluorinated SAMs appear to exhibit superior performance. SAM films should find many applications including in micro- and nanodevices requiring good nanotribological properties and surface protection.

References

- Adamson AW (1990) Physical chemistry of Surfaces, 5th edn. Wiley, New York
- Adao MHVC, Saramago BJV, Fernandes AC (1999) Estimation of the surface properties of styrene-acrylonitrile random copolymers from contact angle measurements. *J Colloid Interface Sci* 217:94–106
- Anonymous (1988) Properties of silicon, EMIS Data Reviews Series No. 4, INSPEC. Institution of Electrical Engineers, London. See also Anonymous, 2002, MEMS Materials Database. <http://www.memsnet.org/material/>
- Barrena E, Kopta S, Ogletree DF, Charych DH, Salmeron M (1999) Relationship between friction and molecular structure: alkylsilane lubricant films under pressure. *Phys Rev Lett* 82:2880–2883
- Bhushan B (1996) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer-Verlag, New York
- Bhushan B (ed) (1998a) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht, Netherlands
- Bhushan B (1998b) Contact mechanics of rough surfaces in tribology: multiple asperity contact. *Tribol Lett* 4:1–35
- Bhushan B (ed) (1999a) Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton, Florida
- Bhushan B (1999b) Chemical, mechanical and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments. *Diam Relat Mater* 8:1985–2015
- Bhushan B (ed) (2001) Modern tribology handbook, vol. 1—principles of tribology; vol. 2—materials, coatings, and industrial applications. CRC Press, Boca Raton, Florida
- Bhushan B (2003) Adhesion and friction: mechanisms, measurement techniques, and methods for reduction. *J Vac Sci Technol, B* 21:2262–2296
- Bhushan B (2007) Nanotribology and nanomechanics of MEMS/NEMS and BioMEMS/BioNEMS materials and devices. *Microelectron Eng* 84:387–412
- Bhushan B (2008) Nanotribology and nanomechanics in nano/biotechnology. *Phil Trans R Soc A* 366:1499–1537
- Bhushan B (2011) Nanotribology and nanomechanics I and II, 3rd edn. Springer-Verlag, Heidelberg, Germany
- Bhushan B (2013a) Principles and Applications of Tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2016) Encyclopedia of nanotechnology, 2nd edn. Springer International, Switzerland
- Bhushan B, Cichomski M (2007) Nanotribological characterization of vapor phase deposited fluorosilane self-assembled monolayers deposited on poly-dimethylsiloxane surfaces for biomedical micro-nanodevices. *J Vac Sci Technol, A* 25:1285–1293
- Bhushan B, Jung YC (2008) Wetting, adhesion, and friction of superhydrophobic and hydrophilic leaves and fabricated micro/nanopatterned surfaces. *J Phys: Condens Matter* 20:225010–1 to -24
- Bhushan B, Liu H (2001) Nanotribological properties and mechanisms of alkylthiol and biphenyl thiol self-assembled monolayers studied by atomic force microscopy. *Phys. Rev. B* 63:245412–1 to -11
- Bhushan B, Peng W (2002) Contact mechanics of multilayered rough surfaces. *Appl Mech Rev* 55:435–480
- Bhushan B, Ruan J (1994) Tribological performance of thin film amorphous carbon overcoats for magnetic recording disks in various environments. *Surf Coat Technol* 68(69):644–650
- Bhushan B, Zhao Z (1999) Macro- and microscale tribological studies of molecularly-thick boundary layers of perfluoropolyether lubricants for magnetic thin-film rigid disks. *J Info Storage Proc Syst* 1:1–21
- Bhushan B, Israelachvili JN, Landman U (1995a) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Kulkarni AV, Koinkar VN, Boehm M, Odoni L, Martelet C, Belin M (1995b) Microtribological characterization of self-assembled and langmuir-blodgett monolayers by atomic and friction force microscopy. *Langmuir* 11:3189–3198
- Bhushan B, Yang L, Gao C, Suri S, Miller RA, Marchon B (1995c) Friction and wear studies of magnetic thin film rigid disks with glass-ceramic, glass and aluminum-magnesium substrates. *Wear* 190:44–59
- Bhushan B, Liu H, Hsu SM (2004) Adhesion and friction Studies of silicon and hydrophobic and low Friction films and investigation of scale effects. *ASME J Tribol* 126:583–590
- Bhushan B, Kasai T, Kulik G, Barbieri L, Hoffmann P (2005) AFM study of perfluorosilane and alkylsilane

- self-assembled monolayers for anti-stiction in MEMS/NEMS. *Ultramicroscopy* 105:176–188
- Bhushan B, Cichomski M, Hoque E, DeRose JA, Hoffmann P, Mathieu HJ (2006a) Nanotribological characterization of perfluoroalkylphosphonate self-assembled monolayers deposited on aluminum-coated silicon substrates. *Microsyst Technol* 12:588–596
- Bhushan B, Hansford D, Lee KK (2006b) Surface modification of silicon surfaces with vapor phase deposited ultrathin fluorosilane films for biomedical devices. *J Vac Sci Technol, A* 24:1197–1202
- Bowden FP, Tabor D (1950) The friction and lubrication of solids, Part I. Clarendon Press, Oxford, UK
- Butter RS, Waterman DR, Lettington AH, Ramos RT, Fordham EJ (1997) Production and wetting properties of fluorinated diamond-like carbon coatings. *Thin Solid Films* 311:107–113
- Cai S, Bhushan B (2008) Meniscus and viscous forces during separation of hydrophilic and hydrophobic surfaces with liquid mediated contacts. *Mater Sci Eng R* 61:78–106
- Callister WD (1997) Materials science and engineering, 4th edn. Wiley, New York
- Chambers RD (1973) Fluorine in Organic Chemistry. Wiley, New York
- Cheng J, Krica LJ (eds) (2001) Biochip technology. Harwood Academic Publishers, New York
- Chou SY, Krauss PR, Renstrom PJ (1996) Imprint lithography with 25-nanometer resolution. *Science* 272:85–87
- Clark ES (1999) The molecular conformations of polytetrafluoroethylene: forms II and IV. *Polymer* 40:4659–4665
- Clear SC, Nealey PF (2001) The effect of chain density on the frictional behavior of surfaces modified with alkylsilanes and immersed in n-alcohols. *J Chem Phys* 114:2802–2811
- Cottrell TL (1958) The strength of chemical bonds, 2nd edn. Butterworths Scientific Publishers, London
- DePalma V, Tillman N (1989) Friction and wear of self-assembled tricholosilane monolayer films on silicon. *Langmuir* 5:868–872
- DeRose JA, Hoque E, Bhushan B, Mathieu HJ (2008) Characterization of perfluorodecanote self-assembled monolayers on aluminum and comparison of stability with phosphonate and siloxy self-assembled monolayers. *Surf Sci* 602:1360–1367
- DeVecchio D, Bhushan B (1997) Localized surface elasticity measurements using an atomic force microscope. *Rev Sci Instrum* 68:4498–4505
- Donnet C, Fontaine J, Grill A, Patel V, Jahnes C, Belin M (1997) Wear-resistant fluorinated diamond like carbon films. *Surf Coat Technol* 94–95:531–536
- Dorfman VF (1992) Diamond-like nanocomposites (DLN). *Thin Solid Films* 212:267–273
- Douglass MR (1998) Lifetime estimates and unique failure mechanisms of the digital micromirror device (DMD). In 1998 International reliability physics proceedings, IEEE catalog No. 98 CH 36173. Presented at the 36th annual international reliability physics symposium, Reno, pp 9–16
- Dubois LH, Zegarski BR, Nuzzo RG (1990) Fundamental studies of microscopic wetting on organic surfaces. 2. Interaction of secondary adsorbates with chemically textured organic monolayers. *J Am Chem Soc* 112:570–579
- Erdemir A, Donnet C (2001) Tribology of diamond, diamond-like carbon, and related films. In: Bhushan B (ed) Modern tribology handbook, vol. 2: materials, coatings, and industrial applications. CRC Press, Boca Raton, FL, pp 871–908
- Eustathopoulos N, Nicholas M, Drevet B (1999) Wettability at high temperature. Pergamon, Amsterdam
- Fujihira M, Tani Y, Furugori M, Akiba U, Okabe Y (2001) Chemical force microscopy of self-assembled monolayers on sputtered gold films patterned by phase separation. *Ultramicroscopy* 86:63–73
- Garcia-Parajo M, Longo C, Servat J, Gorostiza P, Sanz F (1997) Nanotribological properties of octadecyltrichlorosilane self-assembled ultrathin films studied by atomic force microscopy: contact and tapping modes. *Langmuir* 13:2333–2339
- Geyer W, Stadler V, Eck W, Zharnikov M, Golzhauser A, Grunze M (1999) Electron-induced crosslinking of aromatic self-assembled monolayers: negative resists for nanolithography. *Appl Phys Lett* 75:2401–2403
- Good RJ, Oss CJV (1992) Modern approaches to wettability-theory and applications. Plenum Press, New York
- Grischke M, Bewilogua K, Trojan K, Dimigan H (1995) Application-oriented modification of deposition process for diamond-like-carbon based coatings. *Surf Coat Technol* 74–75:739–745
- Grischke M, Hieke A, Morgenweck F, Dimigan H (1998) Variation of the wettability of DLC coatings by network modification using silicon and oxygen. *Diam Relat Mater* 7:454–458
- Hansma H, Motamedi F, Smith P, Hansma P, Wittman JC (1992) Molecular resolution of thin, highly oriented poly(tetrafluoroethylene) films with the atomic force microscope. *Polymer Comm* 33:647–649
- Hein M, Best LR, Pattison S, Arena S (1997) Introduction to general, organic, and biochemistry, 6th edn. Brooks/Cole Publishing Co., Pacific Grove, CA
- Heller MJ, Guttman A (eds) (2001) Integrated microfabricated biodevices. Marcel Dekker, New York, NY
- Hoque E, DeRose JA, Hoffmann P, Mathieu HJ, Bhushan B, Cichomski M (2006a) Phosphonate self-assembled monolayers on aluminum surfaces. *J Chem Phys* 124:174710
- Hoque E, DeRose JA, Kulik G, Hoffmann P, Mathieu HJ, Bhushan B (2006b) Alkylphosphonate modified aluminum oxide surfaces. *J Phys Chem B* 110:10855–10861
- Hoque E, DeRose JA, Hoffmann P, Bhushan B, Mathieu HJ (2007a) Alkylperfluorosilane self-assembled monolayers on aluminum: a comparison

- with alkylphosphonate self-assembled monolayers. *J Phys Chem C* 111:3956–3962
- Hoque E, DeRose JA, Hoffmann P, Bhushan B, Mathieu HJ (2007b) Chemical stability of nonwetting, low adhesion self-assembled monolayer films formed by perfluoroalkylsilazation of copper. *J Chem Phys* 126:114706
- Hoque E, DeRose JA, Bhushan B, Mathieu HJ (2008) Self-assembled monolayers on aluminum and copper oxide surfaces: surface and interface characteristics, nanotribological properties, and chemical stability. In: Bhushan B, Fuchs H, Tomitori M (eds) *Applied scanning probe methods Vol. IX—characterization*. Springer-Verlag, Heidelberg, Germany, pp 235–281
- Hoque E, DeRose JA, Bhushan B, Hipps KW (2009) Low adhesion, non-wetting phosphonate self-assembled monolayer films formed on copper oxide surfaces. *Ultramicroscopy* 109:1015–1022
- Hornbeck LJ (2001) The DMD™ projection display chip: a MEMS-based technology. *MRS Bull* 26:325–328
- Hoshino T (1999) Adsorption of atomic and molecular oxygen and desorption of silicon monoxide on Si(111) surfaces. *Phys Rev B* 59:2332–2340
- Israelachvili JN (1992) *Intermolecular and surface forces*, 2nd edn. Academic Press, London
- Jung C, Dannenberger O, Xu Y, Buck M, Grunze M (1998) Self-assembled monolayers from organosulfur compounds: a comparison between sulfides, disulfides, and thiols. *Langmuir* 14:1103–1107
- Kasai T, Bhushan B, Kulik G, Barbieri L, Hoffmann P (2005) Nanotribological study of perfluorosilane SAMs for anti-stiction and low wear. *J Vac Sci Technol, B* 23:995–1003
- Kester DJ, Brodbeck CL, Singer IL, Kyriakopoulos A (1999) Sliding wear behavior of diamond-like nanocomposite coatings. *Surf Coat Technol* 113: 268–273
- Kim HK, Lee JP, Park CR, Kwak HT, Sung MM (2003) Thermal decomposition of alkylsiloxane self-assembled monolayers in Air. *J Phys Chem B* 107:4351
- Kluth GJ, Sander M, Sung MM, Maboudian R (1998) Study of the desorption mechanism of alkylsiloxane self-assembled monolayers through isotopic labeling and high resolution electron energy-loss spectroscopy experiments. *J Vac Sci Technol, A* 16:932–936
- Koinkar VN, Bhushan B (1996) Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J Vac Sci Technol, A* 14:2378–2391
- Kumar A, Whitesides GM (1993) Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol ink followed by chemical etching. *Appl Phys Lett* 63:2002–2004
- Lee KK, Bhushan B, Hansford D (2005) Nanotribological characterization of perfluoropolymer thin films for BioMEMS Applications. *J Vac Sci Technol, A* 23:804–810
- Lide DR (2004) CRC handbook of chemistry and physics, 85th edn. CRC Press, Boca Raton, Florida
- Lio A, Charych DH, Salmeron M (1997) Comparative atomic force microscopy study of the chain length dependence of frictional properties of alkanethiol on gold and alkylsilanes on mica. *J Phys Chem B* 101:3800–3805
- Liu H, Bhushan B (2002a) Investigation of nanotribological properties of alkylthiol and biphenyl thiol self-assembled monolayers. *Ultramicroscopy* 91:185–202
- Liu H, Bhushan B (2002b) Orientation and RELOCATION of biphenyl thiol self-assembled monolayers. *Ultramicroscopy* 91:177–183
- Liu H, Bhushan B (2003a) Nanotribological characterization of molecularly-thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* 97:321–340
- Liu H, Bhushan B (2003b) Adhesion and friction studies of microelectromechanical systems/nanoelectromechanical systems materials using a novel microtriboapparatus. *J Vac Sci Technol, A* 21:1528–1538
- Liu H, Bhushan B (2004a) Nanotribological characterization of digital micromirror devices using an atomic force microscope. *Ultramicroscopy* 100:391–412
- Liu H, Bhushan B (2004b) Investigation of nanotribological and nanomechanical properties of the digital micromirror device by atomic force microscope. *J Vac Sci Technol, A* 22:1388–1396
- Liu H, Bhushan B, Eck W, Stadler V (2001) Investigation of the adhesion, friction, and wear properties of biphenyl thiol self-assembled monolayers by atomic force microscopy. *J Vac Sci Technol, A* 19:1234–1240
- Man KF, Stark BH, Ramesham R (1998) A resource handbook for MEMS reliability, Rev. A, (JPL Press, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA). See also, Man KF (2002) MEMS reliability for space applications by elimination of potential failure modes through testing and analysis. <http://www-rel.jpl.nasa.gov/Org/5053/atpo/products/Prod-map.html>
- Manz A, Becker H (eds) (1998) *Microsystem technology in chemistry and life sciences*. Springer-Verlag, Heidelberg, Germany
- McDermott MT, Green JBD, Porter MD (1997) Scanning force microscopic exploration of the lubrication capabilities of n-alkanethiolate monolayers chemisorbed at gold: structural basis of microscopic friction and wear. *Langmuir* 13:2504–2510
- McMaster MC, Schroeder SLM, Madix RJ (1993) Molecular propane adsorption dynamics on Pt(110)–(1 × 2). *Surf Sci* 297:253–271
- Miura YF, Takenga M, Koini T, Graupe M, Garg N, Graham RL, Lee TR (1998) Wettability of self-assembled monolayers generated from CF₃-terminated alkanethiols on gold. *Langmuir* 14:5821–5825
- Mohrig JR, Hammond CN, Morrill TC, Neckers DC (1998) *Experimental organic chemistry*. W.H. Freeman, New York

- Nosonovsky M, Bhushan B (2008) Roughness-induced superhydrophobicity: a way to design non-adhesive surfaces. *J Phys: Condens Matter* 20:225009-1 to -30
- Ratajczak-Sitarz M, Katrusiak A, Kaluski Z, Garbaczuk J (1987) 4,4'-biphenyldithiol, *Acta Crystallogr C* 43:2389–2391
- Ren S, Yang S, Zhao Y, Yu T, Xiao X (2003) Preparation and characterization of ultrahydrophobic surface based on a stearic acid self-assembled monolayer over polyethylenimine thin films. *Surf Sci* 546:64–74
- Ruhe J, Novotny VJ, Kanazawa KK, Clarke T, Street GB (1993) Structure and tribological properties of ultra-thin alkylsilane films chemisorbed to solid surfaces. *Langmuir* 9:2383–2388
- Scandella L, Schumacher A, Kruse N, Prins R, Meyer E, Luethi R, Howald L, Scherge M, Schaefer JA (1998) Surface modification and mechanical properties of bulk silicon. In: Bhushan B (ed) *Tribology issues and opportunities in MEMS*. Kluwer Academic, Dordrecht, Netherlands, pp 529–537
- Schonherr H, Vancso GJ (1999) Tribological properties of self-assembled monolayers of fluorocarbon and hydrocarbon thiols and disulfides on Au(111) studied by scanning force microscopy. *Mater Sci Eng C* 8–9:243–249
- Schrader ME, Loeb GI (eds) (1992) *Modern approaches to wettability*. Plenum, New York
- Sexton BA, Hughes AE (1984) A comparison of weak molecular adsorption of organic-molecules on clean copper and platinum surfaces. *Surf Sci* 140: 227–248
- Siepmann JI, McDonald IR (1993) Monte Carlo simulation of the mechanical relaxation of a self-assembled monolayer. *Phys Rev Lett* 70:453–456
- Tambe NS, Bhushan B (2005a) Nanotribological Characterization of Self Assembled Monolayers Deposited on Silicon and Aluminum Substrates. *Nanotechnology* 16:1549–1558
- Tambe NS, Bhushan B (2005b) A New Atomic Force Microscopy Based Technique for Studying Nanoscale Friction at High Sliding Velocities. *J Phys D Appl Phys* 38:764–773
- Tambe NS, Bhushan B (2005c) Friction Model for Velocity Dependence of Nanoscale Friction. *Nanotechnology* 16:2309–2324
- Tambe NS, Bhushan B (2005d) Durability studies of micro/nanoelectromechanical systems materials, coatings and lubricants at high sliding velocities (up to 10 mm/s) using a modified atomic force microscope. *J Vac Sci Technol A* 23:830–835
- Tanner DM, Smith NF, Irwin LW et al (2000) “MEMS reliability: infrastructure, test structure, experiments, and failure modes”, SAND2000-0091. Sandia National Laboratories, Albuquerque, New Mexico
- Tao Z, Bhushan B (2005) Degradation Mechanisms and Environmental Effects on Perfluoropolyether Self Assembled Monolayers and Diamondlike Carbon Films. *Langmuir* 21:2391–2399
- Tao Z, Bhushan B (2006) Surface modification of AFM silicon probes for adhesion and wear reduction. *Tribol Lett* 21:1–16
- Tian F, Xiao X, Loy MMT, Wang C, Bai C (1999) Humidity and temperature effect on frictional properties of mica and alkylsilane monolayer self-assembled on mica. *Langmuir* 15:244–249
- Tian J, Xia Y, Whitesides GM (1998) Microcontact printing of SAMs. In: Ulman A (ed) *Thin films—self-assembled monolayers of thiols*, vol 24. Academic Press, San Diego, CA, pp 227–254
- Tokachichu DR, Bhushan B (2006) Bioadhesion of polymers for BioMEMS. *IEEE Trans Nanotech* 5:228–231
- Tsukruk VV, Bliznyuk VN (1998) Adhesive and friction forces between chemically modified silicon and silicon nitride surfaces. *Langmuir* 14:446–455
- Tsukruk VV, Nguyen T, Lemieux M, Hazel J, Weber WH, Shevchenko VV, Klimenko N, Sheldukov E (1998) Tribological properties of modified MEMS surfaces. In: Bhushan B (ed) *Tribology issues and opportunities in MEMS*. Kluwer Academic, Dordrecht, Netherlands, pp 607–614
- Ulman A (1991) An introduction to ultrathin organic films: from Langmuir-Blodgett to self-assembly. Academic Press, San Diego, CA
- Ulman A (ed) (1995) *Characterization of organic thin films*. Butterworth-Heinemann, Boston, Mass
- Ulman A (1996) Formation and structure of self-assembled monolayers. *Chem Rev* 96:1533–1554
- van der Berg A (ed) (2003) *Lab-on-a-chip: chemistry in miniaturized synthesis and analysis systems*. Elsevier, Amsterdam, Netherlands
- Wasserman SR, Tao YT, Whitesides GM (1989) Structure and reactivity of alkylsiloxane monolayers formed by reaction of alkylchlorosilanes on silicon substrates. *Langmuir* 5:1074–1089
- Xia Y, Whitesides GM (1998) Soft lithography. *Angew Chem Int Ed* 37:550–575
- Xia Y, Kim E, Zhao XM, Rogers JA, Prentiss M, Whitesides GM (1996) Complex optical surfaces formed by replica molding against elastomeric masters. *Science* 273:347–349
- Xiao X, Hu J, Charych DH, Salmeron M (1996) Chain length dependence of the frictional properties of alkylsilane molecules self-assembled on mica studied by atomic force microscopy. *Langmuir* 12:235–237
- Zasadzinski JA, Viswanathan R, Madsen L, Garnaes J, Schwartz DK (1994) Langmuir-Blodgett films. *Science* 263:1726–1733
- Zisman WA (1959) Friction, durability and wettability properties of monomolecular films on solids. In: Davies R (ed) *Friction and wear*. Elsevier, Amsterdam, pp 110–148

Bharat Bhushan

Abstract

Boundary films are formed by physisorption, chemisorption, and chemical reaction. A good boundary lubricant should have a high degree of interaction between its molecules and the solid surface. As a general rule, liquids are good lubricants when they are polar and, thus, able to grip solid surfaces (or be adsorbed). In this chapter, we present an overview of various perfluoropolyethers (PFPEs) and ionic liquid films. PFPEs exhibit the highest chemical and thermal stability. Ionic liquid films possess efficient heat transfer properties and some electrical conductivity of interest. We present a summary of nanodeformation, molecular conformation, and lubricant spreading studies, followed by an overview of the nanotribological properties of polar and nonpolar PFPEs and ionic liquid films studied by atomic force microscopy (AFM), and chemical degradation studies using a high vacuum tribotest apparatus.

Keywords

Boundary lubrication • Molecularly-thick layers • Chemically bonded films • Friction • Wear • Lubrication • Nanotribology

14.1 Introduction

Boundary films are formed by physisorption, chemisorption, and chemical reaction. With physisorption, no exchange of electrons takes place between the molecules of the adsorbate and those of the adsorbent (Bhushan 2013a, b). The physisorption process typically involves van der Waals forces, which are relatively weak. In chemisorption, there is an actual sharing of

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: Bhushan.2@osu.edu

electrons or an electron interchange between the chemisorbed species and the solid surface. The solid surfaces very strongly bond to the adsorption species through covalent bonds. Chemically reacted films are formed by the chemical reaction of the solid surface with the environment. Physisorbed films can be either monomolecularly or polymolecularly thick. Chemisorbed films are monomolecular, but stoichiometric films formed by chemical reaction can have a large film thickness. In general, the stability and durability of surface films decrease in the following order: chemically reacted films, chemisorbed films, and physisorbed films. A good boundary lubricant should have a high degree of interaction between its molecules and the solid surface. As a general rule, liquids are good lubricants when they are polar and, thus, able to grip solid surfaces (or be adsorbed). Polar lubricants contain reactive functional groups with low ionization potential or groups having high polarizability (Bhushan 1993, 1996, 2001, 2013a, b). Boundary lubrication properties of lubricants are also dependent upon the molecular conformation and lubricant spreading (Novotny et al. 1989; Novotny 1990; Mate and Novotny 1991; Mate 1992a).

Perfluoropolyether (PFPE) films, self-assembled monolayers (SAMs), and Langmuir-Blodgett (LB) films can be used as boundary lubricants (Roberts 1990; Ulman 1991; Bhushan 1993, 1996, 1999, 2001, 2013a, b; Bhushan and Zhao 1999; Hoque et al. 2008). PFPEs exhibit the following properties: low surface tension and a low contact angle, which allow easy spreading on surfaces and provide hydrophobic property; chemical and thermal stability which minimize degradation under use; low vapor pressure, which provides low out-gassing; high adhesion to substrate via organofunctional bonds; and good lubricity, which reduces the interfacial friction and wear (Bhushan 1996, 2001). PFPE films are used exclusively for lubrication of magnetic rigid disks and metal evaporated magnetic tapes to reduce friction and wear of the head-medium interface

(Bhushan 1996). The structure of the lubricants employed at the head-medium interface has not changed substantially over the past decade, except with the variations in the end groups. However, the thickness of the PFPE film used to lubricate the disk has steadily decreased from multilayer thicknesses to the sub-monolayer thickness regime (Bhushan and Zhao 1999; Bhushan 2001). Molecularly-thick PFPE films are also being considered for lubrication purposes for micro/nanoelectromechanical systems (MEMS/NEMS) (Bhushan 1998, 2001, 2008a, 2016). It is well known that the properties of molecularly-thick liquid films confined to solid surfaces can be dramatically different from those of the corresponding bulk liquid. In order to efficiently develop lubrication systems that meet the requirements of advanced rigid disk drives and MEMS/NEMS applications, the nanotribology of thin films of PFPE lubricants should be fully understood (Bhushan et al. 1995a; Bhushan 1999, 2008b). It is also important to understand lubricant-substrate interfacial interactions and the influence of operating environment on the nanotribological performance of molecularly-thick PFPE liquid films.

Ionic liquids have been considered as potential lubricants (Bhushan et al. 2008a; Palacio and Bhushan 2008a, 2009a). They possess efficient heat transfer properties. They are also electrically conducting which is of interest in various MEMS/NEMS applications such as atomic force microscope probe-based data recording (Bhushan and Kwak 2008; Bhushan et al. 2008b; Palacio and Bhushan 2008b).

In this chapter, we present an overview of PFPE and ionic liquid films. We first present a summary of nanodeformation, molecular conformation, and lubricant spreading studies and an overview of nanotribological and electrical studies of commonly used polar and nonpolar PFPE and ionic liquid films using atomic force microscopy (AFM) and chemical degradation studies using a high vacuum tribotest apparatus.

14.2 Nanodeformation, Molecular Conformation, Spreading, Nanotribological and Electrical Studies, and Environmental Effects of Commonly Used PFPE Lubricant Films

The molecular structures of two commonly used typical PFPE lubricants (Z-15 and Z-DOL 2000) are schematically shown in Fig. 14.1 (Fomblin, Solvay Solexis Inc., Milan, Italy). Z-15 has nonpolar $-CF_3$ end groups, whereas Z-DOL is a polar lubricant with hydroxyl ($-OH$) end groups. Their typical properties are summarized in Table 14.1. It shows that Z-15 and Z-DOL have almost the same density and surface tension, but Z-15 has a larger molecular weight and higher viscosity. Both of them have low surface tension, low vapor pressure, low evaporation weight loss, and good oxidative stability (Bhushan 1993, 1996). For nanotribological characterization, a single crystal Si(100) wafer with native oxide layer has generally been used as a substrate for deposition of molecularly-thick lubricant films. Z-15 and Z-DOL films can be directly deposited on the Si(100) wafer by a dip coating technique. The clean silicon wafer is vertically submerged into a dilute solution of lubricant in a

hydrocarbon solvent (HT- 70, Solvay Solexis Inc., Milan, Italy). The silicon wafers are vertically pulled up from the solution with a motorized stage at a constant speed for deposition of desired thicknesses of Z-15 and Z-DOL lubricants (Bhushan and Zhao 1999; Bhushan 2001; Liu and Bhushan 2003). The lubricant film thickness obtained in dip coating is a function of concentration and pulling-up speed, among other factors. The Z-DOL film is bonded to the silicon substrate by heating the as-deposited Z-DOL samples in an oven at 150 °C for about 30 min. The native oxide layer of Si(100) wafer reacts with the $-OH$ groups of the lubricants during thermal treatment (Ruhe et al. 1994, 1996; Koinkar and Bhushan 1996a; Liu and Bhushan 2003). Subsequently, fluorocarbon solvent (FC-72, 3 M) washing of the thermal treated specimen removes loosely absorbed species leaving the chemically bonded phase on the substrate. The chemical bonding between Z-DOL molecules and the silicon substrate is illustrated in Fig. 14.2. The bonded and washed Z-DOL film is referred as Z-DOL (fully bonded) in this chapter. The as-deposited Z-15 and Z-DOL films are mobile phase lubricants (i.e., liquid-like lubricant), whereas the Z-DOL (fully bonded) films are fully bonded soft solid phase (i.e., solid-like) lubricants.

Fig. 14.1 Schematics of the molecular structures of Z-15 and Z-DOL 2000. Z-15 has nonpolar $-CF_3$ end groups, whereas Z-DOL is a polar lubricant with hydroxyl ($-OH$) end groups. In this figure, the m/n value, shown in Table 14.1, equals 2/3

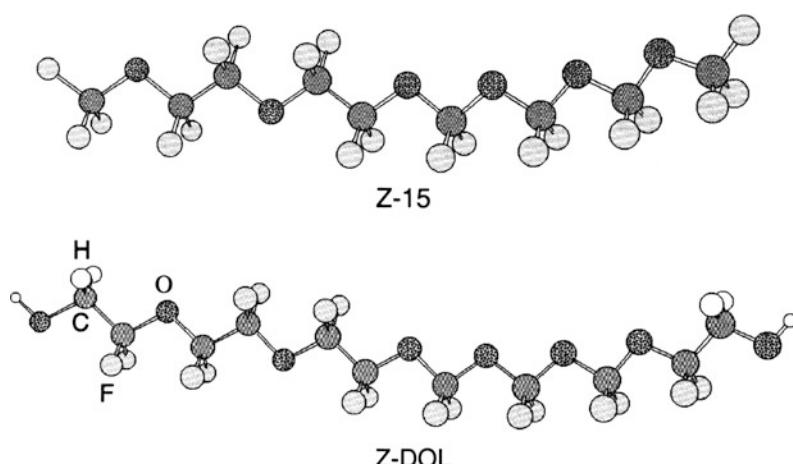


Table 14.1 Typical properties of Z-15 and Z-DOL (Fomblin Z, Solvay Solexis Inc., Milan, Italy)

	Z-15	Z-DOL (2000)
Formula	$\text{CF}_3\text{O}-(\text{CF}_2-\text{CF}_2\text{O})_m-(\text{CF}_2\text{O})_n-\text{CF}_3^a$	$\text{HO}-\text{CH}_2-\text{CF}_2\text{O}-(\text{CF}_2-\text{CF}_2\text{O})_m-(\text{CF}_2\text{O})_n-\text{CF}_2-\text{CH}_2-\text{OH}^a$
Molecular weight (Daltons)	9100	2000
Density (ASTM D891) 20 °C (g/cm ³)	1.84	1.81
Kinematic viscosity (ASTM D445) (cSt)	148 90 25	85 34 —
20 °C		
38 °C		
99 °C		
Viscosity index (ASTM D2270)	320	—
Surface tension (ASTM D1331) (dyn/cm) 20 °C	24	24
Vapor pressure (Torr) 20 °C	1.6×10^{-6}	2×10^{-5}
100 °C	1.7×10^{-5}	6×10^{-4}
Pour point (ASTM D972) °C	-80	—
Evaporation weight loss (ASTM D972) 149 °C, 22 h (%)	0.7	—
Oxidative stability (°C)		320
Specific heat (cal/g °C) 38 °C	0.21	—

^am/n ~ 2/3

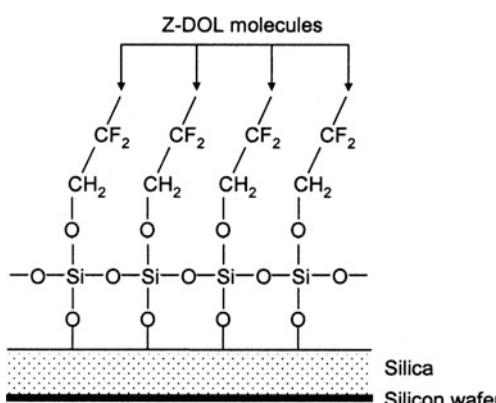


Fig. 14.2 Schematic of Z-DOL molecules that chemically bonded on Si(100) substrate surface (which has native oxide) after thermal treatment at 150 °C for 30 min

14.2.1 Nanodeformation, Molecular Conformation, and Spreading

Nanodeformation behavior of Z-DOL lubricants was studied using an AFM by Blackman et al. (1990a, b). Before bringing a tungsten tip into

contact with a molecular overlayer, it was first brought into contact with a bare clean-silicon surface, Fig. 14.3. As the sample approaches the tip, the force initially is zero, but at point A the force suddenly becomes attractive (top curve) which increases until point B where the sample and tip come into intimate contact and the force becomes repulsive. As the sample is retracted, a pull-off force of 5×10^{-8} N (point D) is required to overcome adhesion between the tungsten tip and the silicon surface. When an AFM tip is brought into contact with an unbonded Z-DOL film, a sudden jump in adhesive contact is also observed. A much larger pull-off force is required to overcome the adhesion. The adhesion is initiated by the formation of a lubricant meniscus surrounding the tip. This suggests that the unbonded Z-DOL lubricant shows liquid-like behavior. However, when the tip was brought into contact with a lubricant film which was firmly bonded to the surface, the liquid-like behavior disappears. The initial attractive force (point A) is no longer sudden as with the liquid film, but rather gradually increases as the tip penetrates the film.

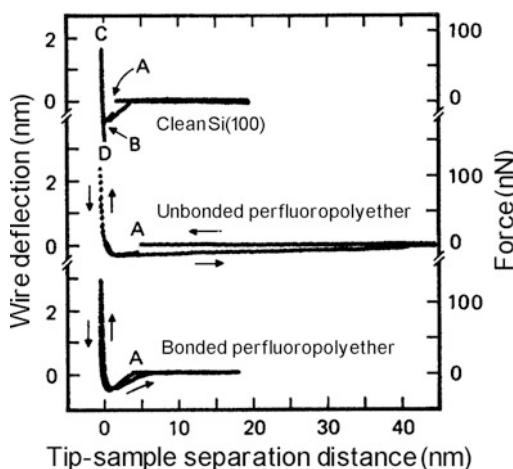


Fig. 14.3 Wire deflection (normal load) as a function of tip-sample separation distance curves comparing the behavior of clean Si (100) surface to a surface lubricated with free and unbonded PFPE lubricant, and a surface where the PFPE lubricant film was thermally bonded to the surface (Blackman et al. 1990a)

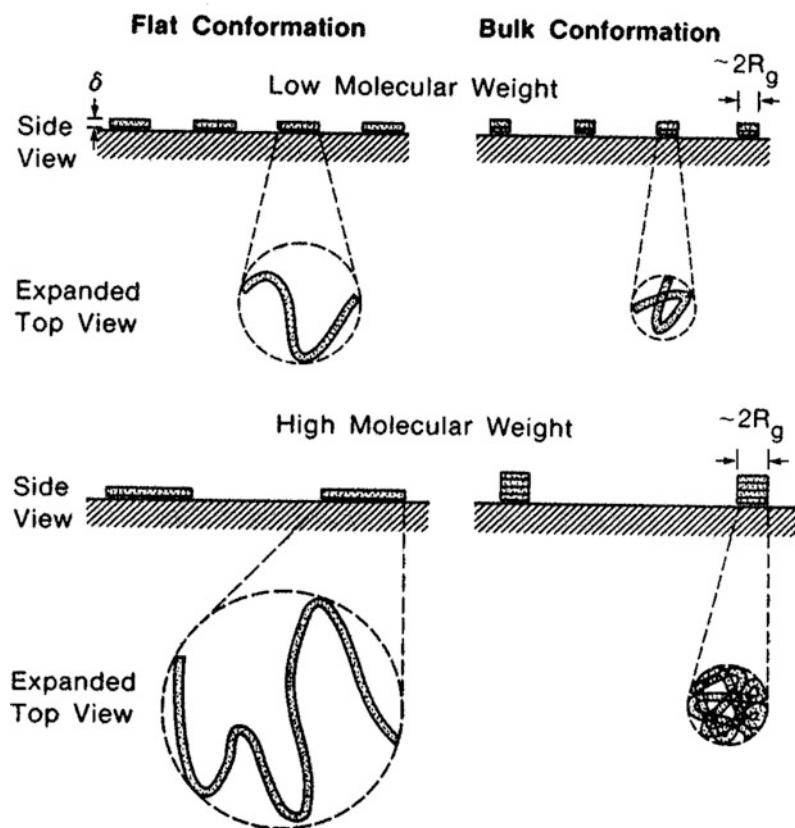
According to Blackman et al. (1990a, b), if the substrate and tip were infinitely hard with no compliance and/or deformation in the tip and sample supports, the line for B–C would be vertical with an infinite slope. The tangent to the force-distance curve at a given point is referred to as the stiffness at that point and was determined by fitting a least-squares line through the nearby data points. For silicon, the deformation is reversible (elastic) since the retracting (outgoing) portion of the curve (C–D) follows the extending (ingoing) portion (B–C). For bonded lubricant film, at the point where slope of the force changes gradually from attractive to repulsive, the stiffness changes gradually, indicating compression of the molecular film. As the load is increased, the slope of the repulsive force eventually approaches that of the bare surface. The bonded film was found to respond elastically up to the highest loads of 5 μ N which could be applied. Thus, bonded lubricant behaves as a soft polymer solid.

Figure 14.4 illustrates two extremes for the conformation on a surface of a linear liquid polymer without any reactive end groups and at submonolayer coverages (Novotny et al. 1989; Mate and Novotny 1991). At one extreme, the molecules lie flat on the surface, reaching no more

than their chain diameter δ above the surface. This would be the case if a strong attractive interaction exists between the molecules and the solid. On the other extreme, when a weak attraction exists between polymer segments and the solid, the molecules adopt conformation close to that of the molecules in the bulk, with the microscopic thickness equal to about the radius of gyration R_g . Mate and Novotny (1991) used AFM to study conformation of 0.5–1.3 nm thick Z-15 molecules on clean Si(100) surfaces. They found that the thickness measured by AFM is thicker than that measured by ellipsometry with the offset ranging from 3 to 5 nm. They found that the offset was the same for very thin submonolayer coverages. If the coverage is submonolayer and inadequate to make a liquid film, the relevant thickness is then the height (h_e) of the molecules extended above the solid surface. The offset should be equal to $2h_e$, assuming that the molecules extend the same height above both the tip and silicon surfaces. They thus concluded that the molecules do not extend more than 1.5–2.5 nm above a solid or liquid surface, much smaller than the radius of gyration of the lubricants ranging between 3.2 and 7.3 nm, and similar to the approximate cross-sectional diameter of 0.6–0.7 nm for the linear polymer chain. Consequently, the height that the molecules extend above the surface is considerably less than the diameter of gyration of the molecules and only a few molecular diameters in height, implying that the physisorbed molecules on a solid surface have an extended, flat conformation. They also determined the disjoining pressure of these liquid films from AFM measurements of the distance needed to break the liquid meniscus that forms between solid surface and the AFM tip. (Also see Mate 1992a.) For a monolayer thickness of about 0.7 nm, the disjoining pressure is about 5 MPa, indicating strong attractive interaction between the liquid molecules and the solid surface. The disjoining pressure decreases with increasing film thickness in a manner consistent with a strong attractive van der Waals interaction between the liquid molecules and the solid surface.

Rheological characterization shows that the flow activation energy of PFPE lubricants is

Fig. 14.4 Schematic representation of two extreme liquid conformations at the surface of the solid for low and high molecular weights at low surface coverage. δ is the cross-sectional diameter of the liquid chain and R_g is the radius of gyration of the molecules in the bulk (Mate and Novotny 1991)



weakly dependent on chain length, and is strongly dependent on the functional end groups (Kim et al. 1999). PFPE lubricant films that contain polar end groups have lower mobility than those with nonpolar end groups of similar chain length (Ruths and Granick 1998). The mobility of PFPE also depends on the surface chemical properties of the substrate. The spreading of Z-DOL on an amorphous carbon surface has been studied as a function of hydrogen or nitrogen content in the carbon film, using scanning microellipsometry (Ma et al. 1999). The diffusion coefficient data presented in Fig. 14.5 is thickness dependent. It shows that the surface mobility of Z-DOL increased as hydrogen content increased, but decreased as nitrogen content increased. The enhancement of Z-DOL surface mobility by hydrogenation may be understood from the fact that the interactions between Z-DOL molecules and the carbon surface can be significantly weakened due to a

reduction of the number of high-energy binding sites on the carbon surface. The stronger interactions between Z-DOL molecules and the carbon surface, as nitrogen content in the carbon coating increases, leads to the lowering of Z-DOL surface mobility.

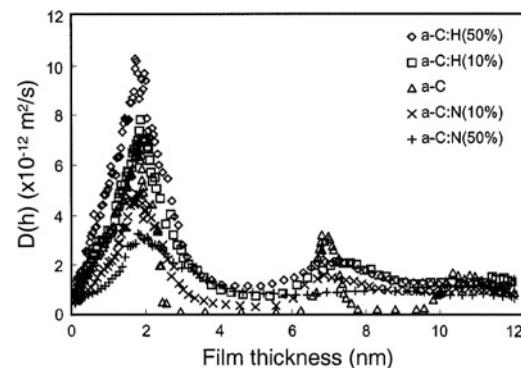


Fig. 14.5 Diffusion coefficient $D(h)$ as a function of lubricant film thickness for Z-DOL on different carbon films (Ma et al. 1999)

Molecularly-thick films may be sheared at very high shear rates, on the order of 10^8 – 10^9 s $^{-1}$ during sliding, such as during magnetic disk drive operation. During such shear, lubricant stability is critical to the protection of the interface. For proper lubricant selection, viscosity at high shear rates and associated shear thinning need to be understood. Viscosity measurements of eight different types of PFPE films show that all eight lubricants display Newtonian behavior and their viscosity remains constant at shear rates up to 10 7 s $^{-1}$ (Jonsson and Bhushan 1995; Hahm and Bhushan 1997).

14.2.2 Nanotribological Studies and Environmental Effects

AFM techniques have been used to investigate the nanotribological performance of PFPEs. Mate (1992b, 1993), O’Shea et al. (1992, 1994), Bhushan et al. (1995a, b), Koinkar and Bhushan (1996a, b), Bhushan and Sundararajan (1998), Bhushan and Dandavate (2000), Liu and Bhushan (2003), and Palacio and Bhushan (2007a, 2009b) used an AFM to provide insight into how PFPE lubricants function at the molecular level. Mate (1992b, 1993) conducted friction experiments on bonded and unbonded Z-DOL, and he found that the coefficient of friction of the unbonded Z-DOL is about two times larger than the bonded Z-DOL (Also see O’Shea et al. 1992, 1994). Koinkar and Bhushan (1996a, b) and Liu and Bhushan (2003) studied the friction and wear performance of a Si(100) sample lubricated with Z-15, Z-DOL, and Z-DOL (fully bonded) lubricants. They found that using Z-DOL (fully bonded) could significantly improve the adhesion, friction, and wear performance of Si(100). They also discussed the lubrication mechanisms on the molecular level. Bhushan and Sundararajan (1998) and Bhushan and Dandavate (2000) studied the effect of tip radius and relative humidity on the adhesion and friction properties of Si(100) coated with Z-DOL (fully bonded).

In this section, we review the adhesion, friction, and wear properties of Z-15 and Z-DOL at various operating conditions (rest time, velocity, relative humidity, temperature, and tip radius). The experiments were carried out using a commercial AFM system with pyramidal Si₃N₄ and diamond tips. An environmentally controlled chamber and a thermal stage were used to perform relative humidity and temperature effect studies.

14.2.2.1 Friction and Adhesion

To investigate the friction properties of Z-15 and Z-DOL (fully bonded) films on Si(100), the friction force versus normal load curves were measured by making friction measurements at increasing normal loads (Liu and Bhushan 2003). The representative results of Si(100), Z-15, and Z-DOL (fully bonded) are shown in Fig. 14.6. An approximately linear response of all three samples is observed in the load range of 5–130 nN. The friction force of solid-like Z-DOL (fully bonded) is consistently smaller than that for Si(100), but the friction force of liquid-like Z-15 lubricant is higher than that of Si(100). Sundararajan and Bhushan (2001) have studied the static friction force of silicon micromotors lubricated with Z-DOL by AFM. They also found that liquid-like lubricant of Z-DOL significantly increased the static friction force; on the contrary a solid-like Z-DOL (fully bonded) coating can dramatically reduce the static friction force. This is in good agreement with the results of Liu and Bhushan (2003). In Fig. 14.6, the nonzero value of the friction force signal at zero external load is due to the adhesive forces. It is well known that the following relationship exists between the friction force F and external normal load W (Bhushan 2013a, b)

$$F = \mu(W + W_a) \quad (14.1)$$

where μ is the coefficient of friction, and W_a is the adhesive force. Based on this equation and the data in Fig. 14.6, we can calculate the μ and W_a values. The coefficients of friction of Si(100), Z-15, and Z-DOL are 0.07, 0.09 and 0.04,

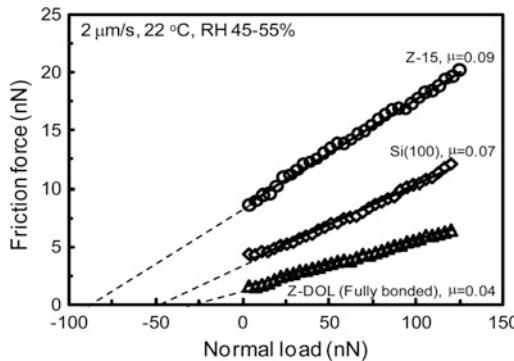


Fig. 14.6 Friction force versus normal load curves for Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film at 2 $\mu\text{m}/\text{s}$, and in ambient air sliding against a Si_3N_4 tip. Based on these curves, coefficient of friction μ and adhesion force of W_a can be calculated (Liu and Bhushan 2003)

respectively. Based on Eq. 14.1, the adhesive force values are obtained from the horizontal intercepts of the friction force versus normal load curves at a zero value of friction force. Adhesive force values of Si(100), Z-15, and Z-DOL are 52, 91, and 34 nN, respectively.

The adhesive forces of these samples were also measured by using a force calibration plot (FCP) technique to obtain force-versus-distance curves. In this technique, the tip is brought into contact with the sample, and the maximum force needed to pull the tip and sample apart is measured as the adhesive force. Figure 14.7 shows the typical FCP curves of Si(100), Z-15, and Z-DOL (fully bonded) (Liu and Bhushan 2003). As the tip approaches the sample within a few nanometers (A), an attractive force exists between the tip and the sample surfaces. The tip is pulled toward the sample, and contact occurs at point B on the graph. The adsorption of water molecules and/or presence of liquid lubricant molecules on the sample surface can also accelerate this so called “snap-in,” due to the formation of meniscus of the water and/or liquid lubricant around the tip. From this point on, the tip is in contact with the surface, and as the piezo extends further, the cantilever is further deflected. This is represented by the slope portion of the curve. As the piezo retracts, at point C the tip goes beyond the zero deflection (flat) line, because of the attractive

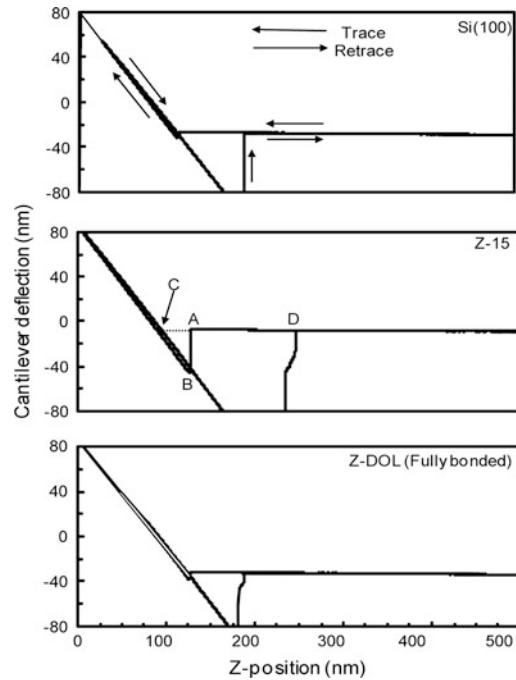
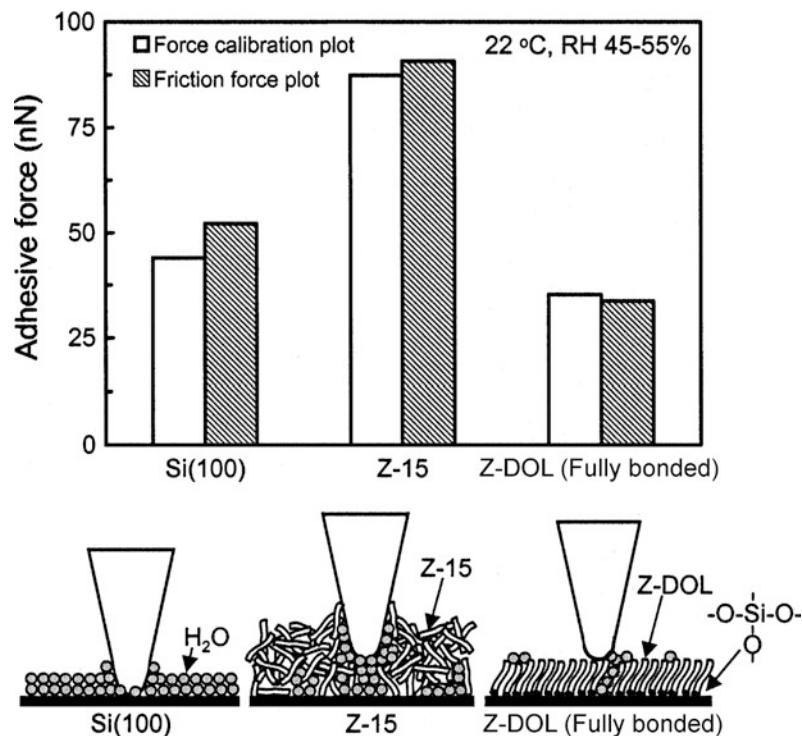


Fig. 14.7 Typical force calibration plots of Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film in ambient air. The adhesive forces can be calculated from the horizontal distance between points C and D, and the cantilever spring constant of 0.58 N/m (Liu and Bhushan 2003)

forces, into the adhesive force regime. At point D, the tip snaps free of the adhesive forces and is again in free air. The adhesive force (pull-off force) is determined by multiplying the cantilever spring constant (0.58 N/m) by the horizontal distance between points C and D, which corresponds to the maximum cantilever deflection toward the samples before the tip is disengaged. Incidentally, the horizontal shift between the loading and unloading curves results from the hysteresis of the PZT tube.

The adhesive forces of Si(100), Z-15, and Z-DOL (fully bonded) measured by FCP and friction force versus normal load plot are summarized in Fig. 14.8 (Liu and Bhushan 2003). The results measured by these two methods are in good agreement. Figure 14.8 shows that the presence of mobile Z-15 lubricant film increases the adhesive force as compared to that of Si(100). Whereas, the presence of solid phase Z-DOL (fully bonded) film reduces the adhesive force as

Fig. 14.8 Summary of the adhesive forces of Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film measured by force calibration plots and friction force versus normal load plots in ambient air. The schematic (bottom) showing the effect of meniscus, formed between the AFM tip and the sample surface, on the adhesive and friction forces (Liu and Bhushan 2003)



compared that of Si(100). This result is in good agreement with the results of Blackman et al. (1990a) and Bhushan and Ruan (1994a). Sources of adhesive forces between the tip and the sample surfaces are van der Waals attraction and long-range meniscus force (Bhushan 1999, 2013a, b). Relative magnitudes of the forces from the two sources are dependent upon various factors including the distance between the tip and the sample surface, their surface roughness, their hydrophobicity, and relative humidity (Stifter et al. 2000). For most surfaces with some roughness, meniscus contribution dominates at moderate to high humidities.

The schematic (bottom) in Fig. 14.8 shows the relative size and source of meniscus. The native oxide layer (SiO_2) on the top of Si(100) wafer exhibits hydrophilic properties, and some water molecules can be adsorbed on this surface. The condensed water will form a meniscus as the tip approaches to the sample surface. In the case of a sphere (such as a single asperity AFM tip) in contact with a flat surface, the attractive Laplace force F_L caused by the capillary is:

$$F_L = 2\pi R\gamma(\cos \theta_1 + \cos \theta_2) \quad (14.2)$$

where R is the radius of the sphere, γ is the surface tension of the liquid against air, and θ_1 and θ_2 are the contact angles between liquid and flat and spherical surfaces, respectively (Israelachvili 1992; Bhushan 2013a, b). As the surface tension value of Z-15 (24 dyn/cm) is smaller than that of water (72 dyn/cm), the larger adhesive force in Z-15 cannot only be caused by the Z-15 meniscus. The non-polar Z-15 liquid does not have complete coverage and strong bonding with Si(100). In the ambient environment, the condensed water molecules will permeate through the liquid Z-15 lubricant film and compete with the lubricant molecules present on the substrate. The interaction of the liquid lubricant with the substrate is weakened, and a boundary layer of the liquid lubricant forms puddles (Koinkar and Bhushan 1996a, b). This dewetting allows water molecules to be adsorbed on the Si(100) surface as aggregates along with Z-15 molecules, and both of them can form meniscus while the tip approaches to the surface. In

addition, as the Z-15 film is pretty soft as compared to the solid Si(100) surface, and penetration of the tip in the film occurs while pushing tip down, this leads to a large area of the tip involved to form the meniscus at the tip-liquid (water aggregates along with Z-15) interface. These two factors of the liquid-like Z-15 film result in higher adhesive force. It should also be noted that Z-15 has a higher viscosity compared to that of water, therefore Z-15 film provides higher resistance to sliding motion and results in a larger coefficient of friction. In the case of Z-DOL (fully bonded) film, both active groups of Z-DOL molecules are strongly bonded on the Si(100) substrate through the thermal and washing treatment, thus the Z-DOL (fully bonded) film has a relatively low free surface energy and cannot be displaced readily by water molecules or readily adsorb water molecules. Thus, the use of Z-DOL (fully bonded) can reduce the adhesive force. We further believe that the bonded Z-DOL molecules can be orientated under stress (behaves as a soft polymer solid), which facilitates sliding and reduces coefficient of friction.

These studies suggest that if the lubricant films exist as liquid-like, such as Z-15 films, they easily form menisci (by themselves and the adsorbed water molecules), and thus have higher adhesive force and higher friction force, whereas if the lubricant film exists in a solid-like phase,

such as Z-DOL (fully bonded) films, they are hydrophobic with low adhesion and friction.

In order to study the uniformity of a lubricant film and its influence on friction and adhesion, friction force mapping and adhesive force mapping of PFPE have been carried out by Koinkar and Bhushan (1996b) and Bhushan and Dandavate (2000), respectively. Figure 14.9 shows gray scale plots of surface topography and friction force images obtained simultaneously for unbonded Demnum type PFPE lubricant film on silicon (Koinkar and Bhushan 1996b). Demnum type PFPE lubricant (Demnum, Daikin, Japan) chains have $-\text{CF}_2\text{--CH}_2\text{--OH}$, a reactive end group on one end, whereas Z-DOL chains have the hydroxyl groups on both ends, as described earlier (Bhushan 2001). The friction force plot shows well distinguished low and high friction regions roughly corresponding to high and low surface height regions in the topography image (thick and thin lubricant regions). A uniformly lubricated sample does not show such a variation in the friction. Figure 14.10 shows the gray scale plots of the adhesive force distribution for silicon samples coated uniformly and nonuniformly with Z-DOL lubricant. It can be clearly seen that there exists a region which has an adhesive force distinctly different from the other region for the nonuniformly coated sample. This

Fig. 14.9 Gray scale plots of the surface topography and friction force obtained simultaneously for unbonded 2.3-nm thick Demnum type PFPE lubricant film on silicon (Koinkar and Bhushan 1996a)

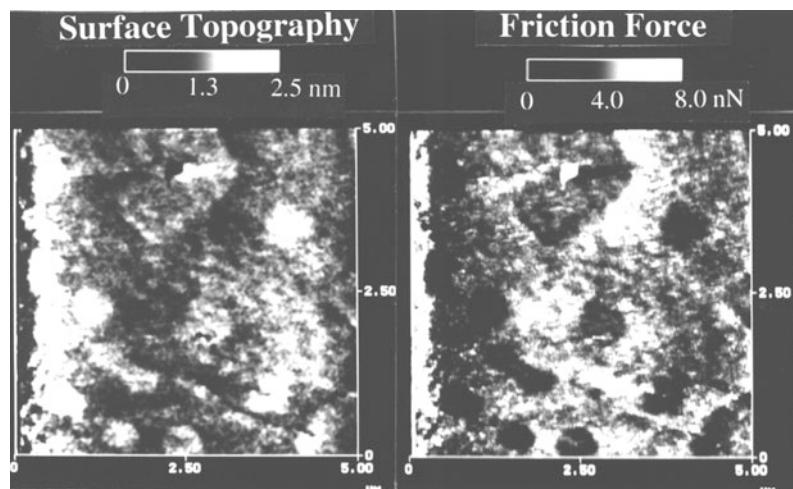
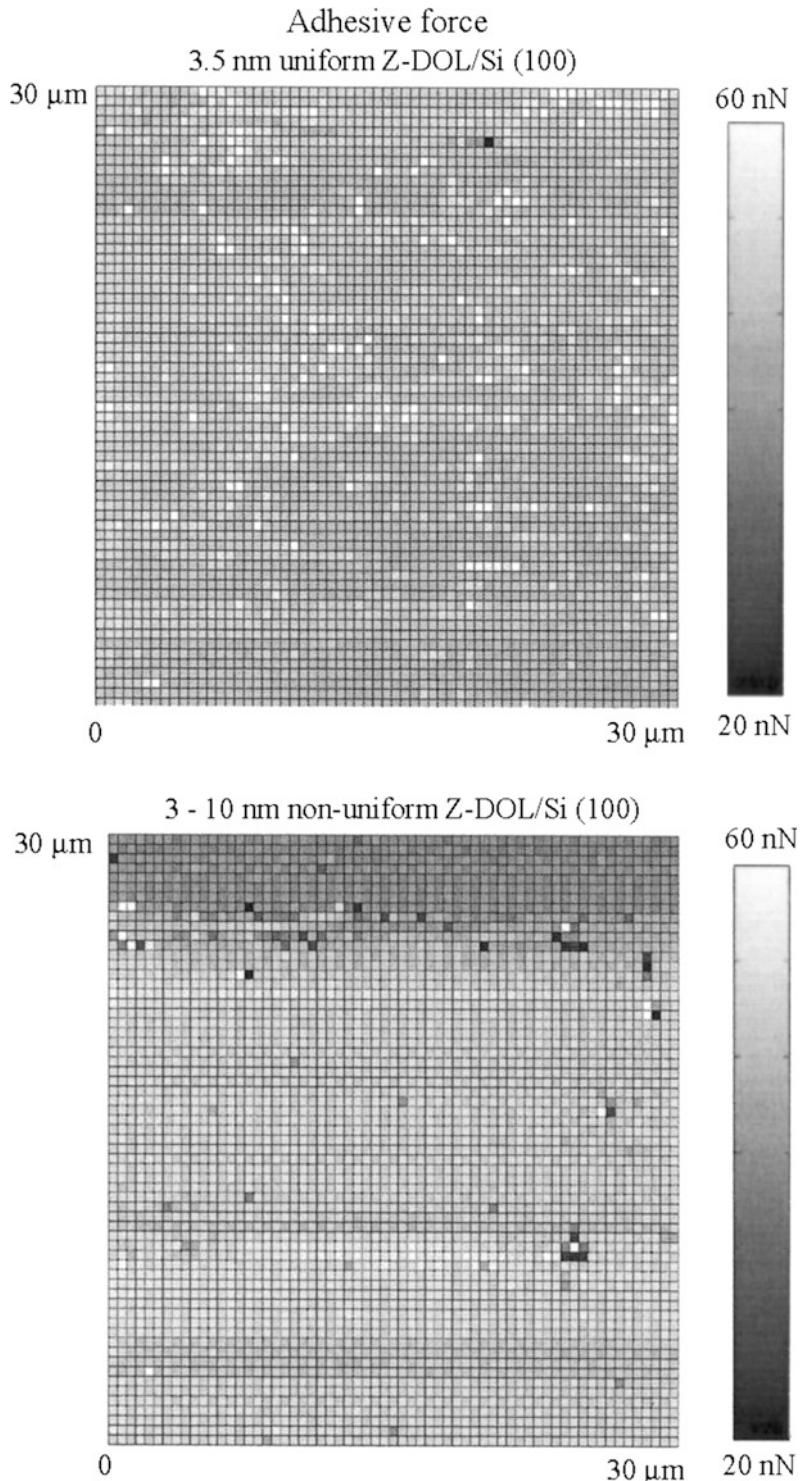


Fig. 14.10 Gray-scale plots of the adhesive force distribution of a uniformly-coated, 3.5-nm thick unbonded Z-DOL film on silicon and 3- to 10-nm thick unbonded Z-DOL film on silicon that was deliberately coated nonuniformly by vibrating the sample during the coating process (Bhushan and Dandavate 2000)



implies that the liquid film thickness is nonuniform giving rise to a difference in the meniscus forces.

14.2.2.2 Rest Time Effect

It is well known that in the computer rigid disk drive, the stiction force increases rapidly with an increase in rest time between the head and magnetic medium disk (Bhushan 1996, 2001). Considering that the stiction and friction are one of the major issues that leads to the failure of computer rigid disk drives and MEMS, it is very important to find out if the rest time effect also exists on the nanoscale. First the rest time effect on the friction force, adhesive force, and coefficient of Si(100) sliding against a Si_3N_4 tip was studied, Fig. 14.11a (Liu and Bhushan 2003). It is found that the friction and adhesive forces logarithmically increase up to a certain equilibrium time after which they remain constant. Figure 14.11a also shows that the rest time does not affect the coefficient of friction. These results suggest that the rest time can result in growth of the meniscus, which causes a higher adhesive force, and in turn a higher friction force. However, in the whole testing range the friction mechanisms do not change with the rest time. Similar studies were also performed on Z-15 and Z-DOL (fully bonded) films. The results are summarized in Fig. 14.11b (Liu and Bhushan 2003). It is seen that similar time effect has been observed on Z-15 film but not on Z-DOL (fully bonded) film.

An AFM tip in contact with a flat sample surface can be treated as a single asperity contact. Therefore, a Si_3N_4 tip in contact with Si(100) or Z-15/Si(100) can be modeled as a sphere in contact with a flat surface covered by a layer of liquid (adsorbed water and/or liquid lubricant), Fig. 14.12a. A meniscus forms around the contacting asperity and grows with time until equilibrium occurs (Chilamakuri and Bhushan 1999). The meniscus force, which is the product of meniscus pressure and meniscus area, depends on the flow of liquid phase toward the contact zone. The flow of the liquid towards the contact zone is governed by the capillary pressure P_c , which draws liquid into the meniscus, and the

disjoining pressure Π , which tends to draw the liquid away from the meniscus. Based on Young and Laplace equation, the capillary pressure, P_c is

$$P_c = 2\kappa\gamma \quad (14.3)$$

where 2κ is the mean meniscus curvature ($=\kappa_1 + \kappa_2$, where κ and κ_2 are the curvatures of the meniscus in the contact plane and perpendicular to the contact plane). Mate and Novotny (1991) have shown that the disjoining pressure decreases rapidly with increasing liquid film thickness in a manner consistent with a strong van der Waals attraction. The disjoining pressure, Π for these liquid films can be expressed as

$$\Pi = \frac{A}{6\pi h^3} \quad (14.4)$$

where A is the Hamaker constant, and h is the liquid film thickness. The driving forces that cause the lubricant flow that result in an increase in the meniscus force are the disjoining pressure gradient due to a gradient in film thickness and capillary pressure gradient due to curved liquid-air interface. The driving pressure, P , can then written as

$$P = -2\kappa\gamma - \Pi. \quad (14.5)$$

Based on these three basic relationships, the following differential equation has been derived by Chilamakuri and Bhushan (1999), which can describe the meniscus at time t

$$\begin{aligned} 2\pi x_0 \left(D + \frac{x_0^2}{2R} - h_0 \right) \frac{dx_0}{dt} \\ = \frac{2\pi h_0^3 \gamma (1 + \cos \theta)}{3\eta} - \frac{Ax_0}{3\eta h} \cot \alpha \end{aligned} \quad (14.6)$$

where η is the viscosity of the liquid, and a is given as

$$a = R(1 - \cos \phi) \sim \frac{R\phi^2}{2} \sim \frac{x_0^2}{2R}. \quad (14.7)$$

The differential Eq. 14.6 was solved numerically using Newton's iteration method. The meniscus force at any time t less than the

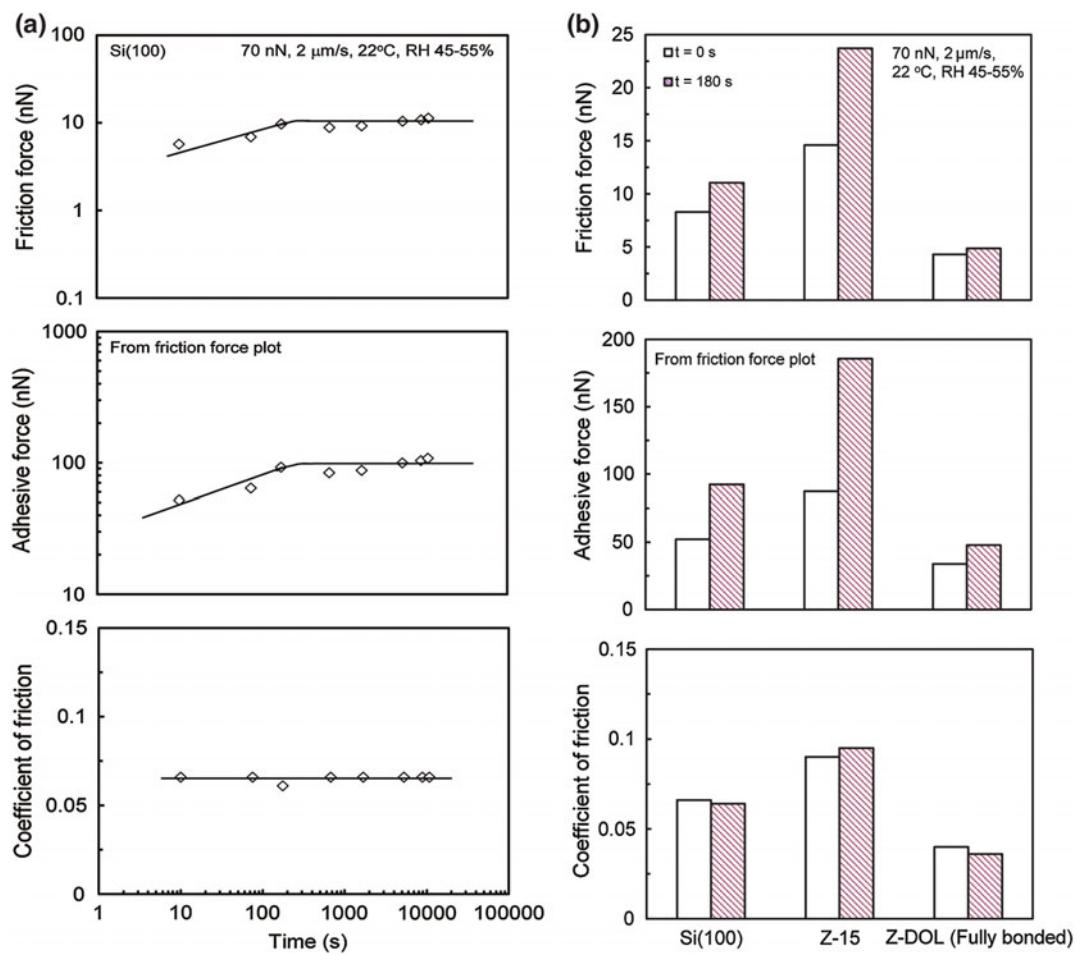


Fig. 14.11 **a** Rest time effect on friction force, adhesive force, and coefficient of friction of Si(100). **b** Summary of the rest time effect on friction force, adhesive force, and coefficient of friction of Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL(fully bonded) film. All of the measurements were carried out at 70 nN, 2 μm/s, and in ambient air (Liu and Bhushan 2003)

equilibrium time is proportional to the meniscus area and meniscus pressure ($2\kappa\gamma$), and it is given by

$$f_m(t) = 2\pi R\gamma(1 + \cos\theta) \left[\frac{x_0}{(x_0)_{eq}} \right]^2 \left[\frac{\kappa}{\kappa_{eq}} \right] \quad (14.8)$$

where $(x_0)_{eq}$ is the value of x_0 at the equilibrium time

$$\left[(x_0)_{eq} \right]^2 = 2R \left\{ \frac{-6\pi h_0^3(1 + \cos\theta)}{A} + (h_0 - D) \right\}. \quad (14.9)$$

This modeling work (at microscale) showed that the meniscus force initially increases logarithmically with the rest time up to a certain equilibrium time, after which it remains constant. Equilibrium time decreases with an increase in liquid film thickness, a decrease in viscosity, and a decrease in the tip radius, Fig. 14.12b. This

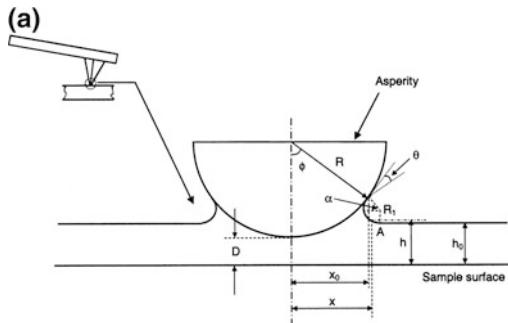


Fig. 14.12 **a** Schematic of a single asperity in contact with a smooth flat surface in the presence of a continuous liquid film when ϕ is large. **b** Results of the single asperity model. Effect of viscosity of the liquid, radius of the asperity and film thickness is studied with respect to the time dependent meniscus force (Chilamakuri and Bhushan 1999)

early numerical modeling work and the data at the nanoscale in Fig. 14.11a are in good agreement.

14.2.2.3 Velocity Effect

To investigate the velocity effect on friction and adhesion, the friction force versus normal load relationships of Si(100), Z-15, and Z-DOL (fully bonded) at different velocities were measured, Fig. 14.13 (Liu and Bhushan 2003). Based on these data, the adhesive force and coefficient of friction values can be calculated by Eq. 14.1. The variation of friction force, adhesive force, and coefficient of friction of Si(100), Z-15, and Z-DOL (fully bonded) as a function of velocity are summarized in Fig. 14.14. It indicates that for Si(100) the friction force decreases logarithmically with increasing velocity. For Z-15, the friction force decreases with increasing velocity up to $10 \mu\text{m/s}$, after which it remains almost constant. The velocity has very little effect on the friction force of Z-DOL (fully bonded), it reduced slightly only at very high velocity. Figure 14.14 also indicates that the adhesive force of Si(100) is increased when the velocity is higher than $10 \mu\text{m/s}$. The adhesive force of Z-15 is reduced dramatically with a velocity increase up to $20 \mu\text{m/s}$, after which it is reduced slightly, and

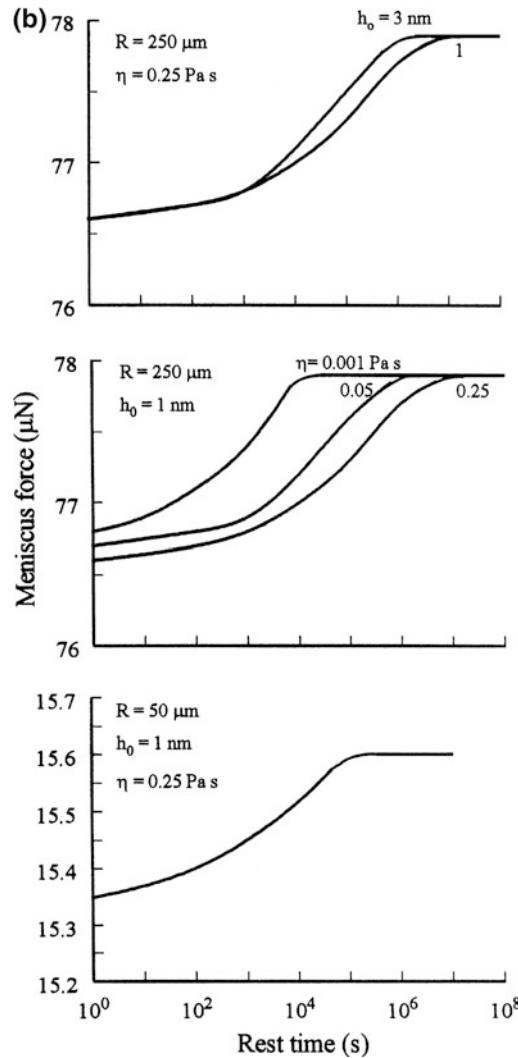


Fig. 14.12 (continued)

the adhesive force of Z-DOL (fully bonded) is also decreased at high velocity. In the testing range of velocity, only the coefficient of friction of Si(100) decreases with velocity, but the coefficients of friction of Z-15 and Z-DOL (fully bonded) remain almost constant. This implies that the friction mechanisms of Z-15 and Z-DOL (fully bonded) do not change with the variation of velocity.

The mechanisms of the effect of velocity on adhesion and friction are explained based on

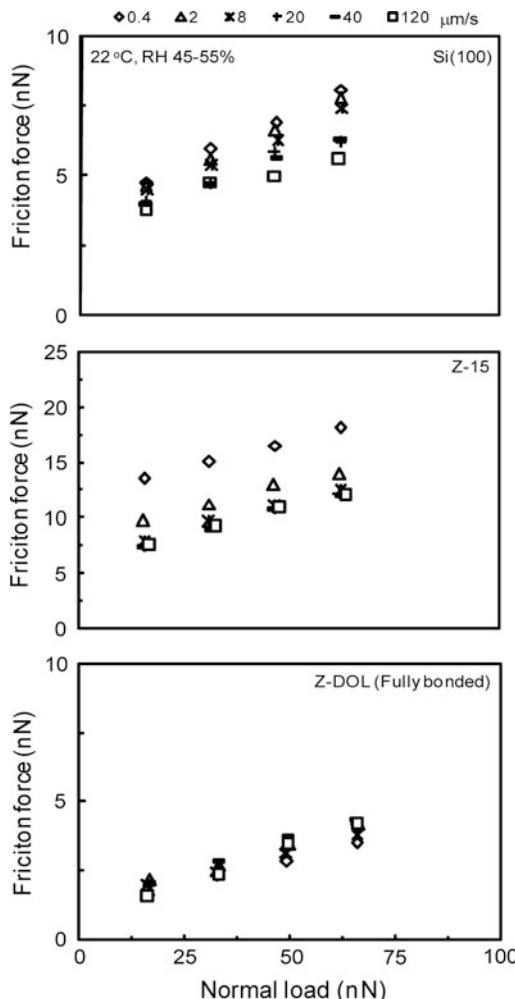
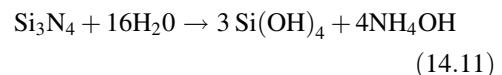
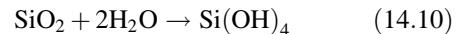


Fig. 14.13 Friction forces versus normal load data of Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film at various velocities in ambient air (Liu and Bhushan 2003)

schematics shown in Fig. 14.14 (right). For Si(100), tribochemical reaction plays a major role. Although, at high velocity, the meniscus is broken and does not have enough time to rebuild, the contact stresses and high velocity lead to tribochemical reactions of the Si(100) wafer and Si_3N_4 tip which have native oxide (SiO_2) layers with water molecules. The following reactions occur:



The $\text{Si}(\text{OH})_4$ is removed and continuously replenished during sliding. The $\text{Si}(\text{OH})_4$ layer between the tip and Si(100) surface is known to be of low shear strength and causes a decrease in friction force and coefficient of friction in the lateral direction (Ishigaki et al. 1986; Fischer 1988; Mizuhara and Hsu 1992; Danyluk et al. 1994; Muratov and Fischer 2000). The chemical bonds of Si–OH between the tip and Si(100) surface induce a large adhesive force in the normal direction. For Z-15 film, at high velocity the meniscus formed by condensed water and Z-15 molecules is broken and does not have enough time to rebuild, therefore, the adhesive force and consequently friction force is reduced. For Z-DOL (fully bonded) film, the surface can adsorb few water molecules in the ambient condition, and at high velocity these molecules are displaced, which is responsible for a slight decrease in friction force and adhesive force. Even at a high velocity range, the friction mechanisms for Z-15 and Z-DOL (fully bonded) films still are shearing of the viscous liquid and molecular orientation, respectively. Thus the coefficients of friction of Z-15 and Z-DOL (fully bonded) do not change with velocity.

Koinkar and Bhushan (1996a, b) have suggested that in the case of samples with mobile films such as condensed water and Z-15 films, alignment of liquid molecules (shear thinning) is responsible for the drop in friction force with an increase in scanning velocity. This could be another reason for the decrease in friction force for Si(100) and Z-15 film with velocity in this study.

14.2.2.4 Relative Humidity and Temperature Effect

The influence of relative humidity on friction and adhesion were studied in an environmentally

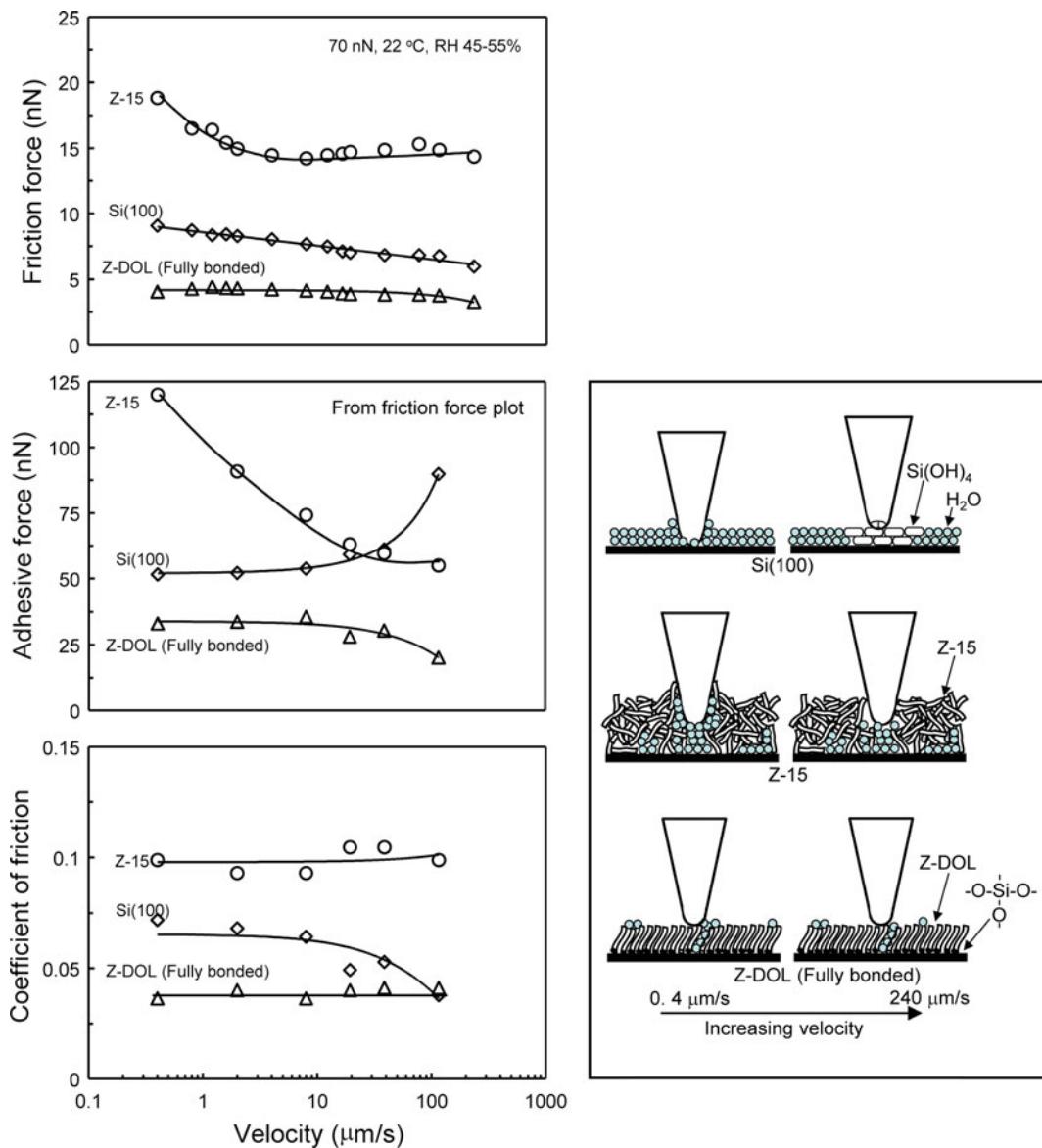


Fig. 14.14 The influence of velocity on the friction force, adhesive force and coefficient of friction of Si(100), and 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film at 70 nN, in ambient air. The schematic (right) shows the change of surface composition (by trabochemical reaction) and change of meniscus while increasing the velocity (Liu and Bhushan 2003)

controlled chamber. The friction force was measured by making measurements at increasing relative humidity, and the results are presented in Fig. 14.15 (Liu and Bhushan 2003). It shows that for Si(100) and Z-15 film, the friction force increases with a relative humidity increase up to

RH 45%, and then it shows a slight decrease with further increase in relative humidity. Z-DOL (fully bonded) has a smaller friction force than Si (100) and Z-15 in the whole testing range, and its friction force shows a relative apparent increase when the relative humidity is higher than RH

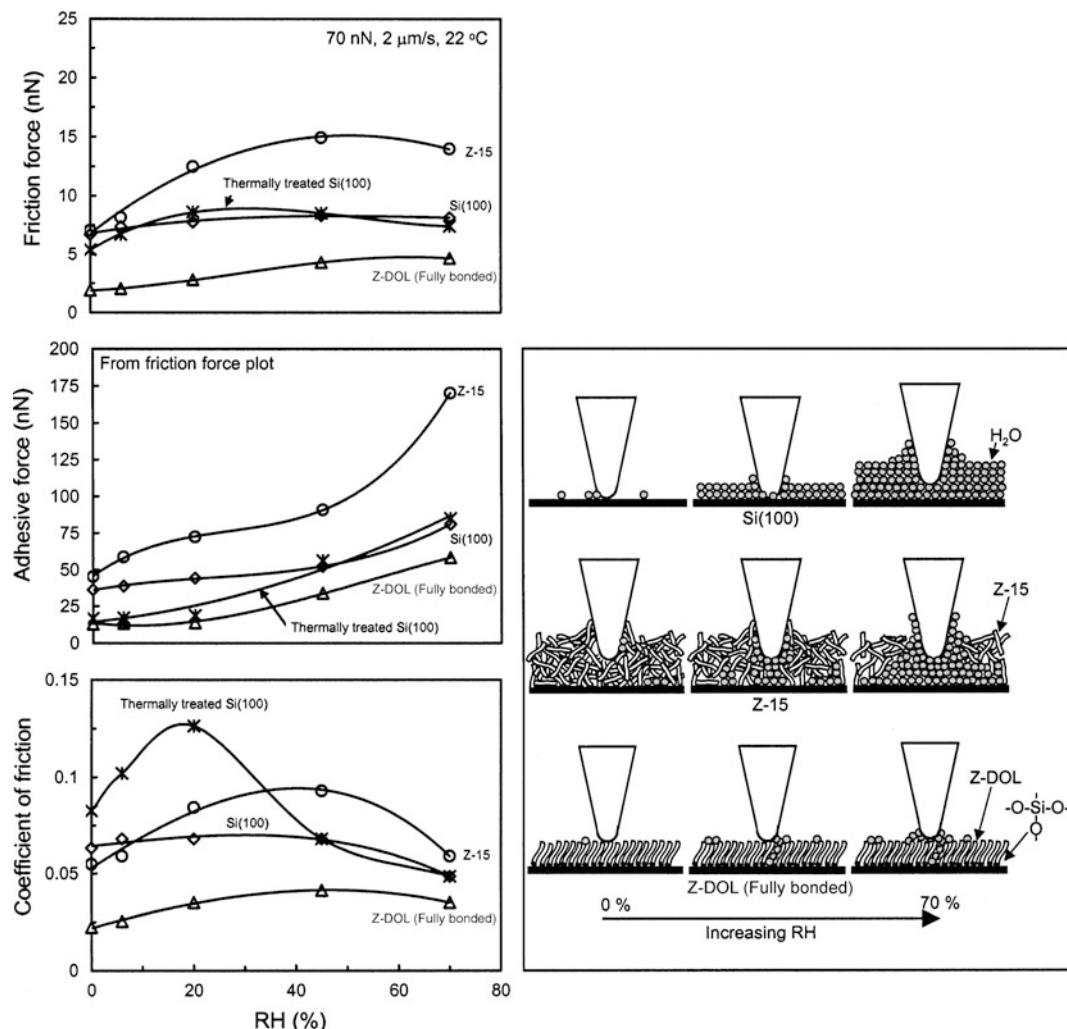


Fig. 14.15 The influence of relative humidity (RH) on the friction force, adhesive force, and coefficient of friction of Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film at 70 nN, 2 μm/s, and in 22 °C air. Schematic (right) shows the change of meniscus while increasing the relative humidity. In this figure, the thermally treated Si(100) represents the Si(100) wafer that was baked at 150 °C for 1 h in oven (in order to remove the adsorbed water) just before it was placed in the 0% RH chamber (Liu and Bhushan 2003)

45%. For Si(100), Z-15, and Z-DOL (fully bonded), their adhesive forces increase with relative humidity, and their coefficients of friction increase with a relative humidity up to RH 45%, after which they decrease with a further increase of relative humidity. It is also observed that the humidity effect on Si(100) really depends on the history of the Si(100) sample. As the surface of Si(100) wafer readily adsorbs water in air, without any pre-treatment the Si(100) used in our

study almost reaches its saturated stage of adsorbing water, and is responsible for less effect during increasing relative humidity. However, once the Si(100) wafer was thermally treated by baking at 150 °C for 1 h, a larger effect was observed.

The schematic (right) in Fig. 14.15 shows that Si(100), because of its high free surface energy, it can adsorb more water molecules with an increase in the relative humidity. As discussed

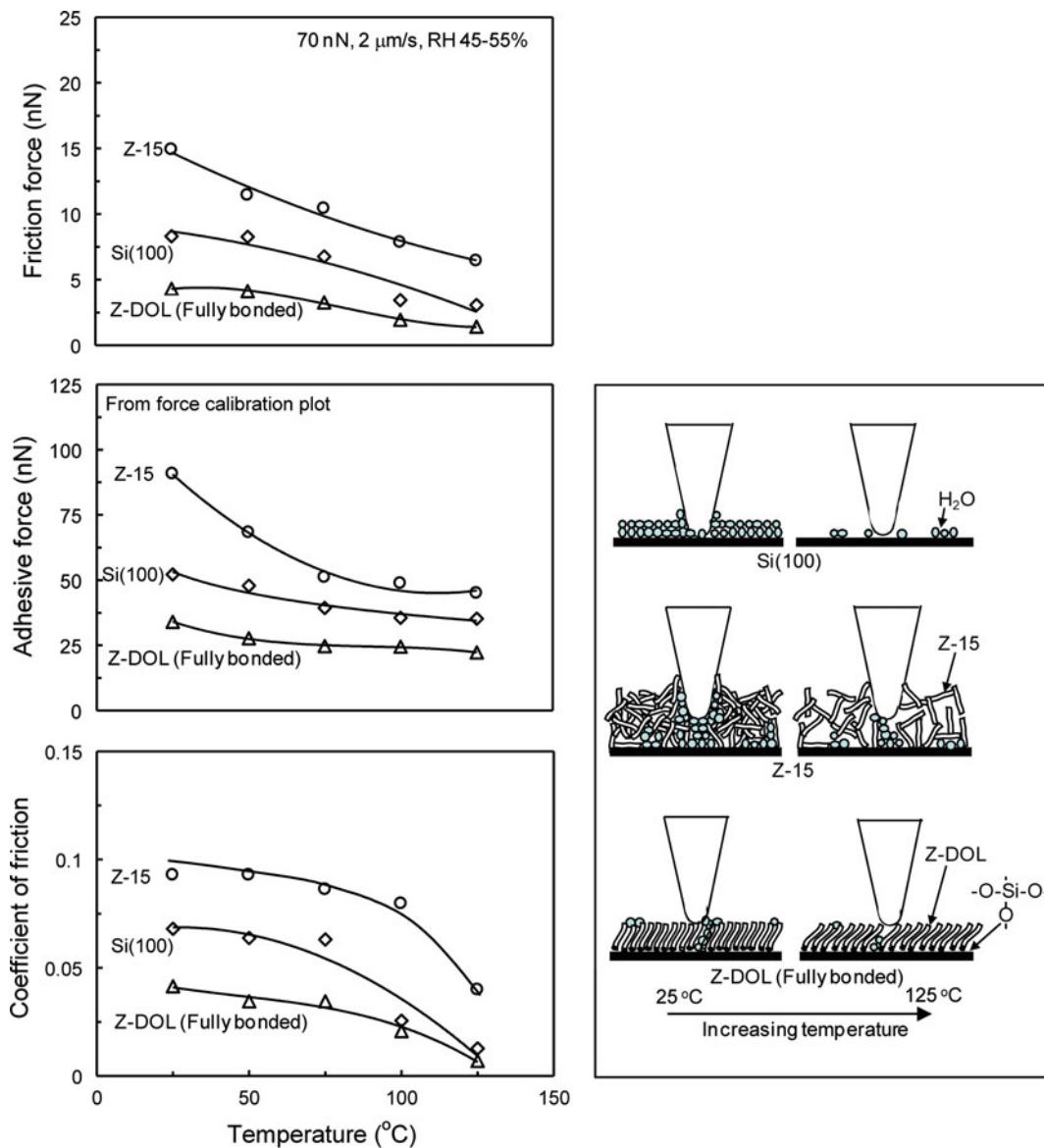


Fig. 14.16 The influence of temperature on the friction force, adhesive force, and coefficient of friction of Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film at 70 nN, at 2 μm/s, and in RH 40–50% air. The schematic (right) shows that at high temperature, desorption of water decreases the adhesive forces. And the reduced viscosity of Z-15 leads to the decrease of coefficient of friction. High temperature facilitates orientation of molecules in Z-DOL (fully bonded) film which results in lower coefficient of friction (Liu and Bhushan 2003)

earlier, for Z-15 film in a humid environment, the condensed water from the humid environment competes with the lubricant film present on the sample surface. Obviously, more water molecules also can be adsorbed on Z-15 surface while increasing relative humidity. The more adsorbed

water molecules in the case of Si(100), along with lubricant molecules in Z-15 film case, form a larger water meniscus, which leads to an increase of friction force, adhesive force, and coefficient of friction of Si(100) and Z-15 with humidity. However, at a very high humidity of

RH 70%, large quantities of adsorbed water can form a continuous water layer that separates the tip and sample surface and act as a kind of lubricant, which causes a decrease in the friction force and coefficient of friction. For Z-DOL (fully bonded) film, because of their hydrophobic surface properties, water molecules can be adsorbed and cause an increase in the adhesive force and friction force only at high humidity (\geq RH 45%).

The effect of temperature on friction and adhesion were studied using a thermal stage attached to the AFM. The friction force was measured by making measurements at increasing temperature from 22 to 125 °C. The results are presented in Fig. 14.16 (Liu and Bhushan 2003). It shows that the increasing temperature causes a decrease of friction force, adhesive force, and coefficient of friction of Si(100), Z-15, and Z-DOL (fully bonded). The schematic (right) in Fig. 14.16 indicates that at high temperature, desorption of water leads to a decrease of friction force, adhesive force and coefficient of friction for all of the samples. Besides that, the reduction of the surface tension of water also has a contribution to the decrease of friction and adhesion. For Z-15 film, the reduction of viscosity at high temperature has an additional contribution to the decrease of friction. In the case of Z-DOL (fully bonded) film, molecules are easier to be oriented at high temperature, which may be also responsible for the low friction.

Using a surface force apparatus, Yoshizawa and Israelachvili (Yoshizawa et al. 1993; Yoshizawa and Israelachvili 1993) have shown that a change in the velocity or temperature induces phase transformation (from crystalline solid-like to amorphous, then to liquid-like) in surfactant monolayers, which are responsible for the observed changes in the friction force. Stick-slip is observed in the low velocity regime of a few $\mu\text{m/s}$, and adhesion and friction first increase followed by a decrease in the temperature range of 0–50 °C. Stick-slip at low velocity and adhesion and friction curves peaking at some particular temperature (observed in their study) have not been observed in the AFM study. It suggests that the phase transformation may not happen in this

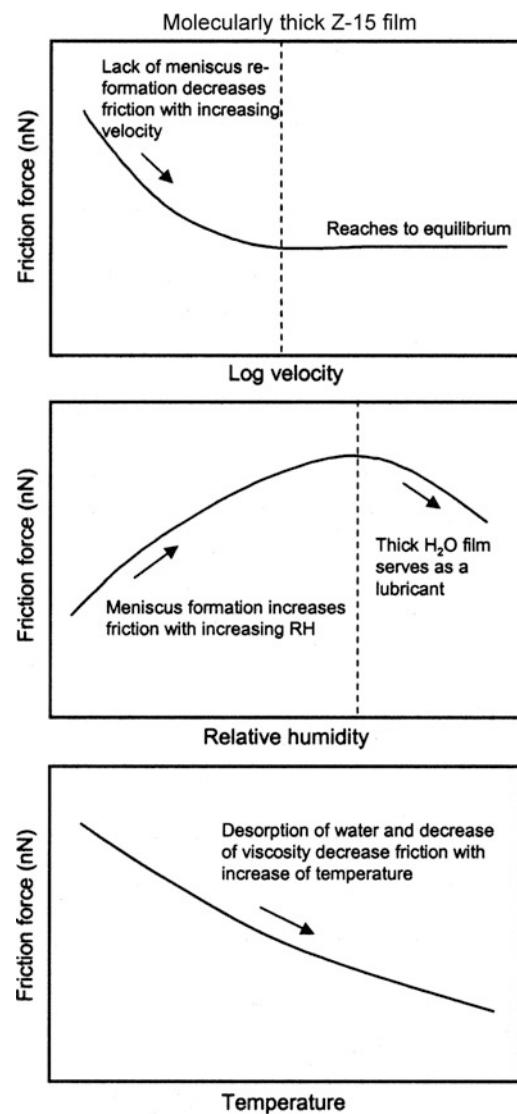


Fig. 14.17 Schematic shows the change of friction force of molecularly thick Z-15 films with log velocity, relative humidity, and temperature (Liu and Bhushan 2003)

study. This is because PFPEs generally have very good thermal stability (Bhushan 1993, 1996).

As a brief summary, the influence of velocity, relative humidity, and temperature on the friction force of Z-15 film is presented in Fig. 14.17. The changing trends are also addressed in this figure.

14.2.2.5 Tip Radius Effect

The tip radius and relative humidity affect adhesion and friction for unlubricated and

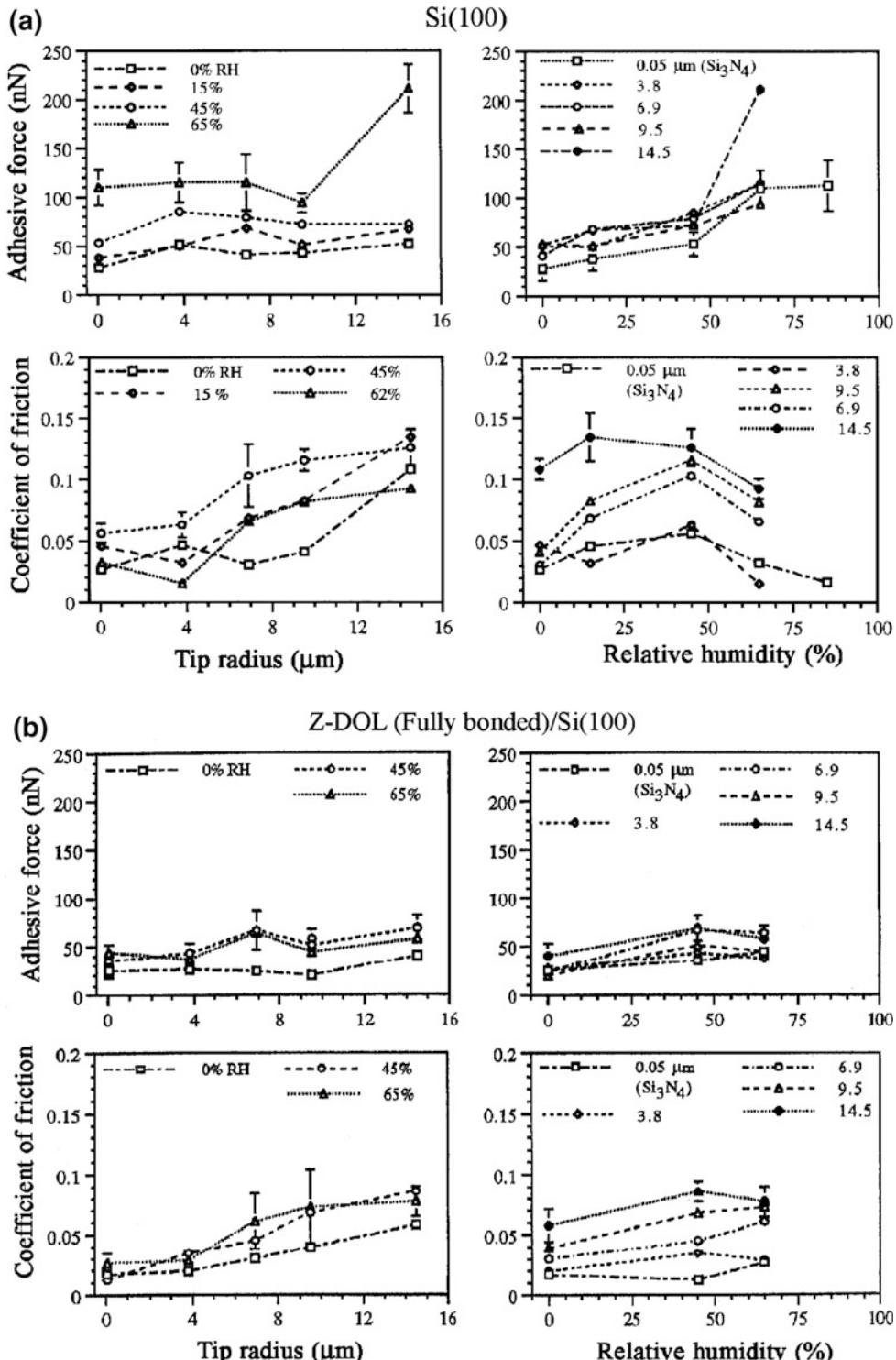


Fig. 14.18 Adhesive force and coefficient of friction as a function of tip radius at several humidities and as a function of relative humidity at several tip radii on **a** Si(100), and **b** 0.5-nm Z-DOL (fully bonded) films (Bhushan and Sundararajan 1998)

lubricated surfaces (Bhushan and Sundararajan, 1998; Bhushan and Dandavate 2000). Figure 14.18a shows the variation of single point adhesive force measurements as a function of tip radius on a Si(100) sample for several humidities. The adhesive force data are also plotted as a function of relative humidity for various tip radii. Figure 14.18a indicates that the tip radius has little effect on the adhesive forces at low humidities, but the adhesive force increases with tip radius at high humidity. Adhesive force also increases with an increase in humidity for all tips. The trend in adhesive forces as a function of tip radii and relative humidity, in Fig. 14.18a, can be explained by the presence of meniscus forces, which arise from the capillary condensation of water vapor from the environment. If enough liquid is present to form a meniscus bridge, the meniscus force should increase with an increase in tip radius based on Eq. (14.2). This observation suggests that the thickness of the liquid film at low humidities is insufficient to form continuous meniscus bridges to affect adhesive forces in the case of all tips.

Figure 14.18a also shows the variation in coefficient of friction as a function of tip radius at a given humidity and as a function of relative humidity for a given tip radius on the Si(100) sample. It can be observed that for RH 0%, the coefficient of friction is about the same for the tip radii except for the largest tip, which shows a higher value. At all other humidities, the trend consistently shows that the coefficient of friction increases with tip radius. An increase in friction with tip radius at low to moderate humidities arises from increased contact area (i.e. higher van der Waals forces) and higher values of shear forces required for larger contact area. At high humidities, similar to adhesive force data, an increase with tip radius occurs because of both contact area and meniscus effects. It can be seen that for all tips, the coefficient of friction increases with humidity to about RH 45%, beyond which it starts to decrease. This is attributed to the fact that at higher humidities, the adsorbed water film on the surface acts as a lubricant between the two surfaces (Liu and Bhushan 2003). Thus the interface is changed at

higher humidities, resulting in lower shear strength and hence lower friction force and coefficient of friction.

Figure 14.18b shows adhesive forces as a function of tip radius and relative humidity on Si(100) coated with a 0.5 nm thick Z-DOL (fully bonded) film. Adhesive forces for all the tips with the Z-DOL (fully bonded) lubricated sample are much lower than those experienced on unlubricated Si(100) shown in Fig. 14.18a. The data also show that even at a monolayer thickness of the lubricant, there is very little variation in adhesive forces with tip radius at a given humidity. For a given tip radius, the variation in adhesive forces with relative humidity indicates that these forces slightly increase from RH 0% to RH 45%, but remain more or less the same with a further increase in humidity. This is seen even with the largest tip which indicates that the lubricant is indeed hydrophobic; there is some meniscus formation at humidities higher than RH 0%, but the formation is very minimal and does not increase appreciably even up to RH 65%. Figure 14.18b also shows the coefficient of friction for various tips at different humidities for the Z-DOL (fully bonded) lubricated sample. Again, all the values obtained with the lubricated sample are much lower than the values obtained on unlubricated Si(100) shown in Fig. 14.18a. The coefficient of friction increases with tip radius for all humidities as was seen on unlubricated Si(100) due to an increase in the contact area. Similar to the adhesive forces, there is an increase in friction from RH 0% to RH 45% due to a contribution from increased number of menisci bridges. However, there is very little additional water film forming due to the hydrophobicity of the Z-DOL (fully bonded) layer thereafter, and consequently, the coefficient of friction does not change appreciably, even with the largest tip. These findings show that even a monolayer of Z-DOL (fully bonded) offers good hydrophobic performance of the surface.

14.2.2.6 Wear

To study the durability of lubricant films at the nanoscale, the friction of Si(100), Z-15, and

Z-DOL (fully bonded) as a function of the number of scanning cycles were measured, Fig. 14.19 (Liu and Bhushan 2003). As observed earlier, friction force of Z-15 is higher than that of Si(100), and Z-DOL (fully bonded) has the lowest value. During cycling, friction force of Si(100) shows a slight variation during the initial few cycles then remains constant. This is related to the removal of the top adsorbed layer. In the case of Z-15 film, the friction force shows an increase during the initial few cycles and then approaches to higher and stable values. This is believed to be caused by the attachment of the Z-15 molecules onto the tip. The molecular interaction between these attached molecules to the tip and molecules on the film surface is responsible for an increase in the friction. However, after several scans, this molecular interaction reaches equilibrium, and after that friction force and coefficient of friction remain constant. In the case of Z-DOL (fully bonded) film, the friction force starts out low and remains low

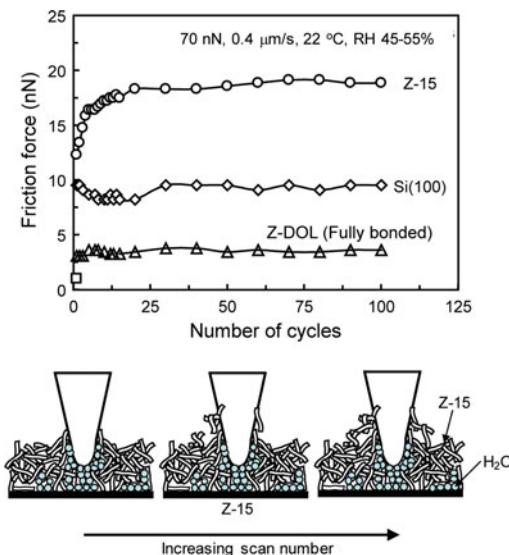


Fig. 14.19 Friction force versus number of sliding cycles for Si(100), 2.8-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film at 70 nN, 0.8 $\mu\text{m}/\text{s}$, and in ambient air. Schematic (bottom) shows that some liquid Z-15 molecules can be attached onto the tip. The molecular interaction between the attached molecules onto the tip with the Z-15 molecules in the film results in an increase of the friction force with multi scanning (Liu and Bhushan 2003)

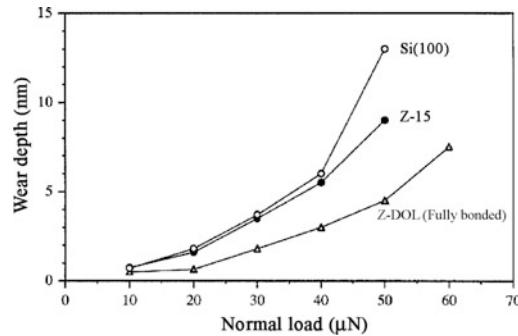


Fig. 14.20 Wear depth as a function of normal load using a diamond tip for Si(100), 2.9-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) after one cycle (Koinkar and Bhushan 1996a)

during the entire test for 100 cycles. It suggests that Z-DOL (fully bonded) molecules do not get attached or displaced as readily as Z-15.

Koinkar and Bhushan (1996a, b) conducted wear studies using a diamond tip at high loads. Figure 14.20 shows the plots of wear depth as a function of normal force, and Fig. 14.21 shows the wear profiles of the worn samples at 40 μN normal load. The 2.3 nm thick Z-DOL (fully bonded) lubricated sample exhibits better wear resistance than unlubricated and 2.9 nm thick Z-15 lubricated silicon samples. Wear resistance of the Z-15 lubricated sample is little better than that of the unlubricated sample. The Z-15 lubricated sample shows debris inside the wear track. Since Z-15 is a liquid lubricant, the debris generated is held by the lubricant, and it becomes sticky, which moves inside the wear track and does damage, Fig. 14.20. These results suggest that Z-DOL (fully bonded) exhibits higher wear resistance of the substrate as compared to that for Z-15.

To study the effect of the degree of chemical bonding, durability tests were conducted on both fully bonded and partially bonded Z-DOL films. Durability results for Z-DOL (fully bonded) and Z-DOL bonded and unwashed Z-DOL (partially bonded) (which is a partially bonded film, and contains both bonded and mobile phase lubricants) with different film thicknesses are shown in Fig. 14.22 (Koinkar and Bhushan 1996b). Thicker films, such as Z-DOL (partially bonded)

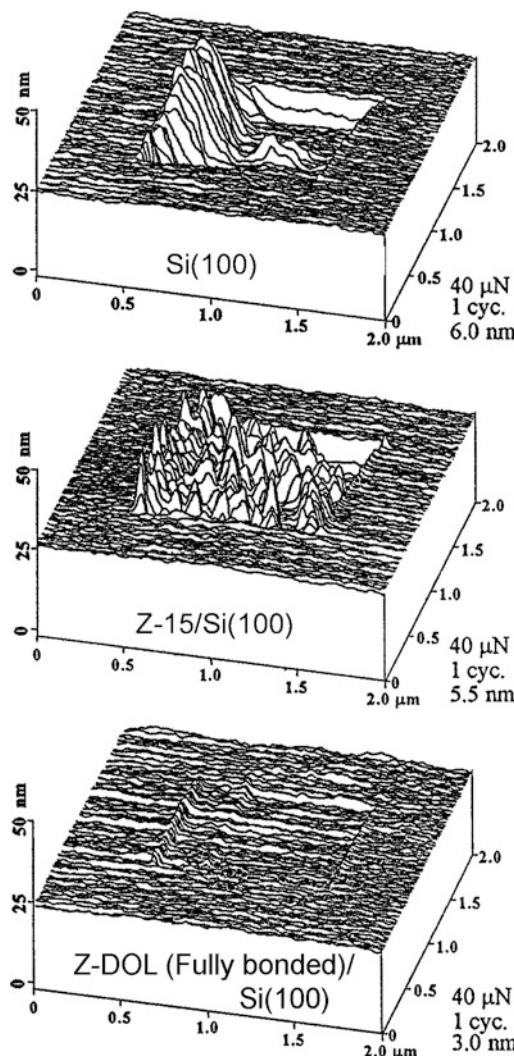


Fig. 14.21 Wear profiles for Si(100), 2.9-nm thick Z-15 film, and 2.3-nm thick Z-DOL (fully bonded) film after wear studies using a diamond tip. Normal force used and wear depths are listed in the figure (Koinkar and Bhushan 1996a)

with a thickness of 4.0 nm (bonded/mobile = 2.3 nm/1.7 nm) exhibit behavior similar to 2.3 nm thick Z-DOL (fully bonded) film. Figure 14.22 also indicates that Z-DOL (fully bonded) and Z-DOL (partially bonded) films with a thinner film thickness exhibit a higher friction value. Comparing 1.0 nm thick Z-DOL (fully bonded) with 3.0 nm (bonded/mobile = 1.0 nm/2.0 nm) thick Z-DOL (partially bonded), the Z-DOL (partially bonded) film exhibits lower and

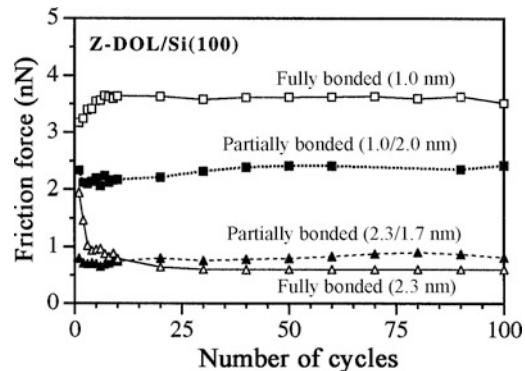


Fig. 14.22 Friction force as a function of number of cycles using Si_3N_4 tip at a normal load of 300 nN for Z-DOL (fully bonded) and Z-DOL (partially bonded) films with different film thicknesses (Koinkar and Bhushan 1996b)

stable friction values. This is because the mobile phase on a surface acts as a source of lubricant replenishment. Similar conclusions have also been reported by Ruhe et al. (1996), Bhushan and Zhao (1999), and Eapen et al. (2002). All of them indicate that using partially bonded Z-DOL films can dramatically reduce friction and improve the wear life.

14.3 Nanotribological, Electrical, and Chemical Degradations Studies and Environmental Effects in Novel PFPE Lubricant Films

Electrical properties of lubricant films are of interest in various MEMS/NEMS applications. Changes in the surface potential and electrical resistance can be measured during sliding using an AFM (DeVecchio and Bhushan 1998; Palacio and Bhushan 2007a, b, 2008a, b, 2009b; Bhushan et al. 2008a). These techniques are also useful for wear detection and for studying the initiation of wear (DeVecchio and Bhushan 1998; Bhushan 2008b). Palacio and Bhushan (2007a, b, 2009b) carried out nanotribological studies on various novel PFPE lubricant films and monitored the electrical properties as well. Chemical degradation studies and environmental effects on various PFPE lubricant films on a Si

(100) wafer and amorphous carbon (diamondlike carbon) coated magnetic tapes were carried out in a macroscale configuration by Tao and Bhushan (2005), Bhushan and Tao (2006), Bhushan et al. (2007).

Structure and properties of several novel PFPE lubricants commonly used in the lubrication of magnetic rigid disks—Z-TETRAOL 2000 and A20H-2000—are presented in Table 14.2 (Tao and Bhushan 2005). Z-DOL 2000 is also included for comparison. Z-TETRAOL (Solvay Solexis Inc.) is a derivate of PFPE. The backbone of Z-TETRAOL is the same as the conventional PFPE lubricant Z-DOL (Solvay Solexis Inc.) as described earlier. The difference between Z-TETRAOL and Z-DOL is that Z-TETRAOL has two hydroxyl groups at each end while Z-DOL has one hydroxyl group at each end. It is believed that two hydroxyl bonds are expected to lead to stronger interaction with the substrate. However, the Z-TETRAOL lubricant film is less mobile because of larger viscosity which may lead to shorter durability. A20H (Moresco, Japan) is a PFPE lubricant with a cyclotriphosphazene at one end and a hydroxyl at the other end. The backbone of A20H is also the same as that of Z-DOL. Phosphazene lubricants (such as X1-P) have been used as additives in the data storage industry because they exhibit better

durability in high humidity environments (Bhushan and Zhao 1999; Bhushan 2001). It is believed that X1-P coats the mating surface and makes it hydrophobic and minimizes stiction and improves durability. Studies have shown that A20H exhibits less thinning or reduced mobility during drive rotation (Tao and Bhushan 2005). The durability of these less mobile lubricants could be less than that of the highly mobile Z-DOL + X1-P systems. A mixture of Z-DOL and A20H is known to provide longer durability and good performance at high humidity.

The lubricants were applied on single-crystal Si(100) with a native oxide layer on the surface using the dip coating technique by Tao and Bhushan (2005) and Palacio and Bhushan (2007a). The wafer was ultrasonicated in acetone followed by methanol for 10 min each. This was then followed by soaking in the solvent HFE 7100 (3 M, St. Paul, MN), which consists of isomers of methoxynonafluorobutane ($C_4F_9OCH_3$). The cleaned wafer was vertically submerged into a beaker containing a dilute solution of the lubricant in HFE 7100 for 10 min and then pulled off. The lubricated sample used without post thermal treatment is referred to as untreated. Partially bonded samples were prepared by heating at 150 °C for 30 min after dip coating, while the fully bonded samples were heated at 150 °C for

Table 14.2 Chemical structures and selected properties of several PFPE lubricants (data obtained from manufacturers' data sheet)

Lubricant	End group (X ^a)	Molecular weight (amu)	Density ($\times 10^3$ kg/m ³ at 20 °C)	Kinematic viscosity (mm ² /s)	Surface tension (mJ/m ²)	Vapor pressure (Torr)	
						20 °C	100 °C
Z-DOL 2000 ^b	–CF ₂ –CH ₂ –OH	2000	1.81	85	24	2×10^{-5}	6×10^{-4}
Z-TETRAOL 2000 ^b	–CF ₂ –CH ₂ –O–CH ₂ –CH(OH)–CH ₂ –OH	2300	1.75	2000	–	5×10^{-7}	2×10^{-4}
A20H-2000 ^c	–CH ₂ –OH and –CH ₂ –N ₃ P ₃ (OC ₆ H ₄ –CF ₃) ₅	3000	1.7	65	22	–	–

^aThe whole molecular chain: X–CF₂–O–(CF₂–CF₂–O)_m–(CF₂–O)_n–CF₂–X (m/n = 2/3)

^bSolvay Solexis, Inc., Thorofare, NJ

^cMoresco Matsumura Oil Research Corp., Kobe-city, Hyogo, Japan

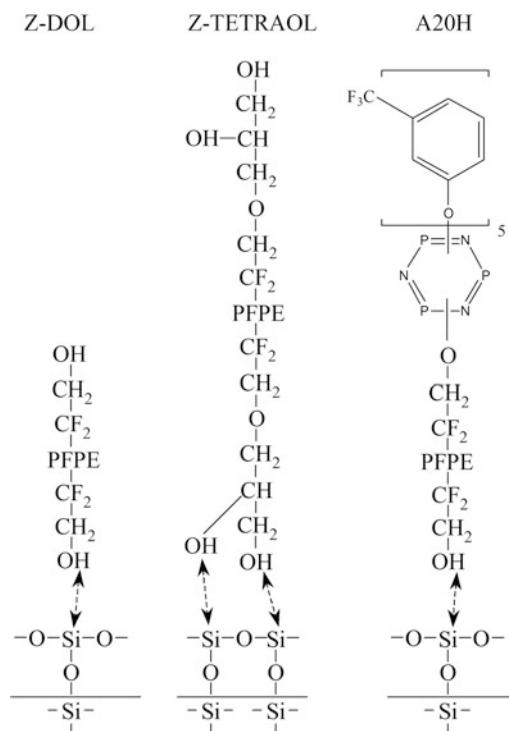


Fig. 14.23 Schematics of Z-DOL, Z-TETRAOL, and A20H molecules bonded on Si substrates

30 min and washed in the HFE 7100 solvent to remove the mobile fraction. The lubricant coated silicon samples were then measured with an ellipsometer and the coating thickness was found to be about 1, 3, and 7 nm for the fully bonded, partially bonded and untreated samples, respectively. A schematic of the bonding of the Z-DOL, Z-TETRAOL, and A20H endgroups to the silicon substrate is shown in Fig. 14.23.

14.3.1 Nanotribological Studies

Adhesive force and coefficient of friction measurements were made using an AFM and the results are presented in Fig. 14.24 (Palacio and Bhushan 2007a). The lubricant-coated samples reduce adhesion compared to the uncoated

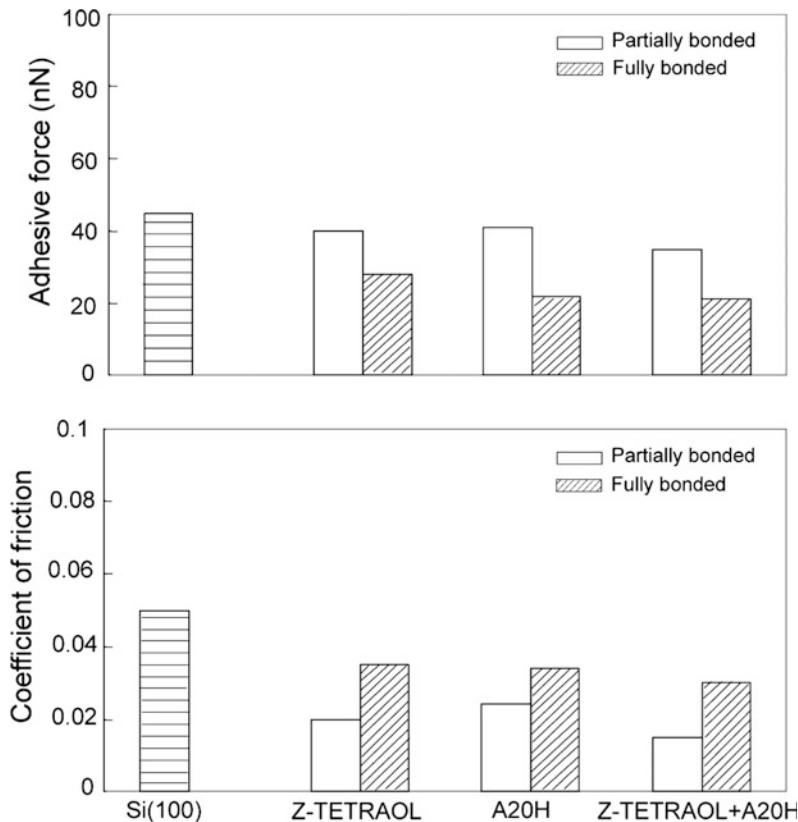
silicon. The adhesive forces measured on the partially bonded lubricant films are higher than the data from their fully bonded counterparts. As discussed earlier, the mobile fraction on the surface of the partially bonded sample facilitates the formation of a meniscus, which increases the tip-surface adhesion. All of the lubricant-coated samples exhibit a reduction in the coefficient of friction relative to the uncoated silicon. The partially bonded samples have a lower coefficient of friction compared to the fully bonded samples, implying that the mobile lubricant molecules in the former facilitate sliding of the tip on the surface. This will be analyzed in more detail in the following subsection in the context of wear, which was monitored using surface potential and resistance measurements.

Palacio and Bhushan (2009b) studied the effect of relative humidity and temperature on the friction and adhesion of novel PFPE films deposited on amorphous carbon (diamondlike carbon) coated magnetic tapes. The trends were similar to that as reported earlier in Sect. 14.2.2.4.

14.3.2 Wear Detection by Surface Potential Measurements

Wear experiments were performed on the lubricated surfaces using a diamond tip over a $5 \mu\text{m} \times 5 \mu\text{m}$ region at a load of $10 \mu\text{N}$ for 20 cycles. Figure 14.25a shows the surface height and surface potential images for the fully bonded and partially bonded lubricants (Palacio and Bhushan 2007a). The corresponding images for the uncoated silicon substrate are shown for comparison. It can be seen from the images that debris is generated around the wear region for all coatings tested. More wear debris is observed in the fully bonded samples than in the partially bonded samples. Samples with the fully bonded lubricant have limited wear protection compared to those with partially bonded lubricants because

Fig. 14.24 Summary of the adhesive force and coefficient of friction of Si (100), Z-TETRAOL, A20H and Z-TETRAOL + A20H films in room temperature (22 °C) and ambient air (45–55% RH) (Palacio and Bhushan 2007a)

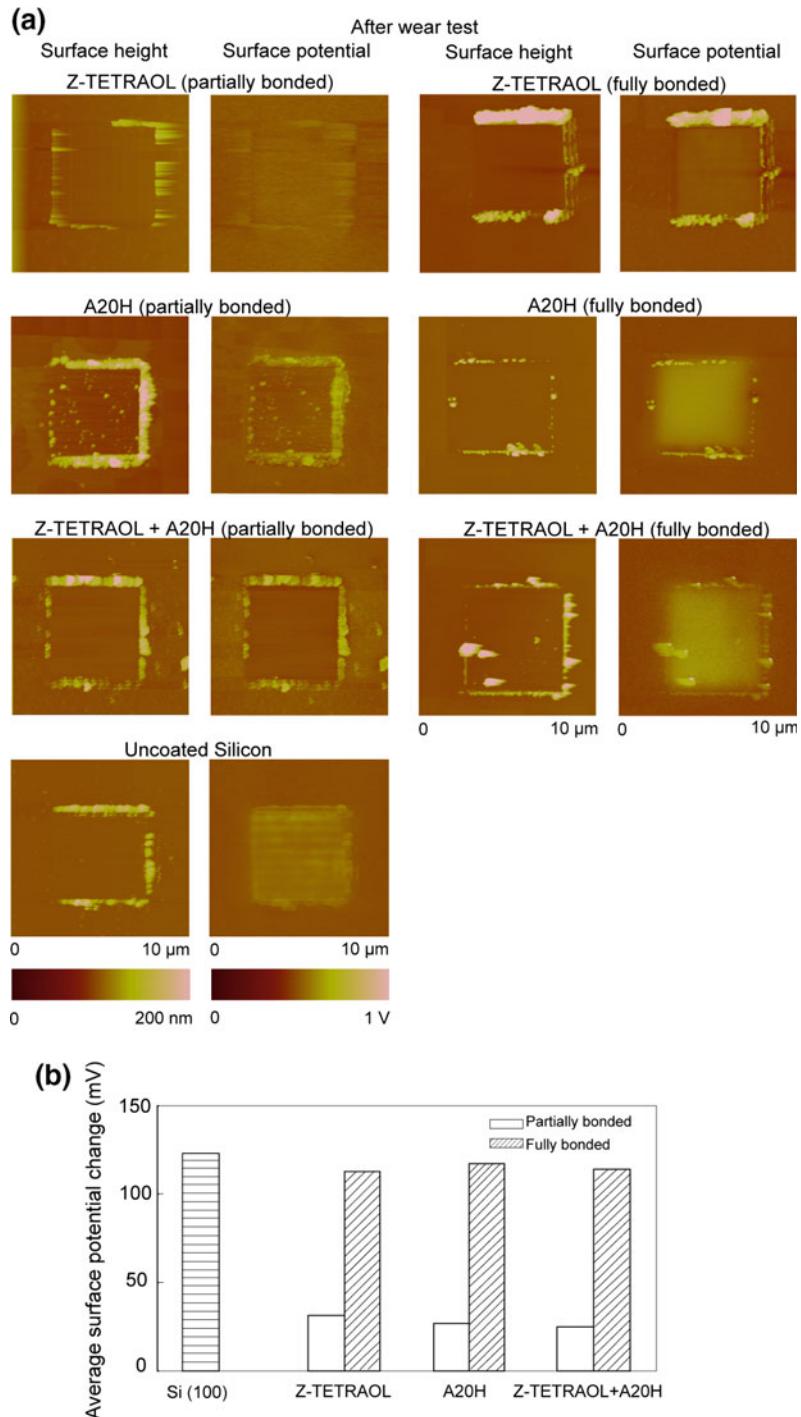


the former has only immobile molecules, while the latter has both immobile and mobile molecules. The mobile fraction of the lubricant can replenish the surface, i.e., it can move to the worn area and protect it after the immobile lubricant molecules have been displaced from the surface after repeated rubbing. Since the partially bonded films are thicker than the fully bonded films, additional wear protection is attributed to a thickness effect.

Figure 14.25b presents a bar chart showing surface potential changes for various samples. The surface potential of the area subjected to the wear test increased, which is more prominent in the fully bonded samples. The partially bonded samples exhibited a smaller change in surface potential, indicating less wear. These findings should be correlated to the result of the wear test

on the uncoated silicon sample, where the increase in surface potential is well-understood. The Kelvin probe method measures the surface potential difference between the tip and sample, which pertains to differences in the work functions between these two materials. For conducting and semiconducting materials, the mechanism is as follows. The surface potential is altered during physical wear because the Fermi energy level is altered. This is the energy required to remove an electron to a point just outside the material surface. Thermodynamic equilibrium is disrupted with a change in the Fermi level, and can only be restored by the flow of electrons either into or from the area subjected to wear. This mechanism does not apply to materials such as SiO_2 (naturally present as a thin layer) and the lubricants, which are both insulators. Physical wear on these

Fig. 14.25 **a** Surface height and surface potential maps of the coatings under investigation after wear testing. Brighter areas correspond to higher values of both the height and surface potential change, and **b** chart showing surface potential change (Palacio and Bhushan 2007a)



materials would not cause a change in the surface potential because charge dissipation is poor. Therefore, a considerable surface potential change would be observed only when: (1) the

lubricant has been fully removed from the substrate; (2) the native SiO_2 layer has been abraded from the surface; and (3) wear has caused subsurface structural changes. However, for

insulators such as the lubricant, electrostatic charges are introduced as it comes into contact with a material with a dissimilar electron affinity (the diamond tip) during the wear test. The charges on the insulating lubricant surface have low mobility and would eventually dissipate into the ambient environment. In the presence of debris, the electrostatic charges become localized and may get trapped (as debris particles are mostly isolated), causing an increase in the surface potential of the debris that formed around the worn area.

14.3.3 Wear Detection by Electrical Resistance Measurements of Z-TETRAOL and the Effect of Cycling

For electrical resistance measurements, lubricant films were applied on evaporated Au film deposited on a Si substrate. The silicon wafer used in the experiments is lightly doped, and it was coated with Au so that a metal-metal contact would be attained once the lubricant is removed. PFPE is insulating. Figure 14.26 presents surface height and resistance images for a fully bonded Z-TETRAOL coating on Au by using a scanning spreading resistance microscopy (SSRM) attachment in an AFM. For comparison, another set of tests was conducted on the same sample, and the post-wear surface potential images were obtained as shown in the same figure. The Au surface is expected to have a low resistance (corresponding to darker areas in the resistance map), while the lubricant should have high resistance (lighter areas) as it is an insulator. The surface height images show a small amount of wear after 1 and 5 cycles. In both cases, the resistance images are featureless, which implies that the lubricant is still present, and the Au is not yet exposed. A surface potential increase is observed, which is attributed to electrostatic charge build-up, but this does not imply full lubricant removal since the resistance of the test area has remained unchanged to this point. After 20 cycles, the contact resistance in the tested area

has suddenly decreased, which indicates exposure of the underlying Au. The wear debris has the same resistance compared to the unworn lubricant, indicating that the current flow is influenced by the inherent difference in conductivity between the probe tip and the material it is in contact with.

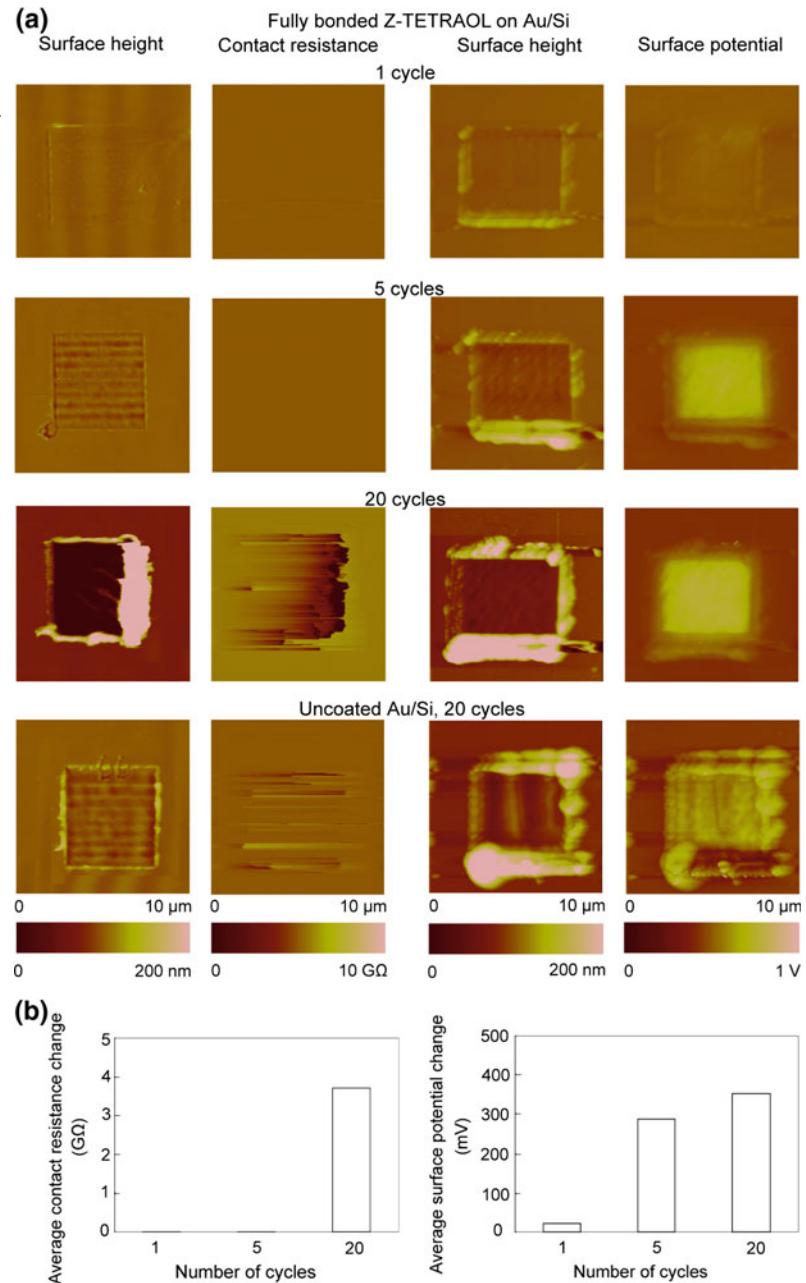
During the scan of the area subjected to the wear test, the current measured by the SSRM sensor corresponds to the contact resistance which exists between the metal-coated tip and the sample. The spreading resistance of the electric current flowing within the semiconductor sample is not measured because this is only present when there is direct contact between the metal-coated tip and a highly doped semiconductor sample.

Because the lubricant is very soft, some signal instability could occur. This comes from possible tip contamination by the lubricant as well as plowing of wear debris during the scan. Since this AFM-based resistance measurement is a contact technique, contamination and plowing are more likely to occur compared to the surface potential measurement, which is a tapping technique. However, this technique is of interest because it provides complementary information to the Kelvin probe method in measuring the extent of wear of conducting films and lubricants, which are materials that can potentially be used in MEMS/NEMS devices.

14.3.4 Chemical Degradation and Environmental Studies

Tao and Bhushan (2005) carried out chemical degradation studies in a high vacuum tribotest apparatus (Bhushan and Ruan 1994b; Bhushan et al. 1995b). In this apparatus, a lubricated wafer mounted on a flexible cantilever beam was slid against an uncoated Si(100) wafer in a macroscale configuration. The system was equipped with a mass spectrometer so that gaseous emissions from the interface could be monitored *in situ* during the sliding in high vacuum and other controlled environments. The

Fig. 14.26 **a** Surface height, contact resistance and surface potential maps of fully bonded Z-TETRAOL coating on Au as a function of the number of wear cycles, and **b** bar chart showing contact resistance and surface potential change (Palacio and Bhushan 2007a)



normal load and friction force at the contacting interface were measured using resistive-type strain gage transducers. For the sliding tests, the lubricated Si(100) sample was glued onto a flat surface at the end of a rotating shaft. The sample was slid against a Si(100) wafer mounted on the flat surface of a slider integrated with a

flexible cantilever used in magnetic rigid disk drives. The sliding speed was 0.3 m/s and applied pressure was 150 kPa. The environmental effects were investigated in high vacuum (2×10^{-7} Torr), argon, dry air (less than 2% RH), ambient air (30% RH) and high humidity air (70% RH).

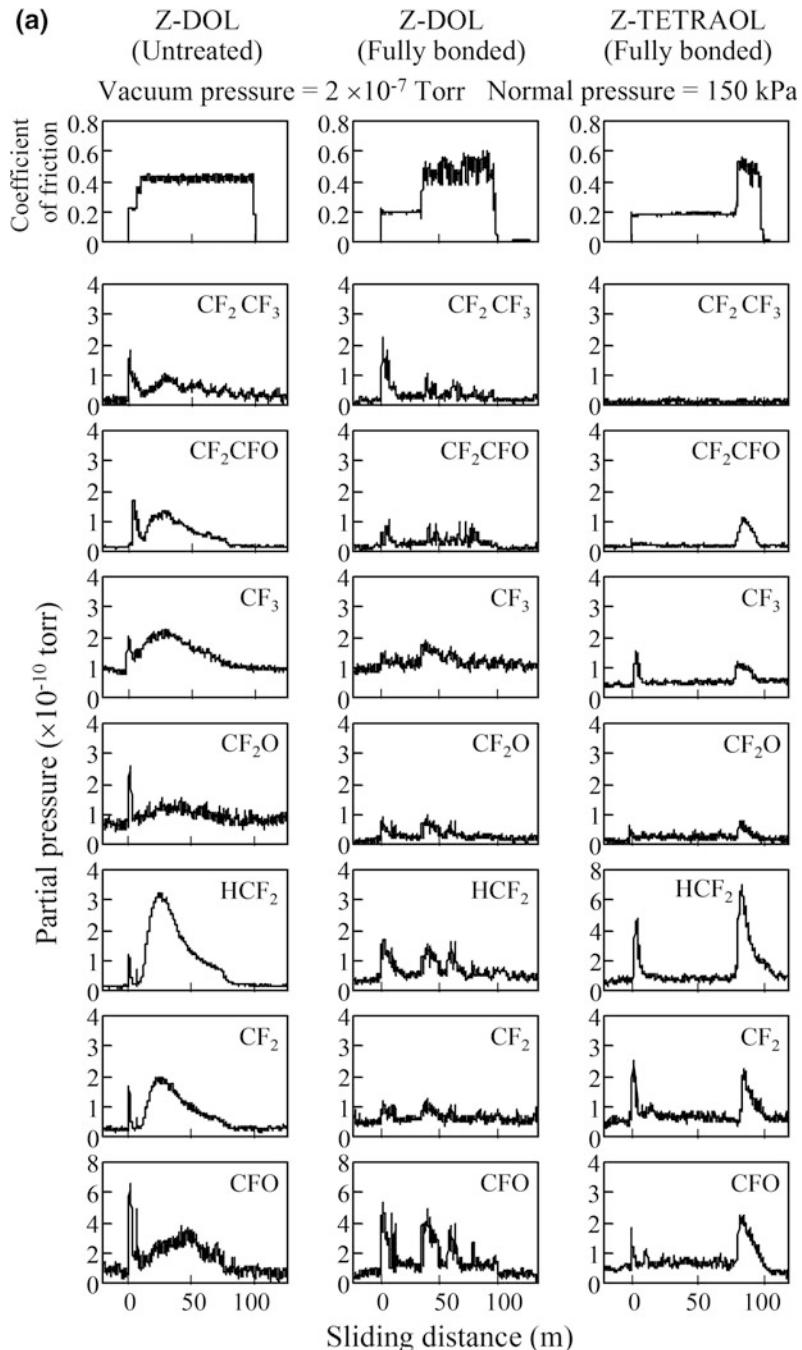
14.3.4.1 Chemical Degradation Studies

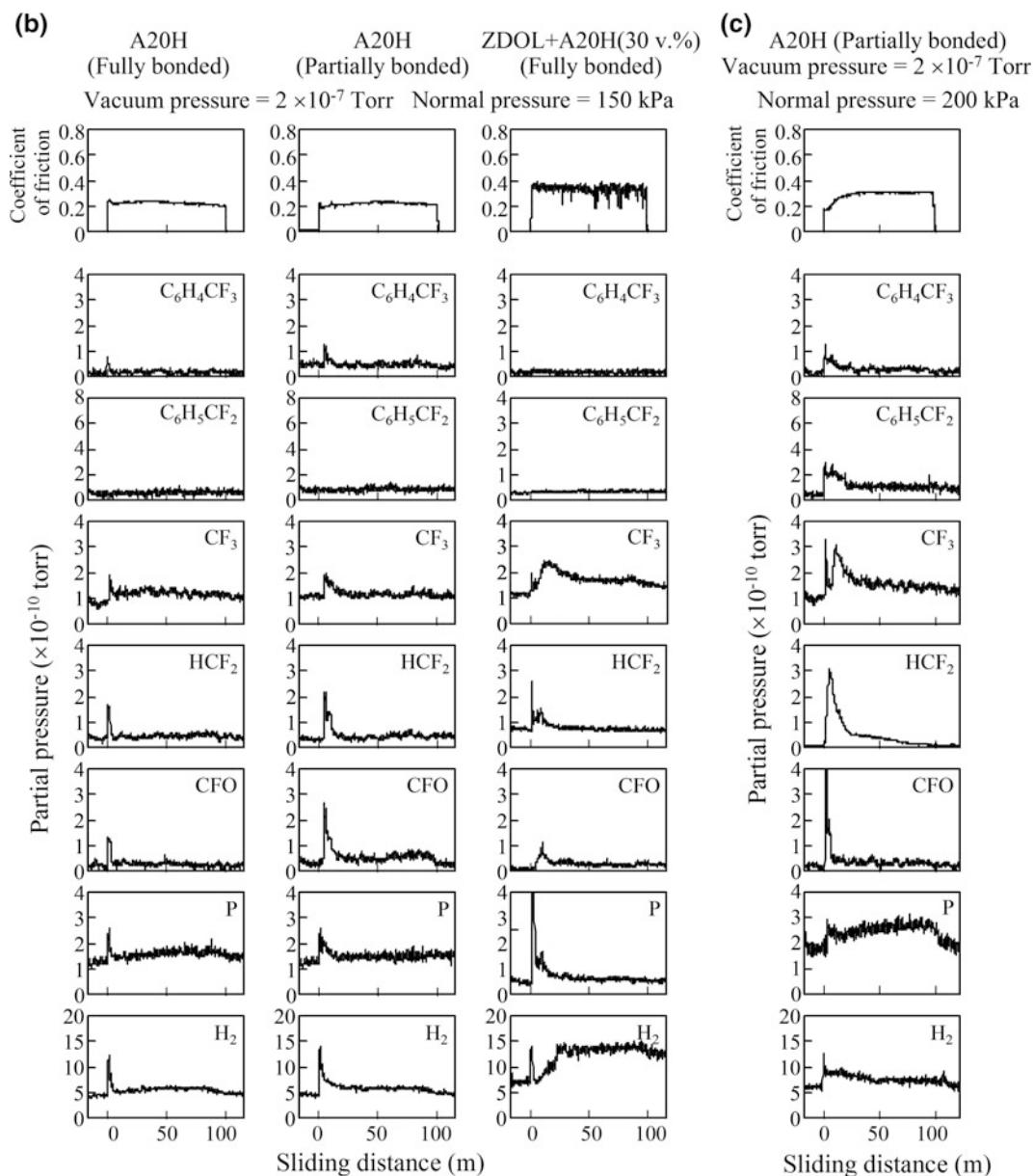
Coefficient of friction and partial pressure of the gaseous products for fully bonded Z-DOL, Z-TETRAOL, A20H, Z-DOL + A20H (30 vol %), and partially bonded A20H during sliding in high vacuum are shown in Fig. 14.27a–c. The

result for untreated Z-DOL is also shown in the figure as a comparison.

In the tests, the sharp increase of friction indicates the failure of the lubricant film. Therefore, the durability of the lubricant film can be obtained from the friction curve. Under the

Fig. 14.27 Coefficients of friction and mass spectra data on **a** untreated (2.2 nm) and fully bonded (2.3 nm) Z-DOL, fully bonded Z-TETRAOL (2.6 nm), **b** fully bonded (0.9 nm) and partially bonded (2.3 nm) A20H, and Z-DOL + A20H (2.4 nm) in high vacuum and under 150 kPa normal pressure, **c** partially bonded A20H in high vacuum and under 200 kPa normal pressure (Tao and Bhushan 2005)



**Fig. 14.27** (continued)

normal pressure of 150 kPa, the untreated Z-DOL failed immediately after sliding. The fully bonded Z-DOL began to fail after sliding of about 40 m. The fully bonded Z-TETRAOL, began to fail at about 80 m. Fully bonded and partially bonded A20H, however, did not fail during the 100 m sliding. The fully bonded Z-DOL + A20H (30 vol.%) failed immediately

after sliding. The results show that fully bonded Z-TETRAOL and A20H are more durable than fully bonded Z-DOL, while the untreated Z-DOL is less durable than the bonded films. The PFPE lubricants can be bonded to the applied surface through the hydroxyl group. Z-TETRAOL, with two hydroxyl groups at each end of the molecular chain, can be bonded more tightly on the silicon

surface than Z-DOL. The high durability of A20H, however, is beyond expectation. To further confirm the result, the normal pressure of 200 kPa was applied on partially bonded A20H. At the normal pressure of 200 kPa, the coefficient of friction was found to increase. At the same time, gaseous products were detected (see Fig. 14.27c). The final coefficient of friction was about 0.3, which is lower than the values (about 0.4) of the other films after failure. This could indicate that the A20H film was only partially worn. The reason for lower durability of Z-DOL + A20H (30 vol.%) in high vacuum is not clear.

In the tests on Z-DOL and Z-TETRAOL films, CF_2CF_3 , CF_2CFO , CF_3 , CF_2O , HCF_2 , CF_2 , and CFO were found to increase when friction increased. In the test of A20H films, $\text{C}_6\text{H}_4\text{CF}_3$, $\text{C}_6\text{H}_5\text{CF}_2$, CF_3 , HCF_2 , CFO, P, and H_2 were detected. For Z-DOL + A20H (30 vol.%), $\text{C}_6\text{H}_4\text{CF}_3$, $\text{C}_6\text{H}_5\text{CF}_2$, CF_3 , HCF_2 , CFO, P, and H_2 were also detected during sliding. Degradation mechanisms for Z-DOL and various model lubricants have been studied by Zhao and Bhushan (2000, 2001) and Zhao et al. (2000). Based on their work, Z-DOL starts to decompose above 350 °C. In the sliding conditions used here, it is not likely that the frictional heat can generate such a high temperature. The possibility of catalytic degradation is very low because of the small contact area. Triboelectrical reaction and mechanical scission are considered to be the dominant mechanisms during the sliding of Si on PFPE films. Electron emission is known to occur during sliding for both metal and nonmetal surfaces. The interaction of electrons with PFPE molecules could create various radicals such as $\cdot\text{CF}_2\text{O}\text{--CF}_2\cdot$, $\cdot\text{CF}_2\text{--CF}_2\text{O}\text{--CF}_2\cdot$, and/or $\cdot\text{CF}_2\text{--CF}_3$. These radicals can decompose or react with each other and/or the remaining PFPE molecules. A detailed description of triboelectrical reactions can be found in Zhao et al. (2000) and Zhao and Bhushan (2001).

Mechanical scission is another mechanism in causing the degradation of PFPE films. The PFPE molecule has a long linear chain structure. The C–C and C–O bonds in the molecular chain are easily subjected to cleaving

by the microasperities on the rubbing surfaces, which results in decomposition of the lubricant in the sliding process.

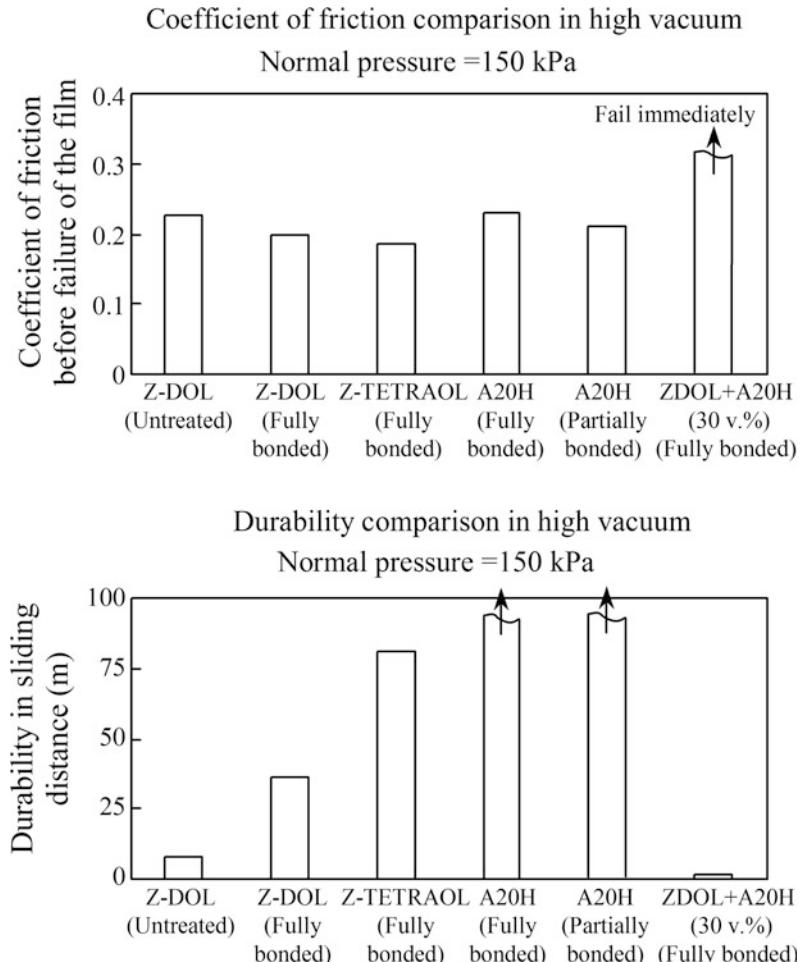
A summary of the coefficients of friction and durability for tested lubricants in high vacuum is shown in Fig. 14.28. The coefficients of friction of fully bonded Z-DOL and Z-TETRAOL are lower than untreated Z-DOL and A20H in high vacuum. Z-TETRAOL shows higher durability than Z-DOL, while A20H (both fully bonded and partially bonded) shows even high durability in high vacuum.

14.3.4.2 Environmental Studies

The wear tests were conducted in high vacuum, argon, dry air (less than 2% RH), air with 30% RH, and air with 70% RH (Tao and Bhushan 2005). The applied normal pressure was 150 kPa for all tested films, which was the same as in the degradation tests. Test results are presented in Fig. 14.29. The coefficients of friction show that friction and durability vary with the environment. Z-DOL films fail more rapidly in high vacuum than in other environments. As described in the previous section, there are very few foreign molecules on the contacting surfaces in high vacuum. This enables an intimate contact between the lubricant film and the counterpart surface. A tendency of chemical bonding occurs between the lubricant film and the counterpart surface. In argon, the Z-DOL films exhibited lower friction and higher durability than in dry air. The chemical effects of oxygen can influence the friction and durability of the lubricant films (Bhushan and Ruan 1994b). The water molecules at a moderate humidity level (ambient air) can act as a lubricant between the contacting surfaces. However, in a high humidity environment, the water molecules can penetrate the Z-DOL film and result in a nonuniform distribution of the Z-DOL molecules. The fully bonded Z-DOL film is less influenced by environment than untreated Z-DOL. The reason is that the untreated Z-DOL has more free hydroxyl end groups. The water molecules in the environment can interact with the hydroxyl group of Z-DOL via hydrogen bond.

Z-TETRAOL is more durable than Z-DOL. Especially, a high humidity environment does

Fig. 14.28 Coefficient of friction and durability comparison of Z-DOL, Z-TETRAOL, A20H, Z-DOL + A20H in high vacuum (Tao and Bhushan 2005)



not seem to have negative effects on Z-TETRAOL, although it has one more hydroxyl end group than Z-DOL and may attract more water molecules in high humidity. This could be because the two-hydroxyl end group provides better attachment than Z-DOL on the surface. The tight bonding reduces the probability of nonuniform distribution of the film. The coefficient of friction of fully bonded Z-TETRAOL is very close to fully bonded Z-DOL in argon and air with various humidity levels. In high vacuum, Z-TETRAOL shows a lower coefficient of friction than Z-DOL.

Both fully bonded and partially bonded A20H were very durable in high vacuum. In argon, fully bonded and partially bonded A20H did not fail

during the 100 m sliding. In air with various humidity levels, however, the fully bonded A20H failed. The durability is lower compared to Z-DOL. The low durability can be from the low thickness of the film, which is only 0.9 nm. For the 2.3-nm thick partially bonded A20H film, the durability in air with various humidity levels is apparently improved. Especially, at high humidity level, partially bonded A20H exhibits high durability. A20H consists of a phosphazene group that is large in size and protects the surface. Phosphazene lubricant has been used as an additive and is known to have high durability in a high humidity environment. The coefficient of friction of A20H, however, is higher (in the range of 5–15%) than that of Z-DOL in various environments.

Fig. 14.29 Coefficient of friction data in high vacuum, argon, and air with different humidity levels of untreated (2.2 nm) and fully bonded (2.3 nm) Z-DOL, fully bonded Z-TETRAOL (2.6 nm), fully bonded (0.9 nm) and partially bonded (2.3 nm) A20H, and Z-DOL + A20H (2.4 nm) (Tao and Bhushan 2005)

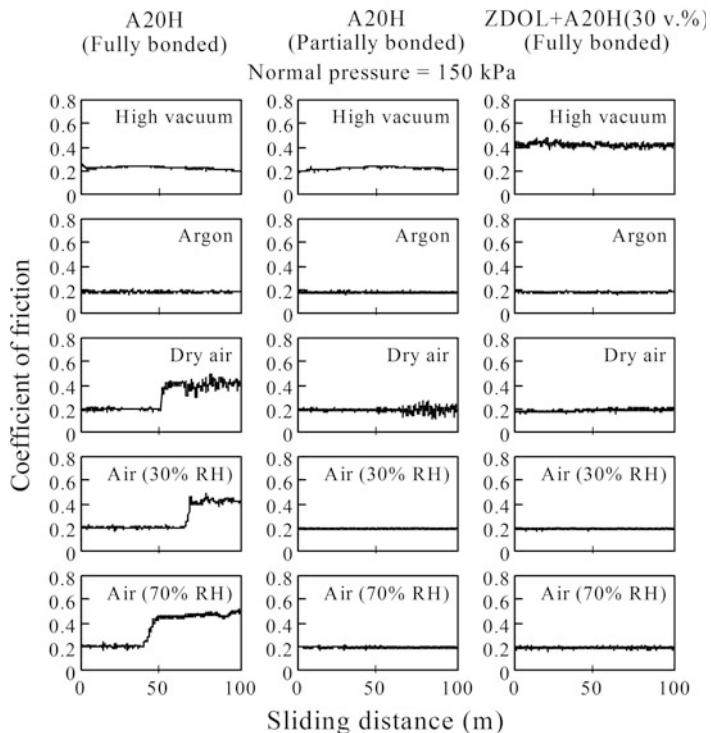
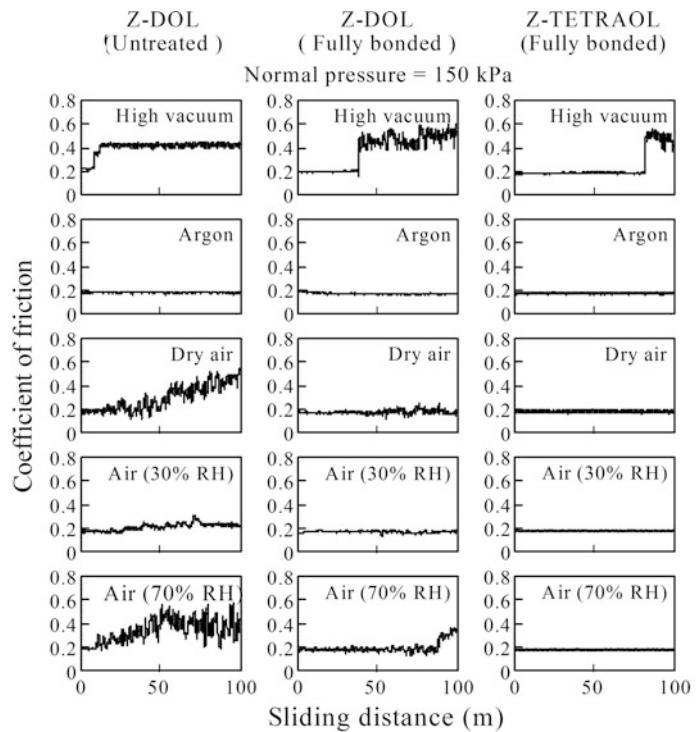
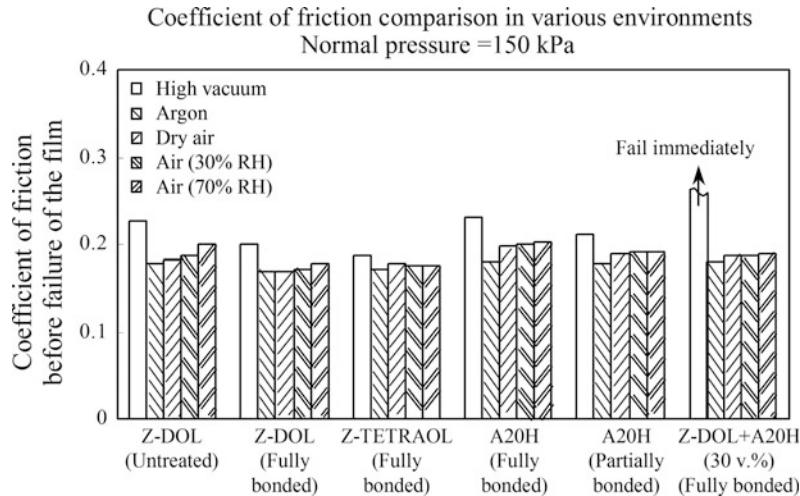


Fig. 14.30 Comparison of coefficient of friction data for Z-DOL, Z-TETRAOL, A20H, and Z-DOL + A20H in high vacuum, argon, and air with different humidity levels (Tao and Bhushan 2005)



Z-DOL + A20H (30 vol.%) shows low durability in high vacuum. In argon and air with various humidity levels, however, the film did not fail during the 100 m sliding. The durability of the mixture is higher than fully bonded Z-DOL. The coefficient of friction of the mixture is higher than that of Z-DOL while lower than that of A20H.

A summary of the coefficients of friction before the failure of the lubricant films in various environments is presented in Fig. 14.30. To summarize the highlights, Z-TETRAOL exhibits higher durability than Z-DOL. A20H exhibits high durability in high vacuum, in argon, and in air with various humidity levels. The mixture of Z-DOL + A20H (30 vol.%) shows low durability

in high vacuum while it shows high durability in argon and air with various humidity level.

In order to further investigate the durability of the lubricant films, a single crystal Si (100) ball (1 mm in diameter, 5×10^{-7} atoms/cm³ boron doped) was used as a slider on the films to accelerate wear. Tests were performed in argon, dry air, ambient air, and high humidity air. The applied load on the silicon ball was 2.5 g, and the sliding speed was 0.3 m/s. The sliding distance was up to 600 m. The results are shown in Fig. 14.31. From Fig. 14.31, in all environments, Z-TETRAOL exhibits higher durability than the other films. The durability of A20H is slightly lower than that of Z-TETRAOL except in ambient air: both are comparable. Tests on Z-TETRAOL and A20H were terminated at 600 m. Z-DOL + A20H shows higher durability than Z-DOL.

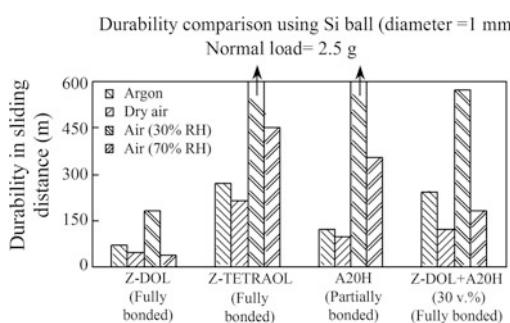


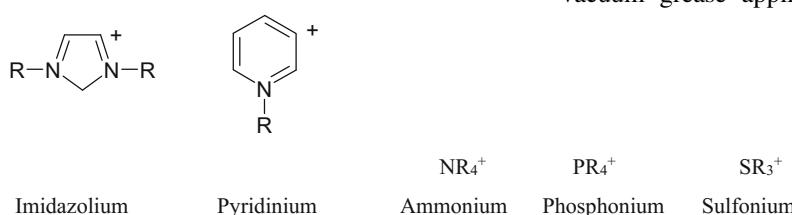
Fig. 14.31 Durability comparison for Z-DOL, Z-TETRAOL, A20H, and Z-DOL + A20H in argon and air with different humidity levels. A single crystal Si (100) ball (1 mm in diameter) was used as a slider on the films (Tao and Bhushan 2005)

14.4 Nanotribological and Electrical Studies of Ionic Liquid Films

An ionic liquid (IL) is a synthetic salt with a melting point below 100 °C. A room temperature ionic liquid is a synthetic molten salt with melting points at or below room temperature. One or both of the ions are an organic species. At least one ion has a delocalized charge such that the

formation of a stable crystal lattice is prevented, and the ions are held together by strong electrostatic forces. As a result of the poor coordination of the ions, these compounds are liquid below 100 °C or even at room temperature (Bhushan et al. 2008a; Palacio and Bhushan 2008a).

The number of combinations of anions and cations that can be used to produce ionic liquids is in the range of one million. Typical cations include imidazolium, pyridinium, ammonium, phosphonium and sulfonium as shown below, where R stands for an organic group.



Typical anions are tetrafluoroborate, BF_4^- , hexafluorophosphate, PF_6^- , bis(trifluorosulfonyl) imide, $(\text{CF}_3\text{SO}_2)_2\text{N}^-$ (“triflamide”) and toluene-4-sulfonate, $\text{C}_7\text{H}_7\text{O}_3\text{S}^-$ (“tosylate”) (Kinzig and Sutor 2005). Dependent upon the substrate wettability and other functional requirements, a set of cations and anions can be combined. The ionic liquids were initially developed for use as electrolytes in batteries and for electrodeposition. Recent applications have geared these compounds as environmentally friendly solvents for chemical synthesis (“green chemistry”) where these liquids are used as substitutes for conventional organic solvents.

Ionic liquids are considered as potential lubricants. Their strong electrostatic bonding, compared to covalently-bonded fluids, leads to very desirable lubrication properties. They also possess desirable properties such as negligible volatility, nonflammability, high thermal stability or high decomposition temperature, efficient heat transfer properties, low melting point, as well as compatibility with lubricant additives. Unlike conventional lubricants that are electrically insulating, ionic liquids can minimize the contact resistance between sliding surfaces because they are conducting, and conducting lubricants are

needed for various electrical applications (e.g., Bhushan and Kwak 2008; Palacio and Bhushan 2008b; Bhushan et al. 2008b). These liquids can also be used to mitigate arcing, which is a cause of electrical breakdown in sliding electrical contacts. In addition, ILs have high thermal conductivity which helps to dissipate heat during sliding. The use of ionic liquids instead of hydrocarbon base oils (such as highly reformed mineral oils) has the potential to dramatically reduce air emissions. Perfluoropolyethers (PFPEs) are used in magnetic rigid disk and vacuum grease applications due to their high

thermal stability and extremely low vapor pressure. However, from the commercial standpoint, ionic liquids are cheaper than PFPEs by a factor of two or so, providing the motivation for comparing the tribological properties of the former with the latter. ILs are being considered for MEMS/NEMS applications because of their high temperature stability, electrical conductivity, and desirable lubrication properties.

Bhushan et al. (2008a) evaluated ionic liquids with the hexafluorophosphate anion deposited on Si(100) wafers, and these were found to exhibit improved friction and wear properties compared to conventional lubricants. The films were evaluated as untreated, partially bonded (by heating at 150 °C for 30 min after dip coating), and fully bonded (thermal treated and washed). The ionic liquid containing the octyl sulfate anion has also been developed and is of interest due to its resistance to hydrolysis. Based on experience, anions are observed to affect tribological performance. Table 14.3 lists the physical and thermal properties of a selected ionic liquid and its properties are compared to the PFPE lubricant Z-TETRAOL.

The durability of ionic liquid films on various metal and ceramic substrates has been investigated

Table 14.3 Physical, thermal and electrical properties of BMIM-PF₆ and Z-TETRAOL

	1-Butyl-3-methylimidazolium hexafluorophosphate (BMIM-PF ₆)	Z-TETRAOL
Cation	C ₈ H ₁₅ N ₂ ⁺	–
Anion	PF ₆ [–]	–
Molecular weight (g/mol)	284 ^a	2300 ^b
T _{melting} (°C)	10 ^c	–
T _{decomposition} (°C)	300 ^c	~320 ^b
Density (g/cm ³)	1.37 ^a	1.75 ^b
Kinematic viscosity (mm ² /s)	281 ^a (20 °C) 78.7 ^d (40 °C)	2000 ^b (20 °C)
Pour point (°C)	<–50 ^e	–67 ^b
Specific heat (J/g K)	1.44 ^f (25 °C)	~0.20 ^b (50 °C)
Thermal conductivity at 25 °C (W/m-K)	0.15 ^g	~0.09 ^b
Dielectric strength at 25 °C (kV/mm)	–	~30 ^b
Volume resistivity (Ω cm)	714 ^h	~10 ^{13b}
Vapor pressure at 20 °C (Torr)	<10 ^{–9}	5 × 10 ^{–7b}
Wettability on Si	moderate ^c	–
Water contact angle	46° (untreated) 39° (partially bonded) 41° (fully bonded)	54° (untreated) 83° (partially bonded) 88° (fully bonded)
Miscibility with isopropanol	Total ^a	–
Miscibility with water	None ^h	–

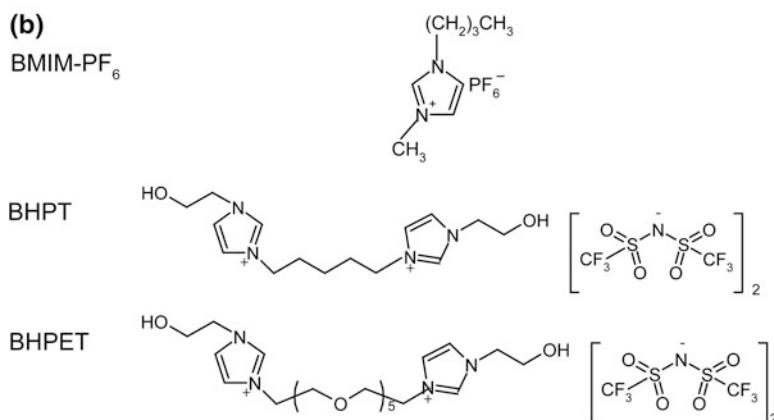
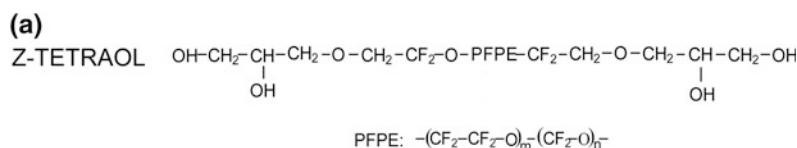
^aMerck Ionic Liquids Database, Darmstadt, Germany, also: <http://ildb.merck.de/ionicliquids/en/startpage.htm>^bZ-TETRAOL Data Sheet, Solvay Solexis Inc., Thorofare, NJ^cKinzig and Sutor (2005)^dReich et al. (2003)^eWang et al. (2004)^fKabo et al. (2004)^gFrez et al. (2006)^hCarda-Broch et al. (2003)

from the standpoint of film formation (wettability) and film removal (friction and wear), where it was found that certain cations and anions exhibit better wetting, friction reduction, and wear resistance properties (Bhushan et al. 2008a). In general, ionic liquids exhibit better wettability on noble metal and ceramic surfaces than on non-noble metal surfaces (Kinzig and Sutor 2005). The flat imidazolium cation shows poorer wettability compared to bulkier cations such as ammonium and sulfonium. Among salts with the imidazolium cation, the presence of longer organic side chains lead to a reduction of the coefficient of friction. An anion effect is also observed, where oxygen-rich anions show better substrate wettability and lower

wear compared to other imidazolium salts. Based on these findings, the ionic liquids of interest are 1-butyl-3-methylimidazolium hexafluorophosphate (BMIM-PF₆) and 1-butyl-3-methylimidazolium octyl sulfate (BMIM-OctSO₄). These were studied by Bhushan et al. (2008a), and the former was found to be superior in tribological performance. The chemical structures of BMIM-PF₆ and Z-TETRAOL for comparison are shown in Fig. 14.32, and a summary of their properties is presented in Table 14.3.

Some dicationic ILs are thermally stable up to 400 °C (Anderson et al. 2005). The adhesion and friction properties of two dicationic IL films on Si(100) substrate, based on the imidazolium

Fig. 14.32 **a** Chemical structures of the Z-TETRAOL molecule, and **b** chemical structures of the BMIM-PF₆, BHPT and BHPET molecules



cation and the bis(trifluoromethanesulfonyl) imide (or “triflamide”) anion, were studied by Palacio and Bhushan (2009a) and compared to the monocationic IL 1-butyl-3-methyl-1*H*-imidazolium hexafluorophosphate (BMIM-PF₆). AFM experiments were also performed at various humidity and temperature conditions in order to investigate the effect of the environment on the nanolubrication properties of these ILs. Micro-scale friction and wear experiments using the ball-on-flat tribometer and wear at ultralow loads using an AFM were carried out. Fourier Transform Infrared (FTIR) spectroscopy and X-ray photoelectron spectroscopy (XPS) were used to determine the chemical species that affect intermolecular bonding and also to elucidate the effect of the environment on the IL film surface in the case of FTIR data.

The dicationic ionic liquids used in this study were 1,1'-(Pentane-1,5-diyl)bis(3-hydroxyethyl-1*H*-imidazolium-1-yl) di[bis(trifluoromethanesulfonyl)imide]] (abbreviated as BHPT) and 1,1'-(3,6,9,12,15-Pentaoxapentadecane-1,15-diyl)bis(3-hydroxyethyl-1*H*-imidazolium-1-yl) di[bis(trifluoromethanesulfonyl)imide]] (abbreviated as BHPET) (Payagala et al. 2007) deposited on Si(100) wafers. The common name “triflamide”

will be used here when referring to the anion. The properties of these dicationic ILs were compared with the conventional monocationic ionic liquid 1-Butyl-3-methyl-1*H*-imidazolium hexafluorophosphate, abbreviated as BMIM-PF₆ (Merck, Germany). Their chemical structures are shown in Fig. 14.32b. These compounds have been applied on single-crystal Si(100) (phosphorus doped) with a native oxide layer on the surface using the dip coating technique. The method and the apparatus used have been described earlier. The films were heat-treated at 150 °C for 30 min after dip coating, such that they were partially bonded (Bhushan et al. 2008a).

14.4.1 Monocationic Liquid Films

In this section, nanotribological data on BMIM-PF₆ and Z-TETRAOL are presented (Bhushan et al. 2008a; Palacio and Bhushan 2008a, b, 2009b).

14.4.1.1 Nanotribological Studies

In Fig. 14.33a, the surface height images for the untreated sample (air dried) are compared with

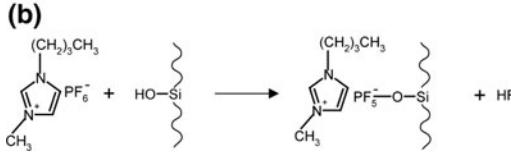
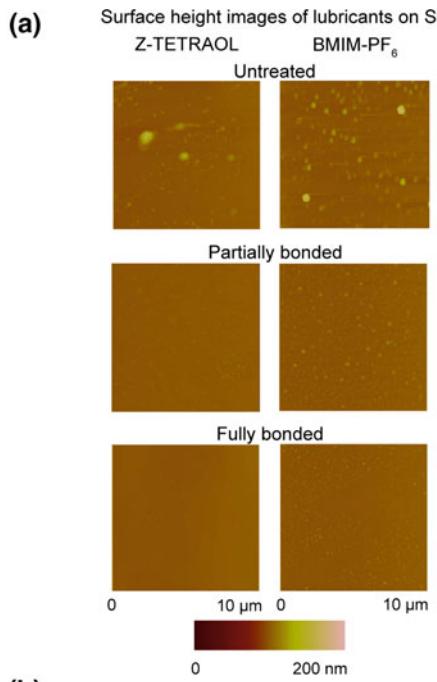


Fig. 14.33 **a** Surface height images for the untreated, partially bonded and fully bonded films of BMIM-PF₆ on silicon substrate, and **b** schematic for the attachment of BMIM-PF₆ to the silicon substrate (Bhushan et al. 2008a)

the two chemically-bonded samples (partially bonded and fully bonded). Aggregates of varying sizes are observed on the untreated Z-TETRAOL and on the ionic liquid coating. These aggregates could have formed initially during preparation of the dilute solution. For Z-TETRAOL, the long PFPE chains can orient in various configurations (such as coils), leading to aggregate formation (Bhushan et al. 2008a). In ionic liquids containing the 1-butyl-3-methylimidazolium cation, it is believed that aggregate formation in solvents that are less polar compared to water aids in minimizing the charge density (charge delocalization) within the ions. These aggregates are subsequently deposited on the silicon surface during

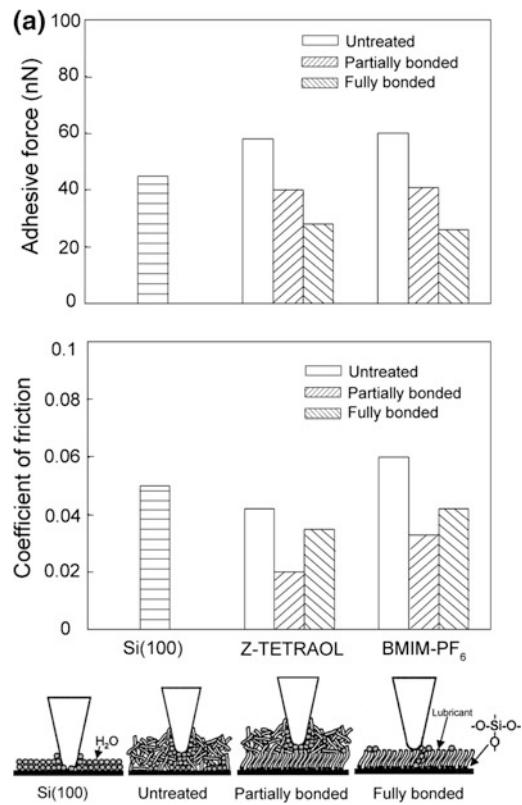


Fig. 14.34 **a** Summary of the adhesive force and coefficient of friction, and **b** durability data after 100 cycles for BMIM-PF₆ at room temperature (22 °C) and ambient air (45–55% RH). Data for the uncoated Si and Z-TETRAOL are shown for comparison. Schematic in (a) shows the effect of chemical bonding treatment and meniscus formation between the AFM tip and sample surface on the adhesive and friction forces (Bhushan et al. 2008a)

the dip coating procedure. It is observed that the untreated lubricant surface has more prominent aggregates compared to the two chemically-bonded samples, implying that the heat treatment promotes bonding to the Si substrate. Without the chemical bonding procedure, the lubricant molecules are less likely to attach to the substrate and would tend to attract each other instead, such that dewetting is more likely. The immobilization of the ionic liquid, which is promoted by thermal treatment, occurs by reaction of the anion with the hydroxyl groups

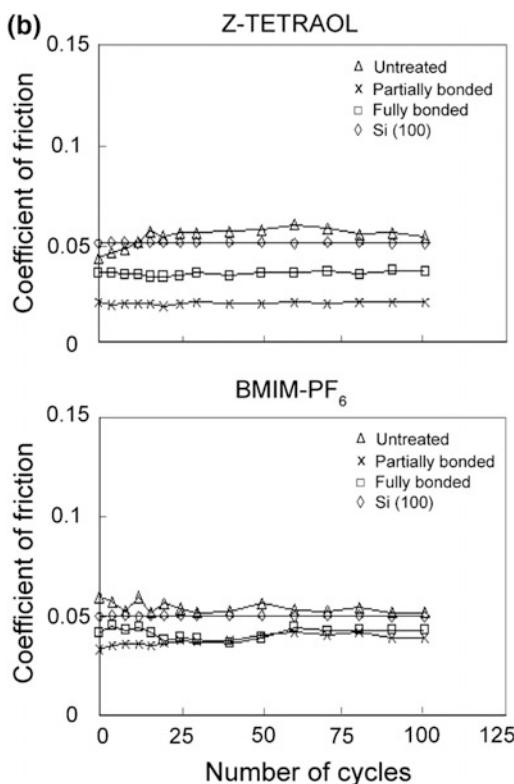


Fig. 14.34 (continued)

present on the silicon surface, as shown in Fig. 14.33b for BMIM-PF₆.

Figure 14.34a is a summary of the adhesive force and coefficient of friction measurements on the ionic liquid. Z-TETRAOL and Si (100) data are provided for comparison. The adhesive force has been observed to decrease in the following order: untreated > partially bonded > fully bonded. This mobile fraction on the untreated sample facilitates the formation of a meniscus, which increases the tip-sample adhesion. The adhesive force is highest in the untreated coating since it has the greatest amount of the mobile fraction among the three samples. Conversely, the sample with no mobile lubricant fraction available (fully bonded) has the lowest adhesive force.

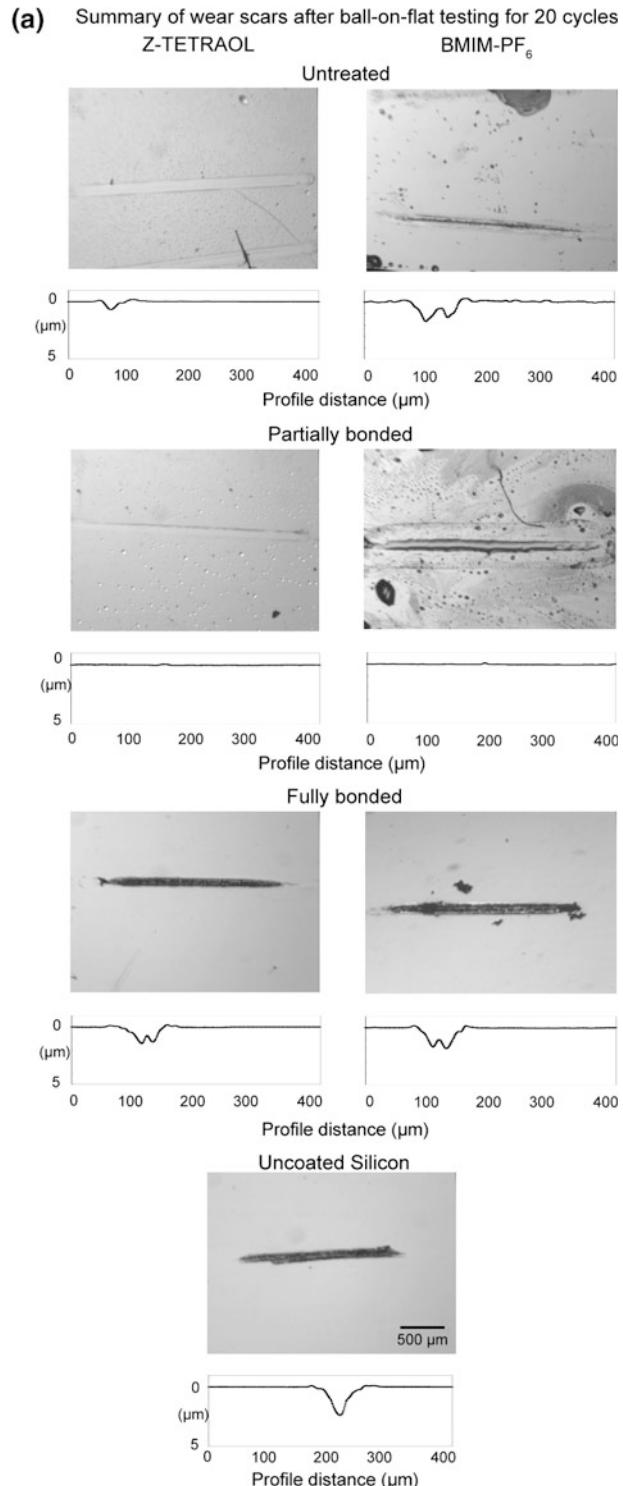
A different trend is observed in the coefficient of friction (μ) data. Both the fully bonded and partially bonded samples have lower μ values compared to the uncoated silicon. Friction forces are lower on the latter, implying that the mobile

lubricant fraction present in the partially bonded samples facilitates sliding of the tip on the surface. However, μ values for the untreated samples are higher than the data for the heat treated coatings. Due to the lack of chemical bonding, the interaction of the lubricant to the substrate is weakened and dewetting can occur. Water and lubricant molecules are more likely to form a meniscus as the tip approaches the surface. This provides greater resistance to tip sliding, leading to higher coefficient of friction values. The lower portion of Fig. 14.34a is a schematic illustrating the role of meniscus formation in the adhesive and friction forces obtained for the uncoated Si and the untreated, partially bonded and fully bonded lubricant-coated Si surfaces.

Figure 14.34b contains plots of the coefficient of friction as a function of the number of sliding cycles at 70 nN normal load. Only a small rise in the coefficient of friction was observed for both Z-TETRAOL and the BMIM-PF₆ surfaces, indicating low surface wear. In the case of untreated Z-TETRAOL, a crossover is observed, where the coefficient of friction increases from its initial value and exceeds the μ of silicon after a certain number of cycles. This is attributed to the transfer of lubricant molecules to the AFM tip and the interaction of the transferred molecules with the lubricant still attached on the Si substrate, which will increase the friction force.

In order to compare friction and wear properties on the nanoscale with that on the microscale, conventional ball-on-flat tribometer experiments were conducted on the same samples. Images and profile traces of the wear scars are shown in Fig. 14.35a. The coefficient of friction and number of cycles to failure are summarized in Fig. 14.35b. Ionic liquid shows enhanced durability compared to both the Z-TETRAOL-coated and the uncoated Si. The nanoscale data presented in Fig. 14.34 can be compared to μ values obtained from ball-on-flat (Fig. 14.35b). The μ values of the untreated lubricant samples obtained by using AFM are lower than the μ obtained from the ball-on-flat tests. This is attributed to the difference in the length scales of the test techniques. An AFM tip

Fig. 14.35 **a** Optical images and height profiles taken after 20 cycles, and **b** summary of the coefficient of friction and number of cycles to failure from ball-on-flat tests on various BMIM-PF₆ coating. Data for the uncoated Si and Z-TETRAOL are shown for comparison (Bhushan et al. 2008a)



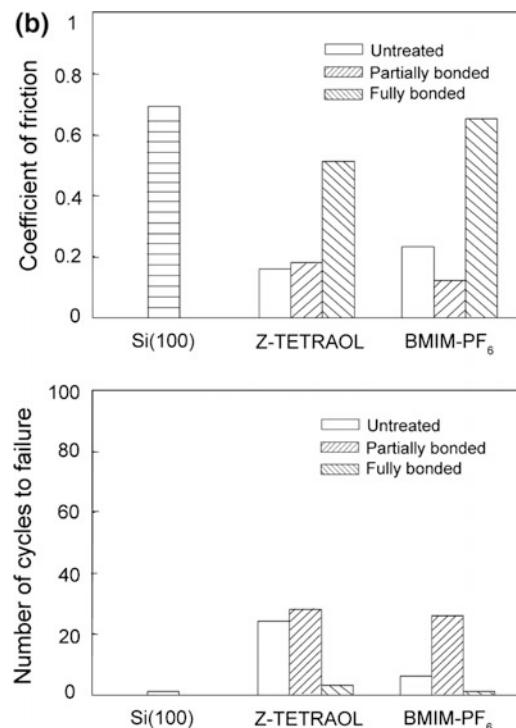


Fig. 14.35 (continued)

simulates a single asperity contact while the conventional friction test involves the contact of multiple asperities present in the test system (Bhushan 2013a, b). With regards to wear, the interface contact of the AFM and ball-on-flat techniques are different from each other. In an AFM, the contact stress is very high such that material can be displaced more easily. For a ball-on-flat test, the ball exerts a lower pressure on the surface, and the coating is in a confined geometry. As a consequence, displacement of the coating is not as easy as in AFM, leading to enhanced wear resistance.

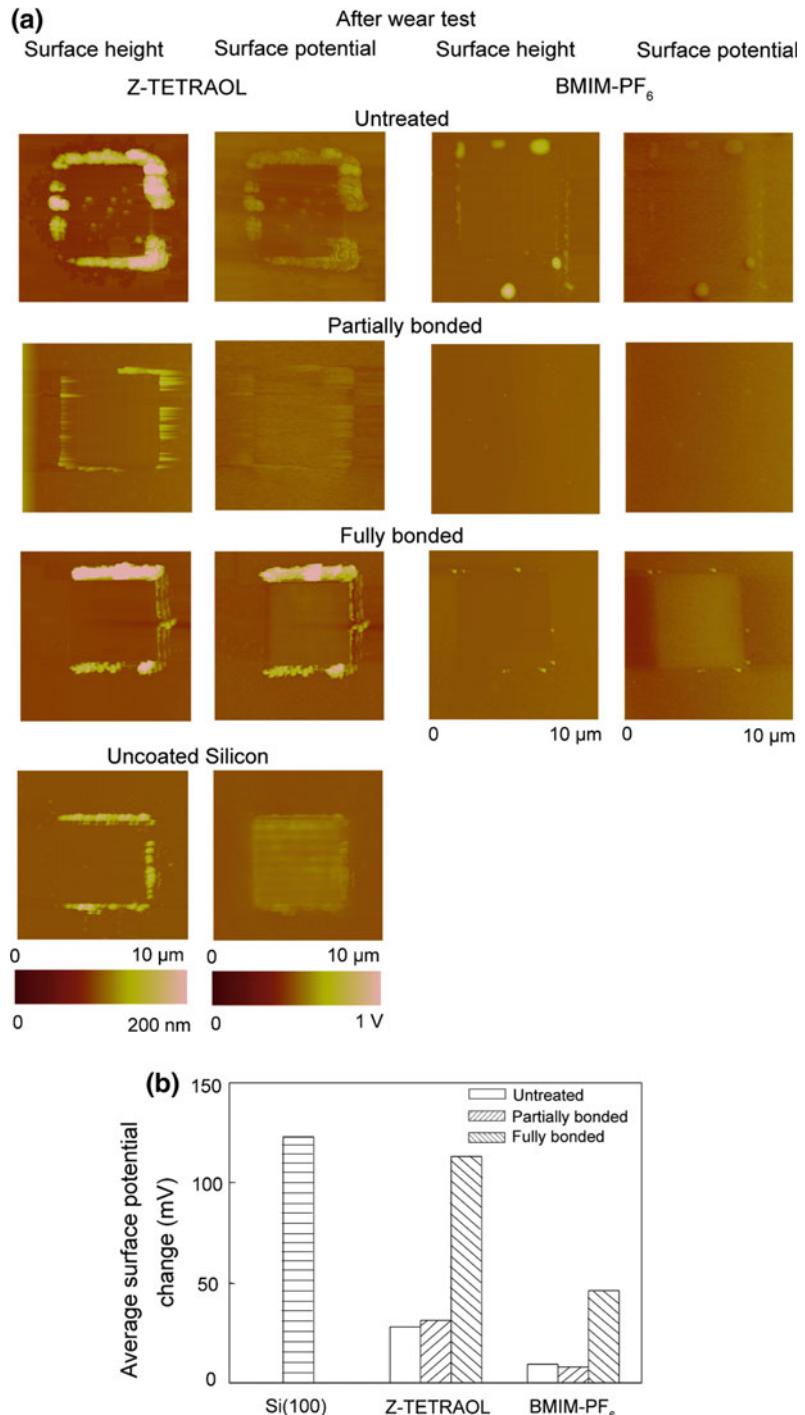
14.4.1.2 Wear Detection by Surface Potential and Electrical Resistance Measurements

Figure 14.36a is a summary of the wear tests and the corresponding surface potential measurements on the ionic liquid. A bar plot summarizing the average surface potential change on the tested area is shown in Fig. 14.36b. In all cases, a

smaller amount of debris was generated compared to the uncoated silicon surface, indicating that the ionic liquid provides wear protection. In general, the samples containing the mobile lubricant fraction (i.e., untreated and partially bonded surfaces) exhibit a lower surface potential change compared to the fully bonded sample, which only has immobile lubricant molecules. This is attributed to lubricant replenishment by the mobile fractions, which can occur in the untreated and partially bonded samples (Tao and Bhushan 2005; Palacio and Bhushan 2007a). From the bar plot in Fig. 14.36b, it is also observed that the change in surface potential is generally lower in the ionic liquid coating compared to the Z-TETRAOL coating and the uncoated silicon. This indicates that any built-up surface charges arising from the wear test were immediately dissipated onto the conducting ionic lubricant coating surface. In the case of Z-TETRAOL and the uncoated silicon, the charges remained trapped in the test area, since both these materials are insulators. Based on these findings, a considerable surface potential change will be observed on the wear region when: (1) the lubricant has been fully removed from the substrate; (2) the native SiO₂ layer has been abraded from the surface; (3) wear has caused subsurface structural changes; and (4) charges build up as they are unable to dissipate into the surrounding material.

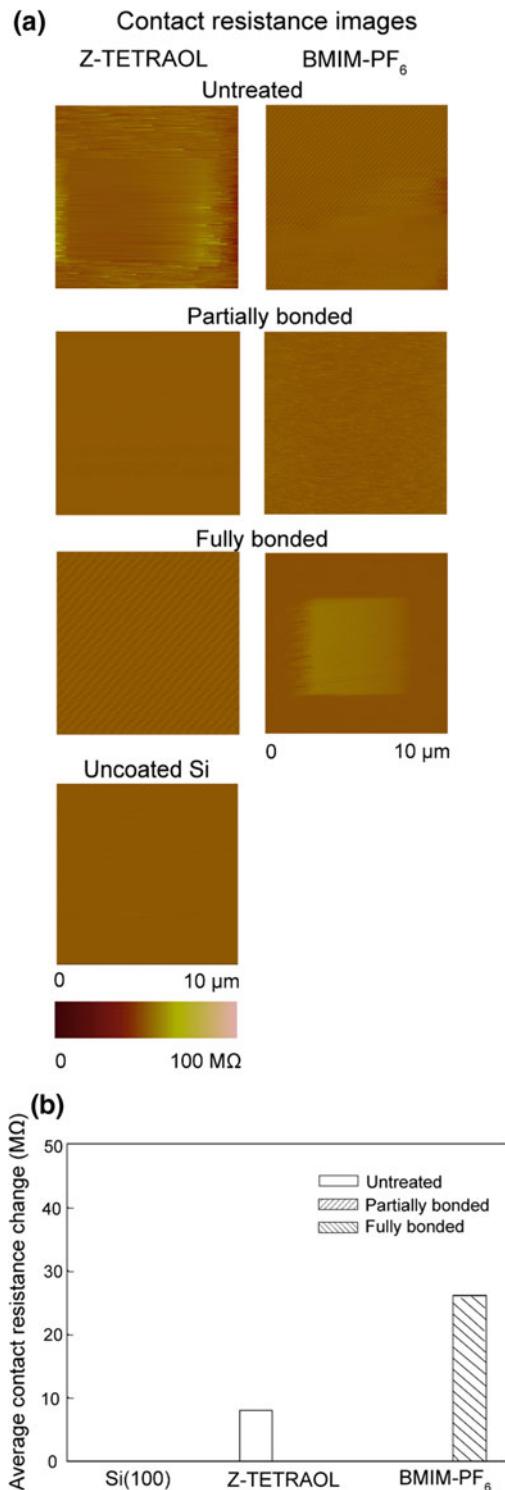
Contact resistance images for the surfaces subjected to the wear tests are presented in Fig. 14.37a. The average change in the contact resistance of the wear region relative to the untested area is summarized in Fig. 14.37b. The fully bonded BMIM-PF₆ has an appreciable contact resistance increase in the wear region. Since silicon is a semiconductor, it has much higher resistance compared to the surrounding ionic liquid. The resistance increase in the worn area implies that the substrate is exposed after the wear test. Partially bonded films did not get worn out from the substrate after the test, as evidenced by the lack of contact resistance change in the tested area. The untreated Z-TETRAOL exhibited an observable resistance change, while BMIM-PF₆ did not. This can be correlated to the

Fig. 14.36 **a** Surface height and surface potential maps after wear tests, and **b** bar chart showing surface potential change for various BMIM-PF₆ coatings. Data for the uncoated Si and Z-TETRAOL are shown for comparison (Bhushan et al. 2008a)



durability data in Fig. 14.34b, where the untreated Z-TETRAOL sample exhibited an increase in the friction force with time due to the transfer of lubricant molecules to the tip. Easier

lubricant removal means that the diamond tip (in the case of Fig. 14.37) can cause substrate wear much sooner, leading to the observed resistance increase in the tested area. However, the



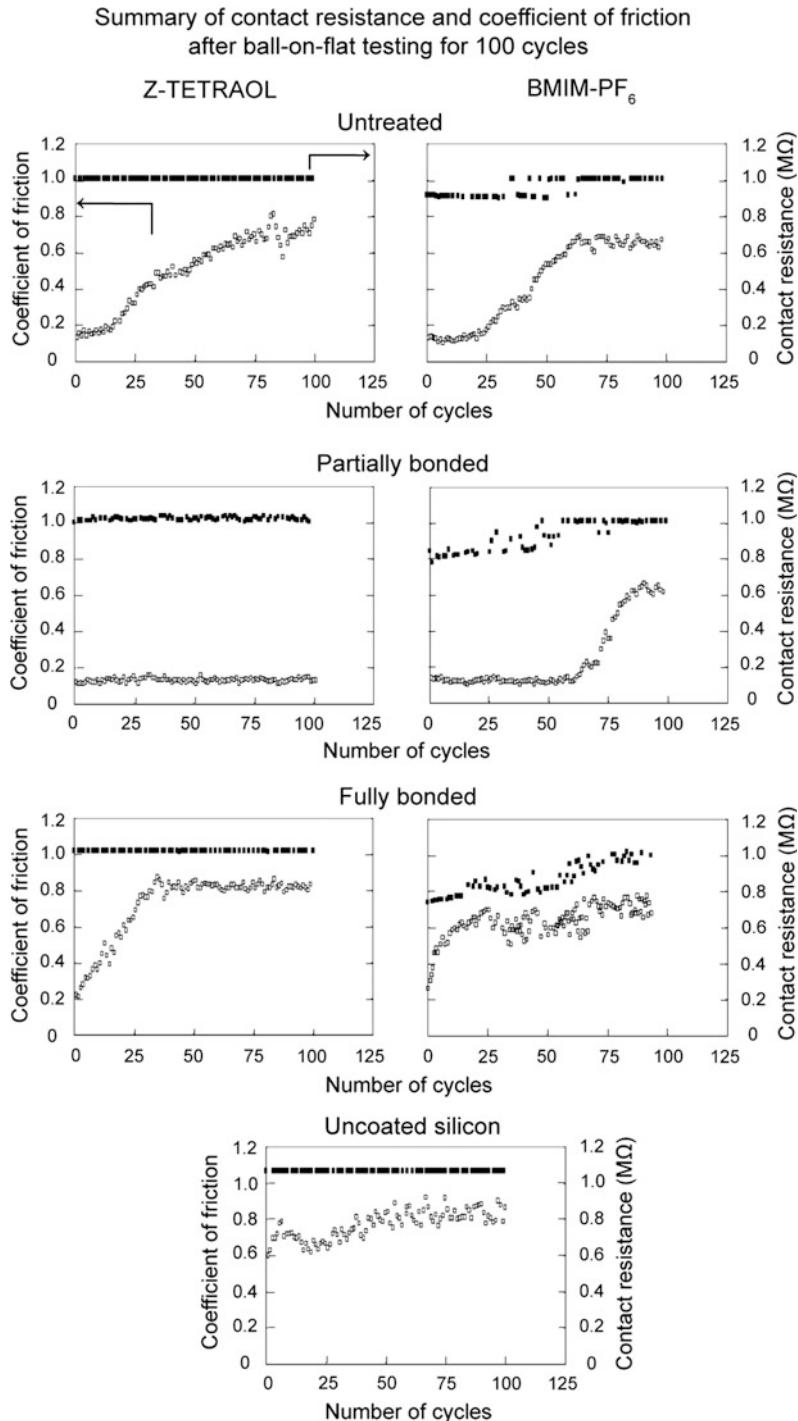
resistance image does not provide a clear contrast between Z-TETRAOL and the newly exposed substrate since both materials have high resistance values.

Microscale contact resistance obtained from ball-on-flat tribometer testing is shown in Fig. 14.38, along with the corresponding coefficient of friction data. For the ionic liquid film, the initial resistance is slightly lower than that of uncoated silicon, confirming their conductive nature. For the Z-TETRAOL samples, the contact resistance is about the same magnitude as the uncoated silicon. But for the conducting ionic liquid, an increase in resistance corresponds to an increase in the coefficient of friction, indicating wear of the lubricant and exposure of the silicon substrate, similar to observations in the nanoscale. These results are consistent with the adhesion, friction, and surface potential results with regards to wear detection and wear protection coming from the mobile and immobile lubricant fractions.

The durability data and trends for the ionic liquid obtained by using a steel ball are inferior to the results in Fig. 14.35, which was measured by using a sapphire ball. In Fig. 14.38, the partially bonded samples still show the best durability, but in this case, Z-TETRAOL has the highest number of cycles to failure (opposite trend compared to Fig. 14.35b), as indicated by the point where the jump in the coefficient of friction is observed. This can be accounted for from the wetting properties of ionic liquids on different surfaces. It has been observed that ionic liquids have a tendency to wet nonmetal surfaces (e.g. Si_3N_4 , SiO_2 , glass) better than conventional metal surfaces (such as 440C, M50 and 52100 steel) (Bhushan et al. 2008a). For wear tests with a steel ball, less wettability means less lubricant retention at the

Fig. 14.37 a Nanoscale contact resistance images after wear tests and b bar chart showing contact resistance change for various BMIM-PF_6 coatings. Data for the uncoated Si and Z-TETRAOL are shown for comparison. Cases without the clear wear scar did not exhibit measurable change (Bhushan et al. 2008a)

Fig. 14.38 Microscale contact resistance and coefficient of friction after ball-on-flat tests for 100 cycles on various BMIM-PF₆ coatings. Data for the uncoated Si and Z-TETRAOL are shown for comparison (Bhushan et al. 2008a)



interface. This material wetting effect is possibly more significant for the ionic liquid than in Z-TETRAOL, but nonetheless, the durability of

the partially bonded BMIM-PF₆ is still close to its Z-TETRAOL counterpart, such that ILs are still viable lubricants comparable to PFPEs.

14.4.2 Dicationic Ionic Liquid Films

In this section, nanotribological data on BHPT and BHPET are presented. Data on BMIM-PF₆ and uncoated Si(100) are also presented for comparisons (Palacio and Bhushan 2009a).

14.4.2.1 Nanotribological Studies

Figure 14.39a is a summary of the contact angle, adhesive force, and coefficient of friction (μ) measurements for the coated and uncoated samples performed at ambient temperature and humidity conditions (22 °C and 50% RH, respectively). The data shown in the bar plots are the averages of three measurements and the error bars represent $\pm 1\sigma$. BHPT is the least hydrophilic as it has the highest contact angle (81°) among the three IL coatings. This coating also exhibited the greatest reduction in the coefficient of friction relative to the uncoated surface. The high contact angle of BHPT leads to minimal meniscus formation between the tip and surface, leading to a large drop in the nanoscale friction force. In addition, BHPT has a pentyl chain which links the two imidazolium cations. This chain can orient the cation molecules on the substrate, thereby facilitating tip sliding on the film surface. In contrast, the polyether chain of BHPET is susceptible to interactions with water molecules which can promote (instead of minimize) meniscus formation. The BMIM-PF₆ film also exhibited a reduction in the adhesive force and coefficient of friction. This comes from the combination of mobile and immobile lubricant fractions. Immobilization of this ionic liquid is possible as a result of the thermal treatment, which promotes the reaction between the hexafluorophosphate anion with the hydroxyl groups present on the silicon substrate surface (Valkenberg et al. 2002; Bhushan et al. 2008a; Palacio and Bhushan 2009a).

Wear tests were conducted by monitoring the change in the friction force on a 2 μm line for 100 cycles. The data shown in Fig. 14.39b is a representative out of three measurements made for each sample. The focus of this experiment is the wear of the lubricant film on the substrate. The μ value of the BHPT film changed

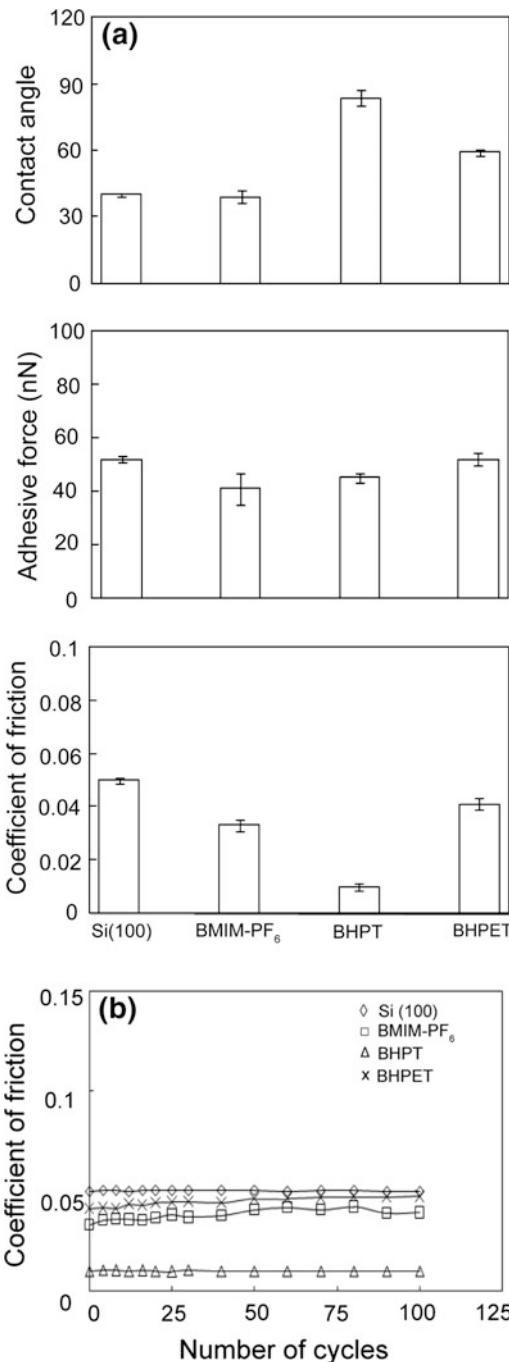


Fig. 14.39 **a** Summary of the contact angle, adhesive force and coefficient of friction, and **b** durability data after 100 cycles for BMIM-PF₆, BHPT and BHPET coatings at room temperature (22 °C) and ambient air (45–55% RH). Data for uncoated Si are shown for comparison. The error bars in **a** represent $\pm 1\sigma$ based on three measurements performed (Palacio and Bhushan 2009a)

minimally during the duration of the experiment, which indicates that the film was not being worn after 100 cycles. On the other hand, the BMIM-PF₆ and BHPT samples exhibited a gradual increase in μ , which means that these films could be undergoing some wear and that their interaction with the silicon substrate is weaker compared to that of BHPET.

Figure 14.40 is a summary of the average contact potential change after the wear tests conducted by creating $5 \times 5 \mu\text{m}^2$ wear scars with a diamond tip, where the height and surface potential maps were imaged afterwards. The data shown in the bar plot is an average of three measurements and the error bars represent $\pm 1\sigma$. A change in the surface potential in the wear region is observed when the following occur: the lubricant has been fully removed from the substrate, the native SiO₂ layer has been abraded from the surface, wear has caused subsurface structural changes, and charges build up as they are unable to dissipate into the surrounding material (Bhushan et al. 2008a; Palacio and Bhushan 2009a). As expected, the uncoated Si exhibited the greatest amount of wear (as evidenced by debris build-up around the edge of the wear test region) and highest increase in the surface potential. The surface potential image for BHPET film also showed an increase, indicating that the film was worn out after the test. This was not seen on tests with the BMIM-PF₆ and BHPT samples. The surface potential change could be absent in the test area if the lubricant was not

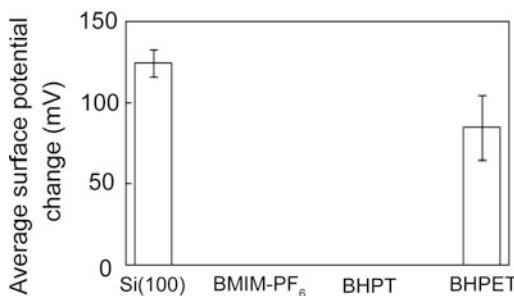


Fig. 14.40 Bar chart showing average surface potential change for BMIM-PF₆, BHPT and BHPET coatings. Data for uncoated Si are shown for comparison. The error bars represent $\pm 1\sigma$ based on three measurements performed (Palacio and Bhushan 2009a)

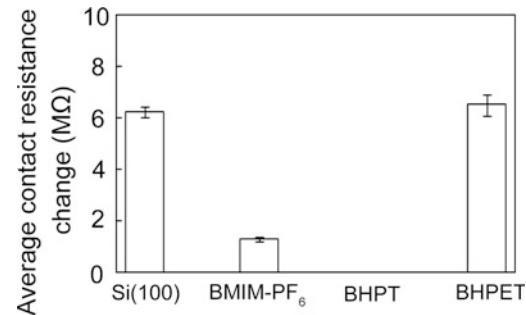


Fig. 14.41 Bar chart showing average contact resistance change for BMIM-PF₆, BHPT and BHPET coatings. Data for uncoated Si are shown for comparison. The error bars represent $\pm 1\sigma$ based on three measurements performed (Palacio and Bhushan 2009a)

removed completely, indicating that these two samples have a stronger interaction with the silicon substrate compared to BHPET.

Figure 14.41 shows a summary of the average contact resistance change after the wear test. As shown in Fig. 14.40, a significant change is observed in the wear region of the Si and BHPET samples. There is also a small amount of localized contact resistance increase in the BMIM-PF₆ sample. These results are consistent with the wear test presented in Fig. 14.39b, where the films are in this order of decreasing durability: BHPT > BMIM-PF₆ > BHPET.

14.4.2.2 Microscale Friction and Wear

In order to compare friction and wear properties at the microscale and the nanoscale, conventional ball-on-flat tribometer experiments were conducted on the same samples. The coefficient of friction data are summarized in Fig. 14.42. The data shown in the bar plot is an average of three measurements. All of the lubricated samples are reported to have less wear scars as a result of the ball having to displace the lubricant before damaging the silicon surface. A reduction in the coefficient of friction arising from application of the lubricant film is observed, which are consistent with the nanoscale adhesion, friction and wear results.

The μ values of the lubricant samples obtained by using AFM are lower than the μ obtained from the ball-on-flat tests. This is

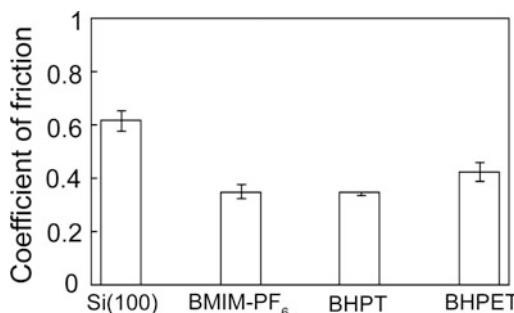


Fig. 14.42 Summary of the coefficient of friction from ball-on-flat tests on uncoated and coated Si samples. The error bars represent $\pm 1\sigma$ based on three measurements performed (Palacio and Bhushan 2009a)

attributed to the difference in the length scales of the test techniques. An AFM tip simulates a single asperity contact while the conventional friction test involves the contact of multiple asperities present in the test system. With regards to wear, the interface contact of the AFM and ball-on-flat techniques are different from each other such that one cannot expect both tests to show the same trend. On an AFM, the tip stress is very high such that material can be displaced more easily. For a ball-on-flat test, the tip exerts a lower pressure on the surface and the coating is in a confined geometry.

14.4.2.3 Relative Humidity and Temperature Effect Measurements

The influence of relative humidity on adhesion and friction is summarized in Fig. 14.43. In general, the adhesive force increases with the relative humidity. The condensed water in the humid environment facilitates meniscus formation between the tip and sample and higher adhesive forces. Since Si(100) is hydrophilic, it readily adsorbs water molecules. For the three ionic liquids, an increase in the adhesive force is also due to increased water adsorption. This comes from attractive electrostatic interactions (ion-dipole forces) between the individual ions and water molecules.

Water adsorption affects the observed coefficient of friction as a function of the relative humidity. In Si, the coefficient of friction is

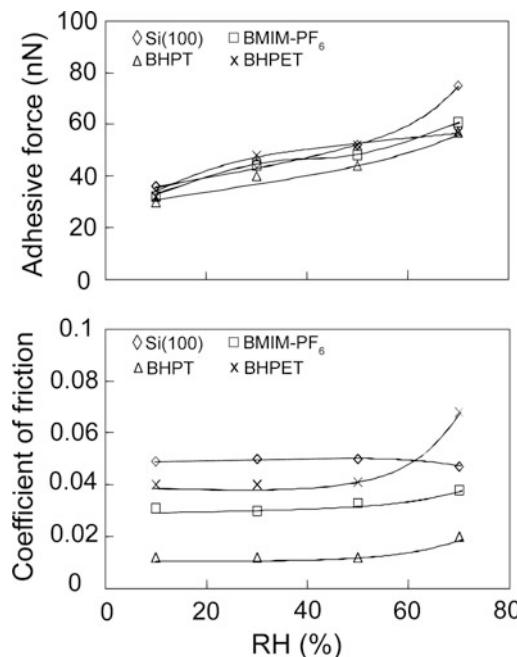


Fig. 14.43 Influence of the relative humidity (RH) on the adhesive force and coefficient of friction for unlubricated and lubricated tapes at 22 °C (Palacio and Bhushan 2009a)

uniform at 10–50% RH, then decreases at 70% RH. The adsorbed water at higher humidity can lead to the formation of a continuous water layer separating the tip and sample surface, which can act as a lubricant. Although the adhesive force increases, the reduction in interfacial strength accounts for the slight decrease of the coefficient of friction at the highest range of humidity level examined. However, the presence of more water molecules at higher humidity has an opposite effect on the ionic liquid surfaces, where the coefficient of friction increases with humidity. The attractive ion-dipole forces between the ions and water are amplified at higher humidity because more water molecules are available. A greater attractive force between tip and surface leads to greater resistance to sliding and a higher coefficient of friction. This is observed on both the monocationic (BMIM-PF₆) and the dicationic (BHPT and BHPET) ionic liquid-coated surfaces. For BHPET, polar interactions between water and the oxygen atoms in the polyether (C–O–C) chain is possible and increases the water

adsorption to the surface. This could account for the larger rise in the coefficient of friction in the BHPET sample from 50 to 70 % RH, compared to BMIM-PF₆ and BHPT.

The effect of temperature on the adhesion and friction properties of the ILs is summarized in Fig. 14.44. The adhesive and friction forces were measured from 22 to 125 °C. As shown in Fig. 14.44, the increase in test temperature leads to a decrease in the adhesive force and the coefficient of friction. The decrease in the adhesive force at higher temperatures is observed in all the samples, while the corresponding drop in the coefficient of friction is seen only for BMIM-PF₆, BHPET and the silicon substrate. At higher temperatures, the surface water molecules are desorbed, leading to the decrease in both the adhesive and friction forces. A reduction in the viscosity at higher temperatures can also facilitate the decrease in the friction force (Liu and Bhushan 2003). In BHPT, the coefficient of

friction was not adversely affected as the test temperature was increased. This implies that at ambient humidity conditions, the BHPT film does not adsorb a large amount of water molecules. Moreover, this also implies that the BHPT surface has weak interactions with surface water molecules, such that friction force during sliding is not very much affected.

14.4.2.4 Fourier Transform Infrared Spectroscopy

Figure 14.45 is a summary of the FTIR spectra obtained for the different ionic liquid coated samples, along with the uncoated Si substrate. The observed peaks are labeled with the chemical bonds which they correspond to. In the case of the Si substrate (with a contact angle of 40°) exposed to water molecules in the ambient, no peaks are observed. C–H stretching vibrations in the coated samples are observed at the 600–800 cm⁻¹ range. For BHPET, the strong peak at ~1060 cm⁻¹ is for the C–O–C vibration, which is prominent due to the presence of the polyether chain in its cation. This peak overlaps with the C–O vibration, which is present in BHPT as the terminal primary alcohol (C–OH). In BMIM-PF₆, a peak appears in this range due to rocking vibrations of the methyl (CH₃) substituent. The peak at 1500–1600 cm⁻¹ comes from the C=N vibrations, which is common to all three ionic liquids since they are all based on the imidazolium cation. However, it is not observed in BHPT as the C=N vibrations may be weak. The wide peak at 3600–4000 cm⁻¹ corresponds to hydrogen bonding, possibly due to water molecules adsorbed on the surface (Morrison and Boyd 1992; Zhang et al. 2007; Palacio and Bhushan 2009a). This is present in BMIM-PF₆ and in BHPET but not in BHPT. This accounts for the much lower contact angle of BMIM-PF₆ and BHPET (39° and 59°, respectively) compared to BHPT (81°). This implies that the surface of BHPT is more hydrophobic compared to either BMIM-PF₆ or BHPET, which is consistent with the observed low coefficient of friction for BHPT (Palacio and Bhushan 2009a).

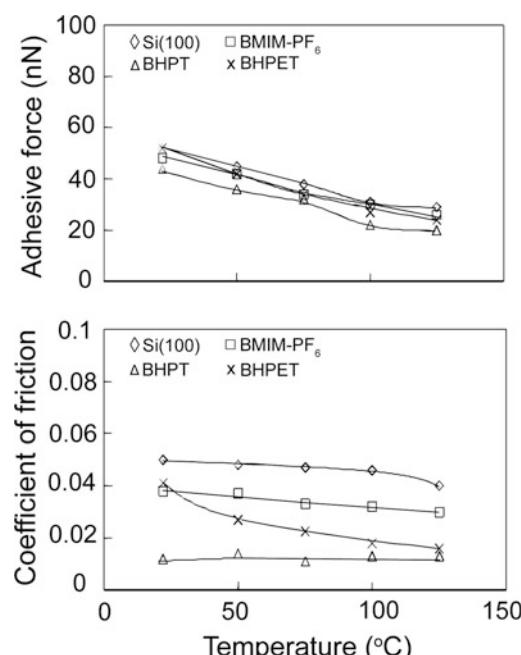
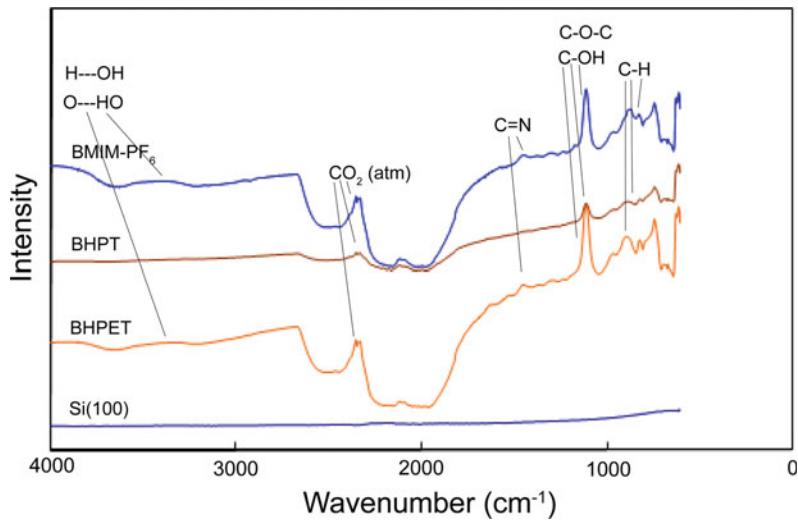


Fig. 14.44 Influence of temperature on the adhesive force and coefficient of friction for uncoated and coated Si samples at 50% RH air (Palacio and Bhushan 2009a)

Fig. 14.45 FTIR spectra of uncoated and coated Si samples. The chemical bonds (or species) were listed above the spectra to indicate the possible bonding modes that correspond to the observed peaks (Palacio and Bhushan 2009a)



14.4.2.5 X-Ray Photoelectron Spectroscopy

The XPS survey spectra obtained on the uncoated and coated Si samples are shown in

Fig. 14.46a. In the uncoated sample, prominent peaks are observed at 99, 151 and 533 eV, corresponding to the binding energies of Si 2p, Si 2s and O 1s electrons, respectively. Additional

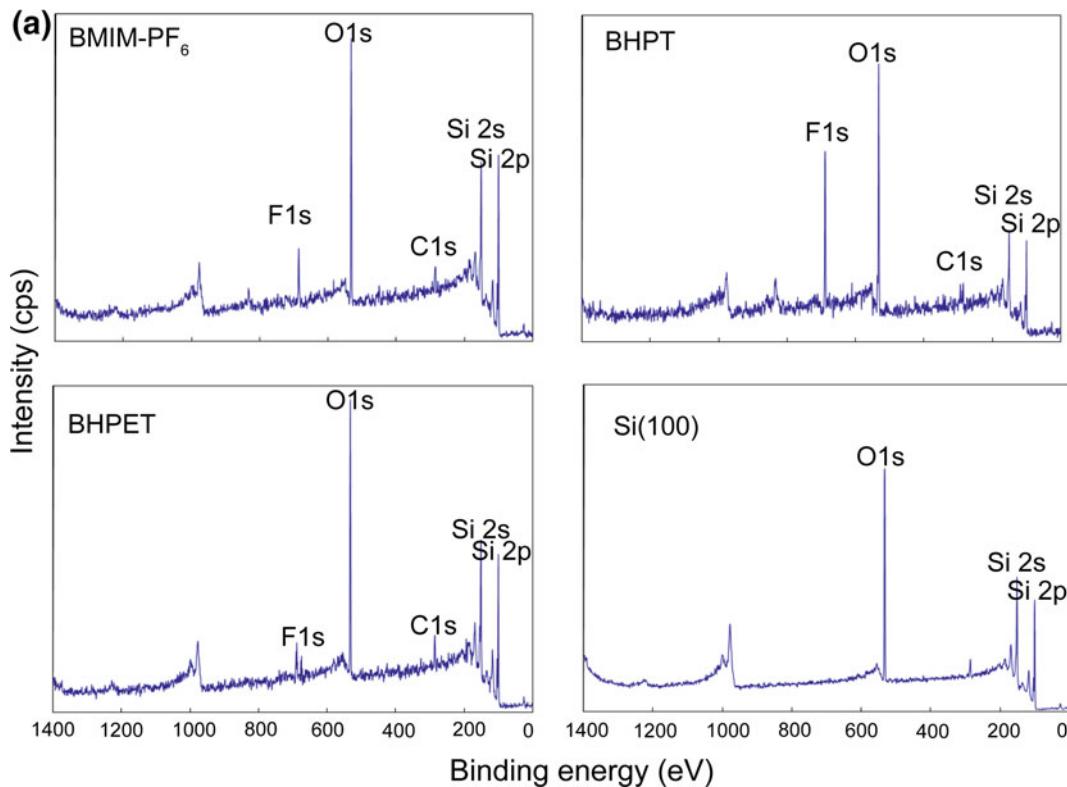


Fig. 14.46 a XPS spectra of uncoated and coated Si samples. Survey scan from 0 to 1400 eV provides the surface elemental composition. b High-resolution (deconvoluted) XPS spectra for Si 2p, C 1s, O 1s and F 1s reveals the different binding environments present on the surface (Palacio and Bhushan 2009a)

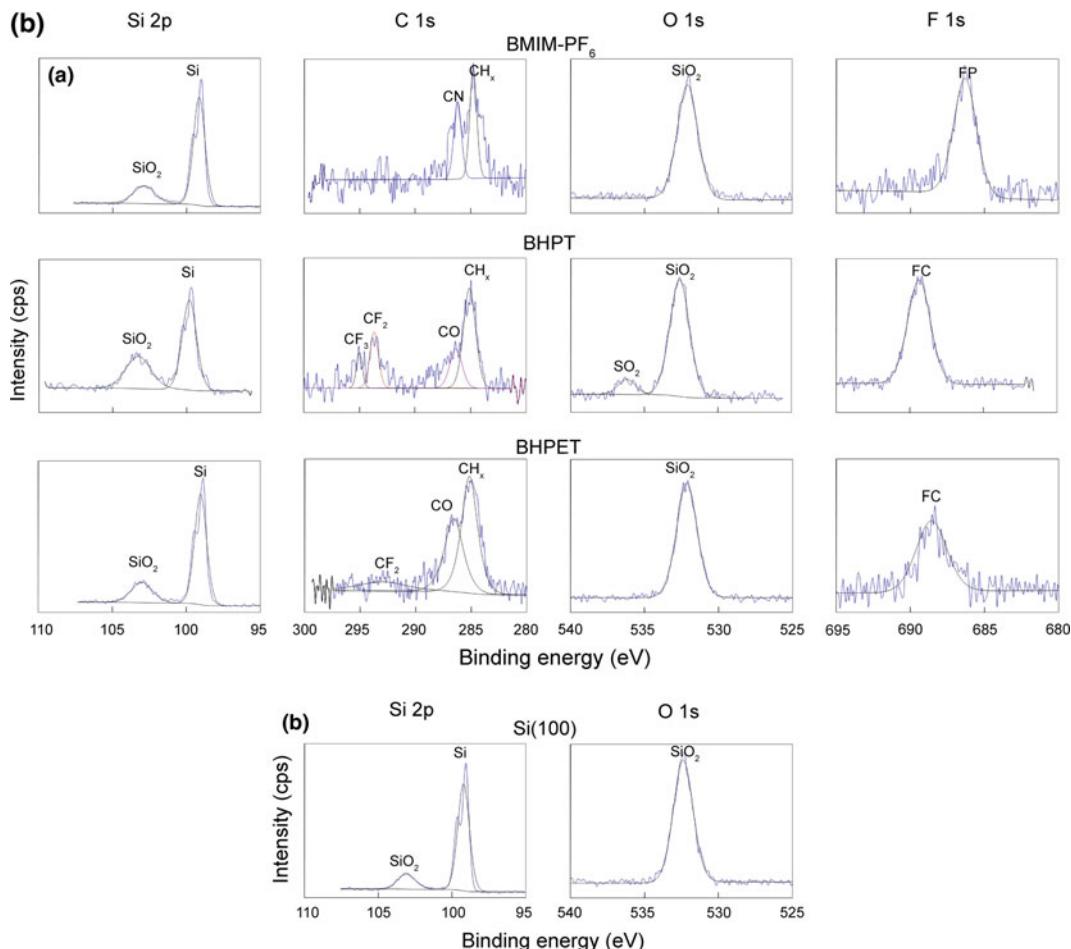


Fig. 14.46 (continued)

peaks are observed on the three coated samples at 285 and 689 eV, which correspond to C 1s and F 1s electrons, respectively. Smaller peaks observed at approximately 1000 eV on all samples are due to Auger lines (KLL transitions) for oxygen (Feldman and Mayer 1986).

The high resolution, best-fit XPS spectra is shown in Fig. 14.46b for the Si 2p, C 1s, O 1s and F 1s electrons. The peaks are labeled with the corresponding chemical bonds, which pertain to either the silicon substrate or groups found on the ionic liquid molecule. One noteworthy exception is the presence peaks at approximately 292 eV, which confirms the presence of CF₂ on the surface. This indicates the immobilization of the BHPT and BHPET ionic liquids, which

occurs by the reaction of the anion with the hydroxyl groups present on the silicon surface (Bhushan et al. 2008a; Palacio and Bhushan 2009a).

14.4.2.6 Relationship Between IL Structure and Adhesion, Friction and Wear Properties

Figure 14.47 is an interpretation of how the IL cations interact with the silicon substrate. For the monocationic BMIM-PF₆, only weak interactions between the imidazolium ring and the silicon surface is expected. For the dicationic ionic liquids, multiple cation attachment schemes are possible. In BHPT, the hydroxyl groups attached

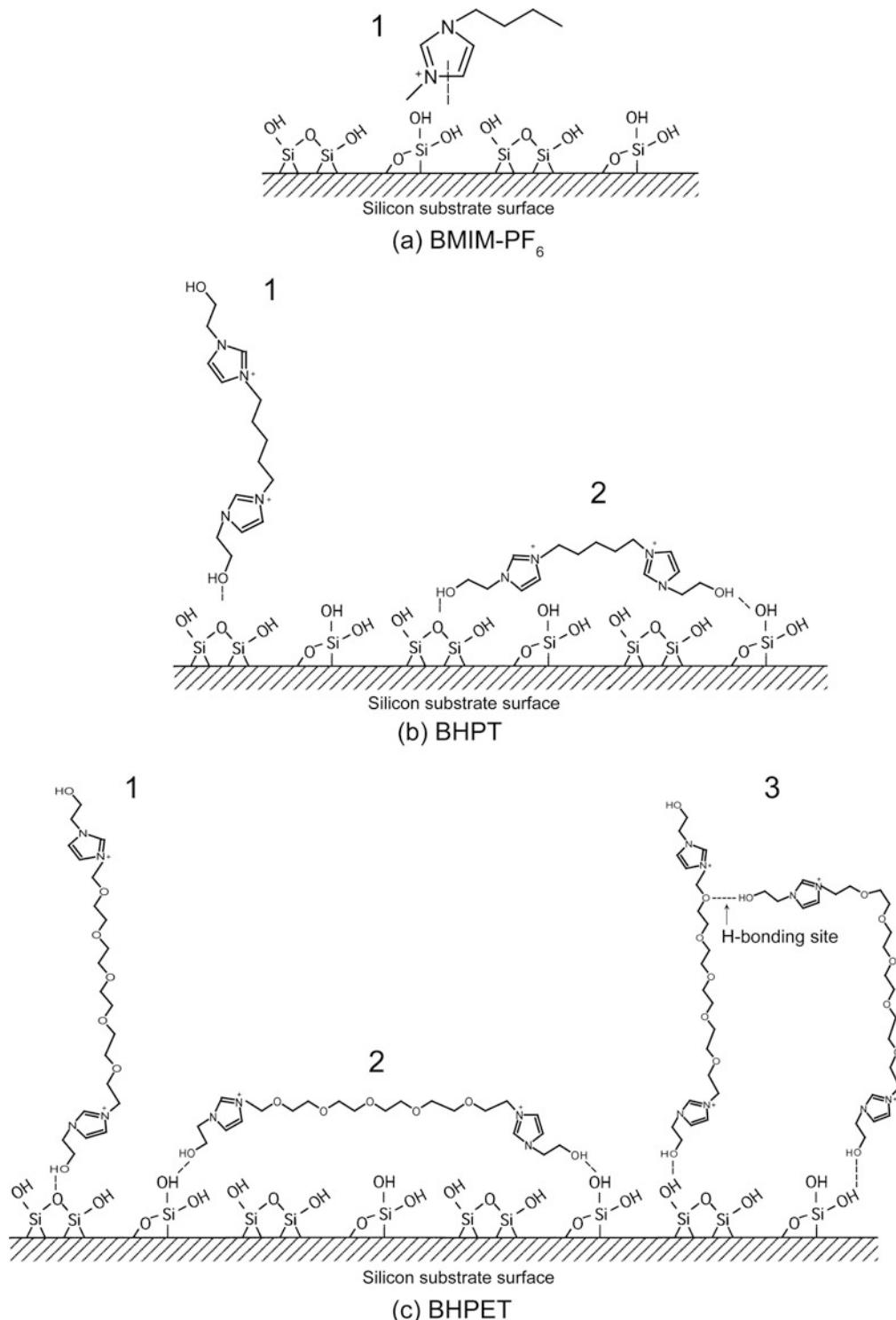


Fig. 14.47 Schematic for the attachment of the cations of **a** BMIM-PF₆, **b** BHPT and **c** BHPET to the silicon substrate (Palacio and Bhushan 2009a)

to the imidazolium cation at the ends of the chain provide a means for strong H-bonding interactions with active sites on the silicon surface. As shown in Fig. 14.47, either one (case 1) or two (case 2) hydroxyl groups can create this bond. The second case is particularly desirable because if the two hydroxyl groups are bonded (i.e., not exposed to the surface), they are not available to interact with water molecules in the ambient, leading to a reduction in the adhesion and friction forces, as well as enhanced wear resistance (Palacio and Bhushan 2009a).

In BHPET, these hydroxyl group attachment schemes are also applicable. However, the additional mechanism of intramolecular hydrogen bonding can also take place (case 3). This is not as desirable as the second case because it depletes the available chain ends with hydroxyl groups which can bond to the silicon surface. The interaction of the lubricant film with the silicon substrate is weakened, and water molecules can displace the lubricant from the substrate. In addition, the polyether chain that links the two cations contains five oxygen atoms in each chain. These oxygen atoms can also form H-bonds with the water molecules in the ambient. This can account for the large difference in the friction properties between BHPET and BHPT. While the application of the BHPT film has lowered the coefficient of friction of the silicon surface, the BHPET film did not have a similar effect due to the molecular interactions described above. This interpretation of lubricant-substrate interaction is corroborated by the adhesion and friction data at varying humidities and temperatures, as well as the FTIR spectra. Both BMIM-PF₆ and BHPET have large peaks in their FTIR spectra that correspond to H-bonds of water molecules, which is consistent with the observed sensitivity of their adhesion and friction to the change in humidity and temperature. Meanwhile, BHPT does not have the aforementioned peak, and its adhesion and friction properties appear to be less sensitive to water molecules compared to the other two ionic liquids (Palacio and Bhushan 2009a).

Figure 14.48 is a schematic showing the interaction of the anions with the silicon substrate. As mentioned previously, XPS spectra

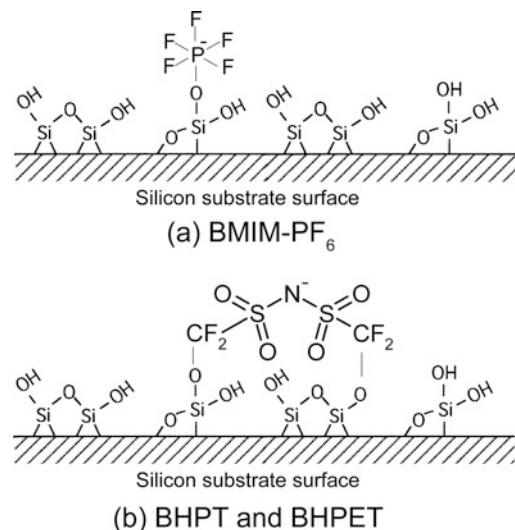


Fig. 14.48 Schematic for the attachment of the anions of **a** BMIM-PF₆, and **b** BHPT and BHPET to the silicon substrate (Palacio and Bhushan 2009a)

indicate immobilization of the ionic liquid, which occurs through the reaction of the anion with the hydroxyl groups on the substrate surface. In BMIM-PF₆, the O attaches to P, while in BHPT and BHPET (which have the same anion), the O bonds to C. Since the reactions involve the creation of new covalent bonds, this implies that the anions of the ILs investigated are more strongly attached to the substrate compared to the cations, which are chemically adsorbed (Palacio and Bhushan 2009a).

Based on the micro- and nanoscale friction and wear measurements, the ionic liquids show strong potential as lubricants for MEMS/NEMS because they have desirable thermal and electrical conductivity as well as desirable tribological properties.

14.5 Closure

Boundary films are formed by physisorption, chemisorption, and chemical reaction. A good boundary lubricant should have a high degree of interaction between its molecules and the solid surface. As a general rule, liquids are good lubricants when they are polar and, thus, able to grip

solid surfaces (or be adsorbed). In this chapter, we presented an overview of various perfluoropolyethers (PFPEs) and ionic liquid films.

Nanodeformation studies have shown that fully bonded PFPE lubricants behave as soft polymer coatings, while the unbonded lubricants behave liquid-like. AFM studies have shown that the physisorbed nonpolar molecules on a solid surface have an extended, flat conformation. The spreading property of PFPE is strongly dependent on the molecular end groups and the substrate chemistry.

Using a solid-like Z-DOL (fully bonded) film can reduce the friction and adhesion of Si(100), while using the liquid-like lubricant Z-15 showed a negative effect. Si(100) and a Z-15 film showed an apparent time effect. The friction and adhesion forces increased as a result of the growth of the meniscus up to an equilibrium time, after which they remained constant. The use of a Z-DOL (fully bonded) film can prevent rest time effects. During sliding at high velocity, the meniscus was broken and did not have enough time to rebuild, which led to a decrease of friction force and adhesive force of Z-15 and Z-DOL (fully bonded). The influence of relative humidity on friction and adhesion is dominated by the amount of adsorbed water molecules. Increasing humidity can either increase friction through increased adhesion by water meniscus or reduce friction through an enhanced water-lubricating effect. Increasing temperature led to desorption of the water layer, a decrease in water surface tension, a decrease in viscosity, and easier orientation of the Z-DOL (fully bonded) molecules. These changes caused a decrease of friction force and adhesion at high temperature. During cycling tests, the molecular interaction between the attached Z-15 molecules to the tip and the Z-15 molecules on the film surface caused the initial increase in friction. Wear tests showed that Z-DOL (fully bonded) can improve the wear resistance of silicon. Partially bonded PFPE films appear to be more durable than fully bonded films. This suggests that partially/fully bonded films are good lubricants for micro/nanoscale devices operating in different environmental conditions.

The surface potential technique has been useful in detecting lubricant removal and the initiation of substrate wear. The increase in surface potential is attributed to the change in the work function of the silicon after wear and electrostatic charge buildup of debris in the lubricant. Coatings with a mobile lubricant fraction were better able to protect the silicon substrate from wear compared to the fully bonded coating. This enhanced protection is attributed to a lubricant replenishment mechanism. The contact resistance technique provides complementary information to surface potential data in detecting exposure of the substrate after wear and is a promising method for studying conducting lubricants.

Degradation of novel PFPE lubricants—Z-DOL, Z-TETRAOL, A20H, and Z-DOL + A20H (30 vol.%)—was studied. The coefficient of friction of Z-TETRAOL was lower than Z-DOL and A20H in high vacuum. Fully bonded Z-TETRAOL exhibited higher durability than fully bonded Z-DOL on Si, and both were more durable than untreated Z-DOL. A20H showed very high durability in high vacuum. Environment influences the friction and durability of PFPE films. Generally, PFPE films are less durable in high vacuum than in other environments because of intimate contact between surfaces. In argon, the PFPE films showed low friction and high durability. Water molecules can act as a lubricant for PFPE films at a moderate humidity level, while at a high humidity level they can penetrate the PFPE films and cause increased friction. The durability of fully bonded Z-TETRAOL and partially bonded A20H is higher than that of Z-DOL in tested environments. The mixture of Z-DOL + A20H (30 vol. %) showed low durability in high vacuum while it showed high durability in argon and air with various humidity levels.

The ionic liquids exhibited low adhesion, friction, and wear properties comparable to the PFPE lubricants. For example, ionic liquid BMIM-PF₆ is comparable to Z-TETRAOL. Based on the surface height, adhesion, and friction data, chemical bonding treatment facilitated

attachment of the ionic liquid to the silicon substrate surface, leading to a more uniform coating and lower adhesion force and coefficient of friction. The partially bonded coatings have the lowest coefficients of friction and longest durability as they possess a desirable combination of lubricant bonded to the substrate, as well as a mobile fraction that facilitates sliding. In micro- and nanoscale experiments, the ionic liquid exhibited comparable durability with the lubricant Z-TETRAOL, which has high thermal stability and extremely low vapor pressure. The low post-wear surface potential change observed on the silicon coated with ionic liquid is indicative of enhanced charge dissipation compared to the Z-TETRAOL-coated and uncoated surfaces, which are poor conductors. Contact resistance data is consistent with the surface potential data with regards to identifying the role of the mobile and immobile lubricant fractions in protecting the surface from wear.

The dicationic liquid BHPT exhibits superior nanoscale friction and wear resistance properties. This is attributed to the presence of a pentyl chain and hydroxyl groups on both chain ends, which facilitate molecular orientation, as well as bonding interactions with the substrate surface. The other dicationic liquid, BHPET, has less desirable adhesion, friction, and wear properties compared to either BHPT or BMIM-PF₆. Intermolecular hydrogen bonding in BHPET reduces the chain ordering on the substrate surface, which accounts for the observed higher adhesive force and coefficient of friction relative to the other ionic liquids investigated. From the nanoscale wear tests, surface potential, and contact resistance imaging, it was found that the films are in this order of decreasing durability: BHPT > BMIM-PF₆ > BHPET. The microscale coefficient of friction was higher than the nanoscale value due to the difference in the length scales and configurations of the two test methods.

Nanotribological experiments performed at various humidities and temperature conditions indicated that the adhesive force and coefficient of friction of the ionic liquid films, especially BMIM-PF₆ and BHPET, are highly sensitive to the amount of water molecules present on the

surface. This was confirmed by FTIR spectroscopy, where the BMIM-PF₆ and BHPET films showed peaks corresponding to hydrogen bonding mainly resulting from water adsorption and partially from intermolecular interactions. The absence of this peak in BHPT indicated that a much smaller amount of water was adsorbed by this ionic liquid, which is consistent with its superior nanotribological properties. The anions are more strongly attached to the substrate compared to the cations. The anions could be covalently bonded, while the cations are chemically adsorbed. XPS analysis confirmed the immobilization of the anion to the silicon substrate.

In summary, PFPE and ionic liquids can be used for nanoscale boundary lubrication in demanding environments.

References

- Anderson JL, Ding R, Ellern A, Armstrong DW (2005) Structure and properties of high stability germinal dicationic ionic liquids. *J Am Chem Soc* 127:593–604
- Bhushan B (1993) Magnetic recording surfaces. In: Glaeser WA (ed) Characterization of tribological materials. Butterworth-Heinemann, Boston, pp 116–133
- Bhushan B (1996) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer, New York
- Bhushan B (1998) tribology issues and opportunities in MEMS. Kluwer Academic Pub, Dordrecht
- Bhushan B (1999) Handbook of micro/nanotribology, 2nd edn. CRC Press, Boca Raton
- Bhushan B (2001) Macro-and microtribology of magnetic storage devices. In: Bhushan B (ed) Modern tribology handbook, vol. 2: materials, coatings, and industrial applications. CRC Press, Boca Raton, pp 1413–1513
- Bhushan B (2008a) Nanotribology and nanomechanics in nano/biotechnology. *Philos Trans R Soc A* 366: 1499–1537
- Bhushan B (2008b) Nanotribology, nanomechanics and nanomaterials characterization. *Philos Trans R Soc A* 366:1351–1381
- Bhushan B (2013a) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2013b) Principles and applications of tribology, 2nd edn. Wiley, New York
- Bhushan B (2016) Encyclopedia of nanotechnology, 2nd edn. Springer International, Switzerland
- Bhushan B, Dandavate C (2000) Thin-film friction and adhesion studies using atomic force microscopy. *J Appl Phys* 87:1201–1210

- Bhushan B, Kwak K (2008) Role of lubricants, scanning velocity, and environment on adhesion, friction and wear of Pt-Ir coated probes for atomic force microscopy probe-based ferroelectric recording technology. *J Phys: Condens Matter* 20, Art. # 325240
- Bhushan B, Ruan J (1994a) Atomic-scale friction measurements using friction force microscopy: Part II—application to magnetic media. *ASME J Tribol* 116:389–396
- Bhushan B, Ruan J (1994b) Tribological performance of thin film amorphous carbon overcoats for magnetic recording disks in various environments. *Surf Coat Technol* 68(69):644–650
- Bhushan B, Sundararajan S (1998) Micro/nanoscale friction and wear mechanisms of thin films using atomic force and friction force microscopy. *Acta Mater* 46:3793–3804
- Bhushan B, Tao Z (2006) Lubrication of advanced metal evaporated tape using novel perfluoropolyether lubricants. *Microsyst Technol* 12:579–587
- Bhushan B, Zhao Z (1999) Macroscale and microscale tribological studies of molecularly thick boundary layers of perfluoropolyether lubricants for magnetic thin-film rigid disks. *J Info Storage Proc Syst* 1:1–21
- Bhushan B, Israelachvili JN, Landman U (1995a) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Yang L, Gao C, Suri S, Miller RA, Marchon B (1995b) Friction and wear studies of magnetic thin-film rigid disks with glass-ceramic, glass, and aluminum-magnesium substrates. *Wear* 190:44–59
- Bhushan B, Cichomski M, Tao Z, Tran NT, Ethen T, Merton C, Jewett RE (2007) Nanotribological characterization and lubricant degradation studies of metal-film magnetic tapes using novel lubricants. *ASME J Tribol* 129:621–627
- Bhushan B, Palacio M, Kinzig B (2008a) AFM-based nanotribological and electrical characterization of ultrathin wear-resistant ionic liquid films. *J Colloid Interf Sci* 317:275–287
- Bhushan B, Kwak K, Palacio M. (2008b) Nanotribology and Nanomechanics of AFM probe based data recording technology. *J Phys: Condens Matter* 20, Art. # 365207
- Blackman GS, Mate CM, Philpott MR (1990a) Interaction forces of a sharp tungsten tip with molecular films on silicon surface. *Phys Rev Lett* 65:2270–2273
- Blackman GS, Mate CM, Philpott MR (1990b) Atomic force microscope studies of lubricant films on solid surfaces. *Vacuum* 41:1283–1286
- Carda-Broch S, Berthod A, Armstrong DW (2003) Solvent properties of the 1-butyl-3-methylimidazolium hexafluorophosphate ionic liquid. *Anal Bioanal Chem* 375:191–199
- Chilamakuri SK, Bhushan B (1999) A comprehensive kinetic meniscus model for prediction of long-term static friction. *J Appl Phys* 15:4649–4656
- Danyluk S, McNallan M, Park DS (1994) Friction and wear of silicon nitride exposed to moisture at high temperatures. In: Jahamir S (ed) *Friction and wear of ceramics*. M. Dekker, New York, pp 61–79
- DeVecchio D, Bhushan B (1998) Use of a nanoscale kelvin probe for detecting wear precursors. *Rev Sci Instrum* 69:3618–3624
- Eapen KC, Patton ST, Zabinski JS (2002) Lubrication of microelectromechanical systems (MEMS) using bound and mobile phase of fomblin Z-DOL. *Tribol Lett* 12:35–41
- Feldman LC, Mayer JW (1986) Fundamentals of surface and thin film analysis. Prentice-Hall, NJ
- Fischer TE (1988) Triboochemistry. *Annu Rev Mater Sci* 18:303–323
- Frez C, Diebold GJ, Tran CD, Yu S (2006) Determination of thermal diffusivities, thermal conductivities and sound speed of room-temperature ionic liquids by the transient grating technique. *J Chem Eng Data* 51:1250–1255
- Hahm C, Bhushan B (1997) High shear rate viscosity measurement of perfluoropolyether lubricants for magnetic thin-film rigid disks. *J Appl Phy* 81:5384–5386
- Hoque E, DeRose JA, Bhushan B, Mathieu HJ (2008) Self-assembled monolayers on aluminum and copper oxide surfaces: surface and interface characteristics, nanotribological properties, and chemical stability. In: Bhushan B, Fuchs H, Tomitori M (eds) *Applied scanning probe methods vol. IX—characterization*. Springer, Heidelberg, Germany, pp 235–281
- Ishigaki H, Kawaguchi I, Iwasa M, Toibana Y (1986) Friction and wear of hot pressed silicon nitride and other ceramics. *ASME J Trib* 108:514–521
- Israelachvili JN (1992) *Intermolecular and surface forces*, 2nd edn. Academic Press, London
- Jonsson U, Bhushan B (1995) Measurement of rheological properties of ultrathin lubricant films at very high shear rates and near-ambient pressure. *J Appl Phys* 78:3107–3109
- Kabo GJ, Blokhin AV, Paulechka YU, Kabo AJ, Shymanovich MP, Magee JW (2004) Thermodynamic properties of 1-butyl-3-methylimidazolium hexafluorophosphate in the condensed state. *J Chem Eng Data* 49:453–461
- Kim CA, Choi HJ, Kono RN, Jhon MS (1999) Rheological characterization of perfluoropolyether lubricant. *Polym Prepr* 40:647–649
- Kinzig BJ, Sutor P (2005) Ionic liquids: novel lubrication for air and space. In: Phase I final report for AFOSR/NL, surfaces research and applications. Inc. Lenexa, KS
- Koinkar VN, Bhushan B (1996a) Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J Vac Sci Technol A* 14:2378–2391
- Koinkar VN, Bhushan B (1996b) Micro/nanoscale studies of boundary layers of liquid lubricants for magnetic disks. *J Appl Phys* 79:8071–8075

- Liu H, Bhushan B (2003) Nanotribological characterization of molecularly-thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* 97:321–340
- Ma X, Gui J, Grannen KJ, Smoliar LA, Marchon B, Jhon MS, Bauer CL (1999) Spreading of PFPE lubricants on carbon surfaces: effect of hydrogen and nitrogen content. *Tribo Lett* 6:9–14
- Mate CM (1992a) Application of disjoining and capillary pressure to liquid lubricant films in magnetic recording. *J Appl Phys* 72:3084–3090
- Mate CM (1992b) Atomic-force-microscope study of polymer lubricants on silicon surface. *Phys Rev Lett* 68:3323–3326
- Mate CM (1993) Nanotribology of lubricated and unlubricated carbon overcoats on magnetic disks studied by friction force microscopy. *Surf Coat Technol* 62:373–379
- Mate CM, Novotny VJ (1991) Molecular conformation and disjoining pressures of polymeric liquid films. *J Chem Phys* 94:8420–8427
- Mizuhara K, Hsu SM (1992) Tribocochemical reaction of oxygen and water on silicon surfaces. In: Dowson D (ed) *Wear particles*. Elsevier Science Press, New York, pp 323–328
- Morrison RT, Boyd RN (1992) Organic chemistry, 6th edn. Wiley, New York
- Muratov VA, Fischer TE (2000) Tribocochemical polishing. *Annu Rev Mater Sci* 30:27–51
- Novotny VJ (1990) Migration of liquid polymers on solid surfaces. *J Chem Phys* 92:3189–3196
- Novotny VJ, Hussla I, Turlet JM, Philpott MR (1989) Liquid polymer conformation on solid surfaces. *J Chem Phys* 90:5861–5868
- O’Shea SJ, Welland ME, Rayment T (1992) Atomic force microscope study of boundary layer lubrication. *Appl Phys Lett* 61:2240–2242
- O’Shea SJ, Welland ME, Pethica JB (1994) Atomic force microscopy of local compliance at solid-liquid interface. *Chem Phys Lett* 223:336–340
- Palacio M, Bhushan B (2007a) Surface potential and resistance measurements for detecting wear of chemically-bonded and unbonded molecularly-thick perfluoropolyether lubricant films using atomic force microscopy. *J Colloid Interf Sci* 315:261–269
- Palacio M, Bhushan B (2007b) Wear detection of candidate MEMS/NEMS lubricant films using atomic force microscopy-based surface potential measurements. *Scripta Mater* 57:821–824
- Palacio M, Bhushan B (2008a) Ultrathin wear-resistant ionic liquid films for novel MEMS/NEMS applications. *Adv Mater* 20:1194–1198
- Palacio M, Bhushan B (2008b) Nanotribological and nanomechanical properties of lubricated PZT thin films for ferroelectric data storage applications. *J Vac Sci Technol A* 26:768–776
- Palacio M, Bhushan B (2009a) Molecularly thick dicationic liquid films for nanolubrication. *J Vac Sci Technol A* 27:986–995
- Palacio M, Bhushan B (2009b) Nanotribological properties of novel lubricants for magnetic tapes. *Ultramicroscopy* 109:980–990
- Payagala T, Huang J, Breitbach ZS, Sharma PS, Armstrong DW (2007) Unsymmetrical dicationic ionic liquids: manipulation of physicochemical properties using specific structural architectures. *Chem Mater* 19:5848–5850
- Reich RA, Stewart PA, Bohaychick J, Urbanski JA (2003) Base oil properties of ionic liquids. *Lub Eng* 49:16–21
- Roberts GG (1990) Langmuir-Blodgett films. Plenum, New York
- Ruhe J, Blackman G, Novotny VJ, Clarke T, Street GB, Kuan S (1994) Thermal attachment of perfluorinated polymers to solid surfaces. *J Appl Poly Sci* 53: 825–836
- Ruhe J, Novotny V, Clarke T, Street GB (1996) Ultrathin perfluoropolyether films—Influence of anchoring and mobility of polymers on the tribological properties. *ASME J Tribol* 118:663–668
- Ruths M, Granick S (1998) Rate-dependent adhesion between opposed perfluoropoly(alkylether) layers: dependence on chain-end functionality and chain length. *J Phys Chem B* 102:6056–6063
- Stifter T, Marti O, Bhushan B (2000) Theoretical investigation of the distance dependence of capillary and Van der Waals forces in scanning probe microscopy. *Phys Rev B* 62:13667–13673
- Sundararajan S, Bhushan B (2001) Static friction and surface roughness studies of surface micromachined electrostatic micromotors using an atomic force/friction force microscope. *J Vac Sci Tech A* 19:1777–1785
- Tao Z, Bhushan B (2005) Bonding, degradation, and environmental effects on novel perfluoropolyether lubricants. *Wear* 259:1352–1361
- Ulman A (1991) An introduction to ultrathin organic films. Academic Press, Boston
- Valkenberg MH, deCastro C, Holderich WF (2002) Immobilization of ionic liquids on solid supports. *Green Chem* 4:88–93
- Wang H, Lu Q, Ye C, Liu W, Cui Z (2004) Friction and wear behaviors of ionic liquid of alkylimidazolium hexafluorophosphates as lubricants for steel/steel contacts. *Wear* 256:44–48
- Yoshizawa H, Israelachvili JN (1993) Fundamental mechanisms of interfacial friction II: stick-slip friction of spherical and chain molecules. *J Phys Chem* 97:11300–11313
- Yoshizawa H, Chen YL, Israelachvili JN (1993) Fundamental mechanisms of interfacial friction I: relationship between adhesion and friction. *J Phys Chem* 97:4128–4140
- Zhang L, Zhang Q, Li J (2007) Electrochemical behaviors and spectral studies of ionic liquid (1-butyl-3-methylimidazolium tetrafluoroborate) based on sol-gel electrode. *J Electroanal Chem* 603:243–248

- Zhao X, Bhushan B (2000) Comparison studies on degradation mechanisms of perfluoropolyether lubricants and model lubricants. *Tribol Lett* 9:187–197
- Zhao X, Bhushan B (2001) Studies on degradation mechanisms of lubricants for magnetic thin-film rigid disks. *Proc Inst Mech Eng Part J: J Eng Tribol* 215:173–188
- Zhao X, Bhushan B, Kajdas C (2000) Lubrication studies of head-disk interfaces in a controlled environment, Part 2: degradation mechanisms of perfluoropolyether lubricants. *Proc Inst Mech Eng Part J: J Eng Tribol* 214:547–559

Part V

Industrial Applications

Micro/Nanotribology and Micro/Nanomechanics of Magnetic Storage Devices

Bharat Bhushan

Abstract

A magnetic recording process involves relative motion between a magnetic medium (tape or disk) against a stationary or rotating read/write magnetic head. For ever-increasing, high areal recording density, the linear flux density (number of flux reversals per unit distance) and the track density (number of tracks per unit distance) should be as high as possible. The size of a single bit dimension for current devices is typically less than 1000 nm^2 . This dimension places stringent restrictions on the defect size present on the head and medium surfaces. Reproduced (read-back) magnetic signal amplitude decreases with a decrease in the recording wavelength and/or the track width. The signal loss results from the magnetic coating thickness, read gap length, and head-to-medium spacing (clearance or flying height). It is known that the signal loss as a result of spacing can be reduced exponentially by reducing the separation between the head and the medium. The need for increasingly higher recording densities requires that surfaces be as smooth as possible and the flying height (physical separation or clearance between a head and a medium) be as low as possible. The ultimate objective is to run two surfaces in contact (with practically zero physical separation) if the tribological issues can be resolved. Smooth surfaces in near contact lead to an increase in adhesion, friction, and interface temperatures, and closer flying heights lead to occasional rubbing of high asperities and increased wear. Friction and wear issues are resolved by appropriate selection of interface materials and lubricants, by controlling the dynamics of the head and medium, and the environment. A fundamental understanding of the tribology (friction, wear, and lubrication) of the magnetic head/medium

B. Bhushan (✉)

Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390, Columbus, OH 43210,
USA
e-mail: bhushan.2@osu.edu

interface, both on macro- and micro/nanoscales, becomes crucial for the growth of this more than \$60 billion a year magnetic storage industry. In this chapter, the general operation of drives and the construction and materials used in magnetic head and medium components are described. Then, micro/nanotribological and micro/nanomechanics studies, including surface roughness, friction, adhesion, scratching, wear, indentation, and lubrication relevant to magnetic storage devices, are presented.

Keywords

Magnetic storage devices • Hard disk drives • Tape drives • Nanotribology • Nanomechanics • Reliability • Adhesion • Stiction • Friction • Wear • Lubrication

15.1 Introduction

15.1.1 Magnetic Storage Devices

Magnetic storage devices used for storage and retrieval are tape and rigid or hard disk drives (HDDs). These devices primarily are used for data storage applications. Tape drives are used for backup and mass storage, given that they have the highest volumetric density and lowest cost per unit storage. The magnetic storage industry is a \$60 billion a year industry in the 1990's, and the total revenue has not changed much in spite of dramatic increases in the volume. Semiconductor storage devices (SSDs) continue to satisfy some of the storage needs. In the data storage industry, magnetic rigid disk drives/media, tape drives/media, and optical disk drive/media account for about \$30B, \$5B, and \$5B, respectively. Magnetic recording and playback involves the relative motion between a magnetic medium (tape or disk) against a read-write magnetic head. Heads are designed so that they develop a (load-carrying) hydrodynamic air film under steady operating conditions to minimize head-medium contact. However, physical contact between the medium and head occurs during starts and stops, referred to as contact-start-stops (CSS) technology (Bhushan 1994, 1996a, 2000, 2001b; Castillo and Bhushan 2006).

In modern magnetic storage devices, the flying heights (head-to-medium separation) are on the order of 5–20 nm and roughnesses of head and medium surfaces are on the order of 1–2 nm RMS. The need for ever-increasing recording densities requires that surfaces be as smooth as possible and the flying heights be as low as possible. Smooth surfaces lead to an increase in adhesion, friction, and interface temperatures, and closer flying heights lead to occasional rubbing of high asperities and increased wear. High stiction (static friction) and wear remain the limiting technology to this industry. The head normally rests on a special laser-textured zone to minimize stiction. Head load/unload (L/UL) technology also is commonly used as an alternative to CSS technology in rigid disk drives that eliminates stiction and wear failure mode associated with CSS. In an L/UL drive, a lift tab extending from the suspension load beam engages a ramp or cam structure as the actuator moves beyond the outer radius of the disk. The ramp lifts (or unloads) the head stack from the disk surfaces as the actuator moves to the parking position. Starting and stopping the disk only occur with the head in the unloaded state (see Fig. 15.3, to be introduced later). Several near contact recording devices are also in use. High stiction and wear are the major impediments to the commercialization of contact recording

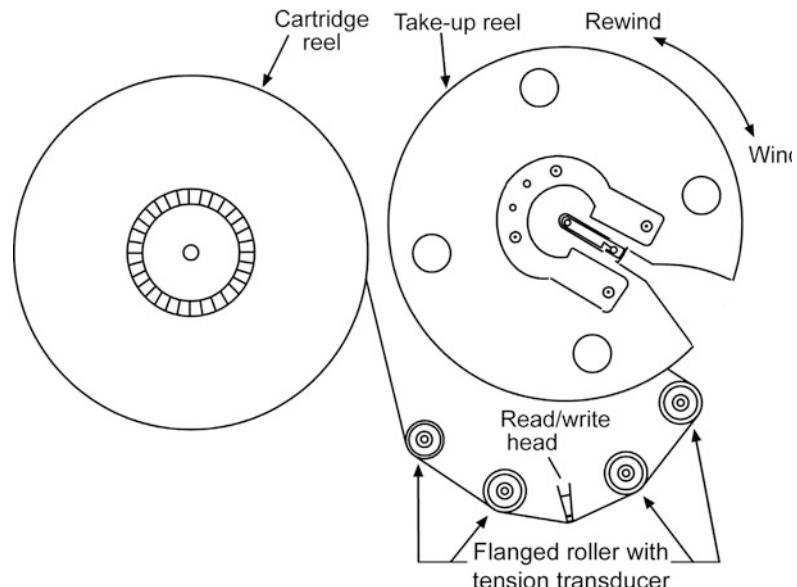
(Bhushan 1993a, 1994, 1995a, b, 1996a, 2000, 2001b, 2003; Castillo and Bhushan 2006).

Magnetic media fall into two categories: particulate media, where magnetic particles ($\gamma\text{-Fe}_2\text{O}_3$, Co- $\gamma\text{Fe}_2\text{O}_3$, CrO_2 , Fe or metal (MP), or barium ferrite) are dispersed in a polymeric matrix and coated onto a polymeric substrate for flexible media (tape); thin-film media, where continuous films of magnetic materials are deposited by vacuum deposition techniques onto a polymer substrate for flexible media or onto a rigid substrate (aluminum or more commonly glass or glass ceramic) for rigid disks. Commonly-used thin magnetic films for tapes are evaporated Co-Ni (82–18 at.%) or Co-O dual layer. Typical magnetic films for rigid disks are metal films of cobalt-based alloys (such as sputtered Co-Pt-Ni, Co-Ni, Co-Pt-Cr, Co-Cr and Co-NiCr). For high recording densities, thin-film media are used. Magnetic heads used are either conventional thin-film inductive magnetoresistive (MR) or giant MR (GMR) heads. The air-bearing surfaces (ABS) of tape heads are generally cylindrical in shape. The rigid-disk heads are supported by a leaf spring (flexure) suspension. The ABS of heads are made of Mn-Zn ferrite, Ni-Zn ferrite, Al_2O_3 -TiC, or calcium titanate (Bhushan 1996a, 2000, 2001b, 2003;

Castillo and Bhushan 2006). In the past, the ABS of some conventional heads were made of plasma-sprayed coatings of hard materials such as Al_2O_3 -TiO₂ and ZrO₂.

Figure 15.1 shows a schematic illustrating the tape path with details of tape guides in a data-processing Linear Tape Open (LTO) drive which uses a rectangular tape cartridge. Figure 15.2a shows the section views of particulate and thin-film magnetic tapes. Almost exclusively, the base film is made of semicrystalline biaxially-oriented poly (ethylene terephthalate) (or PET), poly (ethylene 2,6 naphthalate) (or PEN), or Aramid. The particulate coating formulation consists of binder (typically polyester polyurethane), submicron acicular-shaped magnetic particles (about 50 nm long with an aspect ratio of about 5), submicron head-cleaning agents (typically alumina), and lubricant (typically fatty acid ester). For protection against wear and corrosion and low friction/stiction, the thin-film tape first is coated with a diamondlike carbon (DLC) overcoat deposited by plasma-enhanced chemical vapor deposition, topically lubricated with primarily a perfluoropolyether lubricant. Figure 15.2b shows a schematic of an 8-track (along with 2 servo tracks) thin-film read-write head with MR read and inductive write. The head

Fig. 15.1 Schematic of tape path in an IBM Linear Tape Open (LTO) drive



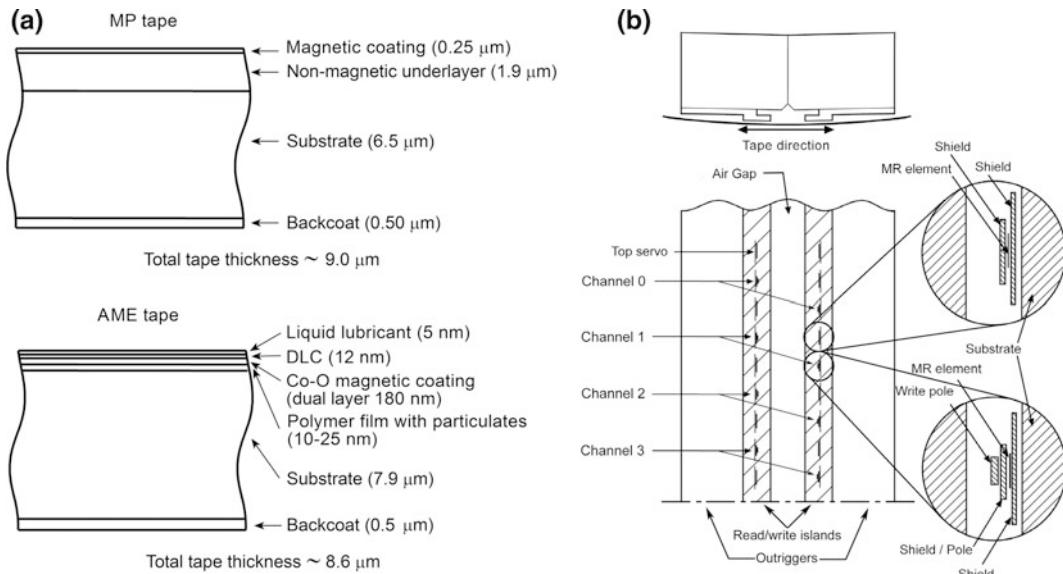


Fig. 15.2 (a) Sectional views of particulate and thin-film magnetic tapes, and (b) schematic of a magnetic thin-film read/write head for an IBM LTO Gen 1 tape drive

steps up and down to provide 384 total data tracks across the width of the tape. The ABS is made of Al_2O_3 -TiC. A tape tension of about 1 N over a 12.7-mm wide tape (normal pressure \sim 14 kPa) is used during use. The RMS roughnesses of the ABS of the heads and tape surfaces typically are 1–1.5 nm and 5–8 nm, respectively.

Figure 15.3 shows schematics of data processing rigid disk drives available with various form factors (21.6-, 27.4-, 48-, 63.5-, 75-, and 95-mm). The figure shows two technologies used to rest the head while the drive is off (to be discussed later). A nonremovable stack of multiple disks mounted on a ball bearing or hydrodynamic spindle, are rotated by an electric motor

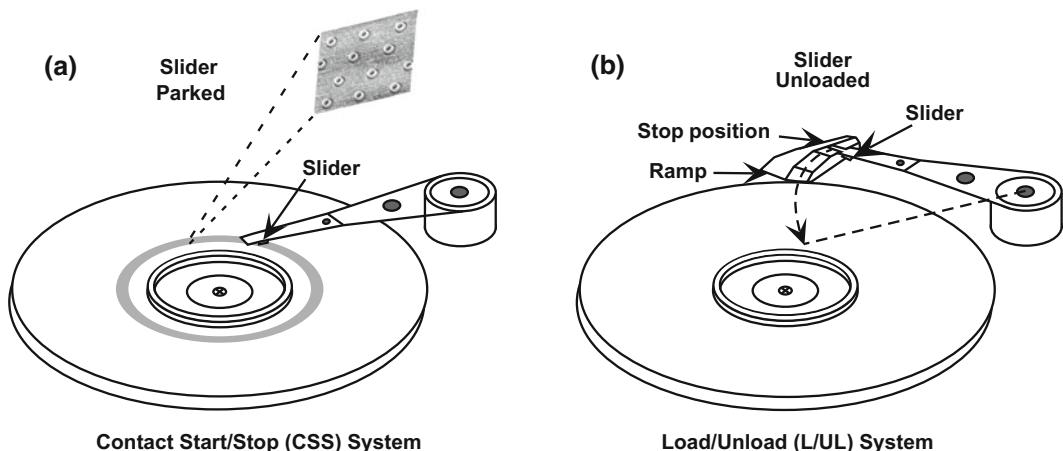
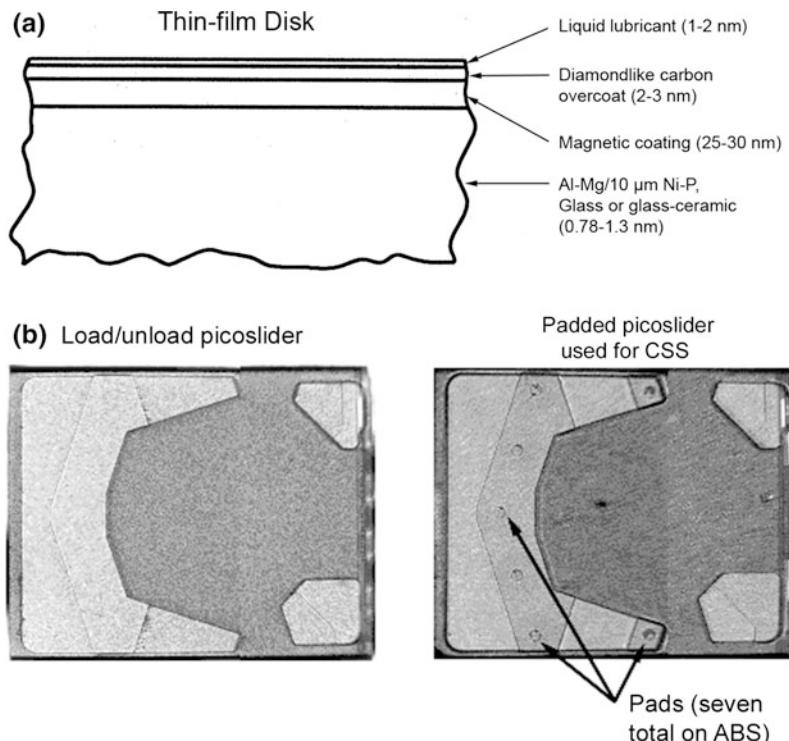


Fig. 15.3 Schematics of data processing magnetic rigid disks. The figure shows two different technologies used to rest the heads while the drive is off. (a) CSS—The heads rest in a special laser-textured zone; (b) L/UL heads rest in the ramp, outside of the disk surface

Fig. 15.4 (a) Sectional views of a thin-film magnetic rigid disk, and (b) schematic of two picosliders—load/unload picoslider and padded picoslider used for CSS



at constant angular speed ranging from about 5000 to in excess of 15,000 RPM, dependent upon the disk size. Head-slider-suspension assembly (allowing one slider for each disk surface) is actuated by a stepper motor or a voice coil motor using a rotary actuator. Figure 15.4a shows the section views of a thin-film rigid disk. The substrate for rigid disks is generally a non-heat-treatable aluminum-magnesium alloy 5086, glass, or glass ceramic. The protective overcoat commonly used for thin-film disks is sputtered DLC, topically lubricated with perfluoropolyether-type lubricants. Lubricants with polar-end groups are generally used for thin-film disks in order to provide partial chemical bonding to the overcoat surface. The disks used for CSS technology are laser textured in the landing zone (Fig. 15.3, left). As mentioned earlier, as an alternative to CSS, L/UL technology is also used commonly (Fig. 15.3, right). Figure 15.4b shows a schematic of two thin-film head picosliders with a step at the leading edge, and GMR read and inductive write. "Pico" refers to the small sizes of 1.25 mm × 1 mm. These

sliders use $\text{Al}_2\text{O}_3\text{-TiC}$ (70–30 wt%) as the substrate material with multilayered thin-film head structure coated and with about 3.5-nm thick DLC coating to prevent the thin film structure from electrostatic discharge. The seven pads on the padded slider are made of DLC and are about 40 µm in diameter and 50 nm in height. A normal load of about 3 g is applied during use.

15.1.2 Micro/Nanotribology and Micro/Nanomechanics and Their Applications

The micro/nanotribological studies are needed to develop fundamental understanding of interfacial phenomena on a small scale and to study interfacial phenomena in (Bhushan et al. 1995a; Bhushan 1995b, 1997, 1999a, c, 2001a, b, 2005, 2008, 2013). Magnetic storage devices operate under low load and encounter isolated asperity interactions. These use multilayered thin film structure and are generally lubricated with molecularly-thin films. Micro/nanotribological

and micro/nanomechanical techniques are ideal to study the friction and wear processes of micro/nanoscale and molecularly thick films. These studies are also valuable in fundamental understanding of interfacial phenomena in macrostructures to provide a bridge between science and engineering. At interfaces of technological applications, contact occurs at multiple asperity contacts. A sharp tip of tip-based microscopes (atomic force/friction force microscopes or AFM/FFM) sliding on a surface simulates a single asperity contact, thus allowing high-resolution measurements of surface interactions at a single asperity contacts. AFMs/FFMs are now commonly used for tribological studies.

In this chapter, we present state-of-the-art of micro/nanotribology and micro/nanomechanics of magnetic storage devices including surface roughness, friction, adhesion, scratching, wear, indentation, and lubrication.

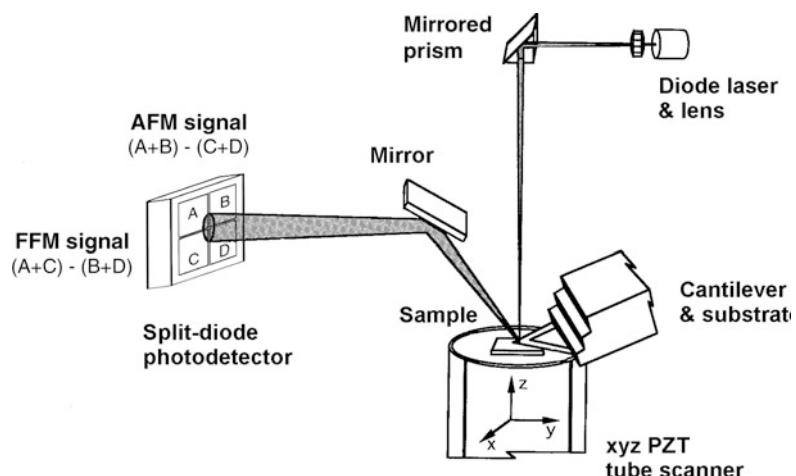
15.2 Experimental

15.2.1 Description of AFM/FFM

AFM/FFMs used in the tribological studies have been described in several papers (Bhushan et al. 1995a; Bhushan 1999a, 2013, 2017). Briefly, in one of the commercial designs, the sample is mounted on a PZT tube scanner to scan the sample in the x-y plane and to move the sample

in the vertical (z) direction, Fig. 15.5. A sharp tip at the end of a flexible cantilever is brought in contact with the sample and the sample is scanned in a raster pattern, Fig. 15.6. Normal and frictional forces being applied at the tip-sample interface are simultaneously measured using a laser beam deflection technique. Surface roughness is measured either in the contact mode or the so-called tapping mode (intermittent contact mode). For surface roughness and friction measurements, a microfabricated square pyramidal Si_3N_4 tip with a tip radius of about 30 nm attached to a cantilever beam (with a normal beam stiffness of about 0.5 N/m) for contact mode or a square-pyramidal etched single-crystal silicon tip with a rectangular silicon cantilever beam (Fig. 15.7) is generally used at normal loads ranging from 10 to 150 nN. A preferred method of measuring friction and calibration procedures for conversion of voltages corresponding to normal and friction forces to force units, are described by Bhushan (1999a, 2013, 2017). The samples are typically scanned over scan areas ranging from $50 \text{ nm} \times 50 \text{ nm}$ to $10 \mu\text{m} \times 10 \mu\text{m}$, in a direction orthogonal to the long axis of the cantilever beam (Ruan and Bhushan 1994a). The scan rate is on the order of 1 Hz. For example, for this rate, the sample scanning speed would be 1 $\mu\text{m}/\text{s}$ for a 500 nm 500 nm scan area. Adhesive force measurements are performed in the so-called friction calibration mode. In this technique, the tip is

Fig. 15.5 Principles of operation of a commercial small sample AFM/FFM



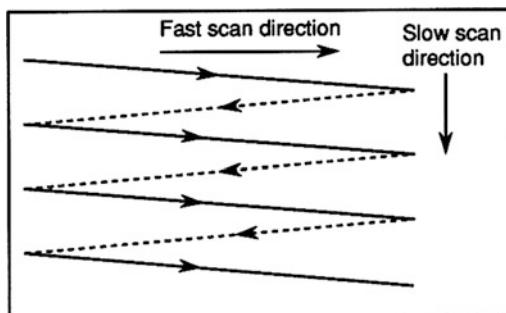
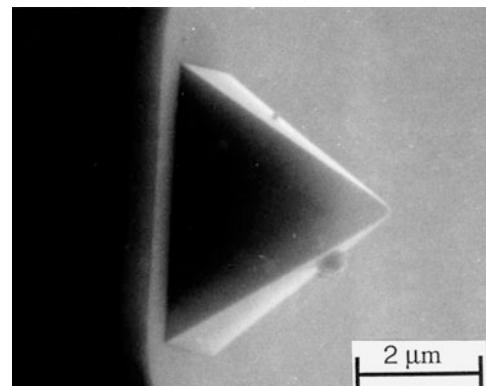


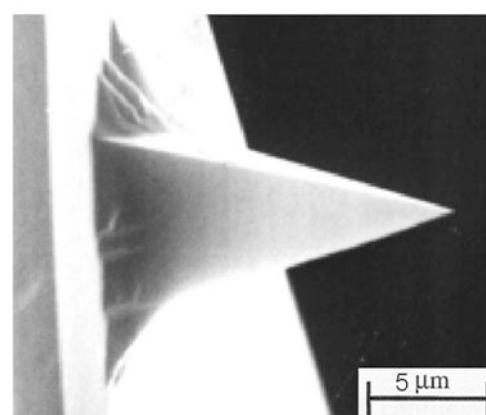
Fig. 15.6 Schematic of triangular pattern trajectory of the AFM tip as the sample is scanned in two dimensions. During imaging, data are recorded only during scans along the *solid scan lines*

brought in contact with the sample and then pulled away. The force required to pull the tip off the sample is a measure of adhesive force.

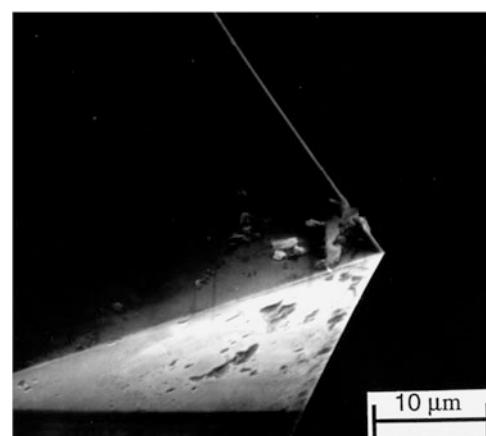
In nanoscale wear studies, the sample is initially scanned twice, typically at 10 nN to obtain the surface profile, then scanned twice at a higher load of typically 100 nN to wear and to image the surface simultaneously, and then rescanned twice at 10 nN to obtain the profile of the worn surface. For magnetic media studied by Bhushan and Ruan (1994), no noticeable change in the roughness profiles was observed between the initial two scans at 10 nN and between profiles scanned at 100 nN and the final scans at 10 nN. Therefore any changes in the topography between the initial scans at 10 nN and the scans at 100 nN (or the final scans at 10 nN) are believed to occur as a result of local deformation of the sample surface. In picoindentation studies, the sample is loaded in contact with the tip. During loading, tip deflection (normal force) is measured as a function of vertical position of the sample. For a rigid sample, the tip deflection and the sample traveling distance (when the tip and sample come into contact) are equal. Any decrease in the tip deflection as compared to vertical position of the sample represents indentation. To ensure that the curvature in the tip deflection-sample traveling distance curve does not arise from PZT hysteresis, measurements on several rigid samples, including single-crystal natural diamond (IIa), were made by Bhushan and Ruan (1994). No curvature was noticed for



Square pyramidal silicon nitride tip



Square pyramidal single-crystal silicon tip



Three-sided pyramidal natural diamond tip

Fig. 15.7 SEM micrographs of a square-pyramidal PECVD Si_3N_4 tip with a triangular cantilever beam (*top*), a square-pyramidal etched single-crystal silicon tip with a rectangular silicon cantilever beam (*middle*), and a three-sided pyramidal natural diamond tip with a square stainless steel cantilever beam (*bottom*)

the case of rigid samples. This suggests that any curvature for other samples should arise from the indentation of the sample.

For microscale scratching, microscale wear and nanoscale indentation hardness measurements, a three-sided pyramidal single-crystal natural diamond tip with an apex angle of 80°E and a tip radius of about 100 nm (determined by scanning electron microscopy imaging) (Fig. 15.7) is used at relatively high loads (1 μ N–150 μ N). The diamond tip is mounted on a stainless steel cantilever beam with normal stiffness of about 25 N/m (Bhushan and Koinkar 1994a, b; Bhushan et al. 1994, 1996). For scratching and wear studies, the sample is generally scanned in a direction orthogonal to the long axis of the cantilever beam (typically at a rate of 0.5 Hz). For wear studies, typically an area of 2 μ m \times 2 μ m is scanned at various normal loads (ranging from 1 to 100 μ N) for selected number of cycles. Scratching can also be performed at ramped loads (Sundararajan and Bhushan 2001). For nanoindentation hardness measurements, the scan size is set to zero and then normal load is applied to make the indents. During this procedure, the diamond tip is continuously pressed against the sample surface for about two seconds at various indentation loads. Sample surface is scanned before and after the scratching, wear, or nanoindentation, to obtain the initial and the final surface topography, at a low normal load of about 0.3 μ N using the same diamond tip. An area larger than the scratched, worn or indentation region is scanned to observe the scratch, wear scars, or indentation marks. Nanohardness is calculated by dividing the indentation load by the projected residual area of the indents (Bhushan and Koinkar 1994b). Nanohardness and Young's modulus of elasticity (stiffness) at shallow depths as low as 5 nm are measured using a depth-sensing capacitance transducer system in an AFM (Bhushan et al. 1996; Bhushan 1999a; Bhushan and Li 2003).

Indentation experiments provide a single-point measurement of the Young's modulus of elasticity (stiffness), localized surface elasticity as well as phase contrast maps (to obtain viscoelastic properties map) can be

obtained using dynamic force microscopy in which an oscillating tip is scanned over the sample surface in contact under steady and oscillating load (DeVecchio and Bhushan 1997; Scherer et al. 1997; Scott and Bhushan 2003; Bhushan and Qi 2003). Recently, a torsional resonance (TR) mode has been introduced (Kasai et al. 2004; Bhushan and Kasai 2004) which provides higher resolution. Stiffness and phase contrast maps can provide magnetic particle/polymer distributions in magnetic tapes as well as lubricant film thickness distribution.

Boundary lubrication studies are conducted using either Si_3N_4 or diamond tips (Bhushan et al. 1995b; Koinkar and Bhushan 1996a, b; Liu and Bhushan 2003). The coefficient of friction is monitored as a function of sliding cycles.

All measurements are carried out in the ambient atmosphere (22 \pm 1 °C, 45 \pm 5% RH, and Class 10,000).

15.2.2 Test Specimens

Data on various head slider materials and magnetic media are presented in the chapter. Al_2O_3 –TiC (70–30 w/o) and polycrystalline and single-crystal (110) Mn–Zn ferrite are commonly used for construction of disk and tape heads. Al_2O_3 , a single-phase material, is also selected for comparisons with the performance of Al_2O_3 –TiC, a two-phase material. A α -type SiC is also selected which is a candidate slider material because of its high thermal conductivity and attractive machining and friction and wear properties (Bhushan 1993b). Single crystal silicon has also been used in some head sliders but its use is discontinued (Bhushan et al. 1992).

Two thin-film rigid disks with polished and textured substrates, with and without a bonded perfluoropolyether are selected. These disks are 95-mm in diameter made of Al–Mg alloy substrate (1.3-mm thick) with a 10- μ m thick electroless plated Ni-P coating, 75-nm thick ($\text{Co}_{79}\text{Pt}_{14}\text{Ni}_7$) magnetic coating, 20-nm thick amorphous carbon or diamondlike carbon (DLC) coating (microhardness \sim 1500 kg/mm² as measured using a Berkovich indenter), and

with or without a top layer of perfluoropolyether lubricant with polar end groups (Z-DOL) coating. The thickness of the lubricant film is about 2 nm. The metal particle (MP) tape is a 12.7-mm wide and 13.2- μm thick (PET base thickness of 9.8 μm , magnetic coating of 2.9 μm with metal magnetic particles and nonmagnetic particles of Al_2O_3 and Cr_2O_3 , and back coating of 0.5 μm). The barium ferrite (BaFe) tape is a 12.7-mm wide and 11 μm thick (PET base thickness of 7.3 μm , magnetic coating of 2.5 μm with barium ferrite magnetic particles and nonmagnetic particles of Al_2O_3 , and back coating of 1.2 μm). Metal evaporated (ME) tape is a 12.7-mm wide tape with 10- μm thick base, 0.2- μm thick evaporated Co–Ni magnetic film and about 10-nm thick perfluoropolyether lubricant and a back-coat. PET film is a biaxially-oriented, semicrystalline polymer with particulates. Two sizes of nearly spherical particulates are generally used in the construction of PET: submicron ($\sim 0.5 \mu\text{m}$) particles of typically carbon and larger particles (2–3 μm) of silica.

15.3 Surface Roughness

Solid surfaces, irrespective of the method of formation, contain surface irregularities or deviations from the prescribed geometrical form. When two nominally flat surfaces are placed in contact, surface roughness causes contact to occur at discrete contact points. Deformation occurs at these points, and may be either elastic or plastic, depending on the nominal stress, surface roughness and material properties. The sum of the areas of all the contact points constitutes the real area that would be in contact, and for most materials at normal loads, this will be only a small fraction of the area of contact if the surfaces were perfectly smooth. In general, real area of contact must be minimized to minimize adhesion, friction and wear (Bhushan 2013).

Characterizing surface roughness is therefore important for predicting and understanding the tribological properties of solids in contact. Various measurement techniques are used to measure surface roughness. The AFM is used to measure

surface roughness on length scales from nanometers to micrometers. A second technique is noncontact optical profiler (NOP) which is a noncontact technique and does not damage the surface. The third technique is stylus profiler (SP) in which a sharp tip is dragged over the sample surface. These techniques differ in lateral resolution. Roughness plots of a glass-ceramic disk measured using an AFM (lateral resolution $\sim 15 \text{ nm}$), NOP (lateral resolution $\sim 1 \mu\text{m}$) and SP (lateral resolution of $\sim 0.2 \mu\text{m}$) are shown in Fig. 15.8a. Figure 15.8b compares the profiles of the disk obtained with different instruments at a common scale. The figures show that roughness is found at scales ranging from millimeter to nanometer scales. Measured roughness profile is dependent on the lateral and normal resolutions of the measuring instrument (Bhushan and Blackman 1991; Oden et al. 1992; Ganti and Bhushan 1995; Poon and Bhushan 1995a, b). Instruments with different lateral resolutions measure features with different scale lengths. It can be concluded that a surface is composed of a large number of length of scales of roughness that are superimposed on each other.

Surface roughness is most commonly characterized by the standard deviation of surface heights which is the square roots of the arithmetic average of squares of the vertical deviation of a surface profile from its mean plane. Due to the multiscale nature of surfaces, it is found that the variances of surface height and its derivatives and other roughness parameters depend strongly on the resolution of the roughness measuring instrument or any other form of filter, hence not unique for a surface (Ganti and Bhushan 1995; Poon and Bhushan 1995a, b; Koinkar and Bhushan 1997a), see for example Fig. 15.9. Therefore, a rough surface should be characterized in a way such that the structural information of roughness at all scales is retained. It is necessary to quantify the multiscale nature of surface roughness.

A unique property of rough surfaces is that if a surface is repeatedly magnified, increasing details of roughness are observed right down to nanoscale. In addition, the roughness at all magnifications appear quite similar in structure, as qualitatively shown in Fig. 15.10. The

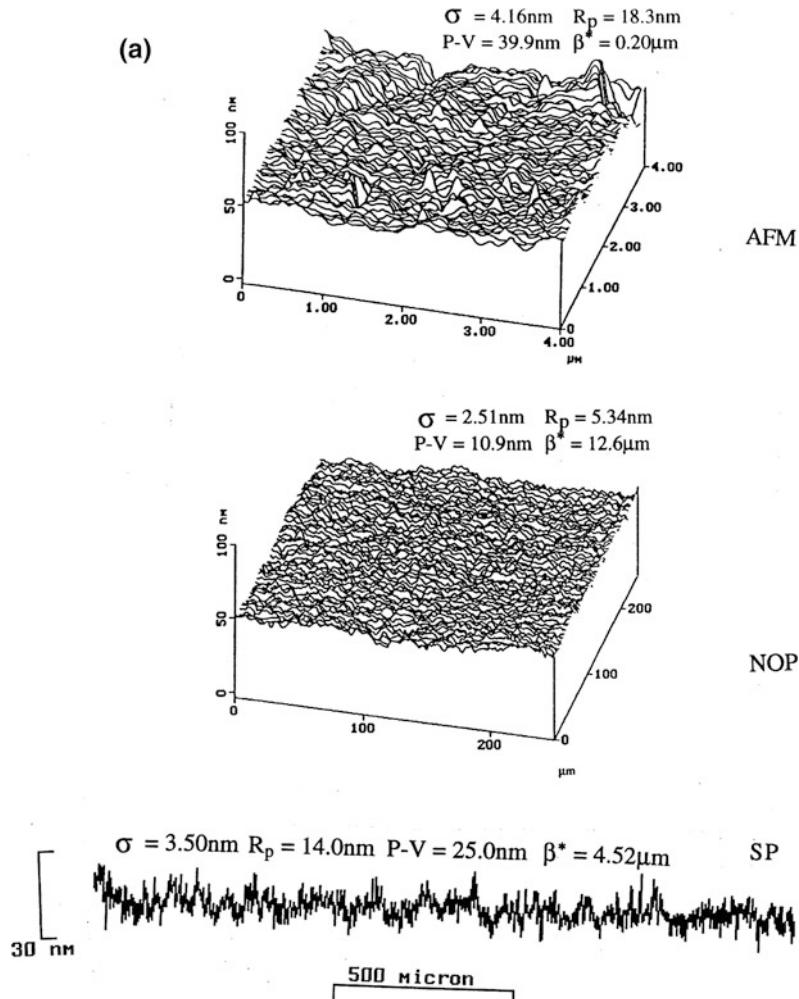


Fig. 15.8 Surface roughness plots of a glass-ceramic disk (a) measured using an atomic force microscope (lateral resolution ~ 15 nm), noncontact optical profiler (NOP) (lateral resolution ~ 1 μm) and stylus profiler (SP) with a stylus tip of 0.2- μm radius (lateral resolution ~ 0.2 μm), and (b) measured using an AFM (~ 150 nm), SP (~ 0.2 μm), and NOP (~ 1 μm) and plotted on a common scale (Poon and Bhushan 1995a)

statistical self-affinity is due to similarity in appearance of a profile under different magnifications. Such a behavior can be characterized by fractal analysis (Majumdar and Bhushan 1990; Ganti and Bhushan 1995). The main conclusion from these studies are that a fractal characterization of surface roughness is *scale independent* and provides information of the roughness structure at all length scales that exhibit the fractal behavior.

Structure function and power spectrum of a self-affine fractal surface follow a power law and can be written as (Ganti and Bhushan 1995)

$$S(\tau) = C\eta^{(2D-3)}\tau^{(4-2D)} \quad (15.1)$$

$$P(\omega) = \frac{c_1\eta^{(2D-3)}}{\omega^{(5-2D)}} \quad (15.2a)$$

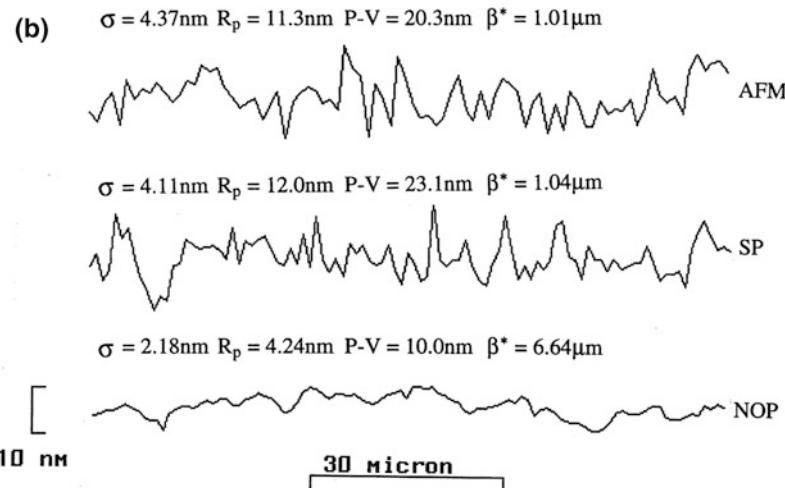


Fig. 15.8 (continued)

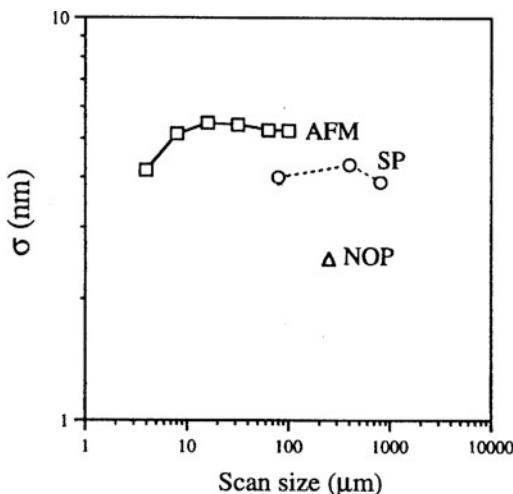


Fig. 15.9 Scale dependence of standard deviation of surface heights for a glass-ceramic disk, measured using atomic force microscope (AFM), stylus profiler (SP), and noncontact optical profiler (NOP)

and

$$c_1 = \frac{\Gamma(5 - 2D) \sin[\pi(2 - D)]}{2\pi} C \quad (15.2b)$$

The fractal analysis allows the characterization of surface roughness by two parameters D and C which are instrument-independent and unique for each surface. D (ranging from 1 to 2 for surface profile) primarily relates to relative

power of the frequency contents and C to the amplitude of all frequencies. η is the lateral resolution of the measuring instrument, τ , is the size of the increment (distance), and ω is the frequency of the roughness. Note that if $S(\tau)$ or $P(\omega)$ are plotted as a function of τ or ω , respectively, on a log-log plot, then the power law behavior would result into a straight line. The slope of line is related to D and the location of the spectrum along the power axis is related to C.

Figure 15.11 present the structure function of a thin-film rigid disk measured using AFM, non-contact optical profiler (NOP), and stylus profiler (SP). Horizontal shift in the structure functions from one scan to another, arises from the change in the lateral resolution. D and C values for various scan lengths are listed in Table 15.1. We note that fractal dimension of the various scans is fairly constant (1.26–1.33); however, C increases/decreases monotonically with σ for the AFM data. The error in estimation of η is believed to be responsible for variation in C. These data show that the disk surface follows a fractal structure for three decades of length scales.

Majumdar and Bhushan (1991) and Bhushan and Majumdar (1992) developed a fractal theory of contact between two rough surfaces. This model has been used to predict whether contacts experience elastic or plastic deformation and to predict the statistical distribution of contact

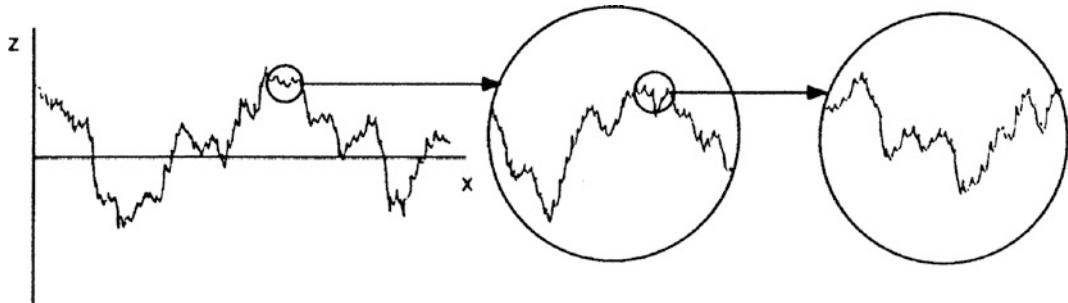


Fig. 15.10 Qualitative description of statistical self-affinity for a surface profile

points. For a review of contact models, see Bhushan (1996b, 1998) and Bhushan and Peng (2002).

Based on the fractal model of elastic-plastic contact, whether contacts go through elastic or plastic deformation is determined by a critical area which is a function of D, C, hardness and modulus of elasticity of the mating surfaces. If contact spot is smaller than the critical area, it goes through the plastic deformations and large spots go through elastic deformations. The critical contact area for inception of plastic deformation for a thin-film disk was reported by

Majumdar and Bhushan (1991) to be about 10^{-27} m^2 , so small that all contact spots can be assumed to be elastic at moderate loads.

The question remains as to how large spots become elastic when they must have initially been plastic spots. The possible explanation is shown in Fig. 15.12. As two surfaces touch, the nanoasperities (detected by AFM type of instruments) first coming into contact have smaller radii of curvature and are therefore plastically deformed instantly, and the contact area increases. When load is increased, nanoasperities in the contact merge, and the load is supported by

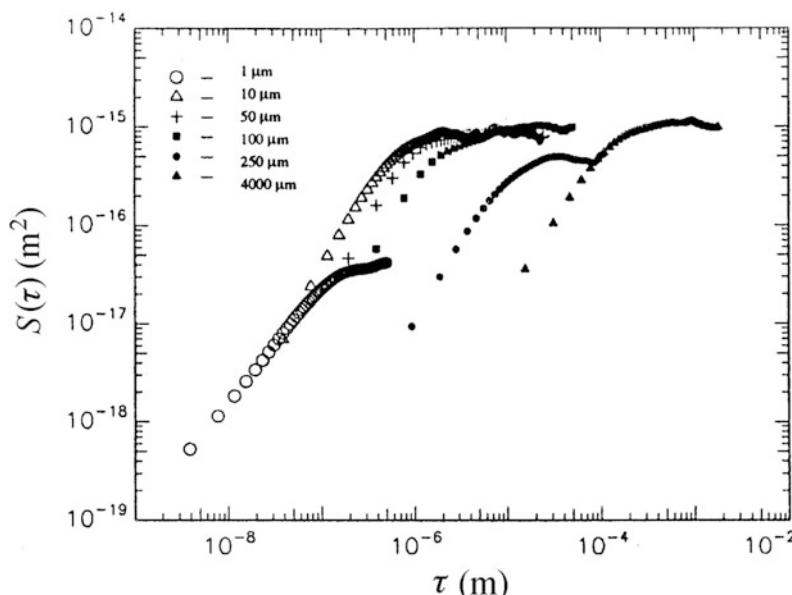


Fig. 15.11 Structure functions for the roughness data measured at various scan sizes using AFM (scan sizes: 1 $\mu\text{m} \times 1 \mu\text{m}$, 10 $\mu\text{m} \times 10 \mu\text{m}$, 50 $\mu\text{m} \times 50 \mu\text{m}$, and 100 μm), NOP (scan size: 250 $\mu\text{m} \times 250 \mu\text{m}$), and SP (scan length: 4000 μm), for a magnetic thin-film rigid disk (Ganti and Bhushan 1995)

Table 15.1 Surface roughness parameters for a polished thin-film rigid disk (Ganti and Bhushan 1995)

Scan size ($\mu\text{m} \times \mu\text{m}$)	σ (nm)	D	C (nm)
1 (AFM)	0.7	1.33	9.8×10^{-4}
10 (AFM)	2.1	1.31	7.6×10^{-3}
50 (AFM)	4.8	1.26	1.7×10^{-2}
100 (AFM)	5.6	1.30	1.4×10^{-2}
250 (NOP)	2.4	1.32	2.7×10^{-4}
4000 (NOP)	3.7	1.29	7.9×10^{-5}

AFM Atomic force microscope

NOP Noncontact optical profiler

elastic deformation of the large scale asperities or microasperities (detected by optical profiler type of instruments) (Bhushan and Blackman 1991).

Majumdar and Bhushan (1991) and Bhushan and Majumdar (1992) have reported relationships for cumulative size distribution of the contact spots, portions of the real area of contact in elastic and plastic deformation modes, and the load-area relationships.

15.4 Friction and Adhesion

Friction and adhesion of magnetic head sliders and magnetic media have been measured by Bhushan and Koinkar (1994a, 1995a, b, c, d), Bhushan and Ruan (1994), Ruan and Bhushan (1994a), Bhushan et al. (1995b), Bhushan (1996a, 1999a, 2001b), Koinkar and Bhushan (1996a, b, 1997a, b), Kulkarni and Bhushan (1997), and Li and Bhushan (1999a, b), and Sundararajan and Bhushan (1999).

Koinkar and Bhushan (1996a, b) and Poon and Bhushan (1995a, b) reported that RMS roughness and friction force increase with an increase in scan size at a given scanning velocity and normal force. Therefore, it is important that while reporting friction force values, scan sizes and scanning velocity should be mentioned. Bhushan and Sundararajan (1998) reported that friction and adhesion forces are a function of tip radius and relative humidity (also see Koinkar and Bhushan 1996b). Therefore, relative humidity should be controlled during the experiments. Care also should be taken to ensure that tip radius does not change during the experiments.

15.4.1 Magnetic Head Materials

Al_2O_3 –TiC is a commonly used slider material. In order to study the friction characteristics of this two phase material, friction of Al_2O_3 –TiC (70–30wt%) surface was measured. Figure 15.13 shows the surface roughness and friction force profiles (Koinkar and Bhushan 1996a). TiC grains have a Knoop hardness of about 2800 kg/mm^2 which is higher than that of Al_2O_3 grains of about 2100 kg/mm^2 . Therefore, TiC grains do not polish as much and result in a slightly higher elevation (about 2–3 nm higher than that of Al_2O_3 grains). Based on friction force measurements, TiC grains exhibit higher friction force than Al_2O_3 grains. The coefficients

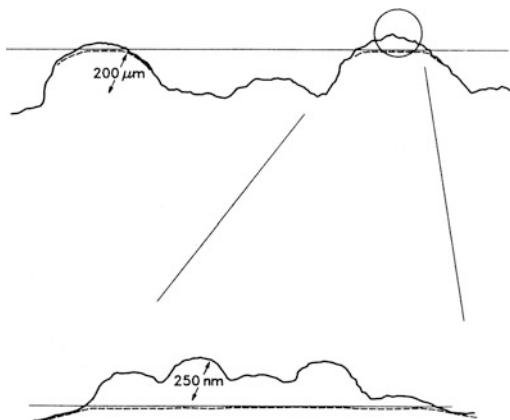


Fig. 15.12 Schematic of local asperity deformation during contact of a rough surface, upper profile measured by an optical profiler and lower profile measured by AFM, typical dimensions are shown for a polished thin-film rigid disk against a flat head slider surface (Bhushan and Blackman 1991)

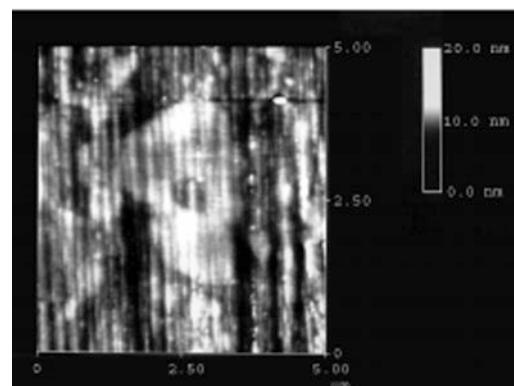
of friction of TiC and Al_2O_3 grains are 0.034 and 0.026, respectively and the coefficient of friction of Al_2O_3 -TiC composite is 0.03. Local variation in friction force also arises from the scratches present on the Al_2O_3 -TiC surface. Thus, local friction values of a two phase materials can be measured. Ruan and Bhushan (1994b) reported that local variation in the coefficient of friction of cleaved HOPG graphite was significant which arises from structural changes occurring during the cleaving process. The cleaved HOPG surface is largely atomically smooth but exhibits line shaped regions in which the coefficient of friction is more than an order of magnitude larger. These measurements suggest that friction measurements can be used for structural mapping of the surfaces.

Surface roughness and coefficient of friction of various head slider materials were measured by Koinkar and Bhushan (1996a). For typical values, see Table 15.2. Macroscale friction values for all samples are higher than microscale friction values as there is less plowing contribution in microscale measurements (Bhushan 1999a, 2013, 2017).

15.4.2 Magnetic Media

Bhushan and coworkers measured friction properties of magnetic media including polished and textured thin-film rigid disks, metal particle (MP), barium ferrite (BaFe) and metal evaporated (ME) tapes, and poly(ethylene terephthalate)(PET) tape substrate (Bhushan 1996a, 1999a, 2001b). For typical values of coefficients of friction of thin-film rigid disks and MP, BaFe and ME tapes, PET tape substrate, see Table 15.3. In the case of magnetic disks, similar coefficients of friction are observed for both lubricated and unlubricated disks, indicating that most of the lubricant (though partially thermally bonded) is squeezed out from between the rubbing surfaces at high interface pressures, consistent with liquids being poor boundary lubricant (Bhushan 2013). Coefficient of friction values on a microscale are much lower than those on the macroscale. When measured for the small

Surface Height



Friction Force

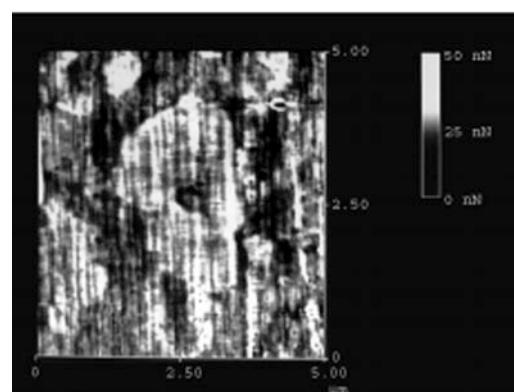


Fig. 15.13 Gray scale surface roughness ($\sigma = 0.97 \text{ nm}$) and friction force map (mean = 7.0 nN, $\sigma = 0.60 \text{ nN}$) for Al_2O_3 -TiC (70 to 30 wt%) at a normal load of 166 nN

contact areas and very low loads used in microscale studies, indentation hardness and modulus of elasticity are higher than at the macroscale (data to be presented later). This reduces the real area of contact and the degree of wear. In addition, the small apparent areas of contact reduces the number of particles trapped at the interface, and thus minimize the “plowing” contribution to the friction force (Bhushan and Koinkar 1994a; Bhushan et al. 1994, 1995a).

Local variations in the microscale friction of rough surfaces can be significant. Figure 15.14 shows the surface height map, the slopes of surface map taken along the sliding direction, and friction force map for a thin-film disk (Bhushan and Ruan 1994; Bhushan et al. 1994, 1995a; Bhushan and Koinkar 1995a, b, c, 1997;

Table 15.2 Surface roughness (σ and P-V distance), micro- and macro-scale friction, microscratching/wear, and nano- and microhardness data for various samples (Koinkar and Bhushan 1996a)

Sample	Surface roughness		Coefficient of friction			Scratch depth at 60 μN (nm)	Wear depth at 60 μN (nm)	Hardness			
	σ (nm)	P-V ^a (nm)	Microscale	Macroscale ^b				(Gpa)	(Gpa)		
				Initial	Final						
	(1 μm \times 1 μm)										
Al ₂ O ₃	0.97	9.9	0.03	0.18	0.2–0.6	3.2	3.7	24.8	15.0		
Al ₂ O ₃ -TiC	0.80	9.1	0.05	0.24	0.2–0.6	2.8	22.0	23.6	20.2		
Polycrystalline Mn-Zn Ferrite	2.4	20.0	0.04	0.27	0.24–0.4	9.6	83.6	9.6	5.6		
Single-crystal (110) Mn-Zn Ferrite	1.9	13.7	0.02	0.16	0.18–0.24	9.0	56.0	9.8	5.6		
SiC (α -type)	0.91	7.2	0.02	0.29	0.18–0.24	0.4	7.7	26.7	21.8		

^aPeak-to-Valley distance^bObtained using silicon nitride ball with 3 mm diameter in a reciprocating mode at a normal load of 10 mN, reciprocating amplitude of 7 mm and average sliding speed of 1 mm/s. Initial coefficient of friction values were obtained at first cycle (0.007 m sliding distance) and final values at a sliding distance of 5 m

Koinkar and Bhushan 1996a, 1997a; Sundararajan and Bhushan 2000). We note that there is no resemblance between the coefficient of friction profiles and the corresponding roughness maps, e.g., high or low points on the friction profile do not correspond to high or low points on the roughness profiles. By comparing the slope and friction profiles, we observe a strong correlation between the two. (For a clearer correlation, see gray-scale plots of slope and friction profiles for FFM tip sliding in either directions, in Fig. 15.15 to be presented in the next paragraph). We have shown that this correlation holds for various magnetic tapes, silicon, diamond, and other materials. This correlation can be explained by a “ratchet” mechanism; based on this mechanism, the local friction is a function of the local slope of the sample surface (Bhushan 1999a, 2013, 2017). The friction is high at the leading edge of asperities and low at the trailing edge. In addition to the slope effect, the collision of tip encountering an asperity with a positive slope produces additional torsion of the cantilever beam leading to higher measured friction force. When encountering an asperity with a negative slope, however, there is no collision effect and hence no effect on friction. The ratchet mechanism and the collision effects thus explain the correlation

between the slopes of the roughness maps and friction maps observed in Fig. 15.14.

To study the directionality effect on the friction, gray scale plots of coefficients of local friction of the thin-film disk as the FFM tip is scanned in either direction are shown in Fig. 15.15. Corresponding gray scale plots of slope of roughness maps are also shown in Fig. 15.15. The left hand set of the figures corresponds to the tip sliding from the left towards right (trace direction), and the middle set corresponds to the tip sliding from the right towards left (retrace direction). It is important to take into account the sign change of surface slope and friction force which occur in the trace and retrace directions. In order to facilitate comparison of directionality effect on friction, the last set of the figures in the right hand column show the data with sign of surface slope and friction data in the retrace direction reversed. Now we compare trace and -retrace data. It is clear that the friction experienced by the tip is dependent upon the scanning direction because of surface topography.

The directionality effect in friction on a macroscale is generally averaged out over a large number of contacts. It has been observed in some magnetic tapes. In a macro-scale test, a 12.7-mm

Table 15.3 Surface roughness (σ), microscale and macro-scale friction, and nanohardness data of thin-film magnetic rigid disk, magnetic tape and magnetic tape substrate (PET) samples

Sample	σ (nm)			Coefficient of microscale friction		Coefficient of macro-scale friction		Nanohardness (Gpa)/normal load (μN)	
	NOP		AFM						
	250 $\mu\text{m} \times$ 250 μm^{a}	1 $\mu\text{m} \times$ 1 μm^{a}	10 $\mu\text{m} \times$ 10 μm^{a}	1 $\mu\text{m} \times$ 1 μm^{a}	10 $\mu\text{m} \times$ 10 μm^{a}	Mn-Zn ferrite	Al_2O_3 -TiC		
Polished, unlubricated disk	2.2	3.3	4.5	0.05	0.06	—	0.26	21/100	
Polished, lubricated disk	2.3	2.3	4.1	0.04	0.05	—	0.19	—	
Textured, lubricated disk	4.6	5.4	8.7	0.04	0.05	—	0.16	—	
Metal-particle tape	6.0	5.1	12.5	0.08	0.06	0.19	—	0.30/50	
Barium-ferrite tape	12.3	7.0	7.9	0.07	0.03	0.18	—	0.25/25	
Metal-evaporated tape	9.3	4.7	5.1	0.05	0.03	0.18	—	0.7 to 4.3/75	
PET tape substrate	33	5.8	7.0	0.05	0.04	0.55	—	0.3/20 and 1.4/20 ^b	

^aScan area; NOP Noncontact optical profiler; AFM Atomic force microscope

^bNumbers are for polymer and particulate regions, respectively

wide MP tape was wrapped over an aluminum drum and slid in a reciprocating motion with a normal load of 0.5 N and a sliding speed of about 60 mm/s. Coefficient of friction as a function of sliding distance in either direction is shown in Fig. 15.16. We note that coefficient of friction on a macro-scale for this tape is different in different directions.

15.5 Scratching and Wear

15.5.1 Nanoscale Wear

Bhushan and Ruan (1994) conducted nanoscale wear tests on MP tapes at a normal load of 100 nN. Figure 15.17 shows the topography of the MP tape obtained at two different loads. For a given normal load, measurements were made twice. There was no discernible difference between consecutive measurements for a given normal load. However, as the load increased from 10 to 100 nN, topographical changes were observed; material (indicated by an arrow) was pushed toward the right side in the sliding direction of the AFM tip relative to the sample. The material movement is believed to occur as a

result of plastic deformation of the tape surface. Similar behavior was observed on all tapes studied. Magnetic tape coating is made of magnetic particles and polymeric binder. Any movement of the coating material can eventually lead to loose debris. Debris formation is an undesirable situation as it may contaminate the head which may increase friction and/or wear between the head and tape, in addition to the deterioration of the tape itself. With disks, they did not notice any deformation under a 100 nN normal load.

15.5.2 Microscale Scratching

Microscratches have been made on various potential head slider materials (Al_2O_3 , Al_2O_3 -TiC, Mn-Zn ferrite and SiC), and various magnetic media (unlubricated polished thin-film disk, MP, BaFe, ME tapes, and PET substrates) and virgin, treated and coated Si(111) wafers at various loads (Bhushan and Koinkar 1994a, 1995a, b, c, d, 1997; Bhushan et al. 1994; Koinkar and Bhushan 1996a, 1997b; Sundararajan and Bhushan 1998, 1999). As mentioned earlier, the scratches are made using a diamond tip.

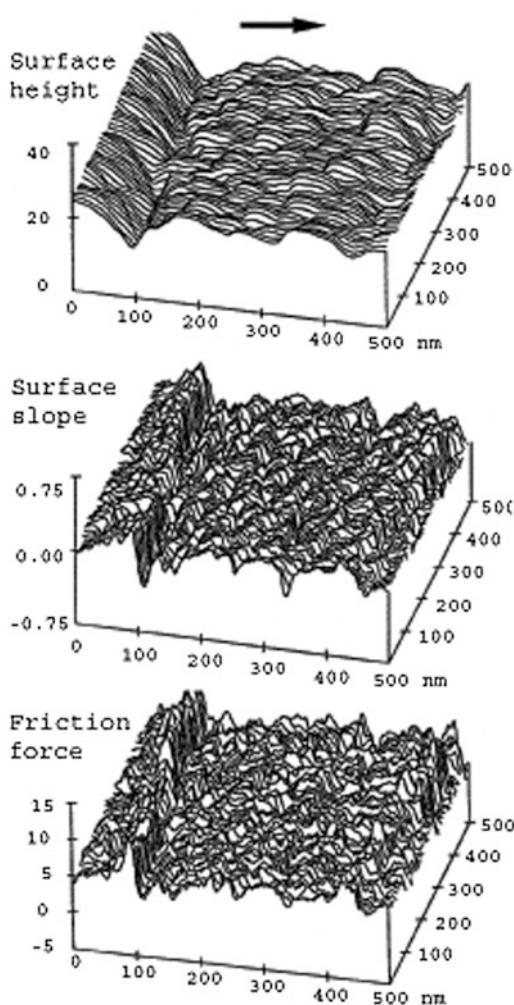


Fig. 15.14 (a) Surface height map ($\sigma = 4.4 \text{ nm}$), (b) slope of the roughness profiles taken in the sample sliding direction (the horizontal axis) (mean = 0.023, $\sigma = 0.18$), and (c) friction force map (mean = 6.2 nN, $\sigma = 2.1 \text{ nN}$) for a textured thin-film rigid disk at a normal load of 160 nN (Bhushan et al. 1994)

15.5.2.1 Magnetic Head Materials

Scratch depths as a function of load and representative scratch profiles with corresponding 2-D gray scale plots at various loads after a single pass (unidirectional scratching) for Al_2O_3 , $\text{Al}_2\text{O}_3\text{-TiC}$, polycrystalline and single-crystal Mn-Zn ferrite and SiC are shown in Figs. 15.18 and 15.19, respectively. Variation in the scratch depth along the scratch is about $\pm 15\%$. The Al_2O_3 surface could be scratched at a normal load of 40 μN . The surface topography of

polycrystalline Al_2O_3 shows the presence of porous holes on the surface. The 2-D gray scale plot of scratched Al_2O_3 surface shows one porous hole between scratches made at normal loads of 40 and 60 μN . Regions with defects or porous holes present, exhibit lower scratch resistance (see region marked by the arrow on 2-D gray scale plot of Al_2O_3). The $\text{Al}_2\text{O}_3\text{-TiC}$ surface could be scratched at a normal load of 20 μN . The scratch resistance for TiC grains is higher than that of Al_2O_3 grains. The scratches generated at normal loads of 80 and 100 μN show that scratch depth of Al_2O_3 grains is higher than that of TiC grains (see corresponding gray scale plot for $\text{Al}_2\text{O}_3\text{-TiC}$). Polycrystalline and single-crystal Mn-Zn ferrite could be scratched at a normal load of 20 μN . The scratch width is much larger for the ferrite specimens as compared with other specimens. For SiC, there is no measurable scratch observed at a normal load of 20 μN . At higher normal loads, very shallow scratches are produced. Table 15.2 presents average scratch depth at 60 μN normal load for all specimens. SiC has the highest scratch resistance followed by $\text{Al}_2\text{O}_3\text{-TiC}$, Al_2O_3 and polycrystalline and single-crystal Mn-Zn ferrite. Polycrystalline and single-crystal Mn-Zn ferrite specimens exhibit comparable scratch resistance.

15.5.2.2 Magnetic Media

Scratch depths as a function of load and scratch profiles at various loads after ten scratch cycles for unlubricated, polished disk and MP tape are shown in Figs. 15.20 and 15.21. We note that scratch depth increases with an increase in the normal load. Tape could be scratched at about 100 nN. With disk, gentle scratch marks under 10 μN load were barely visible. It is possible that material removal did occur at lower load on an atomic scale which was not observable with a scan size of 5 μm square. For disk, scratch depth at 40 μN is less than 10 nm deep. The scratch depth increased slightly at the load of 50 μN . Once the load is increased in excess of 60 μN , the scratch depth increased rapidly. These data suggest that the carbon coating on the disk surface is much harder to scratch than the underlying thin-film magnetic film. This is expected

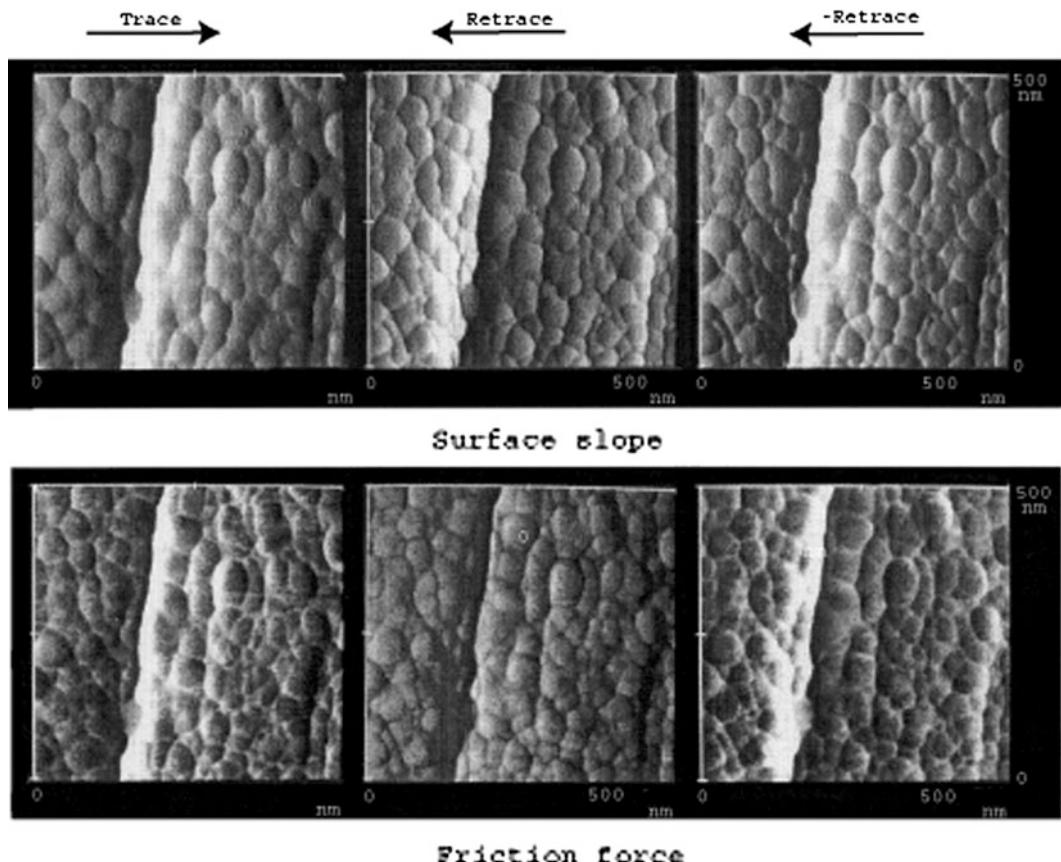


Fig. 15.15 Gray-scale plots of the slope of the surface roughness and the friction force maps for a textured thin-film rigid disk with FFM tip sliding in different directions. Higher points are shown by *lighter color*

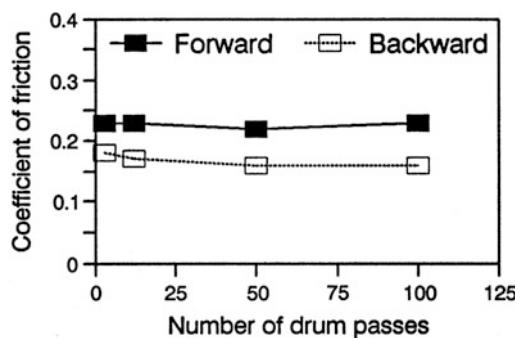


Fig. 15.16 Coefficient of macro-scale of friction as a function of sliding cycles for a metal-particle (MP) tape sliding over an aluminum drum in a reciprocating mode in both directions. Normal load = 0.5 N over 12.7 mm wide tape, sliding speed = 60 mm/s (Bhushan 1995b)

since the carbon coating is harder than the magnetic material used in the construction of the disks.

Microscratch characterization of ultrathin amorphous carbon coatings, deposited by filtered cathodic arc (FCA) direct ion beam (IB), electron cyclotron resonance plasma chemical vapor deposition (ECR-CVD), and sputter (SP) deposition processes have been conducted using a nanoindenter and an AFM (Bhushan and Koinkar 1995d; Koinkar and Bhushan 1997b; Bhushan 1999b; Li and Bhushan 1999a, b; Sundararajan and Bhushan 1999, 2001). Data on various coatings of different thicknesses using a Berkovich tip are compared in Fig. 15.22. Critical loads for various coatings and silicon substrate are summarized in Fig. 15.23. It is clear that, for all deposition methods, the critical load increases with increasing coating thickness due to better load-carrying capacity of thicker coatings as compared to the thinner ones. In

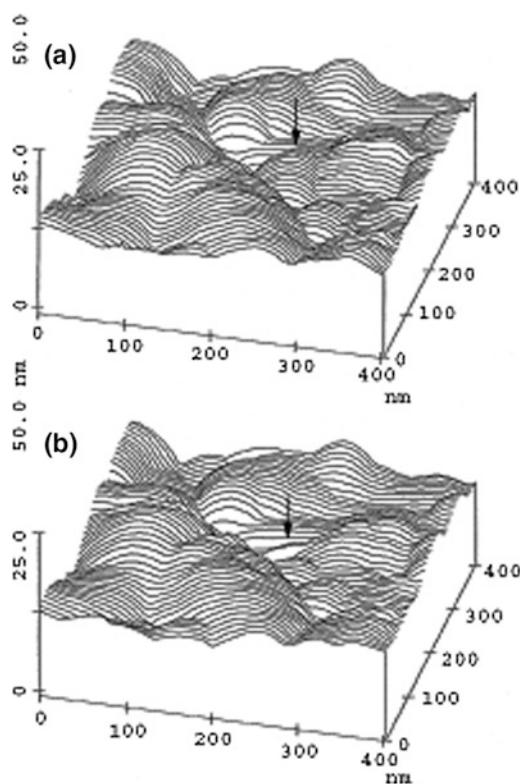


Fig. 15.17 Surface roughness maps of a metal-particle (MP) tape at applied normal load of (a) 10 nN and (b) 100 nN. Location of the change in surface topography as a result of nanowear is indicated by arrows (Bhushan and Ruan 1994)

comparison of the different deposition methods, ECR-CVD and FCA coatings show superior scratch resistance at 20- and 10-nm thicknesses compared to SP coatings. As the coating thickness reduces, ECR-CVD exhibits the best scratch resistance followed by FCA and SP coatings.

Since tapes scratch readily, for comparisons in scratch resistance of various tapes, Bhushan and Koinkar (1995c) made scratches on selected three tapes only with one cycle. Figure 15.24 presents the scratch depths as a function of normal load after one cycle for three tapes—MP, BaFe and ME tapes. For the MP and BaFe particulate tapes, Bhushan and Koinkar (1995c) noted that the scratch depth along (parallel) and across (perpendicular) the longitudinal direction of the tapes is similar. Between the two tapes, MP tape appears to be more scratch resistant than BaFe tape, which depends on the binder, pigment volume concentration (PVC) and the head cleaning agent (HCA) contents. ME tapes appear to be much more scratch resistant than the particulate tapes. However, the ME tape breaks up catastrophically in a brittle mode at a normal load higher than the 50 μ N (Fig. 15.25), as compared to particulate tapes in which the scratch rate is constant. They reported that the hardness of ME tapes is higher than that of particulate tapes, however, a significant difference in the

Fig. 15.18 Scratch depth as a function of normal load after one unidirectional cycle for Al_2O_3 , Al_2O_3 -TiC, polycrystalline Mn-Zn ferrite, single-crystal Mn-Zn ferrite and SiC (Koinkar and Bhushan 1996a)

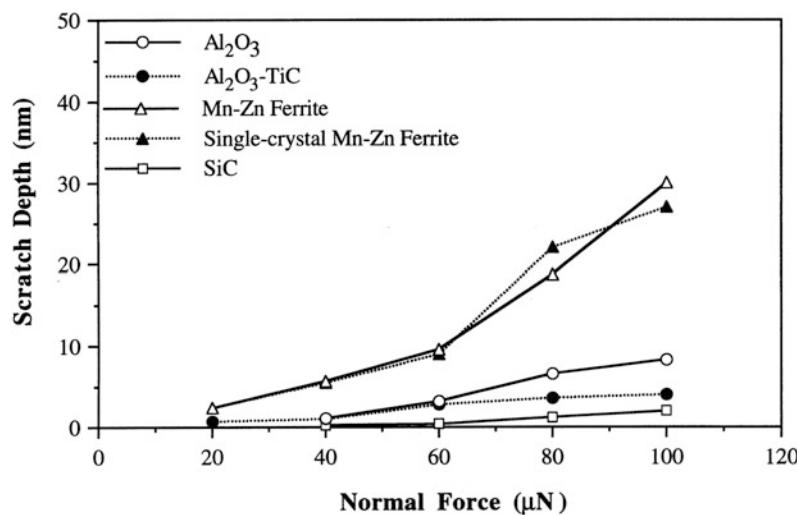
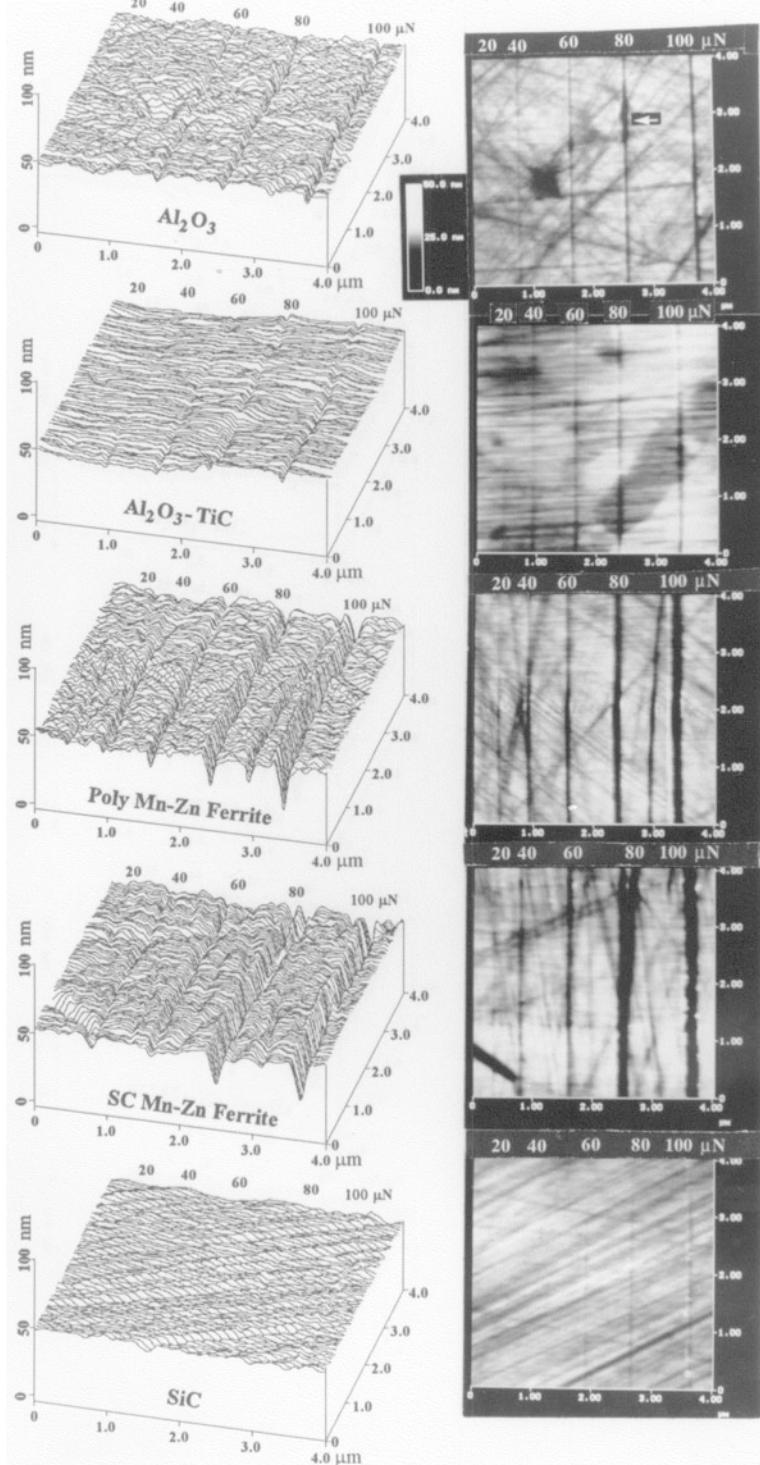


Fig. 15.19 Surface profiles (left column) and 2-D gray scale plots (right column) of scratched Al_2O_3 , $\text{Al}_2\text{O}_3\text{-TiC}$, polycrystalline Mn-Zn ferrite, single-crystal Mn-Zn ferrite, and SiC surfaces. Normal loads used for scratching for one unidirectional cycle are listed in the figure (Koinkar and Bhushan 1996a)



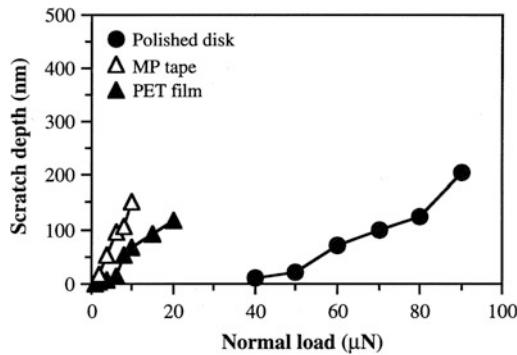


Fig. 15.20 Scratch depth as a function of normal load after ten scratch cycles for unlubricated polished thin-film rigid disk, MP tape and PET film (Bhushan et al. 1994; Bhushan and Koinkar 1995a, c)

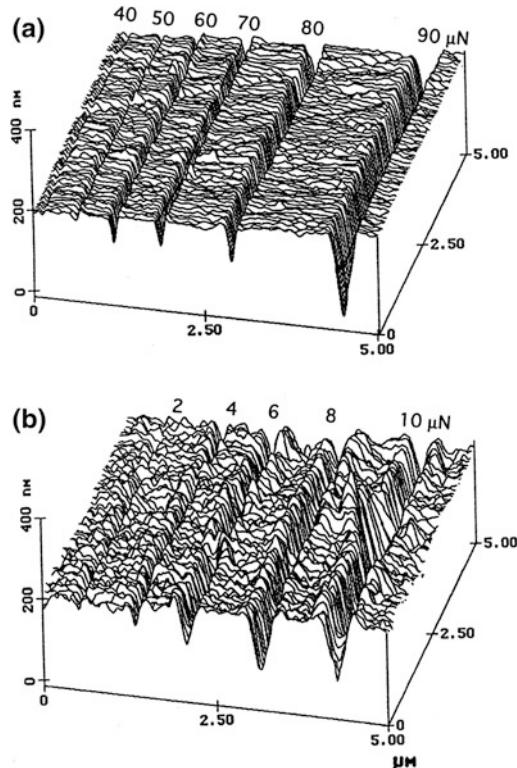


Fig. 15.21 Surface profiles for scratched (a) unlubricated polished thin-film rigid disk and (b) MP tape. Normal loads used for scratching for ten cycles are listed in the figure (Bhushan et al. 1994)

nanoindentation hardness values of the ME film from region to region (Table 15.3) was observed. They systematically measured scratch resistance

in the high and low hardness regions along and across the longitudinal directions. Along the parallel direction, load required to crack the coating was lower (implying lower scratch resistance) for a harder region, than that for a softer region. The scratch resistance of high hardness region along the parallel direction is slightly poorer than that for along perpendicular direction. Scratch widths in both low and high hardness regions is about half ($\sim 2 \mu\text{m}$) than that in perpendicular direction ($\sim 1 \mu\text{m}$). In the parallel direction, the material is removed in the form of chips and lateral cracking also emanates from the wear zone. ME films have columnar structure with the columns lined up with an oblique angle of on the order of about 35° with respect to the normal to the coating surface (Hibst 1993; Bhushan 2000). The column orientation may be responsible for directionality effect on the scratch resistance. Hibst (1993) have reported the directionality effect in the ME tape-head wear studies. They have found that the wear rate is lower when the head moves in the direction corresponding to the column orientation than in the opposite direction.

PET films could be scratched at loads of as low as about $2 \mu\text{N}$, Fig. 15.26. Figure 15.26a shows scratch marks made at various loads. Scratch depth along the scratch does not appear to be uniform. This may occur because of variations in the mechanical properties of the film. Bhushan and Koinkar (1995a) also conducted scratch studies in the selected particulate regions. Scratch profiles at increasing loads in the particulate region are shown in Fig. 15.26b. We note that the bump (particle) is barely scratched at $5 \mu\text{N}$ and it can be scratched readily at higher loads. At $20 \mu\text{N}$, it essentially disappears.

15.5.3 Microscale Wear

By scanning the sample (in 2D) while scratching, wear scars are generated on the sample surface (Bhushan and Koinkar 1994a, 1995a, b, c, d, 1997; Bhushan et al. 1994; Koinkar and Bhushan 1996a, 1997b; Bhushan and Sundararajan 1998; Sundararajan and Bhushan 1998, 1999). The

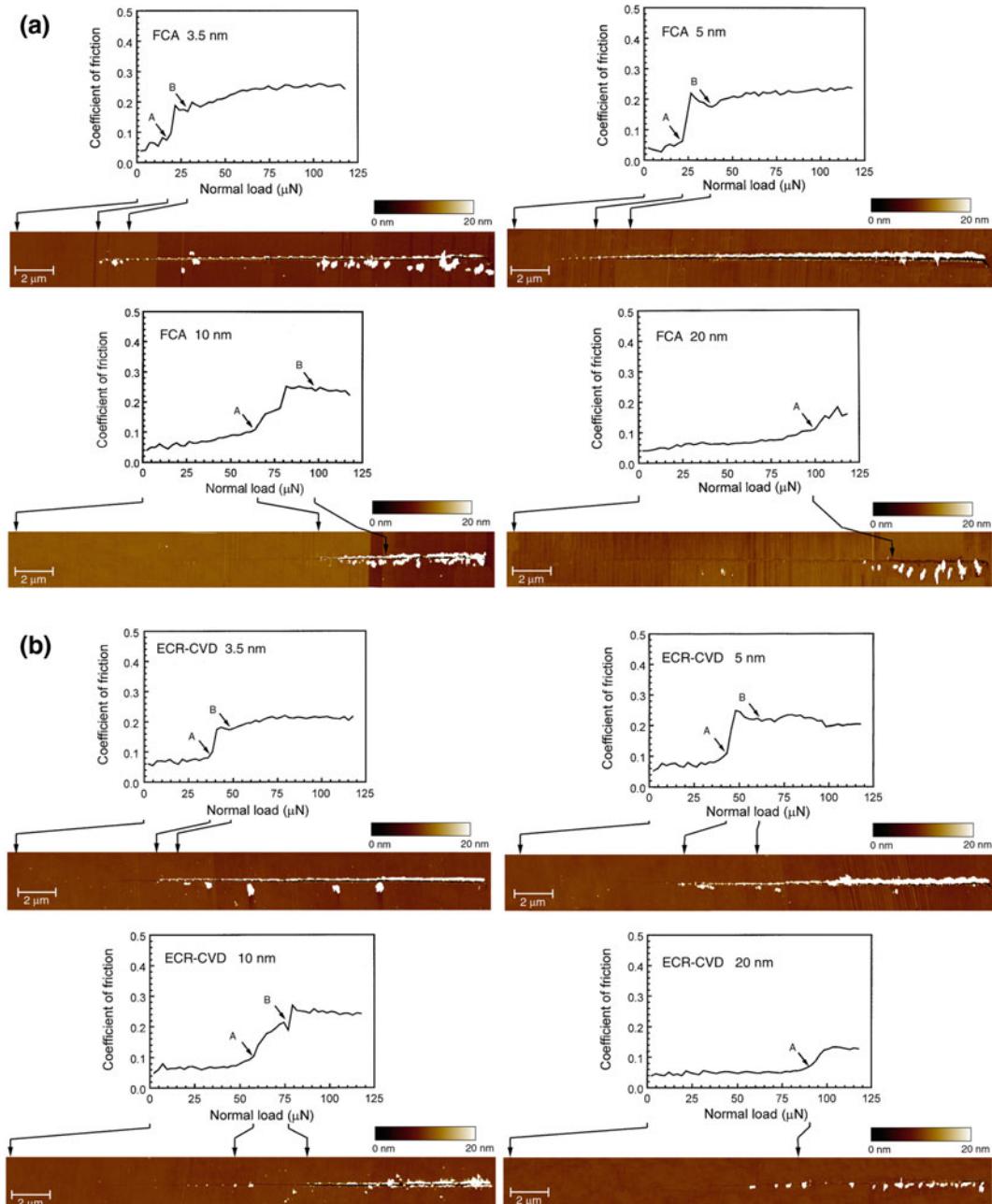


Fig. 15.22 Coefficient of friction profiles during scratch as a function of normal load and corresponding AFM surface height images for (a) FCA, (b), ECR-CVD, and (c) SP-coatings (Sundararajan and Bhushan 2001)

major benefit of a single cycle wear test over a scratch test is that wear data can be obtained over a large area.

15.5.3.1 Magnetic Head Materials

Figure 15.27 shows the wear depth as a function of load for one cycle for different slider materials.

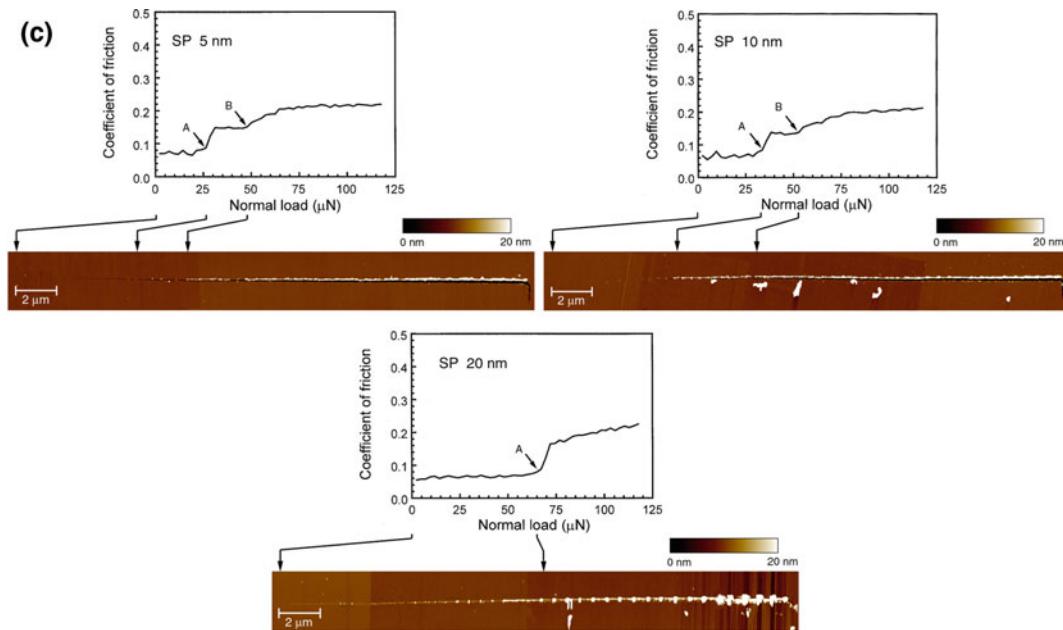


Fig. 15.22 (continued)

Variation in the wear depth in the wear mark is dependent upon the material. It is generally within $\pm 5\%$. The mean wear depth increases with the increase in normal load. The representative surface profiles showing the wear marks (central $2 \mu\text{m} \times 2 \mu\text{m}$ region) at a normal load of $60 \mu\text{N}$ for all specimens are shown in Fig. 15.28. The material is removed uniformly in the wear region for all specimens. Table 15.2 presents average wear depth at $60 \mu\text{N}$ normal load for all

specimens. Microwear resistance of SiC and Al_2O_3 is the highest followed by Al_2O_3 -TiC, single-crystal and polycrystalline Mn-Zn ferrite.

Next, wear experiments were conducted for multiple cycles. Figure 15.29 shows the 2-D gray scale plots and corresponding section plot (on top of each gray scale plot), taken at a location shown by an arrow for Al_2O_3 (left column) and Al_2O_3 -TiC (right column) specimen obtained at a normal load of $20 \mu\text{N}$ and at a different number of scan cycles. The central regions ($2 \mu\text{m} \times 2 \mu\text{m}$) show the wear mark generated after a different number of cycles. Note the difference in the vertical scale of gray scale and section plots. The Al_2O_3 specimen shows that wear initiates at the porous holes or defects present on the surface. Wear progresses at these locations as a function of number of cycles. In the porous hole free region, microwear resistance is higher. In the case of the Al_2O_3 -TiC specimen for about five scan cycles, the microwear resistance is higher at the TiC grains and is lower at the Al_2O_3 grains. The TiC grains are removed from the wear mark after five scan cycles. This indicates that micro-wear resistance of multi-phase materials depends upon the individual grain properties. Evolution

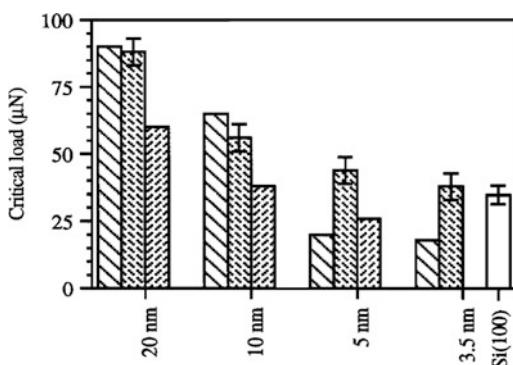


Fig. 15.23 Summary of critical loads estimated from the coefficient of friction profiles and AFM images (Sundararajan and Bhushan 2001)

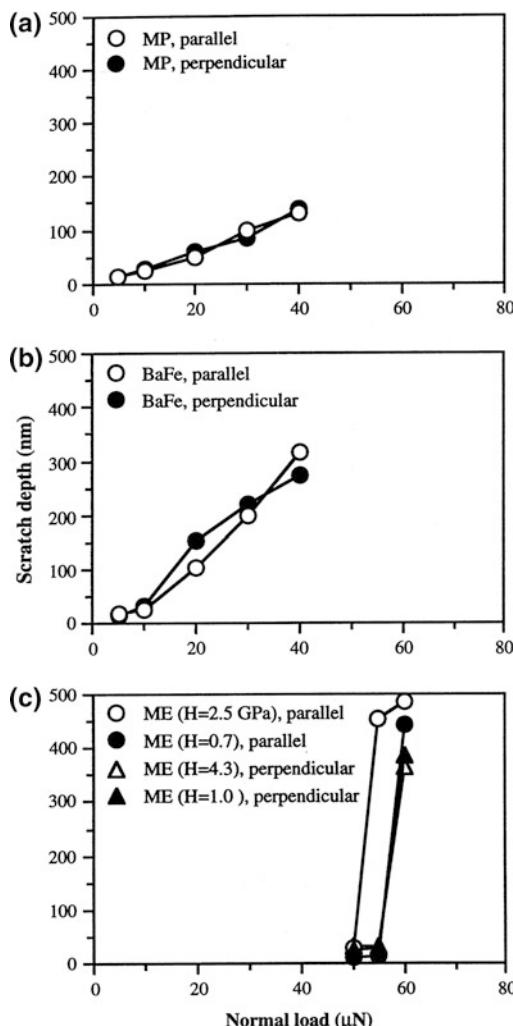


Fig. 15.24 Scratch depth as a function of normal load after one scratch cycle for (a) MP, (b) BaFe, and (c) ME tapes along parallel and perpendicular directions with respect to the longitudinal axis of the tape (Bhushan and Koinkar 1995c)

of wear is uniform within the wear mark for ferrite specimens. Figure 15.30 shows plot of wear depth as a function of number of cycles at a normal load of $20 \mu\text{N}$ for all specimens. The Al_2O_3 specimen reveals highest microwear resistance followed by SiC, Al_2O_3 –TiC, polycrystalline and single-crystal Mn–Zn ferrite. Wear resistance of Al_2O_3 –TiC is inferior to that of Al_2O_3 . Chu et al. (1992) studied friction and wear behavior of the single-phase and

multi-phase ceramic materials and found that wear resistance of multi-phase materials was poorer than single-phase materials. Multi-phase materials have more material flaws than the single-phase material. The differences in thermal and mechanical properties between the two phases may lead to cracking during processing, machining or use.

15.5.3.2 Magnetic Media

Figure 15.31 shows the wear depth as a function of load for one cycle for the polished, unlubricated and lubricated disks (Bhushan et al. 1994). Figure 15.32 shows profiles of the wear scars generated on unlubricated disk. The normal force for the imaging was about $0.5 \mu\text{N}$ and the loads used for the wear were $20, 50, 80$ and $100 \mu\text{N}$ as indicated in the figure. We note that wear takes place relatively uniformly across the disk surface and essentially independent of the lubrication for the disks studied. For both lubricated and unlubricated disks, the wear depth increases slowly with load at low loads with almost the same wear rate. As the load is increased to about $60 \mu\text{N}$, wear increases rapidly with load. The wear depth at $50 \mu\text{N}$ is about 14 nm , slightly less than the thickness of the carbon film. The rapid increase of wear with load at loads larger than $60 \mu\text{N}$ is an indication of the breakdown of the carbon coating on the disk surface.

Figure 15.33 shows the wear depth as a function of number of cycles for the polished disks (lubricated and unlubricated). Again, for both unlubricated and lubricated disks, wear initially takes place slowly with a sudden increase between 40 and 50 cycles at $10 \mu\text{N}$. The sudden increase occurred after 10 cycles at $20 \mu\text{N}$. This rapid increase is associated with the breakdown of the carbon coating. The wear profiles at various cycles are shown in Fig. 15.34 for a polished, unlubricated disk at a normal load of $20 \mu\text{N}$. Wear is not uniform and the wear is largely initiated at the texture grooves present on the disk surface. This indicates that surface defects strongly affect the wear rate.

Hard amorphous carbon coatings are used to provide wear and corrosion resistance to magnetic disks and MR/GMR magnetic heads.

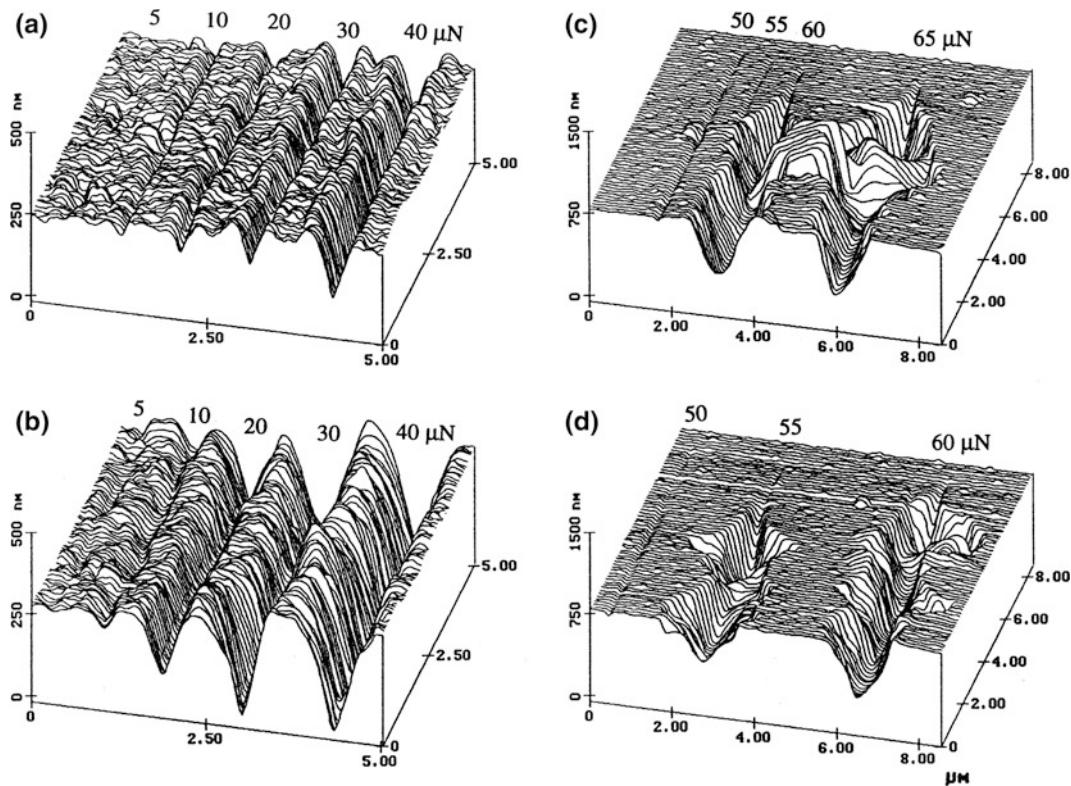


Fig. 15.25 Surface maps for scratched (a) MP, (b) BaFe, (c) ME ($H = 0.7$ GPa), and (d) ME ($H = 2.5$ GPa) tapes along parallel direction. Normal loads used for scratching for one cycle are listed in the figure (Bhushan and Koinkar 1995c)

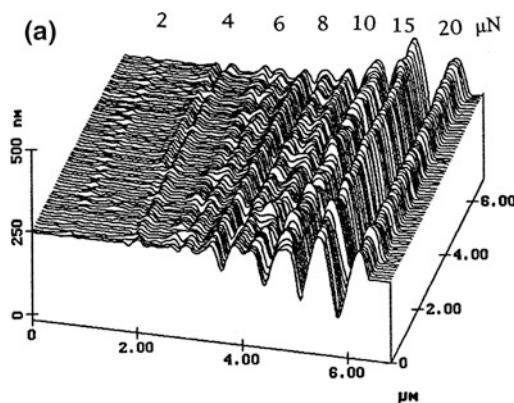


Fig. 15.26 Surface profiles for scratched PET film (a) polymer region, (b) ceramic particulate region. The loads used for various scratches at ten cycles are indicated in the plots (Bhushan and Koinkar 1995a)

A thick coating is desirable for long durability; however, to achieve ever increasing high recording densities, it is necessary to use as thin a

coating as possible. Microwear data on various amorphous carbon coatings of different thicknesses have been conducted by Bhushan and Koinkar (1995d), Koinkar and Bhushan (1997b), and Sundararajan and Bhushan (1999). Figure 15.35 shows a wear mark on an uncoated Si(100) and various 10-nm thick carbon coatings. It is seen that Si(100) wears uniformly, whereas carbon coatings wear nonuniformly. Carbon coating failure is sudden and accompanied by a sudden rise in friction force. Figure 15.36 shows the wear depth of Si(100) substrate and various coatings at two different loads. FCA and ECR-CVD, 20-nm-thick coatings show excellent wear resistance up to 80 μN, the load that is required for the IB 20 nm coating to fail. In these tests, failure of a coating results when the wear depth exceeds the quoted coating thickness. The SP 20 nm coating fails at the much lower load of 35 μN. At 60 μN, the coating hardly

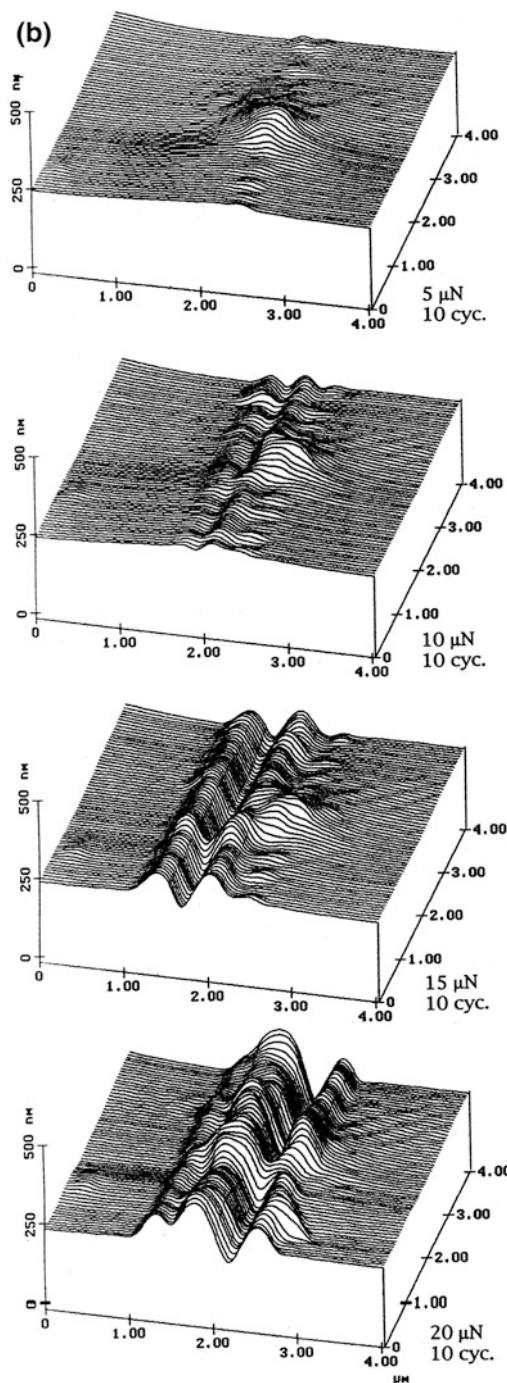


Fig. 15.26 (continued)

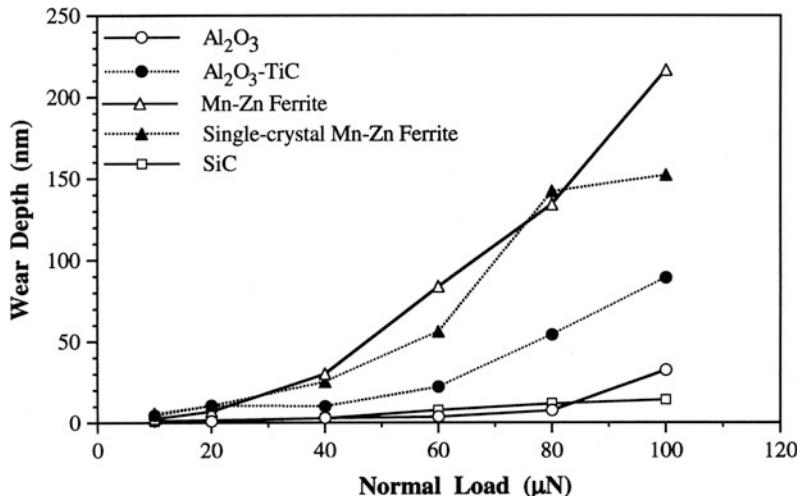
provides any protection. Moving on to the 10 nm coatings, ECR-CVD coating requires about 40 cycles at 60 μ N to fail as compared to IB and

FCA, which fail at 45 μ N. the FCA coating exhibits slight roughening in the wear track after the first few cycles, which leads to an increase in the friction force. The SP coating continues to exhibit poor resistance, failing at 20 μ N. For the 5 nm coatings, the load required to fail the coatings continues to decrease. But IB and ECR-CVD still provide adequate protection as compared to bare Si(100) in that order, failing at 35 μ N compared to FCA at 25 μ N and SP at 20 μ N. Almost all the 20, 10, and 5 nm coatings provide better wear resistance than bare silicon. At 3.5 nm, FCA coating provides no wear resistance, failing almost instantly at 20 μ N. The IB and ECR-CVD coating show good wear resistance at 20 μ N compared to bare Si(100). But IB lasts only about 10 cycles and ECR-CVD about 3 cycles at 25 μ N.

The wear tests highlight the differences in the coatings more vividly than the scratch tests data presented earlier. At higher thicknesses (10 and 20 nm), the ECR-CVD and FCA coatings appear to show the best wear resistance. This is probably due to higher hardness of the coatings (see data presented later). At 5 nm, IB coating appears to be the best. FCA coatings show poorer wear resisting with decreasing coating thickness. SP coatings showed consistently poor wear resistance at all thicknesses. The IB 2.5 nm coating does provide reasonable wear protection at low loads.

Wear depths as a function of normal load for MP, BaFe and ME tapes along the parallel direction are plotted in Fig. 15.37 (Bhushan and Koinkar 1995c). For the ME tape, there is negligible wear until the normal load of about 50 μ N, above this load the magnetic coating fails rapidly. This observation is consistent with the scratch data. Wear depths as a function of number of cycles for MP, BaFe, and ME tapes are shown in Fig. 15.38. For the MP and BaFe particulate tapes, wear rates appear to be independent of the particulate density. Again as observed in the scratch testing, wear rate of BaFe tapes is higher than that for MP tapes. ME tapes are much more wear resistant than the particulate tapes. However, the failure of ME tapes are catastrophic as observed in scratch testing. Wear

Fig. 15.27 Wear depth as a function of normal load after one scan cycle for Al_2O_3 , $\text{Al}_2\text{O}_3\text{-TiC}$, polycrystalline Mn-Zn ferrite, single-crystal Mn-Zn ferrite and SiC (Koinkar and Bhushan 1996a)



studies were performed along and across the longitudinal tape direction in high and low hardness regions. At the high hardness regions of the ME tapes, failure occurs at lower loads. Directionality effect again may arise from the columnar structure of the ME films (Hibst 1993; Bhushan 2000). Wear profiles at various cycles at a normal load of 2 μN for MP and at 20 μN for ME tapes are shown in Fig. 15.39. For the particulate tapes, we note that polymer gets removed before the particulates do (Fig. 15.39a). Based on the wear profiles of the ME tape shown in Fig. 15.39a, we note that most wear occurs between 50 and 60 cycles which shows the catastrophic removal of the coating. It was also observed that wear debris generated during wear test in all cases is loose and can easily be removed away from the scan area at light loads ($\sim 0.3 \mu\text{N}$).

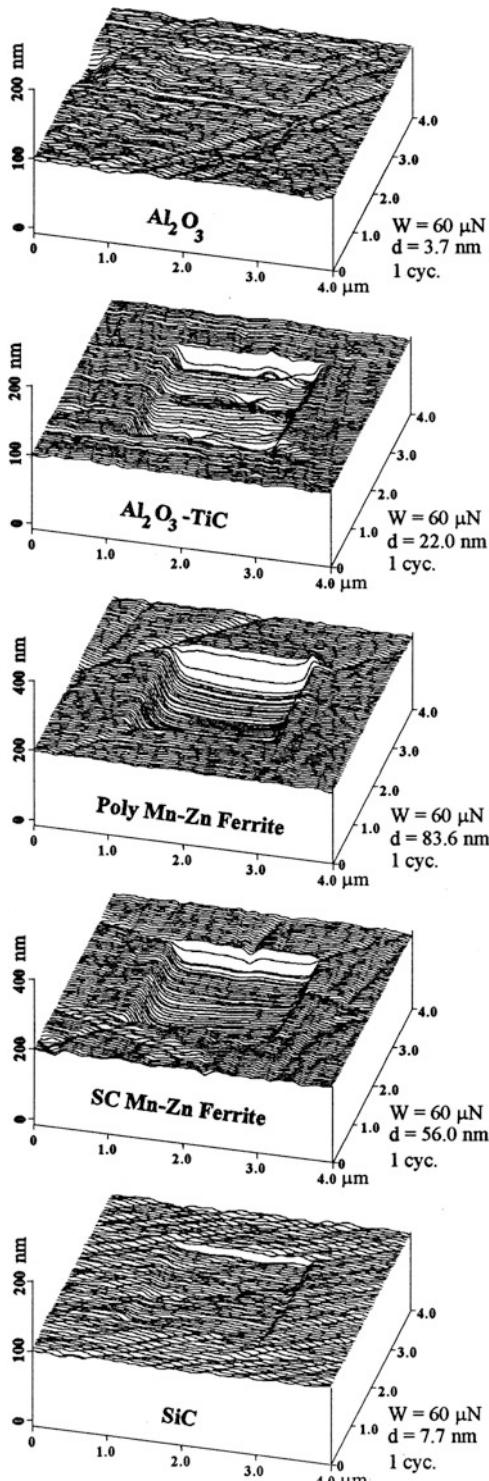
The average wear depth as a function of load for a PET film is shown in Fig. 15.40. Again, the wear depth increases linearly with load. Figure 15.41 shows the average wear depth as a function of number of cycles. The observed wear rate is approximately constant. PET tape substrate consists of particles sticking out on its surface to facilitate winding. Figure 15.42 shows the wear profiles as a function of number of cycles at 1 μN load on the PET film in the nonparticulate and particulate regions (Bhushan and Koinkar 1995a). We note that polymeric

materials tear in microwear tests. The particles do not wear readily at 1 μN . Polymer around the particles is removed but the particles remain intact. Wear in the particulate region is much smaller than that in the polymer region. We will see later that nanohardness of the particulate region is about 1.4 GPa compared to 0.3 GPa in the nonparticulate region (Table 15.3).

15.6 Indentation

15.6.1 Picoscale Indentation

Bhushan and Ruan (1994) measured indentability of magnetic tapes at increasing loads on a picoscale, Fig. 15.43. In this figure, the vertical axis represents the cantilever tip deflection and the horizontal axis represents the vertical position (Z) of the sample. The “extending” and “retracting” curves correspond to the sample being moved toward or away from the cantilever tip, respectively. In this experiment, as the sample surface approaches the AFM tip fraction of a nm away from the sample (point A), the cantilever bends toward the sample (part B) because of attractive forces between the tip and sample. As we continue the forward position of the sample, it pushes the cantilever back through its original rest position (point of zero applied load) entering the repulsive region (or loading portion) of the



◀ **Fig. 15.28** Surface profiles showing the worn region (center 2 $\mu\text{m} \times 2 \mu\text{m}$) after one scan cycles at a normal load of 60 μN for Al_2O_3 , $\text{Al}_2\text{O}_3\text{-TiC}$, polycrystalline Mn–Zn ferrite, single-crystal Mn–Zn ferrite and SiC (Koinkar and Bhushan 1996a)

force curve. As the sample is retracted, the cantilever deflection decreases. At point D in the retracting curve, the sample is disengaged from the tip. Before the disengagement, the tip is pulled toward the sample after the zero deflection point of the force curve (point C) because of attractive forces (van der Waals forces and longer range meniscus forces). A thin layer of liquid, such as liquid lubricant and condensations of water vapor from ambient, will give rise to capillary forces that act to draw the tip towards sample at small separations. The horizontal shift between the loading and unloading curves results from the hysteresis in the PZT tube.

The left portion of the curve shows the tip deflection as a function of the sample traveling distance during sample-tip contact, which would be equal to each other for a rigid sample. However, if the tip indents into the sample, the tip deflection would be less than the sample traveling distance, or in other words, the slope of the line would be less than 1. In Fig. 15.43, we note that line in the left portion of the figure is curved with a slope of less than 1 shortly after the sample touches the tip, which suggests that the tip has indented the sample. Later, the slope is equal to 1 suggesting that the tip no longer indents the sample. This observation indicates that the tape surface is soft locally (polymer rich) but it is hard (as a result of magnetic particles) underneath. Since the curves in extending and retracting modes are identical, the indentation is elastic up to at a maximum load of about 22 nN used in the measurements.

According to Bhushan and Ruan (1994), during indentation of rigid disks, the slope of the deflection curves remained constant as the disks touch and continue to push the AFM tip. The disks were not indented.

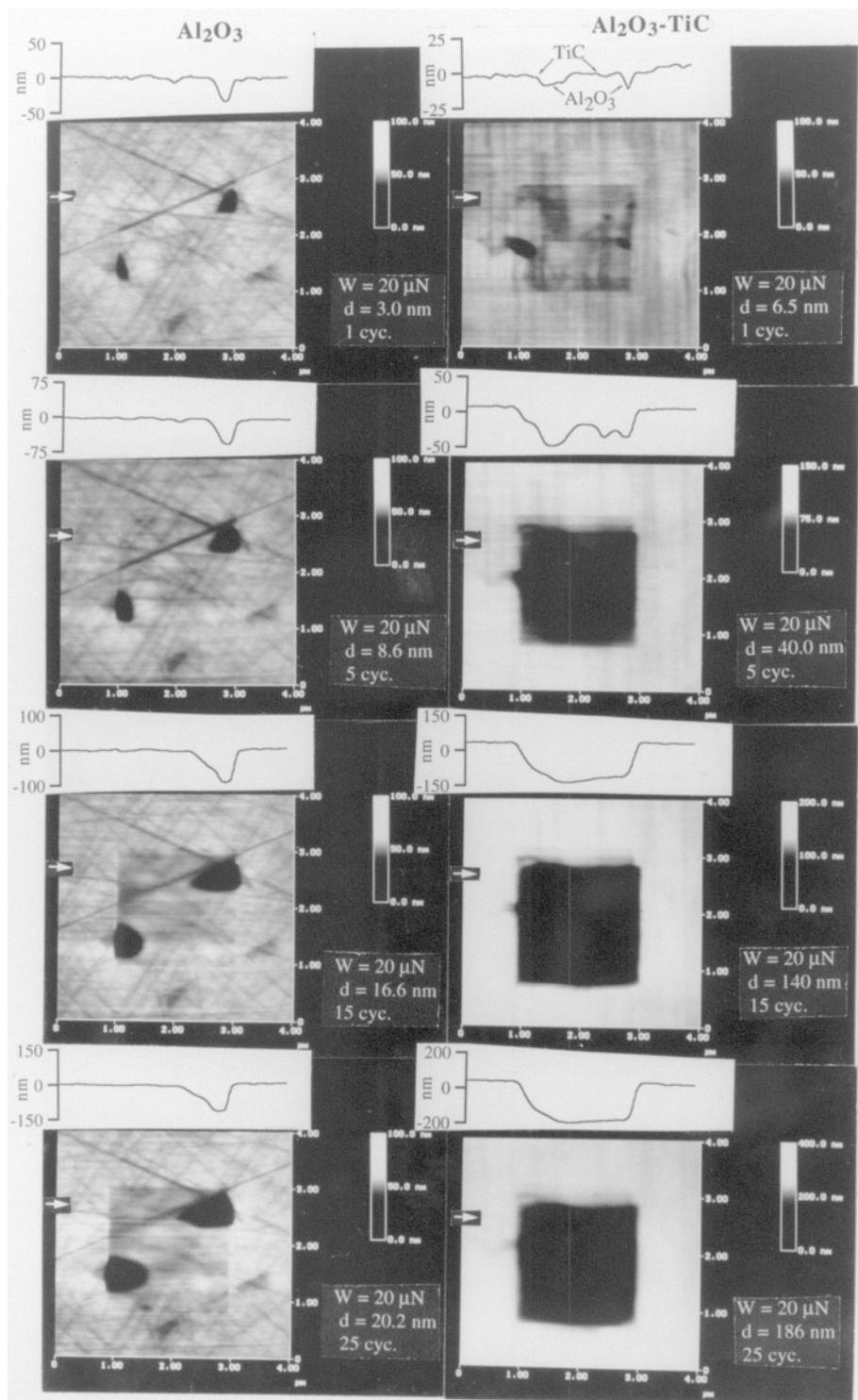


Fig. 15.29 Gray scale 2-D plots showing the worn region (center $2 \mu\text{m} \times 2 \mu\text{m}$) at a normal load of $20 \mu\text{N}$ and different number of scan cycles for Al_2O_3 and $\text{Al}_2\text{O}_3\text{-TiC}$. The 2-D section plots taken at a location shown by an arrow are shown on the top of corresponding gray scale plot. Note the change in vertical scale for gray scale and 2-D section plots (Koinkar and Bhushan 1996a)

Fig. 15.30 Wear depth as a function of number of cycles at a normal load of 20 μN for Al_2O_3 , Al_2O_3 -TiC, polycrystalline Mn-Zn ferrite, single-crystal (SC) Mn-Zn ferrite and SiC (Koinkar and Bhushan 1996a)

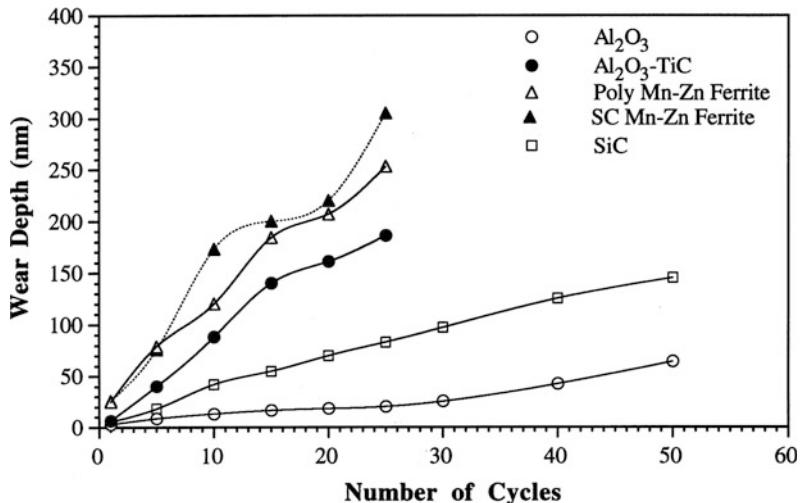
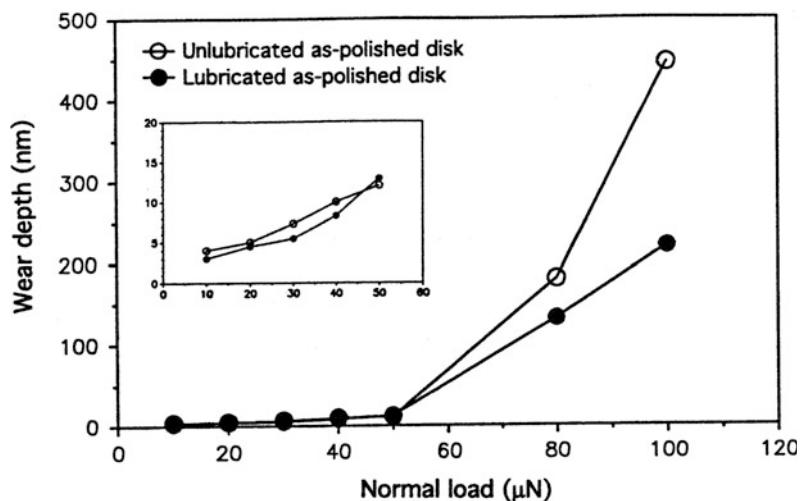


Fig. 15.31 Wear depth as a function of normal load for polished, lubricated and unlubricated thin-film rigid disks after one cycle (Bhushan et al. 1994)



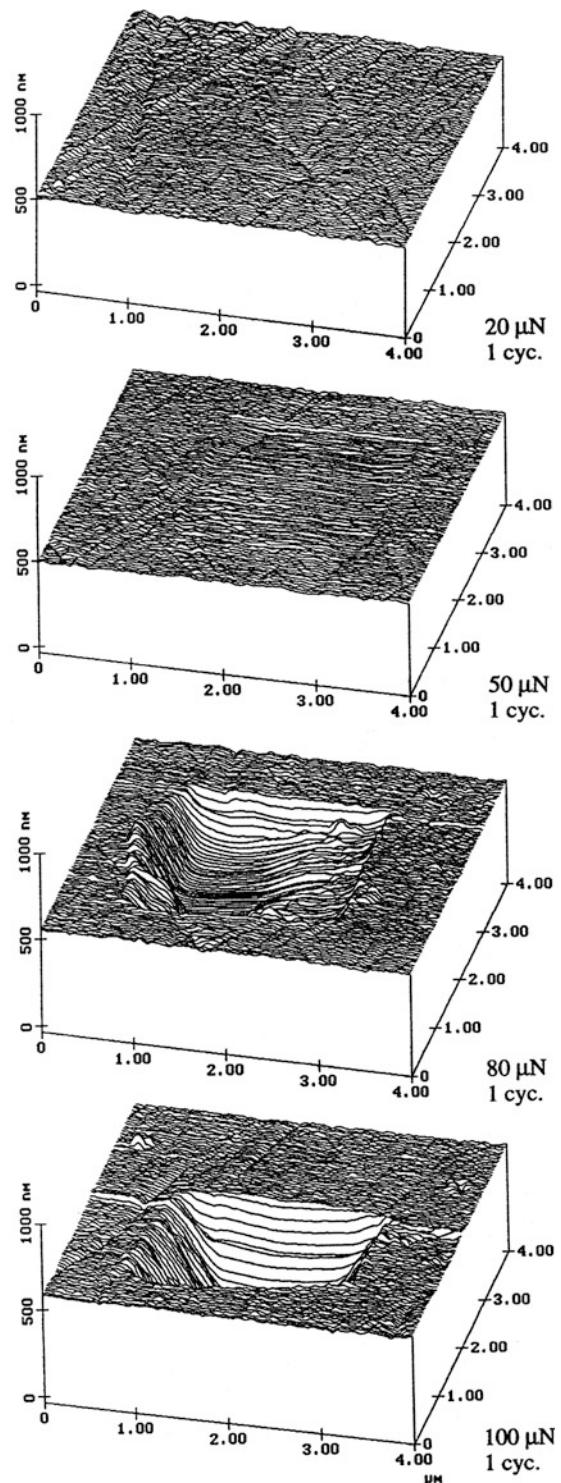
15.6.2 Nanoscale Indentation

Indentation hardness with a penetration depth as low as 5 nm can be measured using AFM. Bhushan and Koinkar (1994b) measured hardness of thin-film disks at load of 80, 100, and 140 μN . Hardness values were 20 GPa (10 nm), 21 GPa (15 nm) and 9 GPa (40 nm); the depths of indentation are shown in the parenthesis. The hardness value at 100 μN is much higher than at 140 μN . This is expected since the indentation depth is only about 15 nm at 100 μN which is smaller than the thickness of carbon coating (~ 30 nm). The hardness value at lower loads is

primarily the value of the carbon coating. The hardness value at higher loads is primarily the value of the magnetic film, which is softer than the carbon coating (Bhushan 1996a). This result is consistent with the scratch and wear data discussed previously.

For the case of hardness measurements made on magnetic thin film rigid disk at low loads, the indentation depth is on the same order at the variation in the surface roughness. For accurate measurements of indentation size and depth, it is desirable to subtract the original (unindented) profile from the indented profile. Bhushan et al. (1994) developed an algorithm for this purpose.

Fig. 15.32 Surface maps of a polished, unlubricated thin-film rigid disk showing the worn region (center 2 μm \times 2 μm) after one cycle. The normal loads are indicated in the figure (Bhushan et al. 1994)



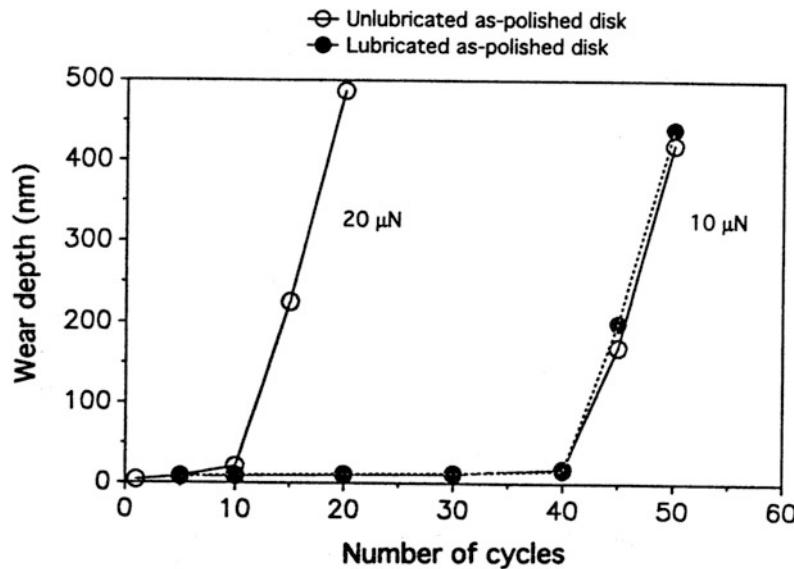
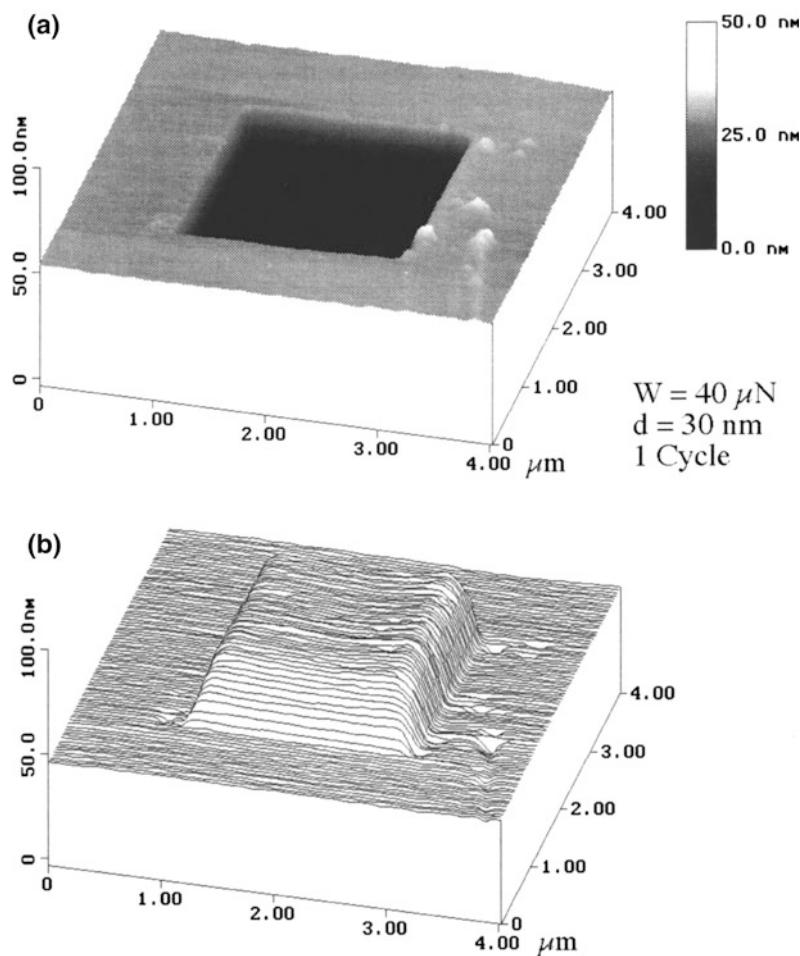


Fig. 15.33 Wear depth as a function of number of cycles for polished, lubricated and unlubricated thin-film rigid disks at 10 μN and for polished, unlubricated disk at 20 μN (Bhushan et al. 1994)

Fig. 15.34 Surface maps of a polished, unlubricated thin-film rigid disk showing the worn region (center 2 μm \times 2 μm) at 20 μN . The number of cycles are indicated in the figure (Bhushan et al. 1994)



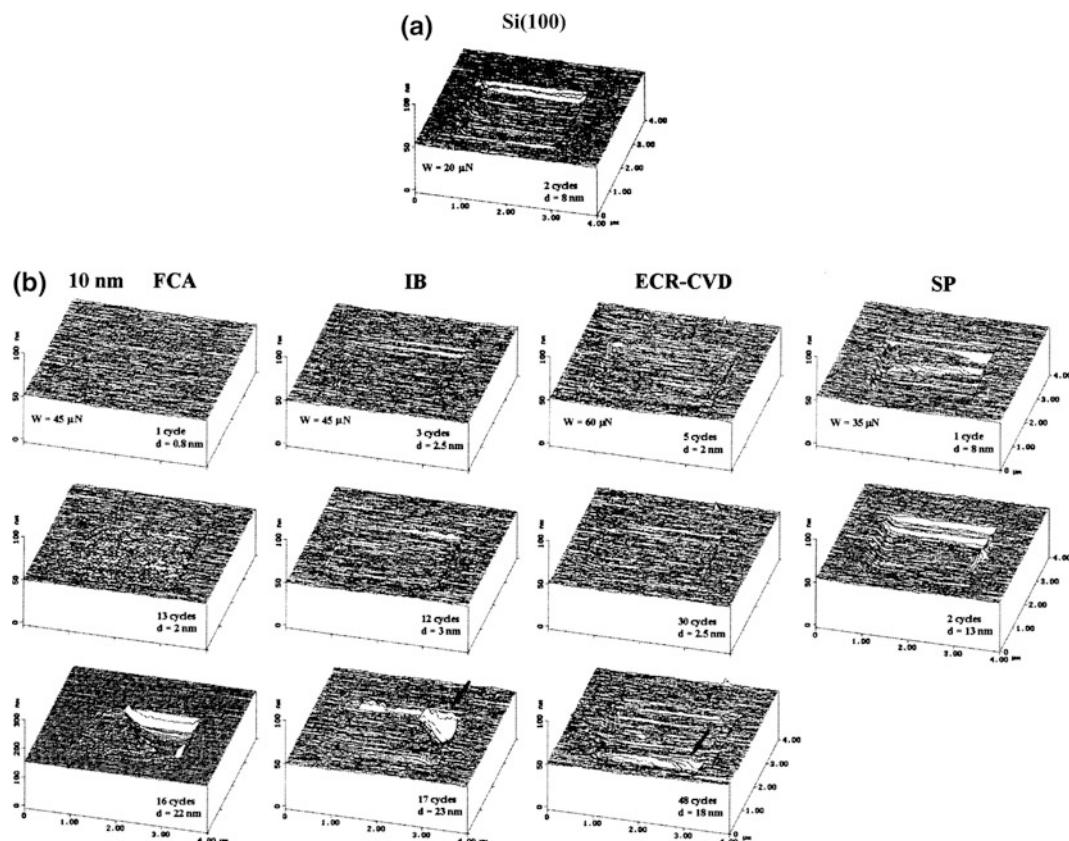


Fig. 15.35 AFM images of wear marks on (a) bare Si(100), and (b) all 10-nm thick amorphous carbon coatings (Sundararajan and Bhushan 1999)

Because of hysteresis, a translational shift in the sample plane occurs during the scanning period, resulting in a shift between images captured before and after indentation. Therefore, the image for perfect overlap needs to be shifted before subtraction can be performed. To accomplish this objective, a small region on the original image was selected and the corresponding region on the indented image was found by maximizing the correlation between the two regions. (Profiles were plane-fitted before subtraction.) Once two regions were identified, overlapped areas between the two images were determined and the original image was shifted with the required translational shift and this subtracted from the indented image. An example

of profiles before and after subtraction is shown in Fig. 15.44. It is easier to measure the indent on the subtracted image. At a normal load of 140 mN the hardness value of an unlubricated, as-polish magnetic thin film rigid disk (rms roughness = 3.3 nm) is 9.0 GPa and indentation depth is 40 nm.

For accurate measurement of nanohardness at very shallow indentation depths, depth-sensing capacitance transducer system in an AFM is used (Bhushan et al. 1996). Figure 15.45a shows the hardness as a function of residual depth for three types of 100-nm thick amorphous carbon coatings deposited on silicon by sputtering, ion beam and cathodic arc processes (Kulkarni and Bhushan 1997). Data on uncoated silicon are also

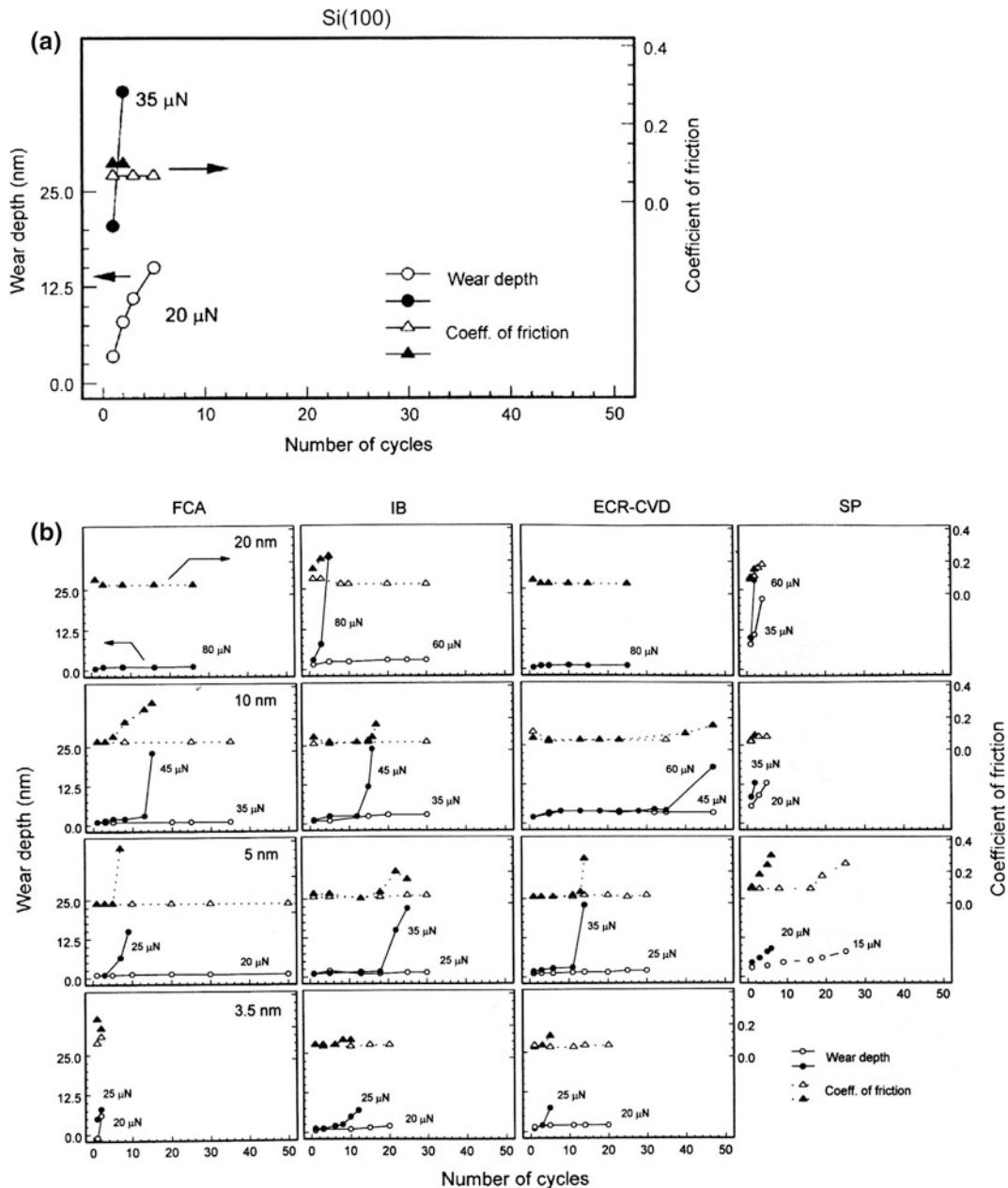


Fig. 15.36 Wear data of (a) bare Si(100) and (b) all amorphous carbon coatings. Coating thickness is constant along each row in (b). Both wear depth and coefficient of friction during wear for a given cycle are plotted (Sundararajan and Bhushan 1999)

included for comparisons. The cathodic arc carbon coating exhibits highest hardness of about 24.9 GPa, whereas the sputtered and ion beam carbon coatings exhibit hardness values of 17.2 and 15.2 GPa respectively. The hardness of Si

(100) is 13.2 GPa. High hardness of cathodic arc carbon coating explains its high wear resistance, reported earlier. Figure 15.45b shows the elastic modulus as a function of residual depth for various samples. The cathodic arc coating exhibits

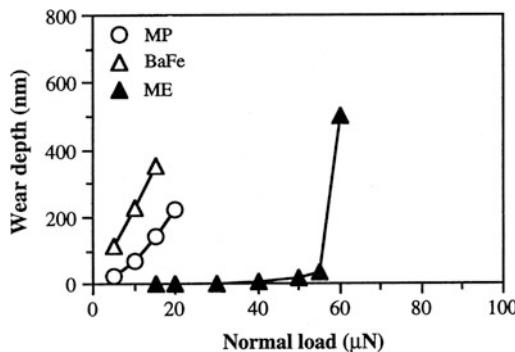


Fig. 15.37 Wear depth as a function of normal load for three tapes in the parallel direction after one cycle (Bhushan and Koinkar 1995c)

the highest elastic modulus. Its elastic modulus decreases with an increasing residual depth, while the elastic moduli for the other carbon coatings remain almost constant. In general, hardness and elastic modulus of coatings are strongly influenced by their crystalline structure, stoichiometry and growth characteristics which depend on the deposition parameters. Mechanical properties of carbon coatings have been known to change over a wide range with $\text{sp}^3\text{-sp}^2$ bonding ratio and amount of hydrogen. Hydrogen is believed to play an important role in the bonding configuration of carbon atoms by helping to stabilize tetrahedral coordination of carbon atoms. Detailed mechanical characterization of amorphous carbon coatings is presented by Li and Bhushan (1999a, b) and Bhushan (1999b).

15.6.3 Localized Surface Elasticity

By using an AFM in a so-called force modulation mode, it is possible to quantitatively measure the elasticity of soft and compliant materials with penetration depths of less than 100 nm (DeVecchio and Bhushan 1997; Scherer et al. 1997). This technique has been successfully used to get localized elasticity maps of particulate magnetic tapes. Elasticity map of a tape can be used to identify relative distribution of hard magnetic/nonmagnetic ceramic particles and the polymeric

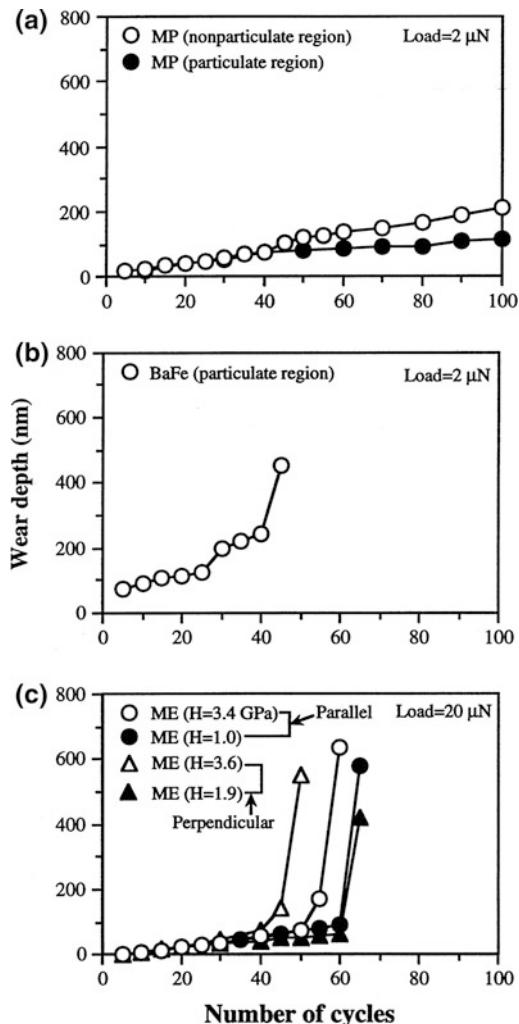


Fig. 15.38 Wear depth as a function of number of cycles for (a) MP, (b) BaFe, and (c) ME tapes in different regions at normal loads indicated in the figure. Note a higher load used for the ME tape in (c) (Bhushan and Koinkar 1995c)

binder on the tape surface which has an effect on friction and stiction at the head-tape interface. Figure 15.46 shows surface height and elasticity maps on an MP tape. The elasticity image reveals sharp variations in surface elasticity due to the composite nature of the film. As can be clearly seen, regions of high elasticity do not always correspond to high or low topography. Based on a Hertzian elastic-contact analysis, the static

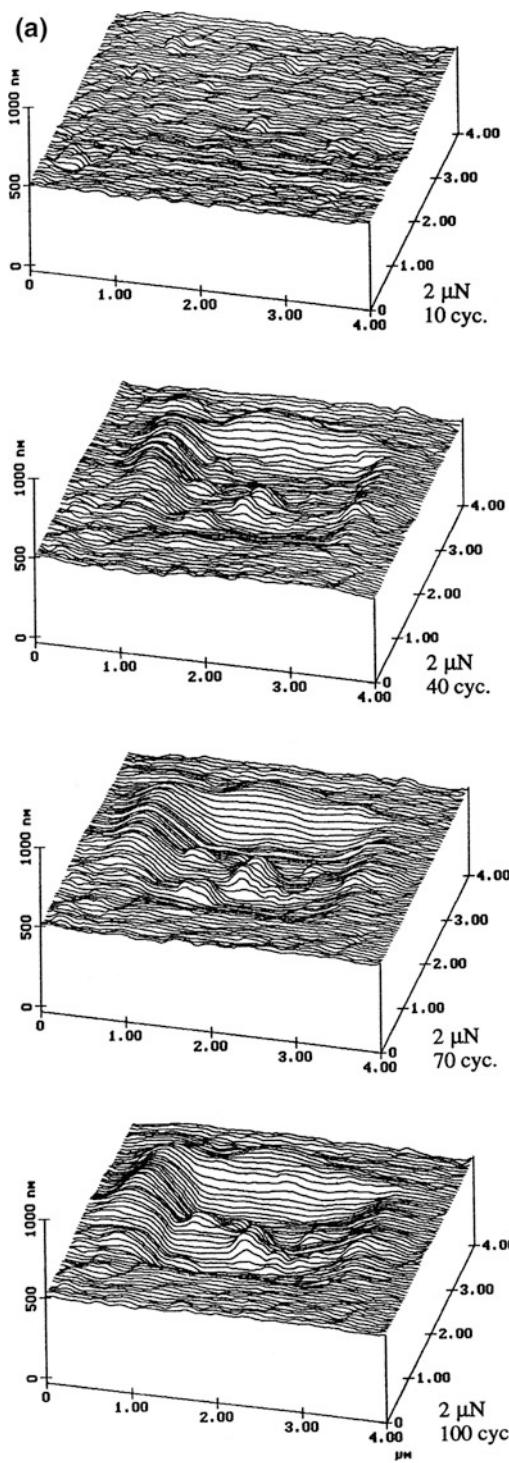


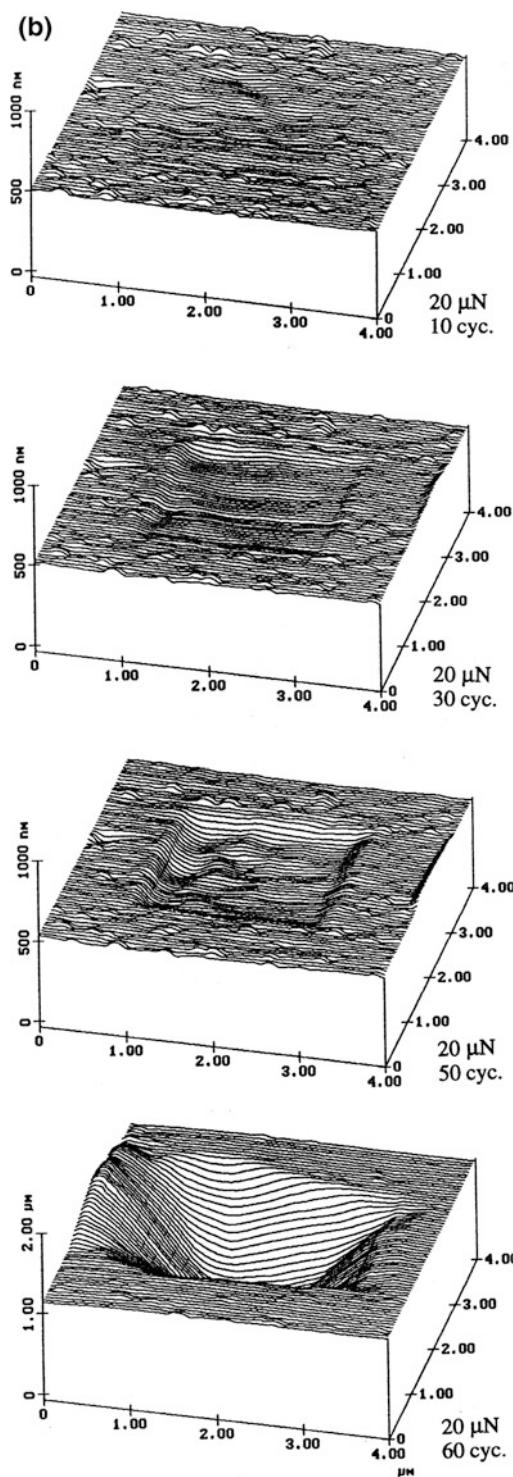
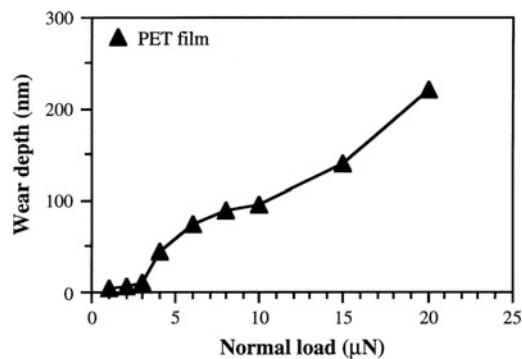
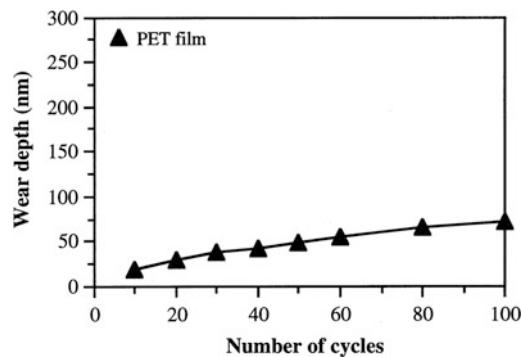
Fig. 15.39 Surface maps showing the worn region (center $2 \mu\text{m} \times 2 \mu\text{m}$) after various cycles of wear at (a) $2 \mu\text{N}$ for MP (particulate region) and at (b) $20 \mu\text{N}$ for ME ($H = 3.4 \text{ GPa}$, parallel direction) tapes. Note a different vertical scale for the bottom profile of (b) (Bhushan and Koinkar 1995c)

indentation depth of these sample during the force modulation scan is estimated to be about 1 nm. The contrast seen is influenced most strongly by material properties in the top few nanometers, independent of the composite structure beneath the surface layer. The trend in number of stiff regions has been correlated to reduced stiction problems in tapes (Bhushan et al. 1997).

Figure 15.47 shows the surface topography and phase image of an alumina particle embedded in the MP tape using a so-called TR mode (Kasai et al. 2004; Bhushan and Kasai 2004). The cross-section view of the particle obtained from the topographic image is shown at the bottom as a visual aid. The edges of the particle show up darker in the TR phase angle image, which suggests that it is less viscoelastic compared to the background. The magnetic particles on top of the alumina particle are clearly visible in the TR phase image. These have a brighter contrast, which is the same as that of the background. Phase contrast mapping appears to provide better resolution than stiffness mapping for magnetic tapes.

15.7 Lubrication

The boundary films are formed by physical adsorption, chemical adsorption, and chemical reaction. The physisorbed film can be either monomolecular or polymolecular thick. The chemisorbed films are monomolecular, but stoichiometric films formed by chemical reaction can have a large film thickness. In general, the stability and durability of surface films decrease in the following order: chemical reaction films, chemisorbed films and physisorbed films.

**Fig. 15.39** (continued)**Fig. 15.40** Wear depth as a function of normal load (after one cycle) for a PET film (Bhushan and Koinkar 1995a)**Fig. 15.41** Wear depth as a function of number of cycles at 1 μN for a PET film (Bhushan and Koinkar 1995a)

A good boundary lubricant should have a high degree of interaction between its molecules and the sliding surface. As a general rule, liquids are good lubricants when they are polar and thus able to grip solid surfaces (or be adsorbed). Polar lubricants contain reactive functional groups with low ionization potential or groups having high polarizability (Bhushan 1993a). Boundary lubrication properties of lubricants are also dependent upon the molecular conformation and lubricant spreading (Novotny et al. 1989; Novotny 1990; Mate and Novotny 1991; Mate 1992).

Mechanical interactions between the magnetic head and the medium in magnetic storage devices are minimized by the lubrication of the magnetic medium (Bhushan 1996a, 2000). The primary

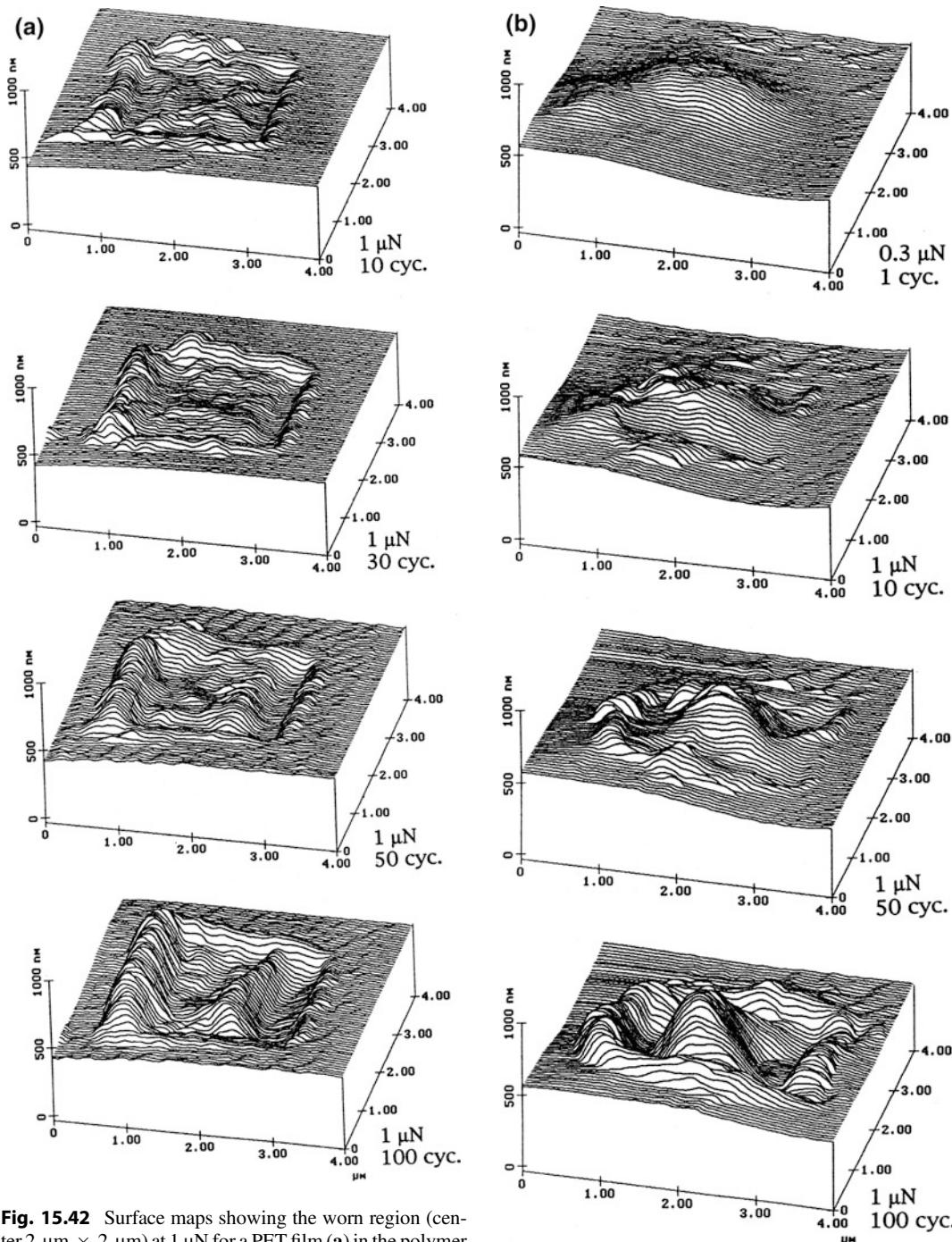


Fig. 15.42 Surface maps showing the worn region (center $2 \mu\text{m} \times 2 \mu\text{m}$) at $1 \mu\text{N}$ for a PET film (a) in the polymer region, (b) in the particulate region. The number of cycles are indicated in the figure (Bhushan and Koinkar 1995a)

function of the lubricant is to reduce the wear of the medium and to ensure that friction remains low throughout the operation of the drive. The

main challenge, though, in selecting the best candidate for a specific surface is to find a material that provides an acceptable wear protection for the

Fig. 15.42 (continued)

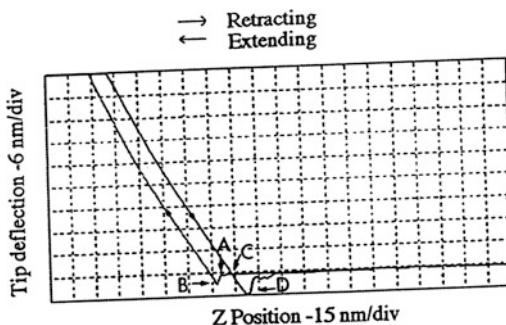


Fig. 15.43 Tip deflection (normal force) as a function of Z (separation distance) curve for a metal-particle (MP) tape. The spring constant of the cantilever used was 0.4 N/m (Bhushan and Ruan 1994)

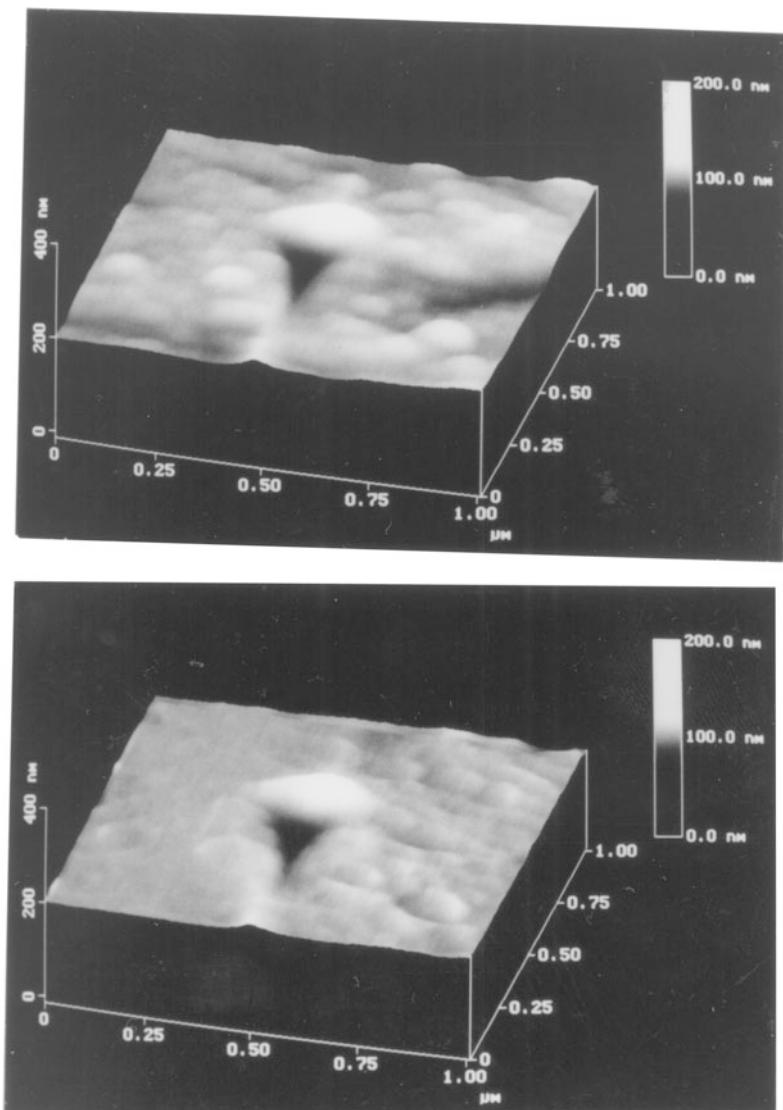
entire life of the product, which can be several years in duration. There are many requirements that a lubricant must satisfy in order to guarantee an acceptable life performance. An optimum lubricant thickness is one of these requirements. If the lubricant film is too thick, excessive stiction and mechanical failure of the head-disk is observed. On the other hand, if the film is too thin, protection of the interface is compromised, and high friction and excessive wear will result in catastrophic failure. An acceptable lubricant must exhibit properties such as chemical inertness, low volatility, high thermal, oxidative and hydrolytic stability, shear stability, and good affinity to the magnetic medium surface.

Fatty acid esters are excellent boundary lubricants, and esters such as tridecyl stearate, butyl stearate, butyl palmitate, buryl myristate, stearic acid, and myristic acid are commonly used as internal lubricants, roughly 1–7% by weight of the magnetic coating in particulate flexible media (tapes and particulate flexible disks) (Bhushan 1996a, 2000). The fatty acids involved include those with acid groups with an even number of carbon atoms between C₁₂ and C₂₂, with alcohols ranging from C₃ to C₁₃. These acids are all solids with melting points above the normal surface operating temperature of the magnetic media. This suggests that the decomposition products of the ester via lubrication chemistry during a head-flexible medium contact may be the key to lubrication.

Topical lubrication is used to reduce the wear of rigid disks and thin-film tapes (Bhushan and Zhao 1999). Perfluoropolyethers (PFPEs) are chemically the most stable lubricants with some boundary lubrication capability, and are most commonly used for topical lubrication of rigid disks. PFPEs commonly used include Fomblin Z lubricants, made by Solvay Solexis Inc., Milan, Italy; and Demnum S, made by Diakin, Japan; and their difunctional derivatives containing various reactive end groups, e.g., hydroxyl or alcohol (Fomblin Z-DOL and Z-TETROL), piperonyl (Fomblin AM 2001), isocyanate (Fomblin Z-DISOC), and ester (Demnum SP). Fomblin Y and Krytox 143AD (made by Dupont USA) have been used in the past for particulate rigid disks. The difunctional derivatives are referred to as reactive (polar) PFPE lubricants. The chemical structures, molecular weights, and viscosities of various types of PFPE lubricants are given in Table 15.4. We note that rheological properties of thin-films of lubricants are expected to be different from their bulk properties. Fomblin Z and Demnum S are linear PFPE, and Fomblin Y and Krytox 143 AD are branched PFPE, where the regularity of the chain is perturbed by –CF₃ side groups. The bulk viscosity of Fomblin Y and Krytox 143 AD is almost an order of magnitude higher than the Z type. Fomblin Z is thermally decomposed more rapidly than Y (Bhushan 1993a). The molecular diameter is about 0.8 nm for these lubricant molecules. The monolayer thickness of these molecules depends on the molecular conformations of the polymer chain on the surface (Novotny et al. 1989; Mate and Novotny 1991).

The adsorption of the lubricant molecules on a magnetic disk surface is due to van der Waals forces, which are too weak to offset the spin-off losses, or to arrest displacement of the lubricant by water or other ambient contaminants. Considering that these lubricating films are on the order of a monolayer thick and are required to function satisfactorily for the duration of several years, the task of developing a workable interface is quite formidable. An approach aiming at

Fig. 15.44 Images with nanoindentation marks generated on a polished, unlubricated thin-film rigid disk at 140 μN (a) before subtraction, and (b) after subtraction (Bhushan et al. 1994)



alleviating these shortcomings is to enhance the attachment of the molecules to the overcoat, which, for most cases, is sputtered carbon. There are basically two approaches which have been shown to be successful in bonding the monolayer to the carbon. The first relies on exposure of the disk lubricated with neutral PFPE to various forms of radiation, such as low-energy X ray (Heidemann and Wirth 1984), nitrogen plasma (Homola et al. 1990), or far ultraviolet (e.g.,

185 nm) (Saperstein and Lin 1990). Another approach is to use chemically active PFPE molecules, where the various functional (reactive) end groups offer the opportunity of strong attachments to specific interface. These functional groups can react with surfaces and bond the lubricant to the disk surface, which reduces its loss due to spin off and evaporation. Bonding of lubricant to the disk surface depends upon the surface cleanliness. After lubrication, the disk is

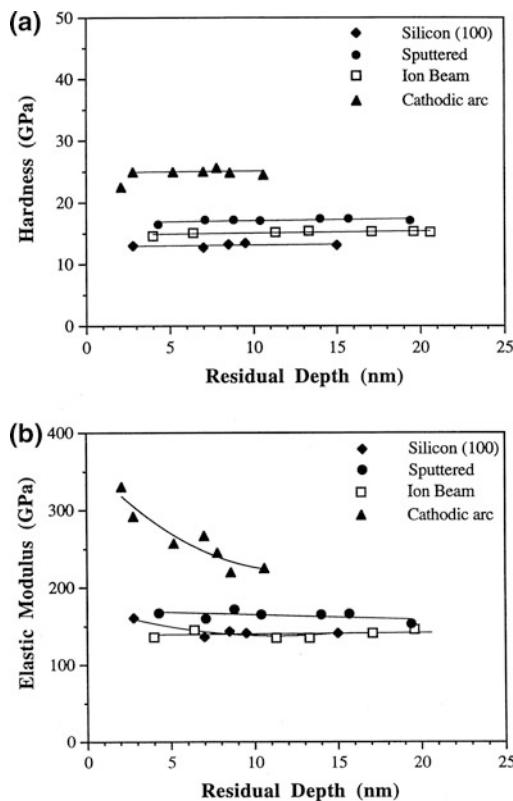


Fig. 15.45 Nanohardness and elastic modulus as a function of residual indentation depth for Si(100) and 100-nm thick coatings deposited by sputtering, ion beam and cathodic arc processes (Kulkarni and Bhushan 1997)

generally heated at 150 °C for 30 min to 1 h to improve the bonding. If only a bonded lubrication is desired, the unbonded fraction can be removed by washing it off for 60 s with a non-freon solvent (FC-72). Their main advantage is their ability to enhance durability without the problem of stiction usually associated with weakly bonded lubricants (Bhushan 1996a).

15.7.1 Boundary Lubrication Studies

Koinkar and Bhushan (1996b) and Liu and Bhushan (2003) studied friction, adhesion, and durability of Z-15 and Z-DOL (bonded and washed, BW) lubricants on Si(100) surface. To investigate the friction properties of Si(100), Z-15, and Z-DOL(BW), the friction force versus

normal load curves were obtained by making friction measurements at increasing normal loads, Fig. 15.48. An approximately linear response of all three samples is observed in the load range of 5–130 nN. From the horizontal intercept at zero value of friction force, adhesive force can be obtained. The adhesive forces for three samples were also measured using the force calibration plot technique. The adhesive force data obtained by the two techniques are summarized in Fig. 15.49, and the trends in the data obtained by two techniques are similar. The friction force and adhesive force of solid-like Z-DOL(BW) are consistently smaller than that for Si(100), but these values of liquid-like Z-15 lubricant is higher than that of Si(100). The presence of mobile Z-15 lubricant film increases adhesive force as compared to that of the Si(100) by meniscus formation. Whereas, the presence of Z-DOL(BW) film reduces the adhesive force because of absence of mobile liquid. See schematics at the bottom of Fig. 15.49. It is well known that in computer rigid disk drives, the stiction force increases rapidly with an increase in rest time between head and the disk (Bhushan 1996a). The effect of rest time of 180 s on the friction force, adhesive force, and coefficient of friction for three samples are summarized in Fig. 15.50. It is seen that time effect is present in Si(100) and Z-15 with mobile liquid present. Whereas, time effect is not present for Z-DOL (BW) because of the absence of mobile liquid.

To study lubricant depletion during micro-scale measurements, nanowear studies were conducted using Si_3N_4 tips. Measured friction as a function of number of cycles for Si(100) and silicon surface lubricated with Z-15 and Z-DOL (BW) lubricants are presented in Fig. 15.51. An area of $2 \mu\text{m} \times 2 \mu\text{m}$ was scanned at a normal force of 70 nN. As observed before, friction force and coefficient of friction of Z-15 is higher than that of Si(100) with the lowest values for Z-DOL (BW). During cycling, friction force and coefficient of friction of Si(100) show a slight decrease during initial few cycles, then remain constant. This is related to the removal of the top adsorbed layer. In the case of Z-15 film, the friction force and coefficient of friction show an increase

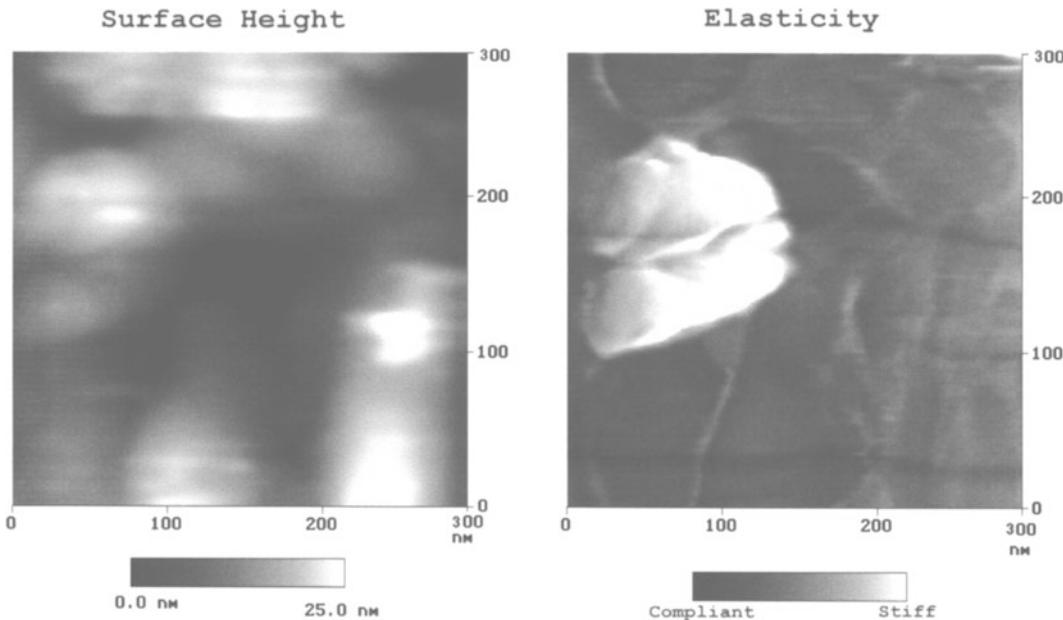


Fig. 15.46 (a) Surface height and elasticity maps for a metal-particle tape A ($\sigma = 6.72 \text{ nm}$ and $P-V = 31.7 \text{ nm}$). σ and $P-V$ refer to standard deviation of surface heights and peak-to-valley distance, respectively. The grayscale on the elasticity map is arbitrary (DeVecchio and Bhushan 1997)

during the initial few cycles and then approach to higher and stable values. This is believed to be caused by the attachment of the Z-15 molecules onto the tip. The molecular interaction between these attached molecules to the tip and molecules on the film surface is responsible for an increase in the friction. But after several scans, this molecular interaction reaches to the equilibrium and after that friction force and coefficient of friction remain constant. In the case of Z-DOL (BW) film, the friction force and coefficient of friction start out to be low and remain low during the entire test for 100 cycles. It suggests that Z-DOL (BW) molecules do not get attached or displaced as readily as Z-15.

15.8 Closure

Atomic force microscopes/friction force microscopes (AFM/FFM) have been successfully used for measurements of surface roughness, friction, adhesion, scratching, wear, indentation, and lubrication on the micro to nanoscales. Commonly-measured roughness parameters are

scale-dependent, requiring the need of scale-independent fractal parameters to characterize surface roughness. A generalized fractal analysis is presented which allows the characterization of surface roughness by two scale-independent parameters. Local variation in microscale friction force is found to correspond to the local surface slope, suggesting that a ratchet mechanism is responsible for this variation. Directionality in the friction is observed on both micro- and macro-scales because of surface topography. Microscale friction is found to be significantly smaller than the macro-scale friction as there is less plowing contribution in microscale measurements.

Wear rates for particulate magnetic tapes and polyester tape substrates are approximately constant for various loads and test durations. However, for magnetic disks and magnetic tapes with a multilayered, thin-film structure, the wear of the DLC overcoat in the case of disks and magnetic layer in the case of tapes, is catastrophic. Breakdown of thin-films can be detected with AFM. Evolution of the wear has also been studied using AFM. We find that the wear is

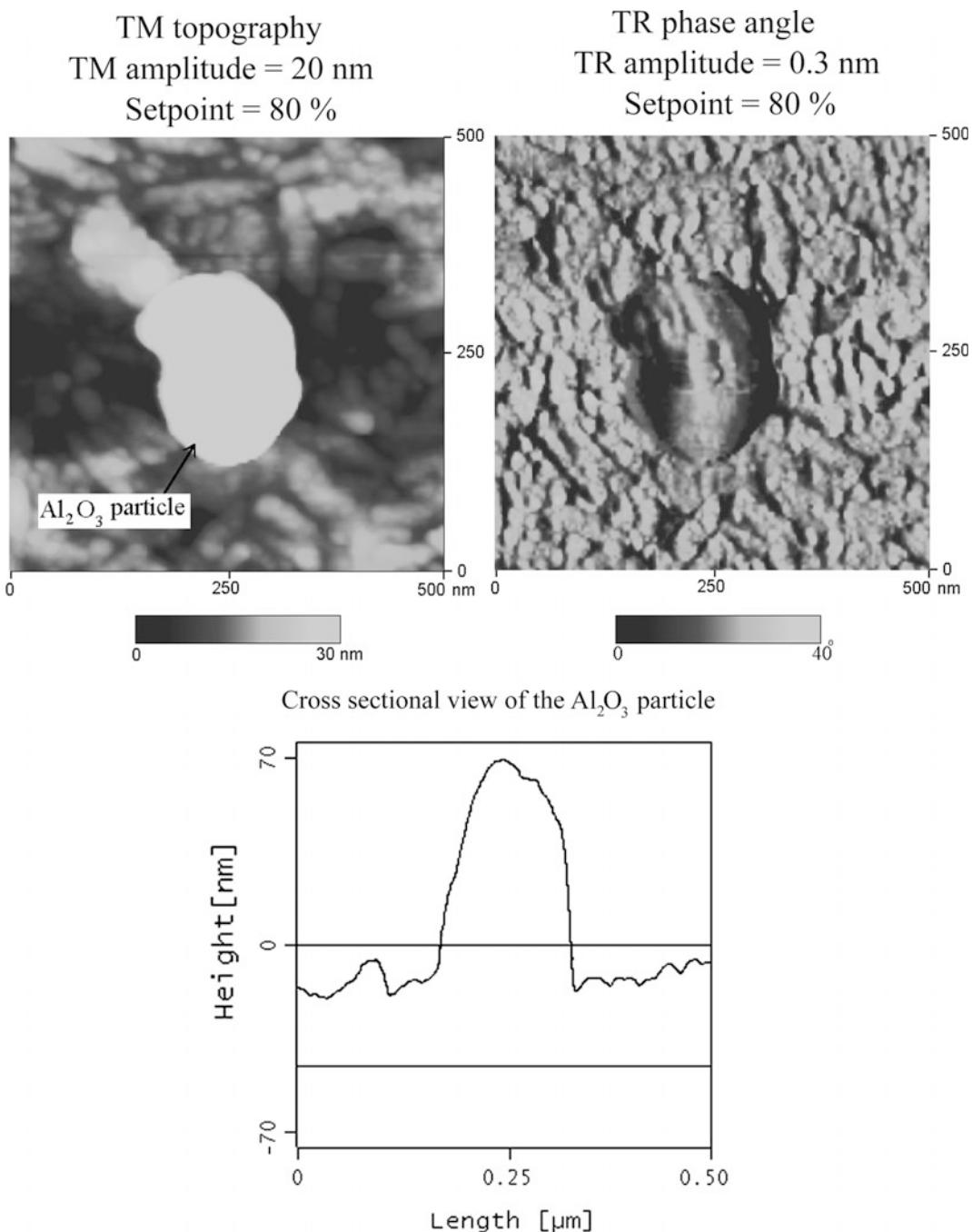


Fig. 15.47 Tapping mode (TM) topography and TR phase angle image of an alumina particle that is used as a head cleaning agent for MP tape. A cross-sectional view of the particle is also shown (Kasai et al. 2004)

initiated at nanoscratches. Amorphous carbon films as thin as 3.5 nm are deposited as continuous films, and exhibit some wear life. Wear life

increases with an increase in film thickness. Carbon coatings deposited by cathodic arc and ECR-CVD processes are superior in wear and

Table 15.4 Chemical structure, molecular weight, and viscosity of perfluoropolyether lubricants

Lubricant	Formula	Molecular weight (Daltons)	Kinematic viscosity cSt (mm ² /s)
Fomblin Z-25	$\text{CF}_3-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_m-(\text{CF}_2-\text{O})_n-\text{CF}_3$	12,800	250
Fomblin Z-15	$\text{CF}_3-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_m-(\text{CF}_2-\text{O})_n-\text{CF}_3$ (m/n ~ 2/3)	9100	150
Fomblin Z-03	$\text{CF}_3-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_m-(\text{CF}_2-\text{O})_n-\text{CF}_3$	3600	30
Fomblin Z-DOL	$\text{HO}-\text{CH}_2-\text{CF}_2-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_m-(\text{CF}_2-\text{O})_n-\text{CF}_2-\text{CH}_2-\text{OH}$	2000	80
Fomblin AM2001	Piperonyl -O-CH ₂ -CF ₂ -O-(CF ₂ -CF ₂ -O) _m -(CF ₂ -O) _n -CF ₂ -O-piperonyl ^a	2300	80
Fomblin Z-DISOC	$\text{O}-\text{CN}-\text{C}_6\text{H}_3-(\text{CH}_3)-\text{NH}-\text{CO}-\text{CF}_2-\text{O}-(\text{CF}_2-\text{CF}_2-\text{O})_n-(\text{CF}_2-\text{O})_m-\text{CF}_2-\text{CO}-\text{NH}-\text{C}_6\text{H}_3-(\text{CH}_3)-\text{N}-\text{CO}$	1500	160
Fomblin YR	$\begin{array}{c} \text{CF}_3 \\ \\ \text{CF}_3-\text{O}- (\text{C}-\text{CF}_2-\text{O})_m(\text{CF}_2-\text{O})_n-\text{CF}_3 \quad (\text{m/n} \sim 40/1) \\ \\ \text{F} \end{array}$	6800	1600
Demnum S-100	$\text{CF}_3-\text{CF}_2-\text{CF}_2-\text{O}-(\text{CF}_2-\text{CF}_2-\text{CF}_2-\text{O})_m-\text{CF}_2-\text{CF}_3$	5600	250
Krytox 143AD	$\begin{array}{c} \text{CF}_3 \\ \\ \text{CF}_3-\text{CF}_2-\text{CF}_2-\text{O}- (\text{C}-\text{CF}_2-\text{O})_m-\text{CF}_2-\text{CF}_3 \\ \\ \text{F} \end{array}$	2600	—

^a3,4 – methylenedioxybenzyl

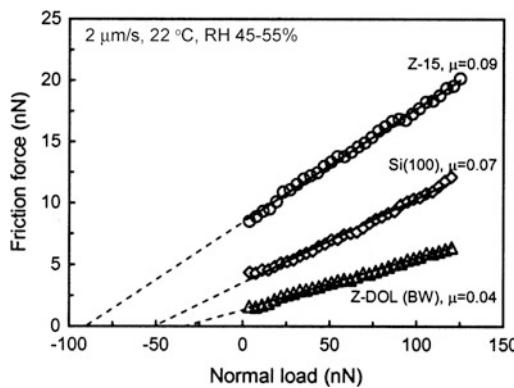


Fig. 15.48 Friction force versus normal load curves for Si(100), 2.8-nm-thick Z-15 film, and 2.3-nm-thick Z-DOL (BW) film at 2 $\mu\text{m}/\text{s}$, and in ambient air sliding against a Si_3N_4 tip. Based on these curves, coefficient of friction (μ) and adhesive force can be calculated (Liu and Bhushan 2003)

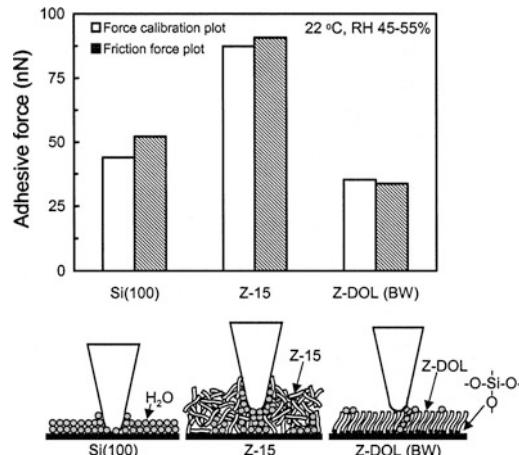


Fig. 15.49 Summary of the adhesive forces of Si(100), 2.8-nm-thick Z-15 film, and 2.3-nm-thick Z-DOL(BW) film. The schematic (bottom) shows the effect of meniscus formation between the AFM tip and the sample surface on the adhesive and friction forces (Liu and Bhushan 2003)

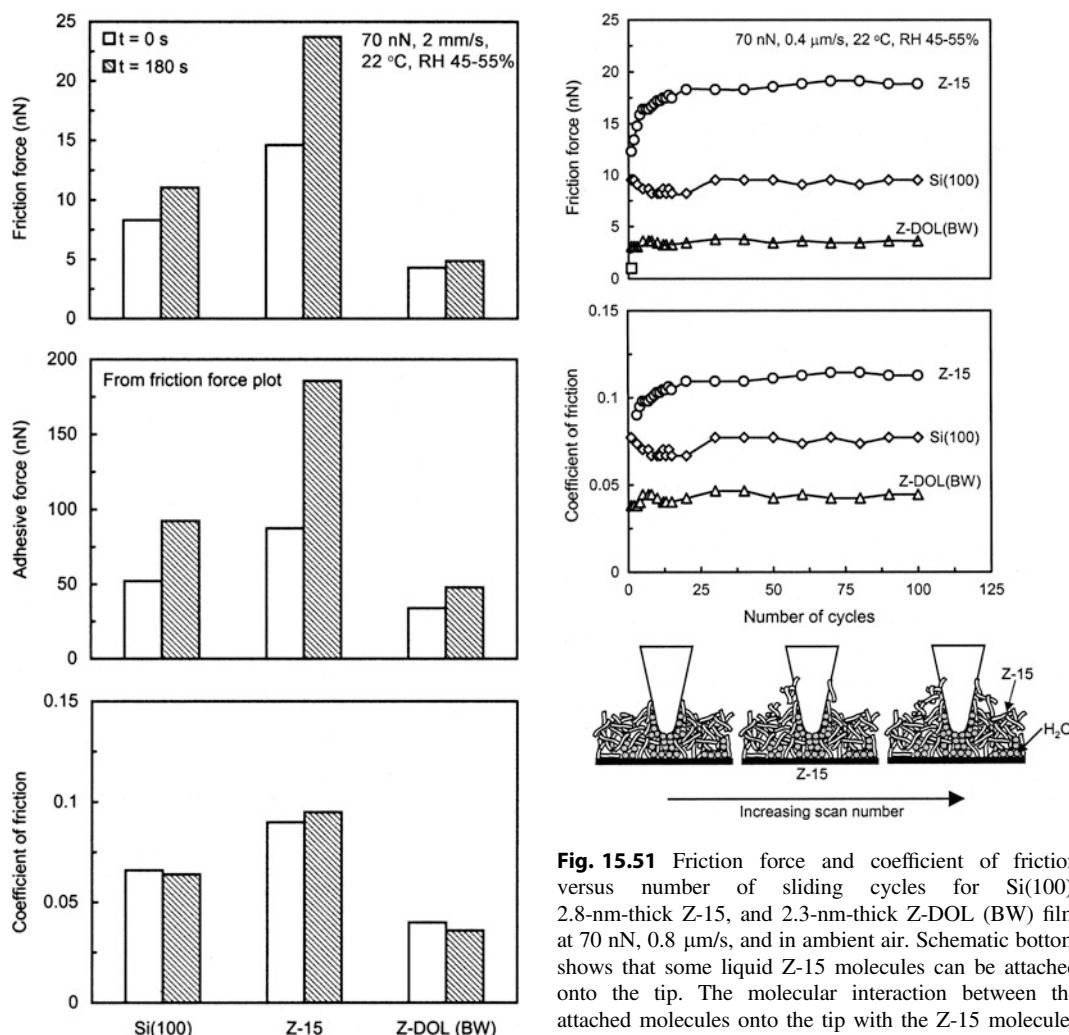


Fig. 15.50 Summary of rest time effect on friction force, adhesive force, and coefficient of friction of Si(100), 2.8-nm-thick Z-15, and 2.3-nm-thick Z-DOL (BW) film (Liu and Bhushan 2003)

mechanical properties followed by ion beam and sputtering processes. AFM has been modified for nanoindentation hardness measurements with depth of indentation as low as 5 nm. Scratching and indentation on nanoscales are the powerful ways of evaluation of the mechanical integrity of ultrathin films.

AFM/FFM friction experiments show that lubricants with polar (reactive) end groups dramatically increase the load or contact pressure that a liquid film can support before solid-solid

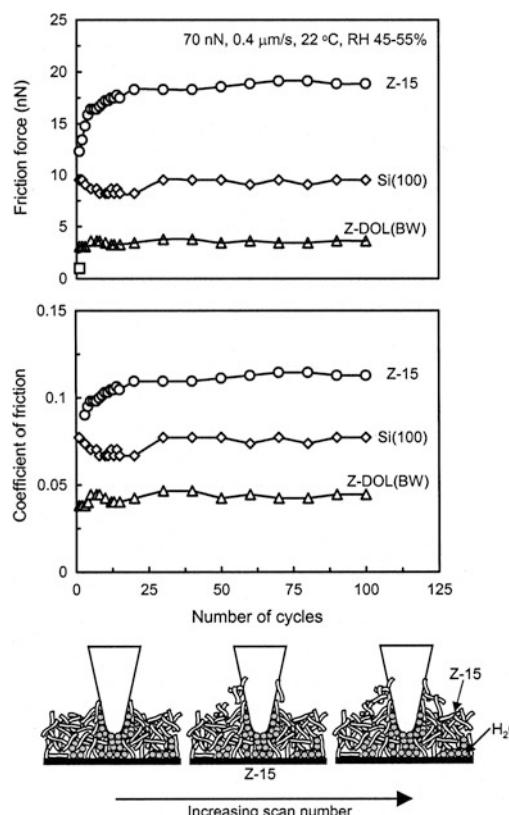


Fig. 15.51 Friction force and coefficient of friction versus number of sliding cycles for Si(100), 2.8-nm-thick Z-15, and 2.3-nm-thick Z-DOL (BW) film at 70 nN, 0.8 $\mu\text{m}/\text{s}$, and in ambient air. Schematic bottom shows that some liquid Z-15 molecules can be attached onto the tip. The molecular interaction between the attached molecules onto the tip with the Z-15 molecules in the film results in an increase of the friction force with multiscanning (Liu and Bhushan 2003)

contact and thus exhibit long durability. The lubricants with the absence of mobile liquid exhibit low friction and adhesion and don't exhibit the rest time effect.

References

- Bhushan B (1993a) Magnetic recording surfaces. In: Glaeser WA (ed) Characterization of tribological materials. Butterworth-Heinemann, Boston, pp 116–133
- Bhushan B (1993b) Magnetic slider/rigid disk substrate materials and disk texturing techniques—Status and future outlook. *Adv Info Storage Syst* 5:175–209

- Bhushan B (1994) Tribology of magnetic storage systems. In: *Handbook of lubrication and tribology*, vol 3. CRC Press, Boca Raton, Florida, pp 325–374
- Bhushan B (1995a) Nanotribology and its applications to magnetic storage devices and MEMS. In: Güntherodt HJ, Anselmetti D, Meyer E (eds) *Forces in scanning probe methods*, vol E 286. Kluwer Academic, Dordrecht, pp 367–395
- Bhushan B (1995b) Micro/nanotribology and its application to magnetic storage devices and MEMS. *Tribol Inter* 28:85–95
- Bhushan B (1996a) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer, New York
- Bhushan B (1996b) Contact mechanics of rough surfaces in tribology: single asperity contact. *Appl Mech Rev* 49:275–298
- Bhushan B (1997) Micro/nanotribology and its applications, NATO ASI series E: applied sciences 330. Kluwer Academic Publishers, Dordrecht, Netherlands
- Bhushan B (1998) Contact mechanics of rough surfaces in tribology: multiple asperity contact. *Tribol Lett* 4:1–35
- Bhushan B (1999a) Micro/nanotribology and micro/nanomechanics of magnetic storage devices. In: *Handbook of micro/nanotribology*, 2nd edn. CRC Press, Boca Raton, Florida, pp 691–762
- Bhushan B (1999b) "Chemical, mechanical and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments", (invited). *Diam Relat Mater* 8:1985–2015
- Bhushan B (1999c) Nanoscale tribophysics and tribomechanics. *Wear* 225–229:465–492
- Bhushan B (2000) Mechanics and reliability of flexible magnetic media, 2nd edn. Springer, New York
- Bhushan B (2001a) Fundamentals of tribology and bridging the gap between the macro- and micro/nanoscales, NATO science series II: mathematics, physics, and chemistry, vol 10. Kluwer Academic Publishers, Dordrecht, Netherlands
- Bhushan B (2001b) Macro- and microtribology of magnetic storage devices. In: *Modern tribology handbook*, Vol. 2—materials, coatings, and industrial applications. CRC Press, Boca Raton, Florida, pp 1413–1513
- Bhushan B (2003) Adhesion and stiction: mechanisms, measurement techniques, and methods for reduction. *J Vac Sci Technol B* 21:2262–2296
- Bhushan B (2005) Nanotribology and nanomechanics. *Wear* 259:1507–1531
- Bhushan B (2008) Nanotribology, nanomechanics, and nanomaterials characterization. *Phil Trans R Soc A* 366:1351–1381
- Bhushan B (2013) Introduction to tribology, 2nd edn. Wiley, New York
- Bhushan B (2017) Springer handbook of nanotechnology, 4th edn. Springer International, Switzerland
- Bhushan B, Blackman GS (1991) Atomic force microscopy of magnetic rigid disks and sliders and its applications to tribology. *ASME J Tribol* 113:452–458
- Bhushan B, Kasai T (2004) A surface topography-independent friction measurement technique using torsional resonance mode in an AFM. *Nanotechnology* 15:923–935
- Bhushan B, Koinkar VN (1994a) Tribological studies of silicon for magnetic recording applications. *J Appl Phys* 75:5741–5746
- Bhushan B, Koinkar VN (1994b) Nanoindentation hardness measurements using atomic force microscopy. *Appl Phys Lett* 64:1653–1655
- Bhushan B, Koinkar VN (1995a) Microtribology of PET polymeric films. *Tribol Trans* 38:119–127
- Bhushan B, Koinkar VN (1995b) Macro and microtribological studies of CrO₂ video tapes. *Wear* 180:9–16
- Bhushan B, Koinkar VN (1995c) Microtribology of metal particle, barium ferrite and metal evaporated magnetic tapes. *Wear* 181–183:360–370
- Bhushan B, Koinkar VN (1995d) Microscale mechanical and tribological characterization of hard amorphous carbon coatings as thin as 5 nm for magnetic disks. *Surf Coat Technol* 76–77:655–669
- Bhushan B, Koinkar VN (1997) Microtribological studies of doped single-crystal silicon and polysilicon films for MEMS devices. *Sens Actuators A* 57:91–102
- Bhushan B, Li X (2003) Nanomechanical characterisation of solid surfaces and thin films. *Int Mater Rev* 48:125–164
- Bhushan B, Majumdar A (1992) Elastic-plastic contact model for bifractal surfaces. *Wear* 153:53–64
- Bhushan B, Peng W (2002) "Contact mechanics of multilayered rough surfaces", (invited). *Appl Mech Rev* 55:435–480
- Bhushan B, Qi J (2003) Phase contrast imaging of nanocomposites and molecularly-thick lubricant films in magnetic media. *Nanotechnology* 14:886–895
- Bhushan B, Ruan J (1994) Atomic-scale friction measurements using friction force microscopy: part II—application to magnetic media. *ASME J Tribol* 116:389–396
- Bhushan B, Sundararajan S (1998) Micro/nanoscale friction and wear mechanisms of thin films using atomic force and friction force microscopy. *Acta Mater* 46:3793–3804
- Bhushan B, Zhao Z (1999) Macro- and microscale studies of molecularly-thick boundary layers of perfluoropolyether lubricants for magnetic thin-film rigid disks. *J Info Storage Proc Syst* 1:1–21
- Bhushan B, Dominiak M, Lazzari JP (1992) Contact-start-stop studies with silicon planar head sliders against thin-film disks. *IEEE Trans Mag* 28:2874–2876
- Bhushan B, Koinkar VN, Ruan J (1994) Microtribology of magnetic media. *Proc Instn Mech Eng Part J: J Eng Tribol* 208:17–29
- Bhushan B, Israelachvili JN, Landman U (1995a) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Miyamoto T, Koinkar VN (1995b) Microscopic friction between a sharp diamond tip and

- thin-film magnetic rigid disks by friction force microscopy. *Adv Info Storage Syst* 6:151–161
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996) Nanoindentation and picoindentation measurements using a capacitance transducer system in atomic force microscopy. *Philos Mag* 74:1117–1128
- Bhushan B, Sundararajan S, Scott WW, Chilamakuri S (1997) Stiction analysis of magnetic tapes. *IEEE Trans Magn* 33:3211–3213
- Castillo J, Bhushan B (2006) Tribology of hard disk drives—magnetic data storage technology. In: Totten E (ed) *Handbook of lubrication and tribology: vol 1—application and maintenance*, 2nd edn. CRC Press, Taylor & Francis Group, Boca Raton, Florida, pp. 16-1–16-44
- Chu MY, Bhushan B, DeJonghe L (1992) Wear behavior of ceramic sliders in sliding contact with rigid magnetic thin-film disks. *Tribol Trans* 35:603–610
- DeVecchio D, Bhushan B (1997) Localized surface elasticity measurements using an atomic force microscope. *Rev Sci Instrum* 68:4498–4505
- Ganti S, Bhushan B (1995) Generalized fractal analysis and its applications to engineering surfaces. *Wear* 180:17–34
- Heideman R, Wirth M (1984) Transforming the lubricant on a magnetic disk into a solid fluorine compound. *IBM Tech Discl Bull* 27:3199–3205
- Hibst H (1993) Metal evaporated tapes and Co-Cr media for high definition video recording. In: Buschow KHJ, Long GJ, Grandjean F (eds) *High density digital recording*, vol E229. Kluwer Academic Publishers, Dordrecht, Netherlands, pp 137–159
- Homola AM, Lin LJ, Saperstein DD (1990) Process for bonding lubricant to a thin film magnetic recording disk. U.S. Patent 4,960,609, October 2
- Kasai T, Bhushan B, Huang L, Su C (2004) Topography and phase imaging using the torsional resonance mode. *Nanotechnology* 15:731–742
- Koinkar VN, Bhushan B (1996a) Microtribological studies of Al_2O_3 , Al_2O_3 -TiC, polycrystalline and single-crystal Mn-Zn ferrite and SiC head slider materials. *Wear* 202:110–122
- Koinkar VN, Bhushan B (1996b) Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J Vac Sci Technol A* 14:2378–2391
- Koinkar VN, Bhushan B (1997a) Effect of scan size and surface roughness on microscale friction measurements. *J Appl Phys* 81:2472–2479
- Koinkar VN, Bhushan B (1997b) Microtribological properties of hard amorphous carbon protective coatings for thin-film magnetic disks and heads. *Proc Instn Mech Eng Part J: J Eng Tribol* 211:365–372
- Kulkarni AV, Bhushan B (1997) Nanoindentation measurements of amorphous carbon coatings. *J Mater Res* 12:2707–2714
- Li X, Bhushan B (1999a) Micro/nanomechanical and tribological characterization of ultra-thin amorphous carbon coatings. *J Mater Res* 14:2328–2337
- Li X, Bhushan B (1999b) Mechanical and tribological studies of ultra-thin hard carbon overcoats for magnetic recording heads. *Z Metallkd* 90:820–830
- Liu H, Bhushan B (2003) Nanotribological characterization of molecularly-thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* 97:321–340
- Majumdar A, Bhushan B (1990) Role of fractal geometry in roughness characterization and contact mechanics of surfaces. *ASME J Tribol* 112:205–216
- Majumdar A, Bhushan B (1991) Fractal model of elastic-plastic contact between rough surfaces. *ASME J Tribol* 113:1–11
- Mate CM (1992) Application of disjoining and capillary pressure to liquid lubricant films in magnetic recording. *J Appl Phys* 72:3084–3090
- Mate CM, Novotny VJ (1991) Molecular conformation and disjoining pressures of polymeric liquid films. *J Chem Phys* 94:8420–8427
- Novotny VJ (1990) Migration of liquid polymers on solid surfaces. *J Chem Phys* 92:3189–3196
- Novotny VJ, Hussla I, Turlet JM, Philpott MR (1989) Liquid polymer conformation on solid surfaces. *J Chem Phys* 90:5861–5868
- Oden PI, Majumdar A, Bhushan B, Padmanabhan A, Graham JJ (1992) AFM imaging, roughness analysis and contact mechanics of magnetic tape and head surfaces. *ASME J Tribol* 114:666–674
- Poon CY, Bhushan B (1995a) Comparison of surface roughness measurements by stylus profiler, AFM and non-contact optical profiler. *Wear* 190:76–88
- Poon CY, Bhushan B (1995b) Surface roughness analysis of glass-ceramic substrates and finished magnetic disks, and Ni-P coated Al-Mg and glass substrates. *Wear* 190:89–109
- Ruan J, Bhushan B (1994a) Atomic-scale friction measurements using friction force microscopy: part I—general principles and new measurement techniques. *ASME J Tribol* 116:378–388
- Ruan J, Bhushan B (1994b) Frictional behavior of highly oriented pyrolytic graphite. *J Appl Phys* 76:8117–8120
- Saperstein DD, Lin LJ (1990) Improved surface adhesion and coverage of perfluoropolyether lubricant following far-UV irradiation. *Langmuir* 6:1522–1524
- Scherer V, Bhushan B, Rabe U, Arnold W (1997) Local elasticity and lubrication measurements using atomic force and friction force microscopy at ultrasonic frequencies. *IEEE Trans Magn* 33:4077–4079
- Scott WW, Bhushan B (2003) Use of phase imaging in atomic force microscopy for measurement of viscoelastic contrast in polymer nanocomposites and molecularly-thick lubricant films. *Ultramicroscopy* 97:151–169
- Sundararajan S, Bhushan B (1998) Micro/nanotribological studies of polysilicon and SiC films for MEMS applications. *Wear* 217:251–261
- Sundararajan S, Bhushan B (1999) Micro/nanotribology of ultra-thin hard amorphous carbon coatings using

- atomic force/friction force microscopy. *Wear* 225–229:678–689
- Sundararajan S, Bhushan B (2000) Topography-induced contributions to friction forces measured using an atomic force/friction force microscope. *J Appl Phys* 88:4825–4831
- Sundararajan S, Bhushan B (2001) Development of a continuous microscratch technique in an atomic force microscopy and its applications to study scratch resistance of ultra-thin hard amorphous carbon coatings. *J Mater Res* 16:437–445

Nanotribology and Nanomechanics of MEMS/NEMS and BioMEMS/BioNEMS Materials and Devices

16

Bharat Bhushan

Abstract

A number of micro/nanoelectromechanical systems (MEMS/NEMS) and BioMEMS/BioNEMS are used in commercial applications and/or are under development. Surface area to volume ratio in MEMS/NEMS is large, and surface forces such as adhesion, friction, and meniscus and viscous forces become very large compared to inertial and electromagnetic forces. Some devices are designed to execute expected functions with short durations, typically in the millisecond to picosecond range. The expected life of the devices for high speed contacts can vary from a few hundred thousand to many billions of cycles, e.g., over a hundred billion cycles for digital micromirror devices (DMDs). This puts serious requirements on materials in systems involving relative motion. Thus, there is a need for a fundamental understanding of adhesion, friction/stiction, wear, lubrication, and the role of surface contamination and environment, all on the nanoscale. Most mechanical properties are known to be scale dependent. Therefore, the properties of nanoscale structures need to be measured. For BioMEMS/BioNEMS, adhesion between biological molecular layers and the substrate, and friction and wear of biological layers can be important. Component-level studies are required to provide a better understanding of the tribological phenomena occurring in MEMS/NEMS. This chapter presents an overview of nanoscale adhesion, friction, and wear studies of materials and lubrication for MEMS/NEMS and BioMEMS/BioNEMS, and component-level studies of stiction phenomena in MEMS/NEMS devices.

B. Bhushan (✉)
Nanoprobe Laboratory for Bio- and Nanotechnology
and Biomimetics, The Ohio State University,
201 W. 19th Ave., W390 Scott Laboratory,
Columbus, OH 43210, USA
e-mail: bhushan.2@osu.edu

Keywords

Microtechnology • Nanotechnology • MEMS • NEMS • Micro/nanodevices • Micro/nanosystems • Biosensors • Nanotribology • Nanomechanics • Lubrication • Molecularly thick films • Reliability • Component level studies

16.1 Introduction

Micro- and nanosystems include Micro/Nano ElectroMechanical Systems (MEMS/NEMS). The acronym MEMS originated in the U.S. The term commonly used in Europe is micro/nanosystem technology (MST/NST), and in Japan it is micro/nanomachines. Another term generally used is micro/nanodevices. Advances in silicon photolithographic process technology since the 1960s led to the development of MEMS in the early 1980s. MEMS combine electrical and mechanical components to create microscopic devices that have a characteristic length between 100 nm and 1 mm. In the early 2000s, with the development of nanofabrication techniques, NEMS were fabricated. NEMS combine electrical and mechanical components to create nanoscopic devices that have a characteristic length of less than 100 nm. In mesoscale devices, if the functional components are on the micro- or nanoscale, they may be referred to as MEMS or NEMS, respectively. These are intelligent miniaturized systems comprised of sensing, processing, and/or actuating functions that combine electrical and mechanical components (Bhushan 2016a). MEMS/NEMS terms are also now used in a broad sense and include electrical, mechanical, fluidic, optical, and/or biological function. MEMS/NEMS for optical applications are referred to as micro/nanooptoelectromechanical systems (MOEMS/NOEMS). MEMS/NEMS for electronic applications are referred to as radio-frequency-MEMS/NEMS or RF-MEMS/RF-NEMS. MEMS/NEMS for biological applications are referred to as BioMEMS/BioNEMS.

To put the dimensions and weights of nanomaterials, MEMS/NEMS, and BioNEMS in

perspective, see Fig. 16.1 and Table 16.1. Examples shown are a single walled carbon nanotube (SWNT) chemical sensor (adapted from Chen et al. 2004), molecular dynamic simulations of carbon-nanotube based gears (adapted from Srivastava 2004), quantum-dot transistor (adapted from van der Wiel et al. 2003), and a digital micromirror device (adapted from Hornbeck 1999). Individual atoms are typically a fraction of a nanometer in diameter, DNA molecules are about 2.5 nm wide, biological cells are in the range of thousands of nm in diameter, and human hair is about 75,000 nm (75 μm) in diameter. The smaller length of a BioNEMS shown in the figure is about 2 nm. NEMS shown range in size from 10 to 300 nm, and the size of MEMS is on the order of 12,000 nm (12 μm). The weight of a micromachined silicon structure can be as low as 1 nN, and NEMS can be built with weight as low as 10^{-20} N with cross sections of about 10 nm. In comparison, the weight of a drop of water is about 10 μN , and the weight of an eyelash is about 100 nN.

These devices are fabricated by manufacturing at the nanoscale, referred to as nanomanufacturing. It is accomplished by using either a “bottom-up” or “top down” approach to the production of nanomaterials, structures, devices, and systems (Madou 2011; Bhushan et al. 2014; Bhushan 2016a).

In the bottom-up approach, nanoscale features are mainly built up from their elemental constituents. Elemental constituents are combined using various nanomanufacturing processes including self-assembly, chemical synthesis, molecular beam epitaxy, nanoimprint lithography, roll-to-roll processing, and dip pen lithography (see, nano.gov). Molecular self-assembly, the spontaneous self-assembly of molecular clusters,

Fig. 16.1 Characteristic dimensions of nanomaterials, MEMS/NEMS, and BioNEMS in perspective (Bhushan 2015). Examples shown are a single walled carbon nanotube (SWNT) chemical sensor (adapted from Chen et al. 2004), molecular dynamic simulations of carbon-nanotube based gears (adapted from Srivastava 2004), quantum-dot transistor (adapted from van der Wiel et al. 2003), and a digital micromirror device (adapted from Hornbeck 1999)

Characteristic Dimensions of MEMS/NEMS

MEMS - less than 1 mm, larger than 100 nm
NEMS - less than 100 nm

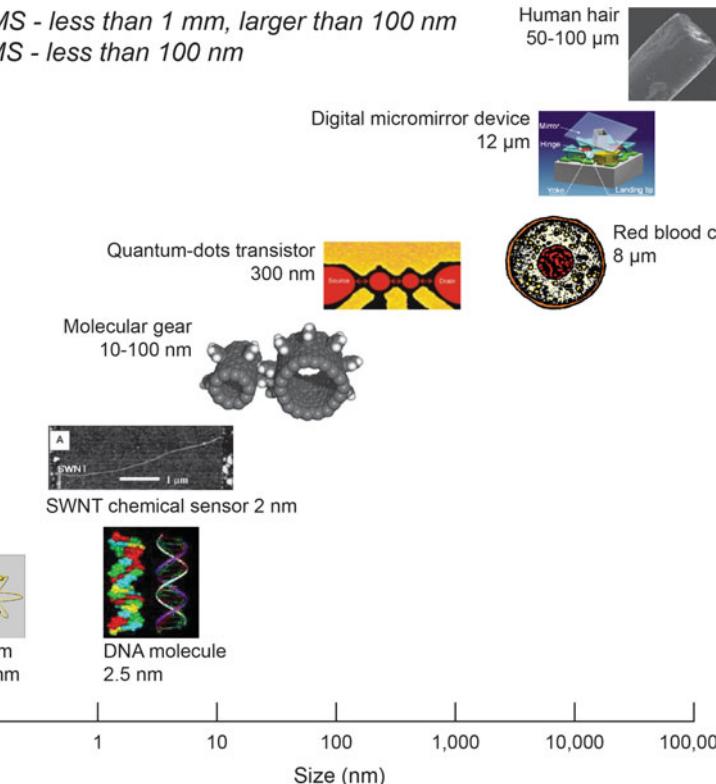


Table 16.1 Characteristic dimensions and weights in perspective

(a) Characteristic dimensions in perspective

NEMS characteristic length: <100 nm

MEMS characteristic length: <1 mm and >100 nm

Single-walled carbon nanotube (SWNT) chemical sensor	~2 nm
Molecular gear	~10 nm
Quantum-dots transistor	300 nm
Digital micromirror	12,000 nm
Individual atoms	Typically fraction of a nm in diameter
DNA molecules	~2.5 nm wide
Biological cells	In the range of thousands of nm in diameter
Human hair	~75,000 nm in diameter

(b) Weights in perspective

NEMS built with cross sections of about 10 nm	As low as 10^{-20} N
Micromachines silicon structure	As low as 1 nN
Eyelash	~100 nN
Water droplet	~10 μN

can occur from simple reagents in solution or from biological molecules (e.g., DNA) used as building blocks for the production of three-dimensional nanostructures. Chemical synthesis is carried out using gaseous precursors and solvents. These chemicals react to produce nanostructures. A variety of vacuum deposition and nonequilibrium plasma chemistry techniques are then used to produce layered nanocomposites and nanotubes. Molecular beam epitaxy is a method for deposition of thin films with a thickness of one atom on a surface. Nanoprint lithography is used to fabricate nanostructures by stamping or printing them on a surface. Roll-to-roll processing is a high volume process used to produce nanodevices on a flexible substrate. Dip pen lithography uses an atomic force microscope tip in a fluid to produce nanoscale features on a surface.

The top down approach uses lithographic and non-lithographic fabrication technologies. Lithographic technology is an underlying technology to make semiconductor chips and components. Continued improvements in lithography for use in the production of nanocomponents have resulted in line widths as small as 10 nm. The top down approach is commonly used to fabricate micro- and nanosystem components, which range in size from micro- to nanometers.

MEMS/NEMS and BioMEMS/BioNEMS are beginning to have a major impact on our lives, comparable to that of semiconductor technology, information technology, or cellular and molecular biology in the 20th century. The integration of micro/nanodevices/systems into larger systems has given breakthrough solutions to many current environmental, medical, and industrial problems including smart materials, nanomanufacturing, electronics, drug delivery, energy, water, biotechnology, information technology, and national security. BioMEMS/BioNEMS are increasingly used in commercial and defense applications. The revenues from nano-enabled products continue growing, with over \$200B in FY 2012 in the US alone, and over \$700B worldwide (Bhushan 2015). MEMS/NEMS and BioMEMS/BioNEMS will have a profound impact on our society. They represent a mega-trend, bringing disruptive innovation.

16.1.1 Introduction to MEMS

The advances in silicon photolithographic process technology since the 1960s led to the development of MEMS in early 1980s. Later, lithographic processes were developed to process nonsilicon materials. The lithographic processes are complemented with nonlithographic processes for the fabrication of components or devices made from plastics or ceramics. Using these fabrication processes, researchers have fabricated a wide variety of devices with dimensions in the range of submicron to a few thousand microns (see e.g., Muller et al. 1990; Fujimasa 1996; Trimmer 1997; Bhushan 1998; Kovacs 1998; Senturia 2000; Elwenspoek and Wiegerink 2001; Gad-el-Hak 2002; Hsu 2002; Hierlemann 2005; Madou 2011). MEMS for mechanical applications include acceleration, pressure, flow, and gas sensors, linear and rotary actuators, and other microstructures or micro-components such as electric motors, gear chains, gas turbine engines, fluid pumps, fluid valves, switches, grippers, and tweezers. MEMS for chemical applications include chemical sensors and various analytical instruments. MOEMS are the devices that include optical components, such as micromirror arrays for displays, infrared image sensors, spectrometers, bar code readers, and optical switches. RF-MEMS include inductors, capacitors, antennas, and RF-switches. High-aspect-ratio MEMS (HARMEMS) have also been introduced.

A wide variety of MEMS devices have been produced and some are used commercially (Fujimasa 1996; Bhushan 1998; Kovacs 1998; Senturia 2000; Elwenspoek and Wiegerink 2001; Hsu 2002; Hierlemann 2005; Madou 2011). A variety of sensors are used in industrial, consumer, defense, and biomedical applications. The largest “killer” industrial applications include accelerometers, pressure sensors, thermal and piezoelectric type inkjet printheads, and digital micromirror devices. Integrated capacitive-type silicon accelerometers have been used in airbag deployment in automobiles since 1991 (Core et al. 1993; Bryzek et al. 1994). Some 90 million units were installed in vehicles in 2004.

Accelerometer technology was over a billion-dollar-a-year industry in 2004, dominated by Analog Devices followed by Freescale Semiconductor (formerly Motorola) and Bosch. It is expected to grow with an annual growth exceeding 30%. Tri-axis accelerometers are needed to describe three-dimensional motion. Multi-axis accelerometers are being used for many other applications such as vehicle stability, rollover control, and gyro sensors for automotive applications. They are used also in various consumer applications including handheld devices, e.g., laptops for free-fall detection (2003), and cellular phones (2004) and tablets for menu navigation, gaming, image rotation, and free-fall detection. Silicon-based piezoresistive pressure sensors were launched in 1990 by GE NovaSensor for manifold absolute pressure (MAP) sensing for engines and for disposable blood pressure sensors, and their annual sales were more than 30 million units and more than 25 million units, respectively, in 2004. MAP sensors measure the pressure in the intake manifold, which is fed to a computer that determines the optimum air/fuel mixture to maximize fuel economy. Most vehicles have these as part of the electronic engine control system. Capacitive pressure sensors for tire pressure measurements were launched by Freescale Semiconductor (formerly Motorola) in early 2000, and are also manufactured by Infineon/SensoNor and GE Novasensor (2003). Piezoresistive type sensors are also used, manufactured by various companies such as EnTire Solutions (2003). The sensing module is located inside the rim of the wheel and relays the information via radio frequency to a central processing unit (CPU) in order to display it to the driver. In 2005, about 9.2 million vehicles were equipped with sensors, which translated to about 37 million units. Their sales have grown rapidly as they are now required in automobiles in the U. S. (starting in 2008), and it affects 17 million vehicles (one in each tire) sold every year. Pressure sensors can be used to detect altitude by measuring air pressure. For example, global positioning systems (GPS) used for navigation have good lateral resolution but poor vertical resolution, which creates problems in

identifying the level in the case of multiple roads stacked up. A pressure sensor is needed to determine the level (altitude) by measuring the air pressure.

Thermal inkjet printers were developed independently by HP and Cannon and were commercialized in 1984 (Aden et al. 1994; Le 1998; Baydo and Groskopf 2001; Lee 2003). Today they are made by Canon, Epson, HP, Lexmark, Xerox, and others. They typically cost less initially than dry-toner laser printers and are the solution of choice for low-volume print runs. Annual sales of thermal inkjet printheads with microscale functional components were more than 500 million units in 2004.

Micromirror arrays are used for digital projection displays. Commercial digital light processing (DLP) equipment, using digital micromirror devices (DMD), were launched in 1996 by Texas Instruments for digital projection displays in computer projectors, high definition television (HDTV) sets, and movie projectors (DLP cinema) (Hornbeck and Nelson 1988; Hornbeck 1999, 2001). Several million projectors had been sold by 2004 (about \$700 million revenue by TI in 2004). Electrostatically actuated, membrane type or cantilever type microswitches have been developed for direct current (DC), RF, and optical applications (Suzuki 2002). There exists two basic forms of RF microswitches: the metal-to-metal contact microswitch (ohmic) and the capacitive microswitch. RF microswitches can be used in a variety of RF applications including cellular phones, phase shifters, smart antennas, multiplexers for data acquisition, and more (Lubecke and Chiao 1999). Optical microswitches are finding applications in optical networking, telecommunications, and wireless technologies (Giles et al. 2001; Suzuki 2002).

Other applications of MEMS devices include chemical, biological, and gas sensors (Hierlemann et al. 2003; Hierlemann 2005), actuators, microresonators, infrared detectors, and focal plane arrays for earth observation, space science, and missile defense applications, pico-satellites for space applications, fuel cells, and many hydraulic, pneumatic, and other consumer

products. MEMS devices are also being pursued in magnetic storage systems (Bhushan 1996a), where they are being developed for super compact and ultrahigh-recording-density magnetic disk drives. Several integrated head/suspension microdevices have been fabricated for contact recording applications (Hamilton 1991). High-bandwidth servo-controlled microactuators have been fabricated for ultrahigh-track-density applications which serve as the fine-position control element of a two-stage, coarse/fine servo system, coupled with a conventional actuator (Horsley et al. 1998; Hirano et al. 2000).

Micro/nano-instruments and micro/nano-manipulators are used to move, position, probe, pattern, and characterize nanoscale objects and nanoscale features (Fukuda et al. 2003). Miniaturized analytical equipment include gas chromatography and mass spectrometry. Other instruments include micro-STM, where STM stands for scanning tunneling microscope.

In some cases, MEMS devices are used primarily for their miniature size, while in others, as in the case of the air bags, because of their low-cost manufacturing techniques. This latter fact has been possible since semiconductor-processing costs have reduced drastically over the last decade, allowing the use of MEMS in many fields.

16.1.2 Introduction to NEMS

NEMS are produced by nanomachining in a typical top-down approach (from large to small) and bottom-up approach (from small to large) largely relying on nanochemistry (see e.g., Drexler 1992; Timp 1999; Dresselhaus et al. 2001; Rietman 2001; Goddard et al. 2002; Nalwa 2002; Poole and Owens 2003). The NEMS field, in addition to the fabrication of nanosystems, has provided the impetus to develop experimental and computation tools. Examples of NEMS include microcantilevers with integrated sharp nanotips for STM and atomic force microscopy (AFM) (Bhushan 1999a, 2005), quantum corral formed using STM by placing atoms one by one (Stroscio and Eigler 1991), AFM cantilever

arrays (Millipede) for data storage (Vettiger et al. 1999), STM and AFM tips for nanolithography, dip-pen nanolithography for printing molecules, nanowires, carbon nanotubes, quantum wires (QWRs), quantum boxes (QBs), quantum transistors (van der Wiel et al. 2003), nanotube based sensors (Stampfer et al. 2004; Hierold 2008), biological (DNA) motors, molecular gears by attaching benzene molecules to the outer walls of carbon nanotubes (Srivastava 2004), devices incorporating nm-thick films (e.g., in giant magnetoresistive or GMR read/write magnetic heads and magnetic media for magnetic rigid disk and magnetic tape drives), nanopatterned magnetic rigid disks, and nanoparticles (e.g., nanoparticles in magnetic tape substrates and nanomagnetic particles in magnetic tape coatings) (Bhushan 1996a, 2000). More than 2 billion read/write magnetic heads were shipped for magnetic disk and tape drives in 2004.

Nanoelectronics can be used to build computer memory using individual molecules or nanotubes to store bits of information (Anonymous 2004), molecular switches, molecular or nanotube transistors, nanotube flat-panel displays, nanotube integrated circuits, fast logic gates, switches, nanoscopic lasers, and nanotubes as electrodes in fuel cells.

16.1.3 Introduction to BioMEMS/BioNEMS

The basic building-block of all life is the cell (Lodish et al. 2003). Lipid bilayer membranes enclose all animal cells. The internal compartment is a hydrophilic space called cytosol, and is filled with the material necessary to sustain life. The human body consists of on the order of 30 trillion human cells, dominated by red blood cells, and on the order of 40 trillion bacterial cells, dominated by colonic bacteria. Bacteria help digest food, but they can also generate toxins (Koch 2001). Proteins are large molecules and consist of long chains of amino acids (Nelson and Cox 2005). Complete proteins contain all eight essential amino acids. Proteins are a vital part of all living things. They are important

to every cell operation, and perform a vast array of functions including catalyzing metabolic reactions, DNA replication, and stimuli response. However, their main function is to heal wounds, fight infection, and build muscle.

BioMEMS/BioNEMS are used for chemical and biochemical analyses (biosensors) in medical diagnostics (e.g., DNA, RNA, proteins, cells, blood pressure and assays, and bacterial pathogens or toxin identification) (Gravesen et al. 1993; van der Berg 2003), tissue engineering (Bhatia and Chen 1999; Lanza et al. 2000; Leclerc et al. 2004), and implantable pharmaceutical drug delivery (Park 1997; Shawgo et al. 2002; Oeberg et al. 2004). They are increasingly used in commercial and defense applications (see e.g., Manz and Becker 1998; Cheng and Kricka 2001; Heller and Guttman 2001; San and Yap 2001; Mastrangelo and Becker 2001; Becker and Locascio 2002; Beebe et al. 2002; van den Berg 2003).

Biosensors, also referred to as biochips, deal with liquids and gases. A large variety of biosensors are based on micro/nanofluidics (Zoval and Madou 2000; Tang and Lee 2001; Taylor et al. 2003; van der Berg 2003; Prakash et al. 2012). Micro/nanofluidic devices offer the ability to work with smaller reagent volumes and shorter reaction times, and perform multiple types of analyses at once. A second type of biosensors includes micro/nanoarrays which perform one type of analysis thousands of times (Raiteri et al. 2001; Bhushan et al. 2005a; Lang et al. 2005; Patolsky and Lieber 2005; Nicholson et al. 2010; Gupta et al. 2011; Casal et al. 2012).

A chip with micro/nanofluidic technology embedded on the disk, called lab-on-a-CD, can test thousands of biological samples rapidly and automatically (Zoval and Madou 2000). An entire laboratory can be integrated onto a single chip, called lab-on-a-chip (Tang and Lee 2001; Taylor et al. 2003; van der Berg 2003). Silicon-based disposable blood-pressure sensor chips were introduced in early 1990s by GE NovaSensor for blood pressure monitoring (about 25 million units in 2004). A blood sugar monitor, referred to as GlucoWatch, was introduced in 2002. It automatically checks blood sugar every 10 min by detecting glucose through

the skin, without having to draw blood. If glucose is out of the acceptable range, it sounds an alarm so the diabetic can address the problem quickly. A variety of biosensors, many using plastic substrates, are manufactured by various companies including ACLARA, Agilent Technologies, CaliperTech, and I-STAT.

Micro/nanoarrays are another type of biochip. They are a tool used in biotechnology research to analyze DNA or proteins to diagnose diseases or discover new drugs. Also called DNA arrays, they can identify thousands of genes simultaneously (San and Yap 2001; Raiteri et al. 2001). They include a microarray of silicon nanowires, roughly a few nm in size, to selectively bind and detect even a single biological molecule such as DNA or protein by using nanoelectronics to detect the slight electrical charge caused by such binding, or a microarray of carbon nanotubes to electrically detect glucose.

After the tragedy of September 11, 2001, concern about biological and chemical warfare has led to the development of handheld units with bio- and chemical sensors for detection of biological toxins, chemical or nerve agents, mustard agents, and chemical precursors to protect subways, airports, water supply, and the population at large (Scott 2003).

BioMEMS/BioNEMS are also being developed for minimally invasive surgery including endoscopic surgery, laser angioplasty, and microscopic surgery. Implantable artificial organs can also be produced. Other applications include implantable drug-delivery devices, e.g., micro/nanoparticles with drug molecules encapsulated in functionalized shells for site-specific targeting applications, and a silicon capsule with a nanoporous membrane filled with drugs for long term delivery (Park 1997; Desai et al. 1999; Hansford et al. 2001; Martin and Grove 2001), nanodevices for sequencing single molecules of DNA in the Human Genome Project (van der Berg 2003), cellular growth using carbon nanotubes for spinal cord repair, nanotubes for nanostructured materials for various applications such as spinal fusion devices, organ growth, and growth of artificial tissues using nanofibers.

16.2 Nanotribology and Nanomechanics Issues in MEMS/NEMS and BioMEMS/BioNEMS

Tribological issues are important in MEMS/NEMS and BioMEMS/BioNEMS requiring intended and/or unintended relative motion. Nanomechanical properties are scale dependent (Bhushan et al. 1996); therefore, these should be measured at relevant scales. In these devices, various forces associated with the device scale down as size decreases. When the length decreases from 1 mm to 1 μm , the surface area decreases by a factor of a million, and the volume decreases by a factor of a billion. As a result, surface forces such as adhesion, friction, meniscus forces, and viscous forces that are proportional to surface area, become a thousand times larger than the forces proportional to the volume, such as inertial and electromagnetic forces. In addition to the consequences of large surface-to-volume ratios, the small tolerances that these devices are designed for make physical contacts more likely, which makes them particularly vulnerable to adhesion between adjacent components. Slight particulate or chemical contamination present at the interface can be detrimental. Further, the small start-up forces and the torques available to overcome retarding forces are small, and the increase in resistive forces such as adhesion and friction become a serious tribological concern that limits the durability and reliability of MEMS/NEMS (Bhushan 1998, 1999a, 2007, 2008b, 2011, 2016a). A large lateral force required to initiate relative motion between two surfaces, large static friction, is referred to as “stiction,” which has been studied extensively in the tribology of magnetic storage systems (Bhushan 1996a, 1999a, 2001, 2003, 2011, 2013a, b). The source of stiction is generally liquid mediated adhesion with the source of liquid being process fluid or capillary condensation of the water vapor from the environment. Adhesion, friction/stiction, wear, and surface contamination affect MEMS/NEMS and BioMEMS/BioNEMS performance and in some cases, can even prevent the devices from

working. Adhesion between a biological molecular layer and the substrate (referred to as “bioadhesion”), reduction of friction and wear of biological layers, biocompatibility, and biofouling in BioMEMS/BioNEMS can be important (Bhushan et al. 2005a, 2006a). Some examples of devices which experience nanotribological problems follow.

16.2.1 MEMS

Figure 16.2 shows examples of several micro-components that can encounter tribological problems. The polysilicon electrostatic micro-motor has 12 stators and a 4-pole rotor, and is produced by surface micromachining. The rotor diameter is 120 μm , and the air gap between the rotor and stator is 2 μm (Tai et al. 1989). It is capable of continuous rotation up to speeds of 100,000 RPM. The intermittent contact at the rotor-stator interface and physical contact at the rotor-hub flange interface result in wear issues, and high stiction between the contacting surfaces limits the repeatability of operation or may even prevent the operation altogether. Next, a bulk micromachined silicon stator/rotor pair is shown with bladed rotor and nozzle guide vanes on the stator with dimensions less than 1 mm (Spearing and Chen 2001; Frechette et al. 2005). These are being developed for a high-temperature micro-gas turbine engine with rotor dimension of 4–6 mm in diameter, and an operating speed of up to 1 million RPM (with a sliding velocity in excess of 500 m/s, comparable to velocities of large turbines operating at high velocities) to achieve high specific power, up to a total of about 10 W. Erosion of blades and vanes and design of the microbearings required to operate at extremely high speeds used in the turbines are some of the concerns. Ultra-short, high-speed micro-hydrostatic gas journal bearings with a length to diameter ratio (L/D) of less than 0.1 are being developed for operation at surface speeds on the order of 500 m/s, which offer unique design challenges (Liu and Spakovszky 2005). Microfabrica, Inc. in the U.S.A., is developing microturbines with an outer diameter as low as

0.9 mm to be used as power sources for medical devices. They use precision ball bearings.

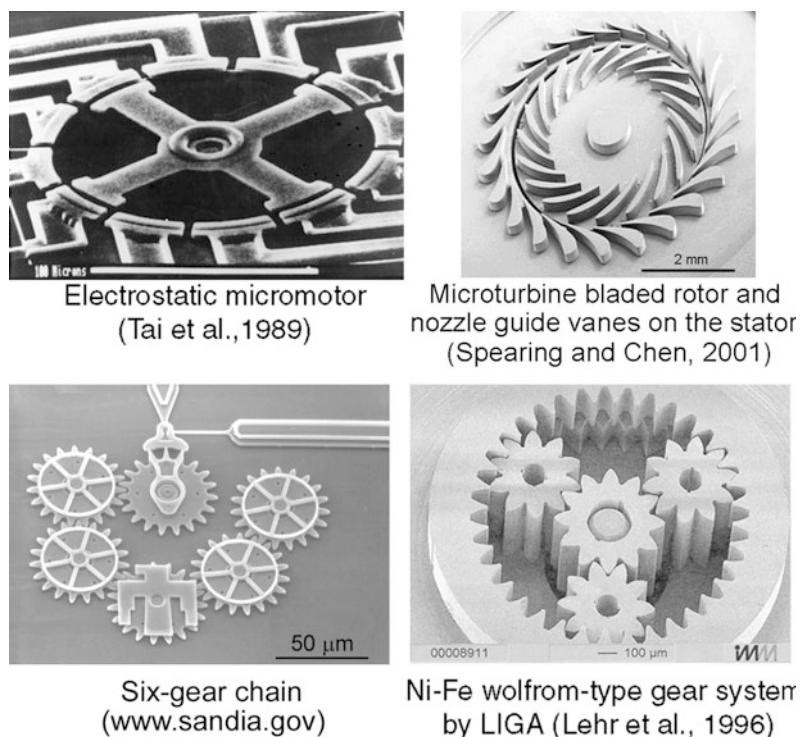
Next in Fig. 16.2 is a scanning electron microscopy (SEM) micrograph of a surface micromachined polysilicon six-gear chain from Sandia National Lab (For more examples of early version, see Mehregany et al. 1988). As an example of non-silicon components, a milligear system produced using the LIGA process for a DC brushless permanent magnet millimotor (diameter = 1.9 mm, length = 5.5 mm) with an integrated milligear box (Lehr et al. 1996, 1997; Michel and Ehrfeld 1998) is also shown. The gears are made of metal (electroplated Ni-Fe) but can also be made from injected polymer materials (e.g., Polyoxy-methylene or POM) using the LIGA process (described in Appendix). Even though the torque transmitted at the gear teeth is small, on the order of a fraction of nN m, because of the small dimensions of gear teeth, the bending stresses are large where the teeth mesh. Tooth breakage and wear at the contact of gear teeth is a concern.

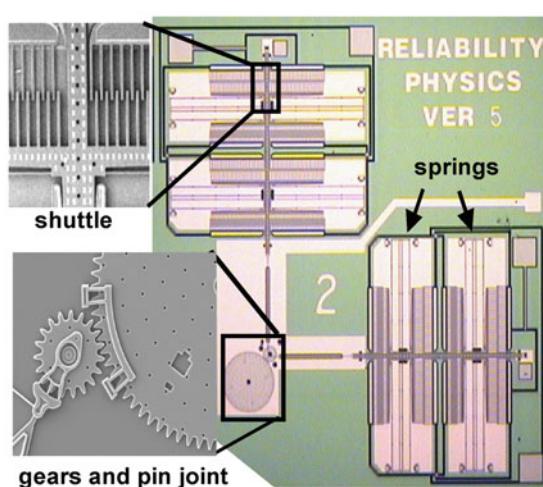
Figure 16.3 shows an optical micrograph of a microengine driven by an electrostatically-

activated comb drive connected to the output gear by linkages. The microengine operates in the kHz frequency range, and can be used as a general drive and power source to drive micromechanisms (Garcia and Sniegowski 1995). Its parts are fabricated from polysilicon. A microgear unit is used to convert reciprocating motion from a linear actuator into circular motion. Another drive linkage oriented at 90° to the original linkage, driven by another linear actuator, allows maintenance of continuous motion. The linkages are connected to the output gear through pin joints that allow relative motion.

One inset shows a polysilicon, multiple microgear speed reduction unit and its components after laboratory wear tests conducted for 600,000 cycles at 1.8% relative humidity (RH) (Tanner et al. 2000). Wear of various components is clearly observed in the figure. Humidity was shown to be a strong factor in the wear of rubbing surfaces. In order to improve the wear characteristics of rubbing surfaces, a 20 nm thick tungsten (W) coating was deposited at 450 °C using the chemical vapor deposition

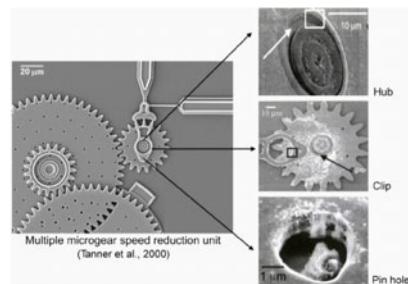
Fig. 16.2 Examples of MEMS components that experience tribological problems



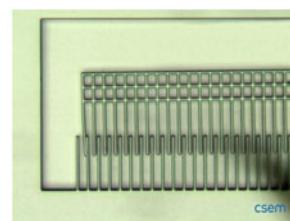


Microengine driven by electrostatically-actuated comb drive

Sandia Summit Technologies (www.mems.sandia.gov)



Microgear unit can be driven at speeds up to 250,000 RPM. Various sliding components are shown after wear test for 600k cycles at 1.8% RH (Tanner et al., 2000)



Stuck comb drive

Fig. 16.3 Optical micrograph of a microengine driven by an electrostatically-actuated comb drive (microengine) fabricated by Sandia Summit Technologies (Garcia and Sniegowski 1995). Top inset shows polysilicon microgear speed reduction unit after laboratory wear test for 600,000 cycles at 1.8% relative humidity (Tanner et al. 2000). The bottom inset shows a stuck comb drive (CSEM)

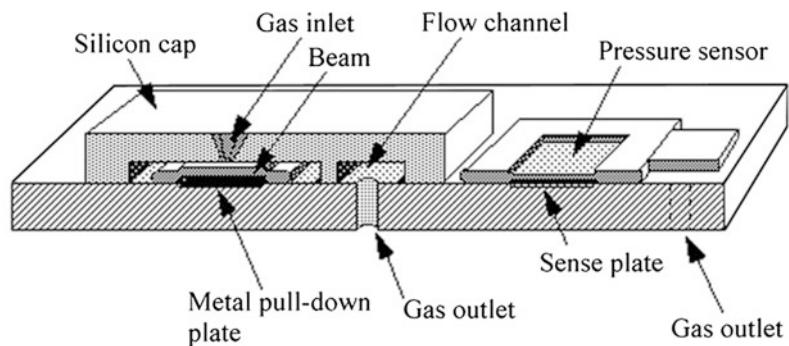
(CVD) technique (Mani et al. 2000). Tungsten-coated microengines tested for reliability showed improved wear characteristics with longer lifetimes than polysilicon microengines. However, these coatings have poor yield. Instead, vapor deposited self-assembled monolayers of fluorinated (dimethylamino) silane have been used (Hankins et al. 2003). The second inset shows a comb drive with a deformed frame, which results in some fingers coming in contact. The contacting fingers can result in stiction.

Figure 16.4 shows a micromachined flow modulator. Several micromachined flow channels are integrated in series with electrostatically actuated microvalves (Robertson and Wise 1998). The flow channels lead to a central gas outlet hole drilled in the glass substrate. Gas enters the device through a bulk micromachined gas inlet hole in the silicon cap. The gas, after passing through an open microvalve, flows parallel to the glass substrate through flow channels

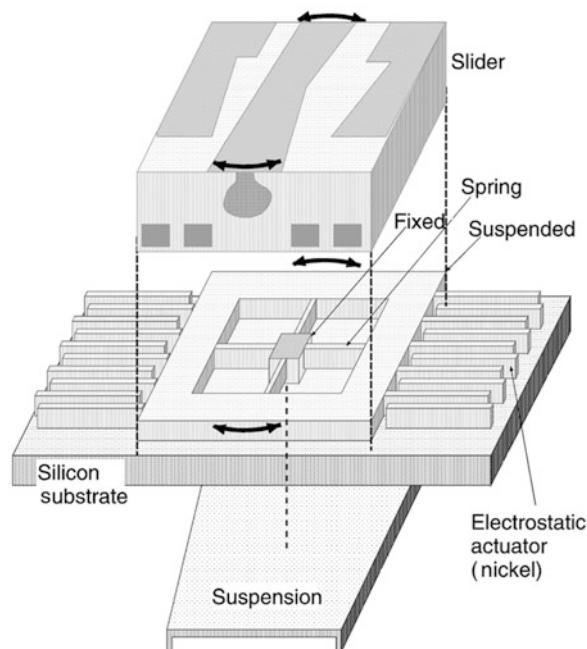
and exits the device through an outlet. The normally open valve structure consists of a free-standing double-end-clamped beam, which is positioned beneath the gas inlet orifice. When electrostatically deflected upwards, the beam seals against the inlet orifice and the valve is closed. In these microvalves used for flow control, the mating valve surfaces should be smooth enough to seal while maintaining a minimum roughness to ensure low adhesion (Bhushan 1996b, 2001, 2007, 2011, 2013a, b).

A second MEMS device shown in Fig. 16.4 is an electrostatically driven rotary microactuator for a magnetic disk drive surface-micromachined by a multilayer electroplating method (Hirano et al. 2000). This high-bandwidth servo-controlled microactuator, located between a slider and a suspension is being developed for ultrahigh-track-density applications, which serves as the fine-position and high-bandwidth control element of a two-stage, coarse/fine servo system

Fig. 16.4 Examples of MEMS devices that experience tribological problems



Low-pressure flow modulator with electrostatically actuated microvalves (Robertson and Wise, 1998)



Electroplated-nickel rotary microactuator for magnetic disk drives (Hirano et al., 2000)

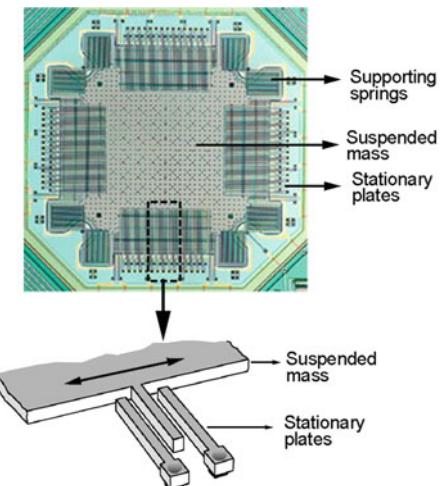
when coupled with a conventional actuator (Horsley et al. 1998; Hirano et al. 2000). A slider is placed on top of the central block of a microactuator, which gives rotational motion to the slider. The bottom of the silicon substrate is attached to the suspension. The radial flexure beams in the central block give the rotational freedom of motion to the suspended mass (slider), and the

electrostatic actuator drives the suspended mass. Actuation is accomplished via interdigitated, cantilevered electrode fingers, which are alternately attached to the central body of the moving part and to the stationary substrate to form pairs. A voltage applied across these electrodes results in an electrostatic force, which rotates the central block. The inter-electrode gap width is about

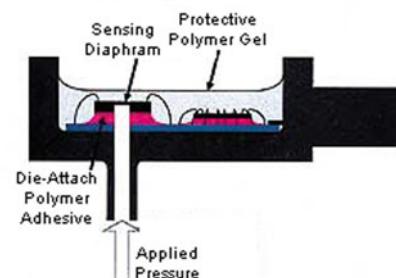
2 μm . Any unintended contacts between the moving and stationary electroplated-nickel electrodes may result in wear and stiction.

Commercially available MEMS devices also exhibit tribological problems. Figure 16.5 shows an integrated capacitive-type silicon accelerometer fabricated using surface micromachining by Analog Devices, a couple of mm in dimension, which is used for the deployment of airbags in automobiles, and for various other devices in the consumer electronics market (Core et al. 1993; Sulouff 1998). The central suspended beam mass (about 0.7 μg) is supported on the four corners by spring structures. The central beam has interdigitated cantilevered electrode fingers (about 125 μm long and 3 μm thick) on all four sides that alternate with those of the stationary electrode fingers as shown, with about a 1.3 μm gap. Lateral motion of the central beam causes a change in the capacitance between these electrodes, which is used to measure the acceleration. Stiction between the adjacent electrodes as well as stiction of the beam structure with the underlying substrate, under isolated conditions, is detrimental to the operation of the sensor (Core et al. 1993; Sulouff 1998). Wear during unintended contacts of these polysilicon fingers is also a problem. A molecularly thick diphenyl siloxane lubricant film, resistant to high temperatures and oxidation, is applied by a vapor deposition process on the electrodes to reduce stiction and wear (Martin and Zhao 1997). For deposition, a small amount of liquid is dispensed into each package before it is sealed. As the package is heated in the furnace, the liquid evaporates which coats the sensor surface. As sensors are required to sense low g accelerations, they need to be more compliant and stiction becomes even a bigger concern.

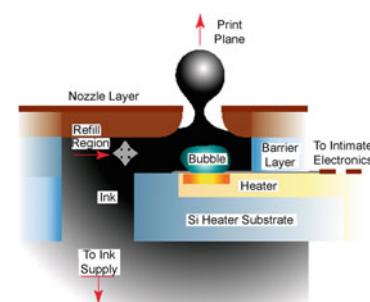
Figure 16.5 also shows a cross-sectional view of a typical piezoresistive type pressure sensor, which is used for various applications including manifold absolute pressure (MAP) and tire pressure measurements in automotive applications, and disposable blood pressure measurements. The sensing material is a diaphragm



Capacitive type silicon accelerometer for automotive sensory applications (Sulouff, 1998)



Piezoresistive type pressure sensor (Parsons, 2001)



Thermal inkjet printhead (Baydo and Groscup, 2001)

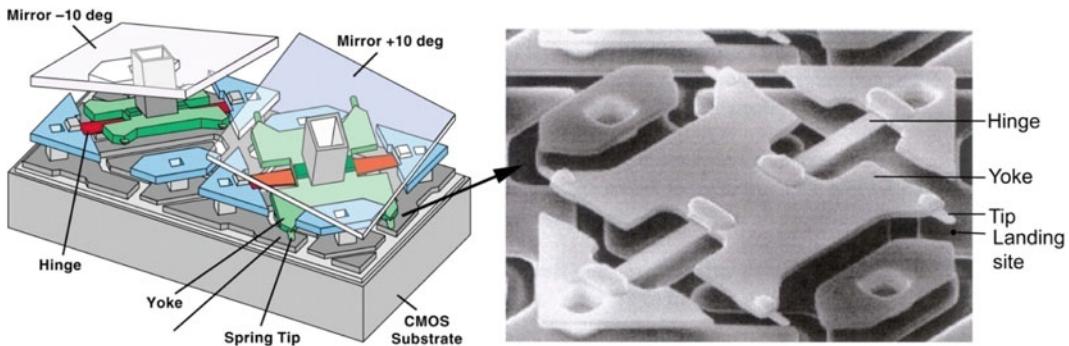
Fig. 16.5 Examples of commercial MEMS devices that experience tribological problems

formed on a silicon substrate, which bends with applied pressure (Smith 1997; Parsons 2001). The deformation causes a change in the band structure of the piezoresistors that are placed on the diaphragm, leading to a change in the resistivity of the material. The MAP sensors are subjected to drastic conditions—extreme temperatures, vibrations, sensing fluid, and thermal shock. Fluid under extreme conditions could cause corrosive wear, while fluid cavitation could cause erosive wear. The protective gel encapsulant generally used can react with sensing fluid and result in swelling or dissolution of the gel. Silicon cannot deform plastically, therefore any pressure spikes leading to deformation past its elastic limit will result in fracture and crack propagation. Pressure spikes could also cause the diaphragm to delaminate from the support substrate. Finally, cyclic loading of the diaphragm during use can lead to fatigue and wear of the silicon diaphragm or delamination.

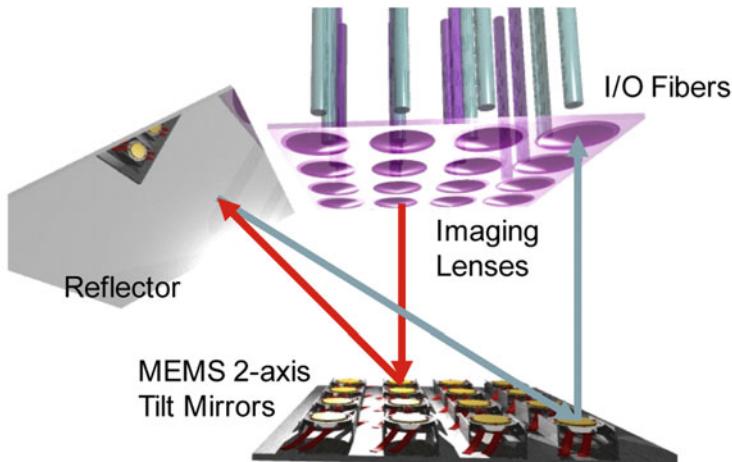
The bottom schematic in Fig. 16.5 shows a cross-section of a thermal printhead chip (on the order of 10–50 cm³ in volume) used in inkjet printers (Baydo and Groscup 2001). They comprise of a supply of ink and an array of elements with microscopic heating resistors on a substrate mated to a matching array of ink-injection orifices or nozzles (about 70 µm in diameter) (Aden et al. 1994; Le 1998; Lee 2003). In each element, a small chamber is heated by the resistor where a brief electrical impulse vaporizes part of the ink and creates a tiny bubble. The heaters operate at several kHz and are therefore capable of high-speed printing. As the bubble expands, some of the ink is pushed out of the nozzle onto the paper. When the bubble pops, a vacuum is created and this causes more ink from the cartridge to move into the printhead. Clogged ink ports are the major failure mode. There are various tribological concerns (Aden et al. 1994). The surface of the printhead where the ink is shot out towards the paper can get scratched and damaged as a result of countless trips back and forth across the pages, which are somewhat rough. As a result of repeated heating and cooling, the heated resistors expand and contract. Over time, these elements will experience fatigue

and may eventually fail. Bubble formation in the ink reservoir can lead to cavitation erosion of the chamber, which occurs when bubbles formed in the fluid become unstable and implode against the surface of the solid and impose impact energy on that surface. Fluid flow through nozzles may cause erosion and ink particles may also cause abrasive wear. Corrosion of the ink reservoir surfaces can also occur as a result of exposure of ink at high temperatures as well as due to ink pH. The substrate of the chip consists of silicon with a thermal barrier layer followed by a thin film of resistive material and then conducting material. The conductor and resistor layers are generally protected by an overcoat layer of a plasma-enhanced chemical vapor deposition (PECVD) α -SiC:H layer, 200–500 nm thick (Chang et al. 1991).

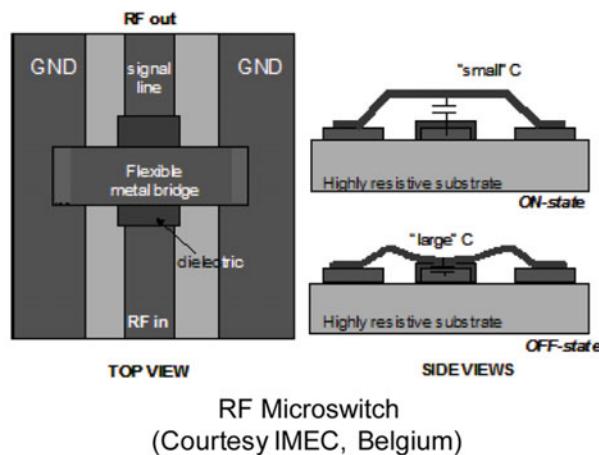
Figure 16.6 shows two DMD pixels used in DLP technology for digital projection displays in computer projectors, HDTV sets, and movie projectors (Hornbeck and Nelson 1988; Hornbeck 1999, 2001). The entire array (chip set) consists of a large number of oscillating aluminum alloy micromirrors as digital light switches which are fabricated on top of a complementary metal-oxide-semiconductor (CMOS) static random access memory integrated circuit. The surface micromachined array consists of half a million to more than two million of these independently controlled reflective micromirrors. Each micromirror is about 12 µm square and with 13 µm pitch, and flip backward and forward at a frequency on the order of 5–7 kHz as a result of electrostatic attraction between the micromirror structure and the underlying electrodes. For binary operation, the micromirror/yoke structure mounted on torsional hinges is oscillated $\pm 10^\circ$ (with respect to the plane on the chip set), and is limited by a mechanical stop. Contact between cantilevered spring tips at the end of the yoke (four present on each yoke) with the underlying stationary landing sites is required for true digital (binary) operation. Stiction and wear during contact between aluminum alloy spring tips and landing sites, hinge memory (metal creep at high operating temperatures), hinge fatigue, shock and vibration failure, sensitivity to particles in the



Digital micromirror device for displays (Hornbeck, 1999)



Tilt mirror arrays for switching optical signal in input and output fiber arrays in optical crossconnect for telecom (Aksyuk et al., 2003)



RF Microswitch
(Courtesy IMEC, Belgium)

Fig. 16.6 Examples of two commercial MOEMS and one RF-MEMS device that experience tribological problems

chip package, and operating environment are some of the important issues affecting the operation reliability of a micromirror device (Henck 1997; Douglass 1998, 2003; Liu and Bhushan 2004a, b). A vapor phase deposited self-assembled monolayer of the fatty acid perfluorodecanoic acid (PFDA) on surfaces of tip and landing sites is used to reduce stiction and wear (Hornbeck 1997; Robbins and Jacobs 2001). However, these films are susceptible to moisture, and to keep moisture out and create a background pressure of PFDA, a hermetic chip package is used. The spring tip uses the stored spring energy to pop up the tip during pull-off. A lifetime estimate of over 100,000 operating hours with no degradation in image quality is typical. At a mirror modulation frequency of 7 kHz, each micromirror element needs to switch about 2.5 trillion cycles.

Figure 16.6 also shows a schematic of a 256×256 -port large optical cross-connects, introduced in 2000 by Glimmerglass, Hayward, California, for optical telecommunication networks in order to be able to manipulate rapidly a larger number of optical signals (Aksyuk et al. 2003). This optical microswitch uses 256 or more movable mirrors on a chip for switching a light beam from an input fiber to a few output fibers. The mirrors are made of gold-coated polysilicon and about 500 μm in diameter. Reliability concerns are the same as those described for DMDs. To minimize stiction, the chipset is hermetically sealed in dry nitrogen (90% N₂, 10% He).

Figure 16.6 also shows a schematic of an electrostatically-actuated, capacitive-type RF microswitch for switching of RF signals at microwave and low frequencies (DeWolf and van Spengen 2002). It is a membrane type microswitch, and consists of a flexible metal (Al) bridge that spans the RF transmission line in the center of a coplanar waveguide. When the bridge is up, the capacitance between the bridge and RF transmission line is small, and the RF signal passes without much loss. When a DC voltage is applied between the RF transmission line and the bridge, the latter is pulled down until it touches a dielectric isolation layer. Thus, the

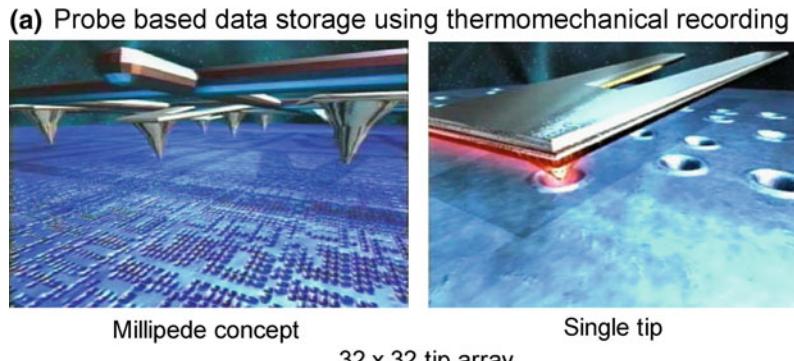
large capacitance created shorts the RF signal to the ground. The failure modes include creep in the metal bridge, fatigue of the bridge, charging and degradation of the dielectric insulator, and stiction of the bridge to the insulator (DeWolf and van Spengen 2002; Suzuki 2002; Zaghloul et al. 2011a, b). Stiction occurs due to capillary condensation of water vapor from the environment, van der Waals forces, and/or charging effects (Bhushan 2003, 2013a, b). If the restoring force in the bridge of the switch is not large enough to pull the bridge up again after the actuation voltage has been removed, the device fails due to stiction. Humidity induced stiction can be avoided by hermetically sealing the microswitch. Some roughness of the surfaces reduces the probability of stiction. Selected actuation waveforms can be used to minimize charging effects.

16.2.2 NEMS

Probe-based data recording technologies have been explored for ultra-high areal density recording where the probe tip (with a radius of about 5 nm) is expected to be scanned at velocities up to 100 mm/s. Major techniques include —thermomechanical (Vettiger et al. 1999), phase change (Bhushan and Kwak 2007), and ferroelectric recording (Bhushan and Kwak 2008; Kwak and Bhushan 2008). We discuss the tribological issues with two of the widely pursued techniques (Bhushan et al. 2008a).

Figure 16.7a shows the thermomechanical recording system which uses arrays of 1024 silicon microcantilevers (“Millipede”) and playback on a ~ 40 nm thick polymer medium with a harder Si substrate (Vettiger et al. 1999). The cantilevers consist of integrated tip heaters with tips of nanoscale dimensions. Sharp tips themselves are also example of NEMS. Thermomechanical recording is a combination of applying a local force to the polymer layer and softening it by local heating. The tip, heated to about 400 °C, is brought in contact with the polymer for recording. Readings are done using the heater cantilever, originally used for recording, as a

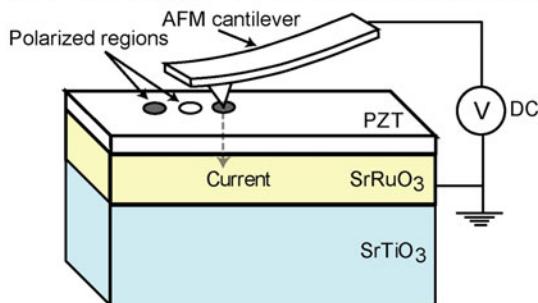
Fig. 16.7 Two example of NEMS devices—
a thermomechanical recording,
 and **b** ferroelectric recording—that experience tribological problems



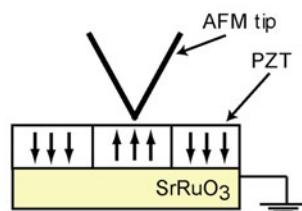
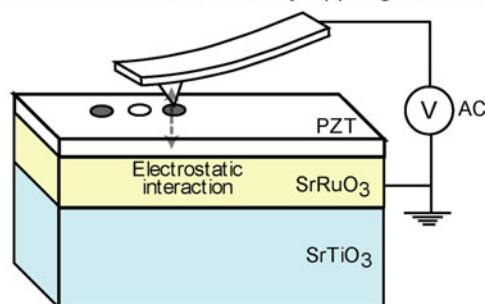
(b)

Probe based data storage
 using ferroelectric recording

Writing by applying DC voltage in contact mode



Reading of piezoelectric response in contact or electrostatic force in noncontact mode by applying AC voltage



thermal readback sensor by exploiting its temperature-dependent resistance. The principle of thermal sensing is based on the fact that the thermal conductivity between the heater and the storage substrate changes according to the spacing between them. When the spacing between the heater and sample is reduced as the tip moves into a bit, the heater's temperature and hence its resistance will decrease. Thus, changes in temperature of the continuously-heated resistor are monitored while the cantilever is scanned over data bits, providing a means of detecting the bits. Erasing for subsequent rewriting is carried out by thermal reflow of the storage field by heating the medium to 150° for a few seconds. The smoothness of the reflowed medium allows multiple rewriting of the same storage field. Bit sizes ranging between 10 and 50 nm have been achieved by using a 32 × 32 (1024) array write/read chip (3 mm × 3 mm). It has been reported that tip wear occurs due to contact between the tip and the Si substrate during writing. Tip wear is a major concern for device reliability.

Figure 16.7b shows the schematic of domain writing and reading in a ferroelectric film. The electrically conductive AFM tips are placed in contact with a piezoelectric (lead zirconate titanate, PZT), film-coated medium (Bhushan and Kwak 2008; Kwak and Bhushan 2008). Ferroelectric domains on PZT film are polarized by applying short voltage pulses (~ 10 V, ~ 100 μ s) that exceed the coercive field of the PZT layer, resulting in local, nonvolatile changes in the electronic properties of the underlying film. The temperature rise during recording is expected to be on the order of 80 °C. Reading out of the polarization states in the ferroelectric film can be carried out using two different methods. In one method, the static surface charge, proportional to the normal component of polarization, can be detected by electrostatic force microscopy in the noncontact mode. In the second method, an AFM is operated in the contact mode and piezoresponse force is measured by applying an alternating-current (AC) voltage. Wear of the conducting tip and the PZT layer at high scanning velocities is a major concern for

device reliability. Various lubricant films are being developed to minimize wear (Bhushan and Kwak 2007, 2008; Bhushan et al. 2008a, b; Kwak and Bhushan 2008; Palacio and Bhushan 2008a, 2009).

In magnetic data storage, magnetic recording is accomplished by relative motion between the magnetic head slider and a magnetic rigid disk (Bhushan 1996a). Magnetic rigid disks and heads used today for magnetic data storage consist of nanostructured films a few nm thick. Figure 16.8a shows the sectional view of a conventional magnetic rigid disk. The superparamagnetic effect poses a serious challenge for the ever increasing areal density of disk drives. One of the promising methods to circumvent the density limitations imposed by this effect is the use of a nanopatterned disk (Fig. 16.8b). In a conventional disk, the thin magnetic layer forms a random mosaic of nanometer scale grains, and each recorded bit consists of many tens of these random grains. In a patterned disk, the magnetic layer is created as an ordered array of highly uniform islands, each island capable of storing an individual bit. These islands may be one or a few grains, rather than a collection of random decoupled grains. This increases the density by a couple of orders of magnitude. Figure 16.8c shows a schematic of an inductive write GMR read head structure. These are constructed from a variety of materials: magnetic alloys, metal conductors, ceramic, and polymer insulators in a complex three dimensional structure. The multilayered thin film structure used to construct the sensor and individual films are only a few nm thick. The head slider surface, which flies on the disk surface, is coated with about 3-nm thick of diamondlike carbon coatings to protect thin-film structure from electrostatic discharge. Any isolated contacts between the disk and sensor and lubricant pickup pose tribological concerns (Bhushan 1996a).

Nano-objects are used in various applications where they come into sliding contact with each other and the surfaces where they are used, which can lead to nano-object deformation (Bhushan et al. 2014). Some examples of these applications include drug delivery for cancer treatment (to be presented in the BioNEMS section), oil

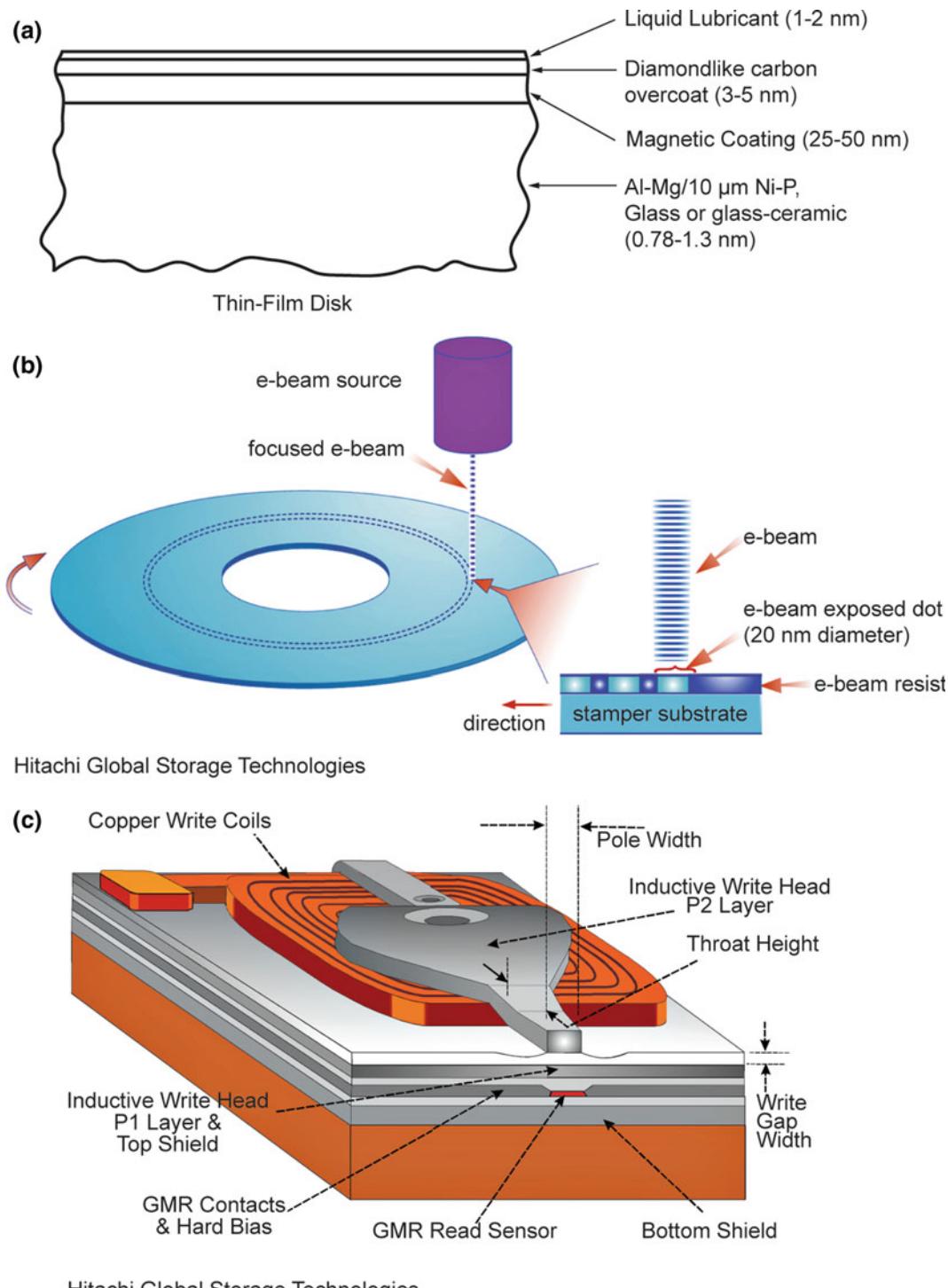


Fig. 16.8 Schematic of **a** sectional view of a conventional magnetic rigid disk, **b** nanopatterned magnetic rigid disk, and **c** an inductive write, GMR read magnetic head structure for magnetic data storage

detection, contaminant removal, catalysis, and tribology on the macro- to nanoscale. For oil detection applications, as in the example shown in Fig. 16.9a, oxidized carbon black nanoparticles with a polyvinyl alcohol shell are coated with an oil detecting agent. This agent is released on contact with hydrocarbons in oil, and its absence on recovery and analysis of the nano-objects is an indication of the presence of oil (Berlin et al. 2011). Agglomeration of

nano-objects can prevent flow through porous media, which results in losses and prevents eventual recovery of the nano-objects (Johnson et al. 1996). Friction on surfaces and during flow in liquids can also hinder recovery. Low friction of nano-objects over surfaces and within liquids is of interest.

In contaminant removal, magnetic shell cross-linked knedel-like nanoparticles (MSCKs) are of interest. These are composites of polymer

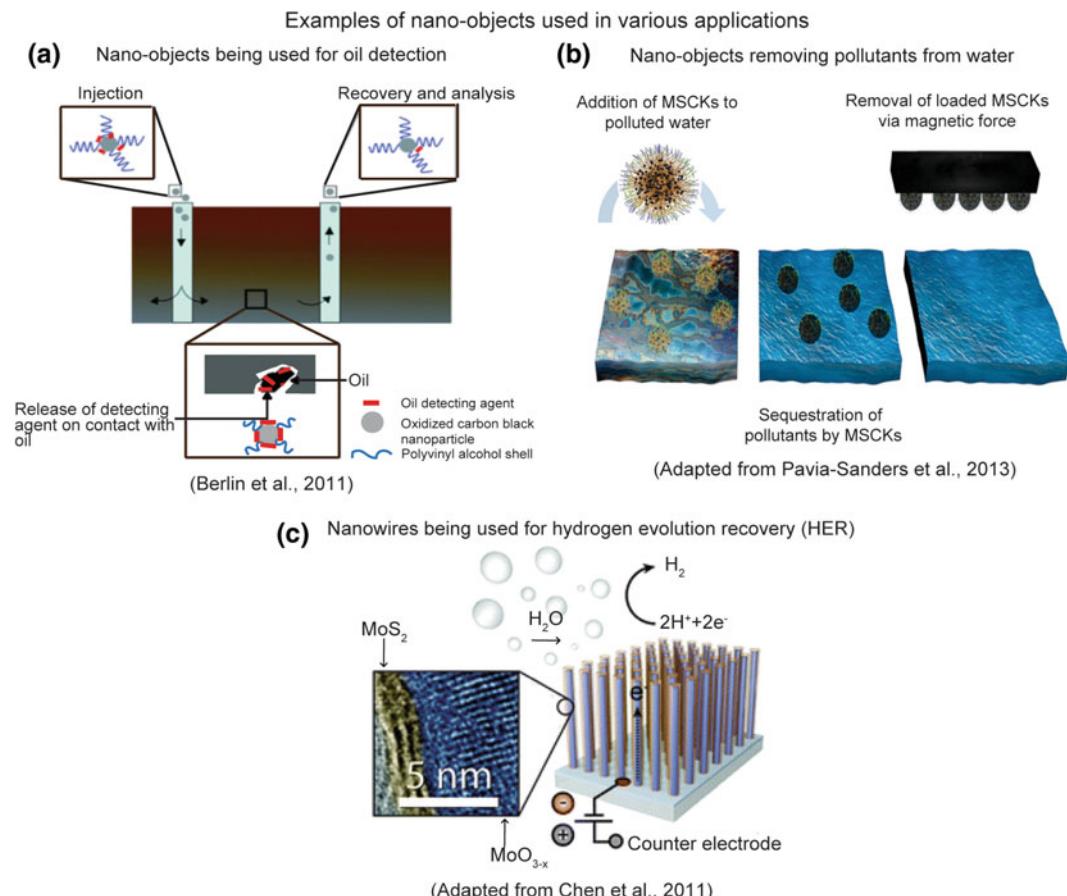


Fig. 16.9 Examples of nano-objects in applications in **a** oil detection, where carbon black nano-objects are coated with an oil detecting agent. After injection into the ground, the agent is released on contact with hydrocarbons, and this is used as an indication of the presence of oil on recovery and analysis of the nanoparticle (Berlin et al. 2011); **b** contaminant removal with magnetic shell cross-linked knedel-like nanoparticles (MSCKs). The MSCKs are added to the polluted water (*left*), where they selectively absorb the oil (*center*) and are finally removed by introduction of a neodymium magnet (Pavia-Sanders et al. 2013) and **c** catalysis for hydrogen evolution recovery (HER), using nanowires. The nanowires consist of an inner MoO₃ Core and outer MoS₂ shell, which is catalytically active. At these sites, the transfer of electrons necessary for recovery of hydrogen from water is enhanced by the MoS₂, which improves the efficiency of the HER (Chen et al. 2011). In these applications, low friction of nano-objects in liquids is desirable

layers on iron oxide nanoparticles. In Fig. 16.9b, the MSCKs are added to polluted water (left), where they selectively absorb oil (center) and are finally removed by introducing a neodymium magnet (Pavia-Sanders et al. 2013). Reducing agglomeration and fluid friction is important for efficient manipulation of nano-objects through liquids to contaminated sites.

In applications involving catalysis for hydrogen evolution recovery (HER), nanowires have been used to aid in the catalytic activity, as shown in Fig. 16.9c. The nanowires consist of an inner molybdenum trioxide (MoO_3) core and outer molybdenum disulfide (MoS_2) shell. The core is used for transporting electrons to edge sites of the outer MoS_2 shell, which is catalytically active. This is preferred over bulk MoS_2 which has a limited number of edge sites for the same volume of material compared to the nanowires. At these sites, the transfer of electrons necessary for the recovery of hydrogen from water is enhanced by the MoS_2 , which improves HER efficiency (Chen et al. 2011). Low friction of water moving over the nanowires is important for efficient flow over edge sites and conversion to hydrogen.

Nano-objects have demonstrated the ability to reduce friction and wear in applications on the macroscale. For applications on the micro/nanoscale, increasing the lifetime and efficiency of individual components of systems is crucial to the commercialization of micro/nanosystems. Adhesive and friction forces can hinder device operation and reliability (Bhushan 2011, 2013a, b). In these environments, nano-objects can be deformed locally or the entire nano-object can be compressed. Knowledge of the mechanical behavior during loading and unloading in these environments is essential for sustained use. Studies in friction, wear, and deformation behavior are critical to successful integration of nano-objects for the various applications.

16.2.3 BioMEMS

Biosensors are devices typically used to detect target biomolecules such as proteins, enzymes, or

nucleic acids, as well as chemical contaminants. Biosensors can be classified into two broad categories: microarray type, which typically consist of (1) microfluidic and nanofluidic sensors that usually involve manipulation of small fluidic volumes (microliters to nanoliters), leading to optical methods for detection, and (2) cantilever or field-effect devices with adsorption of target analytes to sensing elements as the main transduction mechanism.

Advances in micro- and nanofabrication have led to the development of integrated micro- and nanofluidic biosensors that bring together several operations, e.g., sample extraction, purification, sensing, etc., on a simple platform. These sensors are designed either to detect a single or a class of (bio)chemicals, or for system level analytical capabilities for a broad range of (bio)chemical species known as a micro total analysis system (μTas). Detection is generally performed using optical methods in micro- and nanofluidic biosensors and mechanical or electrical methods in microarray-type biosensors (Prakash et al. 2012). Optical methods include fluorescence, optical cavity resonator, and surface plasmon resonance. Methods that are used for driving the fluid flow in micro- and nano channels. Mechanical methods include mechanical changes induced by adsorption of target molecules or analytes to cantilevers or nanowires. Electrical methods include changes to inherent biomolecular charge in the case of field-effect transistors or electrochemical changes. In micro- and nanofluidic biosensors, two methods are commonly used to drive the fluid flow—pressure-driven and electric field driven or electrokinetic.

An example of a wristwatch type biosensor based on microfluidics referred to as lab-on-a-chip system is shown in Fig. 16.10a (Tang and Lee 2001; van der Berg 2003). The chip relies on microfluidics and involves manipulation of tiny amounts of fluids in microchannels using micro-valves. The test fluid is injected into the chip generally using an external pump or syringe for analysis. Some chips have been designed with an integrated electrostatically-actuated diaphragm type micropump. The sample, which can have volume measured in nanoliters, flows through

microfluidic channels via an electric potential and capillary action using microvalves (having various designs including membrane type) for various analyses. The fluid is preprocessed and then analyzed using a biosensor. Another example of a biosensor is the cassette type biosensor used for human genomic DNA analysis, and integrated biological sample preparation is shown in Fig. 16.10b (Taylor et al. 2003). The implementation of micropumps and microvalves allows for fluid manipulation and multiple sample processing steps in a single cassette. Blood or other aqueous solutions can be pumped into the system where various processes are performed.

Microvalves, which are found in most microfluidic components of biosensors, can be classified in two categories: active microvalves (with an actuator) for flow regulation in microchannels, and passive microvalves integrated with micropumps. Active microvalves consist of a valve seat and a diaphragm actuated by an external actuator (Ohnstein et al. 1990; Gravesen et al. 1993; Shoji and Esashi 1994). Different types of actuators are based on piezoelectric, electrostatic, thermopneumatic, electromagnetic, and bimetallic materials, shaped memory alloys, and solenoid plungers. An example of an electrostatic cantilever type active microvalve is shown in Fig. 16.11 (Ohnstein et al. 1990). Passive microvalves used in micropumps include mechanical check valves and a diffuser/nozzle (Gravesen et al. 1993; Shoji and Esashi 1994; Stehr et al. 1996; Woias 2001; Nguyen et al. 2002). Check valves consist of a flap or membrane that is capable of opening and closing with changes in pressure; see Fig. 16.11 for schematics. A diffuser/nozzle uses an entirely different principle and only works with the presence of a reciprocating diaphragm. When one convergent channel works simultaneously with another convergent channel oriented in a specific direction, a change in pressure is possible.

There are four main types of mechanical micropumps: a diaphragm micropump involving mechanical check valves, electrostatic micropumps, valveless rectification pumps with or without diffuser/nozzle type valves, and rotary

micropumps (Gravesen et al. 1993; Shoji and Esashi 1994; Stehr et al. 1996; Woias 2001; Nguyen et al. 2002). A diaphragm micropump consists of a reciprocating diaphragm which can be piezoelectrically driven, working in synchronization with two check valves (Fig. 16.11) (Woias 2001). Electrostatic micropumps have a diaphragm as well, but are driven using two electrodes (Fig. 16.11) (Bustgens et al. 1994). Valveless micropumps also consist of a diaphragm that is piezoelectrically driven, but does not incorporate passive mechanical valves. Instead, these pumps use an elastic buffer mechanism or a variable gap mechanism. Finally, a rotary micropump has a rotating rotor that simply adds momentum to the fluid by the fast moving action of the blades (Fig. 16.11) (Ahn and Allen 1995; Doepper et al. 1997). Rotary micropumps can be driven using an integrated electromagnetic motor or by the presence of an external electric field. All of these micropumps can be made of silicon or a polymer material.

During the operation of the microvalves and micropumps discussed above in which contacts occur due to relative motion, adhesion and friction properties become important. During operation, active mechanical microvalves have an externally actuated diaphragm that comes into contact with a valve seat to restrict the fluid flow. Adhesion between the diaphragm and valve seat will affect the operation of the microvalve. In the diaphragm micropumps, two passive mechanical check valves are incorporated into the design. Passive mechanical check valves also exhibit adhesion when the flap or membrane comes into contact with the valve seat when fluid flow is removed. Adhesion also occurs during the operation of valveless micropumps when the diaphragm, which is piezoelectrically driven, comes into contact with the rigid outlet. Finally, adhesion and friction can also be seen during the operation of rotary micropumps when the gears rotate and come into contact and rub against one another.

If the adhesion between the microchannel surface and the biofluid is high, the biomolecules will stick to the microchannel surface and restrict

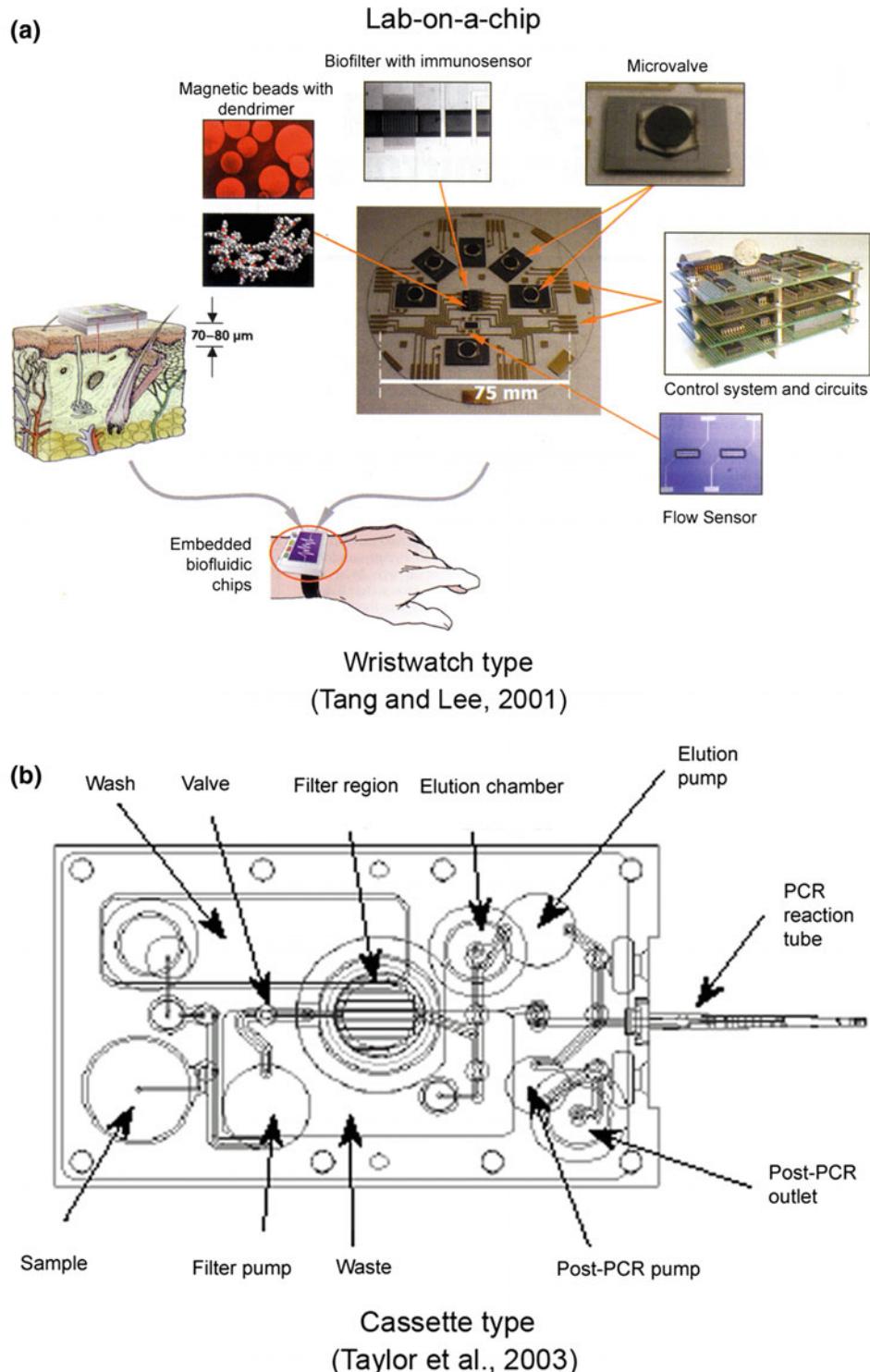


Fig. 16.10 **a** MEMS-based biofluidic chip, commonly known as a lab-on-a-chip, that can be worn like a wristwatch. **b** Cassette type biosensor used for human genomic DNA analysis (PCR-polymerase chain reaction)

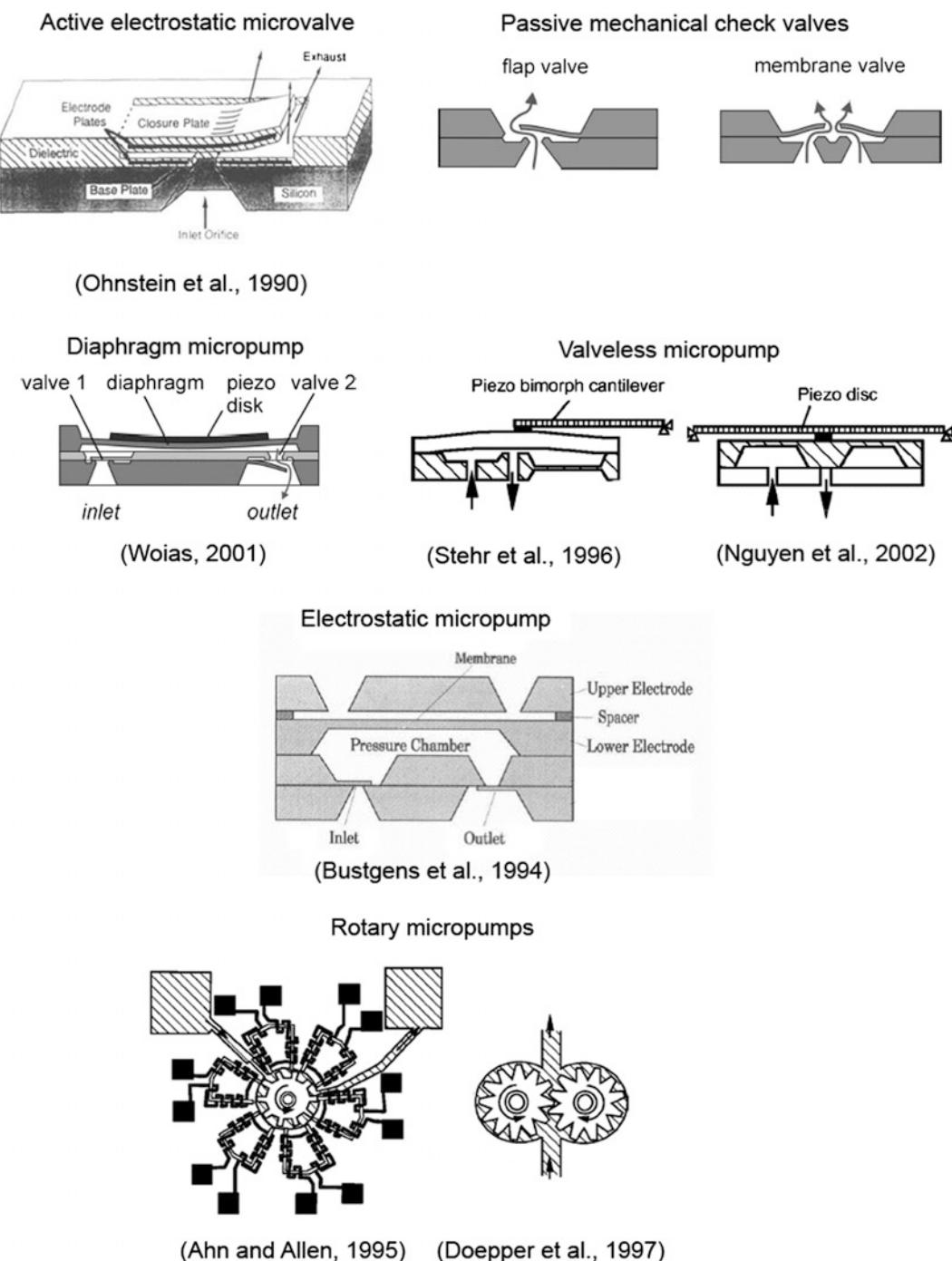


Fig. 16.11 Multiple examples of valves and pumps found in microfluidic devices. Mechanical check valves, diaphragm micropump, valveless micropump, and rotary micropump

flow. In order to facilitate flow, microchannel surfaces with low bioadhesion are required. Fluid flow in polymer channels can produce triboelectric surface potential, which may affect the flow. Polymers are known to generate surface potential, and the magnitude of the potential varies from one polymer to another (Henniker 1962; Sakaguchi and Kashiwabara 1992; Freeman and March 1999). Conductive surface layers on the polymer channels can be deposited to reduce triboelectric effects.

As mentioned earlier, the microfluidic biosensor shown in Fig. 16.10a required the use of micropumps and microvalves. As an example, a microdevice with 1000 channels requires 1000 micropumps and 2000 microvalves, which makes it bulky and poses reliability concerns. To overcome this problem, two methods can be used for driving the flow of fluids in microchannels: pressure driven and electrokinetic. Electrokinetic flow is based on the movement of molecules in an electric field due to their charges. There are two components to electrokinetic flow: electrophoresis, which results from the accelerating force due to the charge of a molecule in an electric field, and electroosmosis, which uses electrically controlled surface tension to drive the uniform liquid flow. Biosensors based on electrokinetic flow have also been developed. In so-called “digital based microfluidics,” based on the electroosmosis process, electrically controlled surface tension is used to drive liquid droplets, thus eliminating the need for valves and pumps (Cho et al. 2003; Wheeler et al. 2005). These microdevices consist of a rectangular grid of gold nanoelectrodes instead of micro/nanochannels. An externally applied electric field enables manipulation of few nanoliter samples through the capillary circuitry.

An example of a microarray type biosensor is based on a field-effect transistor (FET) and is shown in Fig. 16.12a (Bhushan et al. 2005a; Lee et al. 2005a; Nicholson et al. 2010; Gupta et al. 2011; Casal et al. 2012). FETs are sensitive to the electrical field produced by the charge at the surface of the gate insulator. In this sensor, the gate metal electrode of a metal-oxide semiconductor field effect transistor (MOSFET) is

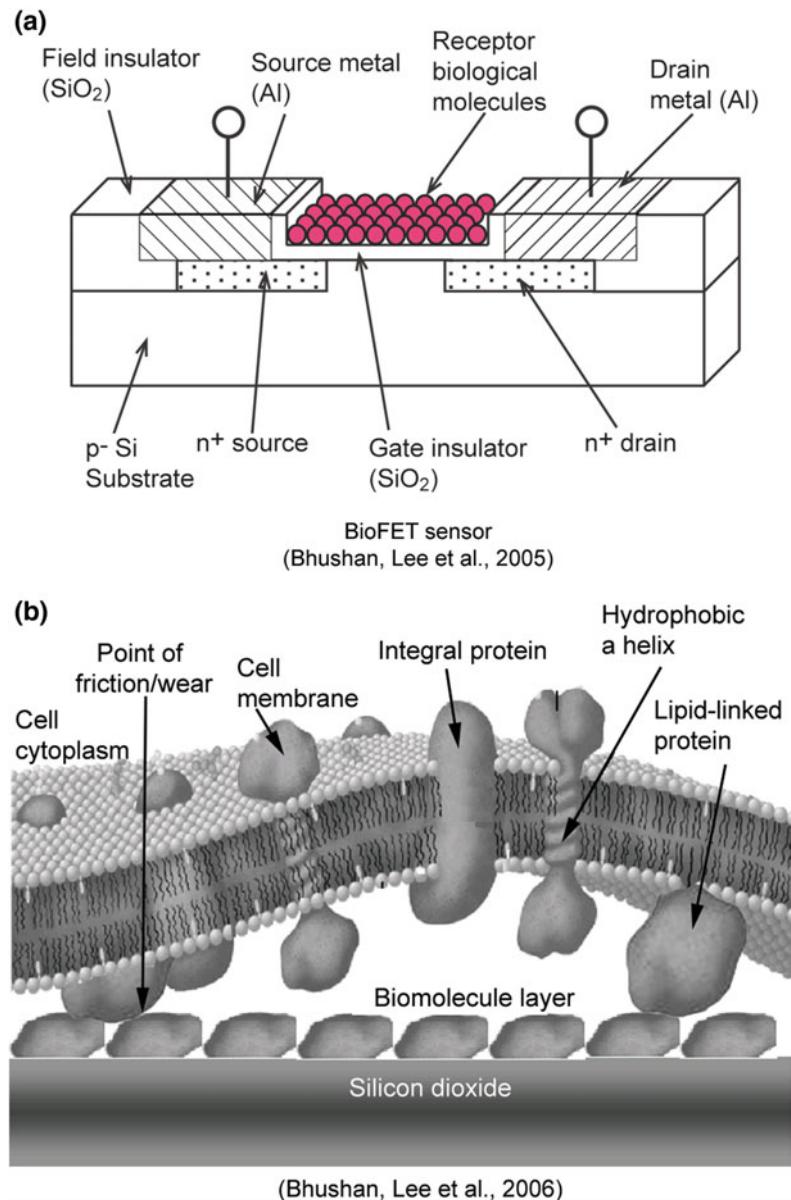
removed and replaced with a protein (receptor layer) whose cognate is the analyte (e.g., virus or bacteria) that is meant to be sensed. Various proteins may have 1–25 (positive or negative) charges per molecule. The binding of the receptor layer with the analyte produces a change in the effective charge, creating a change in the electrical field. This electrical field change may produce a measurable change in the current flow through the device. Adhesion between the protein layer and silica substrate affects the reliability of the biosensor. In the case of implanted biosensors, the biosensors come in contact with the exterior environment, such as tissues and fluids, and any relative motion of the sensor surface with respect to that exterior environment may result in surface damage. A schematic of the generation of friction and wear points when an implanted biosensor surface comes in contact with a living tissue is shown in Fig. 16.12b (Bhushan et al. 2006a). The friction, wear, and adhesion of the biosensor surface may be critical in these applications (Black 1999; Bhushan et al. 2005a, 2006a, 2009).

Polymer BioMEMS have been designed to measure cellular surfaces. For two examples, see Fig. 16.13 (Wei et al. 2005). The device on the left has cantilevers anchored at the periphery of the circular structure, while the device on the right has cantilevers anchored at the two corners on the top and the bottom. The cell adheres to the center of the structure, and the contractile forces generated in the cells cytoskeleton cause the cantilever to deflect. The deflection of the compliant polymer cantilevers is measured optically and related to the magnitude of the forces generated by the cell. Adhesion between cells and the polymer beam is desirable. In order to design the sensors, micro- and nanoscale mechanical properties of polymer structures are needed.

16.2.4 BioNEMS

Micro/nanofluidic devices provide a powerful platform for electrophoretic separations for a variety of chemical and biochemical analysis. Electrophoresis is a versatile analytical method

Fig. 16.12 **a** Schematic of a bioFET sensor (Bhushan et al. 2005a), and **b** schematic showing the generation of friction and wear points due to interaction of implanted biomolecule layer on a biosensor with living tissue (Bhushan et al. 2006a)



which is used for separation of small ions, neutral molecules, and large biomolecules. Figure 16.14 shows an interdigitated micro/nanofluidic silicon array with nanochannels for separation process. Figure 16.15 shows a schematic of an implantable, immunoisolation submicroscopic biocapsule for drug delivery treating significant medical conditions such as Type I diabetes (Desai et al. 1999; Hansford et al. 2001). The purpose of the immunoisolation biocapsule is to create an

implantable device capable of supporting foreign living cells that can be transplanted into humans. It is a silicon capsule consisting of two nanofabricated membranes bonded together with the drug (e.g., encapsulated insulin producing islet cells) contained within the cavities for long-term delivery. The pores or nanochannels in a semi-permeable membrane as small as 6 nm are used as flux regulators for the long-term release of drugs. The nanomembrane also

Fig. 16.13 Schematic of two designs for polymer bioMEMS structures to measure cellular forces (Wei et al. 2005)

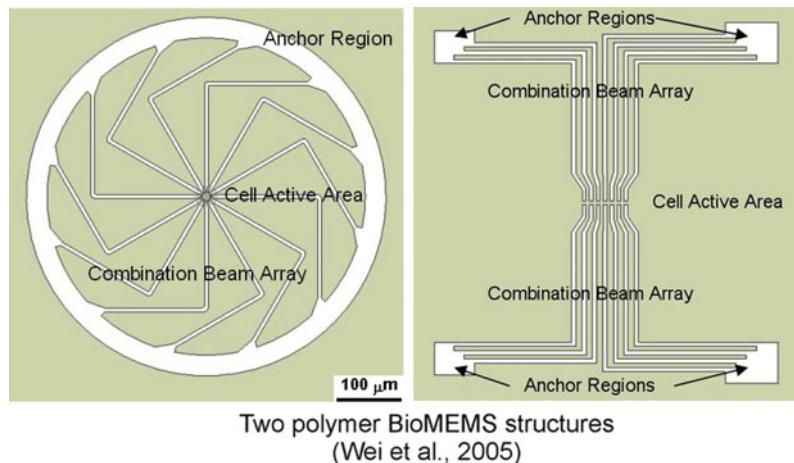


Fig. 16.14 Interdigitated micro/nanofluidic silicon array for separation process (Hansford et al. 2001)

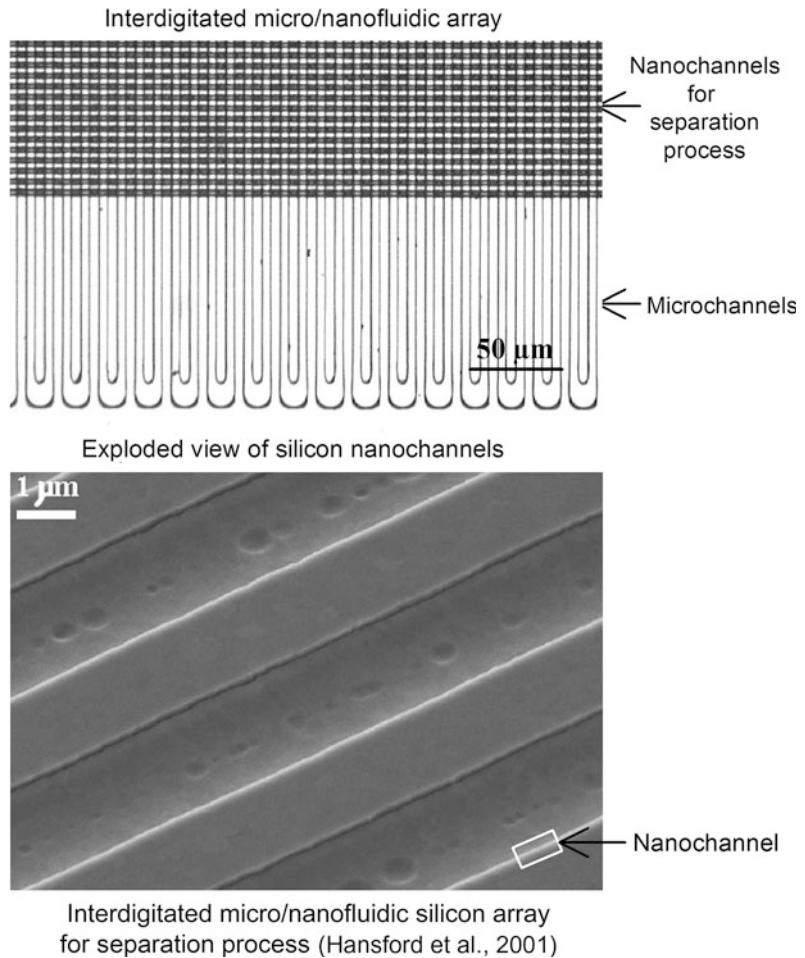
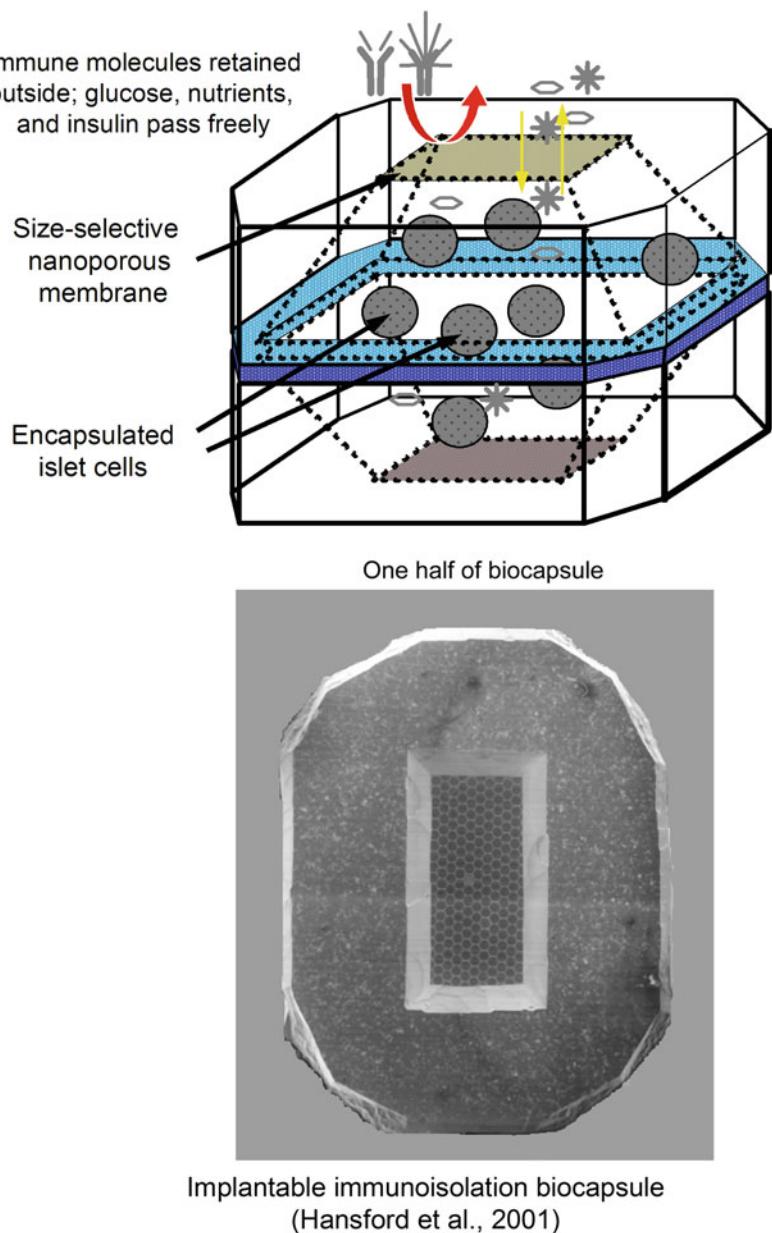


Fig. 16.15 Schematics of implantable, immunoisolation submicroscopic biocapsules (drug delivery device) (Hansford et al. 2001)



protects therapeutic substances from attack by the body's immune system. The pores are large enough to provide the flow of nutrients (e.g., glucose molecules) and drugs (e.g., insulin). However, they are small enough to block natural antibodies, which can penetrate any orifice larger than 18 nm. Pores in silicon can be etched using sacrificial layer lithography, described in Appendix (Hansford et al. 2001).

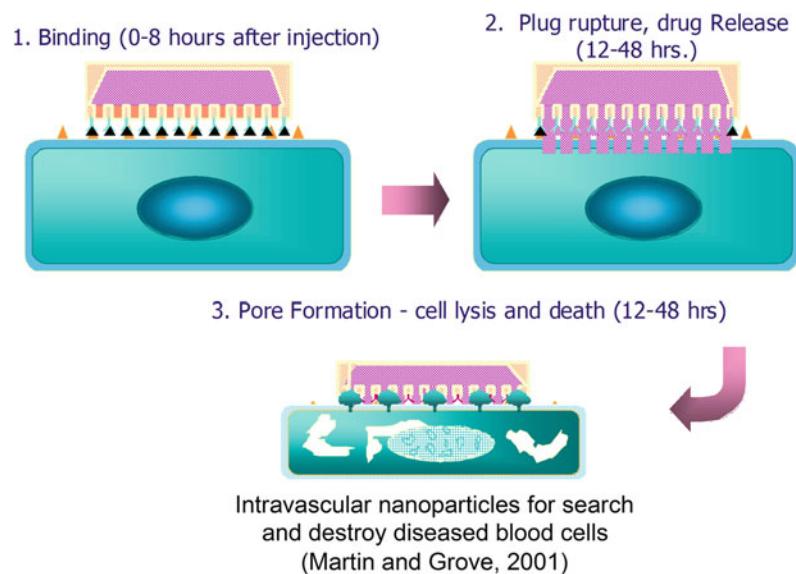
The main reliability concerns in micro/nanofluidic silicon arrays and implantable biocapsules are biocompatibility and potential biofouling (undesirable accumulation of microorganisms) of the channels/membrane by means of protein and cell adsorption from biological fluids. Biofouling can also result in the clogging of the nanochannels/nanopores, which potentially could render the device ineffective. The adhesion of proteins and

cells to an implanted device can also cause detrimental results such as inflammation and excessive fibrosis. Deposition of the self-assembled monolayers of selected organic molecules on the channels implants, which makes them hydrophobic, presents an innovative solution to combat the adverse effects of the biological fluids (Bhushan et al. 2005b, 2006b; Kasai et al. 2005; Lee et al. 2005b).

Figure 16.16 shows a conceptual model of a targeted intravascular drug delivery device using nanoparticles to search-and-destroy disease (tumor) cells (Martin and Grove 2001). The tumor cells have a 1–2 orders of magnitude higher density of receptors than the normal cells and a lower pH. Some receptors are only expressed on tumor cells. Nanoparticles made of various materials including gold (Au), graphene, iron oxide, polymer, and silica have been studied in targeted drug delivery for cancer treatment (Duncan 2003; Murakami et al. 2004; Ferrari 2005; Decuzzi et al. 2009; Panyala et al. 2009; Ferrari 2010; Irvine 2011). With lateral dimensions of 1 μm or less, the particles are smaller than any blood cells. These particles can be injected into the blood stream and travel freely through the circulatory system. In order to direct these drug-delivery nanoparticles to cancer sites, their external surfaces are chemically modified to carry molecules that have lock-and-key binding specificity with molecules that support a growing cancer

mass. As the particles come in close proximity with the diseased cells, the ligands on the particle surfaces attach to the receptors on the cells. As soon as the particles dock on the cells, a compound is released that forms a pore on the membrane of the cells, which leads to cell death and, ultimately, death of the cancer mass that was being nourished by the blood vessel. The adhesive interactions are regulated by specific (ligand-receptor binding) and non-specific (short range van der Waals, electrostatic, steric) interactions (Decuzzi et al. 2005). Adhesion between nanoparticles and disease cells is required. Furthermore, the particles should travel close to the endothelium lining of vascular arteries to facilitate the interaction between the particles and diseased cells. Human capillaries can have radii as small as 4–5 μm . Decuzzi et al. (2005) analyzed the margination of a particle circulating in the blood stream and calculated the speed and time for margination (drifting of particles towards the blood vessel walls) as a function of density and diameter of particle, based on various forces present between the circulating particle and the endothelium lining. They reported that the particles used for drug delivery should have a radius smaller than the critical value (in the range of 100 nm). Studies show that a lateral force on the particles assists them in faster margination towards the endothelium walls. Thus, nonspherical particles are more desirable. In

Fig. 16.16 Intravascular nanoparticles to search and destroy diseased blood cells. Life cycles include (1) target binding, (2) drug release, and (3) cell death (Martin and Grove 2001)



addition, reducing agglomeration and friction of nano-objects moving through a liquid and in contact with surfaces is important for efficient transport to diseased sites.

Because of their unique mechanical and electrical properties, single walled and multiwalled carbon nanotubes (SWNT and MWNT) are being used for thermal management of high power devices, reinforced composites and super strong fiber and sheets, chemical and biological sensors, electromechanical devices, field emission devices, molecular electronics and computing (Bhushan 2008c; Hierold 2008). Figure 16.17 shows a SWNT biosensor (Chen et al. 2004). The conductance of carbon nanotube (CNT) devices changes when proteins adsorb on the surface. The change in electrical resistance is a measure of protein adsorption. For high performance, adhesion should be strong between the adsorbent and SWNT.

In summary, adhesion, stiction/friction, and wear clearly limit the lifetimes and compromise the performance and reliability of MEMS/NEMS and BioMEMS/BioNEMS. Figure 16.18 summarizes tribological problems encountered in some of the MEMS, MOEMS, RF-MEMS, and BioMEMS devices just discussed.

16.2.5 Microfabrication Processes

In addition to in-use stiction, stiction issues are also present in some processes used for the fabrication of MEMS/NEMS. For example, the last step in surface micromachining involves the removal of sacrificial layer(s) called release, since the microstructures are released from the surrounding sacrificial layer(s). The release is accomplished by an aqueous chemical etching, rinsing, and drying processes. Due to meniscus effects as a result of wet processes, the suspended structures can sometimes collapse and permanently adhere to the underlying substrate, as shown in Fig. 16.19 (Guckel and Burns 1989). Adhesion is caused by water molecules adsorbed on the adhering surfaces and/or because of formation of adhesive bonds by silica residues that remain on the surfaces after the water has evaporated. This so-called release stiction is overcome by using dry release methods, such as CO₂ critical point drying or sublimation drying (Mulhern et al. 1993). CO₂ at high pressure is in a supercritical state and becomes liquid. Liquid CO₂ is used to remove wet etchant, and then it is converted back to gas phase.

16.2.6 Tribological Needs

Various MEMS/NEMS are designed to perform expected functions in the millisecond to picosecond range. The life expectancy of devices for high speed contacts can vary from a few hundred thousand to many billions of cycles, e.g., over a hundred billion cycles for DMDs, which puts stringent requirements on materials (Bhushan 1998; Man et al. 1998; Kayali et al. 1999; Tanner et al. 2000; Arney 2001; Man 2001). Adhesion between a biological molecular layer and the substrate (referred to as “bioadhesion”), reduction of friction and wear of biological layers, biocompatibility, and biofouling for BioMEMS/BioNEMS are important (Bhushan et al. 2005a, 2006a). Most mechanical properties are known to be scale dependent (Bhushan et al.

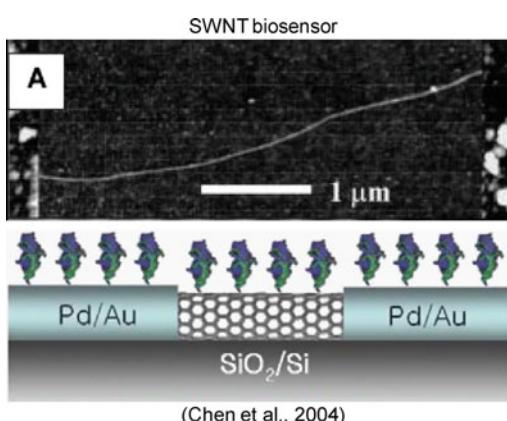
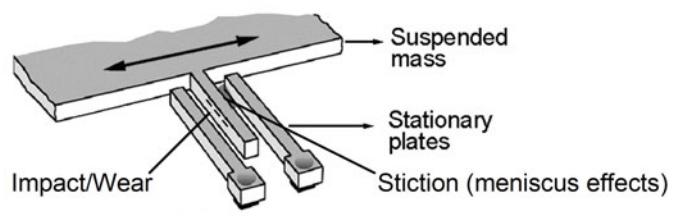


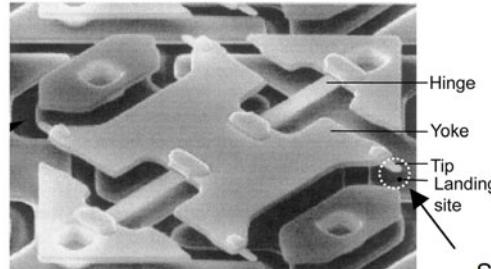
Fig. 16.17 SEM micrograph of SWNT biosensor, bottom schematic shows the adsorption of protein molecules to the SWNT (Chen et al. 2004)

Fig. 16.18 Summary of tribological issues in MEMS, MOEMS, RF-MEMS and BioMEMS device operation

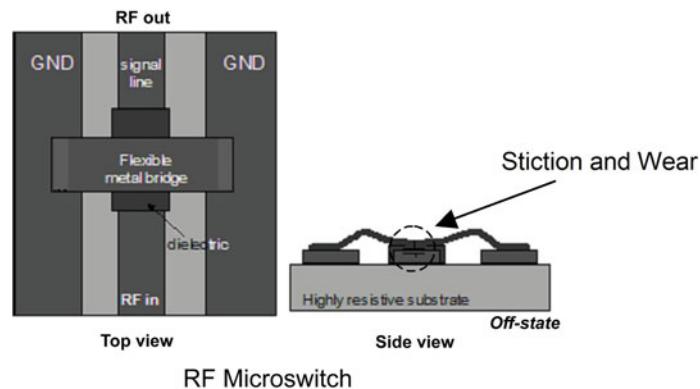
Tribological issues during device operation



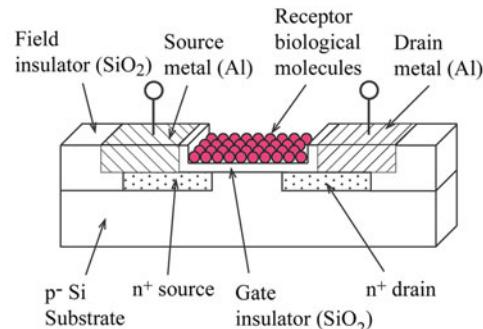
Capacitive type accelerometer



Digital micromirror device



RF Microswitch



BioFET sensor
(Bhushan et al., 2005)

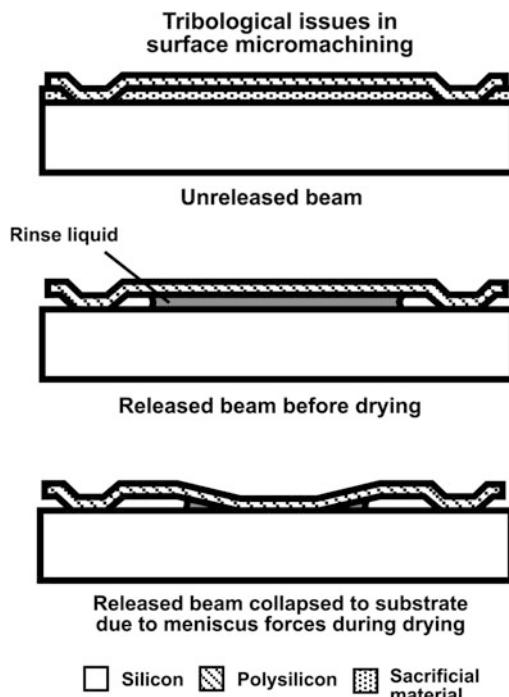


Fig. 16.19 Tribological issues in microfabrication via surface micromachining

1996). Therefore, the properties of nanoscale structures need to be measured (Sundararajan and Bhushan 2002). There is a need to develop a fundamental understanding of adhesion, friction/stiction, wear, and the role of surface contamination, and environment (Bhushan 1998). MEMS/NEMS materials need to possess good mechanical and tribological properties on the micro/nanoscale. There is a need to develop lubricants and identify lubrication methods that are suitable for MEMS/NEMS devices and applications. Methods need to be developed to enhance adhesion between biomolecules and the device substrate, referred to as bioadhesion. Component-level studies are required to provide a better understanding of the tribological phenomena occurring in MEMS/NEMS.

The fields of nanotribology and nanomechanics and atomic force microscopy-based techniques have provided researchers a viable approach to address these problems (Bhushan et al. 1995a; Bhushan 1999a, 2005, 2007, 2008a, b, 2011). The following is an overview of nanoscale adhesion,

friction and wear studies of materials and lubrication for MEMS/NEMS and BioMEMS/BioNEMS, and component-level studies of stiction phenomena in MEMS/NEMS devices.

16.3 Nanotribology and Nanomechanics Studies of Silicon and Related Materials

The materials of most interest for planar fabrication processes using silicon as the structural material are undoped and boron-doped (p^+ -type) single-crystal silicon for bulk micromachining, and doped and undoped phosphorus (n^+ -type) low-pressure chemical vapor deposition (LPCVD) polysilicon films for surface micromachining. Since silicon-based devices lack high-temperature capabilities with respect to both mechanical and electrical properties, SiC is being developed as a structural material for high-temperature microsensor and microactuator applications (Shor et al. 1993; Mehregany et al. 1998; Yang et al. 2001; Young et al. 2004). SiC can also be desirable for high frequency micromechanical resonators, in the GHz range, because of its high modulus of elasticity to density ratio and, consequently, high resonance frequency. Table 16.2 compares selected bulk properties of SiC and Si(100). Researchers have found low-cost techniques for producing single-crystalline 3C-SiC (cubic or β -SiC) films via epitaxial growth on large area silicon substrates for bulk micromachining (Zorman et al. 1995) and polycrystalline 3C-SiC films on polysilicon and silicon dioxide layers for surface micromachining of SiC (Zorman et al. 1998). Single crystalline 3C-SiC piezoresistive pressure sensors have been fabricated using bulk micromachining (Yang et al. 2001; Young et al. 2004). Surface micromachined polycrystalline SiC micromotors have been fabricated and reported to provide satisfactory operation at high temperatures (Yasseen et al. 2000). Because of their importance to MEMS/NEMS applications, the tribology of these materials must be examined and optimized.

Table 16.2 Selected bulk properties^a of 3C (β - or cubic) SiC and Si(100)

Sample	Density (kg/m ³)	Hard-ness (GPa)	Elastic modulus (GPa)	Fracture tough-ness (MPa m ^{1/2})	Thermal conductivity ^b (W/m K)	Coeff. of thermal expansion ^b ($\times 10^{-6}/\text{K}$)	Melting point (°C)	Band-gap (eV)
β -SiC	3210	23.5–26.5	440	4.6	85–260	4.5–6	2830	2.3
Si(100)	2330	9–10	130	0.95	155	2–4.5	1410	1.1

^aUnless stated otherwise, data shown were obtained from Bhushan and Gupta (1991)

^bObtained from Shackelford et al. (1994)

In this section, both macroscale and microscale tribological properties of virgin and treated/coated silicon, polysilicon films, and SiC are presented. As will be shown, bare silicon exhibits inadequate tribological performance and needs to be coated with a solid and/or liquid overcoat or be surface treated (e.g., oxidation and ion implantation, commonly used in semiconductor manufacturing), which reduces friction and wear. SiC films exhibit good tribological performance.

16.3.1 Virgin and Treated/Coated Silicon Samples

16.3.1.1 Nanotribological and Nanomechanical Properties of Silicon and Effect of Ion Implantation

Friction and wear of single-crystalline and polycrystalline silicon samples have been studied, and the effect of ion implantation with various doses of C⁺, B⁺, N₂⁺, and Ar⁺ ion species at 200 keV energy to improve their friction and wear properties has been studied (Gupta et al. 1993, 1994; Gupta and Bhushan 1994). The coefficient of macroscale friction and the wear factor of virgin single-crystal silicon and C⁺-implanted silicon samples as a function of ion dose are presented in Fig. 16.20 (Gupta et al. 1993). The macroscale friction and wear tests were conducted using a ball-on-flat tribometer. Each data bar represents the average value of four to six measurements. The coefficient of friction and wear factor for bare silicon are very high and decrease drastically with ion dose. Silicon samples bombarded above the

ion dose of 10¹⁷ C⁺ cm⁻² exhibit extremely low values of coefficients of friction (typically 0.03–0.06 in air) and wear factors (reduced by as much as four orders of magnitude). Gupta et al. (1993) reported that a decrease in the coefficient of friction and wear factor of silicon as a result of C⁺ ion bombardment occurred because of formation of silicon carbide rather than amorphization of silicon. Gupta et al. (1994) also reported an improvement in friction and wear with B⁺ ion implantation.

Microscale friction measurements were performed using an atomic force/friction force microscope (AFM/FFM) (Bhushan 1999a, 2005, 2007, 2008a, 2011). Table 16.3 shows values of surface roughness and coefficients of macroscale and microscale friction for virgin and doped silicon. There is a decrease in the coefficients of microscale and macroscale friction values as a result of ion implantation. When measured for small contact areas and very low loads used in microscale studies, indentation hardness and elastic modulus are higher than that at the macroscale. This, added to the effect of the small apparent area of contact reducing the number of trapped particles on the interface, results in less plowing contribution and lower friction in the case of microscale friction measurements. Figure 16.21a, b show results of microscale wear resistance studies of ion-implanted silicon samples studied using a diamond tip in an AFM (Bhushan and Koinkar 1994) are shown in. For tests conducted at various loads on Si(111) and C⁺-implanted Si(111), it is noted that the wear resistance of the implanted sample is slightly poorer than that of virgin silicon up to about 80 μN . Above 80 μN , the wear resistance of implanted Si improves. As one continues to run

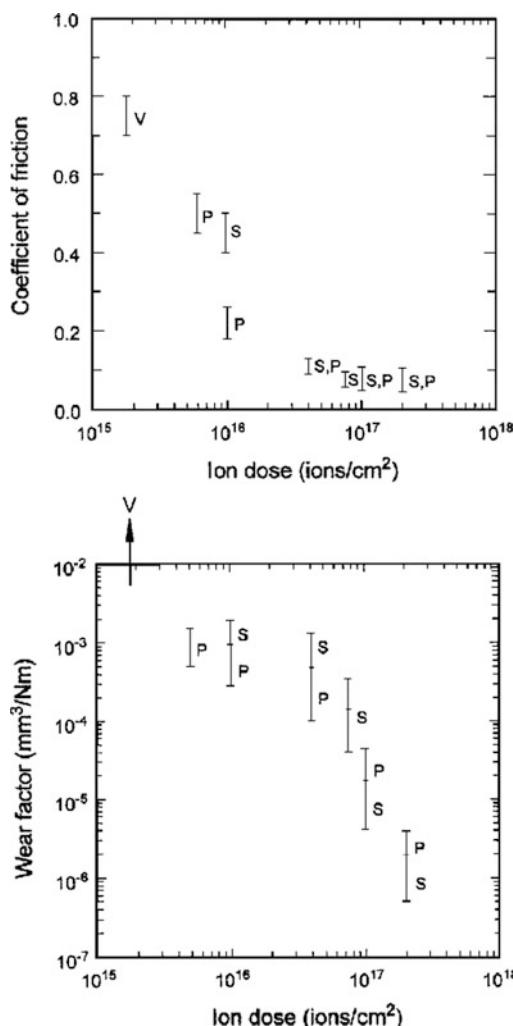


Fig. 16.20 Influence of ion doses on the coefficient of friction and wear factor on C⁺ ion bombarded single-crystal and polycrystalline silicon slid against alumina ball. V corresponds to virgin single-crystal silicon, while S and P denote tests that correspond to doped single- and polycrystalline silicon, respectively (Gupta et al. 1993)

Table 16.3 Surface roughness and micro- and macroscale coefficients of friction of selected samples

Material	RMS roughness (nm)	Coefficient of microscale friction ^a	Coefficient of macroscale friction ^b
Si(111)	0.11	0.03	0.33
C ⁺ -implanted Si (111)	0.33	0.02	0.18

^aVersus Si₃N₄ tip, tip radius of 50 nm in the load range of 10–150 nN (2.5–6.1 GPa) at a scanning speed of 5 μm/s over a scan area of 1 μm × 1 μm in an AFM

^bVersus Si₃N₄ ball, ball radius of 3 mm at a normal load of 0.1 N (0.3 GPa) at an average sliding speed of 0.8 mm/s using a tribometer

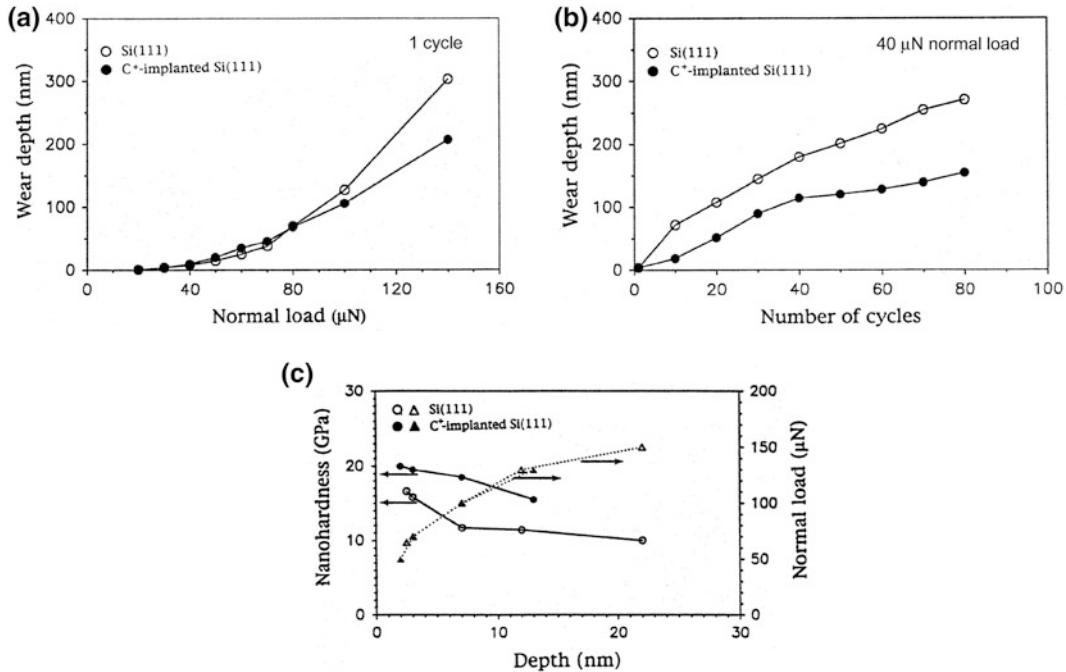


Fig. 16.21 Wear depth as a function of **a** load (after one cycle), and **b** cycles (normal load = 40 mN) for Si(111) and C⁺-implanted Si(111). **c** Nanohardness and normal load as function of indentation depth for virgin and C⁺-implanted Si (111) (Bhushan and Koinkar 1994)

tests at 40 μN for a larger number of cycles, the implanted sample, which forms hard and tough silicon carbide, exhibits higher wear resistance than the unimplanted sample. Damage from the implantation in the top layer results in poorer wear resistance; however, the implanted zone at the subsurface is more wear resistant than the virgin silicon.

Hardness values of virgin and C⁺-implanted Si(111) at various indentation depths (normal loads) are presented in Fig. 16.21c (Bhushan and Koinkar 1994). The hardness at a small indentation depth of 2.5 nm is 16.6 GPa, and it drops to a value of 11.7 GPa at a depth of 7 nm and a normal load of 100 μN . Higher hardness values obtained in low-load indentation may arise from the observed pressure-induced phase transformation during the nanoindentation (Pharr 1991; Callahan and Morris 1992). Additional increases in the hardness at an even lower indentation depth of 2.5 nm reported here may arise from the contribution of complex chemical films (not from native oxide films) present on the silicon surface.

At small volumes there is a lower probability of encountering material defects (dislocations, etc.). Furthermore, according to the strain gradient plasticity theory advanced by Fleck et al. (1994), large strain gradients inherent in small indentations lead to accumulation of geometrically necessary dislocations that cause enhanced hardening. These are some of the plausible explanations for an increase in hardness at smaller volumes. If the silicon material were to be used at very light loads, such as in microsystems, the high hardness of surface films would protect the surface until it is worn.

From Fig. 16.21c, hardness values of C⁺-implanted Si(111) at a normal load of 50 μN is 20.0 GPa with an indentation depth of about 2 nm, which is comparable to the hardness value of 19.5 GPa at 70 μN , whereas the measured hardness value for virgin silicon at an indentation depth of about 7 nm (normal load of 100 μN) is only about 11.7 GPa. Thus, ion implantation with C⁺ results in an increase in hardness in silicon. Note that the surface layer of the

implanted zone is much harder compared with the subsurface, and may be brittle leading to higher wear on the surface. The subsurface of the implanted zone (SiC) is harder than the virgin silicon, resulting in higher wear resistance, which is also observed in the results of the macroscale tests conducted at high loads.

16.3.1.2 Effect of Oxide Films on Nanotribological Properties of Silicon

Macroscale friction and wear experiments have been performed using a magnetic disk drive with bare, oxidized, and implanted pins sliding against amorphous-carbon coated magnetic disks lubricated with a thin layer of perfluoropolyether lubricant (Bhushan and Venkatesan 1993; Venkatesan and Bhushan 1993, 1994; Bhushan 1999b). Representative profiles of the variation of the coefficient of friction with a number of sliding cycles for an Al₂O₃-TiC slider and bare and dry-oxidized silicon pins are shown in Fig. 16.22. For bare Si(111), after an initial increase in the coefficient of friction, it drops to a steady state value of 0.1, as seen in Fig. 16.22. The rise in the coefficient of friction for the Si(111) pin is associated with the transfer of amorphous carbon from the disk to the pin, oxidation-enhanced fracture of pin material followed by tribochemical oxidation of the transfer film, while the drop is associated with the formation of a transfer coating on the pin. Dry-oxidized Si(111) exhibits excellent characteristics and no significant increase was observed over 50,000 cycles (Fig. 16.22). This behavior has been attributed to the chemical passivity of the oxide and lack of transfer of diamond-like carbon (DLC) from the disk to the pin. The behavior of PECVD-oxide (data are not presented here) was comparable to that of dry oxide, but for the wet oxide there was some variation in the coefficient of friction (0.26–0.4). The difference between the dry and wet oxide was attributed to increased porosity of the wet oxide (Bhushan and Venkatesan 1993). Since tribochemical oxidation was determined to be a significant factor, experiments were conducted in dry nitrogen (Venkatesan and Bhushan 1993,

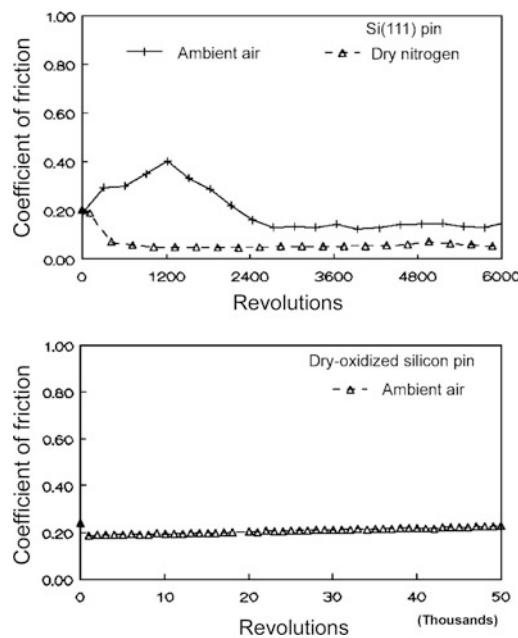


Fig. 16.22 Coefficient of friction as a function of number of sliding revolutions for Si(111) pin in ambient air and dry nitrogen and dry-oxidized silicon pin in ambient air (Bhushan and Venkatesan 1993)

1994). The variation of the coefficient of friction for a silicon pin sliding against a thin-film disk in dry nitrogen is shown in Fig. 16.22. It is seen that in a dry nitrogen environment, the coefficient of friction of Si(111) sliding against a disk decreased from an initial value of about 0.2–0.05 with continued sliding. Based on SEM and chemical analysis, this behavior has been attributed to the formation of a smooth amorphous-carbon/lubricant transfer patch and suppression of oxidation in a dry nitrogen environment. Based on macroscale tests using disk drives, it is found that the friction and wear performance of bare silicon is not adequate. With dry-oxidized or PECVD SiO₂-coated silicon, no significant friction increase or interfacial degradation was observed in ambient air.

Table 16.4 and Fig. 16.23 show surface roughness, microscale friction and scratch data, and nanoindentation hardness for the various silicon samples (Bhushan and Koinkar 1994). Scratch experiments were performed using a diamond tip in an AFM. Results on polysilicon samples are also shown for comparison.

Table 16.4 RMS, microfriction, microscratching/microwear and nanoindentation hardness data for various virgin, coated and treated silicon samples

Material	RMS roughness ^a (nm)	Coefficient of microscale friction ^b	Scratch depth ^c at 40 μN (nm)	Wear depth ^c at 40 μN (nm)	Nanohardness ^c at 100 μN (GPa)
Si(111)	0.11	0.03	20	27	11.7
Si(110)	0.09	0.04	20		
Si(100)	0.12	0.03	25		
Polysilicon	1.07	0.04	18		
Polysilicon (lapped)	0.16	0.05	18	25	12.5
PECVD-oxide coated Si(111)	1.50	0.01	8	5	18.0
Dry-oxidized Si (111)	0.11	0.04	16	14	17.0
Wet-oxidized Si(111)	0.25	0.04	17	18	14.4
C ⁺ -implanted Si(111)	0.33	0.02	20	23	18.6

^aScan size of 500 nm × 500 nm using AFM

^bVersus Si₃N₄ tip in AFM/FFM, radius 50 nm; at 1 μm × 1 μm scan size

^cMeasured using an AFM with a diamond tip of radius 100 nm

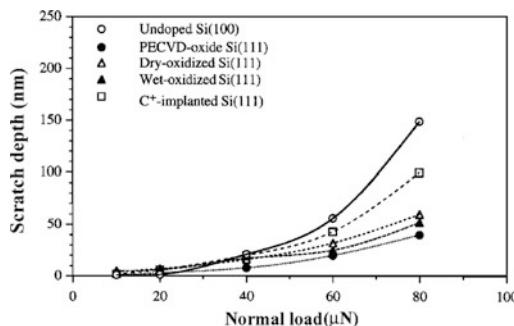


Fig. 16.23 Scratch depth as a function of normal load after 10 cycles for various silicon samples, virgin, treated, and coated (Bhushan and Koinkar 1994)

Coefficients of microscale friction values for all the samples are about the same. These samples showed wear scars from scratching with a 10 μN load. Scratch depth increased with normal load. Crystalline orientation of silicon has little influence on scratch resistance because natural oxidation of silicon in ambient masks the expected effect of crystallographic orientation. PECVD-

oxide samples showed the best scratch resistance, followed by dry-oxidized, wet-oxidized, and ion-implanted samples. Ion implantation with C⁺ does not appear to improve scratch resistance.

Wear data on the silicon samples are also presented in Table 16.4 (Bhushan and Koinkar 1994). PECVD-oxide samples showed superior wear resistance followed by dry-oxidized, wet-oxidized, and ion-implanted samples. This agrees with the trends seen in scratch resistance. In PECVD, ion bombardment during the deposition improves the coating properties such as suppression of columnar growth, freedom from pinhole, decrease in crystalline size, and increase in density, hardness, and substrate-coating adhesion. These effects may help in improving the mechanical integrity of the sample surface. Coatings and treatments improved the nanohardness of silicon. Note that dry-oxidized and PECVD films are harder than wet-oxidized films as these films may be porous. High hardness of oxidized films may be responsible for measured high scratch/wear resistance.

16.3.2 Nanotribological and Nanomechanical Properties of Polysilicon Films and SiC Film

Studies have also been conducted on undoped polysilicon film, heavily doped (n^+ -type) polysilicon film, heavily doped (p^+ -type) single-crystal Si(100), and 3C-SiC (cubic or β -SiC) film (Bhushan et al. 1998; Sundararajan and Bhushan 1998; Li and Bhushan 1999). The polysilicon films studied here are different from the ones discussed previously.

Table 16.5 presents a summary of the tribological studies conducted on polysilicon and SiC films. Values for single-crystal silicon are also

shown for comparison. Polishing of the as-deposited polysilicon and SiC films drastically affects the roughness, as the values reduce by two orders of magnitude. Si(100) appears to be the smoothest followed by polished undoped polysilicon and SiC films, which have comparable roughness. The doped polysilicon film shows higher roughness than the undoped sample, which is attributed to the doping process. Polished SiC film shows the lowest friction followed by polished and undoped polysilicon film, which strongly supports the candidacy of SiC films for use in MEMS/NEMS devices. Macroscale friction measurements indicate that SiC film exhibits one of the lowest friction values as compared to the other samples. Doped polysilicon sample

Table 16.5 Summary of micro/nanotribological properties of the sample materials

Sample	RMS roughness ^a (nm)	P-V distance ^a (nm)	Coefficient of friction Micro ^b Macro ^c	Scratch depth ^d (nm)	Wear depth ^e (nm)	Nanohardness ^f (GPa)	Young's modulus ^f (GPa)	Fracture toughness ^g , K_{IC} MPa m ^{1/2}
Undoped Si (100)	0.09	0.9	0.06	89	84	12	168	0.75
			0.33					
Undoped polysilicon film (as deposited)	46	340	0.05					
Undoped polysilicon film (polished)	0.86	6	0.04	99	140	12	175	1.11
			0.46					
n^+ -type polysilicon film (as deposited)	12	91	0.07					
n^+ -type polysilicon film (polished)	1.0	7	0.02	61	51	9	95	0.89
			0.23					
SiC film (as deposited)	25	150	0.03					
SiC film (polished)	0.89	6	0.02	6	16	25	395	0.78
			0.20					

^aMeasured using AFM over a scan size of 10 $\mu\text{m} \times 10 \mu\text{m}$

^bMeasured using AFM/FFM over a scan size of 10 $\mu\text{m} \times 10 \mu\text{m}$

^cObtained using a 3-mm diameter sapphire ball in a reciprocating mode at a normal load of 10 mN and average sliding speed of 1 mm/s after 4 m sliding distance

^dMeasured using AFM at a normal load of 40 μN for 10 cycles, scan length of 5 μm

^eMeasured using AFM at normal load of 40 μN for 1 cycle, wear area of 2 $\mu\text{m} \times 2 \mu\text{m}$

^fMeasured using Nanoindenter at a peak indentation depth of 20 nm

^gMeasured using microindenter with Vickers indenter at a normal load of 0.5 N

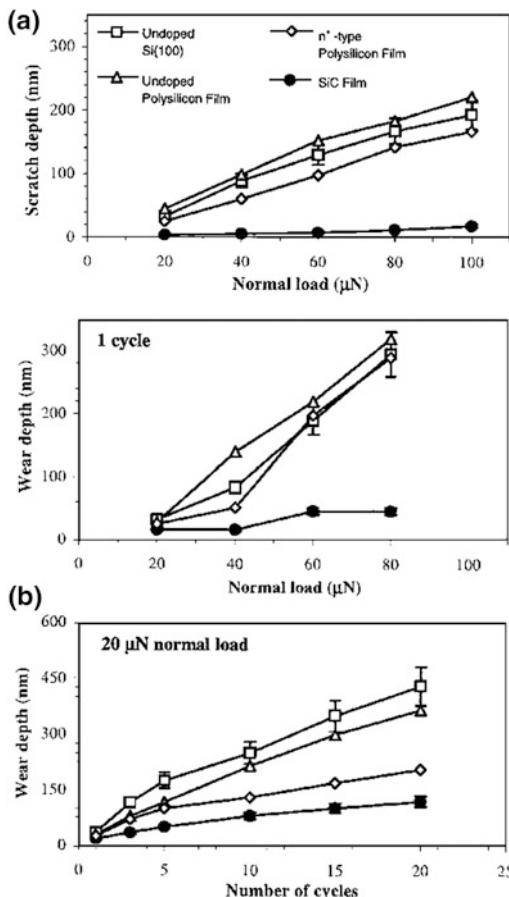


Fig. 16.24 **a** Scratch depths for 10 cycles as a function of normal load and **b** wear depths as a function of normal load and as a function of number of cycles for various samples (Bhushan et al. 1998)

shows low friction on the macroscale as compared to the undoped polysilicon sample possibly due to the doping effect.

Figure 16.24a shows a plot of scratch depth versus normal load for various samples (Bhushan et al. 1998; Sundararajan and Bhushan 1998). Scratch depth increases with increasing normal load. Figure 16.25 shows AFM 3D maps and averaged 2D profiles of the scratch marks on the various samples. It is observed that scratch depth increases almost linearly with the normal load. Si (100) and the doped and undoped polysilicon film show similar scratch resistance. From the data, it is clear that the SiC film is much more scratch resistant than the other samples. Figure 16.24b

shows results from microscale wear tests on the various films. For all the materials, the wear depth increases almost linearly with increasing number of cycles. This suggests that the material is removed layer by layer in all the materials. Here also, SiC film exhibits lower wear depths than the other samples. Doped polysilicon film wears less than the undoped film. Higher fracture toughness and higher hardness of SiC as compared to Si (100) is responsible for its lower wear. Also, the higher thermal conductivity of SiC (see Table 16.2) as compared to the other materials leads to lower interface temperatures, which generally results in less degradation of the surface (Bhushan 1996a, 2013a, b). Doping of the polysilicon does not affect the scratch/wear resistance and hardness much. The measurements made on the doped sample are affected by the presence of grain boundaries. These studies indicate that SiC film exhibits desirable tribological properties for use in MEMS devices.

16.4 Lubrication Studies for MEMS/NEMS

MEMS/NEMS devices may benefit from various lubricants. Several studies of liquid perfluoropolyether (PFPE) lubricant films, self-assembled monolayers (SAMs), and hard diamondlike carbon (DLC) coatings have been carried out for the purpose of minimizing adhesion, friction, and wear (Bhushan et al. 1995a; Bhushan 1999a, b, 2001, 2003, 2005, 2007, 2008a, b, 2011, 2013a, b). Many variations of these films are hydrophobic (low surface tension and high contact angle) and have low shear strength which provides low adhesion, friction, and wear. Relevant details are presented here.

16.4.1 Perfluoropolyether Lubricants

The classical approach to lubrication uses freely-supported multimolecular layers of liquid lubricants (Bhushan 1999a, 2011, 2013a, b). The liquid lubricants are sometimes chemically bonded to improve their wear resistance. Partially

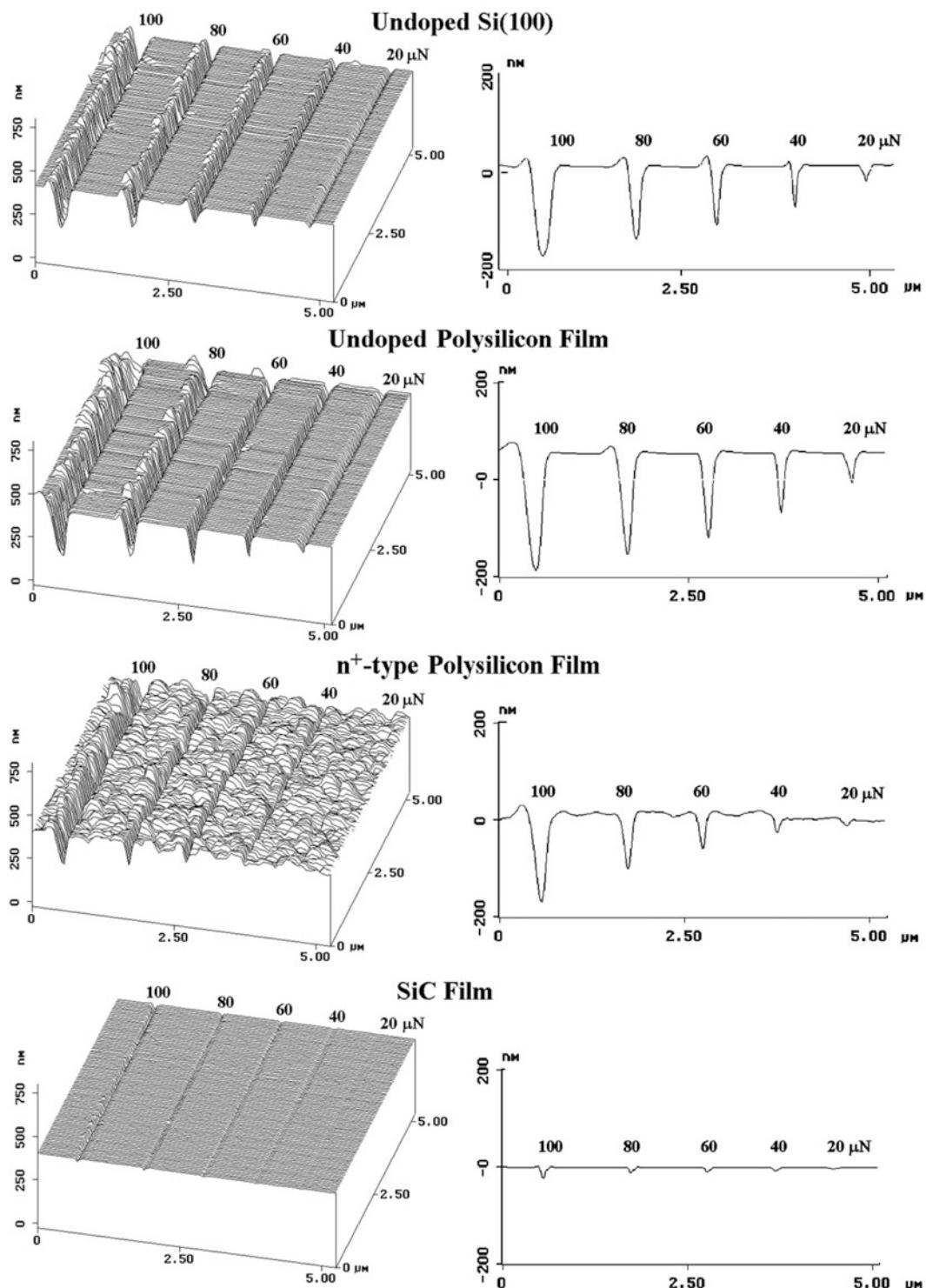


Fig. 16.25 AFM 3D maps and averaged 2D profiles of scratch marks on various samples (Bhushan et al. 1998)

chemically-bonded, molecularly-thick PFPE lubricants are widely used for lubrication of magnetic storage media because of their thermal stability and extremely low vapor pressure (Bhushan 1996a), and are found to be suitable for MEMS/NEMS devices.

Adhesion, friction, and durability experiments have been performed on virgin Si (100) surfaces and silicon surfaces lubricated with various PFPE lubricants (Koinkar and Bhushan 1996a, b; Bhushan 1999a, 2005; Liu and Bhushan 2003a; Tao and Bhushan 2005a; Bhushan et al. 2007; Palacio and Bhushan 2007a, b). Results of two PFPE lubricants will be presented here: Z-15 (with $-\text{CF}_3$ nonpolar end groups), $\text{CF}_3\text{O}-(\text{CF}_2\text{CF}_2\text{O})_m-(\text{CF}_2\text{O})_n-\text{CF}_3$ ($m/n \sim 2/3$), and Z-DOL (with $-\text{OH}$ polar end groups), $\text{HO}-\text{CH}_2-\text{CF}_2\text{O}-(\text{CF}_2\text{CF}_2\text{O})_m-(\text{CF}_2\text{O})_n-\text{CF}_2\text{CH}_2-\text{OH}$ ($m/n \sim 2/3$). The Z-DOL film was thermally bonded at 150° for 30 min, and the unbonded fraction was removed by a solvent (fully bonded) (Bhushan 1996a). The thicknesses of Z-15 and the Z-DOL (fully bonded) films were 2.8 and 2.3 nm, respectively. Nanoscale measurements were made using an AFM. The adhesive forces of Si(100), Z-15, and Z-DOL (fully bonded) measured by force calibration plot, and friction force versus normal load plot are summarized in Fig. 16.26. The results measured by these two methods are in good agreement. Figure 16.26 shows that the presence of mobile Z-15 lubricant film increases the adhesive force as compared to that of Si(100) by meniscus formation (Stifter et al. 2000; Bhushan 2013a, b). Whereas, the presence of solid phase Z-DOL (fully bonded) film reduces the adhesive force as compared that of Si(100) because of the absence of mobile liquid. The schematic (bottom) in Fig. 16.26 shows relative size and sources of meniscus. It is well known that the native oxide layer (SiO_2) on the top of Si(100) wafer exhibits hydrophilic properties, and some water molecules can be adsorbed on this surface. The condensed water will form a meniscus as the tip approaches the sample surface. The larger adhesive force in Z-15 is not just caused by the Z-15 meniscus; the non-polarized Z-15 liquid does not have good wettability and strong bonding with Si(100). Consequently, in

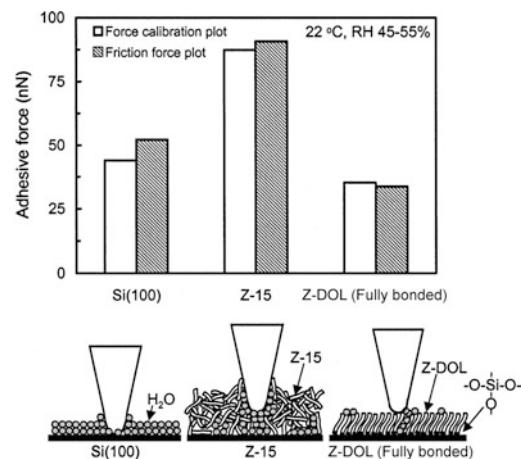


Fig. 16.26 Summary of the adhesive forces of Si(100) and Z-15 and Z-DOL (fully bonded) films measured by force calibration plots and friction force versus normal load plots in ambient air. The schematic (bottom) showing the effect of meniscus, formed between AFM tip and the surface sample, on the adhesive and friction forces (Liu and Bhushan 2003a)

the ambient environment, the condensed water molecules from the environment permeate through the liquid Z-15 lubricant film and compete with the lubricant molecules presented on the substrate. The interaction of the liquid lubricant with the substrate is weakened, and a boundary layer of the liquid lubricant forms puddles (Koinkar and Bhushan 1996a, b). This dewetting allows water molecules to be adsorbed on the Si (100) surface as aggregates along with Z-15 molecules, and both of them can form a meniscus while the tip approaches the surface. Thus, the dewetting of liquid Z-15 film results in higher adhesive force and poorer lubrication performance. In addition, as the Z-15 film is soft compared to the solid Si(100) surface, penetration of the tip in the film occurs while pushing the tip down. This leads to a large area of the tip involved in forming the meniscus at the tip-liquid (mixture of Z-15 and water) interface. It should also be noted that Z-15 has a higher viscosity compared to water, therefore the Z-15 film provides higher resistance to lateral motion and coefficient of friction. In the case of the Z-DOL (fully bonded) film, the active groups of Z-DOL molecules are mostly bonded on Si(100)

substrate, thus the Z-DOL (fully bonded) film has low free surface energy and cannot be displaced readily by water molecules or readily adsorb water molecules. Thus, the use of Z-DOL (fully bonded) can reduce the adhesive force.

To study the relative humidity effect on friction and adhesion, the variation of friction force, adhesive force, and coefficient of friction of Si(100), Z-15, and Z-DOL (fully bonded) as a function of relative humidity are shown in Fig. 16.27. It shows that for Si(100) and Z-15

film, the friction force increases with a relative humidity increase up to 45%, and then it shows a slight decrease with further increase in the relative humidity. Z-DOL (fully bonded) has smaller friction force than Si(100) and Z-15 in the whole testing range, and its friction force shows a relative apparent increase when the relative humidity is higher than 45%. For Si(100), Z-15, and Z-DOL (fully bonded), their adhesive forces increase with relative humidity, and their coefficients of friction increase with a relative humidity

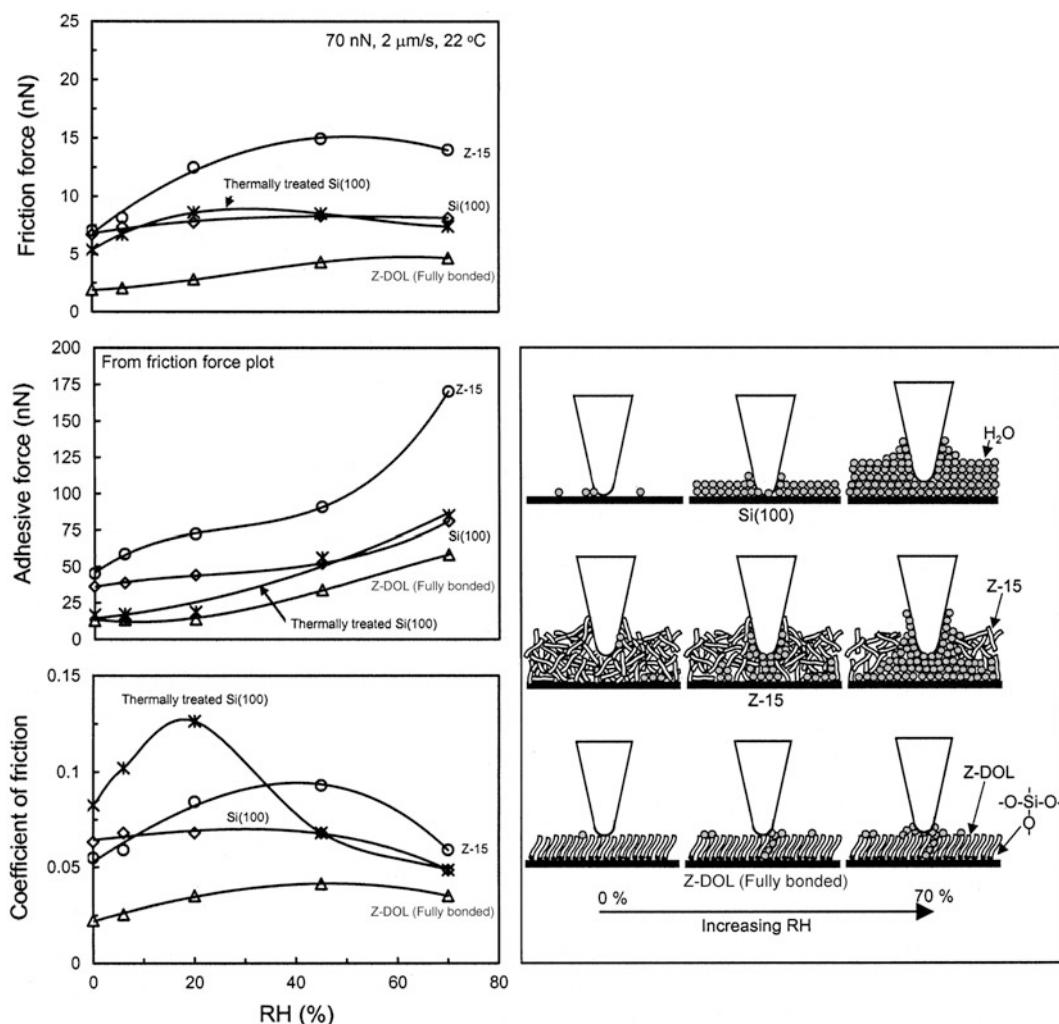


Fig. 16.27 The influence of relative humidity of the friction force, adhesive force, and coefficient of friction of Si(100) and Z-15 and Z-DOL (fully bonded) films at 70 nN, 2 $\mu\text{m}/\text{s}$, and in 22 $^{\circ}\text{C}$ air. Schematic (right) shows the change of meniscus while increasing the relative humidity. In this figure, the thermal treated Si(100) represents the Si(100) wafer that was baked at 150 $^{\circ}\text{C}$ for 1 h in an oven (in order to remove the adsorbed water) just before it was placed in the 0% RH chamber (Liu and Bhushan 2003a)

up to 45%, after which they decrease with further increasing of the relative humidity. It is also observed that the humidity effect on Si(100) really depends on the history of the Si(100) sample. As the surface of the Si(100) wafer readily adsorbs water in air, without any pre-treatment the Si(100) used in our study almost reaches its saturated stage of adsorbed water, and is responsible for less effect during increasing relative humidity. However, once the Si(100) wafer was thermally treated by baking at 150 °C for 1 h, a greater effect was observed.

The schematic (right) in Fig. 16.27 shows that Si(100), because of its high free surface energy, can adsorb more water molecules while increasing relative humidity. As discussed earlier, for the Z-15 film in the humid environment, the condensed water competes with the lubricant film present on the sample surface, and interaction of the liquid lubricant film with the silicon substrate is weakened, and a boundary layer of the liquid lubricant forms puddles. This dewetting allows water molecules to be adsorbed on the Si(100) substrate mixed with Z-15 molecules (Koinkar and Bhushan 1996a, b). Obviously, more water molecules can be adsorbed on Z-15 surface with increasing relative humidity. The more adsorbed water molecules in the case of Si(100), along with lubricant molecules in Z-15 film case, form a bigger water meniscus which leads to an increase of friction force, adhesive force, and coefficient of friction of Si(100) and Z-15 with humidity. But at a very high humidity of 70%, large quantities of adsorbed water can form a continuous water layer that separates the tip and sample surface, acting as a kind of lubricant, which causes a decrease in the friction force and coefficient of friction. For Z-DOL (fully bonded) film, because of their hydrophobic surface properties, water molecules can be adsorbed at a humidity higher than 45%, and cause an increase in the adhesive force and friction force.

To study the durability of lubricant films at the nanoscale, the friction force of Si(100), Z-15, and Z-DOL (fully bonded) as a function of the number of scanning cycles are shown in Fig. 16.28. As observed earlier, the friction force

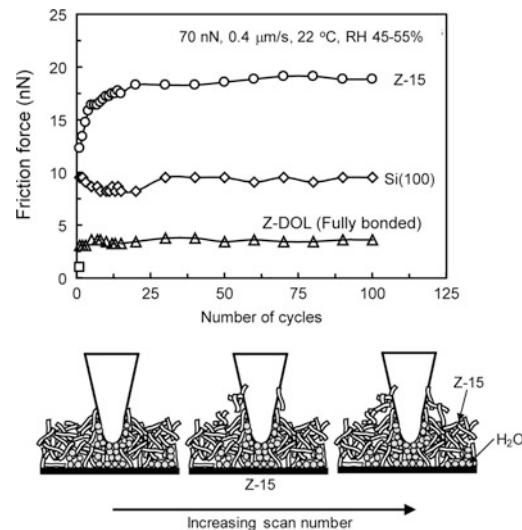


Fig. 16.28 Friction force versus number of sliding cycles for Si(100) and Z-15 and Z-DOL (fully bonded) films at 70 nN, 0.4 $\mu\text{m}/\text{s}$, and in ambient air. Schematic (bottom) shows that some liquid Z-15 molecules can be attached onto the tip. The molecular interaction between the attached molecules onto the tip with the Z-15 molecules in the film results in an increase of the friction force with multi scanning (Liu and Bhushan 2003a)

of Z-15 is higher than that of Si(100), with the lowest values for Z-DOL(fully bonded). During cycling, the friction force of Si(100) shows a slight decrease during initial few cycles then remain constant. This is related to the removal of the top adsorbed layer. In the case of the Z-15 film, the friction force shows an increase during the initial few cycles and then approaches to higher and stable values. This is believed to be caused by the attachment of the Z-15 molecules onto the tip. The molecular interaction between these attached molecules to the tip and molecules on the film surface is responsible for an increase in friction. However, after several scans, this molecular interaction reaches equilibrium, and after that, friction force and the coefficient of friction remain constant. In the case of the Z-DOL (fully bonded) film, the friction force and coefficient of friction start out low and remain low during the entire test for 100 cycles. It suggests that Z-DOL (fully bonded) molecules do not get attached or displaced as readily as Z-15.

16.4.2 Self-assembled Monolayers (SAMs)

For lubrication of MEMS/NEMS, another effective approach involves the deposition of organized and dense molecular layers of long-chain molecules. Two common methods to produce monolayers and thin films are the Langmuir-Blodgett (L-B) deposition and self-assembled monolayers (SAMs) by chemical grafting of molecules. L-B films are physically bonded to the substrate by a weak van der Waals attraction, while SAMs are chemically bonded via covalent bonds to the substrate. Because of the choice of chain length and terminal linking group that SAMs offer, they hold great promise for boundary lubrication of MEMS/NEMS. A number of studies have been conducted on tribological properties of various SAMs (Bhushan et al. 1995b; Bhushan and Liu 2001; Liu et al. 2001; Liu and Bhushan 2002; Bhushan et al. 2005b, 2006b, c; Kasai et al. 2005; Lee et al. 2005b; Tambe and Bhushan 2005a; Tao and Bhushan 2005b).

Bhushan and Liu (2001) studied the effect of film compliance on adhesion and friction. Based on friction and stiffness measurements, SAMs with high-compliance long-carbon chains exhibit low friction; chain compliance is desirable for low friction. The friction mechanism of SAMs is explained by a so-called “molecular spring” model. According to this model, the chemically adsorbed self-assembled molecules on a substrate are just like assembled molecular springs

anchored to the substrate. An asperity sliding on the surface of SAMs is like a tip sliding on the top of “molecular springs or brush.” The molecular spring assembly has compliant features and can experience orientation and compression under load. 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 orientation is determined by the spring constant of a single molecule 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 leads to energy dissipation. The SAMs with high-compliance long carbon chains also exhibit the best wear resistance (Bhushan and Liu 2001; Liu and Bhushan 2002). In wear experiments, the wear depth as a function of normal load curves show a critical normal load. Below the critical normal load, SAMs undergo orientation. At the critical load, SAMs wear away from the substrate due to weak interface bond strengths, while above the critical normal load, severe wear takes place on the substrate.

Bhushan et al. (2005b, 2006c), Kasai et al. (2005), Tambe and Bhushan (2005a), and Tao and Bhushan (2005b) studied various SAMs that were vapor phase deposited on Si and Al substrates with native oxide films, Fig. 16.29. Perfluorodecyltrichlorosilane (PFTS), n-octyldimethyl (dimethylamino) silane (ODMS) ($n = 7$), and n-octadecylmethyl (dimethylamino) silane

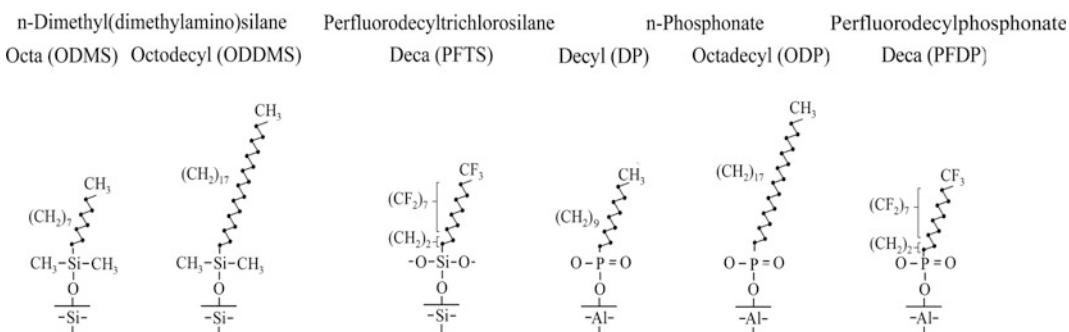


Fig. 16.29 Schematics of the structures of perfluoroalkylsilane and alkylsilane SAMs on Si with native oxide substrates, and perfluoroalkylphosphonate and alkylphosphonate SAMs on Al with native oxide substrates

(ODDMS) ($n = 17$) were deposited on Si substrates. Perfluorodecylphosphonate (PFDP), n-decylphosphonate (DP) ($n = 7$) and n-octadecylphosphonate (ODP) ($n = 17$) were deposited on Al substrates. Figure 16.30 shows the static contact angles, adhesive forces, friction forces, and coefficients of friction of the two substrates with various SAMs under ambient conditions (Bhushan et al. 2005b, 2006c). Based on the data, all films exhibit higher contact angles and lower adhesive forces and coefficients of friction as compared to their corresponding substrates. Among various films, PFTS/Si exhibits a higher contact angle and lower adhesive force as compared to that of ODMS/Si or ODDMS/Si. Longer-chain film ODDMS/Si has superior performance than shorter-chain film ODMS. Trends for films on Al substrates are similar to that on Si substrates. Thus, substrate had little effect. The coefficients of friction for various SAMs are comparable.

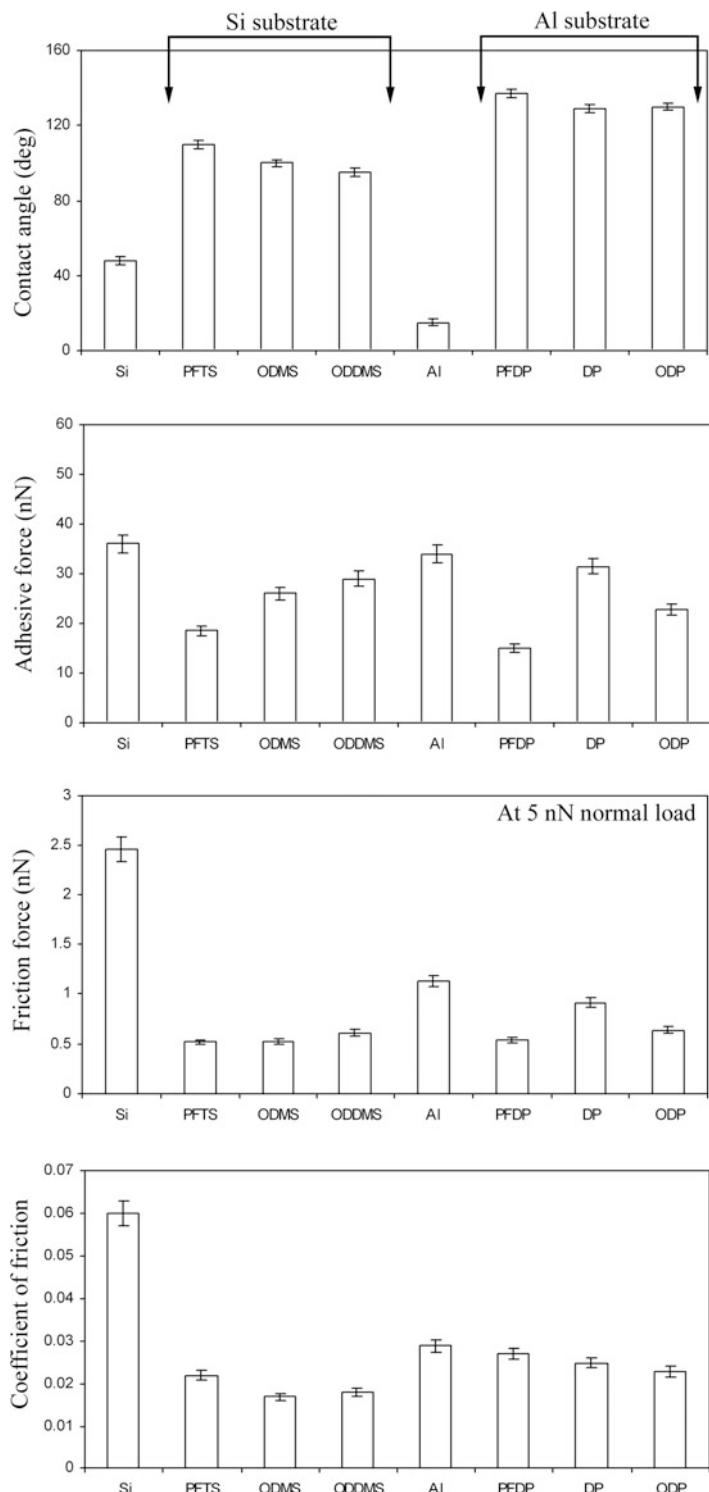
The effect of relative humidity on adhesion and friction for various SAMs on Si substrate was studied. Adhesive force, friction force at 5 nN of normal load, coefficient of friction, and micro-wear data are presented in Fig. 16.31 (Kasai et al. 2005; Bhushan et al. 2006c). Adhesive force for silicon showed an increase with relative humidity. This is expected since the surface of silicon is hydrophilic as shown in Fig. 16.30. More condensation of water at the tip-sample interface at higher humidity increases the adhesive force due to the capillary effect. On the other hand, the adhesive force for the SAMs showed a very weak dependency on the change in humidity. This occurs since the surface of the SAMs is hydrophobic. The adhesive force of ODMS/Si and ODDMS/Si showed a slight increase from 75 to 90% RH. Such increase was absent for PFTS/Si, possibly because of the hydrophobicity of PFTS/Si. The friction force of silicon showed an increase with relative humidity up to about 75% RH and a slight decrease beyond this point. The initial increase possibly results from the increase in adhesive force. The decrease in friction force at higher humidity could be attributed to the lubricating effect of the water layer. This effect is more pronounced in the coefficient of friction. Since the adhesive force increased and

coefficient of friction decreased in this range, those effects cancel out each other, and the resulting friction force showed slight changes. On the other hand, the friction force and coefficient of friction of SAMs showed very small changes with relative humidity like that found for adhesive force. This suggests that the adsorbed water layer on the surface maintained a similar thickness throughout the relative humidity range tested. The differences among the SAM types were small, within the measurement error. However a closer look at the coefficient of friction for ODMDS/Si showed a slight increase from 75 to 90% RH as compared to PFTS/Si, possibly for the same reason as the adhesive force changes. The inherent hydrophobicity of SAMs means that they did not show much relative humidity dependence.

Figure 16.32a shows the relationship between the decrease in surface height as a function of the normal load during wear tests for various SAMs on Si and Al substrates (Kasai et al. 2005; Bhushan et al. 2006c). As shown in the figure, the SAMs exhibit a critical normal load beyond the point of which the surface height drastically decreases. Figure 16.32a also shows the wear behavior of the Al and Si substrates. Unlike the SAMs, the substrates show a monotonic decrease in surface height with increasing normal load, with wear initiating from the very beginning, i.e., even for low normal loads. Si (Young's modulus of elasticity, $E = 130$ GPa (Anonymous 1988), hardness, $H = 11$ GPa (Bhushan 1999b)) is relatively hard in comparison to Al ($E = 77$ GPa, $H = 0.41$ GPa). Hence the decrease in surface height for Al is much larger than that for Si for similar normal loads.

The critical loads corresponding to the sudden failure of SAMs are shown in Fig. 16.32b. Amongst all the SAMs, ODDMS shows the best performance in the wear tests, believed to be due to the longer chain length effect. Fluorinated SAMs —PFTS and PFDP show higher critical load as compared to ODMS and DP with similar chain lengths. ODP shows higher critical load as compared to DP because of its longer chain length. The mechanism of failure of compliant SAMs during wear tests was presented earlier. It is believed that the SAMs fail mostly due to shearing of the

Fig. 16.30 The static contact angle, adhesive force, friction force, and coefficient of friction measured using an AFM for various SAMs on Si and Al substrates (Bhushan et al. 2005b, 2006c)



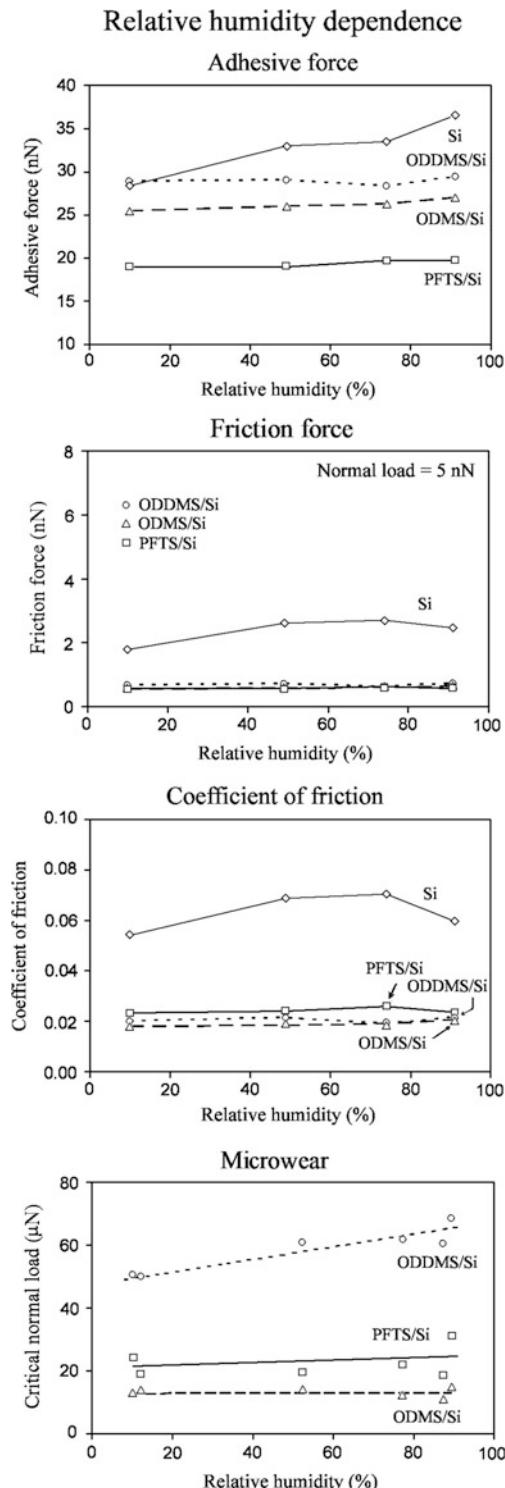


Fig. 16.31 Relative humidity effect on adhesive force, friction force, coefficient of friction, and microwear for various SAMs on Si substrates (Kasai et al. 2005; Bhushan et al. 2006c)

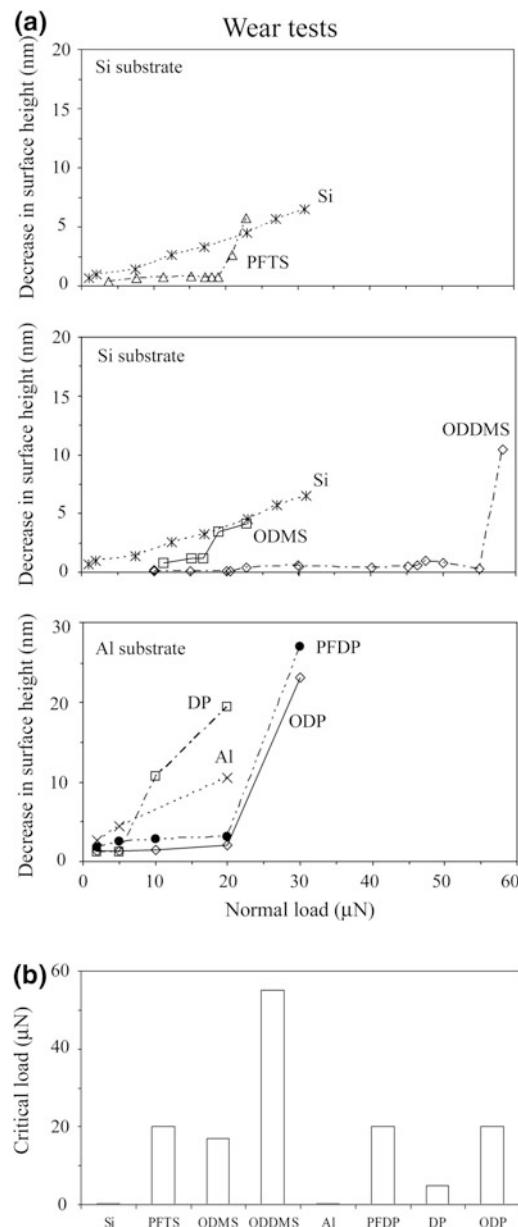


Fig. 16.32 **a** Decrease in surface height as a function of normal load after one scan cycle for various SAMs on Si and Al substrates, and **b** comparison of critical loads for failure during wear tests for various SAMs (Kasai et al. 2005; Bhushan et al. 2006c)

molecule at the head group; that is, by means of shearing of the molecules off the substrate.

To study the effect of relative humidity, wear tests were performed at various humidities. The bottom of Fig. 16.31 shows critical normal load as a function of relative humidity. The critical normal load showed weak dependency on relative humidity for ODMS/Si and PFTS/Si, and was larger for ODMS/Si than PFTS/Si throughout the humidity range. This suggests that water molecules could penetrate into the ODDMS, which might work as lubricant (Liu and Bhushan 2002). This effect was absent for PFTS/Si and ODMS/Si.

Bhushan et al. (2006b) and Lee et al. (2005b) studied various fluoropolymer multilayers and fluorosilane monolayers on Si and a selected fluorosilane on PDMS surfaces. For nanoscale devices such as in nanochannels, monolayers are preferred. They reported that all fluorosilane films increased the contact angle. The fluorosilane monolayer—1H, 1H, 2H, 2H—perfluorodecyltriethoxysilane (PFDTES) resulted in a contact angle of about 100°.

Based on these studies, a perfluoro SAM with a compliant layer should have optimized tribological performance for MEMS/NEMS and BioMEMS/BioNEMS applications.

16.4.3 Hard Diamondlike Carbon (DLC) Coatings

Hard amorphous carbon (a-C), commonly known as diamondlike carbon (DLC) (implying high hardness), coatings are deposited by a variety of deposition techniques, including filtered cathodic arc (FCA), ion beam, electron cyclotron resonance chemical vapor deposition (ECR-CVD), plasma-enhanced chemical vapor deposition (PECVD), and sputtering (Bhushan and Gupta 1991; Bhushan 1999b). These coatings are used in a wide range of applications, including tribological, optical, electronic, and biomedical applications. Ultrathin coatings (3–10 nm thick) are employed to protect against wear and corrosion in magnetic storage applications—thin-film rigid disks, metal evaporated tapes, and thin-film read/write head, Gillette Mach 3 razor blades,

glass windows, and sunglasses. The coatings exhibit low friction, high hardness and wear resistance, chemical inertness to both acids and alkalis, lack of magnetic response, and optical band gap ranging from zero to a few eV, depending upon the deposition technique and its conditions. Selected data on DLC coatings relevant for MEMS/NEMS applications is presented in a later section on microtriboapparatus.

16.5 Nanoscale Friction, Wear, and Mechanical Behavior of Nano-objects

Nano-objects are described as a single material or composite having at least one dimension that is between 1 and 100 nm. They come in a variety of discrete geometries which include spheres, tubes, rods, wires, and pillars. Compared to their bulk material counterparts, many nano-objects exhibit enhanced mechanical, electrical, magnetic, and chemical properties. This makes them attractive for use in many applications on the macro- to nanoscale. Many nano-objects also exhibit low friction and wear properties. These are used in various sliding applications in which they come into sliding contact with each other and the mating surfaces. Contacts can lead to nano-object deformation. For these situations, friction, wear, and nano-object deformation are of interest.

Maharaj and Bhushan (2015a) studied Au NPs and NRs, MoS₂, and WS₂ multi-walled nanotubes (MWNT), and carbon nanohorns (CNH). These nano-objects were chosen since they have either been used in applications (discussed earlier), or have been of interest. They also provide an opportunity to explore the relationship between their friction and wear-reduction ability in relation to their geometry, mechanical properties, and deformation mechanisms under different loading conditions. Figure 16.33 shows TEM images of spherical Au nanoparticles approximately 50 nm in diameter, and Au nanorods 50 nm in diameter and 200 nm in length. Figure 16.34 shows representative TEM images of the MoS₂ and WS₂ MWNTs and

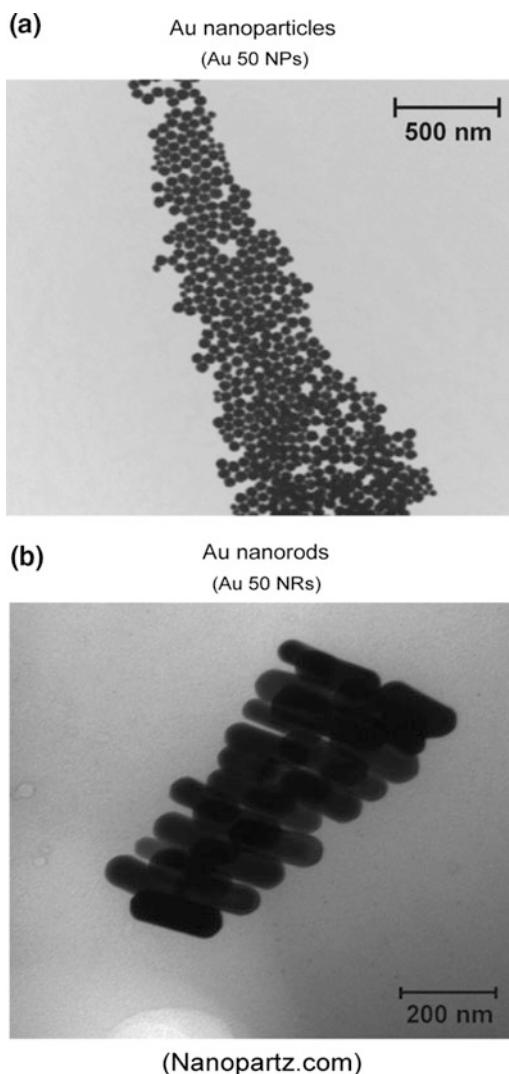


Fig. 16.33 TEM images of **a** spherical Au nanoparticles approximately 50 nm in diameter, and **b** Au nanorods 50 nm in diameter and 200 nm in length ([Nanopartz.com](#))

CNHs. Table 16.6 shows the water contact angle of various thin films. Tables 16.7 and 16.8 show nanomechanical properties of various nanoparticles, MWNTs, films, and bulk materials.

Au nano-objects have been used in applications requiring targeted drug delivery due to their chemical inertness and low to no toxic effects. In tribological studies, their small contact area and mobility due to their spherical shape allows them to roll and slide. This results in reduced friction and wear on the macro- to nanoscale in dry and

submerged-in-liquid environments (Bhushan and Gupta 1991; Moughin et al. 2008; Palacio and Bhushan 2008b; Maharaj and Bhushan 2012; Maharaj et al. 2013; Maharaj and Bhushan 2014b). Au thin films also have been studied to investigate further scale effects on the nanoscale. Their flat surfaces prevent inaccuracies in measuring mechanical properties due to curvature effects of increasingly smaller nano-objects.

MoS₂ and WS₂ have a layered lattice structure with weak bonding between layers which provide low friction and these are known solid lubricants (Bhushan and Gupta 1991; Bhushan 2013a, b). MoS₂ and WS₂ MWNTs have also been studied in tribology on the macro- to nanoscale in dry and liquid environments and increasingly, in other applications such as catalysis in HER. MoS₂ and WS₂ MWNTs have been shown to reduce friction and wear similar to Au nano-objects. In addition, the exfoliation of the outer layers of the nanotubes and subsequent formation of an easily shearable film aids in friction and wear reduction (Bhushan and Gupta 1991; Maharaj and Bhushan 2012, 2013).

A single CNH has a structure similar to a single-walled carbon nanotube with one end closed in a cone-shaped cap (horn). These horns agglomerate through van der Waals attractions to form a spherical nano-object with unique morphology. CNHs have been the subject of myriad of studies due to their unique surface morphology and wide ranging properties of graphene, which includes chemical stability, low surface energy (Kim et al. 2011), high thermal (Sani et al. 2010) and electrical conductivity (Geim and Novoselov 2007), and low friction and wear (Maharaj et al. 2013).

For nano-objects submerged in liquid environments, water, dodecane, and glycerol have been used (Maharaj and Bhushan 2015a). They have been utilized in friction and wear studies as they are simple liquids without additives, and their properties are well known.

Maharaj and Bhushan (2015a) carried out single nano-object contact and multiple nano-object contact studies. In single nano-object contact, Au NPs and NRs, MoS₂ and WS₂ MWNTs, and CNHs have been studied with an

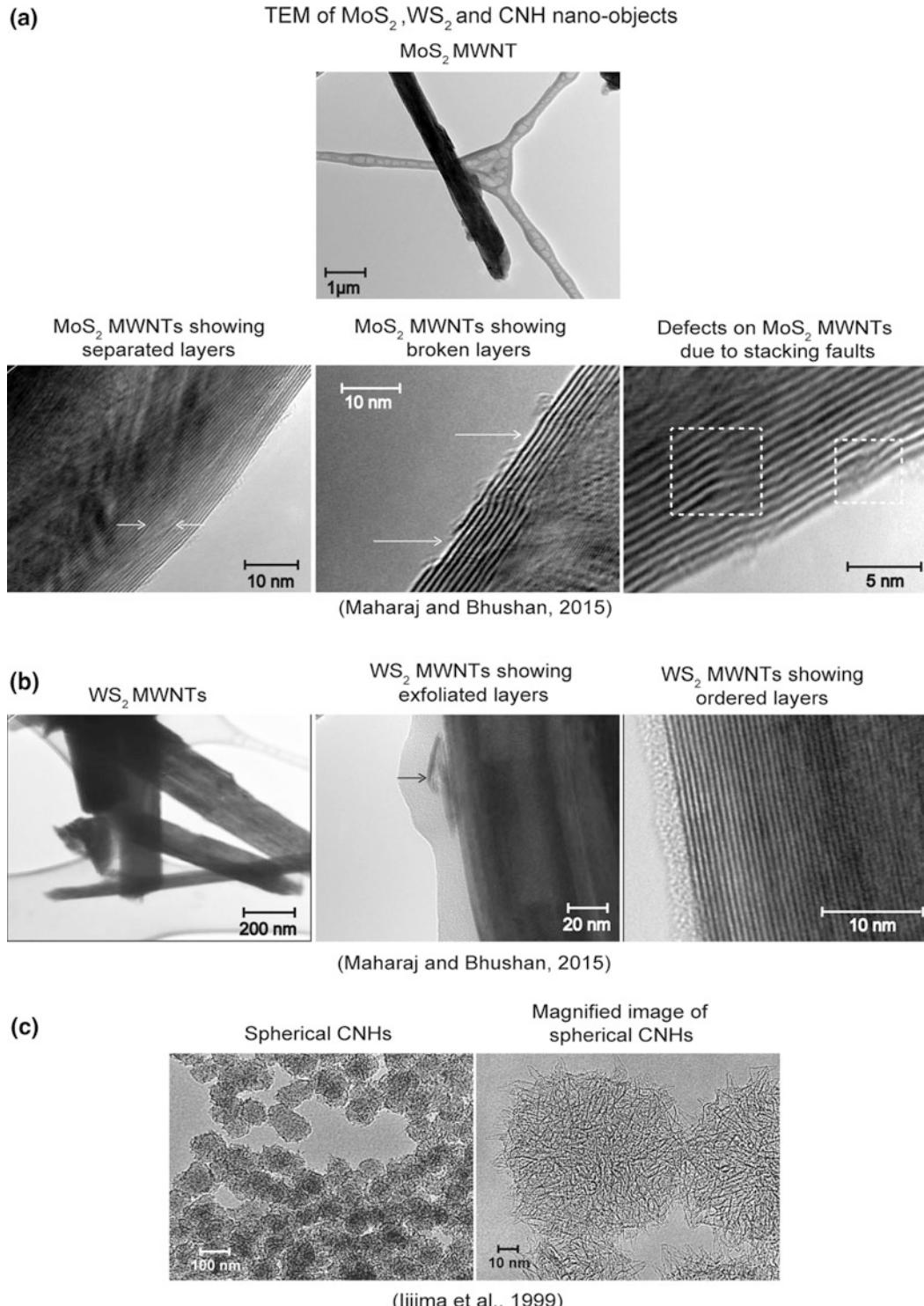


Fig. 16.34 TEM images of **a** MoS_2 MWNT (top), a magnified image showing separated layers between the arrows (left), magnified image showing broken surface layers (center), and a magnified image of a nanotube showing stacking faults highlighted by the dashed lines (right); **b** agglomerated WS_2 MWNTs (left), a magnified image with an arrow pointing to exfoliated layers (center), and a magnified image of a nanotube showing ordered layers (Maharaj and Bhushan 2015b); **c** spherical CNHs (left) with magnified image of CNHs (center), and magnified image showing individual horns (Iijima et al. 1999)

Table 16.6 Water contact angle for thin films of Au, MoS₂, WS₂, and graphene in contact with water (Maharaj et al. 2013)

Films	Water contact angle (°)
Au	50 ^a
MoS ₂	75 ^b
WS ₂	93 ^c
Graphene	127 ^d

^aCayre and Paunov (2004)^bFuerstenaau and Han (2003)^cZhang et al. (2002)^dWang et al. (2009)**Table 16.7** Nanomechanical properties of Au nanoparticles and thin film at a maximum load of 80 μN compared to bulk Au (Maharaj and Bhushan 2014a)

	Nanoparticle diameter/film thickness	Contact depth (nm)	Hardness (GPa)	Young's modulus of elasticity (GPa)	Explanation for enhanced hardness
Au 500 ^a	~500 nm	45 ± 4	0.7 ± 0.1	70 ± 11	Hall-petch, strain gradient plasticity
Au thin film (100 nm) ^a	~100 nm	15 ± 1	1 ± 0.1	76 ± 11	Hall-petch, strain gradient plasticity
Bulk Au ^b	5 mm		0.22	78	

^aMaharaj and Bhushan (2014a)^bSamsonov (1968)**Table 16.8** Nanomechanical properties of MoS₂ MWNTs obtained from the nanoindentation experiments normal to the basal plane

Mechanical properties	MoS ₂ MWNTs ^a (500 nm)	MoS ₂ ^{b, c} (2 mm)	Graphene ^{d, e} (single layer)
Hardness (GPa)	0.6 ± 0.2	0.6	950
Young's modulus (GPa)	12.5 ± 4	52	1000

For comparison, mechanical properties of naturally occurring MoS₂ (normal to the basal planes) and single layer graphene are given (Maharaj and Bhushan 2015b)

^aMaharaj and Bhushan (2015b)^bBhushan and Gupta (1991)^cFeldman (1976)^dLee et al. (2008)^eZhang and Pan (2012)

AFM as illustrated in Fig. 16.35 where they are manipulated by lateral pushing. Friction forces were examined in dry and submerged-in-liquid environments. Friction was found to be dependent on size, shape, material, and roughness. Friction was found to be lower in submerged-in-liquid environments compared to dry conditions. These studies provide an understanding of the nature of the friction mechanism of nano-objects at the initiation of sliding.

In multiple nano-object contact studies, the effect of nano-objects on friction and wear reduction was determined. Sliding takes place over several nano-objects, as illustrated in Fig. 16.35. Real area of contact, rolling, sliding, and exfoliation and tribofilm formation were studied for their effect on friction and wear. Results of studies on the macroscale were compared to the nanoscale to understand scale effects of friction.

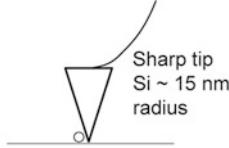
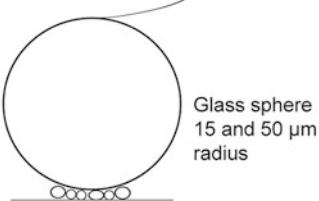
Studies	Observations
Single nano-object contact (Au NPs/NRs, MoS ₂ /WS ₂ MWNTs CNHs) 	Lower friction  <ul style="list-style-type: none"> • Reduced real area of contact of nano-objects • Lower work of adhesion of nano-objects • At submerged in liquid environments compared to dry • At lower viscosity of liquids
Multiple nano-object contact (Au NPs/NRs, MoS ₂ /WS ₂ MWNTs and CNHs) 	Lower friction and wear  <ul style="list-style-type: none"> • Reduced real area of contact between sliding surfaces • Rolling/ Sliding • Exfoliation and tribofilm formation

Fig. 16.35 Illustration of the studies in friction and wear of nano-objects. The studies examined were single and multiple nano-object contact in dry and liquid environments. Observations for each study are given in the column on the right

An overview of friction and wear in dry and liquid environments is presented next, and various mechanisms are presented to explain the friction and wear behavior of nano-objects.

16.5.1 Single Nano-object Friction

For nano-objects that are spherical in nature, manipulation can involve rolling, sliding, and rotation about the vertical axis. Manipulation can also involve rolling, sliding, and in-plane rotation in the case of the MoS₂ and WS₂ MWNTs. In-plane rotation can occur when the nanotubes are not pushed directly at the center of its length.

Figure 16.36a presents the results for the average friction forces for Au 90 NPs, CNHs and MoS₂ and WS₂ MWNTs in dry and submerged in liquid environments. The friction force presented is the result of adhesion between the nano-objects and the silicon substrate. The adhesive force can include both van der Waals

and meniscus forces for dry environments. However, under submerged-in-liquid environments, meniscus forces are eliminated.

Friction forces for the CNHs and MoS₂ and WS₂ MWNTs compared to Au 90 NPs are lower due to a material effect. For the MWNTs it is not believed that exfoliation and tribofilm formation are responsible for the lower values, as discussed by Maharaj et al. (2013), since there is no normal load acting on them.

Figure 16.36b presents the friction forces for the nanoparticles and nanotubes submerged in liquid environments. For Au NPs which were already suspended in DI water, experiments were not carried out in dodecane and glycerol. In dodecane and glycerol, similar to water, adhesive forces are due to van der Waals interactions since meniscus forces are eliminated. The lower friction forces in submerged-in-liquid environments compared to the dry environments can be attributed to both elimination of meniscus forces and nano-objects sliding on a low shear strength

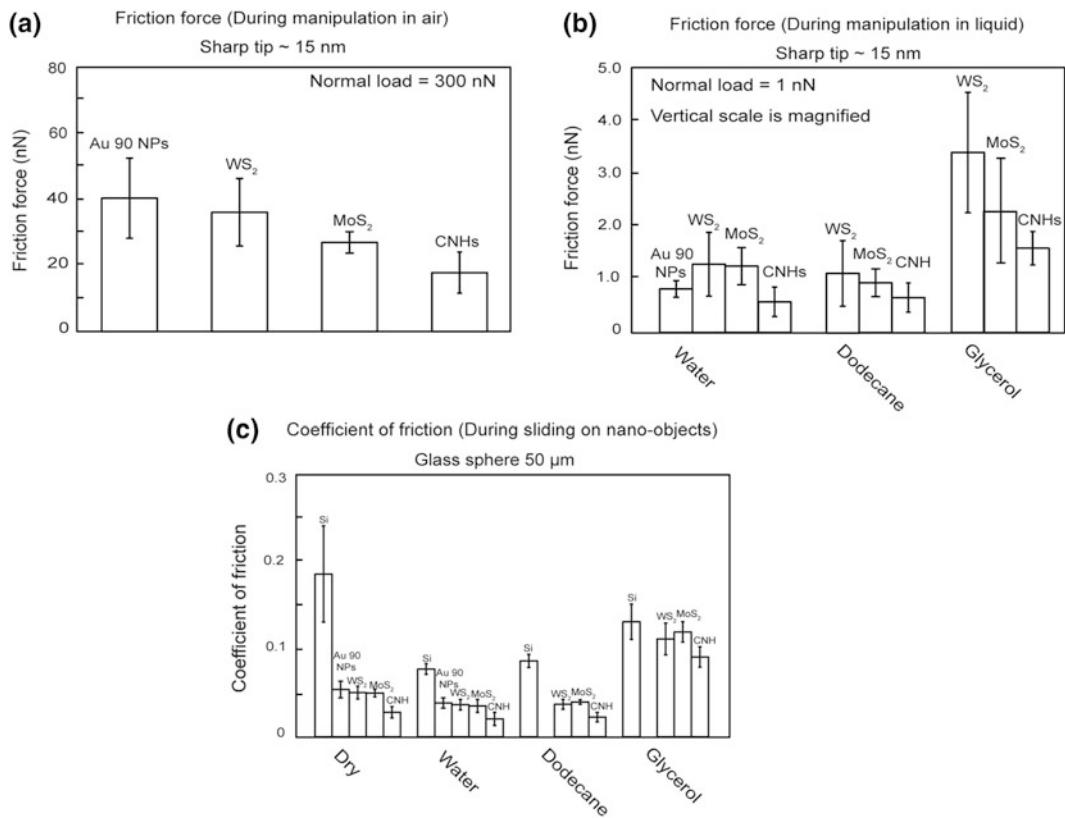


Fig. 16.36 Friction force for Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs on the silicon substrate during manipulation, at normal loads of **a** 300 nN in air and **b** 1 nN in liquids. The *vertical scale* is magnified for data in liquids. **c** Coefficients of friction for both dry and submerged-in-liquid environments, with and without nano-objects. Friction forces and coefficient of friction are lower in submerged-in-water environments (Maharaj et al. 2013)

surface. The lowest friction forces occur with the CNHs due to the low contact area provided by the roughness of the nanohorns. Here, the material effects due to work of adhesion does not contribute to friction. The highest friction force occurs in glycerol, which has the highest viscosity and therefore the greatest contribution to viscous drag force. Similar viscosity of dodecane and water accounts for the similarity in friction

forces observed for both liquids with and without nano-objects.

The characteristics of the nano-objects discussed in single nano-object contact studies will also be investigated in multiple nano-object contact studies in the context of reducing friction and wear under load. As will be discussed in the next section, new nano-object behavior not previously seen will contribute to

friction and wear reduction due to the application of a load.

16.5.2 Multiple Nano-object Contact

16.5.2.1 Nanoscale Friction and Wear

In multiple nano-object contact, the effect of normal load acting on the Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs between two surfaces is studied to determine the effects on friction and wear. Figure 16.36c summarizes the coefficient of friction for dry and submerged-in-liquid environments. Sliding in multiple nano-object contact results in lower friction forces in all conditions compared to sliding on the bare silicon substrate as discussed by Maharaj et al. (2013). For Au nano-objects that were already suspended in DI water, experiments were not

carried out in dodecane and glycerol. In general, the friction forces were lower for sliding in submerged-in-liquid environments as compared to sliding in dry environments for nano-object coated and uncoated surfaces.

Sliding on multiple asperities on nanopatterned surfaces resulted in lower friction. With the immobile asperities friction reduction occurs as a result of the reduced contact area. For Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs, the friction force reduction can be attributed to nano-object mobility in addition to the reduced contact area. The resulting friction reduction mechanism can be attributed to the reduced contact area, the sliding over deformed nano-objects, and the individual nano-objects being dragged with the glass sphere as shown in Fig. 16.37b–c. It is also possible that as the glass sphere encounters a greater number of nano-

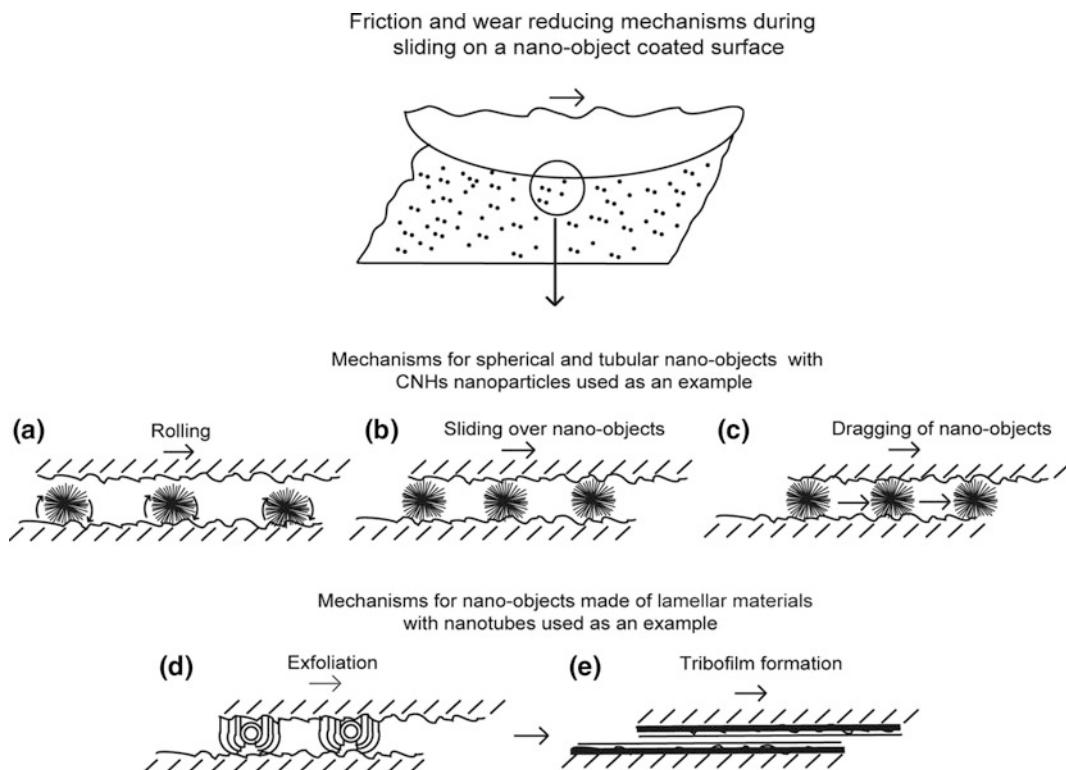


Fig. 16.37 Schematic of friction and wear reducing mechanism of nano-object coated surface by **a** rolling, **b** sliding over nano-objects, **c** dragging of nano-objects, **d** exfoliation, and **e** tribofilm formation. CNHs are used as an example for mechanisms (a) through (c) and WS₂ and MoS₂ MWNTs are used as an example for mechanisms (d) through (e) (Maharaj et al. 2013)

objects, the contact pressure is reduced, allowing undeformed nano-objects to roll between the surfaces as shown in Fig. 16.37a. Exfoliation and tribofilm formation as shown in Fig. 16.37d–e are believed to be possible in dry environments for MoS₂ and WS₂ MWNTs as there is a general trend towards lower friction compared to Au 90. However this is not very significant due to the limited number of MWNTs that would be able to fit under the 50 μm sphere. In dry environments, the coefficient of friction data for the CNHs and the MoS₂ and WS₂ MWNTs are generally lower than the Au 90 NPs. As previously discussed, with the Au 90 NPs there is a higher meniscus force contribution to friction compared to the CNHs and MoS₂ and WS₂ MWNTs. With the CNHs, the reduced contact area due to the roughness of the nanohorns and low meniscus force contribution are responsible for the low friction forces (Maharaj et al. 2013).

For nano-objects submerged in liquid environments, the coefficients of friction were lower compared to dry. For water, as the meniscus forces are eliminated, the coefficient of friction data shows very little difference for Au 90 NPs and MoS₂ and WS₂ MWNTs. The CNHs have the lowest coefficient of friction in all liquid environments due to the nanohorn roughness contributing to lower real area of contact. Similar to single nano-object contact, the highest friction force and coefficient of friction occurs in glycerol, which has the highest viscosity and results in a greater contribution from viscous drag force followed by dodecane. The elimination of the meniscus forces, coupled with sliding on a surface of low shear strength results in lower coefficients of friction (Bhushan and Sundararajan 1998).

The nano-objects were also investigated in wear conditions to determine their effectiveness in protecting the underlying surface which is crucial feature of a potential lubricant. Figure 16.38 summarizes the wear data for sliding on Si, and Si coated with Au 90, CNHs, MoS₂ and WS₂ MWNTs, for 1, 10, and 100 cycles at a normal load of 20 μN under dry environments. A 20 μm × 20 μm area is imaged to show wear scars, which are created over a 10 μm × 10 μm

area. Topography maps along with corresponding height profiles are also shown.

The surface of the uncoated silicon, after 1 cycle, appears rough during the early stages of wear as it becomes damaged. For the Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs coated surfaces, the nano-objects are just beginning to be pushed out of the wear area after 1 cycle. After 10 cycles, the topography profile of the uncoated silicon surface appears smoother than after 1 cycle. This is due to polishing of the surface. The nano-objects are pushed out of the wear area and agglomerate around the edges of the wear scar. After 100 cycles, for the uncoated silicon substrate, a small amount of material is removed, which gives a wear depth of approximately 0.25 nm, as seen in the height profile. For the coated surfaces, a few nano-objects remain in the wear area for MoS₂ and WS₂ MWNTs while they are completely removed for CNHs and Au 90 NPs.

Figure 16.39 shows SEM micrographs of the wear scars in dry environments after 100 cycles (first row). Vertical arrows point to magnified micrographs of the areas within the squares are shown in the second row. Agglomerated Au 90 NPs, CNHs, and broken MoS₂ and WS₂ MWNTs are pointed out by arrows within each micrograph. Agglomeration and breaking of the nano-objects can occur during the wear process due to the high loads being applied. In sample preparation of MoS₂ and WS₂ MWNTs, broken nanotubes also can result during sonication as they collide with one another. Since there is no evidence for crushed nanotubes from the SEM micrographs, it is likely that exfoliation of MoS₂ and WS₂ MWNTs would be limited. Protection of the substrate from wear due to tribofilm formation is not expected to be significant.

After 1 cycle, the damage of the silicon surface is expected to be less than that of an initially uncoated substrate since the Au 90 NPs, CNHs and MoS₂ and WS₂ MWNTs remain in the wear area. In addition to the reduced contact area provided by the nano-objects it is believed that they roll and slide between the glass sphere and the substrate which promotes facile shearing of

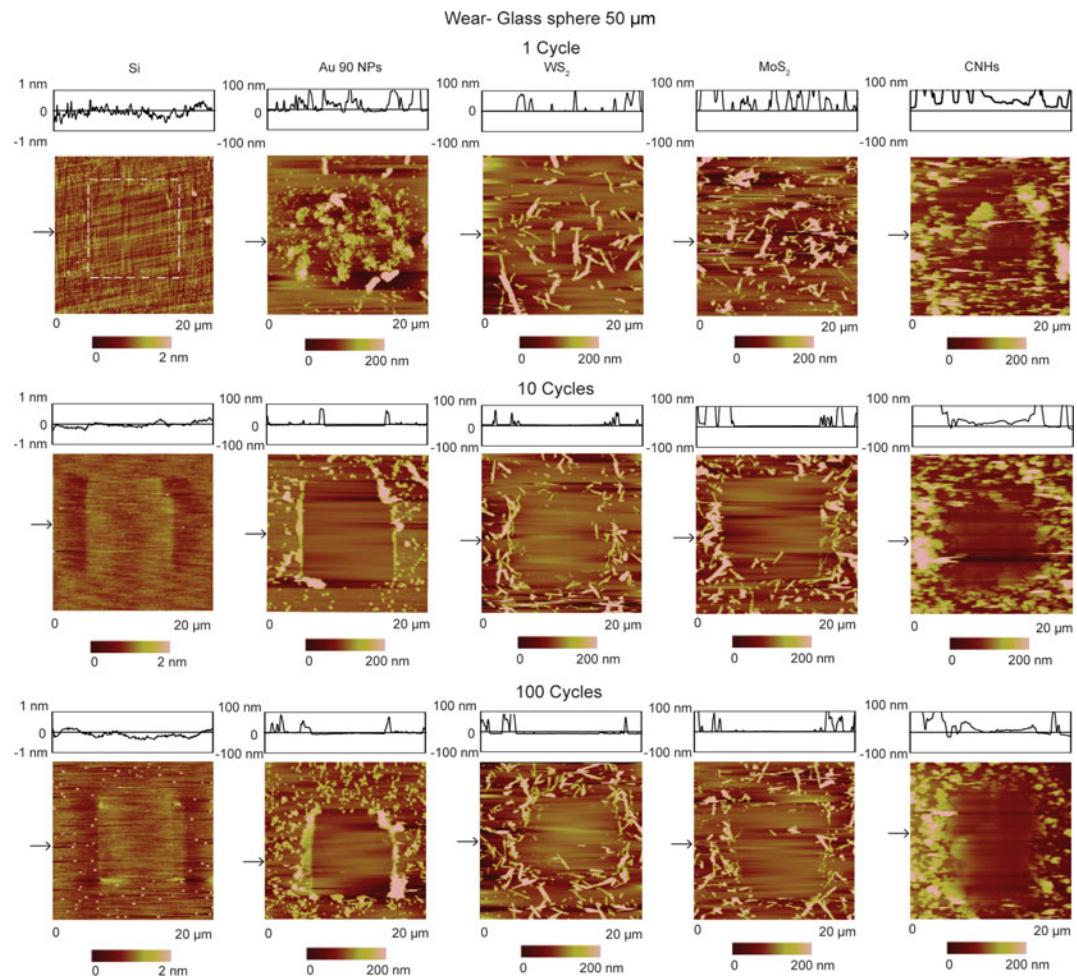


Fig. 16.38 Topography maps and 2-D profiles, at sections shown by the arrows, after sliding at 1, 10 and 100 cycles with a normal load of $20 \mu\text{N}$ on Si and Si coated with Au 90 NPs, CNHs, and MoS_2 and WS_2 nanotubes (Maharaj et al. 2013)

the two surfaces, as shown in Fig. 16.37a–c. Further reductions in wear are expected for the CNHs due to roughness, which results in reduced real area of contact. It is expected that the greatest wear occurs on the bare silicon substrate and the presence of nano-objects provides protection of the substrate with the least wear occurring with the CNHs after 100 cycles.

16.5.2.2 Macroscale Friction and Wear

In macroscale friction and wear tests, a ball on flat tribometer was used. A sapphire ball of 1.5 mm radius was slid over nano-object coated and uncoated silicon substrates under a normal

load of 200 mN for 500 cycles with a stroke length of 10 mm. Figure 16.40 shows optical micrographs of wear scars in dry and liquid environments with and without nano-objects. The widths of the wear scars shown are smaller for sliding in liquid environments compared to sliding in dry environments, as the amount of wear is less. The presence of nano-objects in dry and liquid environments reduces the amount of wear compared to cases without nano-objects.

In liquid environments, as the viscosity of the fluid increases, the width of the wear scar decreases, with the greatest wear protection

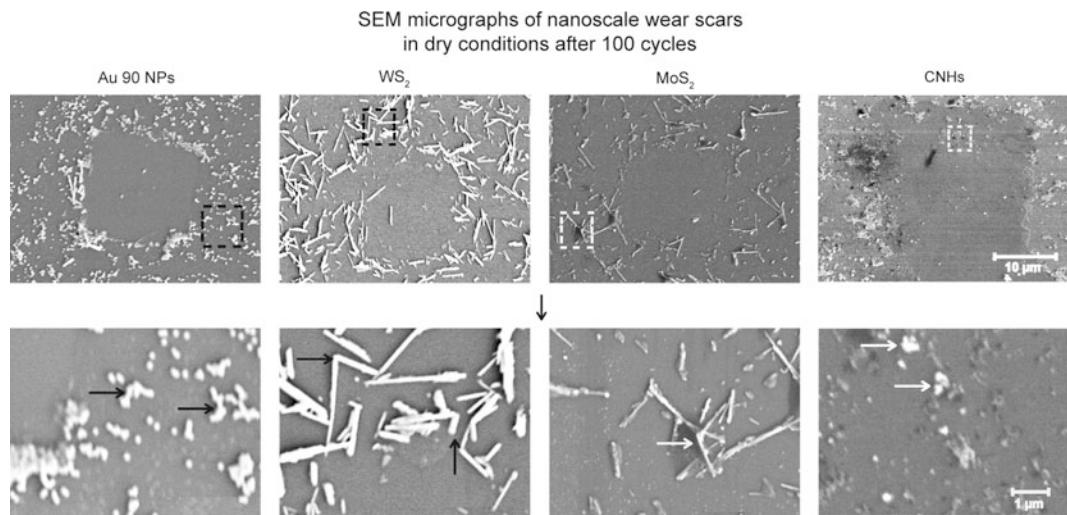


Fig. 16.39 SEM micrographs of nanoscale wear scars for Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs in dry conditions after 100 cycles at a normal load of 20 μN (first row). The nano-objects within the squares are magnified and the vertical arrows point to those micrographs in the second row. The magnified micrographs show agglomerated Au 90 NPs and CNHs, and broken MoS₂ and WS₂ MWNTs as indicated by arrows within the micrographs (Maharaj et al. 2013)

occurring in glycerol. This is due to the ability of the viscous fluid to support the normal load and not be squeezed out of the contact area.

Optical and SEM micrographs of the wear scars for all cases are shown in Fig. 16.41. Nano-objects can be observed within the wear scars for all cases, as seen in Fig. 16.41a. The presence of the nano-objects after the wear test further confirms that reduction in wear and size of the wear scars are due to addition of the nano-objects. Figure 16.41b shows magnified SEM micrographs of wear scars for Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs in dry environments. Agglomerations of nanoparticles within the squares can be observed in the case of the Au 90 NP and CNH wear scar.

Agglomeration occurs during the wear process as nanoparticles are pressed together and are attracted to each other due to van der Waals and attractive meniscus forces. For both MoS₂ and WS₂ wear scars, the SEM micrographs show crushed and broken nanotubes within the squares compared to those in Fig. 16.34a–b, which depicts undeformed nanotubes. It is believed that in the process of being crushed,

the outer layers of the nanotubes break apart leading to exfoliation of the nanotubes and tribofilm formation, as discussed by Maharaj et al. (2013).

In liquid environments, the results are similar to the nanoscale, where the coefficients of friction are lower than in dry environments in general, as reported by Maharaj et al. (2013). This is due to the presence of a surface of low shear strength for sliding. For CNHs submerged in water, the coefficient of friction is the lowest of all four nano-objects due to the reduced contact area provided by the roughness of the nanohorns. For MoS₂ and WS₂ MWNTs after 200 cycles, the coefficient of friction increases and becomes greater than that of the other nano-objects, and reaches that of bare silicon. The increase in friction is due to the oxidation of the MoS₂ and WS₂ MWNTs, as previously discussed. For sliding in dodecane and glycerol, the coefficient of friction is lower sliding on a nano-object coated surface compared to bare silicon. The coefficient of friction is lower for the CNHs compared to the MWNTs, similar to the submerged-in-water case. After approximately 400 cycles for sliding on nanoparticles and

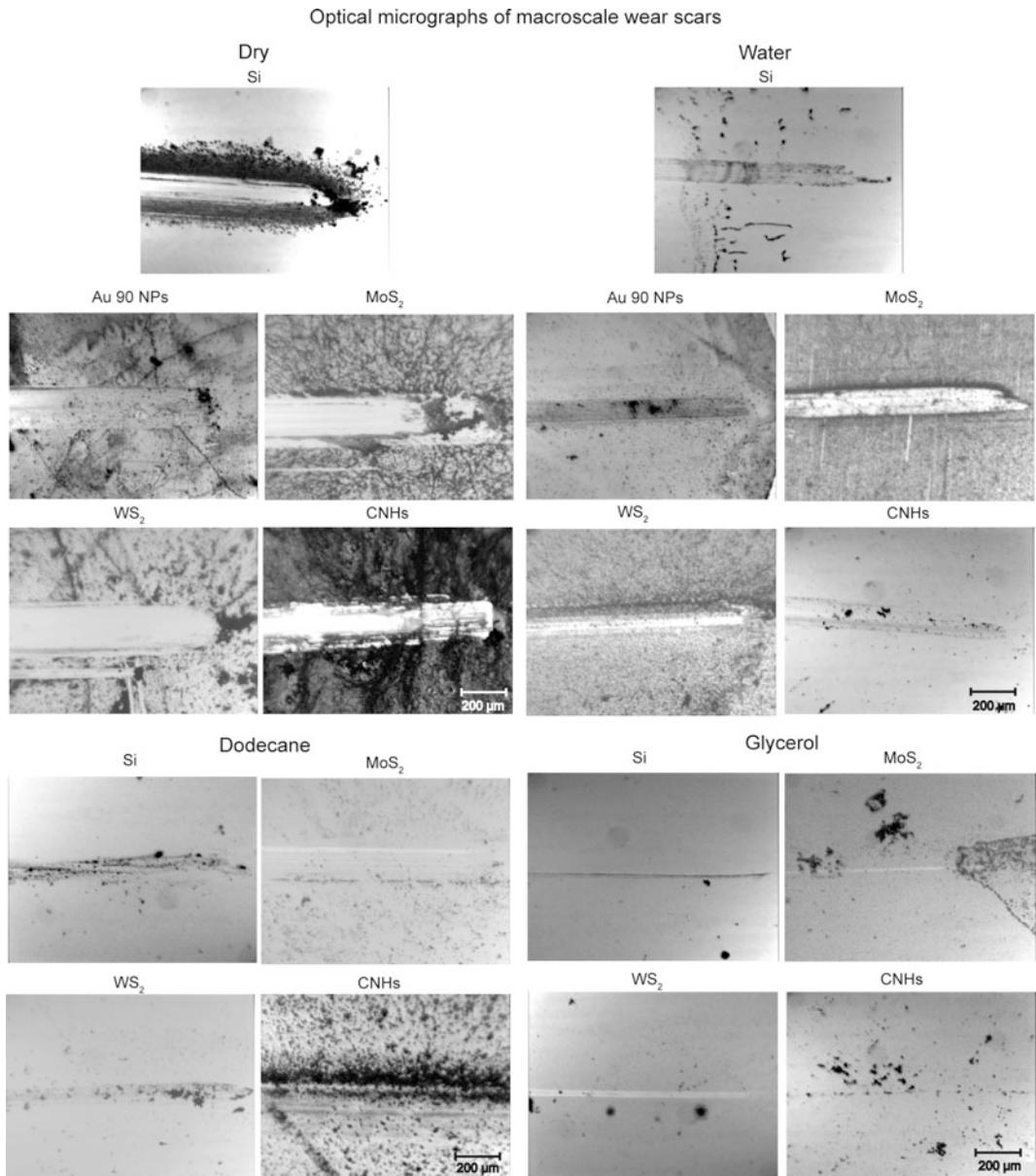


Fig. 16.40 Optical micrographs of macroscale wear scars taken after 500 cycles at a normal load of 200 mN, with 1.5 mm radius sapphire ball sliding in dry, water, dodecane and glycerol conditions with and without Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs (Maharaj et al. 2013). The lowest wear occurs in glycerol with CNHs

nanotubes in dodecane, the coefficient of friction begins to increase as wear of the substrate takes place. In the case of glycerol there is also an increase in coefficient of friction for the nanotubes after 400 cycles. However, the coefficient remains steady for the CNHs. This is due to the

combined effect of the reduced real area of contact for CNHs with the ability of glycerol to better support the load due to its high viscosity, which reduces the wear. The lowest coefficient of friction and least wear takes place in glycerol with CNHs.

Optical micrographs of macroscale wear scars

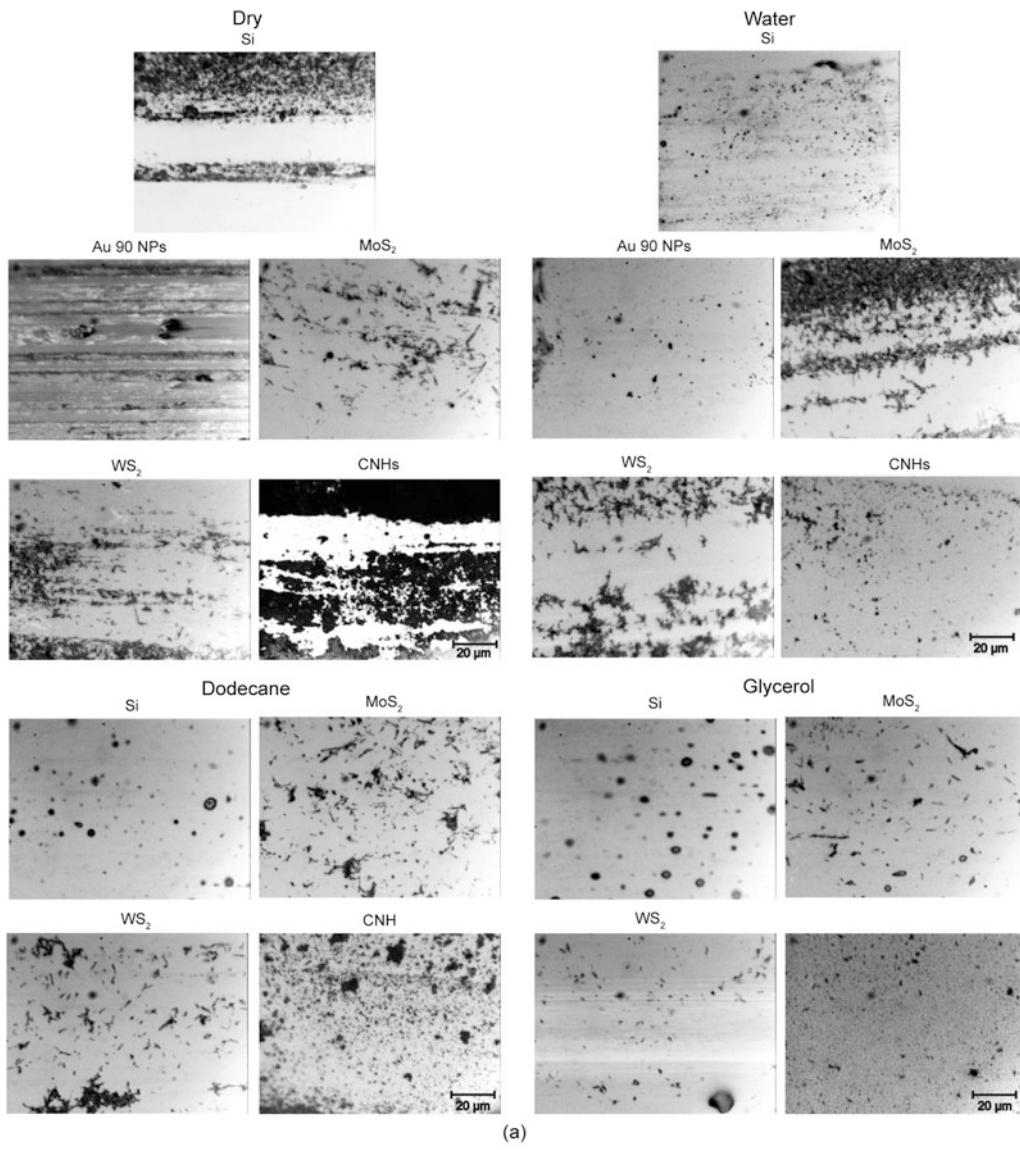
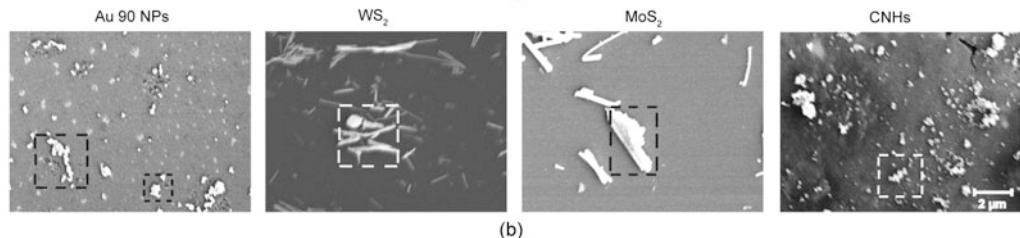
SEM micrographs of macroscale
wear scars in dry conditions

Fig. 16.41 **a** Optical micrographs inside the wear scar on the macroscale taken after 500 cycles at a normal load of 200 mN, in dry, water, dodecane, and glycerol conditions with and without Au 90 NPs, CNHs, and MoS₂ and WS₂ MWNTs, and **b** SEM micrographs of the wear scar for agglomerated Au 90 NPs and CNHs, along with crushed or broken MoS₂ and WS₂ MWNTs in dry environments shown within the squares (Maharaj et al. 2013)

16.5.3 Summary

Au NPs, MoS₂ and WS₂ MWNTs, and CNHs were investigated to understand friction and wear mechanisms of nano-objects and their effect on friction and wear in dry and liquid environments in water, dodecane, and glycerol. Studies were conducted in both single and multiple nano-object contacts with the aid of an AFM. In macroscale studies, a ball on flat tribometer was used.

For single nano-object contact, the lowest friction forces occur with the CNHs compared to Au NPs and MoS₂ and WS₂ MWNTs due to the reduced contact area provided by the roughness of the nanohorns and the low meniscus force contribution of the dry environments. Lower friction forces occur in submerged-in-liquid environments compared to dry environments due to elimination of meniscus force and sliding on a low shear strength surface among the various liquids. The highest friction forces occur in glycerol due to high viscosity.

For multiple nano-object contact, sliding over nano-objects decreased the coefficients of friction compared to sliding on the bare silicon substrate due to reduced contact area and rolling and sliding among the various nano-objects. The lowest coefficient of friction occurs with the CNHs compared to the other nano-objects due to the mechanisms mentioned for single nano-object contact in dry environments. In nanoscale wear experiments in dry environments, the addition of nano-objects reduces the real area of contact and reduces the wear of the substrate through rolling and sliding of the nano-objects.

In macroscale studies, sliding over the various nano-objects reduced friction and wear due to the mechanisms mentioned in multiple nano-object contact on the nanoscale. Additionally, in dry environments, because of exfoliation and tribofilm formation, sliding on the MWNTs decreased the coefficient of friction and reduced wear. The lowest coefficient of friction and least wear in all cases occur with CNHs due to the reduced contact area. The coefficients of friction were also lower in submerged-in-liquid environments compared to dry condition due to

mechanisms mentioned on the nanoscale. The more viscous liquids results in better wear protection and lower coefficient of friction due to their ability to support the load and not be easily squeezed out of the contact area.

16.6 Nanotribological Studies of Biological Molecules on Polystyrene and Silicon Surfaces and Coated Polymer Surfaces

Protein layers are routinely deployed on bioMEMS/NEMS as a functional layer allowing for specific molecular recognition, binding properties, or to facilitate biocompatibility. Examples of these include sensors, immunoisolation capsules, and drug delivery systems (Park 1997; Grayson et al. 2004; Bhushan 2016a). In addition, coating material surfaces with a protein layer facilitates biocompatibility along with implant functionality (Black 1999; Wise 2000). Both the implanted biomaterial and the layer of protein will experience a range of physical and chemical assaults over the life of the device (Ratner et al. 2004). These include degradation by serum or tissue proteases, biochemical dissociation from surfaces, chemical modification, and abrasion by tissue elements (Turbill et al. 1996). In addition, an implanted biomaterial will experience micromotion over time, leading to wear and degradation of the protein layer (Bhushan et al. 2005a, 2006a). These can significantly affect and degrade the performance of the biomaterial, eventually leading to failure (Ratner et al. 2004). Optimizing the adhesion, friction, and wear on the adsorbed protein layer therefore is crucial to prevent device failure.

Protein adsorption, conformation, and adhesion also are of interest in cellular adhesion and growth. Schricker et al. (2011, 2012) and Bhushan and Schricker (2014) studied protein adhesion on block copolymer-based biomaterials used as a scaffold for dental bone generation.

Bioadhesion refers to the adherence of biological material to a natural or synthetic material.

The factors that control bioadhesion or cell adhesion and wear include surface morphology, chemical interactions, physiological factors, physical-mechanical interactions, and fluid environment (Palacio and Bhushan 2012; Bhushan and Utter 2013). Bioadhesion for a given biological material to be bonded to an engineering surface and its wear properties need to be optimized.

In some applications, surfaces with low bioadhesion are required. Fluorinated self-assembled layers have been used to reduce bioadhesion of polymer surfaces.

16.6.1 Nanoscale Adhesion, Friction, and Wear of Protein Layers on Polystyrene Surfaces

Bhushan and Utter (2013) studied nanoscale adhesion, friction, and wear behavior of four common proteins—bovine serum albumin (BSA), collagen, fibronectin, and streptavidin (STA)—adsorbed to a spin coated polystyrene surface using atomic force microscopy. Bovine serum albumin is a homolog of human serum albumin and its primary physiological functions include helping to maintain blood oncotic pressure, the binding and transport of important ligands, and several antioxidant functions (Francis 2010). Type I collagen is the most abundant protein in the human body and is the primary load bearing component of connective tissue (Di Lullo et al. 2002). Fibronectin is a major extracellular component required for clot formation and wound healing found throughout the body (Pankov and Yamada 2002). Finally, STA is used extensively in the biotechnology field for protein purification, biosensors, and the immobilization of biomolecules due to its interaction with biotin (Green 1990). Figure 16.42 shows protein structures. Bovine serum albumin, fibronectin and streptavidin are all globular proteins (Bhushan and Utter 2013). Collagen is the only protein used in this experiment with a linear structure. Table 16.9 compares and contrasts the masses and isoelectric points of the four proteins

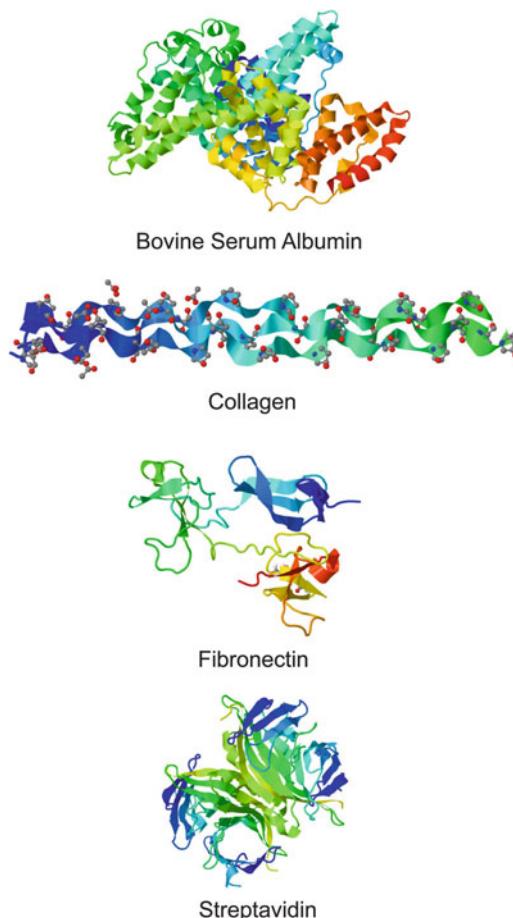


Fig. 16.42 Images showing the physical structure of the bovine serum albumin, collagen, fibronectin, and streptavidin. Proteins are not shown to scale relative to one another. All figures based on the protein data bank's files (Bhushan and Utter 2013)

used (Bhushan and Utter 2013). The isoelectric point is the pH at which the protein will be electrically neutral in solution. Changing the pH of the solution above or below the isoelectric point will change the net charge on each protein. All of the proteins used in this experiment have an acidic isoelectric point giving them a net negative charge in DI water. STA and BSA have masses of 60 and 67 kDa, respectively, with fibronectin and collagen being an order of magnitude more massive.

Polystyrene was selected as the substrate surface due to its extensive use in the biotechnology field. 50 µl of protein solution were

Table 16.9 Summary of the different protein masses and their isoelectric point (Bhushan and Utter 2013)

Protein	Molecular weight (kDa)	Isoelectric point
Bovine serum albumin	~67 ^a	4.7 ^e
Collagen	~300 ^b	4.7 ^f
Fibronectin	~500 ^c	5 ^g
Streptavidin	~60 ^d	5.5 ^d

^aHirayama et al. (1990)

^bDi Lullo et al. (2002)

^cPankov and Yamada (2002)

^dGreen (1990)

^ePeters (1980)

^fHattori et al. (1999)

^gProctor (1987)

added to the surface, and the samples were allowed to sit overnight. Before beginning any experiment, the samples were rinsed with DI water to remove the electrolytes and any unbound protein. Phosphate buffered saline (PBS) (pH 7.4) and the two remaining pH levels were created by combining monosodium phosphate (Aldrich) and disodium phosphate (Aldrich) in DI water to create the proper phosphate buffered solution.

Nanoscale adhesion, friction, and wear measurements were made using an AFM in DI and PBS (pH 5.0, 7.4 and 9.0). Adhesion and friction were investigated for all four proteins by varying the concentration of the protein solution from 1 to 100 µg/mL in DI water and PBS (pH 7.4). Nanoscale wear was also studied at a protein concentration of 100 µg/mL in PBS (pH 7.4). In addition, adhesion and friction were investigated by varying the pH and temperature of the liquid medium for both BSA and STA at 100 µg/mL.

16.6.1.1 Nanoscale Adhesion Measurements

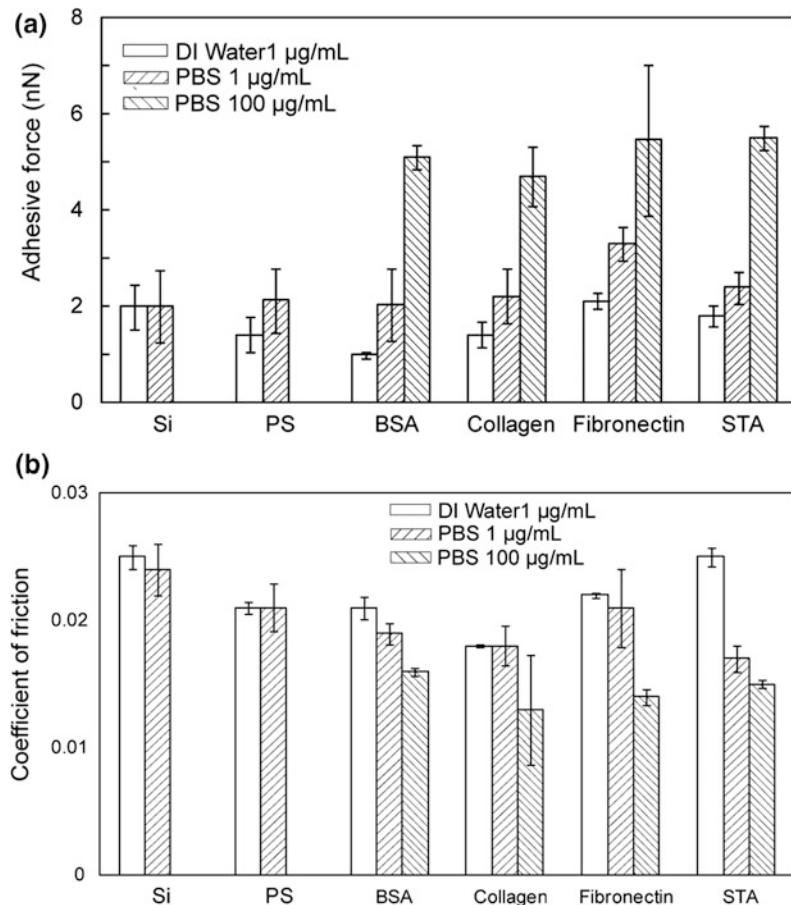
First, the effect of protein concentration on the adhesive force was investigated for all the proteins. Figure 16.43a shows the adhesive force data between the bare silicon nitride AFM tip and each of the various proteins. The adhesive force increases between the protein film and the bare silicon nitride AFM tip as the concentration of protein is increased (Bhushan and Utter 2013). This increase in adhesive force is believed in part to be due to multiple interactions occurring

between the protein film and the AFM tip (Bhushan et al. 2006a). This is also consistent with the observation that increasing the concentration of protein solution increases the density of the immobilized protein on the polystyrene surface. This increase may also be due to the more hydrophobic domains of the protein interacting with the polystyrene surface, leaving the hydrophilic domains to be exposed to the AFM tip. Other studies have shown that near maximum binding will occur at roughly 2 h, and that more protein will directly adsorb depending on solution concentration (Browne et al. 2004; Pamula et al. 2004). It has also been reported that above a concentration of 10 µg/ml, certain proteins will begin to form higher order structures not present at lower concentrations (Bhushan et al. 2006a).

Figure 16.44 lists the basic interactions that may occur during protein adsorption and is only meant to show some of the interactions that can occur (Andrade et al. 1992). The strongest interaction occurs between the hydrophobic domains of the protein and a hydrophobic surface. Next, moderate adsorption occurs between charged protein domains and charged surfaces. Finally, weak adsorption occurs between hydrophilic protein domains and hydrophilic surfaces. Protein adsorption may involve several of these interactions occurring on the same surface and with the same protein.

Given that polystyrene is a hydrophobic surface, it is to be expected that the hydrophobic domains of the proteins interact with the surface

Fig. 16.43 Effect of protein concentration on **a** the adhesive force and **b** the coefficient of friction between the protein film and bare AFM tip. All experiments conducted at ambient temperature 22 ± 1 °C unless otherwise noted. DI water (pH 7.0) and PBS (pH 7.4) were used for all experiments unless otherwise noted. Error bars represent ± 1 standard error (Bhushan and Utter 2013)



during adsorption. The strength of the hydrophobic interaction is strong enough to induce a conformational change in the protein exposing the inner hydrophobic domains to the surface. This allows for the hydrophilic and charged domains to interact with solution. It is also well-known that by varying the pH of the solution, proteins can be made to denature from their native structure, thus allowing the inner hydrophobic domains to interact with the polystyrene surface. Previous studies have shown that chemically modifying the polystyrene surface has an effect on the amount of adsorbed protein, suggesting that the four proteins used in this study have a higher affinity for hydrophobic surfaces (Browne et al. 2004; Pamula et al. 2004). This hydrophobic interaction is believed to be the dominant mechanism for protein adsorption to polystyrene. However, we do not

have the data to show the orientation or packing density of the protein films (Browne et al. 2004). The data from this experiment demonstrates no clear trend to distinguish the proteins from one another. However, what is clear for all of the selected proteins is that an increase in protein concentration leads to an increase in adhesive force between the AFM tip and the protein layer.

Next, the effect of pH on BSA and STA at 100 µg/mL concentration was investigated. Figure 16.45a shows the adhesive force data as a function of pH at 22 °C (Bhushan and Utter 2013). The general trend is that the adhesive force increases for each protein as the pH of the solution becomes more basic. Both proteins exhibit higher adhesive forces than the polystyrene control surface over the range of pH values. As the protein layer is exposed to a more basic solution, the surface will have a much larger net negative

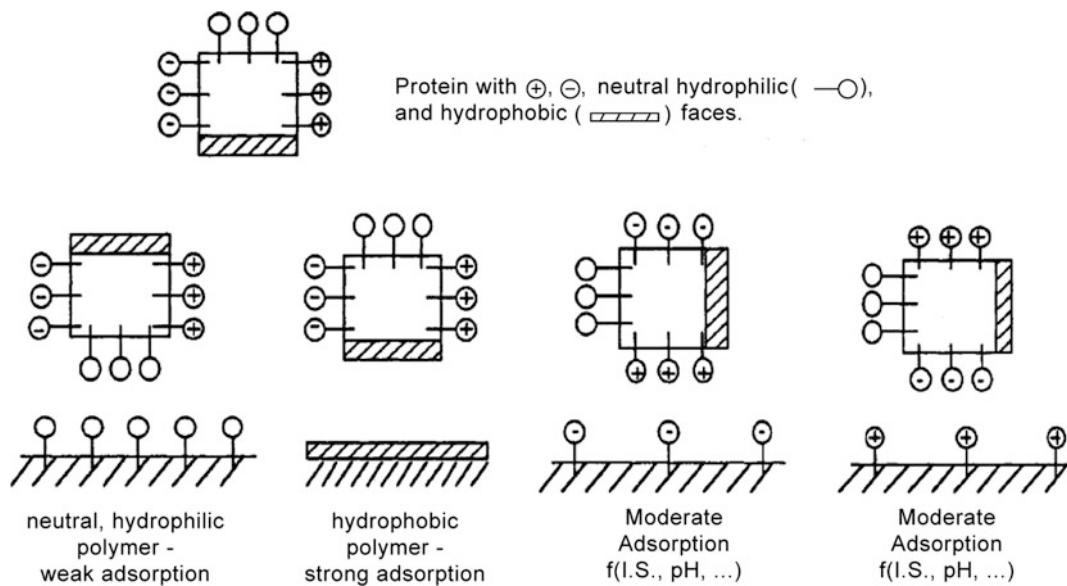


Fig. 16.44 Schematic showing the different protein domains and the potential interactions with the substrate surface (Andrade et al. 1992)

charge. Given that silicon nitride has an isoelectric point of pH 6–7, this may lead to a stronger interaction between the surface and the silicon nitride AFM tip (Kosmulski 2001). For lower pH values (<pH 7.4), STA displays higher adhesive forces than BSA but this trend does not hold for the most basic solution.

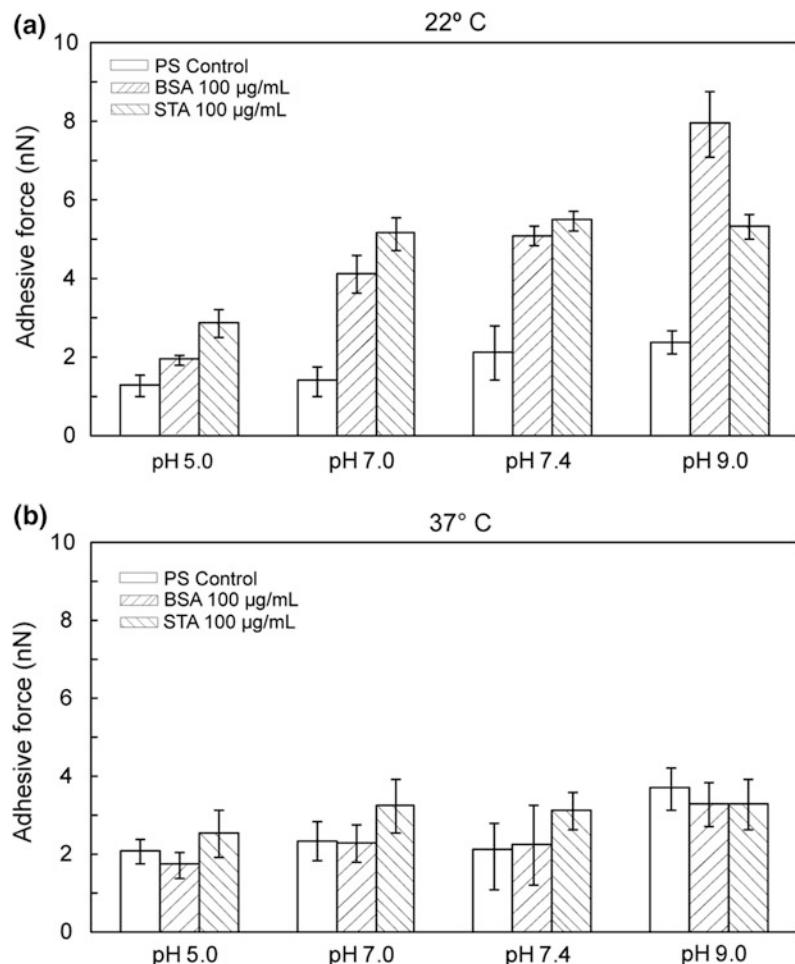
The effect of temperature over the range of pH values also was studied for BSA and STA at 100 $\mu\text{g}/\text{mL}$ concentration. Figure 16.45b shows the adhesive force data as a function of pH at 37 °C (Bhushan and Utter 2013). Adhesive forces for both proteins are much lower than at 22 °C and are not much different than the polystyrene control surface. This decrease in adhesive force at 37 °C may be due in part to the fact that more protein-protein and protein-surface interactions are occurring that result in a lower adhesive force with the AFM tip. The decrease in adhesive force could also be due to a temperature related conformation change in the protein which then interacts more weakly with the AFM tip. However, a slight increase occurs for all samples as the solution becomes more basic. This trend is consistent with the data that was obtained at 22 °C. As before, no clear trend is present to

distinguish BSA from STA. This may be in part due to their similar protein structure, isoelectric point and mass.

16.6.1.2 Nanoscale Friction Measurements

The coefficient of friction was measured for all proteins as a function of protein concentration. Figure 16.43b shows the coefficient of friction between the bare silicon nitride AFM tip and the various proteins (Bhushan and Utter 2013). The friction measurements that were taken show a decreasing coefficient of friction as the protein concentration is increased from 1 to 100 $\mu\text{g}/\text{mL}$. This decrease is believed to be due to the proteins acting as a lubricant film between the polystyrene surface and the silicon nitride AFM tip. At a higher protein concentration, a continuous lubrication film is formed due to a more uniform protein layer being formed on the substrate surface (Bhushan et al. 2006a). Like the adhesion data, there is no clear trend to distinguish the proteins from one another. However, the general trend is that increasing the concentration of these proteins leads to a decrease in the coefficient of friction for each protein tested.

Fig. 16.45 Effect of pH on adhesive forces **a** at 22 °C and **b** 37 °C. Error bars represent ± 1 standard error (Bhushan and Utter 2013)



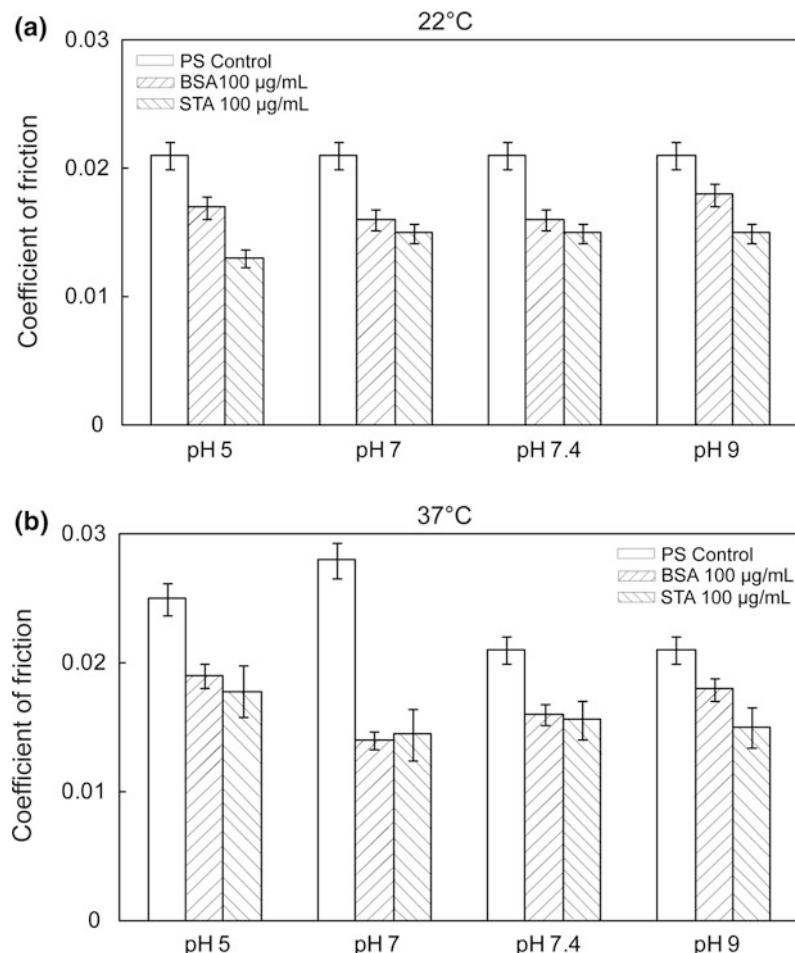
Next, the coefficient of friction was measured for BSA and STA at 100 µg/mL concentration as a function of pH at 22 °C. Figure 16.46a shows the coefficient of friction between the bare silicon nitride AFM tip and two different protein coated surfaces (Bhushan and Utter 2013). Both proteins exhibit lower friction than the polystyrene control, but the data appears to show that the coefficient of friction is fairly constant for all of the samples as the pH of the solution is altered. Also, no clear trend is present to distinguish BSA from STA when the pH varied, but the data from this experiment is in good agreement with what was found earlier by varying the protein concentration.

The effect of temperature also was investigated by heating the sample and solution to 37 °C.

Figure 16.46b shows the coefficient of friction between the bare silicon nitride AFM tip and two different protein coated surfaces (Bhushan and Utter 2013). At an elevated temperature the coefficient of friction increases slightly or stays the same relative to the 22 °C data. Again, the coefficient of friction is lower for the protein coated samples relative to the polystyrene controls, but the coefficient of friction does not significantly change from one pH level to the next.

From these experiments, the coefficient of friction appears to be affected more significantly by the concentration of the adsorbed protein film and not by the pH or temperature of the surrounding medium. As reported by Bhushan et al. (2006a) and Fan et al. (2011), and confirmed with these experiments, protein concentration

Fig. 16.46 Effect of pH on the coefficient of friction **a** at 22 °C and **b** 37 °C. Error bars represent ± 1 standard error (Bhushan and Utter 2013)



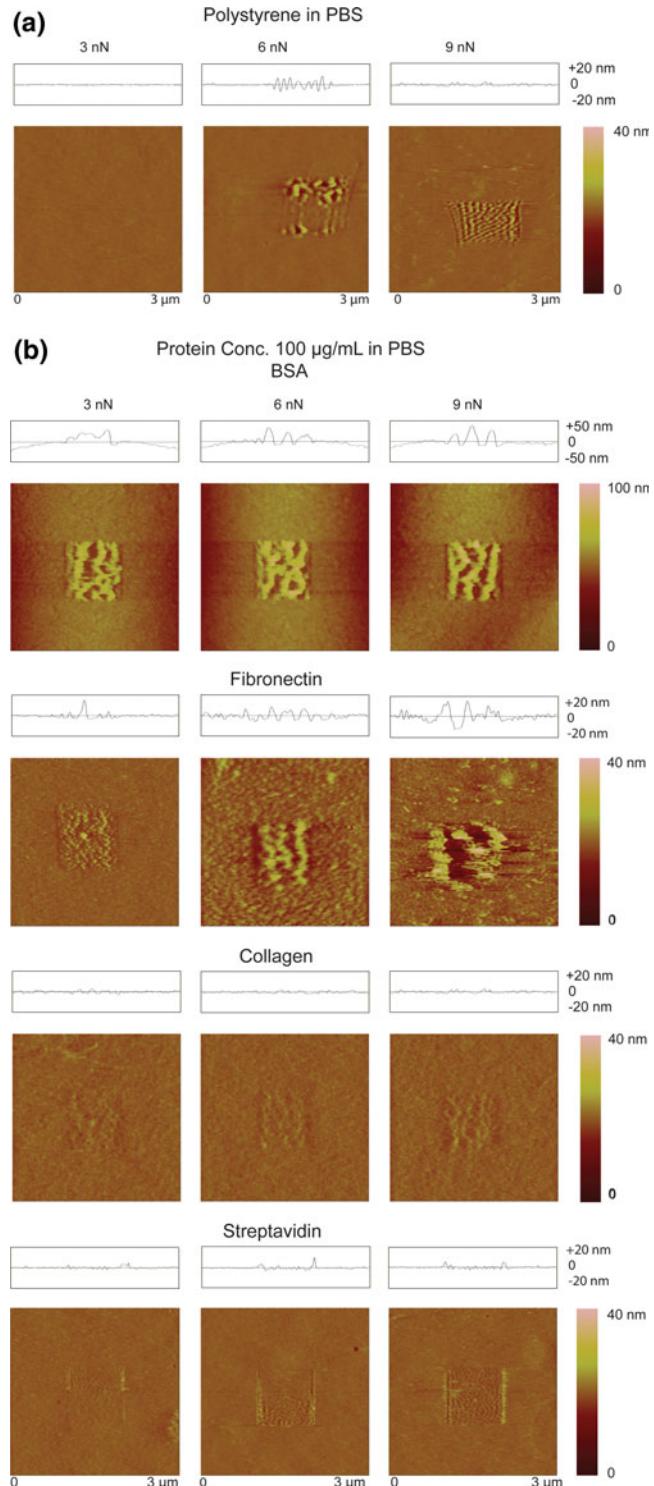
appears to be the primary factor in the formation of a continuous lubrication film. Solution pH and temperature appear to have secondary effects on the friction force experienced by the AFM tip.

16.6.1.3 Nanoscale Wear Tests

Figure 16.47 shows the wear images for each protein and a bare polystyrene control surface (Bhushan and Utter 2013). Each figure shows the height image and cross-sectional profile of a wear mark after conducting the wear scans. For these tests, the protein concentration was kept constant at 100 µg/mL with the normal load increasing from 3 to 9 nN. These tests show that the depth of wear increases with an increasing normal load for all of the samples imaged.

Several key differences become apparent on comparing the wear marks between the different protein coated surfaces. The surface coated with BSA shows the most significant wear relative to all of the other samples. This may be in part due to the smaller mass of BSA and lower isoelectric point. At a pH of 7.4, BSA will have a net negative charge which will interact more strongly with the AFM tip. When the AFM tip comes into contact with the protein film, the net negative charge and small mass allows for more wear to occur relative to the other samples. Streptavidin has approximately the same mass as BSA, but it does not exhibit the degree of wear that BSA undergoes. Therefore, protein mass appears to have a secondary effect on wear over the range of normal forces tested. As expected, the wear loads

Fig. 16.47 AFM height images and corresponding cross-sectional profile of a wear mark in contact mode of **a** bare polystyrene and **b** four proteins at a concentration of 100 $\mu\text{g/mL}$ in PBS. The normal loads used were 3, 6, and 9 nN (Bhushan and Utter 2013)



are of the same order of magnitude as the adhesive forces found in this experiment.

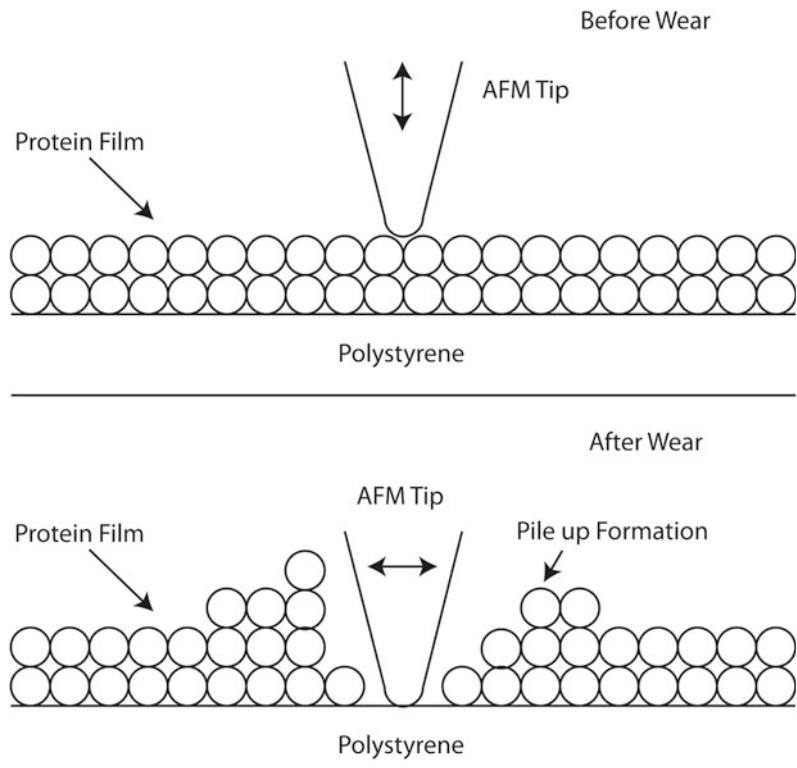
A trend becomes apparent by comparing the wear of the three globular proteins. As the protein isoelectric point becomes more basic and closer to the pH of saline (pH 7.4), wear marks become smaller. The most significant wear occurs with BSA, which has the lowest isoelectric point (pH 4.7). The surface coated with fibronectin (isoelectric pH 5) displays less wear than the BSA sample, but greater wear than the streptavidin (isoelectric pH 5.5) sample. The difference in wear marks for the globular proteins in part may be due to charge-charge repulsion that occurs when the pH of the solution is different than the isoelectric point of the protein.

However, the most interesting surface of all appears to be the one coated with collagen. This surface appears not to undergo any significant wear as compared to the other surfaces on the length scale in which we are interested. Collagen is the only protein tested that is linear in structure compared to the other proteins, and it has the

ability to strongly self-associate (Silver et al. 2003). Collagen also contains more hydrophobic amino acid groups than other proteins, which results in a stronger interaction with the polystyrene surface (Xu et al. 2007). As the protein solution is allowed to sit overnight, slow rearrangements take place in the adsorbed collagen, creating a denser mesh with an increased resistance to wear (Pamula et al. 2004). Given that collagen is the primary load-bearing protein in the body and can be found in tissues in which significant mechanical wear can occur, it seems logical that it could be the most wear resistant protein (Di Lullo et al. 2002). However, as shown by Pamula et al. (2004), and confirmed in the study being presented, as well, the collagen surface will still undergo slight degradation at higher loads. Although fibronectin has the ability to self-associate, it does not appear to form the wear resistant mesh that collagen does due to its globular structure.

Figure 16.48 shows a schematic of the hypothetical wear mechanism for the globular proteins

Fig. 16.48 Schematic of hypothetical wear mechanism of the adsorbed protein film at higher concentration (Bhushan and Utter 2013)



at a higher protein concentration (Bhushan and Utter 2013). Due to the motion of the AFM tip during scanning, the protein molecules displace laterally and begin to interact with the remaining protein molecules. Protein aggregation results leading to a pile up formation that keeps the molecules attached to the surface. The increased normal force causes more protein molecules to displace, leading to an increased degradation of the surface.

16.6.1.4 Summary

Bhushan and Utter (2013) studied the nanoscale adhesion, friction and wear properties of four proteins adsorbed to a spin-coated polystyrene surface using an AFM in DI water and PBS by varying the concentration of the protein solution, pH, and temperature of the liquid medium. They found that under mild operating conditions, all of the proteins had similar adhesion and friction behavior, making a direct comparison between these proteins difficult. They showed that for all four proteins studied, the adhesive force between each protein and the AFM tip increased with increasing protein concentrations. It is believed that this increase in adhesive force in part is due to the formation of multiple interaction sites between the protein film and the AFM tip. Further, as the solution was altered from acidic to basic, adhesive forces also increased for BSA and STA, possibly in part due to the larger net charge of the protein interacting with the AFM tip. Adhesive forces then decreased when the temperature was increased to 37 °C. There was no clear trend present distinguishing BSA from STA, possibly because they have similar structure, isoelectric points, and mass.

It was also shown that for all proteins studied in PBS (pH 7.4), the coefficient of friction decreased as the protein concentration increased. It is believed that this is because the protein layer acts as a continuous lubrication film at higher protein concentrations. As compared to the polystyrene control surface, the coefficient of friction also decreased for BSA and STA when each pH level was tested at 22 °C. For each protein studied, the coefficient of friction remained relatively constant across the range of

pH values. As compared to the polystyrene control surface, increasing the temperature to 37 °C kept the coefficient of friction relatively constant across the range of pH levels for BSA and STA. It appears that the existence of a lubrication film is the most important factor affecting the coefficient of friction. Solution pH and temperature appear only to have secondary effects on the coefficient of friction. There was no clear trend distinguishing BSA from STA when the pH and temperature of the solution are varied.

Wear is the distinguishing factor for the different proteins. Normal load had an effect on wear. As the normal force was increased, there was increased degradation for all of the samples imaged. For all of the proteins tested, the depth of wear marks increased with increased normal load. The globular proteins degraded in the order of their isoelectric points: bovine serum albumin > fibronectin > streptavidin > collagen. It is possible that collagen is more wear resistant because of its linear structure and ability to self-associate. For the normal forces tested, protein mass appears only to have a secondary effect on wear; the primary wear mechanism results from protein structure and isoelectric point.

16.6.2 Adhesion, Friction, and Wear of Biomolecules on Si-Based Surfaces

Proteins on silicon-based surfaces are of extreme importance in various applications, including silicon microimplants, various bioMEMS such as biosensors, and therapeutics. Silicon is a commonly used substrate in microimplants, but it can have undesired interactions with the human immune system. Therefore, to mimic a biological surface, protein coatings are used on silicon based surfaces as passivation layers, so that these implants are compatible in the body and avoid rejection. Whether this surface treatment is applied to a large implant or a bioMEMS, the function of the protein passivation is obtained from the nanoscale 3D structural conformation of the protein. Proteins are also used in bioMEMS because of their function specificity. For

biosensor applications, the extensive array of protein activities provides a rich supply of operations that may be performed at the nanoscale. Many antibodies (proteins) have an affinity to specific protein antigens. For example, pathogens (disease causing agents, e.g., virus or bacteria) trigger the production of antigens which can be detected when bound to a specific antibody on the biosensor. The specific binding behavior of proteins that has been applied to laboratory assays may also be redesigned for *in vivo* use as sensing elements of a bioMEMS. The epitope-specific binding properties of proteins to various antigens are useful in therapeutics. Adhesion between the protein and substrate affects the reliability of an application. Among other things, morphology of the substrate affects the adhesion. Furthermore, for *in vivo* environments, the proteins on the biosensor surface should exhibit high wear resistance during direct contact with the tissue and circulatory blood flow without washing off.

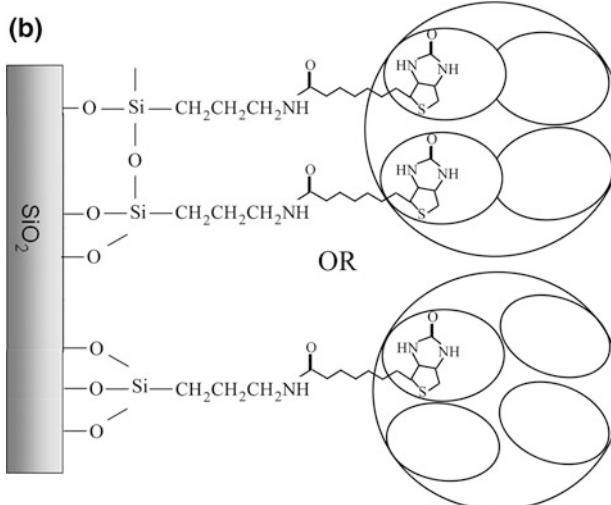
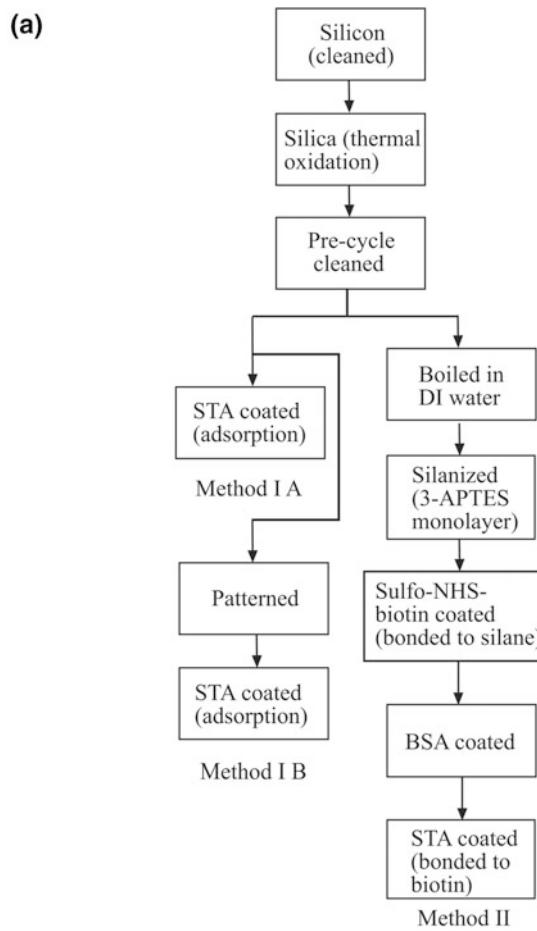
Bhushan et al. (2005a) studied the step by step morphological changes and the adhesion of a model protein—streptavidin (STA)—on silicon based surfaces (see also, Lee et al. 2005a; Eteshola et al. 2008 and Bhushan et al. 2009). Figure 16.49a presents a flow chart showing the sequential modification of a silicon surface. In addition to physical adsorption, they also used nanopatterning and chemical linker methods to improve adhesion. A nanopatterned surface contains large edge surface area, leading to high surface energy which results in high adhesion. In chemical linker method, sulfo-N-hydroxysuccinimido-biotin (sulfo-NHS-biotin) was used as a cross linker because the bonds between the STA and the biotin molecule are one of the strongest non-covalent bonds known (Fig. 16.49b). Sulfo-NHS-biotin was connected to the silica surface through a silane linker, 3-aminopropyltriethoxysilane (3-APTES). In order to make a bond between the silane linker and the silica surface, the silica surface was hydroxylated. Bovine serum albumin (BSA) was used before STA to block non-specific binding sites of the STA protein with silica surface. Figure 16.50 shows the step by step

morphological changes in the silica surface during the deposition process using the chemical linker method. There is an increase in roughness (σ and P-V (peak-to-valley distance)) of the silica surface boiled in de-ionized (DI) water compared to the bare silica surface. After the silanization process, there are many free silane links on the surface which caused higher roughness. Once biotin was coated on the silanized surface, the surface became smoother. Finally, after the deposition of STA, the surface shows large and small clumps. Presumably, the large clumps represent BSA, and the smaller ones represent STA. To measure adhesion between STA and the corresponding substrates, a STA coated tip (or functionalized tip) was used and all measurements were made in PBS, a medium commonly used in protein analysis to simulate body fluid. Figure 16.51 shows the adhesion values of various surfaces. The adhesion value between biotin and STA was higher than that for other samples, which is expected. Edges of patterned silica also exhibited high adhesion values. It appears that both nanopatterned surfaces and chemical linker method increase adhesion with STA.

Bhushan et al. (2006a) studied friction and wear of STA deposited by physical adsorption and chemical linker method (see also, Bhushan et al. 2009). Figure 16.52 shows the coefficient of friction between the Si_3N_4 tip and various samples. The coefficient of friction is less for STA-coated silica samples being adsorbed, as compared to an uncoated sample. The STA coating acts as a lubricant film. The coefficient of friction is found to be dependent upon the concentration of STA, and it decreases with an increase in the concentration. Bhushan et al. (2005a) have reported that the density and distribution of the biomolecules vary with the concentration. At a higher concentration of the solution, the coating layer is more uniform, and covers more of the silica substrate surface than at a lower concentration. This means that the surface forms a continuous lubricant film at a higher concentration.

In the case of samples prepared by chemical linker method, the coefficient of friction increases

Fig. 16.49 **a** Flow chart showing the samples used and their preparation technique, and **b** a chemical structure showing streptavidin protein binding to the silica substrate by the chemical linker method



Streptavidin has four biotin-binding pockets. Two or one may be attached to the biotin on the surface, with the remaining 2 or 3 available to bind the biotin analyte.

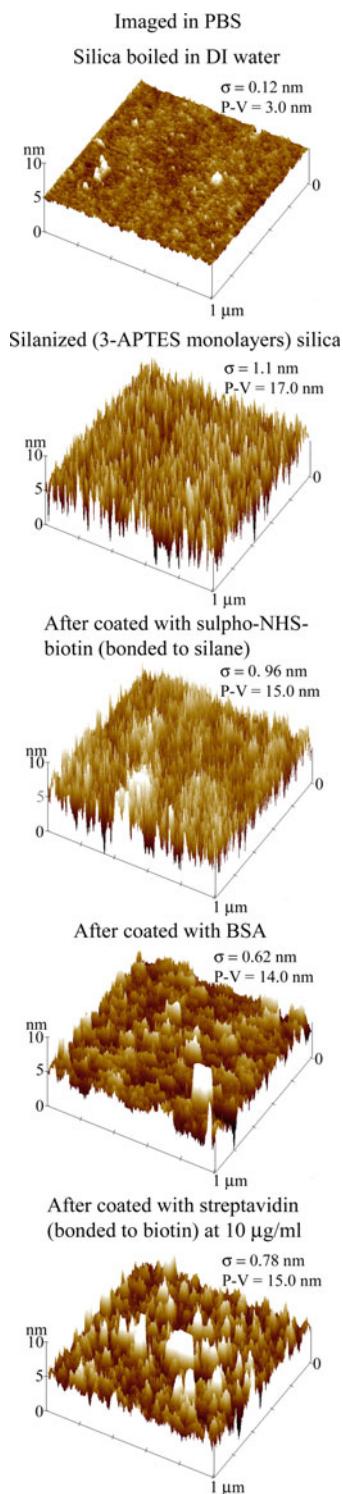


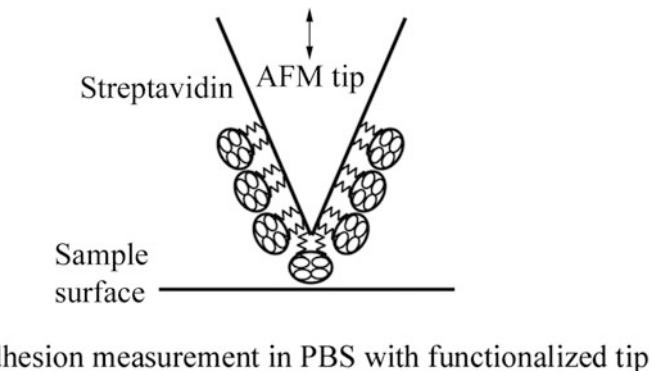
Fig. 16.50 Morphological changes in silica surface during functionalization of silica surface by chemical linker imaged in PBS. Streptavidin is covalently bonded at 10 $\mu\text{g}/\text{ml}$ concentration (Bhushan et al. 2005a)

with an increase in the biomolecular chain length due to increased compliance. When normal load is applied to the surface, the surface gets compressed, resulting in a larger contact area between the AFM tip and the biomolecules. Further, the size of STA is much larger than that of APTES and biotin, resulting in a surface tightly packed with biomolecules that results in very little lateral deflection of the sulfo-NHS-biotin linker. Due to the high contact area and low lateral deflection, the friction force increases for the same applied normal load as compared to samples prepared by the physical adsorption method. These tests reveal that the surfaces coated with biomolecules reduce friction, but if the biomolecular coating of the surface is too thick, or the surface has some cushioning effect, as seen in chemical linker method, the coefficient of friction increases.

Figure 16.53 shows the surface height maps and phase images of wear marks on STA deposited by physical adsorption taken after wear tests at three normal loads. The wear depth increases with the increasing normal load. An increase in normal load causes partial damage to the folding structure of the streptavidin molecules. It is unlikely that the chemical (covalent) bonds within the streptavidin molecule are broken, instead, the folding structure is damaged, leading to the wear mark. When the load is high, about 30% of free amplitude ($\sim 8 \text{ nN}$), the molecules may be removed by the AFM tip due to the indentation effect. Because of this, there is a significant increase in the wear depth from 50% of free amplitude ($\sim 6 \text{ nN}$) to 30% of free amplitude ($\sim 8 \text{ nN}$). The data show that biomolecules will be damaged during sliding.

In summary, for the samples prepared using nanopatterning and chemical linker methods, adhesion is higher as compared to the direct adsorption method. The coefficient of friction is smaller for STA-coated silica prepared using the direct adsorption method as compared to an uncoated silica sample. The coefficient of friction decreases with increasing concentration of STA in the solution because the protein acts as a lubricating film. Friction increases for the STA sample prepared using the chemical linker method due to low lateral deformation and the

Fig. 16.51 Adhesion measurements of silica, patterned silicon, silica boiled in DI water, and sulfo-NHS-biotin using functionalized (with streptavidin) tips obtained from force-distance curves, captured in PBS



Adhesion measurement in PBS with functionalized tip

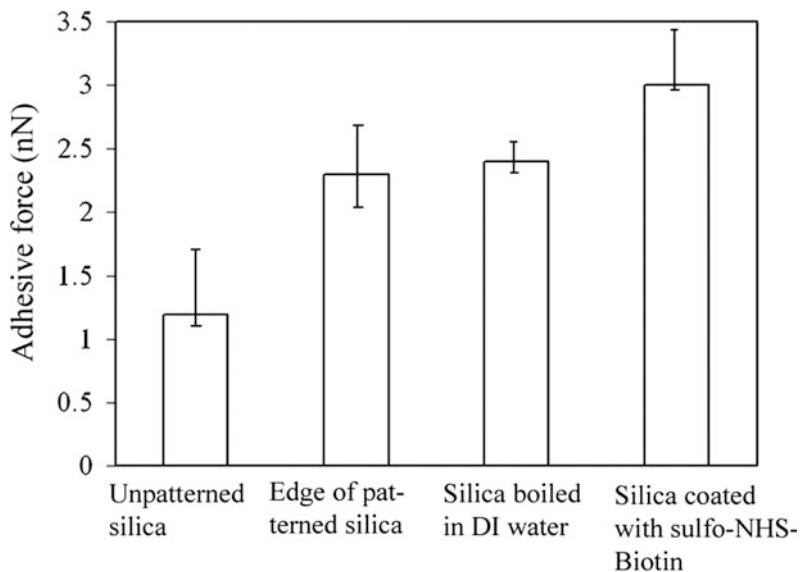
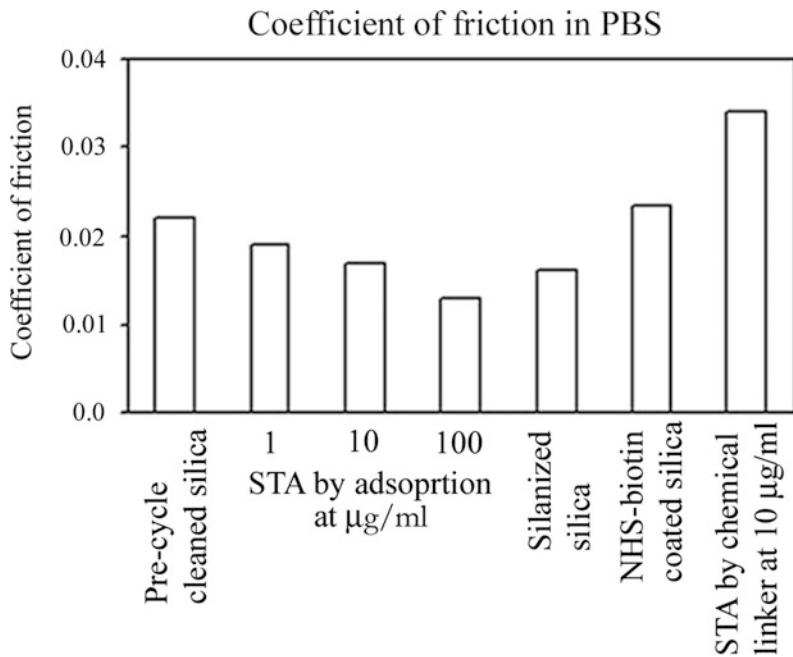


Fig. 16.52 Coefficient of friction for various surfaces with and without biomolecules (Bhushan et al. 2006a)



Streptavidin adsorbed on pre-cycle cleaned silica at 10 µg/ml

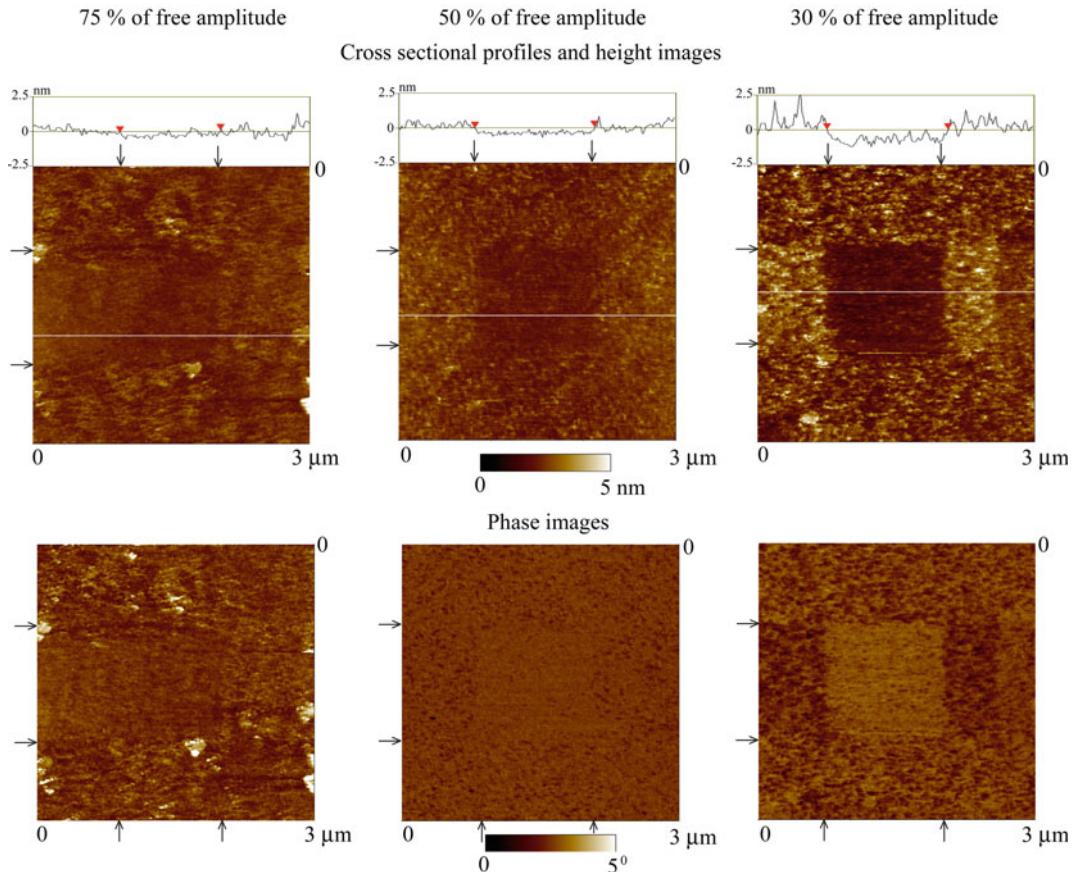


Fig. 16.53 Surface height maps and cross-sectional profiles (top) and phase images (bottom) of wear marks on precycle cleaned silica coated with streptavidin by physical adsorption after wear tests at three normal loads (increasing from left to right). The 75, 50, and 30% of free amplitudes correspond to equivalent normal loads of 3, 6, and 8 nN, respectively (Bhushan et al. 2006a)

cushioning effect. Wear of STA increases with the increasing load.

16.6.2.1 APDMES to Improve Adhesion, Friction and Wear of Biomolecular Films

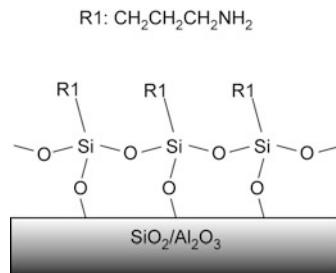
APTES films used in the study reported above were not very smooth, and P-V distances were greater than the summed bond strengths of APTES. Therefore, the Biotin/STA deposited on APTES films were not very robust. APTES films are commonly described as SAMs, though this is often inaccurate (Kallury et al. 1994; Moon et al. 1996; Han et al. 2006). APTES monomers can

form infinite, cross-linked siloxane polymer lattices because APTES is a multi (tri)-valent silane. Further, APTES can form a multilayered structure due to intermolecular polymerization with significant cross-linking between monomers. It should consist of sparse cross-links between the polymer and the substrate, Fig. 16.54 (Bhushan et al. 2009). The propensity to form multilayers and its low mechanical strength make APTES an undesirable interface material. On the other hand, 3-aminopropyltrimethoxysilane (APDMES) cannot polymerize into extensive networks because it is monovalent, forming only siloxane dimers or linkages to substrate oxides

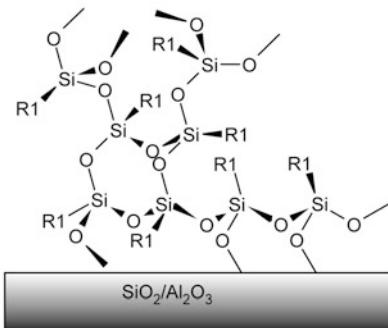
Fig. 16.54 Schematics of idealized and multilayered APTES and idealized APDMES silane polymer films on silicon/SiO₂ and aluminum/Al₂O₃ substrates. Idealized film corresponds to a self-assembled monolayer (Bhushan et al. 2009)

Chemical structure of silane polymer linker

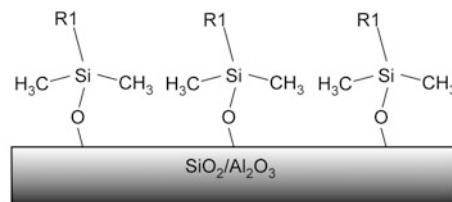
Idealized APTES



Multilayered APTES



Idealized APDMES



(Fig. 16.54). Therefore, APDMES should produce thinner films of greater mechanical robustness than APTES and provide the robustness to the biomolecular layers deposited on its top. Thinner interfaces theoretically would also increase sensor sensitivity to analytes.

Bhushan et al. (2009) examined the thickness and durability of APDMES deposited on SiO₂ and Al₂O₃ substrates, with biotin and biotin/STA bound to them. Figure 16.55a shows the thickness and surface roughness (RMS and P-V) of APTES, APDMES, and SiO₂ substrate for reference. Figure 16.55b shows the surface roughness, adhesive force, and coefficient of friction for biotin and STA-biotin deposited on APTES and APDMES films. The data show that the APTES film was not very smooth. It was shown in Fig. 16.50 that biotin and STA deposited on APTES were also not very smooth with high P-V values. The thickness of the APTES layer was larger than the expected monolayer thickness. The unexpected thickness of the film suggests

multilayering by APTES. APDMES films produced the thinner film, with a thickness comparable to the summed bond lengths of the APDMES polymer (Fig. 16.55a). Bhushan et al. (2009) reported that the APDMES film was more uniform, smoother, nearly continuous, and exhibited a higher contact angle and lower adhesive force as compared to the APTES film. The surface roughness and adhesive force of biotin and STA-biotin on APDMES was also slightly lower than that on APTES film. The coefficient of friction on APDMES was slightly higher than that on APTES (Fig. 16.55b).

Bhushan et al. (2009) also studied wear properties of various films. Contact mode always immediately stripped the surface at low loads; consequently, wear experiments were performed in tapping mode at various loads. AFM surface height and phase angle images and cross-sectional profiles obtained after wear tests for biotin and STA-biotin on APTES and APDMES in PBS are shown in Fig. 16.56a. As

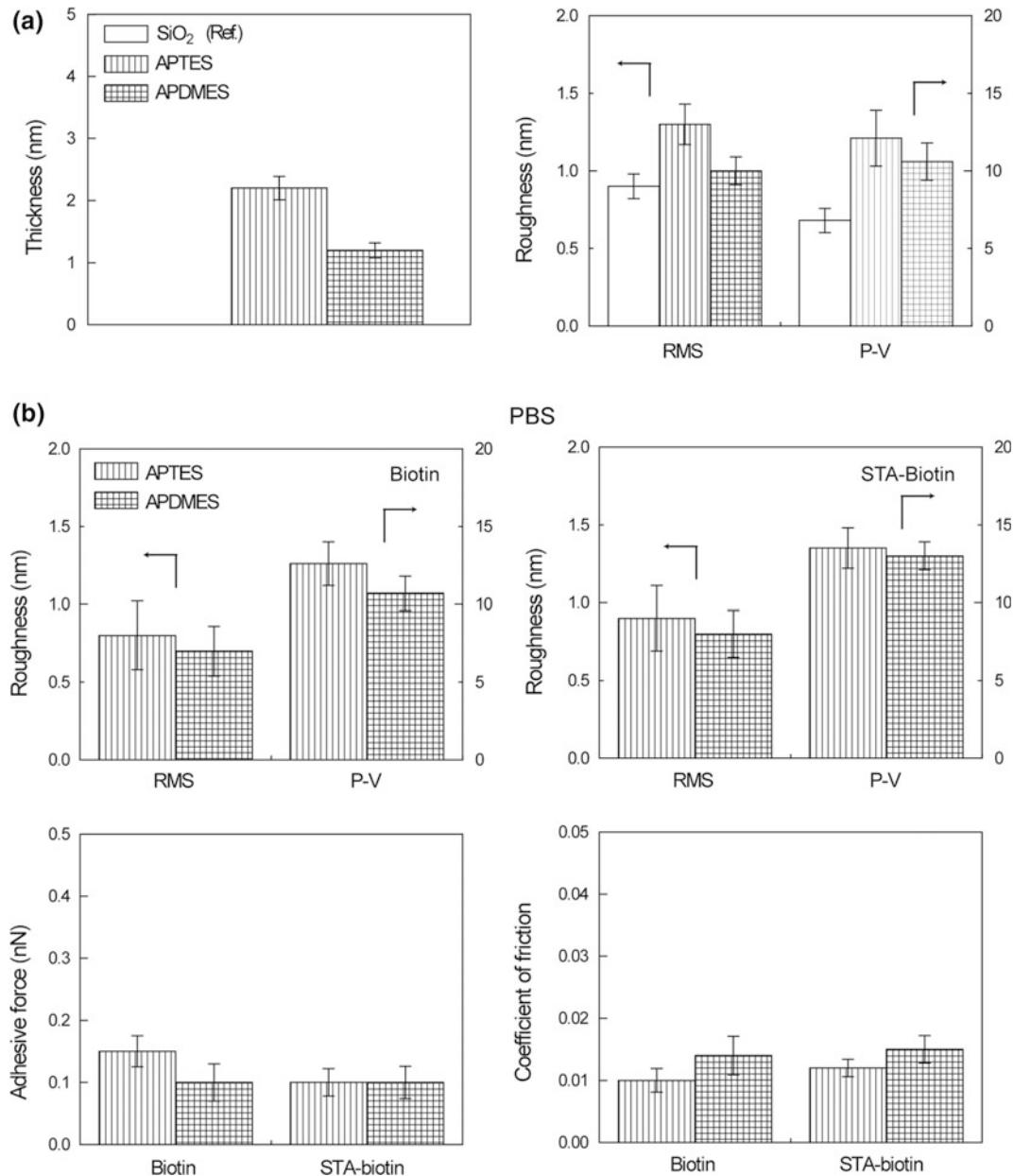


Fig. 16.55 **a** Summary of film thickness and surface roughness (RMS and P-V distance) for APTES and APDMES in air, and **b** summary of the surface roughness (RMS and P-V distance), adhesive force, and coefficient of biotin and STA-biotin on APTES and APDMES films, all in PBS buffer solution. The error bars represent $\pm 1\sigma$ (Bhushan et al. 2009)

controls, the wear experiments of biotin and STA-biotin on SiO₂ without SAM were also carried out. The biomolecular films on APDMES were more robust than that on APTES. Given that each molecule of APDMES must be bonded

to a surface oxide group and that intrapolymer cross-links are not possible in APDMES, it has the highest density of siloxane linkages to the substrate oxide, making it the more robust. Bhushan et al. (2009) also studied the effect of

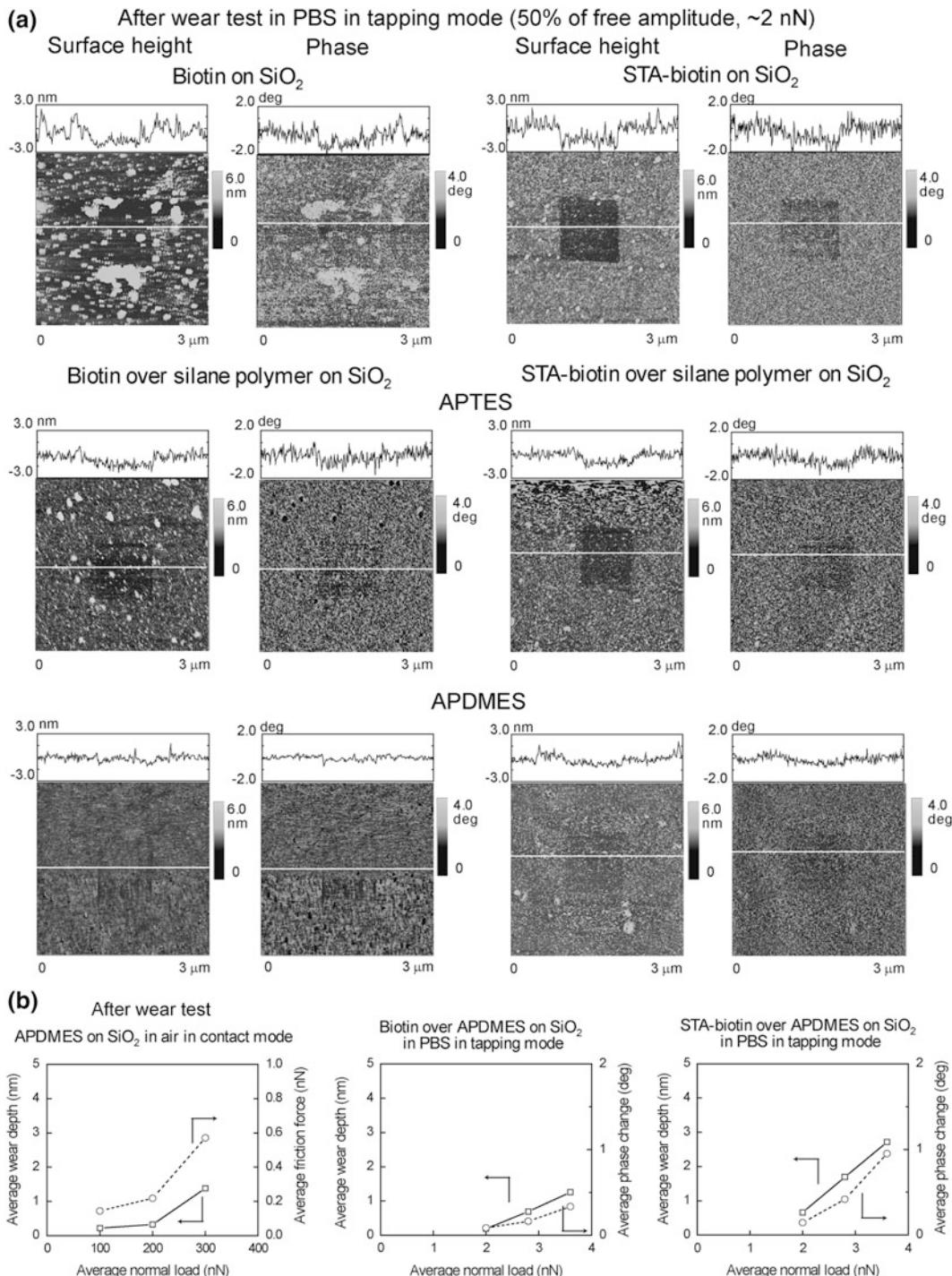


Fig. 16.56 **a** AFM surface height and phase angle images and cross-sectional profiles obtained after wear test in PBS in tapping mode at 50% of free amplitude (~2 nN) on biotin and STA-biotin on SiO₂ and APTES and APDMES films on SiO₂. The white lines indicate the locations of the cross-sections, and **b** plot of average wear depth and average friction force/phase angle as a function of average normal load for APDMES on SiO₂ in air in contact mode, biotin over APDMES on SiO₂ in PBS in tapping mode, and on STA-biotin over APDMES on SiO₂ in PBS in tapping mode (Bhushan et al. 2009)

load on the wear of APDMES and biotin and STA-biotin deposited on APDMES, Fig. 16.56b. As expected, wear increased with an increase in the load. The relationship between the average wear depth (and the coefficient of friction) and the average normal load is generally linear. The slope of the wear depth (and coefficient of friction) to load is steepest for the interface to which STA-biotin was bound because of the cushioning effect of the thick film, as suggested earlier.

In summary, an APDMES film is more uniform, smoother, and nearly continuous as compared to APTES film. These properties of APDMES provide a good interfacial material for biomolecular films providing a smooth and robust structure.

16.6.3 Adhesion of Coated Polymer Surfaces

As mentioned in Appendix, PMMA, PDMS, and other polymers are used in the construction of micro/nanofluidic based biodevices. Adhesion between the moving parts needs to be minimized. Furthermore, if the adhesion between the microchannel surface and the biofluid is high, the biomolecules will stick to the microchannel surface and restrict flow. In order to facilitate flow, a surface with low bioadhesion is required.

Tambe and Bhushan (2005b, c) and Bhushan and Burton (2005) have reported adhesive force data for PMMA and PDMS against an AFM

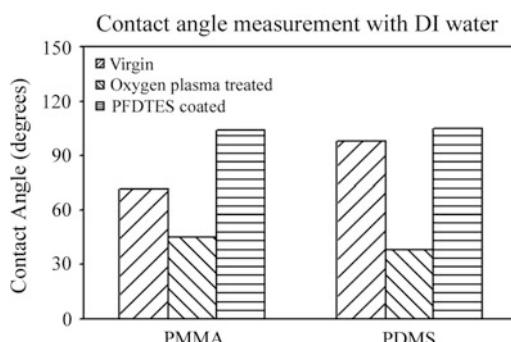


Fig. 16.57 Sessile drop contact angle measurements of virgin, oxygen plasma treated and PFDTES coated PMMA and PDMS surfaces. The maximum error in the data is $\pm 2^\circ$ (Tokachichu and Bhushan 2005)

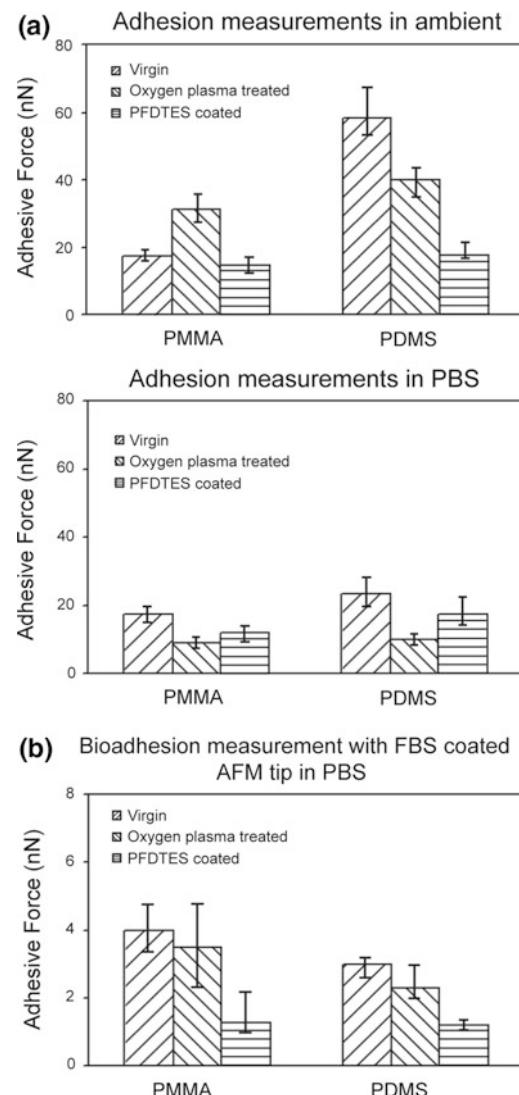


Fig. 16.58 Adhesion measurement of virgin, oxygen plasma treated and PFDTES coated PMMA and PDMS surfaces **a** with bare silicon nitride AFM tip in ambient, and in PBS environment, and **b** dip coated tip with FBS in PBS environment (Tokachichu and Bhushan 2005)

Si_3N_4 tip and a silicon ball. Tokachichu and Bhushan (2005) measured contact angle and adhesion of bare PMMA and PDMS and coated with a perfluoro SAM of perfluorodecyltriethoxysilane (PFDTES). Oxygen plasma treatment was used for hydroxylation of the surface to enhance chemical bonding of the SAM to the polymer surface. They made measurements in

ambient conditions, PBS, and fetal bovine serum (FBS); the latter is a blood component. Figures 16.57 and 16.58 show the contact angle and adhesion data. SAM coated surfaces have high contact angles (Fig. 16.57), as expected. The adhesion value of PDMS in ambient conditions is high because of the electrostatic charge present on the surface. The adhesion values of PDMS are higher than PMMA because PDMS is softer than PMMA (elastic modulus = 5 GPa and hardness = 410 MPa (Wei et al. 2005)), resulting in a higher contact area between the PDMS surface and the AFM tip. Additionally, PMMA does not develop electrostatic charge. When SAM is coated on PMMA and PDMS surfaces, the adhesion values are similar, which shows that the electrostatic charge on virgin PDMS plays no role when the surface is coated. In the PBS solution, there is a decrease in adhesion values because there is no meniscus contribution. The adhesion values for an FBS-coated tip in PBS are generally lower than for an uncoated tip in PBS.

In summary, the adhesion values of SAM coated surfaces are lower than bare surfaces in various environments.

16.7 Trajectory of Submicron Particles for Therapeutic and Diagnostics

For drug delivery, submicron particles are injected into the blood stream in human capillaries (as small as 4–5 μm) and employed to deliver drugs to diseased cells, to locate diseased cells or tumoral masses and estimate the state of disease, and to carry diagnostic agents (fluorescent molecules) to diseased cells to enhance imaging (Decuzzi et al. 2005). Particles exhibiting one of these characteristics are considered smart systems, which can function as purely therapeutic agents or purely diagnostic agents or a combination of both. Small delivery particles include nanocrystals, synthetic vesicles, liposomes, and fabricated silicon.

A schematic of the trajectory of a submicron particle injected intravenously for selectively targeting tumor cells for drug delivery is shown in Fig. 16.59. Particles are reservoirs containing drug or diagnostic agents. These are covered with a layer of adhesive molecules (ligands) for attaching to selected target sites. When a particle

Nanoparticles selectively targeting tumor cells for drug delivery

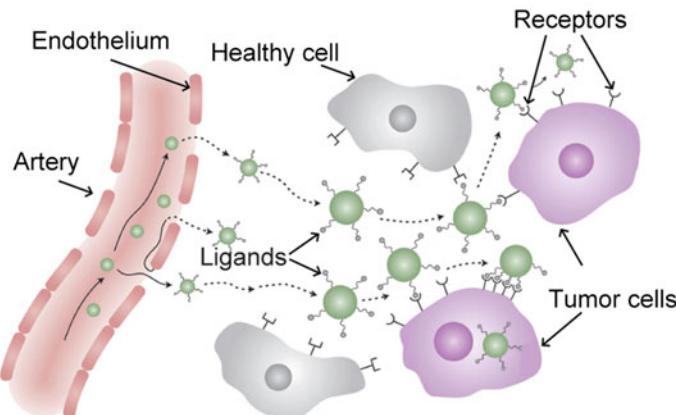


Fig. 16.59 Trajectory of submicron particles injected intravenously for selectively targeting tumor cells for drug delivery (Irvine 2011)

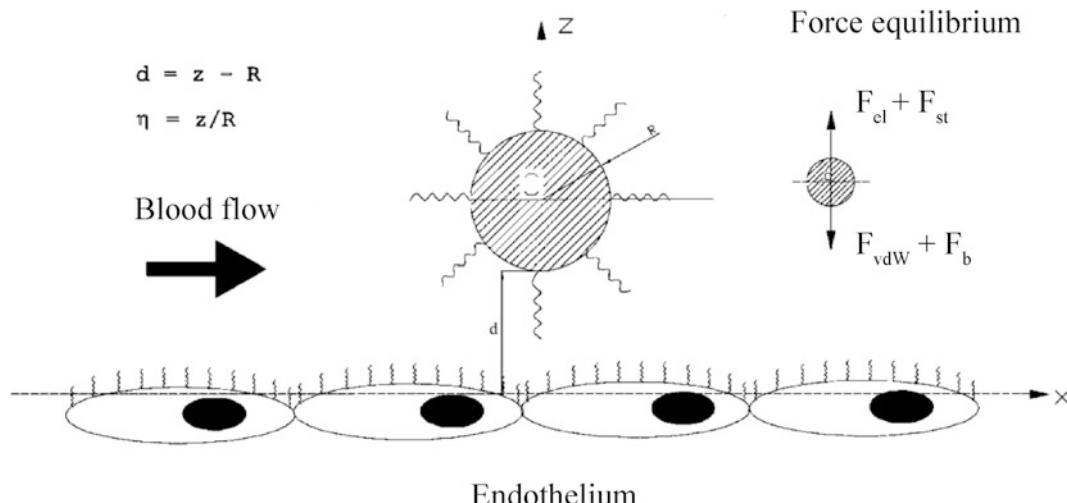


Fig. 16.60 A spherical particle moving close to the endothelium wall in a laminar flow. In the inset, the balance of the forces acting over the particle are sketched. F_{el} , F_{st} , F_{vdW} , and F_b correspond to electrostatic force, steric force, van der Waals force, and buoyancy force, respectively (Decuzzi et al. 2005)

is sufficiently close to select cell surfaces (a few nm), its ligands can interact with the cellular counterpart (receptors), leading to firm attachment. Once the particle is attached to its target, a drug or diagnostic agent is delivered. The speed and the time needed for a particle circulating in the blood stream to reach the endothelium lining of the human capillary is dependent upon the distance and interactive forces. The particle can interact with the endothelium lining with buoyancy, van der Waals forces, electrostatic, and steric forces. These interactions are a function of the material properties of the particles and any coating on it, in particular, relative density (particle density relative to blood), the electrostatic potential, the dielectric constant, and the particle radius. These properties can be optimized.

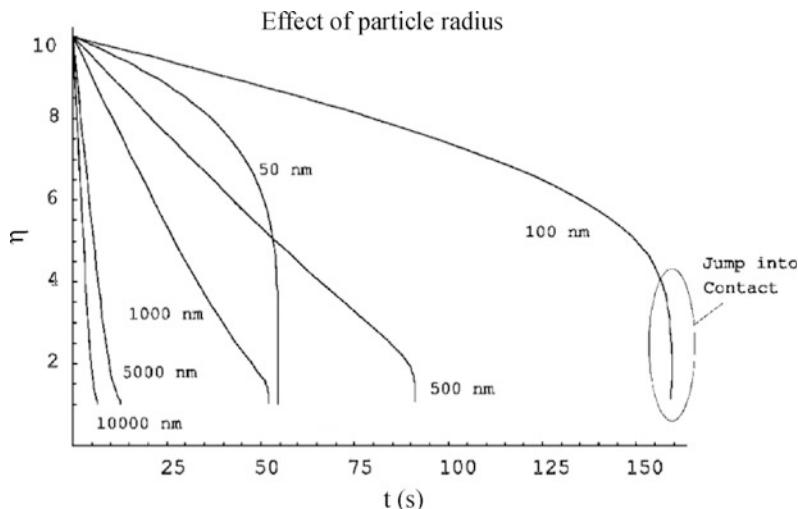
The interacting forces are weak so the particle trajectory should be close to the endothelium lining. Otherwise, the particle may not get

attracted to the lining and perform its intended function. Decuzzi et al. (2005) developed an analytical model to predict the trajectory of a particle freely circulating in the blood stream and associated interaction forces. The model can be used to optimize the particle radii and material properties. Figure 16.60 shows a spherical particle freely circulating in the blood stream with its center at a distance z from the endothelium wall. The particle has a radius R and its trajectory is governed by the forces exerted by the blood stream, gravitation, and electromagnetic interactions. We assume that the particle is sufficiently far from the endothelium wall that specific interactions (such as ligand-receptor interactions) can be ignored. At short range (below 1–3 nm), solvation and other steric forces dominate. However, most important long range forces (5–10 nm) between particles and wall surfaces in the presence of a liquid are buoyancy and van der

Table 16.10 The value of the buoyancy force F_b , van der Waals force F_{vdW} , electrostatic force F_{el} and steric force F_{sr} as a function of the particle radius for a fixed distance from the wall ($\eta = z/R = 2$ or $d = R$) (Decuzzi et al. 2005)

R (nm)	F_b (pN)	F_{vdW} (pN)	F_{el} (pN)	F_{sr} (pN)
10,000	41.092	0.0005	~0	~0
1000	41.092×10^{-3}	0.005	~0	10^{-33}
100	41.092×10^{-6}	0.05	10^{-52}	1.2×10^{-3}
50	51.3×10^{-7}	0.1	10^{-25}	3×10^{-2}

Fig. 16.61 The dimensionless particle position η ($= z/R$) as a function of time t for different values of particle radius ($R = 10,000, 5000, 1000, 500, 100$, and 50 nm). There exists a critical radius R_c at which the travel time is maximum (Decuzzi et al. 2005)



Waals, electrostatic and steric forces. Buoyancy is related to the radius of the particle R and the relative density of the particle relative to blood. The van der Waals interaction (generally attractive) is related to R , its relative position with respect to the endothelium wall z , and the Hamaker constant A , which depends upon the dielectric constants of the media involved. The electrostatic double layer (EDL) interaction (repulsive and attractive) is related to R , z , the ionic concentration, and the characteristic Debye length of the solution. Finally, steric repulsive interaction is related to the unperturbed radius of gyration of polymer chains grafted on the particle surface, R and z . The value of various forces as a function of the particle radius R for a fixed distance from the wall z is presented in Table 16.10. The dominating force is buoyancy when the particle radius is sufficiently large, and van der Waals when the particle is sufficiently small. Electrostatic and steric forces are negligible as long as the distance is larger than the 50 nm considered in this example.

The travel time needed to reach the wall depends upon the particle size, as shown in Fig. 16.61. In this figure, the dimensionless position of the particle center with respect to the endothelium wall η ($= z/R$) is plotted as a function of time t for different particle radii, ranging from $R = 10 \mu\text{m}$ to 50 nm. It was assumed that

particles were initially at a distance d_0 ($z_0 - R$) equal to $9 R$ from the endothelium wall (i.e., the center of the particle is at $z_0 = 10 R$, $\eta_0 = z_0/R = 10$). Thus, the distance traveled scales with the size of the particle. The data show that as the radius of the particle decreases, the time needed to reach the wall increases up to a maximum beyond which it decreases as the radius is further reduced. Larger particles are initially far from the endothelium wall; a $10 \mu\text{m}$ particle is initially $100 \mu\text{m}$ away from the endothelium ($z = 10 R$) where electrostatic and van der Waals interactions are negligible and the particle motion is governed by buoyancy and hemodynamic resistance. As the particle approaches the endothelium wall, van der Waals attraction dominates, and a rapid increase in speed is observed with a *jump into contact* like behavior. Such a behavior is more clearly shown by particles with submicron radius closer to the endothelium wall, so the van der Waals force dominates from the beginning. We note that there is a critical radius at which the time taken for the particle to travel to the wall is maximum. Selected radius should be smaller or larger than the critical radius (possibly smaller so that they can circulate freely even in smaller capillaries).

The effect of the relative density of the particle $\Delta\rho$ on the critical radius R_c at which travel time is maximum is plotted in Fig. 16.62. We note

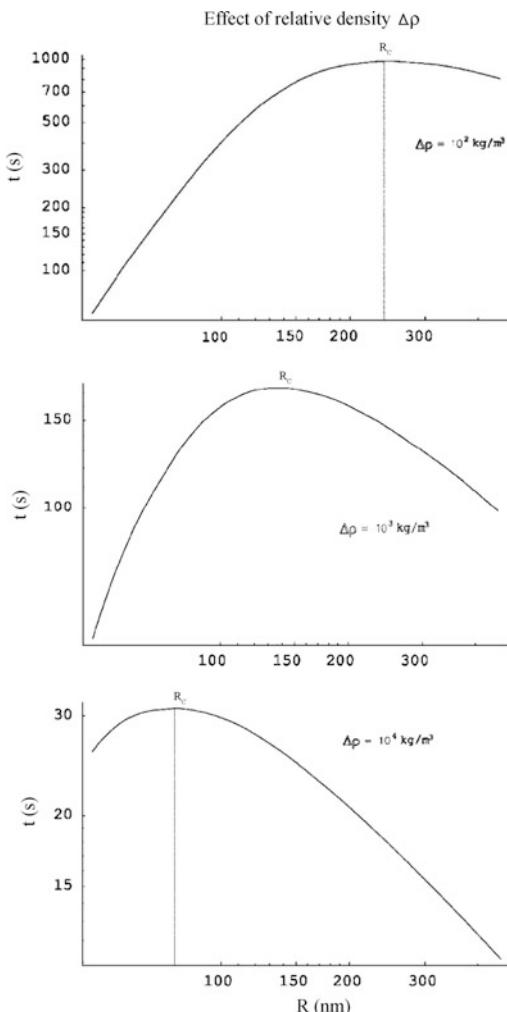


Fig. 16.62 The travel time as a function of particle radius R for different values of the relative density **a** $\Delta\rho = 10^2$, **b** $\Delta\rho = 10^3$, and **c** $\Delta\rho = 10^4 \text{ kg/mm}^3$. The critical radius R_c depends upon $\Delta\rho$ (Decuzzi et al. 2005)

that as $\Delta\rho$ reduces, the R_c increases. As $\Delta\rho$ decreases, the effect of buoyancy becomes less important, and van der Waals attractive forces exert a greater influence as particle radii decrease.

In summary, the interacting forces are weak, so the particle trajectory should be close to the endothelium lining. The trajectory and interaction forces depend upon the particle radii and material properties. These results suggest that particles for therapeutics and diagnostics should have a radius larger or smaller than the critical value (possibly smaller so that they can circulate freely even in smaller capillaries). The material properties of the particle and the polymer chains grafted on it, such as the relative density of the particle, can be tuned specifically to the type of malignant tissue and the state of disease, improving the particle affinity with the diseased cell.

Component Level Studies

16.8.1 Surface Roughness Studies of Micromotor Components

Most of the friction forces resisting motion in a micromotor are concentrated near the rotor-hub interface where continuous physical contact occurs. Surface roughness usually has a strong influence on the friction characteristics on the micro/nanoscale. A catalog of roughness measurements on various components of a MEMS

Table 16.11 Surface roughness parameters and microscale coefficient of friction for various micromotor component surfaces measured using an AFM

	RMS roughness ^a (nm)	Peak-to-valley distance ^a (nm)	Skewness ^a , Sk	Kurtosis ^a , K	Coefficient of microscale friction ^b (μ)
Rotor topside	21 ± 0.6	225 ± 23	1.4 ± 0.30	6.1 ± 1.7	0.07 ± 0.02
Rotor underside	14 ± 2.4	80 ± 11	-1.0 ± 0.22	3.5 ± 0.50	0.11 ± 0.03
Stator topside	19 ± 1	246 ± 21	1.4 ± 0.50	6.6 ± 1.5	0.08 ± 0.01

Mean and $\pm 1\sigma$ values are given

^aMeasured from a tapping mode AFM scan of size $5 \mu\text{m} \times 5 \mu\text{m}$ using a standard Si tip scanning at $5 \mu\text{m/s}$ in a direction orthogonal to the long axis of the cantilever

^bMeasured using an AFM in contact mode at $5 \mu\text{m} \times 5 \mu\text{m}$ scan size using a standard Si_3N_4 tip scanning at $10 \mu\text{m/s}$ in a direction parallel to the long axis of the cantilever

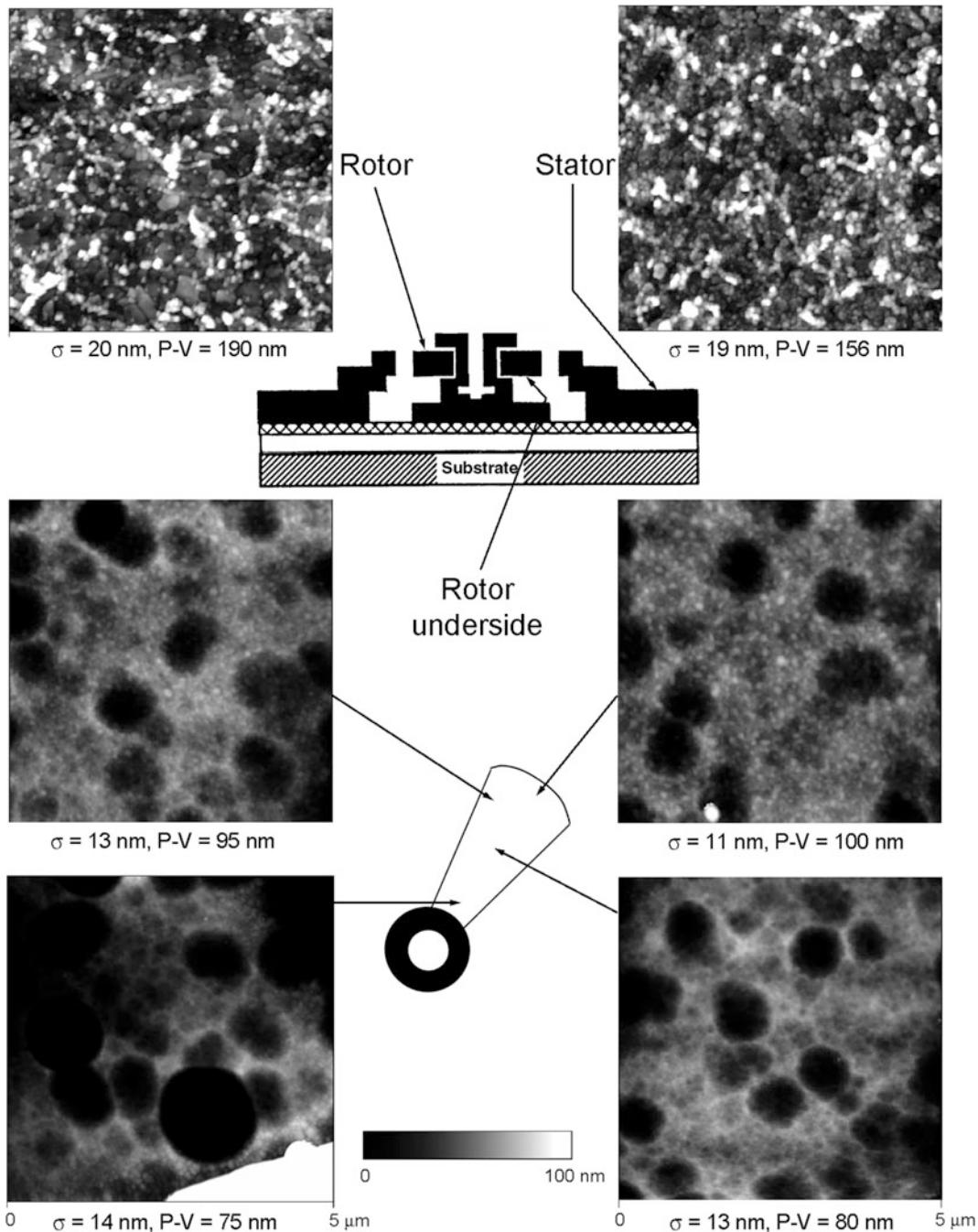


Fig. 16.63 Representative AFM surface height images obtained in tapping mode ($5 \mu\text{m} \times 5 \mu\text{m}$ scan size) of various component surfaces of a micromotor. RMS roughness and P-V values of the surfaces are given. The underside of the rotor exhibits drastically different topography from the topside (Sundararajan and Bhushan 2001)

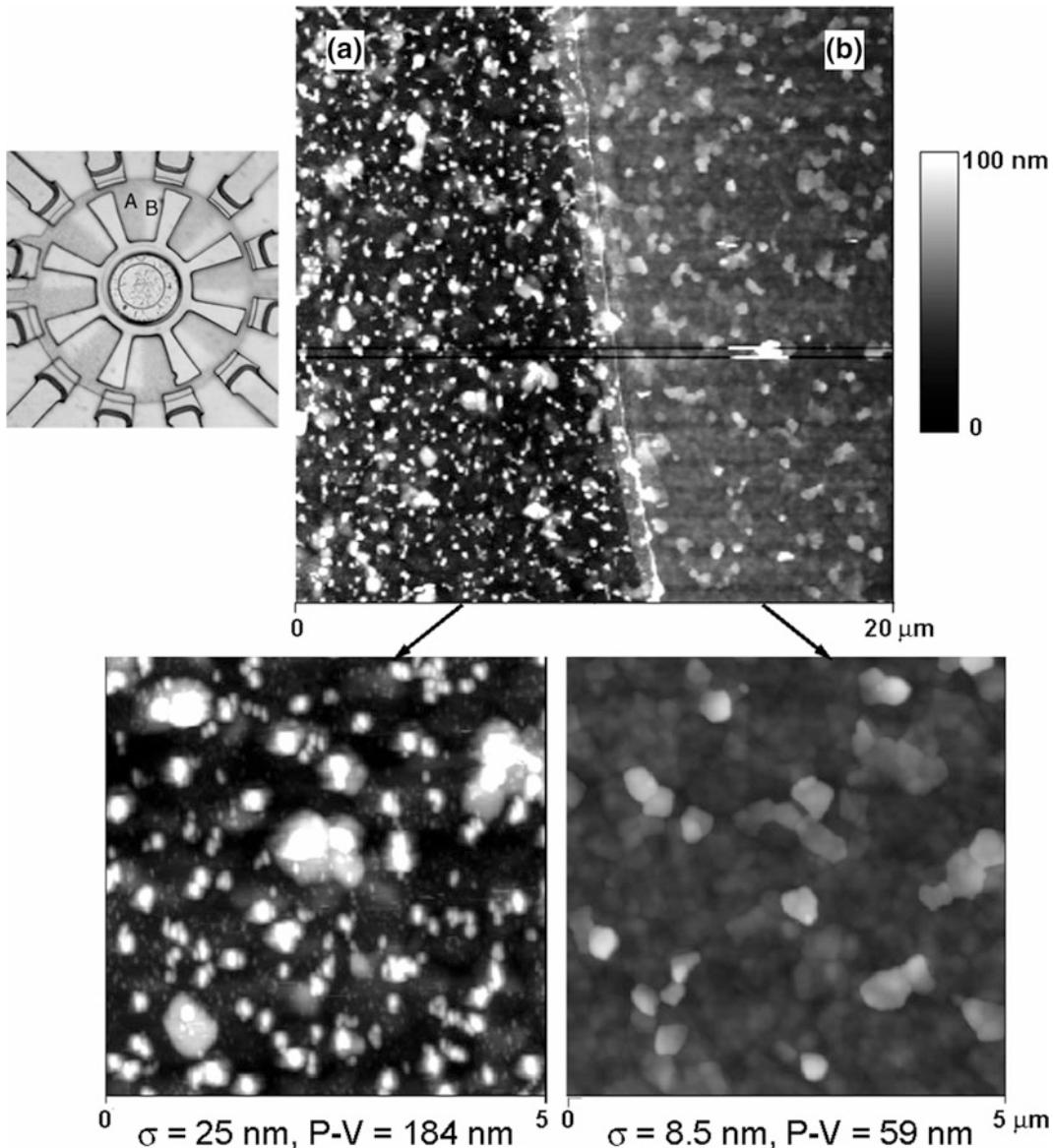


Fig. 16.64 Surface height images of polysilicon regions directly below the rotor. Region A is away from the rotor while region B was initially covered by the rotor prior to the release etch of the rotor. During this step, slight movement of the rotor caused region B to be exposed (Sundararajan and Bhushan 2001)

device does not exist in the literature. Using an AFM, measurements on various component surfaces were made for the first time by Sundararajan and Bhushan (2001). Table 16.11 shows various surface roughness parameters obtained from $5 \times 5 \mu\text{m}$ scans of the various component surfaces of several unlubricated micromotors using

the AFM in tapping mode. A surface with a Gaussian height distribution should have a skewness of zero and kurtosis of three. Although the rotor and stator top surfaces exhibit comparable roughness parameters, the underside of the rotors exhibits lower RMS roughness and P-V distance values. More importantly, the rotor

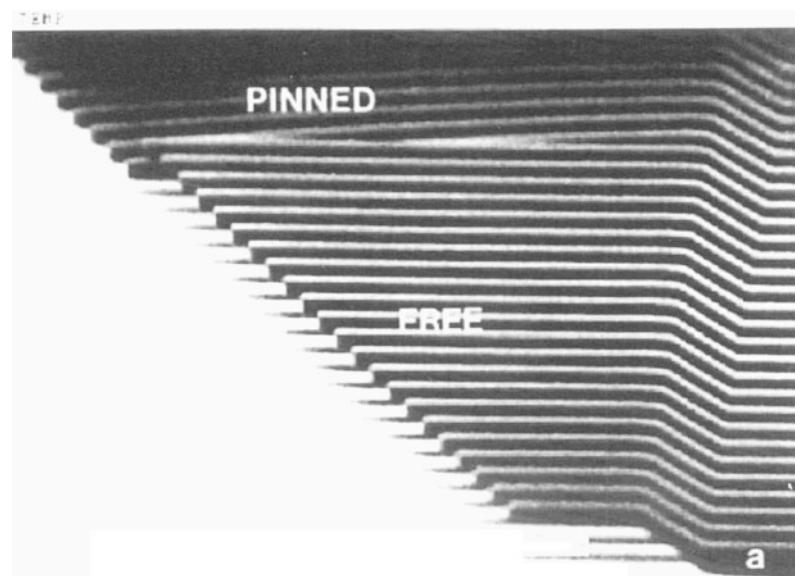
underside shows negative skewness and lower kurtosis than the topsides, both of which are conducive to high real area of contact and hence high friction (Bhushan 2013a, b). The rotor underside also exhibits a higher coefficient of microscale friction than the rotor topside and stator, as shown in Table 16.11. Figure 16.63 shows representative surface height maps of the various surfaces of a micromotor measured using the AFM in tapping mode. The rotor underside exhibits varying topography from the outer edge to the middle and inner edge. At the outer edges, the topography shows smaller circular asperities, similar to the topside. The middle and inner regions show deep pits with fine edges that may have been created by the etchants used for etching of the sacrificial layer. It is known that etching can affect the roughness of surfaces in surface micromachining. The residence time of the etchant near the inner region is high, responsible for larger pits. Figure 16.64 shows roughness of the surface directly beneath the rotors (the base polysilicon layer). There appears to be a difference in the roughness between the portion of this surface that was initially underneath the rotor (region B) during fabrication and the portion that was away from the rotor and hence always

exposed (region A). The former region shows lower roughness than the latter region. This suggests that the surfaces at the rotor-hub interface that come into contact at the end of the fabrication process exhibit large real areas of contact that result in high friction.

16.8.2 Adhesion Measurements of Microstructures

Surface force apparatuses (SFA) and AFMs are used to measure adhesion on micro- to nanoscales between two surfaces. In the SFA, adhesion of liquid films sandwiched between two curved and smooth surfaces is measured. In an AFM, as discussed earlier, adhesion between a sharp tip and the surface of interest is measured. The propensity of adhesion between two surfaces can be evaluated by studying the tendency of microstructures with well-defined contact areas covering a wide spectrum of suspension compliances to stick to the underlying substrate. The test structures that have been used include the cantilever beam array (CBA) technique with different lengths (Mastrangelo and Hsu 1993; Maboudian and Howe 1997; Mastrangelo 1998; De Boer and

Fig. 16.65 SEM micrograph of micromachined array of polysilicon cantilever beams of increasing length. The micrograph shows the onset of pinning for beams longer than 34 μm (Mastrangelo and Hsu 1993)



Michalske 1999) and stand-off multiple dimples mounted on microstructures with a range of compliances, standing above a substrate (Alley et al. 1992). The more-commonly used CBA technique uses an array of micromachined polysilicon beams (for Si MEMS applications) in the mesoscopic length scale that are anchored to the substrate at one end and with different lengths parallel to the surface. CBA relies on peeling and detachment of cantilever beams. Change in free energy or reversible work done to separate unit areas of two surfaces from contact is called work of adhesion. To measure the work of adhesion, electrostatic actuation is used to bring all beams in contact with the substrate, Fig. 16.65 (Mastrangelo and Hsu 1993; Mastrangelo 1998). Once the actuation force is removed, the beams begin to peel themselves off the substrate, which can be observed with an optical interference microscope (e.g., Wyko surface profiler). For beams shorter than a characteristic length, the so-called detachment length, their stiffness is sufficient to free them completely from the substrate underneath. Beams larger than the detachment length remain adhered. The beams at the transition region start to detach and remain attached to the substrate just

at the tips. For this case, by equating the elastic energy stored within the beam and the beam-substrate interfacial energy, the work of adhesion, W_{ad} , can be calculated by the following equation (Mastrangelo and Hsu 1993)

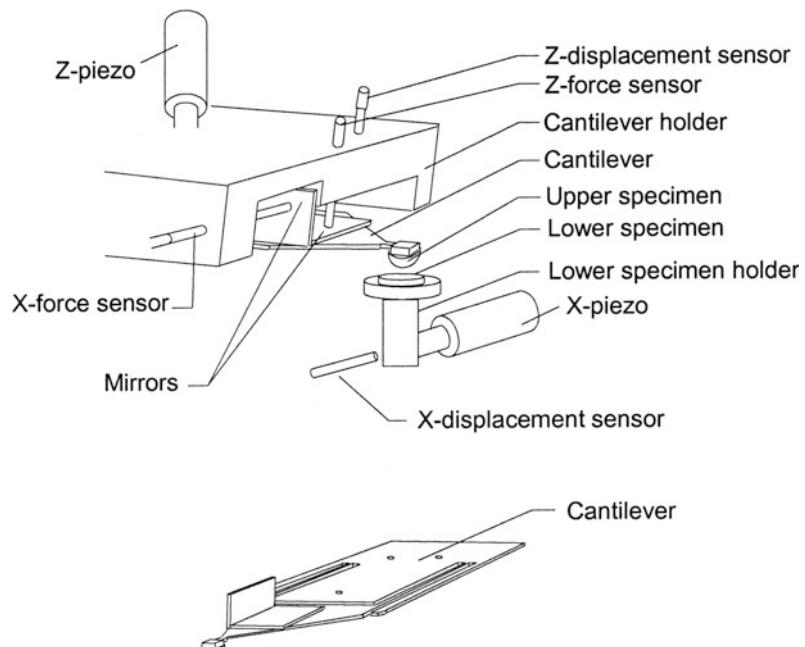
$$W_{ad} = \frac{3Ed^2t^3}{8\ell_d^4} \quad (16.1)$$

where E is the Young's modulus of the beam, d is the spacing between the undeflected beam and the substrate, t is the beam thickness, and P_d is the detachment length. The technique has been used to screen methods for adhesion reduction in polysilicon microstructures.

16.8.3 Microtriboapparatus for Adhesion, Friction, and Wear of Microcomponents

To measure adhesion, friction, and wear between two microcomponents, a microtriboapparatus has been used. Figure 16.66 shows a schematic of a microtriboapparatus capable of

Fig. 16.66 Schematic of the microtriboapparatus including specially designed cantilever (with two perpendicular mirrors attached on the end), lower specimen holder, two piezos (X- and Z-piezos), and four fiber optic sensors (X- and Z-displacement sensors and X- and Z-force sensors) (Liu and Bhushan 2003b)



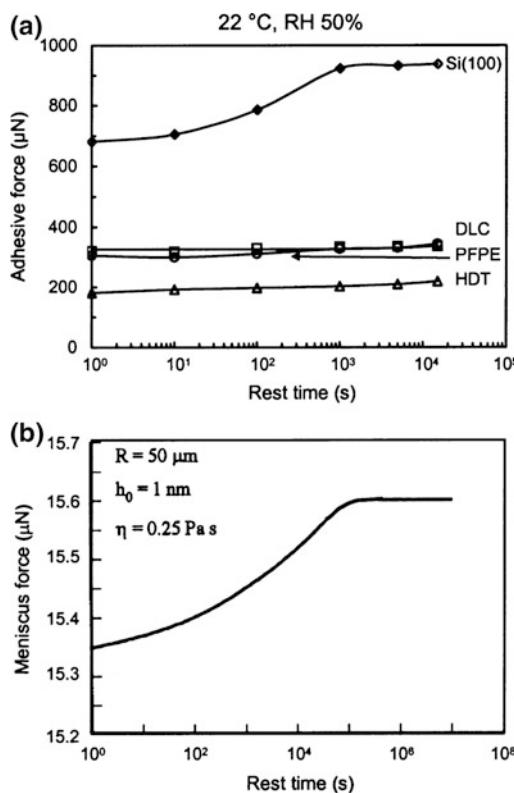


Fig. 16.67 **a** The influence of rest time on the adhesive force of Si(100), DLC, chemically bonded PFPE, and HDT, and **b** single asperity contact modeling results of the rest time effect on the meniscus force for an asperity of radius R in contact with a flat surface with a water film of thickness of h_0 and absolute viscosity of η_0 (Chilamakuri and Bhushan 1999)

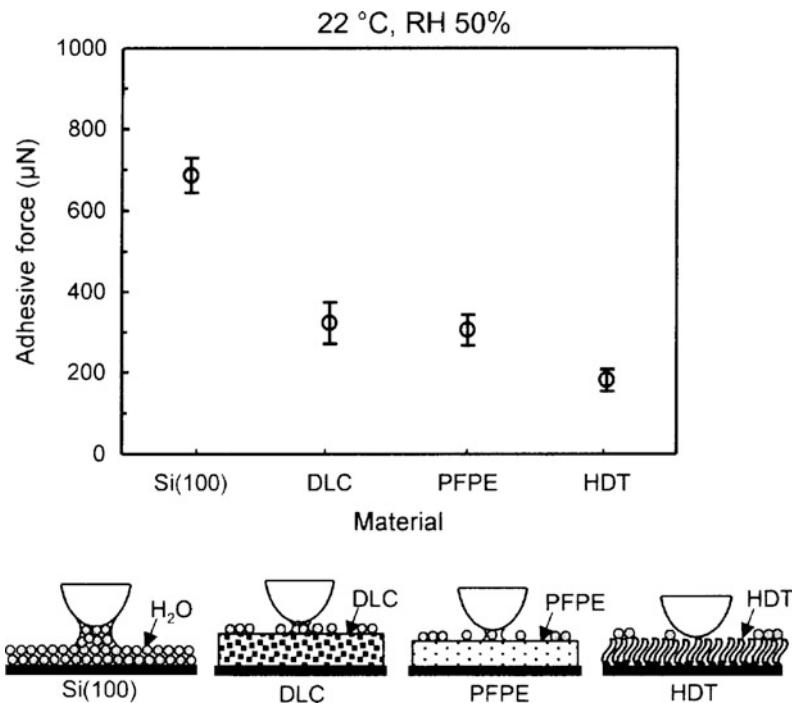
using MEMS components for tests (Liu and Bhushan 2003b). In this apparatus, an upper specimen, mounted on a soft cantilever beam, comes in contact with a lower specimen mounted on a lower specimen holder. The apparatus consists of two piezos (X- and Z-piezos) and four fiber optic sensors (X- and Z-displacement sensors, and X- and Z-force sensors). For adhesion and friction studies, the Z-piezo is used to bring the upper specimen and lower specimen into contact, and the X-piezo is used to apply a relative motion in the lateral direction. The X- and Z-displacement sensors are used to measure the lateral position of the lower specimen and vertical position of the upper specimen, respectively. By monitoring the

deflection of the cantilever, the X- and Z-force sensors are used to measure friction force and normal load/adhesive force between the two specimens, respectively.

As most MEMS/NEMS devices are fabricated from silicon, study of silicon-on-silicon contacts is important. This contact was simulated by a flat single-crystal Si(100) wafer (phosphorus doped) specimen sliding against a single crystal Si(100) ball (1 mm in diameter, 5×10^{17} atoms/cm³ boron doped) mounted on a stainless steel cantilever (Liu and Bhushan 2003b; Bhushan et al. 2004). Both of them have a native oxide layer on their surfaces. With an objective to reduce adhesion and friction and humidity effects, the other materials studied were 10-nm thick DLC deposited by filtered cathodic arc deposition on Si(100), 2.3-nm thick chemically bonded PFPE (Z-DOL, BW) on Si(100), and hexadecane thiol (HDT) monolayer on evaporated Au(111) film to investigate their anti-adhesion performance. For completeness, other approaches to reduce liquid mediated adhesion include the formation of dimples on the contact surfaces to reduce contact area (Mastrangelo 1998; Bhushan 1998, 2008a, b, 2011, 2013a, b) and bioinspired superhydrophobic surfaces (Bhushan 2016b).

It is well known that in computer rigid disk drives, the adhesive force increases rapidly with an increase in rest time between a magnetic head and a magnetic disk (Bhushan 1996a). Considering that adhesion and friction are the major issues that lead to the failure of MEMS/NEMS devices, the rest time effect on the microscale for Si(100), DLC, PFPE, and HDT was studied, and the results are summarized in Fig. 16.67a. It is found that the adhesive force of Si(100) logarithmically increases with the rest time to a certain equilibrium time ($t = 1000$ s), after which it remains constant. Figure 16.67a also shows that the adhesive force of DLC, PFPE, and HDT does not change with rest time. Single asperity contact modeling of the dependence of meniscus force on the rest time has been carried out by Chilamakuri and Bhushan (1999), and the modeling results (Fig. 16.67b) verify experimental observations. Due to the presence of a thin film of

Fig. 16.68 Adhesive forces of Si(100), DLC, chemically bonded PFPE, and HDT at ambient condition and a schematic showing the relative size of water meniscus on different specimens



water adsorbed on Si(100), a meniscus forms around the contacting asperities and grows with time until equilibrium occurs, which causes the rest time effect on its adhesive force. The adhesive forces of DLC, PFPE, and HDT do not change with rest time which suggests that the water meniscus is not present on their surfaces.

The measured adhesive forces of Si(100), DLC, PFPE, and HDT at rest time of 1 s are summarized in Fig. 16.68. It shows that the presence of solid films of DLC, PFPE, and HDT greatly reduced the adhesive force of Si(100), whereas HDT film had the lowest adhesive force. It is well known that the native oxide layer (SiO_2) on the top of a Si(100) wafer exhibits hydrophilic properties, and water molecules, produced by capillary condensation of water vapor from the environment, can easily be adsorbed on this surface. The condensed water will form a meniscus as the upper specimen approaches the lower specimen surface. The meniscus force is a major contributor to the adhesive force. In the case of DLC, PFPE, and HDT, the films were found to be hydrophobic based on contact angle measurements, and the

amount of condensed water vapor was low as compared to that on Si(100). It should be noted that the measured adhesive force is generally higher than that measured in AFM, because the larger radius of Si(100) ball as compared to that of an AFM tip induces larger meniscus and van der Waals forces.

To investigate the velocity effect on friction, the friction force as a function of velocity was measured and is summarized in Fig. 16.69a. For Si(100), the friction force initially decreases with increasing velocity until equilibrium occurred. Figure 16.69a also shows that the velocity had almost no effect on the friction properties of DLC, PFPE, and HDT. This implies that the friction mechanisms of DLC, PFPE, and HDT do not change with the variation of velocity. For Si(100), at high velocity, the meniscus is broken and does not have enough time to rebuild. In addition, it is also believed that tribochemical reaction plays an important role. The high velocity leads to tribochemical reactions of Si(100) (which has native oxide SiO_2) with water molecules to form a $\text{Si}(\text{OH})_4$ film. This film is removed and continuously replenished during

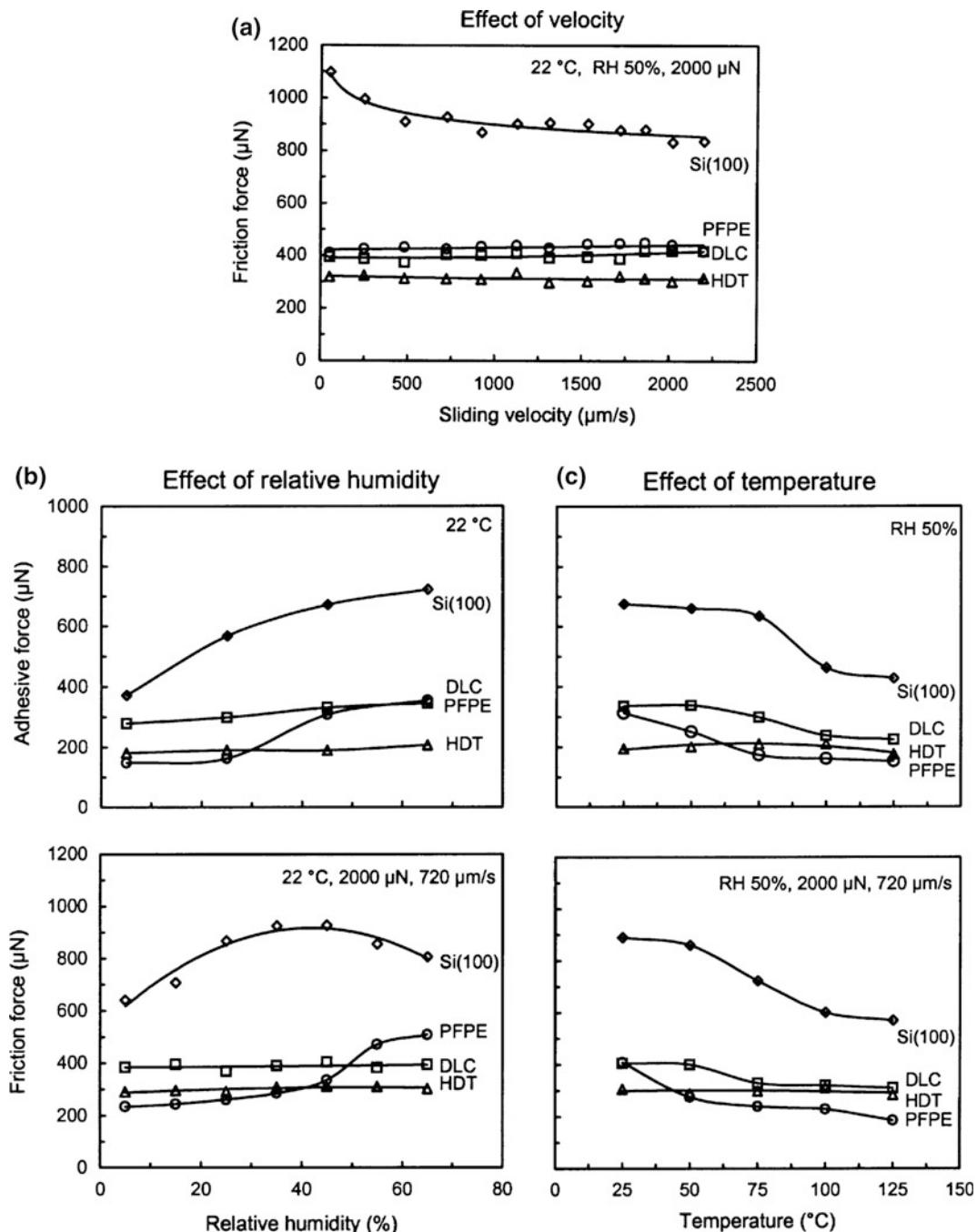


Fig. 16.69 The influence of **a** sliding velocity on the friction forces, **b** relative humidity on the adhesive and friction forces, and **c** temperature on the adhesive and friction forces of Si(100), DLC, chemically bonded PFPE, and HDT

sliding. The $\text{Si}(\text{OH})_4$ layer at the sliding surface is known to be of low shear strength. The breaking of the water meniscus and the formation

of a $\text{Si}(\text{OH})_4$ layer results in a decrease in friction force of Si(100). For DLC, PFPE, and HDT, their surfaces exhibit hydrophobic properties, and only

can adsorb a few water molecules in ambient conditions. The above mentioned meniscus breaking and tribochemical reaction mechanisms do not exist for these films. Therefore, their friction force does not change with velocity.

The influence of relative humidity was studied in an environmentally controlled chamber. The adhesive force and friction force were measured by making measurements at increasing relative humidity, and the results are summarized in Fig. 16.69b. It shows that for Si(100), the adhesive force increased with relative humidity, but the adhesive force of DLC and PFPE only showed a slight increase when humidity was higher than 45%, while the adhesive force of HDT did not change with humidity. Figure 16.69b also shows that for Si(100), the friction force increased with an increase in relative humidity up to 45%, and then it shows a slight decrease with a further increase in the relative humidity. For PFPE, there was an increase in the friction force when humidity was higher than 45%. In the whole testing range, relative humidity did not have any apparent influence on the friction properties of DLC and HDT. In the case of Si(100), the initial increase of relative humidity up to 45% caused more adsorbed water molecules, and formed bigger water meniscus, which led to an increase of friction force. But at very high humidity of 65%, large quantities of adsorbed water can form a continuous water layer that separates the tip and sample surfaces, and acts as a kind of lubricant, which causes a decrease in the friction force. For PFPE, dewetting of the lubricant film at humidity higher than 45% resulted in an increase in adhesive and friction forces. For DLC and HDT, their surfaces show hydrophobic properties, and increasing relative humidity did not play much of a role on their friction force.

The influence of temperature was studied using a heated stage. The adhesive force and friction force were measured by making measurements at increasing temperatures from 22 to 125 °C. The

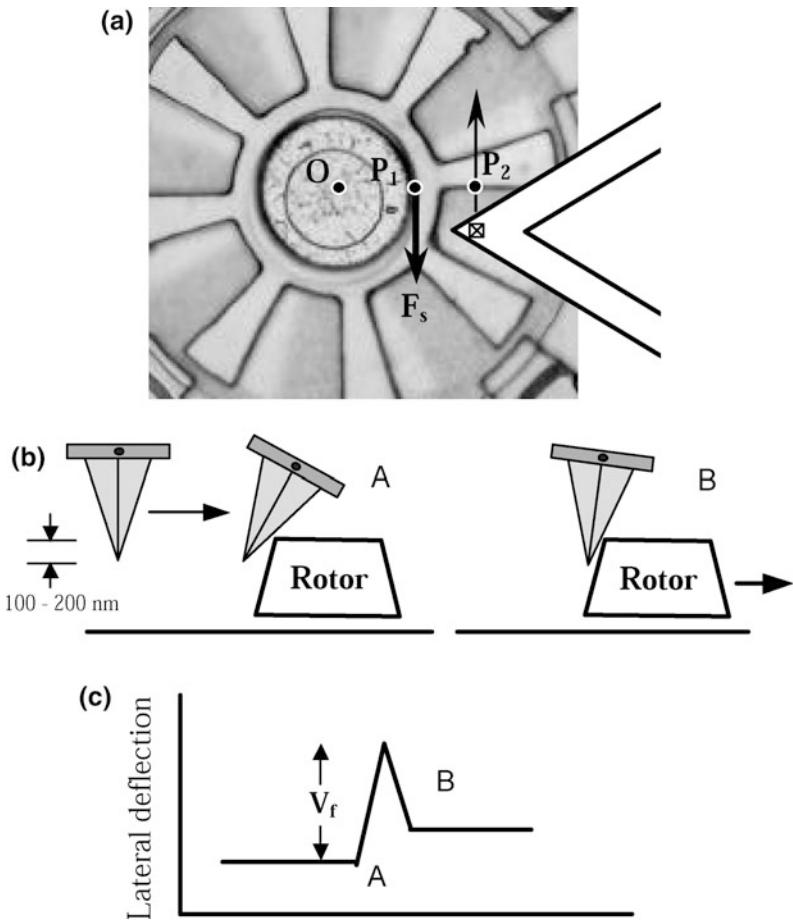
results are presented in Fig. 16.69c. It shows that once the temperature was higher than 50 °C, increasing temperature caused a significant decrease of adhesive and friction forces of Si(100) and a slight decrease in the case of DLC and PFPE. But the adhesion and friction forces of HDT did not show any apparent change with test temperature. At high temperature, desorption of water and reduction of the surface tension of water led to the decrease of adhesive and friction forces of Si(100), DLC, and PFPE. However, in the case of HDT film, as only a few water molecules were adsorbed on the surface, the above-mentioned mechanisms did not play a big role. Therefore, the adhesive and friction forces of HDT did not show any apparent change with temperature. Figure 16.69 shows that in the whole velocity, relative humidity, and temperature test range, the adhesive force and friction force of DLC, PFPE, and HDT were always smaller than that of Si(100), whereas HDT had the smallest value.

To summarize, several methods can be used to reduce adhesion in microstructures. MEMS/NEMS surfaces can be coated with hydrophobic coatings such as PFPEs, SAMs, and passivated DLC coatings.

16.8.4 Static Friction Force (Stiction) Measurements in MEMS

In MEMS devices involving parts in relative motion to each other, such as micromotors, large friction forces become the limiting factor to the successful operation and reliability of the device. It is generally known that most micromotors cannot be rotated as manufactured and require some form of lubrication. Therefore, it is critical to determine the friction forces present in such MEMS devices. To measure in situ the static friction of a rotor-bearing interface in a micro-motor, Tai and Muller (1990) measured the starting torque (voltage) and pausing position for different starting positions under a constant-bias

Fig. 16.70 **a** Schematic of the technique used to measure the force, F_s , required to initiate rotor movement using an AFM/FFM. **b** As the tip is pushed against the rotor, the lateral deflection experienced by the rotor due to the twisting of the tip prior to rotor movement is a measure of static friction force, F_s , of the rotors. **c** Schematic of lateral deflection expected from the above-mentioned experiment. The peak V_f is related to the state of the rotor (Sundararajan and Bhushan 2001)



voltage. A friction-torque model was used to obtain the coefficient of static friction. To measure the in situ kinetic friction of the turbine and gear structures, Gabriel et al. (1990) used a laser-based measurement system to monitor the steady-state spins and decelerations. Lim et al. (1990) designed and fabricated a polysilicon microstructure to measure in situ the static friction of various films. The microstructure consisted of shuttle suspended above the underlying electrode by a folded beam suspension. A known normal force was applied, and lateral force was measured to obtain the coefficient of static friction. Beerschwingen et al. (1994) developed a cantilever-deflection rig to measure friction of LIGA-processed micromotors (see also, Matheison et al. 1996.) These techniques employ indirect methods to determine the friction forces or involve fabrication of complex structures.

A novel technique to measure the static friction force (stiction) encountered in surface micromachined polysilicon micromotors using an AFM was developed by Sundararajan and Bhushan (2001). Continuous physical contact occurs during rotor movement (rotation) in the micromotors between the rotor and lower hub flange. In addition, contact occurs at other locations between the rotor and the hub surfaces and between the rotor and the stator. Friction forces will be present at these contact regions during motor operation. Although the actual distribution of these forces is not known, they can be expected to be concentrated near the hub where there is continuous contact. If we therefore represent the static friction force of the micromotor as a single force F_s acting at point P_1 (as shown in Fig. 16.70a), then the magnitude of the frictional torque about the center of the motor

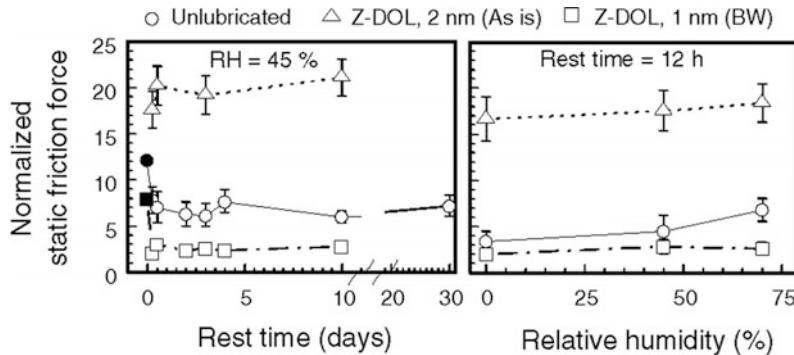


Fig. 16.71 Static friction force values of unlubricated motors and motors lubricated using PFPE lubricants, normalized over the rotor weight, as a function of rest time and relative humidity. Rest time is defined as the time elapsed between a given experiment and the first experiment in which motor movement was recorded (time 0). The motors were allowed to sit at a particular humidity for 12 h prior to measurements (Sundararajan and Bhushan 2001)

(O) that must be overcome before rotor movement can be initiated is

$$T_s = F_s \ell_1 \quad (16.2)$$

where ℓ_1 is the distance OP_1 , which is assumed to be the average distance from the center at which the friction force F_s occurs. Now, consider an AFM tip moving against a rotor arm in a direction perpendicular to the long axis of the cantilever beam (the rotor arm edge closest to the tip is parallel to the long axis of the cantilever beam), as shown in Fig. 16.70a. When the tip encounters the rotor at point P_2 , the tip will twist, generating a lateral force between the tip and the rotor (event A in Fig. 16.70b). This reaction force will generate a torque about the center of the motor. Since the tip is trying to move further in the direction shown, the tip will continue to twist to a maximum value at which the lateral force between the tip and the rotor becomes high enough such that the resultant torque T_f about the center of the motor equals the static friction torque T_s . At this point, the rotor will begin to rotate, and the twist of the cantilever decreases sharply (event B in Fig. 16.70b). The twist of the cantilever is measured in the AFM as a change in the lateral deflection signal (in volts), which is the underlying concept of friction force microscopy (FFM). The change in the lateral deflection signal corresponding to the above-mentioned

events as the tip approaches the rotor is shown schematically in Fig. 16.70c. The value of the peak V_f is a measure of the force exerted on the rotor by the tip just before the static friction torque is matched and the rotor begins to rotate.

Using this technique, the viability of PFPE lubricants for micromotors has been investigated, and the effect of humidity on the friction forces of unlubricated and lubricated devices was studied as well. Figure 16.71 shows static friction forces, normalized over the weight of the rotor, of unlubricated and lubricated micromotors as a function of rest time and relative humidity. Rest time here is defined as the time elapsed between the first experiment conducted on a given motor (solid symbol at time zero) and subsequent experiments (open symbols). Each open symbol data point is an average of six measurements. It can be seen that for the unlubricated motor and the motor lubricated with a bonded layer of Z-DOL(BW), the static friction force was highest for the first experiment, and then dropped to an almost constant level. In the case of the motor with an as-is mobile layer of Z-DOL, the values remained very high up to 10 days after lubrication. In all cases, there was negligible difference in the static friction force at 0 and 45% RH. At 70% RH, the unlubricated motor exhibited a substantial increase in the static friction force, while the motor with bonded Z-DOL showed no increase due to the

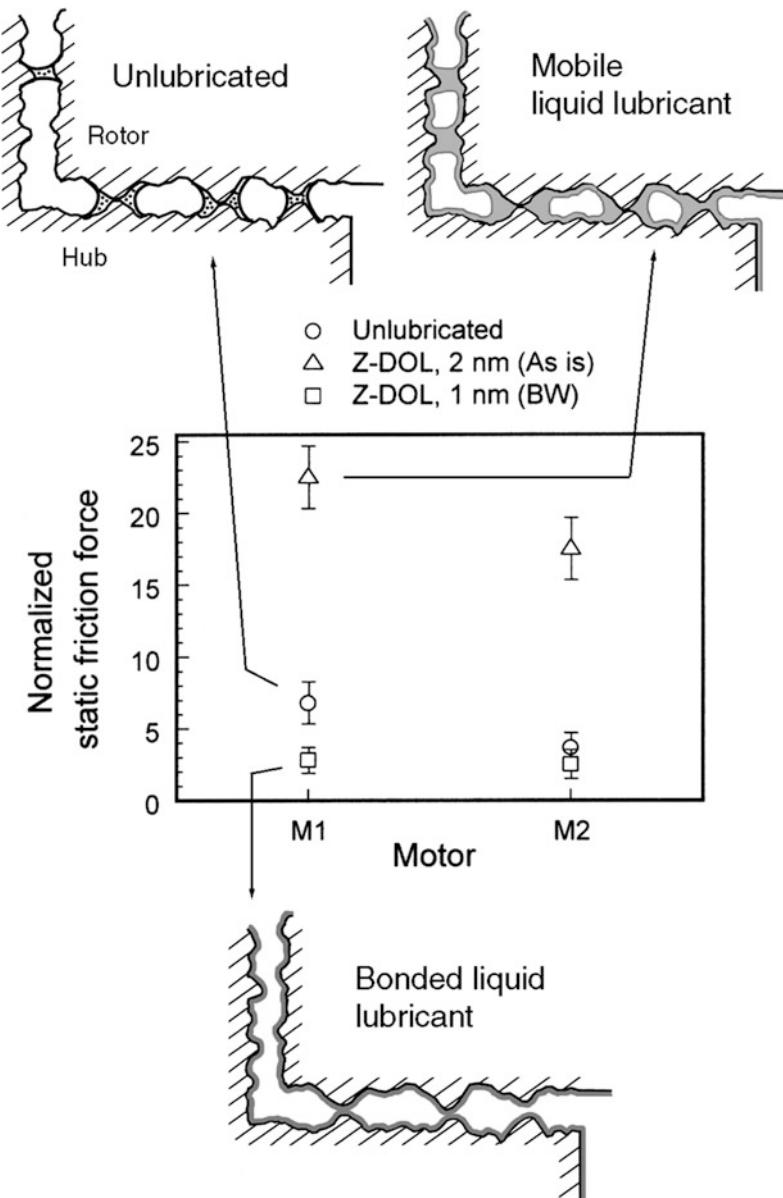


Fig. 16.72 Summary of effect of liquid and solid lubricants on static friction force of micromotors. Despite the hydrophobicity of the lubricant used (Z-DOL), a mobile liquid lubricant (Z-DOL as-is) leads to very high static friction force due to increased meniscus forces whereas a solid-like lubricant (bonded Z-DOL, BW) appears to provide some amount of reduction in static friction force

hydrophobicity of the lubricant layer. The motor with an as-is mobile layer of the lubricant showed consistently high values of static friction force that varied little with humidity.

Figure 16.72 summarizes static friction force data for two motors, M1 and M2, along with

schematics of the meniscus effects for the unlubricated and lubricated surfaces. Capillary condensation of water vapor from the environment resulted in the formation of meniscus bridges between the contacting and near-contacting asperities of two surfaces in close proximity to

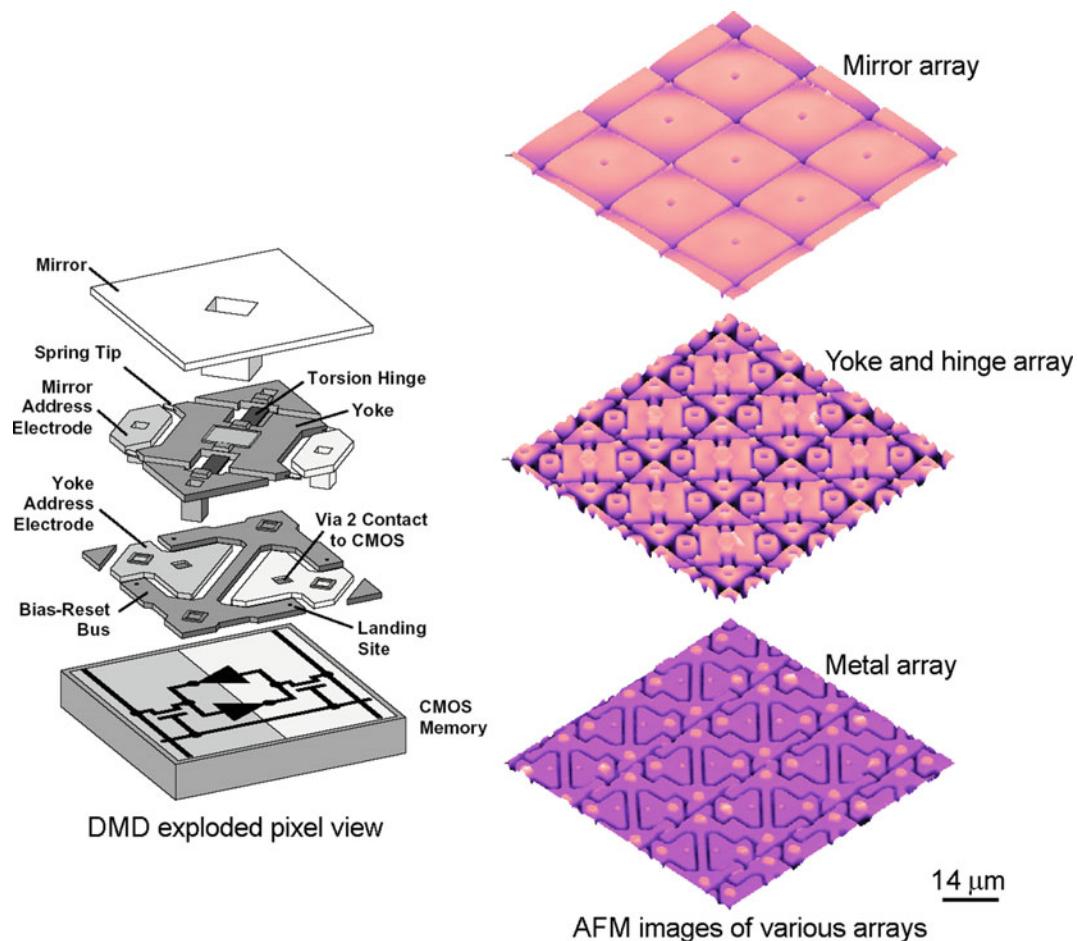


Fig. 16.73 Exploded view of a DMD pixel and AFM surface height images of various arrays. The DMD layers were removed by ultrasonic method (Liu and Bhushan 2004a)

each other, as shown in Fig. 16.72. For unlubricated surfaces, more menisci were formed at higher humidity resulting in higher friction force between the surfaces. The formation of meniscus bridges is supported by the fact that the static friction force for unlubricated motors increases at high humidity (Fig. 16.72). Solid bridging may occur near the rotor-hub interface due to silica residues after the first etching process. In addition, the drying process after the final etch can result in liquid bridging formed by the drying liquid due to meniscus force at these areas (Mastrangelo and Hsu 1993; Maboudian and Howe 1997; Bhushan 2013a, b). Therefore, the initial static friction force will be quite high, as evidenced by the solid data

points in Fig. 16.72. Once the first movement of the rotor permanently breaks these solid and liquid bridges, the static friction force of the motors will drop (as seen in Fig. 16.72) to a value dictated predominantly by the adhesive energies of rotor and hub surfaces, the real area of contact between these surfaces and meniscus forces due to water vapor in the air. At that point, the effect of lubricant films can be observed. Lubrication with a mobile layer, even a thin one, results in very high static friction forces due to meniscus effects of the lubricant liquid itself at and near the contact regions. It should be noted that a motor submerged in a liquid lubricant would result in a fully flooded lubrication regime. In this case, there is no

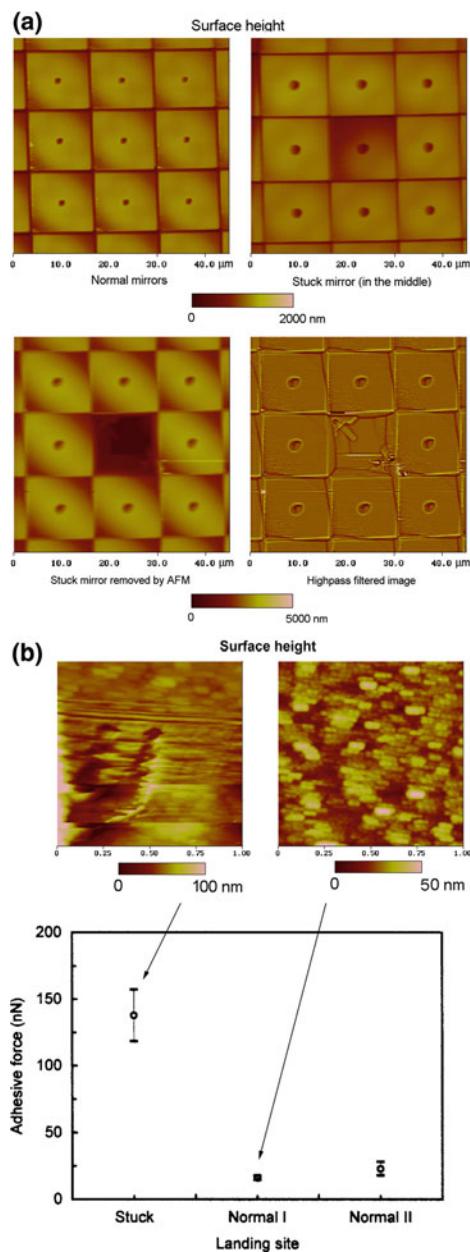


Fig. 16.74 **a** The top row shows AFM surface height images of a stuck micromirror surrounded by eight normal micromirrors. Left image in the bottom row shows the stuck micromirror which was removed by an AFM tip after repeated scanning at high normal load. The right image in the bottom row presents the highpass filtered image showing the residual hinge which sits underneath the removed micromirror is clearly observed. **b** AFM surface height images and adhesive forces of the landing sites underneath the two normal micromirrors and the stuck micromirror (Liu and Bhushan 2004a)

meniscus contribution, and only the viscous contribution to the friction forces would be relevant. However, submerging the device in a lubricant may not be a practical method. A solid-like hydrophobic lubricant layer (such as bonded Z-DOL) results in favorable friction characteristics for the motor. The hydrophobic nature of the lubricant inhibits meniscus formation between the contact surfaces and maintains low friction even at high humidity (Fig. 16.72). This suggests that solid-like hydrophobic lubricants are ideal for lubrication of MEMS, while mobile lubricants result in increased values of static friction force.

16.8.5 Mechanisms Associated with Observed Stiction Phenomena in Digital Micromirror Devices (DMD) and Nanomechanical Characterization

DMDs are used in digital projection displays, as described earlier. The DMD has a layered structure, consisting of an aluminum alloy micromirror layer, yoke and hinge layer, and metal layer on a CMOS memory array (Hornbeck and Nelson 1988; Hornbeck 1999, 2001). A blown-up view of the DMD and the corresponding AFM surface height images are presented in Fig. 16.73 (Liu and Bhushan 2004a). Single-layered aluminum alloy films are used for the construction of micromirrors. Sometimes these also are used for the construction of hinges, spring tips, and landing sites. The aluminum alloy films are overwhelmingly comprised of aluminum; trace elements (including Ti and Si) are present to suppress contact spiking and electromigration, which may occur if current densities become high during electrostatic operation. Multilayered sputtered $\text{SiO}_2/\text{TiN}/\text{Al}$ alloy films are now generally used for the landing site structure to minimize refraction throughout the visible region of the electromagnetic spectrum in order to increase the contrast ratio in projection display systems (Wei et al. 2004a, b). These multilayered films generally also are used for

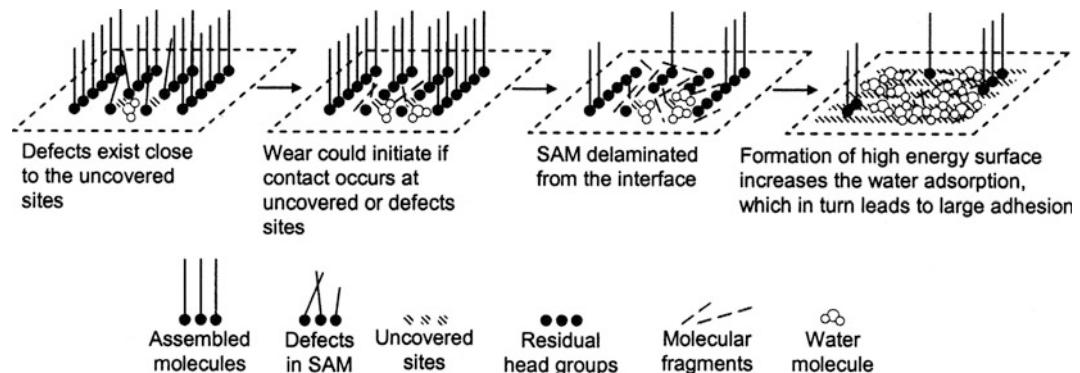


Fig. 16.75 Suggested mechanisms for wear and stiction (Liu and Bhushan 2004a)

hinges and spring tips. A low surface energy SAM is maintained on the surfaces of the DMD, which is packaged in a hermetic environment to minimize stiction during contact between the spring tip and the landing site. A SAM of perfluorinated n-alkanoic acid ($C_nF_{2n-1}O_2H$) (e.g., perfluorodecanoic acid or PFDA, $CF_3(CF_2)_8COOH$) applied by the vapor phase deposition process is used. A getter strip of PFDA is included inside the hermetically sealed enclosure containing the chip, which acts as a reservoir in order to maintain a PFDA vapor within the package.

In order to identify a stuck mirror and characterize its nanotribological properties, the chip was scanned using an AFM (Liu and Bhushan 2004a). It was found that it is hard to tilt the stuck micromirror back to its normal position by adding a normal load at the rotatable corner of the micromirror; thus, called a “hard” stuck micromirror. An example of a stuck micromirror is shown in Fig. 16.74a. Once the stuck micromirror was found, the region was repeatedly scanned at a large normal load, up to 300 nN. After several scans, the stuck micromirror was removed. Once the stuck micromirror was removed, the surrounding micromirrors also could be removed by continuous scanning under a large normal load (Fig. 16.74a bottom row). The adhesive force of the landing site underneath the stuck micromirror

and the normal micromirror are presented in Fig. 16.74b. It clearly indicates that the landing site underneath the stuck micromirror had much larger adhesion. 1 $\mu m \times 1 \mu m$ view of landing sites under stuck and normal micromirrors are also shown in Fig. 16.74b. The landing site under the stuck micromirror had an apparent U-shaped wear mark, which was surrounded by a smeared area.

Liu and Bhushan (2004a) calculated contact stresses to examine whether the stresses were high enough to cause wear at the spring tip landing site interface. The calculated contact stress value was about 33 MPa which is substantially lower than the hardness. Therefore, not much plastic deformation and, consequently, wear was expected. A wear mark was only found on a very few landing sites on the DMD, which means that the SAM coating generally can endure such high contact stresses. Based on data reported in the literature, coverage for vapor-deposited SAMs is expected to be about 97%. The bond strength of the molecules close to the boundary of the uncovered sites is expected to be weak. Thus, the uncovered sites and the adjacent molecules are referred to as defects in the SAM coating. Occasionally, if contact occurs at the defect sites, the large cyclic stress may be close to the critical load, and lead to the initial delamination of the SAM coating at the interface.

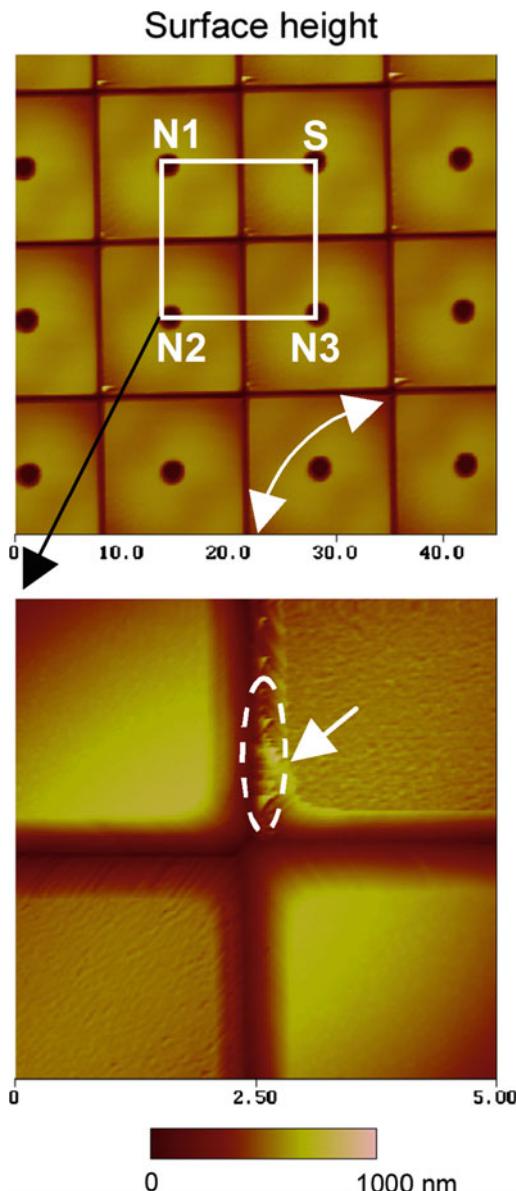


Fig. 16.76 AFM surface height images of normal micromirrors and a soft stuck micromirror. The soft stuck micromirror was labeled as S, and the normal micromirrors studied are labeled as N1, N2, and N3 (Liu and Bhushan 2004b)

Continuous contact leads to the formation of a high surface energy surface by exposure of the fresh substrate and formation of SAM fragments. This eventually leads to an increase in stiction by the formation of large menisci. Once this happens, the stress at the contact area is increased,

which would accelerate the wear. Based on this hypothesis, suggested mechanisms for the wear and stiction of the landing site are summarized in Fig. 16.75. Wear initiates at the defect sites and consequently high stiction can result in high wear. Improving the coverage and wear resistance of SAM coatings could enhance the yield of DMD.

In some cases, the micromirrors are not fully stuck, and can be moved by applying a load at the rotatable corner of the micromirror with a discontinuous motion, thus called “soft” stiction. Soft stuck micromirrors studied by Liu and Bhushan (2004b) were identified in quality inspection. These micromirrors encountered slow transition from one end to the other end (+1/-1). Figure 16.76 shows the AFM surface height images of a location showing a stuck mirror (S) and surrounding normal micromirrors Ni ($i = 1, 2$, and 3). Surprisingly, the images of the stuck and normal micromirror array are almost the same. On the micromirrors of interest, a tilting test was performed at the corner of the micromirrors, the rotatable direction of the microarray is indicated by an arrow bar in Fig. 16.76. The load displacement curve for the stuck micromirror is presented in Fig. 16.77. The curve is not smooth and appears serrated. It clearly indicated that, although the S micromirror can be rotated, it rotates with hesitation. In regimes 1 and 2, as marked in Fig. 16.77, the slopes are much higher. In order to understand the occurrence mechanisms of stiction, stiction of the landing sites of normal and stuck mirrors were measured. Unlike a hard stuck mirror, adhesive forces of soft stuck and normal mirrors were comparable, which suggests that the SAM coating is intact with the soft-stuck mirror. It was found that a high normal load (about 900 nN) and on the order of couple of hundred scans were required to remove the soft stuck micromirrors by an AFM. For comparison, only about 300 nN and about ten scans were required to remove hard stuck mirror. After careful examination of the AFM images of the micromirror sidewalls in Fig. 16.76 (bottom left), it is noted that there are contaminant particles attached to the sidewalls of the S mirror. Therefore, it is believed that during

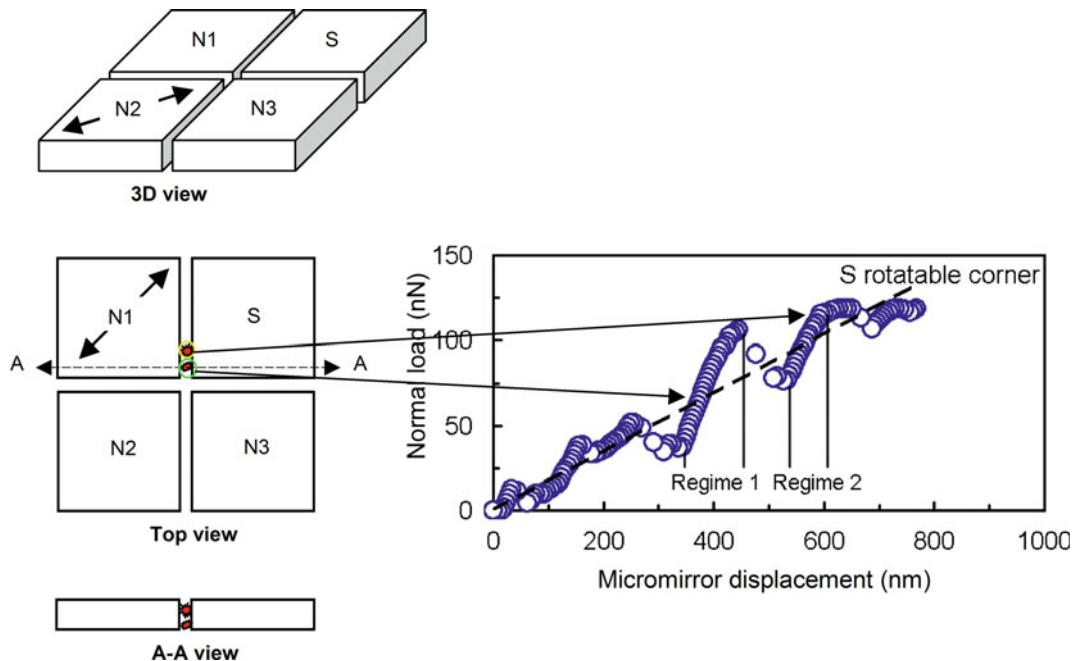
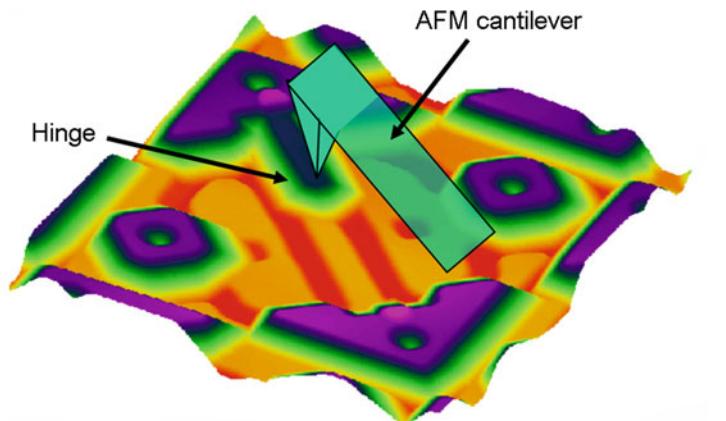


Fig. 16.77 Load-displacement curve obtained on the rotatable corner of S micromirror and schematic to illustrate the suggested mechanism for soft stiction occurrence

Fig. 16.78 AFM surface height image of the residual hinge and schematic diagram of relative position of the hinge and AFM tip during the nanoscale bending and fatigue tests. The tip is located at the free end of the hinge (Liu and Bhushan 2004c)



Hinge length = 3900 nm, width = 780 nm, and thickness = 100 nm

Residual hinge (yoke is removed)

the tilting test for the S micromirror (see schematic in Fig. 16.77), a sharper slope regime will occur in the displacement curve. Extra force is required to overcome the resistance that is induced by the sidewall contamination particles. This is believed to be the reason for the slow

transition of the micromirror during quality inspection.

Finally, nanomechanical characterization of various layers used in the construction of landing sites, hinge, and micromirror materials have been measured by Wei et al. (2004a, b).

Bending and fatigue studies of the hinge have been carried out by Liu and Bhushan (2004c) and Bhushan and Liu (2004) to measure stiffness and fatigue properties. For these studies, the micromirror was removed. During removal, the micromirror/yoke structure was removed simultaneously, leaving the hinge mounted on one end of the array, see Fig. 16.78. The stiffness of the Al hinge was reported to be comparable to the stiffness of bulk Al. The Al hinge exhibited higher modulus than the SiO₂ hinge. The fatigue properties depended upon the preparation of hinge for testing.

Closure

The field of MEMS/NEMS and BioMEMS/BioNEMS has expanded considerably since its inception in the late 1980s. Large surface-to-volume ratio of the devices result in very high surface forces, such as adhesion and friction/stiction that seriously undermine the performance and reliability of the devices. There is a need for a fundamental understanding of adhesion, friction, stiction, wear, and lubrication, and the role of contamination and environment, all on the nanoscale. Most mechanical properties are known to be scale dependent. Therefore, the properties of nanostructures need to be measured. Using AFM-based techniques, researchers have conducted nanotribological and nanomechanics studies of materials and devices. In addition, component-level testing has also been carried out to aid in better understanding of the observed nanotribological phenomena in MEMS/NEMS.

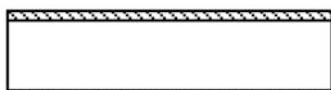
Macroscale and microscale tribological studies of silicon and polysilicon films have been performed. The effect of doping and oxide films and environment on the tribological properties of these popular MEMS/NEMS materials have also been studied. SiC film is found to be a good candidate material for use in high-temperature MEMS/NEMS devices. Perfluoroalkyl SAMs and bonded perfluoropolyether lubricants are well-suited for lubrication of micro/nanodevices under a range of environmental conditions. DLC coatings can also be used for low friction and wear.

Nano-objects exhibit enhanced mechanical, electrical, magnetic, and chemical properties. Many nano-objects also exhibit low friction and wear properties due to their size, shape, and unique properties. Nano-objects of interest include Au nanoparticles, and MoS₂, and WS₂ MWNTs, and CNHs. Nanoscale friction and wear properties have been investigated in single nano-object contact and multiple nano-object contact modes using an AFM. All nano-objects exhibit low friction and wear. Lower friction forces occur in submerged-in-liquid environments, as compared to dry environments.

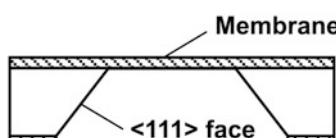
For BioMEMS/BioNEMS, nanoscale adhesion, friction, and wear of a number of protein layers on polystyrene surfaces have been investigated. Adhesion between biological molecular layers and the substrate, and friction and wear of biological layers can be important. Adhesion of biomolecules on Si substrate surfaces for various BioMEMS applications can be improved by nanopatterning and chemical linker method. Friction and wear mechanisms of protein layers have been studied. Trajectory of submicron particles in the blood stream used for therapeutic and diagnostics purposes needs to be optimized in order for them to bond to the target sites on endothelium wall. An analytical model has been developed for this purpose.

Surface roughness measurements of micro-machined polysilicon surfaces have been made using an AFM. The roughness distribution on surfaces is strongly dependent upon the fabrication process. Adhesion and friction of microstructures can be measured using a novel microtriboapparatus. Adhesion and friction measurements on silicon-on-silicon confirm AFM measurements that hexadecane thiol and bonded perfluoropolyether films exhibit superior adhesion and friction properties. Static friction force measurements of micromotors have been performed using an AFM. The forces are found to vary considerably with humidity. A bonded layer of perfluoropolyether lubricant is found to satisfactorily reduce the friction forces in the micro-motor. Tribological failure modes of digital micromirror devices are either “hard” stiction or “soft” stiction. In hard stiction, the tip on the yoke

Bulk Micromachining



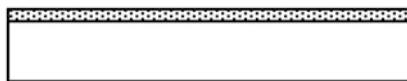
Deposition of silica layers on Si



Patterning with mask and etching of Si to produce cavity

□ Silicon ■ Silica

Surface Micromachining



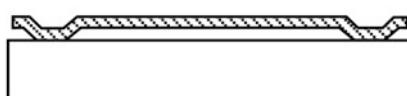
Deposition of sacrificial layer



Patterning with mask



Deposition of microstructure layer



Etching of sacrificial layer to produce freestanding structure

□ Silicon ■ Polysilicon
■ Sacrificial material

Fig. 16.79 Schematic of process steps involved in bulk micromachining and surface micromachining fabrication of MEMS

remains stuck with the landing site underneath. The mechanism responsible for the hard stiction is the localized damage to the SAM on the landing site. Whereas in soft stiction, the mirror-yoke assembly rotates with hesitation. The mechanism responsible for the soft stiction is contaminant particles present at the mirror sidewalls.

AFM/FFM-based techniques show the capability to study and evaluate nanotribology and nanomechanics related to MEMS/NEMS and BioMEMS/BioNEMS devices.

Appendix: Micro/Nanofabrication Techniques

Micro/nanofabrication techniques include top-down methods, in which one builds down from the large to the small, and the bottom-up, in which one builds up from the small to the large.

Top-Down

The top-down fabrication methods used in the construction of MEMS/NEMS include lithographic and nonlithographic techniques to produce micro- and nanostructures. The lithographic techniques fall into three basic categories: bulk micromachining, surface micromachining, and LIGA (a German acronym for Lithographie Galvanoformung Abformung), which translates to lithography, electroplating, and molding. The first two approaches, bulk and surface micromachining, mostly use planar photolithographic fabrication processes developed for semiconductor devices to produce two-dimensional (2D) structures (Jaeger 1988; Bhushan 1998; Voldman et al. 1999; Judy 2001; Madou 2011; Liu 2006). The various steps involved in these two fabrication processes are shown schematically in Fig. 16.79. Bulk micromachining employs anisotropic etching to remove sections through the thickness of a single-crystal silicon wafer, typically 250–500 μm thick. Bulk micromachining is a proven high-volume production process and is routinely used to fabricate microstructures such as accelerometers, pressure sensors, and flow

sensors. In surface micromachining, structural and sacrificial films are alternatively deposited, patterned, and etched to produce a freestanding structure. These films are typically made of low-pressure chemical vapor deposition (LPCVD) polysilicon film with 2–20 μm thickness. Surface micromachining is used to produce sensors, actuators, micromirror arrays, motors, gears, and grippers. The resolution in photolithography is dependent upon the wavelength of light. A commonly-used light source is an argon fluoride excimer laser with 193 nm wavelength (ultraviolet or UV) used in patterning 90 nm lines and spaces. Deep-UV wavelengths, X-ray lithography, electron beam (e-beam) lithography, focused ion beam lithography, maskless lithography, liquid-immersion lithography, and STM writing by removing material atom by atom are some of the recent developments for sub 100 nm patterning.

The fabrication of nanostructures such as nanochannels with sub-10 nm resolution can be accomplished through several routes—e-beam lithography and sacrificial layer lithography (SLL). The process for e-beam lithography is a finely focused electron beam that is exposed over a resist surface, and the exposure duration and location is controlled with the use of a computer (Brewer 1980; Valiev 1992). When the resist is exposed to the electron beam, the electrons either break or join the molecules in the resist, so the local characteristics are changed in such a way that further processes can either remove the exposed part (positive resist) or remove the unexposed part (negative resist). The resist material determines if the molecules will either break or join together; thus, determining if a positive or negative image is produced. E-beam lithography can be used either to create photolithographic masks for replication or to create the devices directly. The masks that are created can be used for either optical or X-ray lithography. One limitation of e-beam lithography is that throughput is drastically reduced since a single electron beam is used to create the entire exposure pattern on the resist. While this technique is slower than conventional lithographic techniques, it is ideal for prototype fabrication because no masks are required.

In SLL process, the use of a sacrificial layer allows the direct control of nanochannel dimensions so long as there exists a method for removing the sacrificial layer with absolute selectivity to the structural layers. A materials system with such selectivity is the silicon/silicon oxide system used widely in the microfabrication of MEMS devices. The use of sidewall deposition of the sacrificial layer and subsequent etching allows for the fabrication of high density nanochannels for biomedical applications. It is based on surface micromachining (Hansford et al. 2001). Figure 16.80 shows a schematic of the process steps in SLL based on Hansford et al. (2001) work on fabrication of polysilicon membranes with nanochannels. As with all the membrane protocols, the first step in the fabrication is the etching of the support ridge structure into the bulk silicon substrate. A low stress silicon nitride (LSN or simply nitride), which functions as an etch stop layer, is then deposited using LPCVD. The base structural polysilicon layer (base layer) is deposited on top of the etch stop layer. The plasma etching of holes in the base layer is what defines the shape of the pores. The buried nitride etch stop acts as an etch stop for the plasma etching of a polysilicon base layer. After the pore holes are etched through the base layer, the pore sacrificial thermal oxide layer is grown on the base layer. The basic requirement of the sacrificial layer is the ability to control the thickness with high precision across the entire wafer. Anchor points are defined in the sacrificial oxide layer to mechanically connect the base layer with the plug layer (necessary to maintain the pore spacing between layers). This is accomplished by using the same mask shifted from the pore holes. This produces anchors in one or two corners of each pore hole, which provides the desired connection between the structural layers while opening as much pore area as possible. After the anchor points are etched through the sacrificial oxide, the plug polysilicon layer is deposited (using LPCVD) to fill in the holes. To open the pores at the surface, the plug layer is planarized using chemical mechanical polishing (CMP) down to the base layer, leaving the final structure with the plug layer only in the

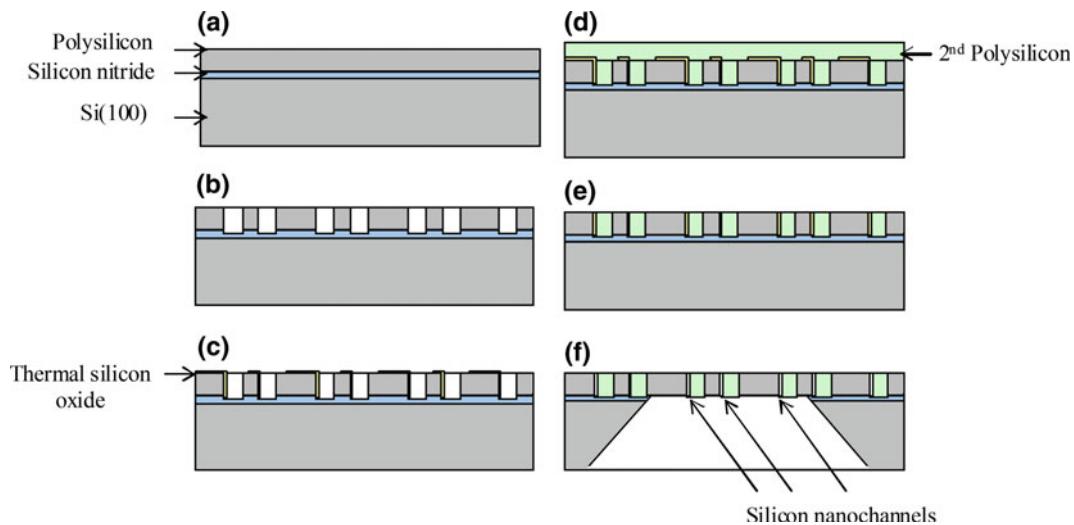


Fig. 16.80 Schematic of process steps involved in sacrificial layer lithography: **a** growth of silicon nitride layer (etch stop) and base polysilicon deposition, **b** hole definition in base, **c** growth of thin sacrificial oxide and patterning of anchor points, **d** deposition of plug polysilicon, **e** planarization of plug layer, and **f** deposition and patterning of protective nitride layer through etch, followed by etching of protective, sacrificial and etch layers final release of the structure in HF (Hansford et al. 2001)

pore hole openings. As the silicon wafer is ready for release, a protective nitride layer is deposited on the wafer (completely covering both sides of the wafer). The backside etch windows are etched in the protective layer, exposing the silicon wafer in the desired areas, and the wafer is placed in a KOH bath to etch. After the silicon wafer is completely removed up to the membrane (as evidenced by the smooth buried etch stop layer), the protective, sacrificial, and etch stop layers are removed by etching in concentrated HF. Etching of the sacrificial layer in polysilicon film defines nanochannels.

The LIGA process is based on the combined use of X-ray lithography, electroplating, and molding processes. X-rays produced by synchrotron radiation are used to prepare the mold. The steps involved in the LIGA process are shown schematically in Fig. 16.81. LIGA is used to produce high-aspect-ratio MEMS (HAR-MEMS) devices that are up to 1 mm in height and only a few microns in width or length (Becker et al. 1986). The LIGA process yields very sturdy 3D structures due to their increased thickness. One of the limitations of silicon microfabrication

processes originally used for fabrication of MEMS devices is the lack of suitable materials that can be processed. With LIGA, a variety of nonsilicon materials such as metals, ceramics and polymers can be processed.

Nonlithographic micromachining processes, primarily in Europe and Japan, are also being used for fabrication of millimeter-scale devices using direct material microcutting or micromechanical machining (such as microturning, micromilling, and microdrilling) or removal by energy beams (such as microspark erosion, focused ion beam, laser ablation, and laser polymerization) (Friedrich and Warrington 1998; Madou 2011). Hybrid technologies including LIGA and high-precision micromachining techniques have been used to produce miniaturized motors, gears, actuators, and connectors (Lehr et al. 1996, 1997; Michel and Ehrfeld 1998; Tanaka 2001). These millimeter-scale devices may find more immediate applications.

A micro/nanofabrication technique, so-called “soft lithography,” is a nonlithographic technique (Manz and Becker 1998; Xia and Whitesides 1998; Becker and Gaertner 2000; Beebe et al.

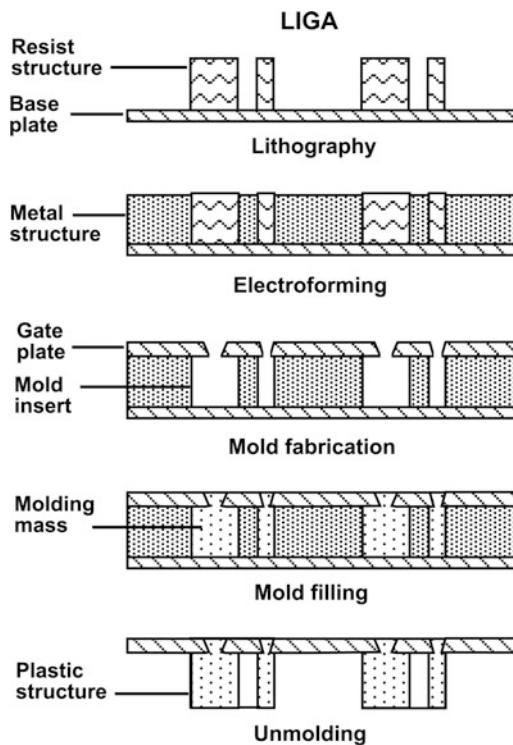


Fig. 16.81 Schematic of process steps involved in LIGA fabrication of MEMS

2002), in which a master or mold is used to generate patterns, defined by the relief on its surface, on polymers by replica molding (Xia et al. 1996), hot embossing (nanoimprint lithography) (Chou et al. 1996), or by contact printing (known as microcontact printing or μ CP) (Kumar and Whitesides 1993). Soft lithography is faster, less expensive, and more suitable for most biological applications than glass or silicon micromachining. Polymers have established an important role in BioMEMS/BioNEMS because of their reduced cost. The use of polymers also offers a wide range of material properties to allow tailoring of biological interactions for improved biocompatibility. Polymer fabrication is believed to be about an order of magnitude cheaper than that silicon fabrication.

Replica molding is the transfer of a topographic pattern by curing or solidifying a liquid precursor against the original patterned mold. The mold or stamp is generally made of a two-part polymer (elastomer and curing agent),

such as poly(dimethylsiloxane) (PDMS) from photolithographically generated photoresist master. Solvent based embossing, or imprinting, uses a solvent to restructure a polymer film. Hot embossing, also called nanoimprint lithography, usually refers to the transfer of pattern from a micromachined quartz or metal master to a pliable plastic sheet. Heat and high pressure allow the plastic sheet to become imprinted. These sheets can then be bonded to various plastics such as polymethyl methacrylate (PMMA). Nanoimprint lithography can produce patterns on a surface having 10-nm resolution. Contact printing uses a patterned stamp to transfer ink (mostly self-assembled monolayer) onto a surface in a pattern defined by the raised regions of a stamp. These techniques can be used to pattern line widths as small as 60 nm.

Replica molding is commonly used for mass-produced, disposable plastic micro/nanocomponents, for example micro/nanofluidic chips, generally made of PDMS and PMMA (Becker and Gaertner 2000; McDonald et al. 2000) and more flexible in choice of materials for construction than conventional photolithography.

To assemble microsystems, microrobots are used. Microrobotics include building blocks, such as steering links, microgrippers, conveyor system, and locomotive robots (Gad-el-Hak 2002).

Bottom-Up Fabrication (Nanochemistry)

Bottom-up approach (from small to large) largely relies on nanochemistry (Drexler 1992; Bhushan 1999a, 2005; Timp 1999; Rietman 2001; Goddard et al. 2002; Nalwa 2002; Poole and Owens 2003). The bottom-up approach includes chemical synthesis, the spontaneous “self-assembly” of molecular clusters (molecular self-assembly) from simple reagents in solution or biological molecules as building blocks to produce three dimensional nanostructures as done by nature, quantum dots (nanocrystals) of arbitrary diameter (about $10\text{--}10^5$ atoms), molecular beam epitaxy (MBE) and organometallic vapor phase epitaxy (OMVPE) to create specialized crystals one

atomic or molecular layer at a time, and manipulation of individual atoms by a scanning tunneling microscope or an AFM or atom optics. The self-assembly must be encoded, that is, one must be able to precisely assemble one object next to another to form a designed pattern. A variety of nonequilibrium plasma chemistry techniques are also used to produce layered nanocomposites, nanotubes, and nanoparticles. Nanostructures can also be fabricated using mechanosynthesis with proximal probes.

References

- Aden JS, Bohorquez JH, Collins DM, Crook MD, Garcia A, Hess UE (1994) The third-generation HP thermal inkjet printhead. *HP J* 45(1):41–45
- Ahn CH, Allen MG (1995) Fluid micropumps based on rotary magnetic actuators. In: MEMS '95: IEEE 8th international workshop on MEMS. IEEE, Piscataway, NJ, pp 408–412
- Aksyuk VA, Pardo F, Carr D, Greywall D et al (2003) Beam-steering micromirrors for large optical cross-connects. *J Lightwave Technol* 21:634–642
- Alley RL, Cuan GJ, Howe RT, Komvopoulos K (1992) The effect of release-etch processing on surface microstructure stiction. In: Mastrangelo CH, Hsu CH (eds) Proceedings of solid state sensor and actuator workshop. IEEE, New York, pp 202–207
- Andrade JD, Hlady V, Wei AP (1992) Adsorption of complex proteins at interfaces. *Pure Appl Chem* 64 (11):1777–1781
- Anonymous (1988) Properties of silicon. EMIS Data Reviews Series No. 4, INSPEC, Institution of Electrical Engineers, London (see also Anonymous, 2002, MEMS Materials Database. <http://www.memsnets.org/material/>)
- Anonymous (2004) International technology roadmap for semiconductors. <http://public.itrs.net/>
- Arney S (2001) Designing for MEMS reliability. *MRS Bull* 26:296–299
- Baydo R, Groskopf A (2001) Getting to the heart of ink jet: printheads. Beyond Recharger, May 10, 10–12. Also visit <http://64.78.37.127.pdf/baydo.pdf>
- Becker H, Gaertner C (2000) Polymer microfabrication methods for microfluidic analytical applications. *Electrophoresis* 21:12–26
- Becker H, Locascio LE (2002) Polymer microfluidic devices. *Talanta* 56:267–287
- Becker EW, Ehrfeld W, Hagmann P, Maner A, Munchmeyer D (1986) Fabrication of microstructures with high aspect ratios and great structural heights by synchrotron radiation lithography, galvanoforming, and plastic moulding (LIGA process). *Microelectron Eng* 4:35–56
- Beebe DJ, Mensing GA, Walker GM (2002) Physics and applications of microfluidics in biology. *Annu Rev Biomed Eng* 4:261–286
- Beerschwinge U, Yang SJ, Reuben RL, Taghizadeh MR, Wallrabe U (1994) Friction measurements on LIGA-processed microstructures. *J Micromech Microeng* 4:14–24
- Berlin JM, Yu J, Lu W, Walsh EE, Zhang L, Zhang P, Chen W, Kan AT, Wong MS, Tomson MB, Tour JM (2011) Engineered nanoparticles for hydrocarbon detection in oil-field rocks. *Energy Environ Sci* 4:505–509
- Bhatia SN, Chen CS (1999) Tissue engineering at the micro-scale. *Biomed Microdevices* 2:131–144
- Bhushan B (1996a) Tribology and mechanics of magnetic storage devices, 2nd edn. Springer, New York
- Bhushan B (1996b) Nanotribology and nanomechanics of MEMS devices. In: Proceedings of ninth annual workshop on micro electro mechanical systems. IEEE, New York, pp 91–98
- Bhushan B (1998) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht
- Bhushan B (1999a) Handbook of micro/nanotribology, 2nd edn. CRC, Boca Raton
- Bhushan B (1999b) Chemical, mechanical and tribological characterization of ultra-thin and hard amorphous carbon coatings as thin as 3.5 nm: recent developments. *Diam Relat Mater* 8:1985–2015
- Bhushan B (2000) Mechanics and reliability of flexible magnetic media, 2nd edn. Springer, New York
- Bhushan B (2001) Modern tribology handbook, vols 1 & 2 (Bhushan B, ed). CRC Press, Boca Raton
- Bhushan B (2003) Adhesion and stiction: mechanisms, measurement techniques, and methods for reduction. *J Vac Sci Technol B* 21:2262–2296
- Bhushan B (2005) Nanotribology and nanomechanics. *Wear* 259:1507–1531
- Bhushan B (2007) Nanotribology and nanomechanics of MEMS/NEMS and BioMEMS/BioNEMS materials and devices. *Microelectron Eng* 84:387–412
- Bhushan B (2008a) Nanotribology, nanomechanics and nanomaterials characterization. *Philos Trans R Soc A* 366:1351–1381
- Bhushan B (2008b) Nanotribology and nanomechanics in nano/biotechnology. *Philos Trans R Soc A* 366:1499–1537
- Bhushan B (2008c) Nanotribology of carbon nanotubes. *J Phys: Condens Matter* 20:365214
- Bhushan B (2011) Nanotribology and nanomechanics, 1–3 edn. Springer, Heidelberg
- Bhushan B (2013a) Principles and applications of tribology. Wiley, New York
- Bhushan B (2013b) Introduction to Tribology. Wiley, New York
- Bhushan B (2015) Governance, policy, and legislation of nanotechnology: a perspective. *Microsyst Technol* 21:1137–1155
- Bhushan B (2016a) Encyclopedia of nanotechnology, 2nd edn. Springer International, Switzerland

- Bhushan B (2016b) Biomimetics: bioinspired hierarchical-structured surfaces for green science and technology. Springer International, Switzerland
- Bhushan B, Burton Z (2005) Adhesion and friction properties of polymers in microfluidic devices. *Nanotechnology* 16:467–478
- Bhushan B, Gupta BK (1991) Handbook of tribology: materials, coatings and surface treatments. McGraw Hill, New York
- Bhushan B, Koinkar VN (1994) Tribological studies of silicon for magnetic recording applications. *J Appl Phys* 75:5741–5746
- Bhushan B, Kwak KJ (2007) Platinum-coated probes sliding at up to 100 mm s^{-1} against coated silicon wafers for AFM probe-based recording technology. *Nanotechnology* 18:345504
- Bhushan B, Kwak KJ (2008) Noble metal-coated probes sliding at up to 100 mm s^{-1} against PZT films for AFM probe-based ferroelectric recording technology (invited). *J Phys: Condens Matter* 20:225013
- Bhushan B, Liu H (2001) Nanotribological properties and mechanisms of alkylthiol and biphenyl thiol self-assembled monolayers studied by AFM. *Phys Rev B* 63:245412
- Bhushan B, Liu H (2004) Characterization of nanomechanical and nanotribological properties of digital micromirror devices. *Nanotechnology* 15:1785–1791
- Bhushan B, Schricker SR (2014) A review of block copolymer-based biomaterials that control protein and cell interactions. *J Biomed Mater Res A* 102A:2467–2480
- Bhushan B, Sundararajan S (1998) Micro/nanoscale friction and wear mechanisms of thin films using atomic force and friction force microscopy. *Acta Mater* 46:3793–3804
- Bhushan B, Utter J (2013) Nanoscale adhesion, friction and wear of proteins on polystyrene. *Colloids Surf B: Biointerfaces* 102:484–491
- Bhushan B, Venkatesan S (1993) Friction and wear studies of silicon in sliding contact with thin-film magnetic rigid disks. *J Mater Res* 8:1611–1628
- Bhushan B, Israelachvili JN, Landman U (1995a) Nanotribology: friction, wear and lubrication at the atomic scale. *Nature* 374:607–616
- Bhushan B, Kulkarni AV, Koinkar VN, Boehm M, Odoni L, Martelet C, Belin M (1995b) Microtribological characterization of self-assembled and langmuir-blodgett monolayers by atomic force and friction force microscopy. *Langmuir* 11:3189–3198
- Bhushan B, Kulkarni AV, Bonin W, Wyrobek JT (1996) Nano/picoindentation measurement using a capacitance transducer system in atomic force microscopy. *Philos Mag* 74:1117–1128
- Bhushan B, Sundararajan S, Li X, Zorman CA, Mehregany M (1998) Micro/nanotribological studies of single-crystal silicon and polysilicon and SiC films for use in MEMS devices. In: Bhushan B (ed) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht, pp 407–430
- Bhushan B, Liu H, Hsu SM (2004) Adhesion and friction studies of silicon and hydrophobic and low friction films and investigation of scale effects. *ASME J Tribol* 126:583–590
- Bhushan B, Tokachichu DR, Keener MT, Lee SC (2005a) Morphology and adhesion of biomolecules on silicon based surfaces. *Acta Biomater* 1:327–341
- Bhushan B, Kasai T, Kulik G, Barbieri L, Hoffmann P (2005b) AFM study of perfluorosilane and alkylsilane self-assembled monolayers for anti-stiction in MEMS/NEMS. *Ultramicroscopy* 105:176–188
- Bhushan B, Tokachichu D, Keener MT, Lee SC (2006a) Nanoscale adhesion, friction, and wear studies of biomolecules on silicon based surfaces. *Acta Biomater* 2:39–49
- Bhushan B, Hansford D, Lee KK (2006b) Surface modification of silicon and PDMS surfaces with vapor phase deposited ultrathin fluorosilane films for biomedical nanodevices. *J Vac Sci Technol, A* 24:1197–1202
- Bhushan B, Cichomski M, Hoque E, DeRose JA, Hoffmann P, Mathieu HJ (2006c) Nanotribological characterization of perfluoroalkylphosphonate self-assembled monolayers deposited on aluminum-coated silicon substrates. *Microsyst Technol* 12:588–596
- Bhushan B, Cichomski M, Tao Z, Tran NT, Ethen T, Merton C, Jewett RE (2007) Nanotribological characterization and lubricant degradation studies of metal-film magnetic tapes using novel lubricants. *ASME J Tribol* 129:621–627
- Bhushan B, Kwak K, Palacio M (2008a) Nanotribology and nanomechanics of AFM probe-based data recording technology. *J Phys: Condens Matter* 20, Art # 365207
- Bhushan B, Palacio M, Kinzig B (2008b) AFM-based nanotribological and electrical characterization of ultrathin wear-resistant ionic liquid films. *J Colloid Interface Sci* 317:275–287
- Bhushan B, Kwak K, Gupta S, Lee SC (2009) Nanoscale adhesion, friction and wear studies of biomolecules on SAM-coated silica and alumina based surfaces. *J Royal Soc Interface* 6:719–733
- Bhushan B, Luo D, Schricker SR, Sigmund W, Zauscher S (2014) Handbook of nanomaterials properties, vols I & II. Springer, Heidelberg
- Black J (1999) Biological performance of materials: fundamentals of biocompatibility. Marcel Dekker, New York
- Brewer G (1980) Electron-beam technology in microelectronic fabrication. Academic Press, New York
- Browne MM, Lubarsky GV, Davidson MR, Bradley RH (2004) Protein adsorption onto polystyrene surfaces studied by XPS and AFM. *Surf Sci* 553:155–167
- Bryzek J, Peterson K, McCulley W (1994) Micromachines on the March. *IEEE Spectrum* May, 20–31
- Bustgens B, Bacher W, Menz W, Schomburg WK (1994) Micropump manufactured by thermoplastic molding. In: Proceedings of IEEE-MEMS workshop. IEEE, Piscataway, NJ, pp 18–21

- Callahan DL, Morris JC (1992) The extent of phase transformation in silicon hardness indentation. *J Mater Res* 7:1612–1617
- Casal P, Wen X, Gupta S, Nicholson T III, Wang Y, Theiss A, Bhushan B, Brillson L, Lu W, Lee SC (2012) ImmunoFET feasibility in physiological salt environments. *Philos Trans R Soc A* 370:2474–2488
- Cayre OJ, Paunov VN (2004) Contact angles of colloid silica and gold particles at air-water and oil-water interfaces determined with the gel trapping technique. *Langmuir* 20:9594–9599
- Chang LS, Gendler PL, Jou JH (1991) Thermal mechanical and chemical effects in the degradation of the plasma-deposited α -SC: H passivation layer in a multilayer thin-film device. *J Mater Sci* 26:1882–1890
- Chen RJ, Choi HC, Bangsaruntip S, Yenilmez E, Tang X, Wang Q, Chang YL, Dai H (2004) An investigation of the mechanisms of electrode sensing of protein adsorption on carbon nanotube devices. *J Am Chem Soc* 126:1563–1568
- Chen Z, Cummins D, Reinecke BN, Clark E, Sunkara MK, Jaramillo TF (2011) Core–shell Mo₃–MoS₂ nanowires for hydrogen evolution: a functional design for electrocatalytic materials. *Nano Lett* 11:4168–4175
- Cheng J, Kricka LJ (eds) (2001) Biochip technology. Harwood Academic Publishers, Philadelphia
- Chilamakuri SK, Bhushan B (1999) A comprehensive kinetic meniscus model for prediction of long-term static friction. *J Appl Phys* 15:4649–4656
- Cho SK, Moon H, Kim C-J (2003) Creating, transporting, cutting, and merging liquid droplets by electrowetting-based actuation for digital microfluidic circuits. *J MEMS* 12:70–80
- Chou SY, Krauss PR, Renstrom PJ (1996) Imprint lithography with 25-nanometer resolution. *Science* 272:85–87
- Core TA, Tsang WK, Sherman SJ (1993) Fabrication technology for an integrated surface-micromachined sensor. *Solid State Technol* 36(Oct):39–47
- De Boer MP, Michalske TA (1999) Accurate method for determining adhesion of cantilever beams. *J Appl Phys* 86:817
- Decuzzi P, Lee S, Bhushan B, Ferrari M (2005) A theoretical model for the margination of particles with blood vessels. *Ann Biomed Eng* 33:179–190
- Decuzzi P, Pasqualini R, Arap W, Ferrari M (2009) Intravascular delivery of particulate systems: does geometry really matter? *Pharm Res* 26:235–243
- Desai TA, Hansford DJ, Kulinsky L, Nashat AH, Rasi G, Tu J, Wang Y, Zhang M, Ferrari M (1999) Nanopore technology for biomedical applications. *Biomed Devices* 2:11–40
- DeWolf I, van Spengen WM (2002) Techniques to study the reliability of metal RF MEMS capacitive switches. *Microelectron Reliab* 42:1789–1794
- Di Lullo GA, Sweeney SM, Korkko J, Ala-Kokko L, San Antonio JD (2002) Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *J Biol Chem* 277(6):4223–4231
- Doepper J, Clemens M, Ehrfeld W, Jung S, Kaemper KP, Lehr H (1997) Micro gear pumps for dosing of viscous fluids. *J Micromech Microeng* 7:230–232
- Douglass MR (1998) Lifetime estimates and unique failure mechanisms of the digital micromirror devices (DMD). In: Proceedings of 36th annual international reliability physics symposium. IEEE, New York, pp 9–16
- Douglass MR (2003) DMD reliability: a MEMS success story. In: Reliability, testing, and characterization of MEMS/MOEMS II, Proceedings of SPIE, vol 4980. SPIE, Bellingham, Washington, pp 1–11
- Dresselhaus MS, Dresselhaus G, Avouris Ph (2001) Carbon nanotubes—synthesis, structure, properties and applications. Springer, Berlin
- Drexler KE (1992) Nanosystems: molecular machinery, manufacturing and computation. Wiley, New York
- Duncan R (2003) The dawning era of polymer therapeutics. *Nat Rev Drug Discov* 2:347–360
- Elwenspoek M, Wiegerink R (2001) Mechanical microsensors. Springer, Berlin
- Eteshola E, Keener MT, Elias M, Shapiro J, Brillson LJ, Bhushan B, Lee SC (2008) Engineering functional protein interfaces for immunologically modified field effect transistor (ImmunoFET) by molecular genetics means. *J R Soc Interface* 5:123–127
- Fan J, Myant CW, Underwood R, Cann PM, Hart A (2011) Inlet protein aggregation: a new mechanism for lubricating film formation with model synovial fluids. *Proc Inst Mech Eng H* 225:696–709
- Feldman JL (1976) Elastic constants of 2H-MoS₂ and 2H-NbSe₂ extracted from measured dispersion curves and linear compressibilities. *J Phys Chem Solids* 37:1141–1144
- Ferrari M (2005) Cancer nanotechnology: opportunities and challenges. *Nat Rev Cancer* 5:161–171
- Ferrari M (2010) Frontiers in cancer nanomedicine: directing mass transport through biological barriers. *Trends Biotechnol* 28:181–188
- Fleck NA, Muller GM, Ashby MF, Hutchinson JW (1994) Strain gradient plasticity: theory and experiment. *Acta Metall Mater* 42:475–487
- Francis GL (2010) Albumin and mammalian cell culture: implications for biotechnology applications. *Cytotechnology* 62(1):1–16
- Frechette LG, Jacobson SA, Breuer KS et al (2005) High-speed microfabricated silicon turbomachinery and fluid film bearings. *J MEMS* 14:141–152
- Freeman GR, March NH (1999) Triboelectricity and some associated phenomena. *Mater Sci Eng* 15:1454–1458
- Friedrich CR, Warrington RO (1998) Surface characterization of non-lithographic micromachining. In: Bhushan B (ed) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht, pp 73–84
- Fuerstenau MC, Han KN (eds) (2003) Principles of mineral processing. Society for mining, metallurgy, and exploration (SME), Littleton

- Fujimasa I (1996) *Micromachines: a new era in mechanical engineering*. Oxford University Press, Oxford
- Fukuda T, Arai F, Dong L (2003) Assembly of nanodevices with carbon nanotubes through nanorobotic manipulations. *Proc IEEE* 91:1803–1818
- Gabriel KJ, Behi F, Mahadevan R, Mehregany M (1990) In situ friction and wear measurement in integrated polysilicon mechanisms. *Sens Actuators A* 21–23:184–188
- Gad-el-Hak M (2002) *The MEMS handbook*. CRC Press, Boca Raton
- Garcia EJ, Sniegowski JJ (1995) Surface micromachined microengine. *Sens Actuators A* 48:203–214
- Geim AK, Novoselov KS (2007) The rise of graphene. *Nat Mater* 6:183–191
- Giles CR, Bishop D, Aksyuk V (2001) MEMS for light-wave networks. *MRS Bull*, April, 328–329
- Goddard WA, Brenner DW, Lyshevski SE, Iafrate GJ (eds) (2002) *Handbook of nanoscience, engineering, and technology*. CRC Press, Boca Raton
- Gravesen P, Branebjerg J, Jensen O (1993) Microfluidics —a review. *J Micromech Microeng* 3:168–182
- Grayson ACR, Shawgo RS, Johnson AM, Flynn NT, Li Y, Cima MJ, Langer R (2004) A BioMEMS review: MEMS technology for physiologically integrated devices. *Proc IEEE* 92:6–21
- Green NM (1990) Avidin and streptavidin. *Method Enzymol* 184:51–67
- Guckel H, Burns DW (1989) Fabrication of micromechanical devices from polysilicon films with smooth surfaces. *Sens Actuators* 20:117–122
- Gupta BK, Bhushan B (1994) Nanoindentation studies of ion implanted silicon. *Surf Coat Technol* 68–69:564–570
- Gupta BK, Chevallier J, Bhushan B (1993) Tribology of ion bombarded silicon for micromechanical applications. *ASME J Tribol* 115:392–399
- Gupta BK, Bhushan B, Chevallier J (1994) Modification of tribological properties of silicon by boron ion implantation. *Tribol Trans* 37:601–607
- Gupta S, Wu H-H, Kwak KJ, Casal P, Nicholson TR III, Wen X, Anisha R, Bhushan B, Berger PR, Lu W, Brillson LJ, Lee SC (2011) Interfacial design and structure of protein/polymer films on oxidized AlGaN surfaces. *J Phys D Appl Phys* 44:034010
- Hamilton H (1991) Contact recording on perpendicular rigid media. *J Mag Soc Jpn* 15(Suppl. S2):483–481
- Han Y, Mayer D, Offenhausser A, Ingebrandt S (2006) Surface activation of thin silicon-oxides by wet cleaning and silanization. *Thin Solid Films* 510:175–180
- Hankins MG, Resnick PJ, Clews PJ, Mayer TM, Wheeler DR, Tanner DM, Plass RA (2003) Vapor deposition of amino-functionalized self-assembled monolayers on MEMS. In: *Proceedings of SPIE*, vol 4980. SPIE, Bellingham, Washington, pp 238–247
- Hansford D, Desai T, Ferrari M (2001) Nano-scale size-based biomolecular separation technology. In: Cheng J, Kricka LJ (eds) *Biochip technology*. Harwood Academic Pub, New York, pp 341–361
- Hattori S, Adachi E, Ebihara T, Shirai T, Someki I, Irie S (1999) Alkali-treated collagen retained the triple helical conformation and the ligand activity for the cell adhesion via $\alpha 2\beta 1$ integrin. *J Biochem* 125:676–684
- Heller MJ, Guttman A (eds) (2001) *Integrated microfabricated biodevices*. Marcel Dekker, New York
- Henck SA (1997) Lubrication of digital micromirror devices. *Tribol Lett* 3:239–247
- Henniker J (1962) Triboelectricity in polymers. *Nature* 196:474
- Hierlemann A (2005) Integrated chemical microsensor systems in CMOS technology. Springer, Berlin
- Hierlemann A, Brand O, Hagleitner C, Baltes H (2003) Microfabrication techniques for chemical/biosensors. In: Casalnuovo S, Brown RB (eds) *Proceedings of the IEEE, chemical and biological microsensors*, vol 91. IEEE, New York, pp 839–863
- Hierold C (2008) *Carbon nanotube devices*. Wiley-VCH, Weinheim
- Hirano T, Fan LS, Kercher D, Pattaik S, Pan TS (2000) HDD tracking microactuator and its integration issues. In: Lee AP, Simon J, Foster FK, Keynton RS (eds) *Proceedings of ASME international mechanical engineer congress and exposition MEMS*, vol 2. ASME, New York, pp 449–452
- Hirayama K, Akashi S, Furuya M, Fukuhara K (1990) Rapid confirmation and revision of the primary structure of bovine serum albumin by ESIMS and Frit-Fab LC/MS. *Biochem Biophys Res Commun* 173:639–646
- Hornbeck LJ (1997) Low surface energy passivation layer for micromechanical devices. U.S. Patent No. 5,602,671, Feb. 11
- Hornbeck LJ (1999) A digital light processingTM update —status and future applications. In: *Proceedings of society for photonics-optics engineering*, vol 3634, Projection displays V, pp 158–170
- Hornbeck LJ (2001) The DMDTM projection display chip: a MEMS-based technology. *MRS Bull* 26:325–328
- Hornbeck LJ, Nelson WE (1988) Bistable deformable mirror device. *OSA Tech Dig Ser Spat Light Modulators Appl* 8:107–110
- Horsley DA, Cohn MB, Singh A, Horowitz R, Pisano AP (1998) Design and fabrication of an angular microactuator for magnetic disk drives. *J Microelectromech Syst* 7:141–148
- Hsu TR (2002) *MEMS and microsystems: design and manufacture*. McGraw-Hill, Boston
- Iijima S, Yudasaka M, Yamada R, Bandow S, Suenaga K, Kokai F, Takahashi K (1999) Nano-aggregates of single-walled graphitic carbon nano-horns. *Chem Phys Lett* 309:165–170
- Irvine DJ (2011) Drug delivery: one nanoparticle, one kill. *Nat Mater* 10:342–343
- Jaeger RC (1988) *Introduction to microelectronic fabrication*, vol 5. Addison-Wesley, Reading
- Johnson PR, Sun N, Elimelech M (1996) Colloid transport in geochemically heterogeneous porous

- media: modeling and measurements. *Environ Sci Technol* 30:3284–3293
- Judy JW (2001) Microelectromechanical systems (MEMS): fabrication, design, and applications. *Smart Mater Struct* 10:1115–1134
- Kallury K, MacDonald PM, Thompson M (1994) Effect of surface water and base catalysis on the silanization of silica by (aminopropyl)alkoxysilanes studied by X-ray photoelectron spectroscopy and ^{13}C cross-polarization/magic angle spinning nuclear magnetic resonance. *Langmuir* 10:492–499
- Kasai T, Bhushan B, Kulik G, Barbieri L, Hoffmann P (2005) Nanotribological study of perfluorosilane SAMs for anti-stiction and low wear. *J Vac Sci Technol, B* 23:995–1003
- Kayali S, Lawton R, Stark BH (1999) MEMS reliability assurance activities at JPL. *EEE Links* 5:10–13
- Kim K-S, Lee H-J, Lee C, Lee S-K, Jang H, Ahn J-H, Kim J-H, Lee H-J (2011) Chemical vapor deposition-grown graphene: the thinnest solid lubricant. *ACS Nano* 5:5107–5114
- Koch A (2001) Bacterial growth and form. Springer-Science+Business Media, B.V., Dordrecht, Germany
- Koinkar VN, Bhushan B (1996a) Micro/nanoscale studies of boundary layers of liquid lubricants for magnetic disks. *J Appl Phys* 79:8071–8075
- Koinkar VN, Bhushan B (1996b) Microtribological studies of unlubricated and lubricated surfaces using atomic force/friction force microscopy. *J Vac Sci Technol, A* 14:2378–2391
- Kosmulski M (2001) Chemical properties of material surfaces. Marcel Dekker, New York
- Kovacs GTA (1998) Micromachined transducers sourcebook. WCB McGraw-Hill, Boston
- Kumar A, Whitesides GM (1993) Features of gold having micrometer to centimeter dimensions can be formed through a combination of stamping with an elastomeric stamp and an alkanethiol ink followed by chemical etching. *Appl Phy Lett* 63:2002–2004
- Kwak KJ, Bhushan B (2008) Platinum-coated probes sliding at up to 100 mm/s against lead zirconate titanate films for atomic force microscopy probe-based ferroelectric recording technology. *J Vac Sci Technol, A* 26:783–793
- Lang HP, Hegner M, Gerber C (2005) Cantilever array sensors. *Mater Today*, April, 30–36
- Lanza RP, Langer R, Vacanti J (eds) (2000) Principles of tissue engineering, 2nd edn. Academic Press, San Diego
- Le H (1998) Progress and trends in ink-jet printing technology. *J Imaging Sci Technol* 42:49–62
- Leclerc E, Furukawa KS, Miyata F, Sakai T, Ushida T, Fujii T (2004) Fabrication of microstructures in photosensitive biodegradable polymers for tissue engineering applications. *Biomaterials* 25:4683–4690
- Lee ER (2003) Microdrop generation. CRC Press, Boca Raton
- Lee SC, Keener MT, Tokachichu DR, Bhushan B, Barnes PD, Cipriany BR, Gao M, Brillson LJ (2005a) Protein binding on thermally grown silicon dioxide. *J Vac Sci Technol, B* 23:1856–1865
- Lee KK, Bhushan B, Hansford D (2005b) Nanotribological characterization of perfluoropolymer thin films for BioMEMS applications. *J Vac Sci Technol, A* 23:804–810
- Lee C, Wei X, Kysar JW, Hone J (2008) Measurement of the elastic properties and intrinsic strength of monolayer graphene. *Sci* 32:385–388
- Lehr H, Abel S, Doppler J, Ehrfeld W, Hagemann B, Kamper KP, Michel F, Schulz Ch, Thurigen Ch (1996) Microactuators as driving units for micro-robotic systems. In: Sulzmann A (ed) Proceedings of microrobotics: components and applications, vol 2906. SPIE, Bellingham, Washington, pp 202–210
- Lehr H, Ehrfeld W, Hagemann B, Kamper KP, Michel F, Schulz Ch, Thurigen Ch (1997) Development of micro-millimotors. *Minim Invasive Ther Allied Technol* 6:191–194
- Li X, Bhushan B (1999) micro/nanomechanical characterization of Ceramic Films for Microdevices. *Thin Solid Films* 340:210–217
- Lim MG, Chang JC, Schultz DP, Howe RT, White RM (1990) Polysilicon microstructures to characterize static friction. In: Proceedings of IEEE micro electro mechanical systems. IEEE, New York, pp 82–88
- Liu C (2006) Foundations of MEMS. Pierson Prentice Hall, Upper Saddle River
- Liu H, Bhushan B (2002) Investigation of nanotribological properties of self-assembled monolayers with alkyl and biphenyl spacer chains. *Ultramicroscopy* 91:185–202
- Liu H, Bhushan B (2003a) Nanotribological characterization of molecularly-thick lubricant films for applications to MEMS/NEMS by AFM. *Ultramicroscopy* 97:321–340
- Liu H, Bhushan B (2003b) Adhesion and friction studies of microelectromechanical systems/nanoelectromechanical systems materials using a novel microtriboapparatus. *J Vac Sci Technol, A* 21:1528–1538
- Liu H, Bhushan B (2004a) Nanotribological characterization of digital micromirror devices using an atomic force microscope. *Ultramicroscopy* 100:391–412
- Liu H, Bhushan B (2004b) Investigation of nanotribological and nanomechanical properties of the digital micromirror device by atomic force microscope. *J Vac Sci Technol, A* 22:1388–1396
- Liu H, Bhushan B (2004c) Bending and fatigue study on a nanoscale hinge by an atomic force microscope. *Nanotechnology* 15:1246–1251
- Liu LX, Spakovszky ZS (2005) Effect of bearing stiffness anisotropy on hydrostatic micro gas journal bearing dynamic behavior. In: Proceedings of ASME turbo expo 2005, Reno, Nevada, Paper No. GT-2005-68199
- Liu H, Bhushan B, Eck W, Stadler V (2001) Investigation of the adhesion, friction, and wear properties of biphenyl thiol self-assembled monolayers by atomic force microscopy. *J Vac Sci Technol, A* 19:1234–1240
- Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipursky L, Darnell J (2003) Molecular cell biology, 5th edn. W. H. Freeman & Company, New York

- Lubecke VM, Chiao JC (1999) MEMS technologies for enabling high frequency communication circuits. In: Proceedings of IEEE 4th international conference on telecommunications. In: Modern Satellite, Cable and Broadcasting Services, Nis, Yugoslavia. IEEE, New York, pp 1–8
- Maboudian R, Howe RT (1997) Critical review: adhesion in surface micromechanical structures. *J Vac Sci Technol*, B 15:1–20
- Madou M (2011) Fundamentals of microfabrication and nanotechnology, 3rd edn. CRC Press, Boca Raton
- Maharaj D, Bhushan B (2012) Effect of spherical Au nanoparticles on nanofriction and wear reduction in dry and liquid environments. *Beilstein J Nanotechnol* 3:759–772
- Maharaj D, Bhushan B (2013) Effect of MoS₂ and WS₂ nanotubes on nanofriction and wear reduction in dry and liquid environments. *Tribol Lett* 49:323–339
- Maharaj D, Bhushan B (2014a) Scale effects of nanomechanical properties and deformation behavior of Au nanoparticle and thin film using depth sensing nanoindentation. *Beilstein J Nanotechnol* 5:822–836
- Maharaj D, Bhushan B (2014b) Nanomanipulation, nanotribology and nanomechanics of Au nanorods in dry and liquid environments using an AFM and depth sensing nanoindenter. *Nanoscale* 6:5838–5852
- Maharaj D, Bhushan B (2015a) Friction, wear and mechanical behavior of nano-objects on the nanoscale. *Mater Sci Eng, R* 95:1–43
- Maharaj D, Bhushan B (2015b) Nanomechanical behavior of MoS₂ and WS₂ multi-walled nanotubes and carbon nanohorns. *Sci Rep Nat* 5:8539
- Maharaj D, Bhushan B, Iijima S (2013) Effect of carbon nanohorns on nanofriction and wear reduction in dry and liquid environments. *J Colloid Interface Sci* 400:147–160
- Man KF (2001) MEMS reliability for space applications by elimination of potential failure modes through testing and analysis. http://www.rel.jpl.nasa.gov/Org_5053/atop/products/Prod-map.html
- Man KF, Stark BH, Ramesham R (1998) A resource handbook for MEMS reliability, Rev. A. JPL Press, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California
- Mani SS, Fleming JG, Walraven JA, Sniegowski JJ et al (2000) Effect of W coating on microengine performance. In: Proceedings of 38th annual international reliability physics symposium. IEEE, New York, pp 146–151
- Manz A, Becker H (eds) (1998) Microsystem technology in chemistry and life sciences. Topics in current chemistry, vol 194. Springer, Heidelberg
- Martin FJ, Grove C (2001) Microfabricated drug delivery systems: concepts to improve clinical benefits. *Biomed Microdevices* 3:97–108
- Martin JR, Zhao Y (1997) Micromachined device packaged to reduce stiction. U.S. Patent No. 5,694,740, Dec. 9
- Mastrangelo CH (1998) Surface force induced failures in microelectromechanical systems. In: Bhushan B (ed) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht, pp 367–395
- Mastrangelo CH, Becker H (eds) (2001) Microfluidics and BioMEMS. In: Proceedings of SPIE, vol 4560. SPIE, Bellingham, Washington
- Mastrangelo CH, Hsu CH (1993) Mechanical stability and adhesion of microstructures under capillary forces—Part II: experiments. *J Microelectromech Syst* 2:44–55
- Matheison D, Beerschwinger U, Young SJ, Rueben RL, Taghizadeh M, Eckert S, Wallrabe U (1996) Effect of progressive wear on the friction characteristics of nickel LIGA processed rotors. *Wear* 192:199–207
- McDonald JC, Duffy DC, Anderson JR, Chiu DT, Wu H, Schueller OJA, Whitesides GM (2000) Fabrication of microfluidic systems in poly(dimethylsiloxane). *Electrophoresis* 21:27–40
- Mehregany M, Gabriel KJ, Trimmer WSN (1988) Integrated fabrication of polysilicon mechanisms. *IEEE Trans Electron Devices* 35:719–723
- Mehregany M, Zorman CA, Rajan N, Wu CH (1998) Silicon carbide MEMS for harsh environments. *Proc IEEE* 86:1594–1610
- Michel F, Ehrfeld W (1998) Microfabrication technologies for high performance microactuators. In: Bhushan B (ed) Tribology issues and opportunities in MEMS. Kluwer Academic, Dordrecht, pp 53–72
- Moon JH, Shin JW, Kim SY, Park JW (1996) Formation of uniform aminosilane thin layers: an imine formation to measure relative surface density of the amine group. *Langmuir* 12:4621–4624
- Mougin K, Gnecco E, Rao A, Cuberes MT, Jayaraman S, McFarland EW, Haidara H, Meyer E (2008) Manipulation of gold nanoparticles: influence of surface chemistry, temperature, and environment (vacuum versus ambient atmosphere). *Langmuir* 24:1577–1581
- Mulhern GT, Soane DS, Howe RT (1993) Supercritical carbon dioxide drying of microstructures. In: Proceedings international conference on solid-state sensors and actuators. IEEE, New York, pp 296–299
- Muller RS, Howe RT, Senturia SD, Smith RL, White RM (1990) Microsensors. IEEE Press, New York
- Murakami T, Ajima K, Miyawaki J, Yudasaka M, Iijima S, Shiba K (2004) Drug-loaded carbon nanohorns: adsorption and release of dexamethasone in vitro. *Mol Pharm* 1:399–405
- Nalwa HS (ed) (2002) Nanostructured materials and nanotechnology. Academic Press, San Diego
- Nelson DL, Cox MM (2005) Lehninger's principles of biochemistry, 4th edn. W. H. Freeman & Company, New York
- Nguyen NT, Huang X, Chuan TK (2002) MEMS-micropumps: a review. *ASME J Fluids Eng* 124:384–392
- Nicholson TR III, Gupta S, Wen X, Wu HH, Anisha R, Casal P, Kwak KJ, Bhushan B, Berger PR, Lu W, Brillson LJ, Lee SC (2010) Rational enhancement of nanobiotechnological device function illustrated by partial optimization of a protein sensing field effect transistor. *Proc I Mech E, Part N: J Nanoeng Nanosyst* 223:149–161

- Oeberg PA, Togawa T, Spelman FA (2004) Sensors in medicine and health care. Wiley, New York
- Ohnstein T, Fukiura T, Ridley J, Bonne U (1990) Micromachined silicon microvalve. In: Proceedings of IEEE-MEMS workshop. IEEE, New York, pp 95–98
- Palacio M, Bhushan B (2007a) Surface potential and resistance measurements for detecting wear of chemically-bonded and unbonded molecularly-thick perfluoropolyether lubricant films using atomic force microscopy. *J Colloid Interface Sci* 315:261–269
- Palacio M, Bhushan B (2007b) Wear detection of candidate MEMS/NEMS lubricant films using atomic force microscopy-based surface potential measurements. *Scr Mater* 57:821–824
- Palacio M, Bhushan B (2008a) Ultrathin wear-resistant ionic liquid films for novel MEMS/NEMS applications. *Adv Mater* 20:1194–1198
- Palacio M, Bhushan B (2008b) A nanoscale friction investigation during manipulation of nanoparticles in controlled environments. *Nanotechnology* 19:315710
- Palacio M, Bhushan B (2009) Molecularly thick ionic liquid films for nanolubrication. *J Vac Sci Technol A* 27:986–995
- Palacio MLB, Bhushan B (2012) Bioadhesion: a review of concepts and applications. *Philos Trans R Soc A* 370:2321–2347
- Pamula E, De Cupere V, Dufrêne YF, Rouxhet PG (2004) Nanoscale organization of adsorbed collagen: influence of substrate hydrophobicity and adsorption time. *J Colloid Interface Sci* 271:80–91
- Pankov R, Yamada KM (2002) Fibronectin at a glance. *J Cell Sci* 115:3861–3863
- Panyala NR, Pena-Mendez EM, Havel J (2009) Gold and nano-gold in medicine: overview, toxicology and perspectives. *J Appl Biomed* 7:75–91
- Park K (ed) (1997) Controlled drug delivery: challenges and strategies. American Chemical Society, Washington, D.C.
- Parsons M (2001) Design and manufacture of automotive pressure sensors. *Sensors* 18:32–46
- Patolsky F, Lieber C (2005) Nanowire nanosensors. *Mater Today*, April 20–28
- Pavía-Sanders A, Zhang S, Flores JA, Sanders JE, Raymond JE, Wooley KL (2013) Robust magnetic/polymer hybrid nanoparticles designed for crude oil entrapment and recovery in aqueous environments. *ACS Nano* 7:7552–7561
- Peters T (1980) Albumin: an overview and bibliography. Miles Laboratory Research Products Division, Elkhart
- Pharr GM (1991) The anomalous behavior of silicon during nanoindentation. In: Nix WD, Bravman JC, Arzt E, Freund LB (eds) Thin films: stresses and mechanical properties III, vol 239. Materials Research Soc., Pittsburgh, pp 301–312
- Poole CP, Owens FJ (2003) Introduction to nanotechnology. Wiley, Hoboken
- Prakash S, Pinti M, Bhushan B (2012) Theory, fabrication and applications of microfluidic and nanofluidic biosensors. *Philos Trans R Soc A* 370:2269–2303
- Proctor RA (1987) Fibronectin: a brief overview of its structure, function, and physiology. *Rev Infect Dis* 9: S317–S321
- Raiteri R, Grattarola M, Butt M, Skladal P (2001) Micromechanical cantilever-based biosensor. *Sens Actuators B: Chemical* 79:115–126
- Ratner BD, Hoffman AS, Schoen FJ, Lemons JE (2004) Biomaterials science: an introduction to materials in medicine. Elsevier, San Diego
- Rietman EA (2001) Molecular engineering of nanosystems. Springer, New York
- Robbins RA, Jacobs SJ (2001) Lubricant delivery for micromechanical devices. U.S. Patent No. 6,300,294 B1, Oct. 9
- Robertson JK, Wise KD (1998) An electrostatically actuated integrated microflow controller. *Sens Actuators A* 71:98–106
- Sakaguchi M, Kashiwabara H (1992) A generation mechanism of triboelectricity due to the reaction of mechaniradicals with mechanoions which are produced by mechanical fracture of solid polymer. *Colloid Polym Sci* 270:621–626
- Samsonov GV (ed) (1968) Handbook of the physiochemical properties of the elements. Plenum, New York
- San CLP, Yap EPH (eds) (2001) Frontiers in human genetics. World Scientific, Singapore
- Sani E, Barison S, Pagura C, Mercatelli L, Sansoni P, Fontani D, Jafrancesco D, Francini F (2010) Carbon nanohorns-based nanofluids as direct sunlight absorbers. *Opt Express* 18:5179–5187
- Schricker SR, Palacio MLB, Bhushan B (2011) Antibody sensed protein surface conformation. *Mater Today* 14:616–621
- Schricker S, Palacio MLB, Bhushan B (2012) Designing nanostructured block copolymer surfaces to control protein adhesion. *Philos Trans R Soc A* 370:2348–2380
- Scott M (2003) MEMS and MOEMS for national security applications. In: Reliability, testing, and characterization of MEMS/MOEMS II, Proceedings of SPIE, vol 4980. SPIE, Bellingham, Washington, pp xxxvii–xliv
- Senturia SD (2000) Microsystem design. Kluwer Academic, Boston
- Shackelford JF, Alexander W, Park JS (eds) (1994) CRC material science and engineering handbook, 2nd edn. CRC Press, Boca Raton
- Shawgo RS, Grayson ACR, Li Y, Cima MJ (2002) BioMEMS for drug delivery. *Curr Opin Solid State Mater Sci* 6:329–334
- Shoji S, Esashi M (1994) Microflow devices and systems. *J Micromech Microeng* 4:157–171
- Shor JS, Goldstein D, Kurtz AD (1993) Characterization of n-type β -SiC as a piezoresistor. *IEEE Trans Electron Devices* 40:1093–1099
- Silver FH, Freeman JW, Seehra GP (2003) Collagen self-assembly and the development of tendon mechanical properties. *J Biomech* 36:1529–1553
- Smith G (1997) The application of microtechnology to sensors for the automotive industry. *Microelectron J* 28:371–379

- Spearing SM, Chen KS (2001) Micro-gas turbine engine materials and structures. *Ceram Eng Sci Proc* 18:11–18
- Srivastava D (2004) Computational nanotechnology of carbon nanotubes. In: Meyyappan M (ed) *Carbon nanotubes: science and applications*. CRC Press, Boca Raton, pp 25–63
- Stampfer C, Jungen A, Hierold C (2004) Fabrication of discrete carbon nanotube based nanoscaled force sensor. In: *Proceedings IEEE sensors 2004*, Vienna. IEEE, New York, pp 1056–1059
- Stehr M, Messner S, Sandmaier H, Zenergle R (1996) The VAMP—a new device for handling liquids or gases. *Sens Actuators A* 57:153–157
- Stifter T, Marti O, Bhushan B (2000) Theoretical investigation of the distance dependence of capillary and van der Waals forces in scanning force microscopy. *Phys Rev B* 62:13667–13673
- Stroscio JA, Eigler DM (1991) Atomic and molecular manipulation with a scanning tunneling microscope. *Science* 254:1319
- Sulouff RE (1998) MEMS opportunities in accelerometers and gyros and the microtribology problems limiting commercialization. In: Bhushan B (ed) *Tribology issues and opportunities in MEMS*. Kluwer Academic, Dordrecht, pp 109–120
- Sundararajan S, Bhushan B (1998) Micro/nanotribological studies of polysilicon and SiC films for MEMS applications. *Wear* 217:251–261
- Sundararajan S, Bhushan B (2001) Static friction and surface roughness studies of surface micromachined electrostatic micromotors using an atomic force/friction force microscope. *J Vac Sci Technol A* 19:1777–1785
- Sundararajan S, Bhushan B (2002) Development of AFM-based techniques to measure mechanical properties of nanoscale structures. *Sens Actuators A* 101:338–351
- Suzuki K (2002) Micro electro mechanical systems (MEMS) micro-switches for use in DC, RF, and optical applications. *Jpn J Appl Phys* 41:4335–4339
- Tai YC, Muller RS (1990) Frictional study of IC processed micromotors. *Sens Actuators A* 21–23:180–183
- Tai YC, Fan LS, Muller RS (1989) IC-processed micro-motors: design, technology and testing. In: *Proceedings of IEEE micro electro mechanical systems*, pp 1–6
- Tambe NS, Bhushan B (2005a) Nanotribological characterization of self assembled monolayers deposited on silicon and aluminum substrates. *Nanotechnology* 16:1549–1558
- Tambe NS, Bhushan B (2005b) Identifying materials with low friction and adhesion for nanotechnology applications. *Appl Phys Lett* 86:061906-1 to -3
- Tambe NS, Bhushan B (2005c) Micro/nanotribological characterization of PDMS and PMMA used for BioMEMS/NEMS applications. *Ultramicroscopy* 105:238–247
- Tanaka M (2001) Development of desktop machining microfactory. *Riken Rev* 34:46–49
- Tang WC, Lee AP (2001) Defense applications of MEMS. *MRS Bull* 26:318–319. Also see www.darpa.mil/mto/mems
- Tanner DM, Smith NF, Irwin LW, Eaton WP et al (2000) MEMS reliability: infrastructure, test structures, experiments, and failure modes, SAND2000-0091. Sandia National Laboratories, Albuquerque, New Mexico. Download from www.prod.sandia.gov
- Tao Z, Bhushan B (2005a) Bonding, degradation, and environmental effects on novel perfluoropolyether lubrications. *Wear* 259:1352–1361
- Tao Z, Bhushan B (2005b) Degradation mechanisms and environmental effects on perfluoropolyether, self assembled monolayers, and diamondlike carbon films. *Langmuir* 21:2391–2399
- Taylor MR, Nguyen P, Ching J, Peterson KE (2003) Simulation of microfluidic pumping in a genomic DNA blood-processing cassette. *J Micromech Microeng* 13:201–208
- Timp G (ed) (1999) *Nanotechnology*. Springer, New York
- Tokachichu DR, Bhushan B (2005) Bioadhesion of polymers for BioMEMS. *IEEE Trans Nanotech* 5:228–231
- Trimmer WS (ed) (1997) *Micromachines and MEMS, classic and seminal papers to 1990*. IEEE Press, New York
- Turbill P, Beugeling T, Poot AA (1996) Proteins involved in the Vroman effect during exposure of human blood plasma to glass and polyethylene. *Biomaterials* 17:1279–1287
- Valiev K (1992) *The physics of submicron lithography*. Plenum Press, New York
- van der Berg A (ed) (2003) *Lab-on-a-chip: chemistry in miniaturized synthesis and analysis systems*. Elsevier, Amsterdam
- van der Wiel WG, De Franceschi S, Elzerman JM, Fujisawa T, Tarucha S, Kouwenhoven LP (2003) Electron transport through double quantum dots. *Rev Modern Phys* 75:1–22
- Venkatesan S, Bhushan B (1993) The role of environment in the friction and wear of single-crystal silicon in sliding contact with thin-film magnetic rigid disks. *Adv Info Storage Syst* 5:241–257
- Venkatesan S, Bhushan B (1994) The sliding friction and wear behavior of single-crystal, polycrystalline and oxidized silicon. *Wear* 171:25–32
- Vettiger P, Brugger J, Despont M, Drechsler U, Duerig U, Haeberle W et al (1999) Ultrahigh density, high data-rate NEMS based AFM data storage system. *Microelectron Eng* 46:11–27
- Voldman J, Gray ML, Schmidt MA (1999) Microfabrication in biology and medicine. *Annu Rev Biomed Eng* 1:401–425
- Wang S, Zhang Y, Abidi N, Cabrales L (2009) Wettability and surface free energy of graphene films. *Langmuir* 25:11078–11081

- Wei G, Bhushan B, Jacobs SJ (2004a) Nanomechanical characterization of digital multilayered thin film structures for digital micromirror devices. *Ultramicroscopy* 100:375–389
- Wei G, Bhushan B, Jacobs SJ (2004b) Nanoscale indentation fatigue and fracture toughness measurements of multilayered thin film structures for digital micromirror devices. *J Vac Sci Technol, A* 22:1397–1405
- Wei G, Bhushan B, Ferrell N, Hansford D (2005) Microfabrication and nanomechanical characterization of polymer microelectromechanical systems for biological applications. *J Vac Sci Technol, A* 23:811–819
- Wheeler AR, Moon H, Bird CA, Loo RRO, Kim C-J, JA JA, Garrell RL (2005) Digital microfluidics with in-line sample purification for proteomics analysis with MALDI-MS. *Anal Chem* 77:534–540
- Wise DL (2000) Biomaterials and bioengineering handbook. CRC Press, New York
- Woias P (2001) Micropumps—summarizing the first two decades. In: Mastrangelo CH, Becker H (eds) Proceedings of SPIE—microfluidics and BioMEMS, vol 4560. SPIE, Bellingham, Washington, pp 39–52
- Xia Y, Whitesides GM (1998) Soft lithography. *Angew Chem Int Ed* 37:550–575
- Xia Y, Kim E, Zhao XM, Rogers JA, Prentiss M, Whitesides GM (1996) Complex optical surfaces formed by replica molding against elastomeric masters. *Science* 273:347–349
- Xu S, Yamanaka J, Miyata I, Yonese M (2007) Effects of denaturation and association of collagen on adsorption behavior: two-dimensional nanostructure and its property. *Colloid Polym Sci* 8:899–906
- Yang YT, Ekinci KL, Huang XMH, Schiavone LM, Roukes ML, Zorman CA, Mehregany M (2001) Monocrystalline silicon carbide nanoelectromechanical systems. *Appl Phys Lett* 78:162–164
- Yasseen AA, Wu CH, Zorman CA, Mehregany M (2000) Fabrication and testing of surface micromachined polycrystalline SiC micromotors. *IEEE Electron Device Lett* 21:164–166
- Young D, Du J, Zorman CA, Ko WH (2004) High-temperature single crystal 3C-SiC capacitive pressure sensor. *IEEE Sens J* 4:464–470
- Zaghoul U, Papaioannou G, Bhushan B, Coccetti F, Pons P, Plana R (2011a) On the reliability of electrostatic NEMS/MEMS devices: review of present knowledge on the dielectric charging and stiction failure mechanisms and novel characterization methodologies. *Microelectron Reliab* 51:1810–1818
- Zaghoul U, Papaioanou GJ, Bhushan B, Coccetti F, Pons P, Plana R (2011b) New insight into reliability of electrostatic capacitive RF MEMS switches. *Int Microw Wireless T* 3:571–586
- Zhang Y, Pan C (2012) Measurements of mechanical properties and number of layers of graphene from nano-indentation. *Diam Relat Mater* 24:1–5
- Zhang S, Zeng XT, Tang ZG, Tan MJ (2002) Exploring the antisticking properties of solid lubricant thin films in transfer molding. *Int J Mod Phys B* 16:1080–1085
- Zorman CA, Fleischmann AJ, Dewa AS, Mehregany M, Jacob C, Nishino S, Pirouz P (1995) Epitaxial growth of 3C-SiC films on 4 in. diam Si(100) silicon wafers by atmospheric pressure chemical vapor deposition. *J Appl Phys* 78:5136–5138
- Zorman CA, Roy S, Wu CH, Fleischman AJ, Mehregany M (1998) Characterization of polycrystalline silicon carbide films grown by atmospheric pressure chemical vapor deposition on polycrystalline silicon. *J Mater Res* 13:406–412
- Zoval JV, Madou MJ (2000) Centrifuge-based fluidic platforms. *Proc IEEE* 92:140–153

Index

A

Abrasive wear, 538, 545, 580
Accelerated friction, 230, 630
Acceleration energy, 601
Accelerometer, 23, 255, 598
Acetylene (C₂H₂), 602
Acoustic emission, 188, 225
Adhesion, 4, 7, 16, 23, 24, 26, 35, 75, 85, 96, 142, 151, 155, 156, 163, 166, 177–179, 184, 221, 224, 230, 232, 244, 245, 277, 282, 295, 309, 310, 320, 325, 327, 330, 334, 352, 359, 373–376, 381, 389, 404, 411, 414, 418, 438, 440–442, 444, 448–450, 457, 458, 460, 461, 465, 471–473, 475–481, 483, 485–489, 491, 492, 494, 502, 503, 510, 542, 553, 554, 556, 565–567, 571, 573, 574, 576, 577, 582, 595, 596, 598, 601, 602, 610, 620, 621, 625, 626, 632, 641, 642, 644, 646, 652, 656–662, 664, 673, 674, 683, 684, 690, 692, 695, 698, 702, 703, 707, 713, 728, 732, 735, 737, 741–743, 749, 750, 754, 757, 761, 788, 793
complex, 330
controlled friction, 485
energy, 471, 481, 487, 492, 510
Hamaker constant, 416, 466
hysteresis, 205, 458, 478, 480, 487, 491, 492, 494, 502, 503, 512
measurement, 221, 479
of polymer, 309
primary minimum, 469, 479
rate-dependent, 457, 481
test, 178, 660
tip-surface, 4, 309, 713
Adhesion force, 457, 458, 465, 477–480, 486, 487, 512, 663, 742, 743, 761
quantized, 470
Adhesive
contact, 457, 466, 478, 482, 484, 486, 490, 492, 494, 503, 583, 692
Adhesive force, 12, 13, 49, 52, 53, 62, 65, 113, 115, 118, 120, 328, 347, 381, 389, 408, 411–416, 439–444, 447, 448, 465, 485, 535, 660–662, 664, 671, 673, 674, 676, 683, 684, 695–698, 702, 703, 706, 709, 713, 714, 728, 734, 736, 737, 742, 743, 755, 788, 791

measurement, 62, 65, 411, 447, 461
Adsorbate, 87, 340, 531, 689
Adsorbed water, 413, 441, 442, 606, 674, 684, 698, 700, 705, 706, 709, 742
AFAM. *See* Atomic force acoustic microscopy
AFM. *See* Atomic force microscopy
Air
damping, 21, 36, 54, 199, 408, 530, 540
induced oscillation, 66
pocket, 144, 147
Al₂O₃, 22, 395, 522, 653, 756, 761, 765, 771, 774
Al₂O₃-TiC, 395, 752, 753, 756, 761, 764, 766, 767, 771, 775
Alkanethiols, 346, 349
Alkylphosphonate SAM, 659, 670, 683
Alkylsiloxane film, 657
Ammonium, 481, 531, 724, 725
Amontons' law, 484, 488, 489, 491, 494, 535
Amorphous carbon
coating, 593, 596, 603–606, 611, 612, 768, 773, 782, 783
Amplitude modulation (AM), 43
AM. *See* Amplitude modulation
Anisotropy
of friction, 531
Antibody, 151, 154, 158, 159, 162–165
APCVD. *See* Atmospheric pressure chemical vapor deposition
Arc discharge, 599
Artifacts, 58, 104, 131, 137, 140, 141, 151, 390
As-deposited film, 446, 665
Atmospheric pressure chemical vapor deposition (APCVD), 256, 264
Atomic force acoustic microscopy (AFAM)), 35, 385
Atomic force microscopy (AFM), 690
adhesion, 7
binnig design, 36, 45
cantilever, 9
carbon nanotube tip, 58
commercial, 7, 9
contact mode, 9
control electronics, 45, 84, 85
dynamic mode, 42
feedback loop, 8
for UHV application, 524

- friction, 7, 24, 54
 instrumentation, 7, 33, 35, 66, 107, 161, 167
 interferometer, 69
 manufacturer, 9, 45, 166, 167
 mode, 9, 11, 35, 42, 47, 135
 piezo creep, 140
 probe construction, 54
 static mode, 42, 74
 surface height map, 621–623, 625
 surface imaging, 96, 309, 390
 test, 257, 258, 415, 425, 730, 736
 thermal drift, 8, 39, 45, 54, 18, 196
 thermal noise, 69, 76, 108, 110, 142
 tip, 4–6, 11, 24, 34, 42, 44, 47, 49, 54, 57, 59, 95, 96, 98, 107, 136, 139, 140, 142, 147, 148, 151, 152, 154, 157, 161–164, 167, 188, 258, 309, 318, 332, 373, 375, 376, 389, 392, 403, 406, 409, 410, 414, 415, 421, 425, 432, 439, 542–545, 575, 654, 664, 665, 673, 676, 679, 692, 693, 697, 700, 727, 728, 755, 764, 777, 780, 791
 tip sliding, 65, 124, 392, 448, 664
 variable temperature, 554
 vibration isolation, 21, 141
 wear, 178, 309, 373, 375, 376, 389, 411, 422, 426, 448, 449, 628, 641, 657, 660, 668, 684, 716, 726, 730, 764, 780, 792
- Atomic resolution, 5, 7, 8, 33, 35, 39, 43, 51, 54, 302, 375–377, 390
 atomic-scale, 43, 302, 390
 image, 5, 6, 33, 39, 43, 51
 imaging, 6, 43, 375, 377, 390
- Atomic-scale
 dissipation, 16, 17, 101, 107, 306, 333, 347, 348, 350, 356, 387, 391, 411, 418, 438, 481, 487, 499, 503, 527, 531, 540
 force measurement, 6, 21, 34, 65
 friction, 301, 302, 327, 331, 333, 338, 340, 361, 391, 394, 448, 530, 545
 image, 36
 stick-slip, 19, 335, 337, 392, 410, 530
- Attraction
 long-range, 473
- Au(111), 353, 444, 542, 543, 657, 659, 661, 662, 664, 667, 681
- Average contact resistance, 735
- Avidin, 142, 151, 152
- B**
- Background noise, 527
- Ball-on-flat
 techniques, 730, 736
 test, 728–730, 735, 736
 tribometer, 630, 726, 728, 735
 tribometer testing, 732
- Bare surface, 328, 329, 457, 693
- Barrier length scale, 145
- BDCS. *See* Biphenyldimethylchlorosilane
- Beam
 deflection FFM, 522
- Bending
 stiffness, 15, 41, 54, 57
- Berkovich
 pyramid, 186
 tip, 188, 189, 196, 232, 621, 768
- Bias voltage, 6, 8, 9, 36, 38, 39, 383, 526, 540, 601, 602, 612
- Bimorph, 18, 386
- Biocapsule, 645
- Biofluidic chip, 22
- Biofouling, 642, 645
- Biological membrane, 136, 471, 473
- Biomedical device, 256, 642, 645
- Biomolecule, 25, 136, 139, 144
- BioNEMS, 24, 253–255, 293, 642
- Biosensor, 22, 24, 25, 255, 642, 645
- Biphenyl, 445, 659, 661, 665, 683
- Biphenyldimethylchlorosilane (BDCS), 446, 665
- Birefringent crystal, 72
- Block-like debris, 231, 631
- BMIM-PF₆, 725–728, 730, 734, 735, 737, 739, 741–743
- Bond-order potential, 304, 318, 320, 327, 336, 337
- Boundary
 film, 504, 689, 741, 784
 lubrication, 20, 356, 373, 375, 389, 438, 443, 448, 450, 482, 483, 498, 643, 675, 690, 756, 784, 786, 788
 slip, 473
- Broken coating chip, 621
- Bubble, 295, 472, 473
- Buckling, 230, 238, 241, 242, 244, 320, 324, 613, 614, 618–621, 625, 627, 631
 stress, 244, 619, 620
- C**
- C₆₀, 39, 40, 355, 356, 390, 432, 491, 536, 595
 film, 39, 40, 356, 390, 391
 island, 356
 ultralow friction, 356
- C₇₀, 594
- CaF₂
 tip, 317
- Calibration, 119, 194, 258
- Cantilever, 42, 50, 55, 57, 60, 72, 73, 85, 109–111, 122, 128, 129, 137
 axis, 128, 130
 bending, 283
 deflection, 9, 42–44, 46, 49, 52, 53, 59, 64–66, 73, 83, 101, 107, 108, 112, 120, 130, 142, 333, 389, 432, 520, 521, 696, 777
 deflection calculation, 120
 effective mass, 101, 105, 106, 408, 522
 eigenfrequency, 525
 flexible, 11, 42, 43, 47, 96, 376, 378, 661, 716, 754
 force sensing, 44
 free oscillating, 542

- material, 64, 98, 99, 101
microbeam, 199, 274, 275
microfabricated, 44, 55, 75, 98
motion, 42
mount, 49, 50
Q-factor, 101, 151, 152
resonance behavior, 68
resonance frequency, 79, 105–107, 151, 382, 388, 521, 522, 537
spring constant, 42–44, 52, 54, 55, 57, 61, 64–66, 99, 137, 141, 145, 167, 268, 376, 408, 521, 522, 526, 538, 660
stainless steel, 56, 102, 257, 380, 382, 621, 627, 755, 756
stiffness, 52, 64, 66, 73, 85, 95, 98–102, 104–107, 110, 113, 121, 122, 660
thermal noise, 76
thickness, 55, 64, 521
trajectory, 62
triangular, 54–57, 68, 75, 98, 100, 101, 107, 127, 131
untwisted, 12, 48, 62, 96
V-shaped, 54, 55, 57, 96, 523
- Cantilever beam
array (CBA), 257
fabrication procedure, 277, 294
lateral bending, 283, 294, 296
torsional vibration, 388
- Cantilever stiffness, 73, 95, 98, 101, 102, 385, 660
- Cantilever-tip assembly, 11, 47, 100, 378, 386–388, 401
- Capacitance detection, 44, 77, 520
- Capacitive
detection, 80
detector, 42
displacement sensor, 180, 181, 184, 385
- Capacitive-type silicon, 23
- Capillary
force, 7, 34, 458, 471, 475, 477, 664
pressure gradient, 700
- Capped nanotube, 323
tip, 323
- Carbon, 16, 24, 58, 59, 240, 320, 327, 352, 357, 359, 361, 407, 428, 446
crystalline, 594
magnetron sputtered, 602
unhydrogenated coating, 608
- Carbon-carbon distance, 320
- Carbon nanotube (CNT), 24, 58, 320, 355–358
bending, 358
tip, 59
- Carboxylates (RCOO⁻), 653
- Carrier gas, 606, 655
- Catalysis, 218
- Cathodic arc carbon, 612–614, 616, 626, 634, 782
- Cation attachment scheme, 739
- Cell death, 26
- Ceramic
nanobeam, 268
slider, 633
- Chain length effect, 674, 679, 680, 684
- Charge
exchange interaction, 464
- Chemical
binding, 144
bond, 325, 534, 543
bonding, 22, 43
bonding force, 43
characterization, 5, 95, 593
degradation, 641
heterogeneity, 481
interaction, 163
- Chemical bonding, 595
- Chemical vapor deposition (CVD), 59, 595
- Chemisorption, 147, 643, 658, 689, 741
- Chord theorem, 475
- Clay nanoparticle, 277
- Cluster, C₆₀molecule, 390
- Coagulation, 156, 472
- Coated
Si ball-on-flat test, 730
silicon, 123, 296, 713
tip, 52, 147, 149, 150, 716
- Coating, 22, 98, 225, 231, 241, 295, 610, 612, 621, 635, 691, 699, 730, 764, 773, 774, 780, 782, 786
continuity, 635
damage, 241, 243, 618
durability, 633
failure, 230, 241
friction and wear behavior, 411
hardness, 609, 611, 622, 626, 630, 632
microstructure, 604, 618
substrate interface, 232, 611, 613, 622, 624, 626
thickness, 618, 621, 627, 631, 632, 635, 713, 749, 768, 774, 781, 792
- Coefficient of friction (COF), 2, 3, 12, 49, 60, 62, 64, 65, 110, 111, 113–116, 118–120, 123, 127, 129, 189, 227, 230, 257, 266, 267, 279, 401, 406, 409, 412–423, 440–443, 484, 485, 489, 495, 505, 509, 523, 535, 556, 565–569, 572, 573, 574, 579, 582, 633, 657, 660, 661, 667, 673–676, 679, 680, 683, 695, 698, 700, 703, 706, 709, 710, 718, 720, 721, 725, 736, 743, 764, 789
average, 63, 661
influence of temperature, 443, 706
lubricant, 509
relationship, 409
- COF. *See* Coefficient of friction
- Coherence length, 71, 83
- Cold welding, 464
- Colloidal
force, 459
probe, 118, 120, 121, 123, 128–130
- Comb drive, 261
- Compression, 199, 220, 221, 223, 240, 242, 316, 325, 328, 350, 358, 445, 477, 664
- Compressive
forces nanotube, 59
stress, 230, 244, 291

- Computational study, tribological process, 301
 Computer simulation, 469, 487, 490, 498, 499, 505
 force, 508
 friction, 511
 Concentration
 correlation, 188
 critical, 539
 Constant
 current mode, 8, 36, 38, 526
 force mode, 74, 84, 137
 height mode, 8, 36–38, 137
 NVE, 308
 Constant amplitude (CA) mode, 525
 Contact
 adhesive, 13, 53, 413–415
 analysis, 13, 212, 437, 563, 566, 783
 apparent macroscopic, 490, 494
 area, 3, 13–17, 179, 186, 191–197, 212, 215, 238,
 283, 311, 314, 316, 318, 326, 330, 331, 338,
 346, 347, 352, 358, 359, 385, 390, 409, 412,
 413, 463, 477–479, 484–486, 489, 490, 491,
 494, 505, 507, 510, 525, 534, 535, 537–539,
 554, 556, 563–565, 567, 569, 570, 573, 576,
 579, 581, 582, 613, 675, 709, 720, 757, 761,
 762
 conductance, 538
 elastic, 16, 17, 186, 192, 204, 215, 217, 314, 316,
 413–415, 419, 563–568, 573, 575, 577, 578
 mechanics, 302, 334, 457, 477, 478, 490
 resistance, 716, 717, 724, 730, 732, 733, 735, 742,
 743
 stiffness, 215–217, 241–244, 262, 277, 279, 385–387,
 389, 402, 436, 520, 537, 610, 616–618, 620
 stress, 212, 220, 278, 409, 730
 Continuum
 model, 530
 Coulomb
 law of friction, 534
 Critical
 concentration, 539
 Cu(111), 539
- D**
- 2-D
 histogram technique, 536
 Damage
 mechanism, 622, 627, 631, 632
 Damped harmonic oscillator, 54, 107, 108, 196
 Damping, 21, 36, 54, 199, 403, 408, 410, 528, 530, 539,
 540
 constant, 54, 196
 mechanism, 851
 pneumatic, 802
 Deborah number, 497
 Debris, 225, 227–230, 366, 411, 420, 423, 621, 630, 631,
 715, 730, 735, 742, 764, 775
 Debye
 frequency, 307
 interaction, 464
 length, 467, 469, 472
 Decylphosphonate (DP), 839
 Defect
 motion, 103
 nucleation, 186
 production, 166
 Deflection, 00
 measurement, 108, 110, 197, 231
 noise, 21, 108
 Degradation, 157, 427, 680, 690
 Degradation of PFPE film
 mechanical scission, 720
 Deionized (DI), 279
 Delamination, 228, 239, 265, 430, 609, 614, 621, 627
 Delivery particle, 874
 Demnum-type PFPE lubricant, 447
 Density of fiber, 4, 71, 80
 Dependence
 indentation, 5
 MD simulation, 303
 Depletion
 attraction, 475
 force, 459, 471, 475
 interaction, 474
 stabilization, 475
 Deposition
 rate, 598, 599, 601
 technique, 208, 596, 599, 610, 644, 741, 843
 Deposition (PECVD)
 etching, 55, 97, 262
 Depth-sensing indentation, 177, 432, 436
 Derjaguin approximation, 469, 471
 Derjaguin–Muller–Toporov (DMT)
 adhesion force, 479, 486, 577
 model, 476, 483
 theory, 478
 Design rule, 67
 Detachment, 140, 145, 878, 880
 Detection system, 44, 69, 74, 76, 407, 528
 Dewetting, 439, 482, 697, 727, 836
 DFM. *See* Dynamic force microscopy
 DFS. *See* Dynamic force spectroscopy
 DI. *See* Deionized
 Diagnostics, 22, 136, 803, 877
 Diamond, 41, 56, 58, 102, 179, 192, 318, 334, 536, 529,
 605, 608, 843
 coating, 597, 606, 608, 610, 636
 film, 606, 607
 friction, 361
 nanoindentation, 327, 367, 376, 612, 613, 615, 756,
 787, 831
 tip, 399, 406, 408, 410, 545, 829, 836, 838
 Diamond-like carbon (DLC)
 coating, 644
 coating microstructure, 206, 242, 599, 626, 637, 644,
 834
 matrix, 632, 644, 751
 Diblock copolymer, 513

- Dicationic ionic liquid film nanotribological study, 734
Dielectric breakdown, 465, 467, 811, 875
Dielectric constant, 465, 467, 874, 875
Diffusion
 parameter, 143, 483, 694
 retardation, 465, 466
 rotational, 130, 807
 thermally activated, 394, 534, 545
Diffusion coefficient, translational, 695
Digital
 feedback, 8, 45, 84
 micromirror device (DMD), 23
 signal processor (DSP), 38
Digital mirror device (DMD) pixel, 23, 645, 801, 825, 829, 889, 890, 891
Dilation, 487, 500
Dimension, 5, 7, 12, 36, 44, 48, 53, 98, 105, 107, 109, 111, 227, 259, 305, 402, 485, 503, 529, 580, 755, 798, 799, 800, 804, 824, 876
Dimensionless particle position, 876
Dip coated tip, 874
Dip coating, 691, 712, 724, 726, 727
Direct-current plasma-enhanced CVD (DC-PECVD), 55, 57, 58, 64, 97, 262, 380, 597, 599, 604
Directed self-assembly, 665, 798, 899
DI silicon nitride, 379
Disjoining pressure, 459, 693, 700
Dislocation
 line tension, 552, 556
 motion, 483, 556, 559, 560
 nanoindentation, 318, 326
 nucleation, 185, 205, 330, 560, 596
 stick-slip, 330, 331, 336, 338
Dispersion
 force, 459, 464, 466
 interaction, 464, 465, 466
Displacement sensor, 179, 464, 881, 882
Dissipation
 image, 540
 measurement, 437, 438, 448
Dissipative interaction, 520
Dissociation constant, 144
Distance, cut-off, 465
Distortion, 45, 82, 83, 86, 87, 140, 324, 333
DLC. *See* Diamond-like carbon
DMT. *See* Derjaguin–Muller–Toporov
DNA
 AFM image, 802
Domain pattern, 143, 232, 295
Double-layer
 force, 460
 interaction, 467, 468
Doubly clamped beam, 806
DP. *See* Decylphosphonate
Drag, 2, 337, 495, 496, 502, 545, 546, 645, 676, 757, 848, 849, 850
Drift state, 824
Driving frequency, 69, 151
Droplet evaporation, 820
Drug
 delivery, 23–25, 216, 823, 824, 855, 874
 delivery device, 616
Dry
 sliding friction, 329, 338
 superadhesive, 11, 12, 48, 62, 112, 115
Dry surface
 force, 409
 friction, 409, 493
DTSSP. *See* 3,3'-dithio-bis (sulfosuccinimidylpropionate)
Dupre' equation, 482, 662
Durability, 55, 228, 389, 438 442, 450, 621, 632, 643, 657, 681, 690, 710
Dynamic
 AFM, 377
 AFM mode, 392
 interaction, 459
 light scattering (DLS), 525
 mode, 42, 43, 196
 operation mode, 110, 127, 130, 389
Dynamic AFM. *See* Dynamic atomic force microscope
Dynamical response, 500
Dynamic atomic force microscope (dynamic AFM), 377
Dynamic force microscopy (DFM), 11, 47, 387, 756
Dynamic force spectroscopy (DFS), 145
- E**
- EAM. *See* Embedded atom method
EBD. *See* Electron beam deposition
e-beam lithography, 896
ECR-CVD. *See* Electron cyclotron resonance chemical vapor deposition
EELS. *See* Electron resonance loss spectroscopy
Effective
 damping constant, 54, 196
 mass, 65, 101, 104–106, 125
 shear stress, 538
 spring constant, 52–54, 143, 165, 166, 310, 528
 viscosity, 495, 497, 498, 501, 509
Effect of cycling, 716
EHD or EHL. *See* Elastohydrodynamic lubrication
Elastic
 contact, 202, 213, 215, 403, 413–415, 417–419, 437, 550, 563–568, 572, 573, 575, 577, 578, 582, 783
 deformation, 16, 17, 193, 280, 329, 391, 422, 438, 449, 522, 535, 573, 574, 760, 761
 limit, 21, 318, 429, 430, 431, 809
 modulus, 176, 182, 184, 190, 193, 197, 207, 209, 210, 213, 214, 217, 218, 220, 223, 231, 238, 242, 254, 257, 262, 263, 264, 267, 268, 273, 276–280, 283, 293, 577, 613, 616, 620, 782, 788, 872
 tip-surface interaction, 4, 151, 318, 332, 374, 528
Elasticity, 16, 85, 123, 177, 190, 193, 346, 385, 389, 432, 437, 438, 499, 534, 556, 577, 756, 760, 783, 827

- Elastohydrodynamic lubrication (EHD or EHL), 495
 Electrical resistance, 24, 716, 825
 Electric force gradient, 52
 Electrochemical
 AFM, 53
 STM, 40
 Electrokinetic flow, 820
 Electromagnetic force, 464
 Electron
 cyclotron resonance chemical vapor deposition (ECR-CVD), 206, 238, 242, 597, 599, 600, 602, 611, 612, 618, 622, 624, 628–631, 635, 768, 769, 774, 793, 843
 energy loss spectroscopy (EELS), 601, 603–605, 610, 637
 interaction, 720
 tunneling, 36
 Electron cyclotron resonance chemical vapor deposition (ECR-CVD), 206, 238, 242, 597, 599, 600, 602, 611, 612, 618, 622, 624, 628–631, 635, 768, 769, 774, 793, 843
 coating, 618, 628, 630, 631, 635, 774
 Electronegativity, 646
 Electronic noise, 45
 Electrostatic
 actuation, 182, 183, 385, 880
 force, 104, 182, 464, 467, 470, 807, 875, 876
 interaction, 304, 466, 467, 736
 Electrostatic double layer (EDL), 467, 469, 875
 Embedding energy, 305
 Endurance limit, 239, 431, 616, 617
 Energy
 dissipation, 107, 329, 333, 343, 347, 418, 481, 487, 499, 527, 531, 539, 540, 554, 556, 665, 677
 landscape, 144–146
 MD simulation, 311
 Energy dissipation, 107, 329, 333, 343, 347, 418, 481, 487, 499, 527, 531, 539, 540, 554, 556, 665, 677
 Entangled state, 483
 Exchange
 carrier plate, 525
 interaction, 464
 External
 vibrations, 8, 39, 140
- F**
 Fabricated
 pillar array, 255, 809
 silicon, 874
 Face-centered cubic (Fcc), 594
 Failure mechanism, 275, 422, 427, 630
 Fatigue
 crack, 242, 270, 272, 620, 631
 damage, 241–244, 617, 618
 failure, 238
 life, 238, 242, 243, 270, 610, 612, 613, 618, 620
 measurement, 610
 resistance, 16, 19, 232, 254
 strength, 253, 259, 270, 293
 test, 240, 241, 244, 246, 262, 264, 270, 616, 619
 FBS. *See* Fetal bovine serum
 FCA. *See* Filtered cathodic arc
 Fcc. *See* Face-centered cubic
 Fe-coated tip, 596, 751, 805
 Feedback
 circuit, 11, 45, 61, 83, 96
 loop, 8, 39, 45, 47, 74, 75, 84, 85, 87, 137, 138
 network, 36
 signal, 378
 FEM. *See* Finite element modeling
 Fermi
 level, 714
 point, 714
 Ferromagnetic probe, 86
 FET. *See* Field effect transistor
 Fetal bovine serum (FBS), 872
 FFM. *See* Friction force microscopy
 FIB. *See* Focused ion beam
 Fiber
 axial stress, 231
 buckling condition, 619–621, 624, 627, 631
 elastic modulus, 611, 613, 616, 620, 626
 fracture, 612, 626, 828, 830
 material, 166
 model, 72, 120
 morphology, 856, 865
 optical interferometer, 69, 71
 stiffness, 526
 Fiber-optic sensor, 881, 882
 FIB-milled probe, 58
 Field Effect Transistor (FET), 820
 Filling angle, 165, 294
 Film
 C₆₀, 39, 40, 356, 391
 nanoindentation thickness, 312, 314
 on substrate, 199
 substrate interface, 230, 613, 626
 surface interface viscosity, 409
 Film ionic liquid, 439, 689, 690, 723–725, 732, 741, 743
 Filtered cathodic arc (FCA)
 coating, 241, 611, 613, 618, 621, 630, 632, 768
 deposition, 599, 611
 Finite Element Method (FEM), 190, 199, 201, 253
 Finite Element Modeling (FEM), 257, 259
 First principle
 calculation, 333, 545
 MD simulation, 303, 309
 simulation, 331
 Flat punch, 188, 190, 195, 217–219, 221
 Flexible cantilever, 11, 42, 43, 47, 96, 376, 378, 661, 717, 754
 Flexible crosslinker
 property, 55, 116
 spring constant, 44, 99
 Flocculation, 469
 Flory

- radius, 474
temperature, 474
- Flow
nanoliters, 816
rate, 602, 603, 609
- Fluctuation interaction, 464
- Fluid
low viscosity, 509, 497
mechanics, 2
multicomponent, 573
Newtonian, 509
- Fluorescence
correlation spectroscopy (FCS), 620
fluctuation, 605
recovery after photobleaching (FRAP), 817
spectroscopy, 603, 605, 626, 737
- Fluoride, 895
- Fluorophore, 162
- FM-AFM, 43
- Focused ion beam (FIB), 41, 57, 58, 71, 73, 80, 83, 120
- Force
adhesion, 142, 458, 460, 465, 476, 477, 487, 587, 663, 743, 761
adhesion, quantized, 471
attractive, 15, 42, 43, 52, 60, 61, 113, 144, 317, 330, 381, 458, 459, 460, 488, 646
between macroscopic bodies, 462, 464, 465, 468
between surfaces in liquid, 458, 466
between surfaces in vacuum, 457
calibration, 12, 47
calibration mode, 51, 52, 381, 389, 754
cantilever-based, 519
capillary, 7, 458, 471, 475, 477, 664
Casimir, 459
charge fluctuation, 522
chemical force, 308
detection, 135, 162
electrostatic force, 7, 35, 104, 108, 109, 182, 464, 467, 470, 807, 875, 876
gradient, 42, 43, 52, 53, 144, 158, 208, 409, 435, 460, 511, 543, 545, 551
long-range, 473, 697, 875
measurement, 6, 85, 162, 188, 381, 403, 406, 407, 526, 620, 709, 761
microscopy, 11, 47, 527
modulation, 385–387, 437, 783
resolution, 141, 194
sensing tip, 44, 80
sensor, 11, 47, 80, 178, 245, 384, 524, 542, 882
spectroscopy (FS), 85, 136, 140, 143, 145, 148, 153, 164, 167
van der Waals force, 52, 381, 412, 414, 415, 459, 464, 466, 467, 470, 471, 595, 662, 777, 811, 874–876
- Force calibration plot (FCP), 439, 660, 696, 697, 836
- Force-distance
curve, 52, 53, 120, 163, 165, 328, 381, 693, 868
profile, 52, 142, 143
- Force-driven dissociation, 146, 147
- Force field spectroscopy, 3-D, 85, 140, 142, 144, 145, 148, 153, 164, 167, 542
- Force-free dissociation, 146
- Force-probe experiment, 147
- Fourier-transform infrared spectroscopy (FTIR), 727, 738, 740, 743
- Four-quadrant photodetector, 520
- Fracture
failure, 621
strength, 234
stress, 257, 258, 293
surface, 271, 272, 281, 293
toughness, 175, 176, 208, 210, 225, 230, 232–234, 237, 238, 242–244, 251, 253, 255, 258, 259, 263, 264, 269, 270, 610, 612, 613, 615, 616, 620, 622, 626, 637, 833
toughness measurement, 234, 258, 610
- Frank–Read source, 558, 560
- Free surface energy, 439, 698, 705, 836, 837
- Freezing-melting transition, 504
- Frenkel–Kontorova–Tomlinson (FKT) model, 529, 530
- Frequency
AFM image, 45, 51, 52, 66, 119, 140, 141, 158, 227–229, 331, 411, 423, 424, 428, 432, 621–623, 628, 667, 668, 770, 780, 892
frequency modulation (FM), 43
measurement precision, 80
- Friction
adhesive component, 418
anisotropy, 350, 448, 531
coefficient, 2, 7, 75, 307, 327, 331, 333, 334, 337, 338, 344, 352, 353, 355, 357, 358, 482, 483, 490–492, 494, 495, 497, 506, 508, 509, 530, 540, 541, 552, 658
dominant, 410
experiments on atomic scale, 301, 302, 327, 332, 338, 340, 520, 524, 530, 534, 545
image, 446
kinetic, 3, 486, 487, 489, 494, 497, 499, 502, 504, 506, 508, 509, 528, 529, 544, 886
load dependence, 420, 488, 530, 579
loop, 117–120, 127, 345, 347, 524, 528, 530, 538, 539, 545
lubrication, 305, 482
map, 309, 392, 404, 410, 411, 497, 530, 532, 763
measurement, 54, 55, 57, 59, 65, 100, 110, 337, 399, 401, 463, 526, 695, 728, 788, 829, 860
measurement method, 59, 110
measuring technique, 500
melting, 501, 507
metal surface, 331, 524, 544
mode, 96, 378
molecular dynamics (MD) simulation, 303, 308, 324, 330, 331, 337, 341, 344–346, 554
of biomolecule, 865
plane, 397, 401, 425
profile, 12, 48, 63, 64, 225, 227, 228, 264, 266, 278, 391, 393, 400, 621–626, 668, 763, 769, 770
scale dependence, 415, 418–420, 549, 550, 553, 556

- self-assembled monolayer (SAM), 527
 sliding, 329, 330, 338
 surface, 345, 683
 test, 660, 730, 736
- Friction and adhesion
 relative humidity, 411
 temperature effect, 442, 675, 695
 velocity effect, 440, 676, 702
- Friction force
 calibration, 12, 48
 curve, 86
 map, 391–397, 405, 537, 672, 673, 698, 761, 763, 765
 measurement, 403, 406, 407
 microscopy (FFM), 485
 profile, 225, 393, 405
 of Si(100), 838
- Friction force microscopy (FFM)
 signal, 12, 115, 406
 tip, 337
- Friction model
 cobblestone model, 485–489
 coulomb model, 489
 distance-dependent model, 504
 interlocking asperity model, 486
 phase transitions model, 505
 rate-and state-dependent, 507
- Friction-torque model, 886
- FS. *See* Force spectroscopy
- FTIR-TRS. *See* Fourier-transform infrared time-resolved spectroscopy
- Fullerene
 film, 356, 432
- Fused silica, 194, 202, 203, 207
- G**
- GaAs/AlGaAs
 heterostructure, 208, 211
- GaP(110), 77, 203, 271
 gap stability, 36
 gauge marker, 16
- Gaussian surface, 419, 563, 568
- Gd on Nb(110), 77
- Geometrically Necessary Dislocations (GND), 556, 557, 582
- Germanium, 531
- Giant MR (GMR), 751
- Glass
 glassiness, 511
 transition temperature, 254, 279, 293, 474
- GMR. *See* Giant MR
- GND. *See* Geometrically necessary dislocations
- Gold, 9, 25, 39–41, 51, 57, 147, 150, 217, 218, 239, 252, 254, 304, 312–315, 425, 529, 540, 653, 657, 820, 824
- Grahame equation, 467
- Grain
 boundary, 176, 266, 313, 314
 nanoindentation boundary, 313–315
- size, 314, 596
- Graphite, 37, 51, 125, 128, 331, 356, 357, 390, 394, 410, 448, 530, 534, 576, 579, 594, 605, 633, 762
 cathode, 599
 flake, 356
 sheet, 125, 128, 356
- Green chemistry, 724
- Griffith fracture theory, 273
- H**
- Hall-Petch behavior, 846
- Hamaker constant, 466, 470, 700, 875
- Hard amorphous carbon coating, 597, 773
- Hard coating, 221
- Hardness
 scale effect, 208, 218, 251
- HARMEMS. *See* High-aspect-ratio MEMS
- Harmonic oscillator, 107, 108, 196
- Harpooning
 interaction, 459
- HDT. *See* Hexadecane thiol
- Heavy-ion bombardment, 831
- Height profile BMIM-PF₆ coating, 729, 731–733
- Hemodynamic resistance, 876
- Hertzian contact model, 437, 488, 575, 783
- Hertz model, 166, 478
- Hertz-plus-offset relation, 535, 536
- Hertz theory, 166, 258, 415, 478
- Heterodyne interferometer, 70, 73
- Hexadecane thiol (HDT), 444, 678
- High-aspect-ratio MEMS (HARMEMS), 800
- Highly oriented pyrolytic graphite (HOPG), 52, 390
- High-pressure phase
 silicon, 318
- High-resolution
 FM-AFM, 6, 7, 140, 158, 159, 376, 754
 imaging, 6
 spectroscopy, 7, 140
- High-speed data collection, 407
 High-temperature
 MEMS/NEMS, 893
 operation STM, 895
 superconductivity (HTCS), 825
- High-temperature superconductivity (HTCS), 825
- Hill coefficient, 116
- H₂O (water), 527, 633, 634, 648, 653, 675, 703
- Homodyne interferometer, 69, 73
- Honeycomb-chained trimer (HCT), 356
- Hooke's law, 141
- HOPG. *See* Highly oriented pyrolytic graphite
- Horizontal coupling, 8
- Horizontally arranged nanotube, 357
- HPI. *See* Hexagonally packed intermediate
- HTCS. *See* High-temperature superconductivity
- H-terminated diamond, 333
- Human genome project, 803
- Humidity, 356, 414, 416, 441, 442, 632, 658, 660, 674, 709, 734, 805, 842, 884, 888

- HUVEC. *See* Human umbilical venous endothelial cell
- Hybrid
- continuum-atomistic thermostat, 308
 - nanotube tip, 59
- Hybridization, 351
- Hydration
- force, 459, 470, 472, 473, 496
 - regulation, 472
- Hydrocarbon
- precursor, 597, 598, 653
 - Hydrodynamic
 - damping, 410
 - force, 459, 495
 - radius, 126
 - stress, 201
- Hydrodynamic
- bonding, 326, 473, 737, 743
 - bond network, 140, 345, 650, 720
 - concentration, 604, 606
 - content, 603, 609, 610, 613, 694
 - flow rate, 609
- Hydrogenated
- carbon, 603, 608
 - coating, 602, 603, 608
- Hydrogen-terminated diamond, 322–324, 327, 333, 334, 336, 339–341, 350, 359
- Hydrophilic, 352, 353, 407, 409, 439, 460, 471, 472, 476, 477, 493, 494, 509, 527, 651, 652, 653, 802, 836, 839
 - surface, 409, 471, 472, 476, 858
- Hydrophilicity, 471, 494, 654
- Hydrophobic
- attraction, 459
 - force, 472, 473
 - interaction, 472, 473
 - surface, 407, 442, 449, 472, 473, 476
- Hydrophobicity, 412, 414, 643–645, 652, 662, 674, 683, 697, 709, 841, 887, 888
- Hysteresis
- adhesion, 458, 481, 487, 491, 492, 494, 503, 511
 - contact angle, 483
 - loop, 45, 202, 505
- I**
- IB *See* Ion beam
- IBD *See* Ion beam deposition
- IL structure, 729
- Image
- effects, 6, 7, 33, 43, 164
 - processing software, 6, 43, 87, 138, 151, 158, 164, 184
 - topography, 152
- Imaging
- atomic-scale, 43, 58, 375, 390
 - bandwidth, 54
- Imidazolium, 724, 725, 734, 737, 739, 741
- Immunoisolation, 654, 821, 823
- Indentation
- creep, 210, 212, 215, 244, 255, 278, 437
 - depth, 5, 6, 176–179, 182, 187, 188, 190–194, 199, 206–208, 212, 214, 215, 234, 237, 244, 255, 262, 263, 315, 434, 610, 780, 830
 - fatigue damage, 242, 243, 618, 619
 - hardness, 35, 176, 177, 190, 207, 218, 384, 436, 450, 559, 756, 780, 793, 831
 - induced compression, 240
 - rate, 312, 317
 - size effect (ISE), 208
- Inelastic regime, 251, 254
- Initial contact, 42, 309
- In situ sharpening of the tips, 38
- Instability, 716
- Integrated tip, 55, 75, 98, 120, 121, 130, 387, 388, 811
- Interaction
- charge exchange, 459, 464
 - Debye, 464
 - depletion, 474
 - Derjaguin–Landau–Verwey–Overbeek (DLVO), 469
 - dispersion, 464–466
 - double-layer, 467, 468
 - dynamic, 459
 - electrostatic, 304, 466, 467, 736
 - energy, 462, 464, 465, 467, 468, 471, 472
 - energy gradient, 144
 - harpooning, 459
 - hydrophobic, 472, 473, 858
 - Keesom, 464
 - London dispersion, 464
 - nonequilibrium, 481
 - orientational, 464
 - osmotic, 475
 - polymer, 474
 - rate-dependent, 481
 - van der Waals, 462, 464, 466, 470, 658, 693, 848, 875, 876
- Interaction force
- cantilever deflection, 98, 140
- Interatomic
- attractive force, 61, 113
 - force, 43, 302, 303, 306, 318, 394, 414, 448
 - spring constant, 43
- Interdiffusion, 481, 487
- Interdigitation, 481, 487
- Interfacial
- defect, 240, 242, 618, 620
 - force, 13, 376, 489
 - friction, 13, 410, 457, 485, 487, 492, 494, 495, 498, 500
 - stress, 618, 625, 626, 632
- Interferometric detection sensitivity, 73
- Interleave scanning, 386
- Intermediate or mixed lubrication, 482
- Intermittent contact mode, 11, 47, 378

Internal stress, 223, 230, 613
 Intramolecular hydrogen bonding, 741
 Intrinsic
 stress, 608
 Ion
 implantation, 41, 205, 435, 601, 828–831
 source, 601
 Ion Beam (IB)
 coating, 630, 774
 Ion Beam Deposition (IBD), 596, 597, 600, 601
 Ionic bond, 44, 317, 333, 459, 539
 Ionic liquid (IL)
 dicationic, 726, 734
 film, 689, 690, 724, 726, 732, 734, 742, 744
 lubrication property, 726

J

Johnson–Kendall–Roberts (JKR)
 model, 481, 536
 relation, 535, 536
 theory, 536, 478, 479, 490
 Joule
 dissipation, 526, 540
 heating, 308, 526
 Jump into contact, 876
 Jump-to-contact (JC), 310

K

KBr, 531, 538, 539
 KBr(100), 538
 Keesom interaction, 464
 Kelvin
 equation, 477
 Kinetic
 friction, 3, 486, 487, 489, 494, 497, 499, 502, 504, 506, 507, 508, 509, 528, 544, 886
 Knife-edge blocking, 83
 Knoop hardness, 395, 761
 Kramers diffusion model, 146
 Kurtosis, 584, 877, 878

L

Lab-on-a-chip, 276, 803, 818
 system, 816
 Langevin dynamics approach, 307
 Langmuir–Blodgett (LB), 329, 644
 Laplace
 force, 662
 pressure, 458, 475, 476, 642
 Large-sample AFM, 96, 97
 Laser
 ablation, 898
 deflection technique, 42
 Lateral
 contact stiffness, 520, 537
 cracks, 204, 205, 210, 434

deflection, 100, 119, 120, 182, 869, 886, 887
 force, 5, 11, 34, 43, 47, 54, 55, 57, 68, 74, 75, 80, 95, 99, 100, 101, 110, 112, 118, 122, 125, 127, 129, 130, 186, 187, 262, 282, 332, 336, 345, 357, 358, 376, 401, 410, 448, 485, 486, 520, 523, 528, 530, 533, 535, 539
 resolution, 5, 10, 36, 38, 43, 47, 51, 135, 137, 141, 154, 180, 374, 378, 437, 447, 525, 643, 757, 760
 spring constant, 54, 125, 521, 538
 stiffness, 85, 99, 100, 101, 110, 122, 127, 180, 388, 390, 528, 537
 Lattice imaging, 43
 LB. *See* Langmuir–Blodgett
 Lennard–Jones (LJ) potential, 305, 415, 416
 LiF
 surface, 333
 Lifshitz theory, 465, 470, 471
 Lift mode, 53, 382, 386
 Lift scan height, 386, 387
 LIGA technique, 254
 Light beam deflection galvanometer, 73
 Linearization, active, 84
 Linear Variable Differential Transformer (LVDT), 83
 Line scan, 404, 405
 Liquid
 bridging, 888
 lubricant, 26, 326, 342, 358, 438, 442, 493, 495, 505, 506, 696, 700, 710, 777, 834, 836, 888
 lubricated surface, 495
 Liquid-air interface, 700
 Liquid film thickness, 446, 447, 551, 700, 701
 Liquid-like Z–15, 698, 788
 Lithographie galvanoformung abformung (LIGA) fabrication, 895
 Lithography, 294, 382, 644, 798, 800, 823, 895–898
 Live cell, 136, 138, 164, 166–168
 Load
 critical, 220, 223, 225, 227–230, 238, 239, 242, 243, 258, 263–266, 422, 446, 449, 612, 613, 621, 622, 624, 627, 630, 667–669, 677, 679, 768, 839, 842, 891
 dependence of friction, 530
 Load-carrying capacity, 632, 768
 Load-controlled friction, 485, 488, 491, 492, 494
 Load curve, 235–238, 279, 314–316, 318, 381, 385, 434, 446, 536, 667, 695, 696, 788, 791, 839
 Load-displacement, 176, 182, 186–189, 192, 201, 208, 292, 318, 385, 431, 434–437, 610, 613, 615, 894
 curve, 182, 186, 188, 189, 192, 201–206, 211–213, 218–220, 222, 235–237, 243, 254, 275, 281, 435, 436, 449, 610, 613–615, 894
 profile, 281
 Loading rate, 142, 145–147, 157, 182, 185, 186, 212–214, 255, 437
 Local deformation, 217, 218, 373, 377, 383, 421, 755
 Localized surface elasticity, 373, 385, 389, 437, 450, 783
 Local stiffness, 85, 446, 658, 665, 667, 683

- London dispersion interaction, 464–466
- Longitudinal piezo-resistive effect, 77
- Long-range attraction, 464
meniscus force, 381, 697
- Lotus (*Nelumbo nucifera*) effect, 643
- Low-noise measurement, 525
- Low pressure chemical vapor deposition (LPCVD), 254, 262, 827, 895
- Low-stress silicon nitride (LSN), 896
- Low temperature AFM/STM, 44
- Low-viscosity fluid, 497
- LPCVD. *See* Low pressure chemical vapor deposition
- Lubricant film, 359, 404, 439, 442, 447, 450, 493, 495, 504, 505, 507, 633, 683, 691, 693, 694, 696, 697, 706, 709, 711–713, 718, 720, 723, 734, 735, 741, 756, 785, 788, 808, 813, 834, 836, 838, 867, 884
thickness, 450, 694, 756
- flow, 700
- fraction, 728, 730, 743
- liquid, 463, 889
- MEMS/NEMS application, 253, 439, 690, 711, 724, 827, 843
- nanotribological performance, 695
- perfluoropolyether (PFPE), 22, 329, 330, 404, 407, 438, 447, 633, 690, 724, 751, 753, 756, 757, 786, 791, 830, 834, 893, 895
- solidlike, 691, 888
- Z-15, 439, 695–697, 710, 788, 834
- Z-DOL, 404, 691, 692, 698
- Lubricated silicon, 403, 710
surface, 375, 411, 464, 484, 485, 495, 709, 888
- Lubricating film, 641, 788, 870
thin film, 341
- Lubrication, 2–4, 7, 20, 301, 302, 305, 308, 326, 328, 342, 353, 374, 376, 389, 438, 439, 448, 450, 459, 482, 484, 493, 495, 497, 594, 643, 658, 675, 690, 695, 724, 726, 743, 750, 756, 784, 790, 834, 838, 863, 865, 889, 893
- elastohydrodynamic (EHD), 482, 497
- intermediate or mixed, 482
- mobile layer, 887
- nanoscale boundary, 389, 743
- LVDT. *See* Linear variable differential transformer
- M**
- Macrohardness, 435
- Macromolecule, 462, 468, 474
- Macroscale friction, 403, 630, 656, 762, 828, 829, 830, 851
- Magnetic disk, 27, 208, 214, 225, 598, 616, 633, 695, 762, 787, 792, 802, 806, 830, 882
- disk drive, 806, 830, 598
- force, 7, 35, 52, 53
- force gradient, 53
- storage device, 21, 26, 176, 374, 596, 637, 642, 750, 753, 754
- thin-film head, 634
- Magnetron sputtered carbon, 602
- Mass effective, 66, 101, 104–106, 125, 131, 408, 522
- Material lubricant, 355
structural, 275, 827
- Maximum adhesive force, 657
- Mechanical resonance, 508
scission, 720
- Mechanics nonadhesive, 457
of cantilever, 66
resonance spectroscopy, 385
- Mechanism-based strain gradient (MSG), 556
- Melted SAM, 675
- Membrane deflection method, 11, 44, 47, 85, 378
- Memory distance, 458
- Meniscus formation, 412–414, 439, 645, 709, 727, 728, 734, 787, 889
of liquid, 448
- Meniscus force, 52, 381, 408, 412–415, 447, 550, 556, 572, 573, 582, 697, 700–702, 709, 777, 804
- Metal evaporated (ME) tape, 597, 757
oxide, 16, 24, 462, 653, 809, 820
particle (MP), 634, 757, 765, 766, 786
particle (MP) tape, 634, 757, 765, 766, 786
tip, 36, 40, 310, 311, 321, 330, 331, 460
- Metal-evaporated (ME), 428
- Metallic bonding, 304, 361, 458
- Metal-oxide-semiconductor field-effect transistor (MOSFET), 24, 820
- Methylene stretching mode, 656, 657
- Magnetic Force Microscopy (MFM), 7, 35, 52, 53
- Mica, 6, 13, 15, 16, 18, 19, 21, 39, 40, 51, 52, 138, 139, 142, 143, 152, 158, 159, 331, 359, 375, 390, 391, 462, 466, 469–472, 476, 478, 479, 480, 489–494
- muscovite, 538
- surface, 12–16, 18, 51, 52, 138, 139, 158, 159, 165, 466, 469, 470, 471, 476, 478–480, 489–494, 497, 498, 502, 505–508, 511, 539
- Microcantilever, 24, 802, 811
- Microchannel, 816, 817, 820, 872
- Microcrystalline graphite, 607
- Microdevice, 252, 254, 645, 802, 820
- Microelectromechanical systems (MEMS)

- device, 252, 277, 800, 802, 806, 807, 808, 810, 825, 826, 834, 886, 896, 897, 898
- Microengine, 805, 806
- Microfabrication, 254, 825, 827, 896, 897
- Microfluidic, 22, 157, 816, 817, 819, 820
- Microfriction, 399, 832
- Microgear, 645, 805, 806
- Microimplant, 865
- Microindentation, 178, 230, 234, 664
- Micromachine, 24, 799, 804, 806, 809, 827, 878, 881, 886, 898
- Micromachined
- array, 24, 809, 881
 - polysilicon beam, 878
 - silicon, 804
- Micromirror, 22–24, 253, 259, 645, 670, 798–801, 809, 811, 889–891
- Micromirror device (DMD), 23
- Micro-/nanoelectromechanical system (MEMS/NEMS), 5, 21, 176, 252, 374, 553, 597, 642, 690, 797, 798
- device, 218, 239, 553, 617, 716, 797, 812, 827, 832, 834, 881, 893, 895
 - lubricant, 906
- Micro-/nanooptoelectromechanicalsystems (MOEMS/NOEMS), 798
- Micropattern, 446, 644, 664–666, 683
- Micropatterned SAM, 446, 665
- Micropump, 816, 817, 819, 820
- Micro-Raman spectroscopy, 603–605
- Microscale
- contact resistance, 732, 733
 - friction, 394, 395, 448, 449, 576, 656, 735, 762, 763, 792, 829, 831, 832, 878
 - scratching, 382, 421, 756, 764
- Microscratch, 220, 421, 423, 620, 621, 630, 637, 667, 764, 768
- Micsensor, 827
- Microslip, 419, 556, 559–561, 567, 582
- Microspark erosion, 898
- Microstructure, 22, 559, 603, 604, 800, 825, 878, 881, 886, 895
- characterization, 604
 - surface, 886
- Microsystem Technology (MST), 252, 798
- Microtip, 38, 411–413
- Microtriboapparatus, 843, 881, 882, 895
- Microtribometer, 355
- Microwear, 425, 620, 621, 627, 628, 630, 637, 673, 674, 771–773, 775, 832, 839, 841
- Millimeter-scale device, 898
- Millipede, 802, 811
- Misfit angle, 530
- Mobile lubricant, 354, 712, 713, 728, 742, 889
- Modulus
- of elasticity, 34, 177, 184, 190, 199, 213, 284, 373, 385, 413, 432, 437, 450, 608, 756, 760, 827, 846
- MOEMS/NOEMS. *See* Micro-/nanooptoelectromechanicalsystems
- Molding process, 294, 897
- Molecular
- chain, 345, 353, 445, 646, 652, 653, 665, 712, 720, 839
 - cohesion, 144
 - conformation, 643, 690, 784
 - dynamics (MD), 157, 302, 303, 308, 324, 330, 331, 337, 341, 344–346, 354, 355, 505, 534
 - dynamics simulation (MDS), 303, 308, 324, 330, 331, 337, 341, 344–346
 - imaging, 9
 - layer, 23, 27, 438, 443, 480, 502, 797, 838, 899
 - shape, 345, 457, 498
 - spring, 444–446, 450, 544, 658, 664, 665, 667, 674, 683, 839
 - spring model, 444–446, 450, 658, 664, 665, 683
- Molecular-beam epitaxy, 798, 800, 899
- Molecular Dynamics (MD), 157, 302, 303, 309, 324, 330, 331, 337, 341, 344, 345, 346, 354, 355, 534
- simulation, 303, 307, 324, 330, 331, 337, 341, 344, 345, 346
- Moment of inertia, 67, 99, 125, 126, 131, 257, 262, 288
- Monocationic, 726, 736, 739
- MoO₃, 332, 333, 355, 815, 816
- MoO₃ nanoparticle, 355
- Magnetoresistive, 802
- MOSFET. *See* Metal-oxide-semiconductor field-effect transistor
- MR. *See* Magnetoresistive
- MSG. *See* Mechanism-based strain gradient
- Multilayer
- thin-film, 22, 238, 616, 634, 753, 813
- Multilayered APTES, 870
- Multimode AFM, 50, 52
- Multimolecular layer, 438, 643, 644, 834
- Multipole interaction potential, 305
- Multiwalled Carbon Nanotube (MWCNT), 379
- Multiwalled Nanotube (MWNT), 58, 59, 380, 381, 825, 843, 844, 846, 848–850, 852–855, 893
- MWCNT. *See* Multiwalled carbon nanotube
- MWNT. *See* Multiwalled nanotube
- N**
- NaCl
- island on Cu, 540
- NaCl(100), 537
- Nanobeam, 257–262, 268, 270, 271, 286, 287–, 289–293
- Nanochannel, 642, 645, 820, 821, 823, 842, 895–897
- Nanochemistry, 24, 802
- Nanoclay, 277, 284, 294
- Nanocomposite material, 355
- Nanocrystal, 874, 899
- Nanocrystallite, 603
- Nanodeformation, 689–692, 742
- Nanodevice, 24, 59, 684, 800, 803

- Nanoelectromechanical systems (NEMS), 21, 24, 27, 251, 252, 642, 690
Nanoelectronics, 802, 803
Nanofabrication
 parameter, 432
Nanofatigue, 216, 232, 238, 616
Nanofluidic
 device, 252 803, 820
 silicon array, 822, 824
Nanohardness, 178, 208, 218, 384, 608, 610, 756, 763, 765, 775, 782, 821, 829–833
Nanoimprint lithography (NIL), 798, 898
Nanoindentation relax
 surface atom, 312
Nanoindenter, 176, 178, 179, 181–186, 190, 195, 196, 198, 206, 208, 210, 212, 230, 231, 242, 251, 254, 255, 261–263, 274, 275, 279, 293, 294, 436, 610, 616, 621, 626, 630, 637, 768, 833
Nano–Kelvin probe, 382, 427
Nanolithography, 24, 802
Nanolubrication, 726
Nanomachining, 24, 34, 374, 382, 421, 432, 449, 803
Nanomanipulation, 96, 130
Nanomechanical characterization, 175, 176
Nanomechanics, 1, 4, 6, 26, 27, 33, 373, 374, 750, 753, 754, 804, 827, 893, 895
Nanometer, 42, 52, 85, 104, 135, 162, 174, 176, 182, 376, 381, 406, 432, 437, 447, 458, 468, 471, 480, 481, 490, 526, 534, 536, 579, 618, 696, 757, 783, 798, 800, 813
Nanometer-scale
 device, 301
 friction, 331, 358
 indentation, 182, 301, 320, 361
Nanopattern, 802, 814, 865, 867, 870, 894
Nano/picoindentation, 385
Nanopores, 823
Nanorheology, 499
Nanoscale
 adhesion, 735, 797, 827, 856, 857
 bending, 256, 261, 894
 contact resistance, 732
 device, 742
 friction, 410, 411, 449, 479, 490, 643, 734, 741, 743, 843, 849, 859, 893
 indentation, 215, 756, 780
 test, 656
Nanoscope I, 7, 38
Nanoscopic device, 252
Nanoscratch, 76, 86, 219, 223, 243, 254, 255, 277, 425, 449, 792
Nanotribology, 1, 4, 5, 21, 26, 27, 33, 373, 519, 545, 642, 690, 753, 754, 804, 827
Nanotube
 bundle, 320, 357
Nanowear map, 411
Nanowire, 6, 124, 139, 177, 311, 312, 802, 815, 816
 NbSe_2 , 538, 541, 542
NC-AFM. *See* Noncontact atomic force microscopy
Negative
 contact force, 350
NEMS. *See* Nanoelectromechanical system
Newtonian
 flow, 497, 501
 fluid, 497
NHS. *See* N-hydroxysuccinimidyl
N-octadecyltrichlorosilane ($n\text{-C}_{18}\text{H}_{37}\text{SiCl}_3$, OTS), 657
Noise
 source, 76
Nomarski interferometer, 72, 73
Nonadhesive contact, 478
Nonconducting film, 41
Noncontact
 AFM, 42
 atomic force microscopy (NC-AFM), 42, 525, 526, 539, 540
 friction, 540
 imaging, 42, 54
 mode, 42, 58, 381, 813
Noncontact atomic force microscopy (NC-AFM), 42, 525, 526, 539, 540
Non-Newtonian flow, 483, 484
Nonpolar end group, 694, 834
Nonsilicon MEMS, 800
Nonspherical tip, 536
Nonwetting, 472
Normal friction, 484, 490, 491, 494
Normal load, 3, 12, 16, 17, 43, 49, 62, 66, 112–114, 116, 123, 125, 130, 177, 187, 220, 223, 225, 227, 229, 255, 260, 263–266, 276, 281, 326, 340, 345, 372, 378, 379, 382, 384, 385, 389–391, 393–397, 401–404, 410, 411, 413, 414, 420, 421, 423–426, 432–434, 439, 445, 446, 449, 480, 489, 492, 496, 551, 563–566, 575–579, 580, 582, 621, 623–625, 626, 630, 657, 660, 661, 664–669, 671, 674, 679, 692, 693, 695–697, 702, 710, 711, 718, 728, 756, 757, 761, 763–766, 768, 769, 771–778, 782, 785, 788, 791, 829–834, 836, 839, 841, 842, 849–853, 862, 863, 865, 869, 873, 873, 881, 889–892
Normal stiffness, 98–102, 104, 106–108, 109, 110, 116, 122, 123, 125, 130, 131, 378, 379, 387, 388, 390, 393, 537, 628
Nose–Hoover thermostat, 308
NSOM. *See* Near-field scanning optical microscopy
NTA. *See* Nitrilotriacetate numerical modeling

O

- Octadecyldimethyl(dimethylamino)silane (ODDMS), 669, 670
Octadecylphosphonate (ODP), 670, 839
Octadecyltrichlorosilane (OTS), 354, 657
Octyldimethyl (dimethylamino)silane (ODMS), 669, 670, 839

- Operation, 3, 7–11, 24, 34, 36, 37, 38, 42, 45–47, 49, 84, 87, 96, 97, 151, 153, 166, 375, 376, 378, 385, 432, 525, 695, 750, 754, 785, 803, 804, 808, 809, 816, 817, 826, 827, 865, 886, 889
- Optical
 detector, 42
 head, 49, 50
 lever, 73, 74, 75, 76, 80, 85
 microswitch, 801, 811
- Optical lever
 deflection method, 85
- Order
 parameter, 483, 502
- Organic
 compound, 646, 647, 650, 652, 661
 monolayer film, 395
- Organofunctional bond, 690
- Organometallic vapor-phase epitaxy, 899
- Oscillating
 cantilever, 53, 70, 72, 138, 377
 tip, 11, 378, 385, 526, 756
- Oscillation
 amplitude, 52, 53, 151, 153, 508, 526
 cycle, 151, 525
- Oscillatory
 force, 459, 469, 470–473, 479, 480, 508
 shear, 16, 17, 19
- Osmotic
 force, 475
 interaction, 475
 pressure, 469, 475
- Oxide
 layer, 425, 426, 439, 476, 524, 653, 691, 697, 712, 726, 881, 883, 897
- Oxygen content, 635
- P**
- Paraboloid load displacement, 188, 192
- Passive
 linearization, 82, 83
 structure, 81
- PBS. *See* Phosphate-buffered saline
- PDMS. *See* Poly(dimethylsiloxane)
- PDP group. *See* 2-pyridylidithiopropionyl group
- Peak indentation load, 186, 206, 208, 210, 236–238, 614
- Peak-to-peak load, 255
- Peak-to-valley (P-V), 236, 487, 789, 865, 878
- Peak-to-valley (P-V) distance, 878, 762, 789, 865, 878
- PECVD. *See* Plasma enhanced chemical vapor deposition
- PEG. *See* Polyethylene glycol
- Perfluorinated SAM, 672
- Perfluorodecanoic acid (PFDA), 811
- Perfluorodecylphosphonate (PFDP), 669, 678, 839
- Perfluorodecyltricholorosilane (PFTS), 669
- Perfluorodecyltriethoxysilane (PFDTES), 842, 872–874
- Perfluoropolyether (PFPE), 438, 690
 lubricant film, 691, 693, 694, 698
 chemical degradation, 690, 711
- surface topography, 698
- Perfluoropolyether (Z-DOL), 404
- Periodic
 boundary condition (PBC), 305
- Permanent dipole moment, 464
- Perpendicular scan, 47, 59, 111, 116
- Perturbation approach, 141, 151
- PES. *See* Photoemission spectroscopy
- PFDA. *See* Perfluorodecanoic acid
- PFDP. *See* Perfluorodecylphosphonate
- PFDTES. *See* Perfluorodecyltriethoxysilane
- PFPE. *See* Perfluoropolyether
- PFTS. *See* Perfluorodecyltricholorosilane
- Phase
 imaging, 784, 869
 transformation, 204, 208, 230, 318, 410, 411, 425, 595, 707
- Phonon excitation, 343, 526
- Phosphate-buffered saline (PBS), 857, 862, 864, 867, 868, 871–874
- Phosphazene lubricant, 712
- Photolithography (PL), 294, 296, 644, 645, 899
- Photoresist
 master, 294, 295
- Physical wear, 714
- Physisorption, 148, 474, 643, 681, 741
- Picoindentation system, 385, 434
- Piezoelectric
 ceramic material, 82, 86
 excitation, 406
 hysteresis, 45, 123
 stack, 84
 tube, 8, 38, 39, 51, 52, 61–63, 81, 82, 113, 128
- Piezoelectric
 drive, 406, 817
 leg, 525
 scanner, 42, 45
- Piezoresistive
 cantilever, 76, 112
 detection, 44, 76
- Piezoscanner, 86
- Piezotube calibration, 8, 39
- Pile-up
 nanoindent, 279
 surface atom, 312
- Pinning, 324, 532, 652, 881
- Pin-on-disk tribotester, 656
- PL. *See* Photolithography
- Plants, 155, 168
- Plasma enhanced chemical vapor deposition (PECVD), 380, 597–600, 809, 843
 carbon sample, 605
- Plastic deformation
 surface, 201, 205
- Plug polysilicon, 897
- PMMA. *See* Poly(methyl methacrylate)
- Poisson's ratio, 57, 68, 192, 199, 200, 242, 288, 432, 552, 577, 618
- Polishing, 3, 41, 426, 596, 832, 850, 897

- Poly(dimethylsiloxane) (PDMS), 252, 892
Poly(methyl methacrylate) (PMMA), 252, 254, 262, 274–279, 281, 293, 295, 872, 873, 899
Poly(propyl methacrylate) (PPMA)
 physical property, 254
Polycrystalline graphite, 604–606
Polydimethylsiloxane (PDMS), 295, 296, 496, 498, 645, 658, 872, 873, 898, 899
Polyethylene glycol (PEG), 148, 150, 160–163, 165, 167
Polyethylene terephthalate (PET), 428, 430, 751, 757, 762, 763, 768, 770, 772, 775, 785
Polymer
 bioMEMS, 254, 820, 824
 brush, 166, 496
 cantilever, 296, 820
 end-adsorbed, 475
 grafted, 343
 interaction, 474
 microbeam, 294, 295
Polymer beam, mechanical property, 283
Polymer (cont.) mushroom
 nonadsorbing, 475
 physisorbed, 352, 353, 356, 476, 644, 693, 742, 784
Polymeric
 magnetic tape, 401, 403, 421, 432, 434, 437
 microbeam, 275
Polysilicon
 cantilever beam, 881
 deposition, 897
 film, 254, 262, 263, 265, 266, 653, 827, 828, 831–833, 893, 895
 microstructure, 881
 undoped, 262–264, 266, 293, 832, 833
Polystyrene (PS)
 physical property, 276, 496
Polytetrafluoroethylene (PTFE)
 coated Si-tip, 355
Polyvinyl alcohol (PVA), 296
Pop-ins, 205, 434
Position, accuracy, 86, 180
Potential, 16, 36, 70, 144, 167, 304, 308, 318, 332, 339, 359, 392, 409, 415, 425, 468, 483, 529, 532, 545, 600, 640, 690, 713, 730, 784, 817, 874
Power dissipation, 525
Power spectra
 MD simulation, 338
PPMA. *See* Poly(propyl methacrylate)
PP. *See* Polypropylene
Pressure
 contact, 312, 410, 415, 417, 449, 488, 577, 578, 595, 850
 Laplace, 458, 475, 476, 642
 osmotic, 469, 475
Probability force distribution, 146
Probe
 colloidal, 118, 120–123, 128–130
 FIB-milled, 58
 surface, 314
Properties of a coating, 597
Property
 protective nitride layer, 897
 protein coating, 865
 protrusion force, 459
PS/clay. *See* Polystyrene/nanoclay composite
PS. *See* Polystyrene
PTFE. *See* Polytetrafluoroethylene
Pt-Ir. *See* Platinum-iridium
Pull-off
 cycle, 24, 811
 force. *See* Adhesion force, 478
Pull-off adhesion force, 478
Pulsed force mode, 385
PUR. *See* Polyurethane
PVA. *See* Polyvinyl alcohol
P-V distance. *See* Peak-to-valley (P-V) distance
PVS. *See* Polyvinylsiloxane pyramidal tip
PZT (lead zirconate titanate)
 scanner, 11, 47, 96, 378, 385
 tube scanner, 38, 49, 378, 754
- Q**
- QCM. *See* Quartz crystal microbalance
Q-control system
QELS. *See* Quasi-elastic light scattering
Q-factor, 101, 151, 152
Quad photodetector, 11, 47, 378
Quadrant detector, 74, 75, 83, 522
Quality factor Q, 54, 67, 101
Quantum
 box (QB), 802, 24
 corral, 802, 24
 dot transistor, 798, 799
 Hall regime
 tunneling, 4–6, 8, 33, 36, 41, 42, 44, 45, 69, 311, 374, 376, 802
Quartz Crystal Microbalance (QCM), 302, 340, 341
Quasi-static
 bending test, 255, 259, 261, 270
 mode, 872, 877, 878
- R**
- Radial cracking, 235, 237, 614
Radius
 of gyration, 474, 693, 694
Raleigh's method, 66
Raman
 spectra, 604–607
 spectroscopy, 603–605
Random
 rough surface, 563
Ranking of various SAMs, 661
Ratchet mechanism, 333, 398, 418, 549, 554, 556, 571–573, 576, 577, 763, 792
Readout electronics, 74, 75
Rebinding, 146
Receptor-ligand

- bond, 151
 interaction, 459
- Recognition
 imaging, 135, 150–153, 155, 167
- Rectangular cantilever, 54, 56, 57, 98, 101, 106, 107, 131, 522
- Reduced modulus, 192, 194
- Relative Humidity (RH)
 influence, 705, 736
- Relative stiffness, 387
- Relaxation time, 483, 484, 497, 509, 511
- Remote detection system, 74
- Residual
 stress, 175–177, 193, 228–231, 240–243, 253, 424, 430, 608, 610–613, 618–621, 625, 631
- Resonance
 curve detection, 54
- Retardation effect, 466
- Retract
 curve, 142, 143
- RICM. *See* Reflection interference contrast microscopy
- RMS. *See* Root mean square
- Room Temperature (RT), 16, 39, 76, 148, 210, 212, 258, 295, 311, 330, 534, 593, 596, 597, 637, 714, 723, 724, 727, 734
- Rotor-hub interface, 878, 888
- Rotor-stator interface, 804
- Roughness
 angle, 419, 420, 568, 569
 of surface, 582
 structure, 759
- RT. *See* Room temperature
- Rupture force, 145, 146
- S**
- Sacrificial
 layer, 296, 823, 825, 877, 896, 897
 layer lithography (SLL), 823, 896, 897
 oxide, 897
- SAM. *See* Scanning acoustic microscopy; self-assembled monolayer
- Sample holder, 64, 102, 600
- SATP. *See* (S-acetylthio)propionate
- Scan
 area, 12, 49, 381, 411, 432, 635, 660, 667, 754, 775
 direction, 12, 49, 50, 87, 330, 332, 352, 353, 381, 523
 frequency, 86, 389, 660
 head, 51
 range, 8, 39, 40, 81, 83
 rate, 9, 12, 39, 49, 50, 352, 375, 381, 389, 660, 754
 size, 9, 38, 39, 49, 50, 81, 381, 384, 389, 390, 394, 406, 423, 562, 660, 756, 761, 768
 speed, 83, 137
- Scanner, piezo, 8, 39, 63, 82
- Scanning
 acoustic microscopy (SAM), 7, 35
 capacitance microscopy (SCM), 7
 chemical potential microscopy (SCPM), 35, 308
 electrochemical microscopy (SECM), 35
 electron microscopy (SEM), 136, 264, 756
 electrostatic force microscopy (SEFM), 7, 35
 head, 8, 39, 525
 ion conductance microscopy (SICM), 35
 Kelvin probe microscopy (SKPM), 7, 35, 382
 magnetic microscopy (SMM), 7, 35
 near field optical microscopy (SNOM), 7, 35
 probe microscope, 33, 376
 probe microscopy (SPM), 7, 27, 34, 95
 speed, 12, 49, 50, 54, 85, 381, 389, 432, 754
 spreading resistance microscopy (SSRM), 716
 system, 84
 thermal microscopy (SThM), 35
 tunneling microscopy (STM), 311
- Scanning force
 acoustic microscopy (SFAM), 7, 35
 microscopy (SFM), 353
- Scanning Force Microscopy (SFM), 7, 35
- Scanning Probe Microscopy (SPM), 7, 27, 34, 95
- Scanning Tunneling Microscopy (STM), 311
 cantilever, 55
 probe construction, 40
 thermal drift, 8, 39, 45
 tip, 34, 45
 vibration isolation, 21
- Scratch
 critical load, 630, 632
 damage mechanism, 627
 depth, 255, 264–266, 276, 278, 331, 334, 765–768, 770, 831
 resistance, 175, 177, 187, 228, 255, 264
 test, 220, 223, 228, 254, 262, 275, 277, 425, 621, 774
- Scratching
 force, 331, 334
 measurement, 421
- Screening, 230, 462, 465, 473, 630
- Self-assembled
 growth, 527
 monolayer (SAM), 24, 147, 322, 443, 527, 644, 664, 811, 834, 898
- Self-Assembled Monolayer (SAM)
 chemically adsorbed, 444
 coated surface, 872
 deposition, 653
 on copper, 658
- Sensitivity, 18, 24, 34, 54, 57, 69, 73, 75, 76, 79, 122, 127, 130, 385, 522
- SFA. *See* Surface force apparatus
- SFG. *See* Sum-frequency generation
- SFM. *See* Scanning force microscope

- Shear
force. *See* Kinetic, friction; Static, friction, 16, 17, 20, 137, 342, 346, 353, 413, 457, 463
modulus, 57, 65, 98, 521, 576
strength, 334, 359, 414, 418, 419, 535, 556, 559, 561, 562, 565–568, 570, 572–575, 576, 579, 582, 703
stress, 19, 157, 231, 314, 345, 346, 355, 358, 409, 483–487, 492, 497, 509, 538, 576, 666
critical, 19, 483–487, 492
effective, 497, 538, 576
Shot noise, 73, 76
Si(001)(2 x 1), 338
Si(100)
coefficient of friction, 440, 442, 701, 705, 707, 792, 838
substrate surface, 692
Si(111) (7x7), 209, 434
Si-based surface, 865
Si cantilever
gold-coated, 123
Silane
bubbler, 654
polymer film, 870
Silanization process, 867
Silanol group (SiOH), 658
Silicon (Si)
adatom, 542
beam, 271
cantilever, 55, 387, 754
friction force, 728
microimplant, 865
micromotor, 695
nanobeam, 271
nitride layer, 55, 57, 147, 148, 421, 857, 897
surface
terminated tip, 351
tip, 333
single-crystal, 379
tribological performance, 695
uncoated, 236, 713, 714, 730, 732, 850
Siloxane polymer lattice, 870
Single
asperity, 49, 373, 414, 448, 575, 582, 630, 700, 754
Single crystal
silicon, 432, 611, 621, 653, 755
silicon cantilever, 55, 387, 388
Single-crystal Si(100), 208, 210, 831, 882
Single-crystal silicon
hardness, 613
Single molecule
packing density, 445, 665, 839
spring constant, 142, 665, 839
Single-walled carbon nanotube (SWNT)
biosensor, 24, 26, 825
Sink-in, 193, 326
Sintering, 464
Si₃N₄ tip, 42, 58, 380, 389, 657, 704, 789, 832, 878
SiO₂
beam, 266, 268, 270, 271
film surface, 652
nanobeam, 266, 268, 284, 293
Si surface
micropatterned, 728
Skewness, 584, 877, 878
Sliding
contact, 215, 217, 335, 336, 338, 347, 359, 560, 813, 843
direction, 112, 114, 330, 349, 350, 355, 396, 402, 494, 529, 763
distance, 229, 335, 401, 463, 511, 580, 586, 633, 634, 680, 762, 833
friction, 329, 358, 404
induced chemistry, 332
velocity, 331, 375, 408–411, 490, 492, 503, 504, 534, 580, 674, 684, 885
Slip
length, 531
plane, 463, 503
Slipping and sliding regime, 501, 502
Smart
system, 874
Smooth
nanobeam, 286, 289, 291, 293
sliding, 492, 493, 494, 506, 532
Snap-in, 381, 447, 448, 696
S-N (stress-life) diagram, 431
Soft
lithography (SL), 294, 644, 645, 898
stiction, 891, 894, 895
stuck micromirror, 891–893
substrate, 199, 200, 326
surface, 11, 378, 387, 414
Solid
lubricant, 12, 342, 358, 646, 888
phase Z-DOL, 696, 836
surface, 5, 139, 201, 645, 409, 461, 482, 492, 505, 511, 544, 693, 784
thin film, 342
Solid-like
behavior, 19, 511
lubricant, 691, 888
Solid-lubricant interaction, 342
Solvation force, 459, 469, 470, 473, 481
Sp³, 00
bonded carbon, 243, 597, 598, 601, 605, 620
bonding, 603, 610, 613, 644
Spacer chain, 646, 652–654, 666, 669, 683
Spherical
particle, 103, 105, 121, 476, 875
tip, 107, 128, 190, 220–223, 387, 410, 415, 536
Split-diode photodetector, 388

- SPM. *See* Scanning probe microscopy
- Spring constant
- calculation, 165, 166
 - effective, 52–54, 143, 165, 310, 528
 - lateral, 2, 54, 125, 538
 - measurement, 64
 - vertical, 44, 57
- SP-STM. *See* Spin-polarized STM
- Sputtering
- deposition, 602
 - power, 604, 609, 610
- SSD. *See* Statistically stored dislocation
- STA-biotin wear, 871, 872
- STA. *See* Streptavidin
- Static
- friction, 3, 19, 20, 494, 507, 886–889
 - friction force, 501, 528, 642, 695, 886–888
 - friction force (stiction), 886
 - indentation, 177, 178, 437, 782
 - interaction, 304, 467, 736
 - mode, 42, 43
- Static contact angle, 659, 662, 663, 670, 671, 839, 842
- Stepping motor, 524, 525
- Steric
- force, 66, 874–876
 - repulsion, 459
- Stick-slip, 19, 330, 331, 336, 338, 342, 352, 355, 391, 392, 394, 404, 410, 448, 492, 494, 498–501, 504, 506, 508, 512, 527, 534, 707
- behavior, 331, 336, 352, 393, 507
 - friction, 210, 330, 338, 512
 - movement, 391
 - phenomenon, 392, 506
- Stick-slip sliding, 499, 500, 504, 507, 508
- Stiction, 21, 23, 24, 437, 642, 643, 645, 683, 700, 712, 750, 751, 784, 788, 789, 797, 804, 806, 808, 809, 811, 825, 886, 889, 891–895
- phenomena, 889
- Stillinger–Weber potential, 304, 317, 318
- STM. *See* Scanning tunneling microscope
- Strain energy difference, 238
- Stray capacitance, 78, 79
- Streptavidin (STA)
- protein binding, 865
- Stress
- distribution, 267, 269–286, 291, 315
 - field, 223, 258, 259
 - intensity, 232, 238, 258, 259, 616
 - maximum, 197, 431, 608
 - measurement, 230, 231, 375
- Strubeck curve, 482, 495, 498
- Structural
- force, 459, 469, 471, 473, 479
 - material, 271, 275, 827
- STS. *See* Scanning tunneling spectroscopy
- Stuck comb drive, 806
- Stuck micromirror, 889, 890–893
- Submicron particle, 874, 875
- Sulfonium, 724, 725
- Superlubricity, 528, 530
- Superlubric state, 342
- Surface
- characterization, 432, 449
 - charge, 459, 467, 469, 730
 - charge density, 467
 - energy, 271, 413, 425, 439, 478, 491, 595, 643, 670, 698, 836, 889, 891
 - film, nanotribology, 643
 - free energy, 670
 - friction, 683
 - height, 377, 399, 419, 423, 428, 446, 622, 665, 666, 679, 715, 731, 757, 841, 873, 894
 - profile, 5, 255, 478, 755, 768, 880
- height map, 621, 624, 762, 764, 869
- imaging, 36, 96, 209
- interaction, 4, 5, 151, 318, 352, 374, 528, 859
- lubricated, 693, 789
- material, 266
- micromachining, 804, 808, 825, 827, 878, 895, 896
- potential, 382, 383, 425, 426, 428, 449, 467, 601, 711, 713, 714, 717, 730, 735, 742, 820
- potential map, 383, 425, 428, 715, 717, 731
- surface roughness, 11, 47, 96, 266, 284, 286, 344, 375, 377–390, 394, 395, 401, 414, 448, 471, 562, 581, 596, 661, 697, 752, 757, 758, 760–762, 765, 780, 790, 829, 871, 872, 877, 878, 895
- protection, 643, 684
- roughness, 11, 47, 266, 284, 286, 377–390, 401, 414, 448, 471, 691, 697, 757, 758, 760, 765, 790, 829, 871, 877, 878, 895, 896
- structure, 163, 166, 314, 317, 319, 479, 509
- temperature, simulation, 317
- tension, 413, 438, 465, 481, 535, 552, 572, 660, 690, 707, 712, 742, 820, 884
- topography, 43, 53, 384, 392, 400, 447, 520, 538, 698, 756, 765, 784
- unlubricated, 693, 785
- Surface Force Apparatus (SFA), 5, 12, 374
- Surfactant monolayer, 481, 484, 492, 502, 506, 507
- Suspended
- beam, 261, 279, 280, 294, 295, 808
- SWNT. *See* Single-walled carbon nanotube
- Synthetic vesicle, 874

T

- Tape
- lubricated, 736
- Tapping Mode (TM), 11, 43, 47, 48, 53, 55, 58, 98, 137, 138, 170, 266, 378–383, 386–389, 422, 423, 430, 667, 754, 790, 872, 873, 877–879
- AFM, 11, 43, 47, 389, 422, 667, 872, 877, 878
 - etched silicon probe (TESP), 55, 98, 378, 388
- TA. *See* Tilt angle
- Transmission Electron Microscopy, 317, 395, 424, 603
- Temperature, 7, 9
- critical, 540, 541

- dependence of friction, 530–536, 538, 575
flory, 474
theta, 474, 475
- Tip, 332, 333, 336–338, 345–357, 374, 375, 376–415, 425, 432, 434, 438, 439, 441–449, 460, 520, 522, 534, 538–551, 566, 575, 577, 611, 616, 620, 626–630
apex, 147, 356, 392, 528, 539
atom, 393
cantilever assembly, 54, 95
carbon nanotube, 320, 323, 327, 355–359, 380
conductive, 34, 41, 813, 820
focused ion beam, 41, 57, 219, 895
geometry, 54, 129, 130, 141, 184, 189, 190, 194, 221–223, 390, 415, 432, 610
jumping, 205, 336, 393
mount, 100, 138, 376
multiwalled carbon nanotube, 379
oscillation, 151, 526, 531, 539, 542
oscillation amplitude, 53, 526
preparation in UHV, 524
preparation method, 40
radius, 41, 54, 140, 187, 230, 255, 273, 345, 379, 389, 413–418, 610, 620, 621, 627, 630, 660, 695, 754, 756, 761
radius effect, 707
- Tip–sample, 7, 34, 43, 62, 107, 115, 138, 151, 311, 378, 385–390, 410, 415, 525, 526, 540, 542–545, 674, 693, 754, 841
adhesion, 728
distance, 415, 525, 540, 542
interaction, 138, 525, 526
interface, 7, 34, 43, 62, 115, 378, 410, 674, 841
separation, 447, 544
separation distance, 693
- Tip–sample interaction, 138, 525, 526
- Tip–surface
distance, 522, 543
interaction, 4, 151, 318, 332, 374, 528
interface, 310, 311
- TIR. *See* Total internal reflection
- TM. *See* Tapping mode
- Tomanek–Zhong–Thomas model, 392
- Tomlinson model, 409, 519, 527, 533
- Top-down method, 895
- Topographical
asymmetry, 399
image, 138–141, 430
- Topography
induced effect, 116, 399, 405
measurement, 11, 43, 54
scan, 382, 425
- Topography and recognition (TREC) imaging, 136, 152
- Torsional Resonance (TR), 126–130
amplitude, 388
mode, 389
- Torsional stiffness, 68, 100, 110, 120, 123–127, 130
- Tosylate, 724
- Total internal reflection, 167
- Trace and retrace, 399, 405, 763
- Transmission Electron Microscopy (TEM), 317, 395, 424, 603
- Traveling direction of the sample, 59, 112
- TREC imaging. *See* Topography and recognition
- Triangular cantilever, 55, 68, 75, 100, 107, 110, 127, 379, 380
- Tribochemical reaction, 359, 440, 676, 703, 883
- Tribochemistry, 340
- Triboelectrical emission, 680
- Tribometer, 355, 630, 726, 732, 828, 855
- Tribotest apparatus, 660, 690
- Triflameide, 724, 726
- TR. *See* Torsional resonance
- Tungsten
sphere, 522
tip, 41, 526, 692
- Tunneling
current, 38–44, 376
tip, 8, 38, 44
- U**
- UHV. *See* Ultrahigh vacuum
- Ultrahigh vacuum (UHV), 9, 45, 309, 334, 524, 526, 519, 530–534, 538
environment, 9, 45
- Ultrasmooth surface, 414, 683
- Ultrathin DLC coating, 228
- Unbinding
force, 142–145, 153, 164
- Unbonded Z-DOL film, 448, 699
- Uncoated Si, 227–229, 236, 628, 631, 713, 714, 727–735, 773, 782, 850, 870
- Undoped polysilicon, 262–266, 832, 833
- Undulation force, 459
- Unfilled polymer, 282
- Unloading curve, 52, 190, 192, 195, 203, 205, 208, 314–316, 381, 385, 434, 615, 698, 777
- Unlubricated
motor, 888
Si(100), 709
- V**
- van der Waals, 52, 113, 305, 331, 341, 381, 412–415, 447, 458, 460, 462, 464–467, 469, 473, 474, 486, 526, 595, 657, 658, 663, 697, 700, 777, 811, 824, 838, 844, 848, 852, 874–877
attraction, 470
attractive force, 43, 381, 459, 777, 877
force, 52, 381, 412–415, 459, 460, 464–467, 471, 662, 777, 811, 874–876
interaction, 464, 466, 470, 658, 693, 848, 875, 876
interaction energy, 462
- vdW. *See* van der Waals
- Velocity
critical, 342, 408, 506, 507
dependence of friction, 530, 533, 534, 536, 538

- rescaling, 306
 Vertical coupling, 8, 39
Vibration
 amplitude, 388, 390, 401, 404
 external, 8, 39, 140
Vickers
 hardness, 177
 indentation, 210, 233–255, 264
 indenter, 189, 192, 201, 235, 833
Viscoelastic
 mapping, 373, 377, 385, 387, 437
 property, 210, 213, 379, 388, 438, 450
Viscoelasticity, 123, 385, 438, 503
 mapping, 438
Viscosity, 17, 18, 106, 294, 345, 389, 409, 439, 442, 463, 466, 484, 495, 499, 508, 691, 695, 698, 702, 712, 737, 836, 852, 855, 883
Viscous
 damping, 403, 408
 dissipation, 16
 drag, 495, 848, 850
 force, 459, 495, 797, 804
V-shaped cantilever, 55, 57, 523
- W**
Water
 contact angle, 725, 843, 846
 films, 409, 466, 473
 vapour, 412, 414, 595, 606, 633, 643, 709, 804, 811, 883, 889
 vapour content, 632
Wear
 damage, 632
 damage mechanism, 631
 debris, 328, 342, 423, 425, 426, 449, 494, 587, 713, 716, 775
 depth, 422, 446, 629, 660, 667, 710, 762, 771, 773–775, 781, 785, 829, 832–834, 850, 869, 872
 mapping, 411
 mark, 411, 423–426, 621, 628, 771, 773, 780, 862, 863, 869
 measurement, 660, 679, 684, 741
 process, 5, 425, 630, 754, 850
- profile, 710, 711, 775
 region, 411, 423, 426, 428, 449, 713, 730, 735
 resistance, 187, 228, 229, 254, 435, 446, 450, 628, 630, 632, 667, 669, 684, 710, 725, 730, 741, 771, 774, 782, 830–833, 839, 865
 test, 421, 425, 610, 612, 630, 660, 677, 679, 714, 715, 720, 723, 730, 734, 764, 771, 774, 805, 841, 852, 869, 872
 tracks, 229, 331, 630, 710, 774
Wearless friction, 341, 458
Wear mark, AFM image, 411, 423, 424, 628, 780
Wet environment, 449
Wettability, 439, 662, 724, 732
Wire cantilever, 54
W tip, 41
- X**
Xenon, 34
XPS spectra, 636, 738, 739
X-ray
 lithography, 895, 897
 photoelectron spectroscopy (XPS), 603, 726
- Y**
Yield point, 19, 202, 205, 483, 489
Yield stress, 19, 187, 317, 499, 500
Young's modulus, 54, 66, 85, 98, 99, 163, 166, 190, 521, 846
 elasticity, 175, 177, 284, 373, 385, 432, 679, 756, 842, 846
- Z**
Z-15
 coefficient of friction, 439–443
Z-15 lubricant, 439, 695–697, 788, 834
Z-DOL, 403, 404, 416, 439–445, 447, 448, 691, 692, 694–713, 718–723, 742, 786, 788, 791, 792, 834, 836, 838, 882, 887–889
 film, 439, 448, 691, 699, 711, 712, 720, 834
 molecule, 439, 692–694, 698, 720, 836
Zisman plot, 670, 671