Nanometre scale indentation: effect of very sharp indenters on adhesion, plasticity, and electronic transport.

Mehdi El Ouali
Department of Physics
McGill University
Montréal, Québec

August 2010

A thesis submitted to McGill University
in partial fulfillment of the requirements of
the degree of Doctor of Philosophy

Acknowledgements

As I prepare to submit this thesis, numerous people come to mind that I would like to thank for supporting me, inspiring me, and helping me grow both academically and personally. I have been fortunate enough that there are far more wonderful people to thank than can fit in these pages and I hope that those who have been left out unintentionally do know how much their contributions meant to me.

First, I would like to thank Professor Peter Grütter for inviting me to join his research group and for his wonderful support throughout my time at McGill University. I am also incredibly indebted to Dr. Yoichi Miyahara for his help and for being a tremendous resource throughout the years. A warm thank you to Dr. David Oliver for providing me with great support when I needed it, and for the numerous insightful conversations we exchanged on nanoindentation. I would also like to thank Dr. Henrik Mortensen, who introduced me to the FIM system and continued to be a valuable resource after he moved on to other endeavours.

I would also like to extend my most sincere thanks to my fellow doctoral students and close collaborators, Till Hagedorn and William Paul. Till has been a relentless partner in improving the reliability of our experimental system over the years, and bringing it to a state such that it can be expected to produce novel science on a regular basis. Will has been an incredible person to work with; his impressive work ethic, dedication, and genuine talent have been truly inspiring.

The Grütter group has been a wonderful group to work with, and I would like to thank all those who over the years, have enriched my knowledge, made the lab an enjoyable workplace, and provided helping hands at multiple times with little to no notice. In no particular order: Tanya, Yan, Anne-Sophie, Shawn, Antoni, Jeff, Jessica, Sarah, Lynda, Tiger, Antoine, Manuel and everyone else. A special thank you to Professor Michel Godin and Professor Vincent Tabard-Cossa, who were then senior doctoral students in the group, for introducing me to the art of evaporating gold and for their valuable advice. Also a very special thank you to Aleks Labuda, for the numerous and passionate conversations we have had, from signal processing and noise to damping and the thermal capacity of coffee. I am also greatly appreciative of his invaluable and timely help with machining.

I would also like to express my profound gratitude to our collaborators. Dr Graham Cross has been a fantastic resource in the interpretation of our colossal mass of data. Professor Rolf Möller has provided us with invaluable help during his visit in the summer of 2008, particularly in terms of improving the reliability of the STM and

Statement of Originality

❖ Nanometre scale indentation experiments of Au(111) were conducted with unprecedented resolution and level of control, in Ultra High Vacuum, and with atomic scale characterization of the indenters before and after indentation. For the first time, a combination of simultaneous measurement of the electronic current between the tip and sample over a wide range, residual impression imaging, and very sharp atomically characterized indenters was achieved.

❖ The use of very sharp tips led to the discovery of a striking size effect, with the measurement of mean contact pressures above 25GPa, which is almost an order of magnitude higher than the largest values of hardness of Au(111) reported in the literature. A concurrent size effect was observed with respect to adhesion, such that no appreciable neck formation was observed when using a very sharp indenter.

❖ The simultaneous measurement of electronic current over a very wide range during indentation led to the discovery that cyclic loading of a nanometre scale volume leads to a progressive and dramatic loss of electronic transport properties. This surprising, and previously unreported, result has since been corroborated by preliminary transport calculations by our collaborators.

❖ An in-depth analysis of cantilever beam statics and dynamics has led to the development of design criteria for cantilever beams that have dramatically enhanced the accuracy and precision of static force measurement.

❖ A novel, comprehensive, and integrated set of software tools for data acquisition, management, and analysis was developed, allowing in particular, the seamless exploration of very large data sets, and tracking of the frequency domain evolution of simultaneously acquired signals.

Abstract

Nanometre scale indentation: effect of very sharp indenters on adhesion, plasticity, and electronic transport.

The mechanics of nanometre scale volumes were investigated experimentally by indentation of sharp W(111) tips of radii 3.6nm, 6.1nm, and 8.5nm, into clean and annealed Au(111) surfaces. Experiments were conducted in Ultra High Vacuum (UHV) and at room temperature. Tips were characterized at the atomic scale *in situ* by Field Ion Microscopy (FIM), and gold surfaces were sputtered, annealed, and characterized chemically by Auger Electron Spectroscopy (AES) and topographically by a combination of Scanning Tunnelling Microscopy/Atomic Force Microscopy (STM/AFM). High resolution load-displacement curves were obtained with simultaneous current measurement in the range 10pA-1mA. Residual impressions were imaged by STM/AFM, revealing symmetry consistent with the substrate. FIM imaging of the majority of tips after indentation showed mild gold adsorption and no damage to the underlying tip structure.

Results showed striking size effects whereby nanometre scale indentation with a very sharp indenter led to mean contact pressures above 25GPa, almost an order of magnitude higher than the largest hardness values reported in the literature for Au(111). Size effects were also observed with regards to adhesion, whereby the sharpest tip led to no observable neck formation, in contrast with blunter tips. Cyclic loading of the nanometre scale volumes led to a progressive and dramatic drop in electronic transport properties as compressive loading experiments were repeated, a result that has since been corroborated by preliminary electronic transport calculations.

Résumé

Nanoindentation à l'échelle nanométrique: effet de pénétrateurs très pointus sur l'adhésion, la plasticité et le transport électronique.

Les propriétés mécaniques de volumes à l'échelle nanométrique ont été étudiées expérimentalement par indentation de pénétrateurs pointus de W(111) de rayons de courbure de 3.6nm, 6.1nm et 8.5nm, dans des surfaces propres et recuites d'or Au(111). Les expériences ont été menées à température ambiante et sous Ultra Haut Vide (UHV). Les pointes ont été caractérisées à l'échelle atomique *in situ* par microscopie à champ ionique (FIM), et les surfaces d'or ont été nettoyées par bombardement ionique, recuites, et caractérisées chimiquement par spectroscopie Auger (AES) et topographiquement par une combinaison de microscopie à effet tunnel et de microscopie à force atomique (STM/AFM). Des courbes de charge-pénétration de haute résolution ont été obtenues, avec mesure simultanée de courants électriques dans la gamme 10pA-1mA. Les impressions résiduelles ont été caractérisées par STM/AFM, révélant une symétrie congruente à celle du substrat. La caractérisation par FIM de la majorité des pointes après indentation a révélé une faible adsorption d'or, et l'absence de dommage à la structure sous-jacente de la pointe.

Les résultats expérimentaux ont montré des effets de taille surprenants, selon lesquels l'indentation à l'échelle nanométrique avec un pénétrateur fortement pointu mène à des pressions de contact moyennes au-dessus de 25GPa, soit presque un ordre de grandeur de plus que les plus grandes valeurs de dureté reportées dans la littérature pour la surface d'or Au(111). Des effets de taille ont également été observés relativement à l'adhésion, selon lesquels le pénétrateur le plus pointu n'a pas présenté de col observable, contrairement aux autres. La charge cyclique des volumes nanométriques a mené à une suppression progressive et dramatique des propriétés de transport électronique à mesure que les charges étaient répétées, un résultat qui a depuis été corroboré par des calculs préliminaires de transport électronique.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Since Richard Feynman's inspiring lecture[1] *"There's Plenty of Room at the Bottom. An Invitation to Enter a New Field of Physics"* in 1959, immense progress has been made in the conquest of the nanoworld. Nanotechnology promises to change our lives in unprecedented ways, pervading all areas of society and tackling some of the defining issues facing mankind [1]. In particular, nanotechnology is expected to provide solutions to energy scarcity and climate change, to revolutionise food production, and to provide personalised and highly targeted treatment for ailments that are presently deemed incurable.

However, before it can live up to the tremendous expectations placed on it by society, nanotechnology will need to meet challenges of its own. One such challenge is developing a thorough understanding of the way matter interacts at the atomic and molecular scale. At that scale, electronic and mechanical properties of materials become fundamentally different from their macroscopic counterparts. Understanding them is not only extremely interesting from a scientific perspective, it is a critical prerequisite to enable the design and construction of nanoscale electronic and mechanical devices that will lead the nanotechnology revolution.

Significant progress has been made in the last two decades, largely empowered by revolutionary instruments and techniques such as the Scanning Tunnelling Microscope (STM) [2], and the Atomic Force Microscope (AFM) [3], which have dramatically altered the way in which we interact with matter at the nanometre scale.

In the present thesis, a combined STM/AFM is used to further explore these interactions, and in particular, to investigate the mechanics and electronics of nanometre scale volumes under high compressive loads. Experimental investigations of such small volumes in a highly controlled environment has been elusive in the field of

---

[1]Transcript available at www.zyvex.com/nanotech/feynman.html, retrieved in August 2010

contact mechanics, and the insights that they elicit are essential in developing an understanding of the fundamental phenomena that are responsible for the wear and yielding behaviour of materials. They allow the exploration of important questions, such as the limits of mechanical fabrication and the level of applicability of contact mechanics models at the near-atomic scale.

It has long been known that macroscopic plasticity of materials is determined by the presence and mobility of defects, particularly dislocations [4]. The probing of nanometre volumes allows the investigation of a very limited number of dislocations, which allows one to build a fundamental understanding of their behaviour and eventually to predict the macroscopic response of materials. In addition, the experimental study of volumes involving no more than a few thousand atoms offers the possibility of direct comparison with computer modelling, which can substantially enhance our understanding of incipient plasticity, defect formation, and electronic and mass transport at the nanoscale.

The experiments described in this thesis involve the mechanical probing of clean Au(111) surfaces with very sharp, atomically characterized, W(111) indenters, with tip radii between 3nm and 9nm. The surfaces are prepared and investigated in Ultra High Vacuum (UHV), with a high load and displacement resolution (on the order of 1nN and 10pm respectively). Indentation of these surfaces at depths of 1-4nm allows the probing of very small volumes, with simultaneous force and current measurement. This leads to the observation of dramatic size effects, with respect to adhesion and plasticity. In particular, we will see that indentation with a very sharp tip leads to the measurement of contact pressures that are far higher than those reported by previous studies of the Au(111) surface. Another striking result reported here is the spectacular drop in the ability of these volumes to conduct electronic currents when they are subjected to cyclic mechanical loading.

In chapter 2, a detailed review of the instrument will be given, primarily from the point of view of force measurement. The principle of our deflection measurement system will be detailed, and noise and vibration considerations will be explored.

Chapter 3 will provide an opportunity to review the details of sample preparation. Surface preparation and characterization techniques will be presented. In particular, a new simple software analysis tool will be introduced to describe the evolution of the power spectral density of a signal.

In chapter 4, a phenomenological overview of the results will be provided, along with a brief description of a comprehensive analysis and data management system developed in the Matlab^{TM} environment. Results of single and cyclic loading exper-

iments will be presented, with topographic scanning probe imaging of the residual impression left after indentation. The effect of cyclic loading on the contact resistance will also be introduced, along with associated preliminary results from electronic transport calculations by the group of Professor Guo at McGill University.

Chapter 5 will introduce contact mechanics models. Standard analysis techniques of nanoindentation data will be outlined and applied, leading to the measurement of very high mean contact pressures.

Finally, chapter 6 will describe observed tip size effects, which show dramatically different behaviour for very sharp tips (below 5nm radius).

A brief summary, along with suggestions of future work will then be provided in chapter 7. It will be followed by appendices outlining important concepts of Digital Signal Processing (DSP) referred to throughout the thesis, and describing useful software enhancements.

Chapter 2

Instrumentation

The instrument used for these investigations enable experiments that are unique. It allows the simultaneous measurement of forces and currents, from the onset of tunnelling to deep indentation, with atomically defined tips. These measurements are performed in a Scanning Probe Microscopy (SPM) configuration at temperatures between liquid nitrogen and room temperature, and in Ultra High Vacuum (UHV).

The system, as it was initially designed, has been documented by Alain Stalder in his PhD thesis [5]. Since this initial description of the system, several enhancements have been implemented, such as the commissioning of a new preparation chamber [6], the implementation of a new coarse approach system [7], and others which will be discussed here.

After a brief overview of the system and its two modes of operation, the focus of this chapter will be on important aspects related to force measurement in the experimental setup. In particular, the various design criteria for the cantilever beams, which fulfill the function of force transducers, will be reviewed.

2.1 Experimental System

2.1.1 Instrument Overview

The experimental system is based on a combination of several surface analysis techniques. However, before describing more specific aspects of the experimental system, I will provide a brief overview of the system as a whole.

Scanning Probe Microscopy

At its core, the system is a Scanning Probe Microscope. Scanning Probe Microscopy (SPM) is based on the measurement of a very localised interaction between a probe and a sample surface. The probe is raster-scanned across an area of interest, which is typically in the range of a few nanometres to several microns. As the tip is scanned, the interaction is recorded in a two-dimensional matrix which produces a map of the interaction. In the majority of cases however, the interaction is kept constant by varying the tip–sample separation in a controlled manner, and the vertical displacement is recorded, providing a topographical map of a surface in which the interaction is equal to a specific reference value, which is referred to as the *setpoint*. Several instructive references are available on the subject, such as the Springer Handbook of Nanotechnology [8], with a section dedicated to SPM, as well as books by Meyer, Hug and Bennewitz [9], Foster and Hofer [10], and by Sarid [11]. Finally, a comprehensive review of SPM was written by Hofer [12].

Scanning Tunnelling Microscopy (STM) was invented by Binnig *et al.* [2] in 1982. The interaction of interest in STM is the electronic tunnelling current between a sharp metallic tip and a conducting sample to which a bias voltage has been applied. A great introductory reference to the topic is a book by Chen [13].

Four years later, Atomic Force Microscopy (AFM) was invented by Binnig *et al.* [3]. The interactions that are measured by AFM are the forces between the probe and the sample, which can be both repulsive or attractive and are varied in nature. A number of useful references exist on the topic, including another book by Sarid [14], and one edited by Morita, Wiesendanger and Meyer [15], both of which are of great practical and theoretical interest.

FIM/STM/AFM system

The experimental system is divided into three vacuum chambers, each of which fulfill different roles.

Airlock Chamber: The airlock is a small chamber held at high vacuum ($\sim 1 \cdot 10^{-7}$mbar), which is used to load samples and tips without venting the other chambers. Once a sample/tip has been loaded and a high vacuum is attained, the sample is moved to the preparation chamber.

Preparation Chamber: Before a tip/sample is moved to the measurement chamber, it is subjected to a number of additional preparation steps which are

completed in the preparation chamber. This chamber is kept at a very high vacuum ($\sim 1 \cdot 10^{-10} - 1 \cdot 10^{-9}$mbar) by a combination of pumps (mechanical, turbomolecular, ion getter, and titanium sublimation). It is equipped in particular with a heatable sample preparation stage, an ion bombardment gun, an e-beam evaporator, a residual gas analyser, and an Auger Electron Spectrometer (AES). This combination of tools allows us to prepare and chemically characterize samples that are freshly cleaned and annealed before they are transferred to the measurement chamber.

Measurement Chamber: The measurement chamber is the *cleanest* of the three chambers, maintaining an Ultra High Vacuum (UHV) on the order of $10^{-10} - 10^{-11}$mbar. In addition to hosting the core measurement instruments (FIM, AFM and STM), it features a cryostat for liquid nitrogen temperatures, a magnetically levitated turbomolecular pump, a cryogenic pump, and a Titanium Sublimation Pump (TSP). It also includes a magnetic damping system and a damping bellow that are used to isolate the SPM system from external sources of vibrations.

The FIM/STM/AFM system has two main modes of operation, as illustrated by figure 2.1. One, described briefly in section 2.1.3, is a combined STM/AFM mode, which allows for simultaneous measurement of the forces and currents between the sample and the probe. The second mode, described in section 2.1.2, is used to determine the atomic structure of the apex of the probe, which is a tungsten tip.

The typical experimental sequence will first involve sample and tip preparation, followed by tip apex imaging by FIM, and STM imaging of the sample surface. During STM imaging, a simultaneous recording of the measurable forces between tip and sample is acquired. After locating a region of interest on the sample, force–distance curves (or current–voltage spectroscopy curves) are obtained, before a new scan of the surface is performed. After the experiments are complete, the tip is generally imaged a second time by FIM to investigate any possible changes.

2.1.2 Field Ion Microscopy

One original feature of our instrument, as compared to traditional Scanning Probe Microscopes, is that it allows for full atomic characterization of the probe. Our probes are electrochemically etched tungsten tips prepared from polycrystalline or single crystal tungsten wire. Our tip preparation procedures have been documented
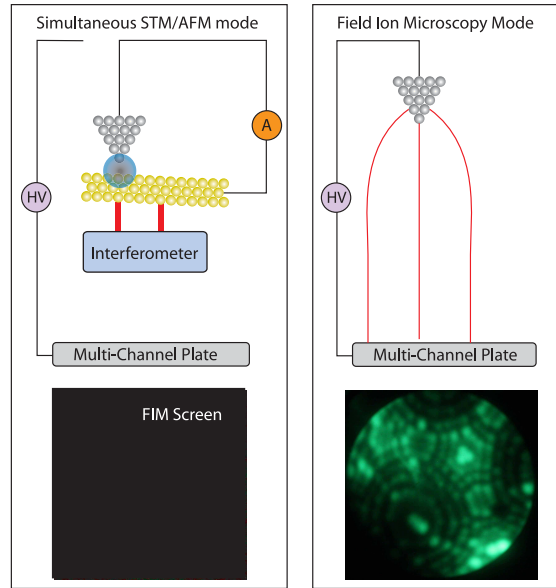
Figure 2.1: In STM/AFM mode, after a sharp tungsten tip ($3nm < R < 10nm$) is brought to tunnelling contact with a Au(111) surface, current and forces are simultaneously recorded. Both surface scans and approach curves are obtained in this mode. In FIM mode, the sample is removed and the tip is imaged at an atomic level by FIM. A 3-D reconstruction and the tip radius are later extracted from the FIM images [16, 17]. In the example above, the tip is a single crystal tungsten tip with a trimer at the apex.

extensively by Lucier [16, 17] and by Hagedorn [7]. More details on tip preparation and characterization relevant to the present studies are described in section 3.4.

The tips obtained through our procedures are very sharp, with a radius of curvature of a few nanometres at the apex. These tips are sufficiently sharp to allow characterization by FIM. FIM principles and experimental details are also described in the previous references. Great introductions to FIM and related techniques are provided by Tsong [18] and by Miller *et al.* [19].

In short, an imaging gas (helium) is introduced into the measurement chamber, and a high positive voltage is applied to the tip. The local sharpness of the tip then enhances the electric field locally. The field gradients polarise and attract imaging gas atoms towards the tip, and the high fields eventually ionise them. The positive ions are then strongly repelled by the electric field lines towards a multichannel plate, which allows the imaging of the incoming beams on a phosphor screen. An image similar to that of figure 2.1 is then obtained. The image is a projection of the tip apex, and each bright dot corresponds to a single atom. The analysis of such images allows the extraction of important information, such as crystal orientation and local radii of curvature, as well as the production of 3-D reconstructions of the tip [16].

2.1.3   STM/AFM Mode

Principle

In STM/AFM mode, the probe is attached to a piezoelectric scanner tube, which allows sub-angstrom lateral and vertical motion of the tip. A small voltage bias (typically on the order of 30–100mV) is applied to the sample, which is a flexible cantilever beam that consists of a gold covered mica strip, clamped at one end. The cantilever beam is flexible which allows it to serve as a force transducer. This arrangement is very unusual in AFM, where the sample is typically rigid and the tip is attached to a flexible cantilever beam. Our unique configuration provides us with a great level of flexibility with regards to our probes, such as local etching and annealing.

The tip is approached carefully towards the sample using a coarse approach system based on a piezo-walker that relies on the stick-slip motion of shear piezoelectric actuators. When the tip–sample distance is within the vertical range of the piezoelectric tube (~ 800nm), the piezotube extends to form a tunnelling junction.

From that point on, the tip can be raster-scanned across the sample to obtain images of a typical lateral range of 20–150nm. While the tip is scanned, its vertical position is controlled by a digital feedback control system that aims to stabilise the tunnelling current at a given setpoint, which is typically 100pA in our experiments. In addition to the current, several channels are acquired simultaneously. In particular, the deflection of the cantilever is acquired to provide a measure of the interaction force between the tip and sample.

One important point to highlight is that the experiments described in this thesis are carried out in static mode. This means that no excitation is applied by the user, and that the interactions are measured directly. While this is typical in STM, it is in contrast to a wide range of AFM studies in the literature that are based on dynamic techniques. These maintain the cantilever beam in oscillation and detect changes in frequency, amplitude, and phase [8, 11, 15, 20, 21]. A piezoelectric shaker mounted underneath the sample stage does afford us the capability, in principle, to operate in dynamic mode. However, for the present investigation, the shaker was used only for testing and the reported experiments were carried out in static mode, which has the distinct advantage of providing a direct measurement of force.

Tunnelling Current

In STM, the tunnelling current is a complex function of a number of parameters including, most importantly, the local density of states at the Fermi energy, the tip–sample separation, and the sample bias voltage [9, 13]. Equation (2.1) highlights the tunnelling current's dependence on each of these parameters.

$$I_t \ \propto \ V \ \rho_s \ e^{-1.025 \ z[\text{Å}]\sqrt{\phi[\text{eV}]}} \quad (2.1)$$

In equation (2.1), V is the bias voltage, ρ_s is the local density of states of the sample at the Fermi level, z is the tip–sample separation (in angstroms) and ϕ is the apparent barrier height (in eV). Typical values for ϕ are close to 4–5eVfor metals, which means that an increase in the tip–sample separation of one angstrom will result in a drop in the tunnelling current of almost an order of magnitude. This extraordinarily sensitive exponential dependence of the tunnelling current on the tip–sample separation is what makes it a great channel to use for feedback and what affords STM its very high resolution.

Prior to being recorded, the tunnelling current is converted into a voltage by a current–voltage (I-V) converter. I-V converter designs are commonly based on operational amplifiers and are almost exclusively linear, with a conversion factor usually on the order of $0.1V/$nA. Most acquisition systems are limited to a range of 1mV–10V, which limits the range that currents can usually be measured to 10pA–100nA at most. While this is adequate for most STM configurations, it is insufficient if one requires currents to be measured from the onset of tunnelling (a few pAs) through to a nanoindentation experiment, when the conductance associated with a contact area of a few hundred atoms can reach several hundred conductance quanta ($G_0 \simeq 77.5\mu$S), which corresponds to a few milliamperes for a bias voltage of 50mV. For that purpose, we use a logarithmic (logI-V) converter, based on the design of Dürig *et al.* [22].

Figure 2.2 shows the calibration function of our logI-V converter, which converts currents in the 10pA–5mA range to voltages between 1mV–10V, and can therefore span nine orders of magnitude, with a measurement bandwidth on the order of 1kHz.

Although we refer to our converter as logarithmic, the actual calibration function is complex, and logarithmic by interval only. In particular, for low values of current (below 2.5nA), the behaviour is linear, with a conversion factor of approximately $1.1V/$nA. While this makes the choice of setpoint easier, the linear behaviour for currents in the range of STM imaging ($10\text{pA} - 2\text{nA}$), will cause the feedback error signal (the tunnelling current) to vary exponentially with tip–sample separation. To

linearise this relationship, and therefore enhance the stability of STM imaging, an additional electronic component which performs a log operation is added between the logI-V converter and the acquisition system.

STM imaging with the FIM/STM/AFM system allows us to ensure that the sample location of interest is flat and free of large surface defects. For the present investigations carried out on the Au(111) surface, the resolution and stability allows us to routinely resolve the herringbone reconstruction.
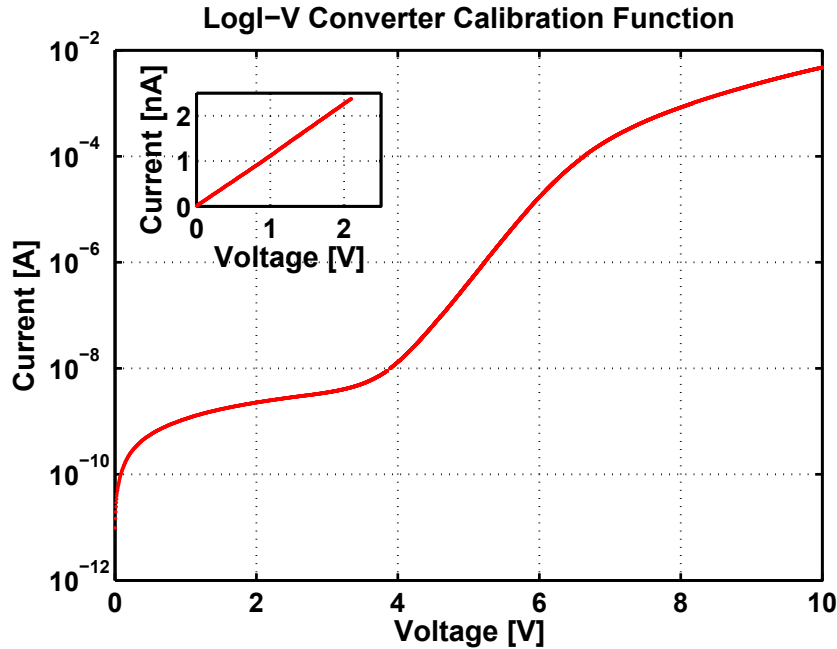
Figure 2.2: Calibration function for our logI-V converter (real data). As shown in the inset, the current–voltage relation starts off linear for low current values, below 2.5nA. For higher currents, the relationship is logarithmic ($V \propto \log I$) for currents up to 10μA, and become quasi logarithmic with a different slope beyond that. This allows the measurement of currents over nine order of magnitudes, enabling the continuous measurement of currents between the onset of tunnelling and deep nanoindentation.

Forces

The forces involved in STM and AFM have a wide variety of ranges, strengths, and physical origins. Numerous authors have categorised these forces in detail, such as Burnham *et al.* [23], Meyer*et al.* [9] (in particular section 3.2), and Foster*et al.* [10]. These have also been discussed in a number of reviews [8, 12, 24, 25]. The most important forces are briefly reviewed below, restricting the discussion to issues relevant to clean surfaces in Ultra High Vacuum.

Van der Waals: Van der Waals forces are long range in nature, involving the mesoscopic structure of tips and samples. Several types of forces are referred to as van der Waals forces, and they are generally related to dipole–dipole interactions, dipole fluctuations, and induced dipoles. For a sphere–plane geometry, they can be written $F_{\mathrm{vdW}} = HR/6z^2$ [9, 23], where the Hamaker constant $H \simeq 10^{-19} J$, R is the tip radius, and z the distance between the sphere and the plane. For our sharp tips with radii $R \simeq 5$nm, this implies a contribution below 1nN at a distance of $5\mathring{A}$. Solutions for a truncated cone geometry are given by Saint Jean *et al.* [26].

Electrostatic: A conductive tip and sample at different potentials can be seen as a capacitor of capacitance C, and the related electrostatic forces are given by $F_{\mathrm{elect}} = \partial C/\partial z \ (U_{\mathrm{bias}} - U_{\mathrm{CPD}})^2$, where U_{CPD} is the contact potential difference and is due to the differences in work function between tip and sample. Electrostatic forces are also below 1nN at small distance [9].

Short Range: Short range forces are the most relevant forces for our studies. They arise at distances below 1nm, and are due to an overlap of electronic wave function and repulsion of the ionic cores. They have both attractive and repulsive components, and can be modelled by Morse or Lennard-Jones potentials [8, 9].

A number of other forces are listed in the literature, such as capillary and magnetic forces. These were omitted from this review because their relevance to our atomically clean Ultra High Vacuum (UHV) conditions is extremely limited.

Till Hagedorn [7] has studied the tip-sample configuration from tunnelling to contact formation. In the present thesis, we are interested in investigating the mechanics after a complete elastic and plastic contact has formed, in which case short range forces are most relevant and are used for computer simulations of contact mechanics and nanoindentation. A number of contact models that describe the forces involved are described in section 5.2.

Acquisition System

The data presented here was acquired using the Scanita [27] acquisition and control software and related hardware. The data is digitized and acquired through six channels using 16-bit Analog-Digital-Converters at 200kHz. The data is then averaged and downsampled twice, once at the hardware level and once at the software level.

The hardware oversampling is typically set to twenty, which means that the data processed by the digital controller arrives at a rate of 10kHz, which is referred to as the *control frequency*. During scanning, the software oversampling is usually set to twenty, which results in a *data frequency* of 500Hz. For our common scan size of 256x256, and considering that both the forward and a backward traces are recorded, this represents a line frequency just below 1Hz. Therefore, a complete image takes slightly over four minutes to acquire.

Our acquisition system allows us to acquire curves at a given location on the sample, typically varying a quantity to measure the effect on another. This allows the investigation of force–distance and current–voltage relationships. The corresponding curves, referred to as *spectroscopy* curves, are acquired by the software in a way that precludes explicit user control over (and knowledge of) the software oversampling rate, and information on the time spacing of data points can be lost. To circumvent this limitation, the software can be forced to acquire a number of points high enough to implicitly require that no software oversampling be performed at all. All the curves acquired with Scanita that are presented in this thesis are taken in that configuration, with a hardware oversampling of 20. All our spectroscopy curves are therefore acquired with a sampling frequency of 10kHz.

In most experiments, the channels that are acquired are as follows:

Current: As discussed previously, the current channel records a voltage that is a quasi-logarithmic function of the actual tunnelling current, as described by figure 2.2. The bandwidth of the logI-V converter being below 2kHz, the recorded signal has very little spectral content at high frequencies, and aliasing is therefore limited when acquired at 10kHz.

Log-Current: A logarithmic version of the previous channel, which is used solely for feedback control.

Deflection: A signal that is proportional to the deflection of the cantilever beam, and therefore to the force between tip and sample. For anti-aliasing purposes, this channel is usually low-pass filtered by hardware at 2.5kHz when acquired with Scanita (at 10kHz), and 10kHz or higher when acquired with Labview$^{\text{TM}}$ or Matlab$^{\text{TM}}$ (usually at 50kHz or higher). Deflection measurement is discussed at length in section 2.3.

Amplified Deflection: An amplified version of the deflection channel, using the

LUPE components[1] of our controller. In recent experiments however, this channel has been abandoned as the deflection signal strength has been greatly enhanced by the implementation of a new preamplifier box mounted on the chamber.

Before exploring the details of our deflection measurement system, the basics of cantilever beam statics and dynamics, as well as their effect in terms of choice of design parameters will be reviewed.

2.2 Cantilever Beam Design

In order to be suitable force transducers for investigations, the cantilever beams need to fulfill a number of requirements, which are described in section 2.5.2. In short, their spring constant should be as low as possible to allow maximum force resolution, while remaining above the maximum force gradient between the tip and sample, to avoid jump-to-contact[2]. The resonance frequency should be as high as possible and remain far from all system frequencies (both resonance frequencies and external excitations), as described in section 2.4.2. However, a large resonance frequency will generally imply a high spring constant, and a compromise is therefore necessary. Typically, the first resonance frequency of our cantilevers is between 500Hz–3kHz, and the spring constant, as measured at the end of the cantilever, is between 2–35N/m.

2.2.1 Cantilever Statics

Basics of Cantilever Statics

In order to describe the static bending of a cantilever beam, we use the Euler-Bernoulli beam theory, which is based on the following assumptions commonly described in engineering textbooks [28]:

◇ The beam is symmetric, and both supported and loaded symmetrically, with respect to the plane of bending.

---

[1]LUPE (magnifying glass in German) components are amplifier cards attached to our acquisition system which allow offset subtraction and amplification of a signal by factors 1, 2, 4, 8, 16, 64, or 128.

[2]With our tips, this requirement is generally met with an operating spring constant higher than 50N/m for a bias voltage between 10mV–100mV.

◇ There exists a longitudinal plane, perpendicular to the bending plane, on which the strain remains zero when the beam deforms. This plane is called the neutral surface and its intersection with the bending place is the neutral axis (dotted line in figure 2.3 and in figure 2.4).

◇ Plane cross sections which are perpendicular to the axis of the undeformed beam remain plane, and are perpendicular to the deflection curve of the beam.

◇ Deformation in the plane of a cross section (x–y directions in figure 2.3) is negligible.
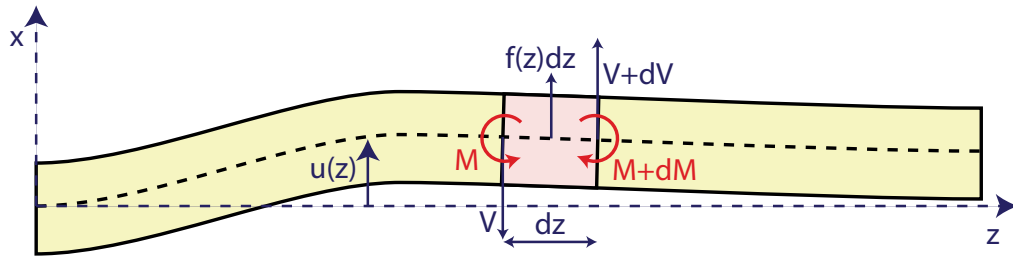
Figure 2.3: Static equilibrium of the cantilever beam

As shown in figure 2.3, the deformation of the cantilever beam along the z axis is given by the deflection curve $u(z)$. In order to derive expressions for the bending of the beam, let us assume that it is subjected to a load distribution defined by a force density $f(z)$. The forces applied to a small volume of thickness dz are therefore the external load $f(z)\ dz$ and the internal shear force $V(z)$, produced by the deformation of the beam, and exerted on each of the two lateral cross sections (V and $V + dV$ respectively). The weight of the beam is neglected, and can be incorporated in $f(z)$. The bending also causes a bending moment $M(z)$.

The static equilibrium of a small volume element of thickness dz with regards to translation can be written:

$$(V + dV) + f(z)\ dz - V \quad = \quad 0 \quad (2.2)$$

This leads to the following important result:

$$\boxed{\frac{\mathrm{d}V}{\mathrm{d}z} \quad = \quad -f(z) \quad (2.3)}$$

Writing the equilibrium with regards to rotation by equating the sum of moments at the point on the deflection curve at $z + dz$ with zero, we obtain:

$$M - (M + dM) + V\ dz + O(dz^2) \quad = \quad 0 \quad (2.4)$$

This can be written in the following form:

$$\frac{\mathrm{d}M}{\mathrm{d}z} = V(z) \quad (2.5)$$
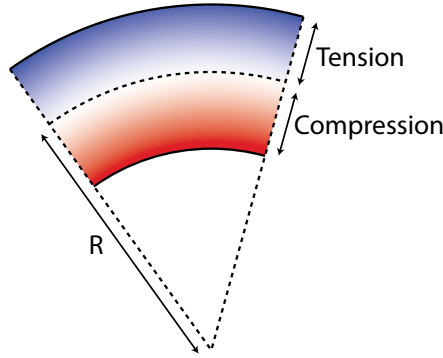
Figure 2.4: As the beam bends, the neutral axis remains strain free. The consequence is that one side of the beam slab is subjected to compressive strain while the other is subjected to tensile strain.

Bending of the beam causes compressive strain on one portion of the dz slab and tensile strain on the other (Fig. 2.4). Assuming that the bending can be described locally by a radius of curvature R, writing the strain distribution and computing the associated stress using Hooke's Law, one can calculate an expression for the bending moment [14], as shown in equation (2.6).

$$\frac{1}{R} = \left| \frac{M}{EI} \right| \quad (2.6)$$

In equation (2.6), E is Young's modulus, and I is the moment of inertia of the cross section, as defined by equation (2.7), where A is the cross section area and x is the distance from the neutral axis. Equation (2.8) gives the moment of inertia for a rectangular bar of width w and thickness τ.

$$I = \int_A x^2 \, dA \quad (2.7)$$

$$I_{\text{rect}} = \frac{w\tau^3}{12} \quad (2.8)$$

Noting that the curvature of a plane curve can be expressed by equation (2.9), and choosing the right sign to make $u(z)$ locally concave for a positive moment, we obtain expression (2.10).

$$\frac{1}{R} \; = \; \frac{\left|\dfrac{\mathrm{d}^2 u}{\mathrm{d}z^2}\right|}{\left(1 + \left(\dfrac{\mathrm{d}u}{\mathrm{d}z}\right)^2\right)^{3/2}} \simeq \left|\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}\right| \quad (2.9)$$

$$\boxed{M \; = \; EI\frac{\mathrm{d}^2 u}{\mathrm{d}z^2} \quad (2.10)}$$

Combining equations (2.3), (2.5), and (2.10) one obtains the Euler-Bernoulli beam equation (2.11).

$$\boxed{\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left(EI\frac{\mathrm{d}^2 u}{\mathrm{d}z^2}\right) = -f(z) \qquad (2.11)}$$

Static Deflection Model

Using the Euler-Bernoulli beam equation, one can derive expressions for the static deflection of a cantilever beam with a rectangular cross section, of length L, width w, and thickness τ. Assuming single point loading by a tip at position $z_{\mathrm{tip}} = t$, and that the load points downwards with an amplitude F, one can write the force density $f(z) = -F\,\delta(z - t)$. The Euler-Bernoulli equation can thus be written:

$$u^{(4)}(z) = \frac{F}{EI}\delta(z - t) \qquad (2.12)$$

Case 1 ($z > t$): In this case, $u^{(3)}(z) = 0$. After integrating three times, and using the boundary condition $M(L) = 0$ (which implies that $M(z) = 0$ for all $z > t$), and writing $u^{(1)}(t) = u'_t$ and $u(t) = u_t$, we obtain the following expression:

$$u(z > t) = u'_t\,(z - t) + u_t \qquad (2.13)$$

Case 2 ($z < t$): In this case, $u^{(3)}(z) = F/EI$, with the following boundary conditions: $u^{(2)}(t) = 0$, $u^{(1)}(0) = 0$, and $u(0) = 0$. The first is due to the fact that $M(z \to t) = 0$, as per case 1. The last two express the fact that the beam is fully anchored on one side. We thus obtain the following expression:

$$u(z < t) = -\frac{F}{6EI}z^2(3t - z) \qquad (2.14)$$

Rearranging these equations, and using equation (2.8), we obtain very convenient expressions for the deflection, as summarised in equations (2.15), (2.16), and (2.17).

$$u_t = -\frac{4t^3 F}{Ew\tau^3} \tag{2.15}$$

$$u(z \leq t) = u_t \left[\frac{3}{2}\frac{z^2}{t^2} - \frac{1}{2}\frac{z^3}{t^3} \right] \tag{2.16}$$

$$u(z \geq t) = u_t \left[\frac{3}{2}\frac{z}{t} - \frac{1}{2} \right] \tag{2.17}$$

The resulting deflection curve and the bending moment along the cantilever are illustrated in figure 2.5. As outlined in section 2.3.2, equations (2.16) and (2.17) are key results used to convert the electronic signal obtained from the interferometer into a calibrated deflection measurement.
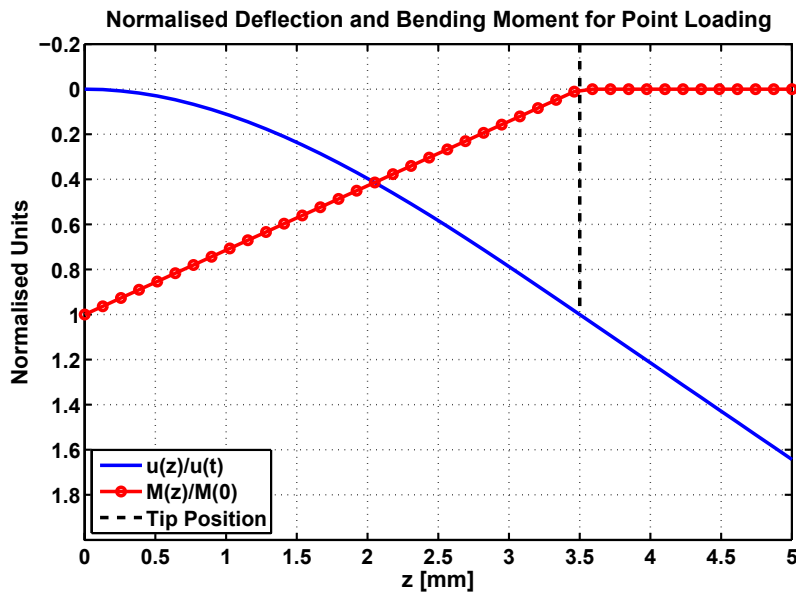
Figure 2.5: Illustration of the normalised deflection and bending moment distribution for a point load applied by a tip at position t=3.5mm.

Cantilever Beam Spring Constant

Remembering that u_t is the deflection of the cantilever at the tip position, equation (2.15) shows that the deflection at the tip is proportional to the force. By writing $F = -k_t u_t$, we obtain an expression for k_t, the cantilever beam spring constant at the point of loading (equation (2.18)).

$$k_t = \frac{Ew\tau^3}{4t^3} \quad (2.18)$$

When taking $t = L$, we retrieve the well-known expression for a cantilever beam spring constant at the free end, k_0 (equation (2.19)).

$$k_0 = \frac{Ew\tau^3}{4L^3} \quad (2.19)$$

Combining equations (2.18) and (2.19), we obtain a scaling relation between k_t, k_0, t, and L(equation (2.20)).

$$k_t = k_0 \frac{L^3}{t^3} \quad (2.20)$$

Equation (2.20) is of great practical importance, as the spring constant of the cantilever at the tip, which is necessary to convert the deflection into force, is computed from k_0 (obtained experimentally prior to putting the sample in UHV) and optical photography measurements of the distances L and t.

2.2.2   Cantilever Dynamics

The analysis performed in the previous section for the static deflection can be extended using Newton's second law to obtain an equation of motion for the beam [13, 14, 29, 30], as shown in equation (2.21) where ρ_l is the density of the beam per unit length.

$$\rho_l \frac{\partial^2}{\partial t^2} u(z,t) + EI \frac{\partial^4}{\partial z^4} u(z,t) + f(z,t) = 0 \qquad (2.21)$$

Using standard variable separation and decomposition methods, one can derive solutions for the beam dynamics. Writing $u(z,t) = v(z)Y(t)$, and $f(z,t) = f(z)Y(t)$, one can obtain two related differential equations: with respect to time (equation (2.22)), and position (equation (2.23)).

$$\ddot{Y} + \omega^2 Y = 0 \quad (2.22)$$
$$\frac{\mathrm{d}^4 v(z)}{\mathrm{d}z^4} - k^4 v(z) - f(z) = 0 \quad (2.23)$$
$$k^4 EI - \omega^2 \rho_l = 0 \quad (2.24)$$

Free Vibrations

In the following, let us assume that $f(z) = 0$ for all z. Equation (2.22) simply states that the time evolution of the system is that of a simple harmonic oscillator. The spatial modes are determined by equation (2.23), which solutions span all linear combinations of $\sin(kz)$, $\cos(kz)$, $\sinh(kz)$, and $\cosh(kz)$. To find the appropriate solutions for our problem, we refer to the following boundary conditions:

$$v(0) = 0 \quad (2.25)$$
$$v'(0) = 0 \quad (2.26)$$
$$v^{(2)}(L) = 0 \quad (2.27)$$
$$v^{(3)}(L) = 0 \quad (2.28)$$

Equations (2.25) and (2.26) result from the beam being anchored at the $z = 0$ extremity, while equations (2.27) and (2.28) respectively indicate no bending moment or shear force at the free end.

The free vibration modes are thus described by equations (2.29) to (2.32), where A_n is a normalisation factor determined by equation (2.32). Table 2.1 lists approximate values for A_n and k_n different values of n. Figure 2.6 illustrates the shapes of the first three modes.

$$v_n^0(z) = A_n \left[(\sin(k_n z) - \sinh(k_n z)) + B_n (\cosh(k_n z) - \cos(k_n z)) \right] \quad (2.29)$$

$$B_n = \frac{\sin(k_n L) + \sinh(k_n L)}{\cos(k_n L) + \cosh(k_n L)} \quad (2.30)$$

$$0 = 1 + \cos(k_n L) \cosh(k_n L) \quad (2.31)$$

$$\delta nm = \int_0^L v_n^0(z) v_m^0(z) dz \quad (2.32)$$

n	1	2	3	4	5	$n \geq 6$
$k_n L$	1.875	4.694	7.855	10.996	14.137	$\simeq (2n-1)\pi/2$
$A_n\sqrt{L}$	.7341	1.0185	.9992	$\simeq 1$	$\simeq 1$	$\simeq 1$
ω_n/ω_1	1	6.267	17.55	34.39	56.85	$(k_n L/k_1 L)^2$

Table 2.1: Parameters for cantilever beam free vibration modes. Adapted from [14, 29]

The resonance frequency for each mode is obtained from equations (2.8) and (2.24), where $\rho_l = \rho w \tau$, with ρ being the mass per unit volume. The result is given in equation (2.33), where the first factor is a constant that depends only on the mode ($k_n L$ is given by table 2.1), the second ($\sqrt{E/\rho}$) depends only on the material
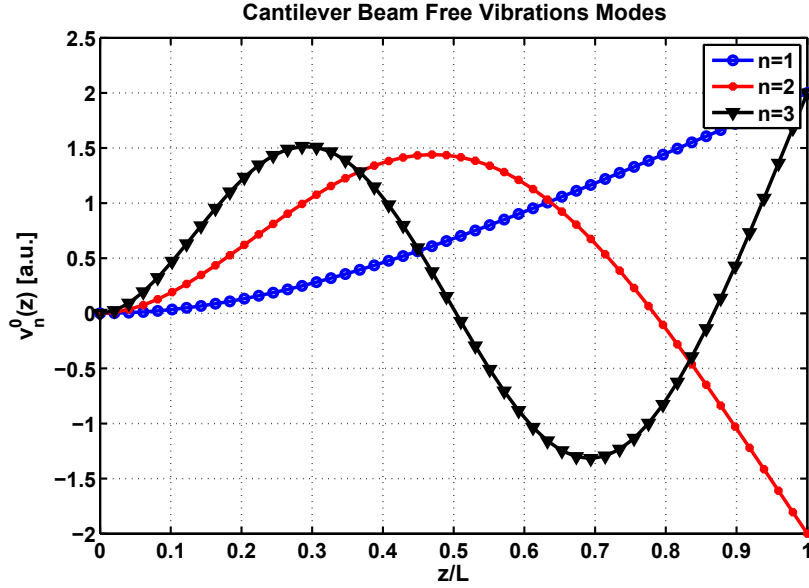
Figure 2.6: Cantilever beam vibration modes.

properties and is close to the speed of sound[3], and the third is strictly a function of the beam dimensions. It is interesting to note that the resonance frequencies are independent of the cantilever width w. Finally, as shown in table 2.1, the spacing between the frequencies for consecutive modes is not equal.

$$f_n = \frac{\omega_n}{2\pi} = \frac{(k_n L)^2}{4\pi\sqrt{3}} \sqrt{\frac{E}{\rho}} \frac{\tau}{L^2} \propto \frac{\tau}{L^2} \qquad (2.33)$$

To obtain a rule of thumb for the frequency of the first mode of our mica cantilevers, we have estimated Young's modulus based on static deflection calibration and found $E \simeq 110$GPa, which is in the range of values reported in the literature. Using the literature value $\rho = 2.7$, we thus obtain equation (2.34).

$$f_1^{\mathrm{mica}} \simeq 1030\mathrm{Hz} \frac{\tau_{\mu m}}{L_{mm}^2} \qquad (2.34)$$

For a typical cantilever of thickness $\tau = 40\mu m$ and length $L = 5mm$, equation (2.34) yields immediately $f_1 \simeq 1650$Hz, which is consistent with observed results. Using table 2.1, the second mode's frequency is thus found to be $f_2 \simeq 6.27 f_1$, etc.

---

[3]The speed of sound in a material is $\sqrt{E/(3(1-2\nu)\rho)}$, where ν is the Poisson ratio.

Cantilever Excitation

The solutions derived above for free vibrations can be used as a basis for solutions to the problem of a cantilever beam subject to an external excitation. When $f(z) \neq 0$, it is easy to see from the previous equations that one can write the following equations with the notation introduced earlier.

$$v(z) = \sum_{n=1}^{\infty} V_n \, v_n^0(z) \qquad (2.35)$$

$$V_n = \frac{c_n}{k_n^4 - k^4} \qquad (2.36)$$

$$c_n = -\int_0^L \frac{f(z)}{EI} \, dz \qquad (2.37)$$

In the case of a damped cantilever, Rast *et al.* [29] introduces a complex elastic modulus, $\hat{E} = E(1 + \imath/Q)$, which generally depends on temperature and frequency, and concludes the following:

$$\hat{V}_n^{Damped} = \frac{c_n EI}{\rho_l \left(\omega_n^2 (1 - \frac{\imath}{Q}) - \omega^2 \right)} \qquad (2.38)$$

In the case of a point force $P = P_0 \delta(z - L)$ applied at the end, Rast *et al.* obtain the following amplitude and phase at the end of the cantilever, which shows that the result is equivalent to a harmonic oscillator, with a mode dependent spring constant D_n, defined by equation (2.41). It is easy to see that $D_1 \simeq 1.03 D_0$, where $D_0 = k_0$ is the static spring constant defined in equation (2.19).

$$|\hat{v}_n(L)| = \frac{P_0 \omega_n^2}{D_n \sqrt{(\omega_n^2 - \omega^2)^2 + \frac{\omega_n^2 \omega^2}{Q^2}}} \qquad (2.39)$$

$$\phi_n = -\arctan \left(\frac{\omega \omega_n}{Q(\omega_n^2 - \omega^2)} \right) \qquad (2.40)$$

$$D_n = \omega_n^2 \rho_l L = \frac{(k_n L)^4 EI}{v_n^0(L)^2 L^3} \qquad (2.41)$$

Thermal Noise

Numerous authors have reported the effect of a finite temperature on cantilever vibrations [30–32]. Briefly, the mean square vibration amplitude of a given mode at a temperature T is given by the equipartition theorem (equation (2.42)).

$$\frac{1}{2}k_BT \;=\; \frac{1}{2}D_n\langle u^2\rangle \quad (2.42)$$

Extending the analysis of the excited cantilever with damping, Rast *et al.* [31] conclude that the spectral density of the nth mode can be written as follows:

$$S_n(\omega) \;=\; \frac{2k_B\omega_n^3T/QD_n}{\left((\omega^2-\omega_n^2)^2+\frac{\omega_n^2\omega^2}{Q^2}\right)} \quad (2.43)$$

Comparing equation (2.43) and the square of equation (2.39) suggests a simple method to establish whether a peak in the measured power spectrum is thermal or due to coupled mechanical vibrations. The first step is to fit to the peak being investigated, a general equation with three[4] parameters A, Q, and ω_n of the form: $A\left((\omega^2-\omega_n^2)^2+\frac{\omega_n^2\omega^2}{Q^2}\right)^{-1}$. The value obtained for Q can then be used to compute the expected thermal peak and compare with the measurement. An example of this procedure is shown in figure 3.17.

2.3 Deflection Measurement

2.3.1 Principle

The measurement of the current is described in section 2.1.3. Assuming negligible piezoelectric non-linearities and time dependent issues (creep and drift), the vertical piezoelectric tube extension is obtained readily by multiplying the voltage applied to the tube by its vertical sensitivity. The last key observable that needs to be measured is the deflection of the cantilever. As illustrated in figure 2.7, once the deflection of the cantilever at the tip is known, the rigid tip–sample separation can be obtained by subtracting the z-piezo extension from the deflection. The displacement (or depth, h) is the negative of the tip–sample separation, and will be used primarily throughout this document, as the majority of the results reported here are obtained in nanoindentation mode, and thus have negative tip–sample separations.

In order to obtain a measurement of the deflection, an interferometer (section 2.3.2) is used in conjunction with the results of section 2.2.1 to compute the value of the deflection of the laser beam at the tip.

Using equations (2.15) to (2.17) and figure 2.8, we can write the difference in beam deflection δu between the locations $z_2 > z_1$ illuminated by the two beams, as

---

[4]In certain cases, when the baseline of the optoelectronic noise is too high, an additional vertical offset parameter may be required
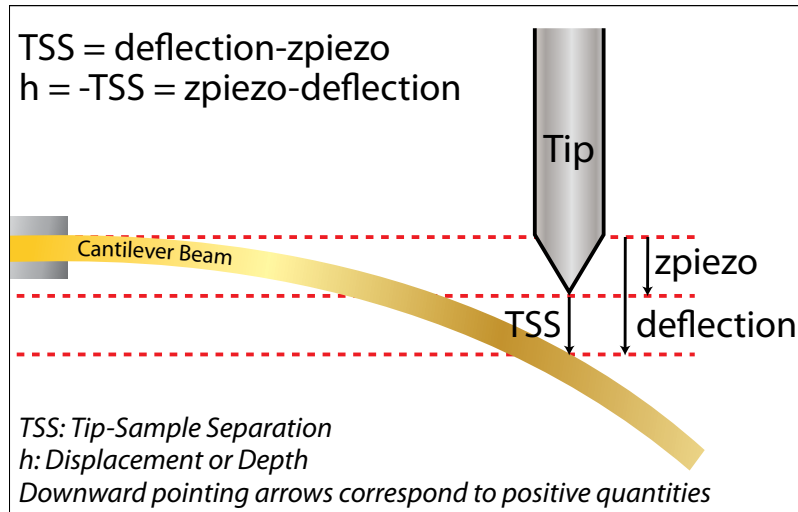
Figure 2.7: The tip–sample separation is obtained by subtracting the extension of the z-piezo from the cantilever beam deflection. A further additive constant is then added for zero-point adjustment. Unless otherwise specified, the zero point is set as the separation at tunnelling contact, generally at a tunnelling resistance of 500MΩ, corresponding to a setpoint of 100pA at 50mV.
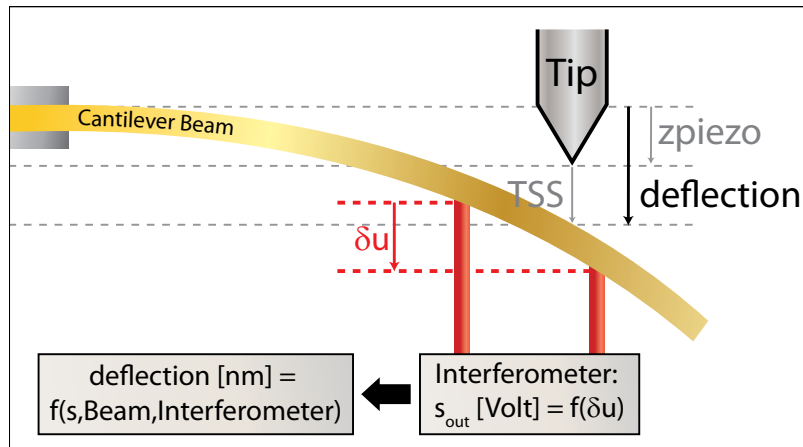
Figure 2.8: Determination of the deflection from the interferometer signal. The interferometer outputs a signal s, which is proportional to the path difference between two beams reflecting off the sample at two different positions along the cantilever. As drawn, $\delta u > 0$ (arrow pointing down), as $u_t > 0$.

a function of the beam deflection at the tip u_t, z_2/t, and z_1/t, where t is the tip position. The result is given by equations (2.44) and (2.45), where Θ is the Heaviside step function ($\Theta(x < 0) = 0$, $\Theta(x > 0) = 1$, and $\Theta(0) = 1/2$). The ζ function is convenient both for data processing and discussion and is plotted in figure 2.9.

$$\delta u = u_t \left[\zeta \left(\frac{z_2}{t} \right) - \zeta \left(\frac{z_1}{t} \right) \right] \tag{2.44}$$

$$\zeta(x) = \Theta(x-1) \left[\frac{3}{2}x - \frac{1}{2} \right] + \Theta(1-x) \left[\frac{3}{2}x^2 - \frac{1}{2}x^3 \right] \tag{2.45}$$
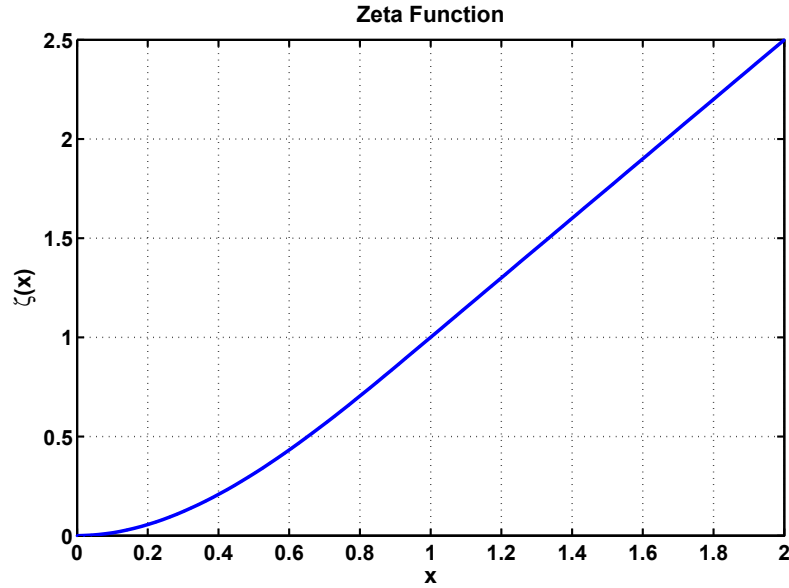
Figure 2.9: Plot of the ζ function (see equation (2.45)).

With knowledge of the beam positions, the tip position, and δu, the cantilever beam deflection at the tip can easily be computed from equation (2.44). In our system, the beams are fixed at a distance $d_b = 1$mm from one another, while the sample can be moved laterally with respect to the beams. The sample can be moved to enhance the sensitivity. The best sensitivity for a given t and u_t is thus attained when δu is maximised by the beam positions.

As illustrated by figure 2.10, the sensitivity for a fixed u_t is highest for $z_2 > z_1 > t$, in which case equation (2.44) can be simplified to equation (2.46). In the worst case, $z_1 = 0$, and equation (2.44) become equation (2.47).

$$\frac{\delta u_{\text{SDM}}^{\text{max}}}{u_t} = \frac{3}{2} \frac{d_b}{t} \tag{2.46}$$

$$\frac{\delta u_{\text{SDM}}^{\text{min}}}{u_t} = \zeta \left(\frac{d_b}{t} \right) \tag{2.47}$$

In equations (2.46) and (2.47), SDM stands for *Static Deflection Model*, which is the model derived from the Euler-Bernoulli beam theory (see section 2.2.1). In particular, it assumes that the beam is properly anchored at one extremity. As we will
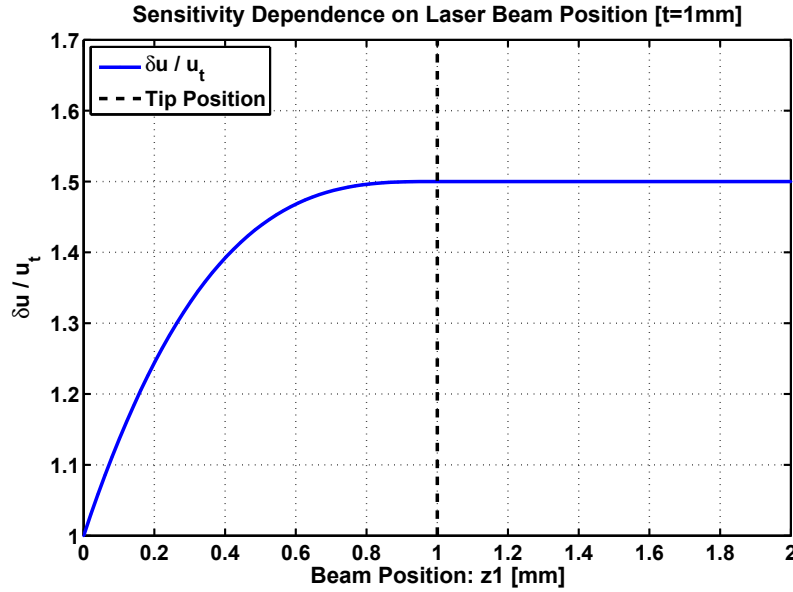
Figure 2.10: Deflection sensitivity dependence on beam position, for a fixed tip position and tip deflection. z_1 is the position of the laser beam closest to the fixed end. The second laser beam is positioned at $z_2 = z_1 + 1\text{mm}$.

see in section 3.3.1, poor clamping, which does not ensure that $u'(0) = 0$, will result in a completely different deformation of the cantilever and the previous expression will no longer hold.

Comparison with Direct Measurement

As illustrated in figure 2.11, another possible approach is to employ a *Direct Measurement* (DM) configuration, whereby one laser beam is below the tip and the second one reflects off a fixed surface, which can be the base of the cantilever at the fixed end, or a separate reference point[5].

In a Direct Measurement configuration, we have:

$$\boxed{\frac{\delta u_{\text{DM}}}{u_t} \quad = \quad 1 \quad (2.48)}$$

It is interesting to note that when the tip is sufficiently close to the fixed end of the sample $t < 3d_b/2$ and the laser beams are optimally positioned, the Static Deflection Model configuration provides a better sensitivity for a given deflection at the tip than the Direct Measurement configuration, as illustrated in figure 2.12,

---

[5]A direct configuration measurement with a separate reference point is hard to achieve experimentally with manually prepared cantilevers, as the two reflecting surfaces need to be perfectly aligned to allow the interferometer beam to recombine properly.
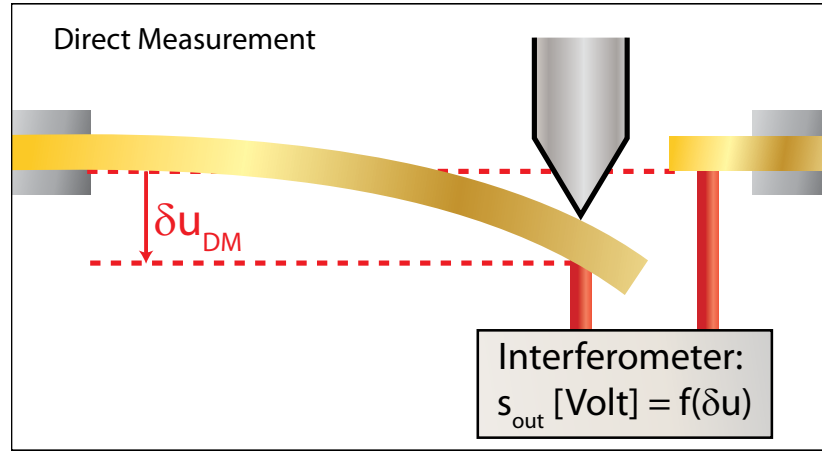
Figure 2.11: Configuration of interferometer for direct measurement of the deflection. Under these conditions, $\delta u_{DM}/u_t = 1$, where DM stands for *Direct Measurement*

which is plotted for our instrumentally constrained value of $d_b = 1$mm. Furthermore, if $t < d_b$, the sensitivity for fixed u_t is better irrespective of the beam positions.
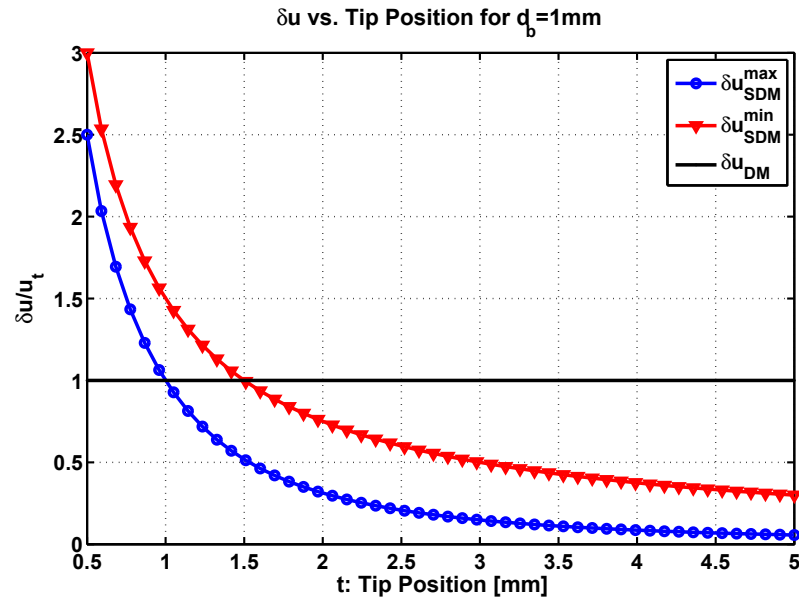
Figure 2.12: Sensitivity at fixed u_t as a function of tip position for various configurations, assuming $d_b = 1$mm.

2.3.2   Interferometer

Our interferometer is based on Schönenberger and Alvarado's design [33], and was described by Stalder [5]. Its principle is illustrated in figure 2.13.
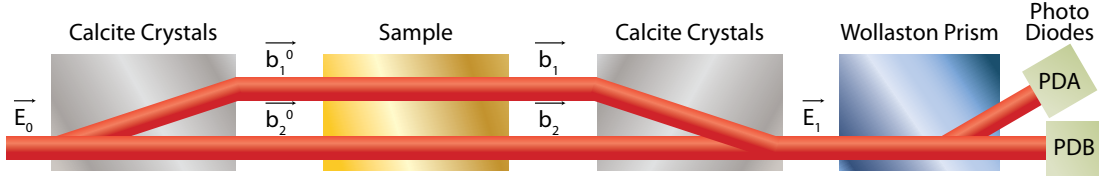
Figure 2.13: Principle of operation of the interferometer (simplified *"unfolded"* linear drawing): A laser beam traversing a calcite crystal in split into two parallel beams separated by 1mm and with orthogonal polarisations. As the beams reflect off the bent sample, a phase difference between the two beams is produced before they recombine in the calcite crystal. A Wollaston prism with axes at 45° with respect to the calcite is then used to split the beams again. The measurement and subtraction of the intensities of these beams allows the measurement of the phase difference, which can be related to the cantilever bending. See text for details.

A calcite crystal is a birefringent crystal. As such, light travelling through it is split into two parallel beams separated by certain distance d_b which depends on the calcite length. The two outgoing beams are polarised linearly and their polarisations are mutually orthogonal.

A Wollaston prism is composed of two birefringent crystal prisms glued to one another at a angle such that a beam travelling through a Wollaston prism will be decomposed into two diverging beams, also linearly polarised with orthogonal polarisations.

The principle of the interferometer is relatively simple. An incoming laser beam is split into two parallel beams separated by a distance $d_b = 1$mm. The two beams are of orthogonal linear polarisations (along orthonormal vectors $\hat{c}_1$ and $\hat{c}_2$ respectively). The two beams, whose electric fields are denoted $\overrightarrow{b_1^0}$ and $\overrightarrow{b_2^0}$ respectively, are then reflected off the sample. Due to the cantilever deflection difference δu (see section 2.3.1 and figure 2.8) between the locations attained by the two beams, one of the beams will travel a path difference of $2\delta u$ more than the other[6]. This path difference will result in an additional phase factor $e^{i\phi}$ for one of the beams. The two beams ($\overrightarrow{b_1}$ and $\overrightarrow{b_2}$) are then recombined into one by travelling in the reverse direction in the calcite crystal described earlier: $\overrightarrow{E_1} = \overrightarrow{b_1} + \overrightarrow{b_2}$. The beam is then split in two again, this time using a Wollaston prim, with axes rotated 45° with respect to the calcite. The outgoing beams, of linear polarisations $\hat{w}_A$ and $\hat{w}_B$ are then detected by two photodiodes PDA and PDB respectively. A difference of the signals from the two photodiodes then provides a measure of δu. To understand why, let us assume that

---

[6]The factor 2 in $2\delta u$ is due to the fact that the beam travels to the sample and back.

the initial electric field is $\overrightarrow{E}_0$, with polarisation vector $\hat{n}$.

$$\overrightarrow{E}_0 = E_0 e^{i(\overrightarrow{k}\cdot\overrightarrow{r}-\omega t)}\hat{n} \qquad (2.49)$$

In equation (2.49), both E_0 and $\hat{n}$ can be complex numbers. After the calcite, the two beams are such that $\overrightarrow{b_k^0} \propto \overrightarrow{E_0} \cdot \hat{c}_k \ \hat{c}_k$, where the proportionality constant is a complex number describing potential gains in phase and losses in amplitude, and is assumed the same for the two beams. Pursuing this reasoning and projecting against $\hat{w}_k$, we obtain $\overrightarrow{E_A}$ and $\overrightarrow{E_B}$ at the photodiodes. The results are summarised in equations (2.50) and (2.51), where T is an overall transmission coefficient, α is a phase which is common to both beams, and ϕ is the additional phase of beam one due to the extra path $2\delta u$.

$$\overrightarrow{E_A} = e^{i\alpha}\sqrt{T}\left[\left(\overrightarrow{E_0}\cdot\hat{c}_1\right)\ (\hat{c}_1\cdot\hat{w}_A)\ e^{i\phi} + \left(\overrightarrow{E_0}\cdot\hat{c}_2\right)\ (\hat{c}_2\cdot\hat{w}_A)\right]\ \hat{w}_A \quad (2.50)$$

$$\overrightarrow{E_B} = e^{i\alpha}\sqrt{T}\left[\left(\overrightarrow{E_0}\cdot\hat{c}_1\right)\ (\hat{c}_1\cdot\hat{w}_B)\ e^{i\phi} + \left(\overrightarrow{E_0}\cdot\hat{c}_2\right)\ (\hat{c}_2\cdot\hat{w}_B)\right]\ \hat{w}_B \quad (2.51)$$

Noting that the Wollaston prism's axes are at a 45° angle with respect to the calcite angles (Fig. 2.14), we have $\hat{c}_1 \cdot \hat{w}_A = \hat{c}_2 \cdot \hat{w}_A = \hat{c}_1 \cdot \hat{w}_B = -\hat{c}_2 \cdot \hat{w}_B = \sqrt{2}/2$.
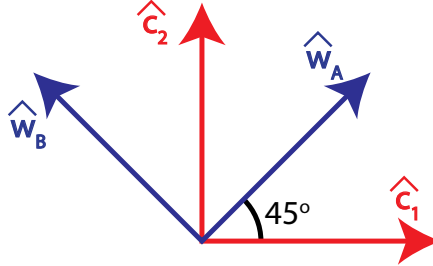
Figure 2.14: Calcite and Wollaston axes.

In addition, writing $n_1 = \hat{n} \cdot \hat{c}_1$, $n_2 = \hat{n} \cdot \hat{c}_2$, and denoting I_0 the intensity[7] of the incoming light, we can write simple expressions for the light intensity incident on the photodiodes (equations (2.52) and (2.53)).

$$I_A = \frac{TI_0}{2}\left|n_1\ e^{i\phi} + n_2\right|^2 \quad (2.52)$$

$$I_B = \frac{TI_0}{2}\left|-n_1\ e^{i\phi} + n_2\right|^2 \quad (2.53)$$

---

[7]The light intensity can be written $I_0 = c\varepsilon_0|E_0|^2/2$, where c is the speed of light in vacuum, and ε_0 the permittivity of vacuum.

Polarisation of Incoming Beam

For a linearly polarised incoming beam, n_1 and n_2 are real. Remembering that $n_1^2 + n_2^2 = 1$ and $\cos^2\phi + \sin^2\phi = 1$, we obtain equations (2.54) and (2.55).

$$
\begin{aligned}
I_A^{\text{linear}} &= \frac{TI_0}{2}(1 + 2n_1n_2\cos\phi) && (2.54) \\
I_B^{\text{linear}} &= \frac{TI_0}{2}(1 - 2n_1n_2\cos\phi) && (2.55)
\end{aligned}
$$

For a right-handed circularly polarised incoming beam, $n_1 = 1/\sqrt{2}$ and $n_2 = i/\sqrt{2}$. This leads to equations (2.56) and (2.57). In the case of left-handed polarised light, the result is the same with opposite signs in front of the sine function.

$$
\begin{aligned}
I_A^{\circlearrowright} &= \frac{TI_0}{2}(1 + \sin\phi) && (2.56) \\
I_B^{\circlearrowright} &= \frac{TI_0}{2}(1 - \sin\phi) && (2.57)
\end{aligned}
$$

It is interesting to look at the the two photodiode intensities' sum and difference in both cases. In both sets of equations, $I_A + I_B = TI_0$. In the absence of losses ($T = 1$) the full light intensity is recovered. The difference is given by equations (2.58) to (2.60).

$$
\begin{aligned}
\Delta I^{\text{linear}} &= 2TI_0n_1n_2\cos\phi && (2.58) \\
\Delta I^{\circlearrowright} &= TI_0\sin\phi && (2.59) \\
\Delta I^{\circlearrowleft} &= -TI_0\sin\phi && (2.60)
\end{aligned}
$$

It becomes apparent that the phase angle can be easily obtained from a measurement of the difference of the photocurrents from the two diodes. It is important to recognise that the absolute value of the factor in front of the sine and cosine functions can be interpreted as an amplification factor. The higher it is, the higher the sensitivity. In the case of circularly polarised light, this factor is always TI_0. In the case of linear polarisation, this factor can take any value between 0 and TI_0. The maximum is attained when $|n_1| = |n_2| = 1/\sqrt{2}$, which corresponds to polarisations of $\pm 45^o$ and $\pm 135^o$. This is intuitive to understand in the sense that the best sensitivity will be obtained when the incoming beam is split evenly along the two calcite directions. In addition, when the incoming beam is linearly polarised along one of the calcite axes, it passes through unaffected and the second beam is completely extinguished.

Phase and Path Difference Measurement

In order to measure changes in phase, it is convenient to offset the phase by a constant so that the intensity difference is proportional to $\sin\phi$. This is done experimentally by bending the calcite crystal lightly with a piezotube. The s signal, which is proportional to the deflection is obtained by amplifying the ratio of the difference to the sum of the photocurrents.

$$s \quad \propto \quad \frac{\Delta I}{\sum I} = \frac{U_{\text{p2p}}}{2} \, \sin\phi \quad (2.61)$$

In equation (2.61), U_{p2p} is the peak-to-peak value of the s signal for sufficient phase variation. During a measurement, when the phase is very small, the $\sin(x) \simeq x$ approximation holds, and it can be estimated from equation (2.62).

$$\phi_{\text{small}} \quad = \quad \frac{2s}{U_{\text{p2p}}} \quad (2.62)$$

The path difference being $2\delta u$, we can write $\phi = k \cdot 2\delta u = 4\pi \, \delta u/\lambda$, where k is the wave vector of the laser beam and λ is the wavelength. In our experimental system, $\lambda = 780$nm, which gives $\lambda/4\pi \simeq 62$nm. For a range of $\delta u = \pm 20$nm, which is far in excess of our experimental needs, we have $\phi = 4\pi \, \delta u/\lambda < 1/3$, which yields an error smaller than 2% in the approximation of equation (2.62). We can therefore conclude that equation (2.63) holds in our experimental conditions.

$$\delta u \quad = \quad \frac{s}{U_{\text{p2p}}} \frac{\lambda}{2\pi} \quad (2.63)$$

Experimental Considerations

Our experimental implementation of the interferometer is illustrated in figure 2.15. The size of the complete assembly is approximately 5cmx5cmx1cm. It is situated in UHV, immediately below the STM stage. The incident beam is generated by a temperature stabilised laser diode outside the chamber and guided into the interferometer by optical fiber. A number of lenses allow the beam collimation to be adjusted so that the spots on the sample are sharp. After passing through the lenses, the beam is reflected by a mirror towards a mini-prism, which reflects it towards the calcite crystal. The beam is then split in two before it reflects off the cantilever and recombines inside the same calcite crystal. The angles of the cantilever are adjusted remotely by controlling piezomotors on the sample stage, so that the recombined

beam reaches the left side of the mini-prism and therefore attains a second corner mirror which aims it at the Wollaston prism. It is then split into two beams which are independently analysed by photodiodes.
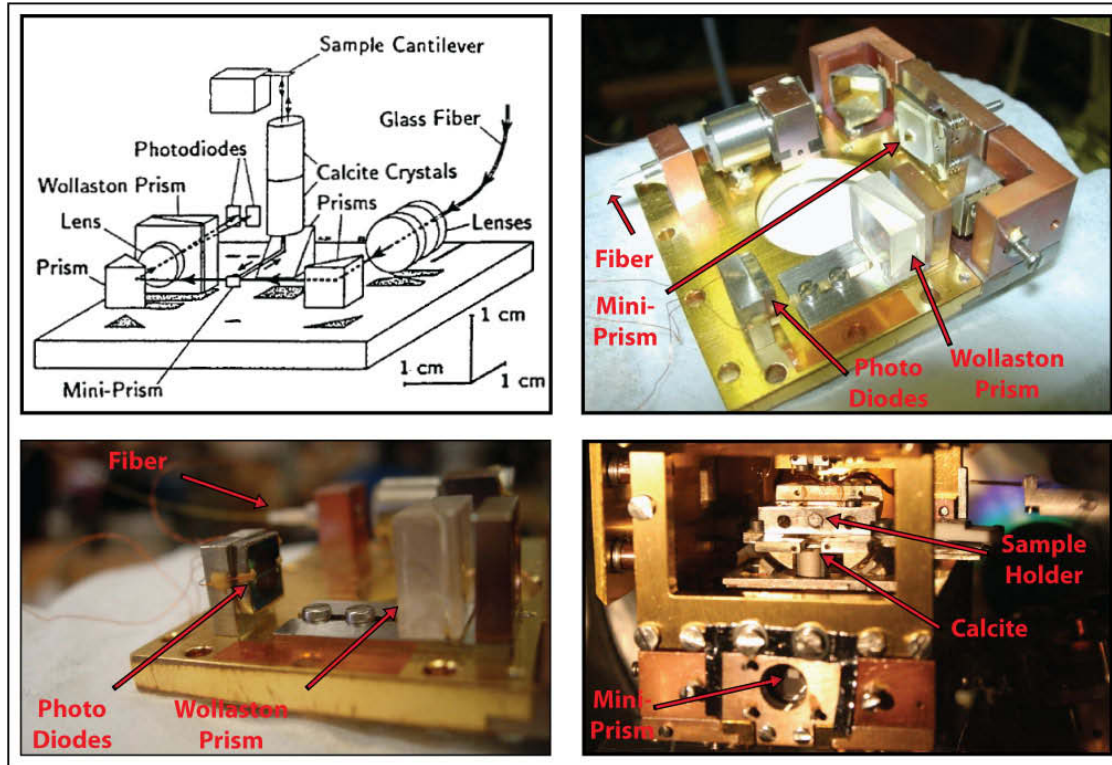
Figure 2.15: Interferometer diagram and photographs. The diagram in the upper left corner is taken from the PhD thesis of Stalder [5].

To change the phase angle discussed above, either to adjust the offset or to determine U_{p2p}, the calcite crystal can be bent gently by applying an appropriate voltage to a piezoelectric tube in which the calcite is mounted. In order to facilitate this bending motion, the crystal was split into two parts that have been fixed at opposite ends of the piezoelectric tube.

As discussed earlier, the polarisation state of the incoming beam is of great importance. For the experiments reported in this thesis, the polarisation was adjusted simply by imposing a small deformation on the external part of the optical fibre until the two beams incident on a milk glass sample were of equal brightness (Fig. 2.16). Fine adjustments can also be made during the cantilever alignment procedure described above.
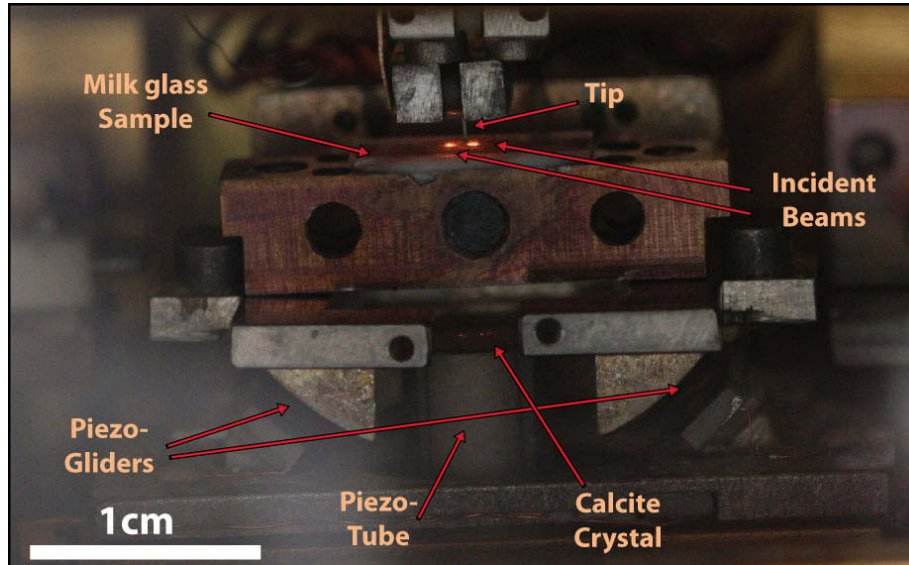
Figure 2.16: Visualization of the interferometer beams on a *milk glass* sample.

2.4   Noise and Vibrations

2.4.1   Noise Sources

In the rest of this document, we use the term *noise* to refer to all signals other than the specific signal we are attempting to measure, and that are added to it by various sources. In Scanning Probe Microscopy (SPM), noise sources can generally be divided into two main categories: optoelectronic and mechanical noise. Optoelectronic noise encompasses all sources of noise that are related to the use of lasers, photodetectors, and electronic components in general.

Mechanical noise refers to mechanical vibrations coupled into the cantilever beam by other vibrating objects, such as buildings and vacuum pumps. Some of the most relevant sources of noise in our context are reviewed below.

Shot Noise: One of the primary sources of noise, shot noise, arises from the statistics of photons incident on a photodetector. The power spectral density of the shot noise in volts is frequency independent and given by equation (2.64), after Sarid [14].

$$S_{\text{shot}}(f) = G^2 \, 2eI \qquad (2.64)$$

In equation (2.64), e is the electronic charge, I is the photocurrent, and G is the gain that transforms the photocurrent into a voltage.

When the signal is measured over a bandwidth Δf, we get:

$$\sqrt{\Delta V_{\text{shot}}^2} = G\sqrt{2eI\Delta f} \qquad (2.65)$$

The power of the signal being G^2I^2, it is immediate to see that the signal-to-noise ratio[8] is proportional to I. It is therefore advantageous to maximise the photocurrent.

Johnson Noise: The Johnson, or Johson-Nyquist noise, is generated by the thermal agitation of electric charge in the load resistor used to convert the photocurrent into a voltage [14]:

$$\sqrt{\Delta V_{\text{Johnson}}^2} = \sqrt{4k_B T G \Delta f} \qquad (2.66)$$

The Johnson noise is also frequency independent and will contribute to the white noise baseline.

Laser Noise: Fluctuations in the laser intensity and phase will also contribute to the noise baseline. In both cases, we have $\sqrt{\Delta V_{\text{laser}}^2} \propto GI$.

Electronic Noise: Active electronic components will generally contribute noise, both in terms of a white (frequency independent) and a pink (1/f) component. Additional peaks at specific positions are also observable, for example, at multiples of the line frequency (60Hz).

External Vibrations: The coupling of external vibrations to the system will generally have a 1/f componentas well as specific peaks that can interact with the cantilever modes.

Thermal Mechanical Noise: As discussed in section 2.2.2, the non-zero temperature of the measurement will cause thermomechanical fluctuations of the cantilever. The corresponding power spectral density is given by equation (2.43).

Typical Noise Spectra

Noise spectra are taken to characterize the dynamic response of cantilever beams, and to determine the noise characteristics of the system before each acquisition. Beyond providing an estimate of the measurement uncertainty, this allows us to identify possible sources of mechanical vibrations that may cause the cantilever to

---

[8]The signal-to-noise ration is defined as the ratio of signal power to noise power.

enter a level of vibration that would preclude stable STM/AFM, and possibly damage valuable tips.
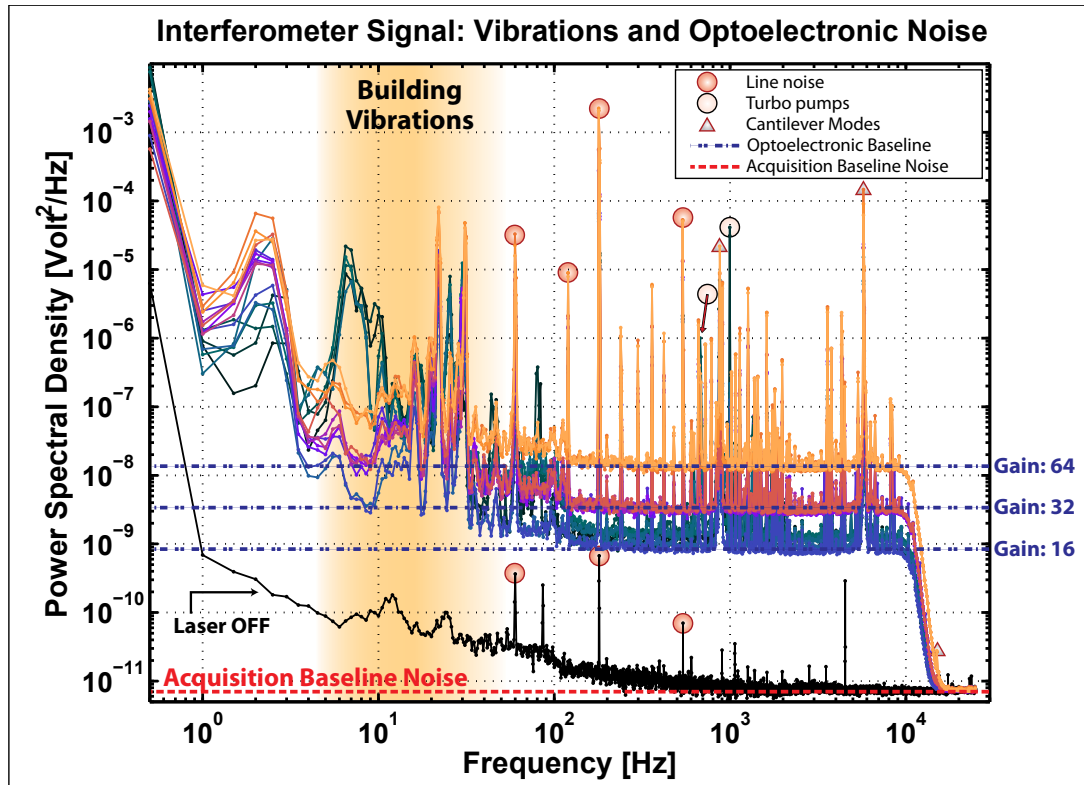
Figure 2.17: Various noise spectra taken with the same cantilever beam of resonance frequency $f_1 = 877$Hz, in different conditions. See text for details.

Figure 2.17 shows a series of example spectra taken with the same cantilever beam of first mode frequency $f_1 = 877$Hz. The spectra were taken in different vibration isolation conditions and with various electronic gains. The signal was filtered by a hardware order 8 low pass filter, and acquired with at 50kHz for a duration of 60s, and the one-sided power spectra were obtained using Welch's method [34], averaging 37 segments with a *hanning* window to obtain a resolution of 0.5Hz. For display, the data was decimated geometrically. The bottom spectrum, taken with the laser off and with a gain of sixteen shows the noise limit imposed by our electronic system. The figure also shows the extent of $1/f$ noise at low frequencies, both produced by the acquisition system and the laser noise.

The first three cantilever beam vibration modes are visible, as well as a number of electronic and mechanical peaks. Turbomolecular pumps induce vibrations at multiples of 753Hz, 1000Hz, and 860Hz, for the airlock, preparation chamber and measurement chamber pumps respectively. Building vibrations are visible below $\sim 40Hz$, and are visibly higher for spectra obtained when vibration isolation was not

used. A large number of electronic peaks are visible, including multiples of 60Hz. In particular, the 180Hz is found to be large and ubiquitous in all our measurements. Substantial, and fairly successful, efforts have been made to mitigate line noise, notably by eliminating ground loops. It is also interesting to note that the 120Hz peak comes from the laser, as it is visible in all spectra except the one acquired with the laser switched off.

When the gain is increased, the baseline of the optoelectronic noise is increased, but the baseline noise beyond the 10kHz line noise remains constant. The gain is applied in several stages. The first stage is applied by the photocurrent preamplifier and converter box, commonly referred to as the *s-box*. In a recent redesign of this amplifier, offset correction and high gains (up to 100) can be applied locally, on the UHV flange, before the signal travels to the acquisition system and acquires additional electronic noise. Higher s-box gain will increase the signal to noise ratio. Signal to noise ratios were not substantially affected by gain increases in this case, because the gains were increased on amplifiers located on the acquisition rack, and therefore amplifying signal which already have a substantial level of noise. The primary aim of these (LUPE) amplifiers is to ensure that the full dynamic range of the acquisition system is used.

In figure 2.17, the power spectrum displayed is the raw spectrum acquired directly from the interferometer signal, in volts. The data can be converted to nanometres by applying the procedures described in section 2.3. However, it is important to keep in mind that this conversion relies on the assumption that the deformation of the cantilever beam is described by the static deflection model introduced earlier. This assumption no longer holds when one attempts to extract quantitative mechanical vibration values from vibration spectra, as the various vibration modes differ substantially from the static deflection solution. A notable exception is the first mode of vibration: the first mode and the static solution are very similar, and differ by a maximum a of 4% in the quarter of the cantilever closest to the free end. Therefore, in addition to providing us with a root-mean-square values of the noise in the measurement of the static signal, converting similar spectra into nanometres will allow us to get a quantitative measure of the vibration amplitude of the first mode. In particular, this allows quantitative thermal peak fitting, as discussed in section 2.2.2.

Another important point to note is that the spectra shown in figure 2.17 are taken without explicit excitation[9] and without loading by the tip. Changes in spectra upon

---

[9]Excitation can be performed by a piezoelectric shaker below the sample holder which can be
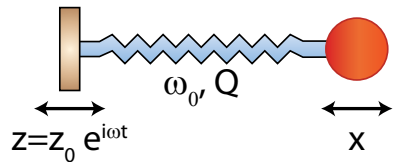
loading are highlighted in section 3.3.2.

2.4.2 Vibration Isolation

As suggested by figure 2.17, a number of external excitations can couple into the cantilever vibrations. In order to carry out effective measurements, it is of the utmost importance that a proper vibration isolation system be used. Such systems have been studied in detail by a number of researchers since the invention of the STM in 1982 [2] and the AFM in 1986 [3]. Models of vibration isolation systems as coupled damped oscillators [35, 36] can easily be adapted to our context. Before such a model is outlined, it is useful to review important results from the analysis of a damped harmonic oscillator.

Damped Harmonic Oscillator

The motion of the extremity of a spring driven at the other end by $z = z_0 e^{\imath \omega t}$ can be described by figure 2.18 and by the differential equation (2.67).

$$\ddot{x} + \frac{\omega_0}{Q}(\dot{x} - \delta \ \dot{z}) + \omega_0^2 = \omega_0^2 z \qquad (2.67)$$

$$\delta \in \{0; 1\}$$

Figure 2.18: Damped Harmonic oscillator

In equation (2.67), δ can be zero or one, depending on the nature of the damping. In our analysis, we will take $\delta = 1$, as it is the most common case [35]. Introducing the ansatz $x = A e^{\imath \phi}$, the transfer function giving x from z is obtained immediately. In equation (2.68), $\Omega = \omega / \omega_0$.

$$H_x \ = \ \frac{x}{z} = \frac{A e^{\imath \phi}}{z_o} = \frac{\omega_0^2 + \delta \ \imath \dfrac{\omega_0 \omega}{Q}}{(\omega_0^2 - \omega^2) + \imath \dfrac{\omega_0 \omega}{Q})} = \frac{1 + \delta \ \imath \Omega/Q}{1 - \Omega^2 + \imath \Omega/Q} \quad (2.68)$$

If the quantity of interest is the difference between x and z, the transfer function is given by equation (2.69).

$$H_{x-z} \ = \ \frac{x - z}{z} = H_x - 1 = \frac{\Omega^2 + \delta \ \imath \Omega/Q}{1 - \Omega^2 + \imath \Omega/Q} \quad (2.69)$$

excited by a lock-in amplifier to obtain the amplitude and phase response of the cantilever beam to specific excitation.
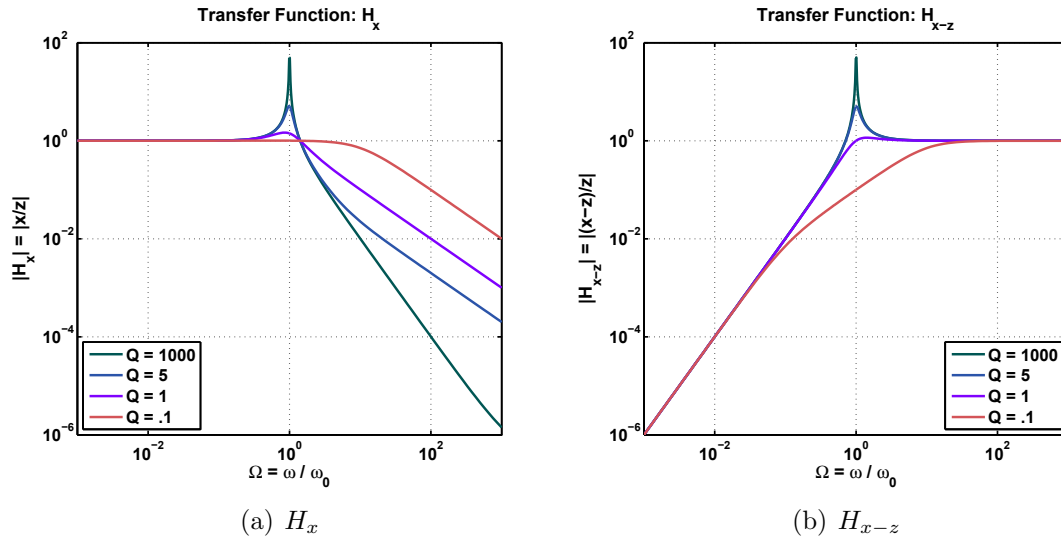
Figure 2.19: Transfer functions for a Damped Harmonic Oscillator, for various Q-factors.

The amplitude dependence of the transfer functions, shown in figure 2.19, has important implications for the design of a vibration isolation system. When one aims to decouple the vibrations of x from z (e.g. when one wishes to isolate the system from building vibrations) H_x is the quantity of interest, and the oscillator is therefore a low pass filter. In order to maximise the damping effect, it is important to choose a resonance frequency as small as possible. The Q-factor is required to be small enough to mitigate the resonance peak at ω_0, but large enough that the high frequency filtering is not excessively affected (see figure 2.19, part a).

Conversely, when the aim is to reduce the fluctuations between z and x, as is the case when designing an STM stage, H_{x-z} is the relevant quantity. In this case, the system acts as a high pass filter, and the resonance frequency of the assembly needs to be as high as possible, as suggested by figure 2.19, part b.

In a real implementation, both transfer functions are relevant as the system will comprise both decoupling oscillators and rigid elements, as described below.

Model of the Vibration Isolation System

Following the approach described by Pohl [35], we model the system as described in figure 2.20.

Our vibration isolation system is composed of several stages. The first stage of isolation consists of four springs attached to the ceiling which support the entire vacuum system. In the real system, these springs are four bungee cords that have been custom ordered to accommodate the weight of the system (approximately 900kg)
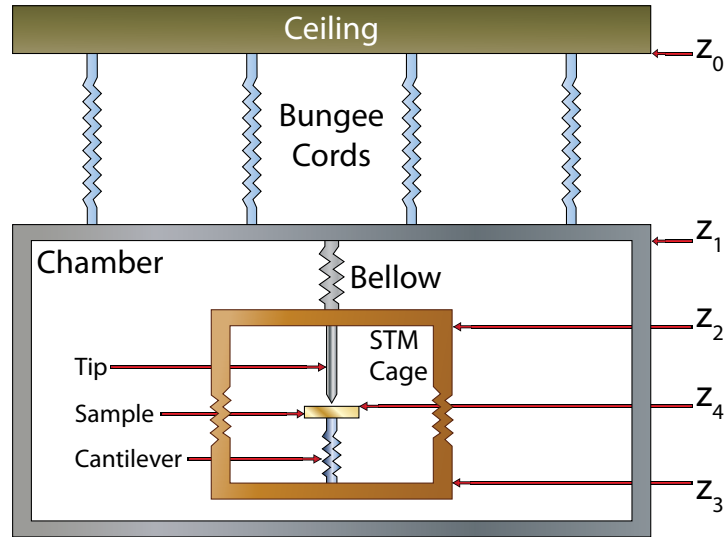
Figure 2.20: Vibration Isolation Model. See text for details.

within the space available and to ensure resonance frequencies of a couple of Hertz. In the following, we model the bungee cords by a single damped oscillator of resonance frequency f_{bungee} and Q-factor Q_{bungee}. The second stage is performed by a system base, composed primarily of a bellow attached to the top flange of the measurement chamber, which supports the STM cage. This system was described in detail by Stalder [5, 37]. A magnetic eddy current damping component further enhances vibration damping between the bellow and the cage. For simplicity, we model the whole assembly by a single pair of frequency and Q-factor: f_{bellow} and Q_{bellow}. The STM cage is designed to be as rigid as possible, but its resonance frequency is on the order of 1.5kHz. The main vertical vibration mode of the cage is modelled by f_{cage} and Q_{cage}. The vertical component of the vibration of the cantilever beam is the last step that affects the rigidity of the tip–sample junction. It is modelled as a damped harmonic oscillator of frequency and Q-factor f_{CB} and Q_{CB} respectively, taken as those of the first deflection vibration mode described in section 2.2.2.

Assuming that the positions of the various elements are described by the quantities illustrated in figure 2.20, one can use equation (2.68) to relate the vibrations of the ceiling z_0 to the vibrations of the cantilever z_4. For that purpose, we write the vibrations of the chamber $z_1 = H_{\text{bungee}}z_0$, the vibrations of the top of the STM cage $z_2 = H_{\text{bellow}}z_1$, the vibrations of the bottom of the cage $z_3 = H_{\text{cage}}z_2$, and the vibrations of the cantilever $z_4 = H_{\text{CB}}z_3$. Up to a constant, we can write the measured cantilever deflection u, and the tip–sample separation tss, as equations (2.70) and (2.71) respectively, which can be used to predict the effect of external vibrations on

the system.

$$u = z_4 - z_3 = (H_{\text{CB}} - 1)H_{\text{cage}} \ H_{\text{bellow}} \ H_{\text{bungee}} \ z_0 \quad (2.70)$$

$$\text{tss} = z4 - z_2 = (H_{\text{CB}} \ H_{\text{cage}} \ - 1)H_{\text{bellow}} \ H_{\text{bungee}} \ z_0 \quad (2.71)$$
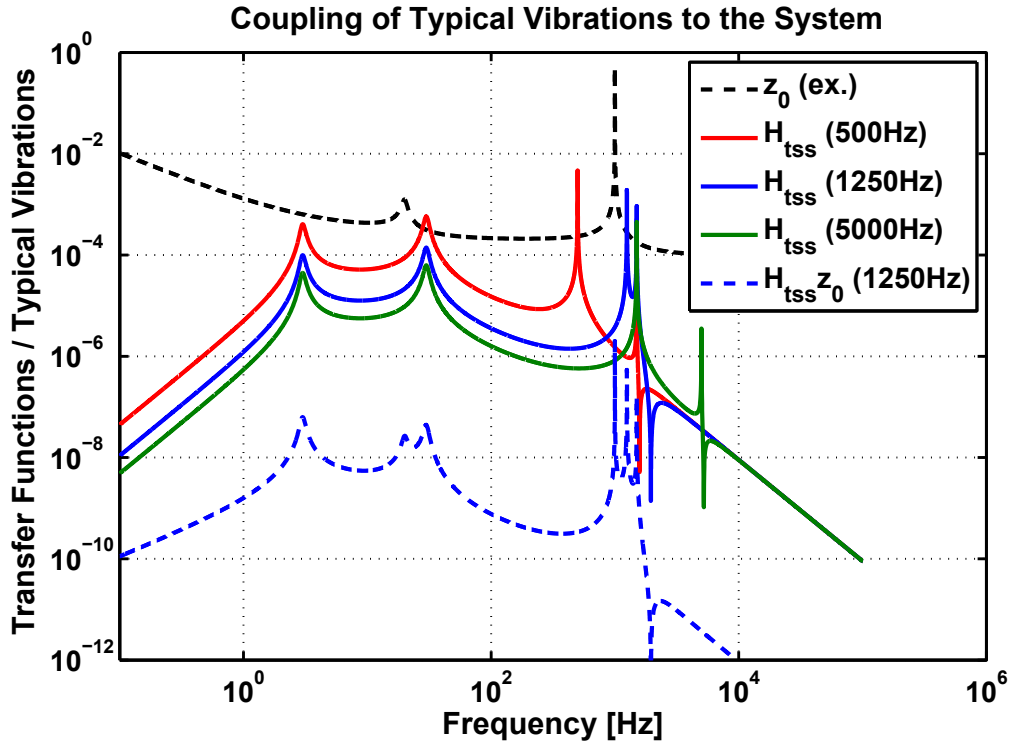
Figure 2.21: Coupling of vibrations into the tip–sample separation for various cantilevers. The solid lines are transfer functions calculated for different cantilever resonance frequencies. A hypothetical excitation spectrum with components of white noise, pink noise, and sporadic peaks (top dashed line) couples into the tip–sample separation (bottom dashed line). See text for details.

Figure 2.21 shows various transfer functions (in solid lines) calculated assuming specific approximate values for the frequencies and Q-factors of the different components: $f_{\text{bungee}} = 3\text{Hz}$, $Q_{\text{bungee}} = 10$, $f_{\text{bellow}} = 30\text{Hz}$, $Q_{\text{bellow}} = 10$, $f_{\text{cage}} = 1500Hz$, $Q_{\text{cage}} = 1000$. The difference between the three solid lines is the cantilever resonance frequency which was taken to be 500Hz, 1250Hz, and 5kHz respectively, with a Q-factor of 1000. The top dashed line corresponds to a hypothetical excitation spectrum with components of white and pink noise, and sporadic peaks which mimic building vibrations of $\sim 20\text{Hz}$, and turbomolecular pump vibrations at 1kHz. The bottom dashed line corresponds to the coupling of the input spectrum into the tip–sample separation for the most realistic cantilever of $f_1 = 1250\text{Hz}$. The advantage of having cantilevers with high frequencies is evident in the figure, as the cantilever acts as

a high pass filter for the tip–sample junction. A higher high pass frequency filter will cause a higher attenuation of low frequency building vibration peaks, which will greatly improve the stability of the tunnelling gap.

Another advantage of a high cantilever beam frequency is that specific system frequencies are avoided. As illustrated by figure 2.22, and discussed in section 3.3.2, when an external vibration has a frequency close to the cantilever resonance, the vibration amplitude of the cantilever will increase dramatically, which can have disastrous effects on the tunnelling gap. Figure 3.15 shows an experimental example of one such interaction.

Effect of External Vibration on First Mode Amplitude
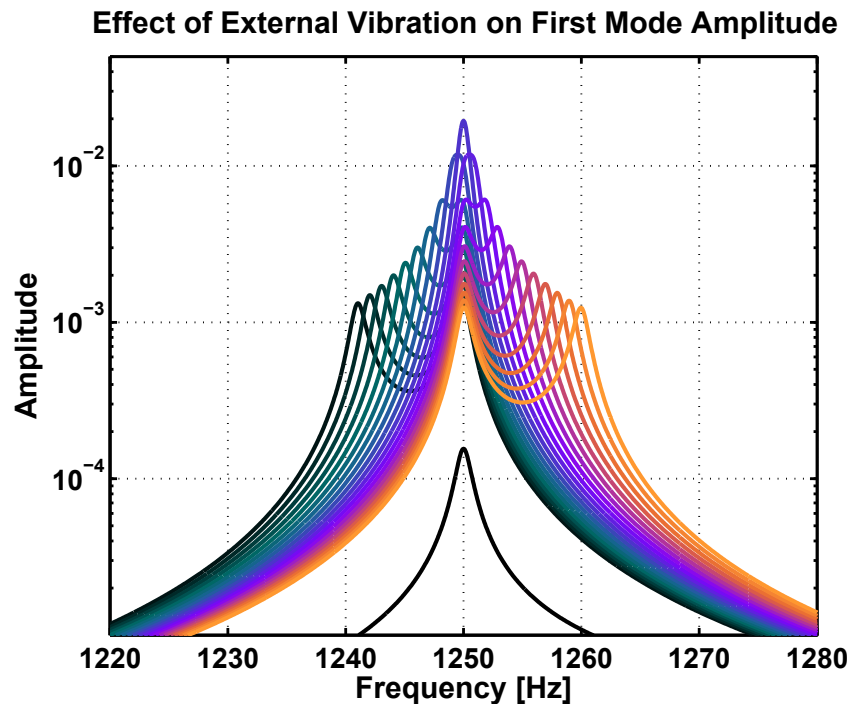
Figure 2.22: Amplitude response of the cantilever near the first mode, with external excitations at nearby frequencies: Each curve describes the amplitude response of a 1250Hz cantilever to a different external excitation. The coloured curves show the response when the frequency of the external excitation varies between 1241Hz and 1260Hz, while the bottom trace corresponds to an external excitation at 1350Hz.

Vibration Isolation Performance

The performance of the vibration isolation system is illustrated by the deflection noise spectra of figure 2.23. For that specific measurement, sixteen spectra were taken to evaluate the effect of the vibration isolation systems and of the various pumps attached to the chamber. The effect is easily appreciated visually. When

integrating the full spectra, and comparing values, one notices a very narrow distribution of noise values for spectra taken in our standard experimental conditions (fully operational vibration isolation and with both preparation chamber and load lock turbomolecular pumps switched off, and with the magnetically levitated pump on). In that case, integrating the noise spectra in a bandwidth of 0.5Hz–25kHz yields $\Delta u_L{}^2 = .043 \pm .002$nm^2 (standard error based on 7 measurements). The turbomolecular pump attached to the measurement chamber is magnetically levitated and does not appear to contribute to noise significantly. The effect of having all turbomolecular pumps switched on appears to have raised the noise to $\Delta u_L{}^2 = .13$nm^2. Finally, the distribution of noise values when different isolation systems were short-circuited was relatively wide and showed no consistent trend, although the noise levels were significantly higher: $\Delta u_L{}^2 = .29 \pm .05$nm^2.
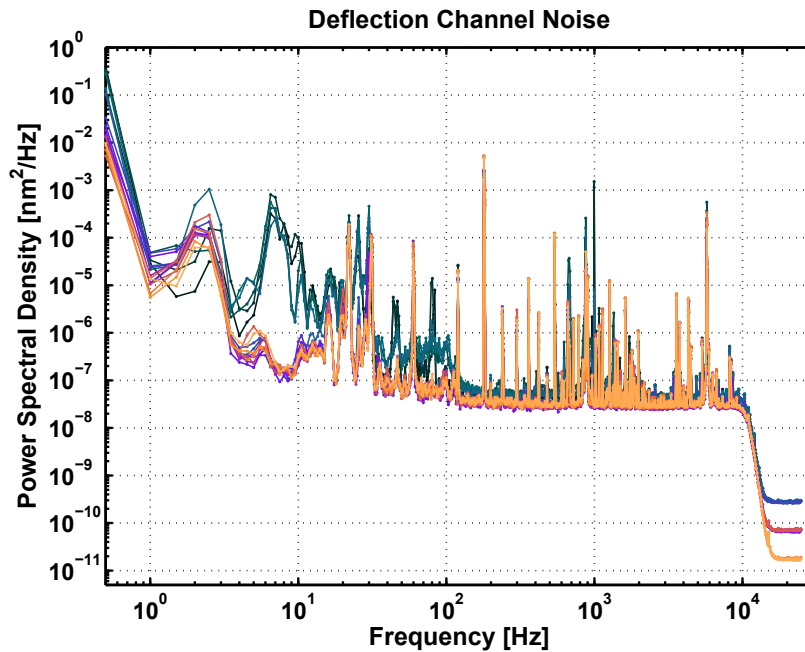
Figure 2.23: Deflection calibrated noise spectra from figure 2.17. Darker (higher) spectra where taken without vibration isolation.

2.5 Resolution and Design Parameters

2.5.1 Error Estimation

Based on equations derived in previous sections, we can write equations (2.72) to (2.77).

$$\delta u = \gamma s, \qquad \gamma = \frac{\lambda}{2\pi U_{\text{p2p}}} \tag{2.72}$$

$$\beta = \zeta(z_2/t) - \zeta(z_1/t), \qquad \beta^{\text{Best}} = \frac{3d_b}{2t} \tag{2.73}$$

$$u_t = \frac{\delta u}{\beta}, \qquad u_t^{\text{Best}} = \frac{2t}{3d_b}\delta u \tag{2.74}$$

$$k_t = \frac{Ew\tau^3}{4t^3}, \qquad k_0 = \frac{Ew\tau^3}{4L^3} \tag{2.75}$$

$$F = k_t u_t = \frac{k_t}{\beta}\delta u = \frac{k_t}{\beta}\gamma s \tag{2.76}$$

$$\frac{F^{\text{Best}}}{\delta u} = \frac{Ew\tau^3}{6t^2 d_b} = \frac{2k_0 L^3}{3t^2 d_b} = \frac{2k_t t}{3d_b} \tag{2.77}$$

As pointed out earlier, all other things being equal, the best resolution is attained when the interferometer laser beams are placed on the free end side of the tip. For simplicity, we assume that this is the case in the following, and we drop the *"Best"* superscript. Uncertainties of the measurement of δu, u_t, and F are given by equations (2.78), (2.79), and (2.80) respectively.

$$\frac{\Delta \delta u^2}{\delta u^2} = \frac{\Delta \gamma^2}{\gamma^2} + \frac{\Delta s^2}{s^2} \tag{2.78}$$

$$\frac{\Delta u_t^2}{u_t^2} = \frac{\Delta \delta u^2}{\delta u^2} + \frac{\Delta t^2}{t^2} + \frac{\Delta d_b^2}{d_b^2} \tag{2.79}$$

$$\frac{\Delta F^2}{F^2} = \frac{\Delta k_0^2}{k_0^2} + 3\frac{\Delta L^2}{L^2} + 2\frac{\Delta t^2}{t^2} + \frac{\Delta \delta u^2}{\delta u^2} + \frac{\Delta d_b^2}{d_b^2} \tag{2.80}$$

The most significant source of error is likely to the be the determination of the effective clamping point of the cantilever, which determines the offsets used in computing the tip position t and the cantilever length L. The uncertainty of this measurement is dominated by user error, and is on the order of $\sqrt{\Delta L^2} \simeq \sqrt{\Delta t^2} \simeq .1$mm.

The spring constant calibration procedure is described in section 3.3.1. Numerous measurements of cantilever spring constants typically yield a standard deviation between 1 and 2N/m for cantilever spring constants in the range of 5 to 40N/m. Since the measurement of the spring constant is repeated 5 times for each calibration, the error is multiplied by a factor of $\sqrt{1/5}$. This leads to an uncertainty of $\sqrt{\Delta k_0^2} \leq$ 1N/m. This leads to a relative error inferior to 10% for $k_0 > 10$N/m, and inferior to 4% for $k_0 > 30$N/m.

In practical terms, the effects of uncertainties on γ and d_b are difficult to separate. Their combined effect becomes visible during the very large force experiments with flat tips described in section 3.3.1, when the cantilever deflection is expected to match

the tip motion. A combined uncertainty of approximately 5% allows us to estimate upper bounds of 4% each for γ and d_b.

The last term that needs to be estimated is the uncertainty of the signal as acquired. This can be estimated directly by integrating noise spectra such as the one from figure 2.17. An estimate of the uncertainty on δu can also be readily obtained from such spectra. In the case of adequate damping, the spectra of figure 2.23 yield $\sqrt{\Delta \delta u^2} \simeq .05$nm. This value will naturally depend on the conditions of acquisition. Figure 2.24 shows a sample estimation of the force error, based on equation (2.80). For this cantilever, $k_0 = 17$N/m and $L = 6.7$mm. Using the uncertainty values outlined above, with in particular $\sqrt{\Delta \delta u^2} \simeq .05$nm, one obtains an estimate of the error, which depends on the tip position and the applied force.
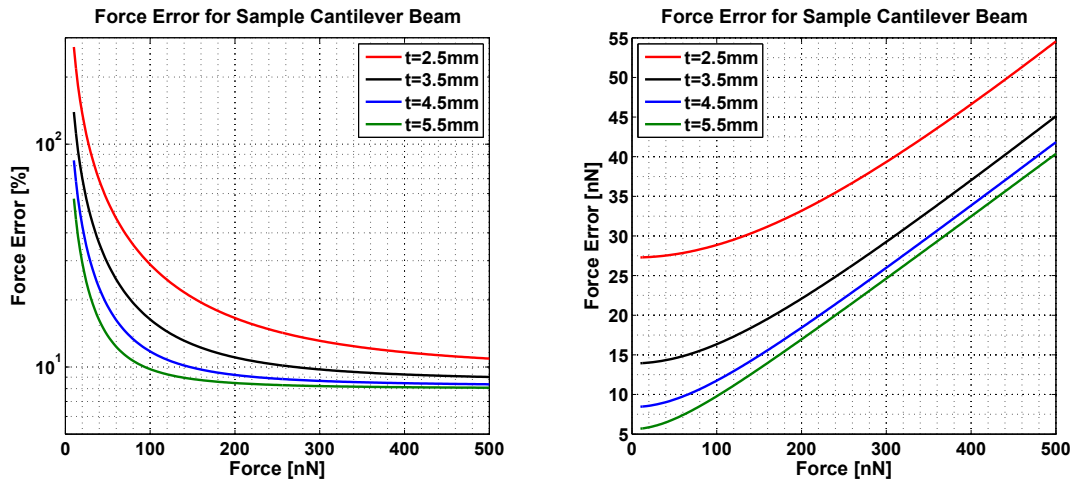
Figure 2.24: Error estimation on the force channel for a typical cantilever $k_0 = 17$N/m, $L = 6.8$mm.

The figure shows the expected relative and absolute errors on the force. In accordance with equation (2.80) and with the arguments made above, the relative error decreases with increasing tip position. For low forces, the relative error is very high, but it decreases with increased force, eventually plateauing around 10%. For different cantilevers, the relative error would also decrease with increasing k_0 and t.

A similar error analysis can be carried out on the deflection, yielding similar trends, as illustrated by figure 2.25.

2.5.2   Design Parameters

Several parameters are relevant when designing a cantilever beam for the FIM/STM/AFM system. These parameters are geometric parameters (w, L, τ, t) and material prop-
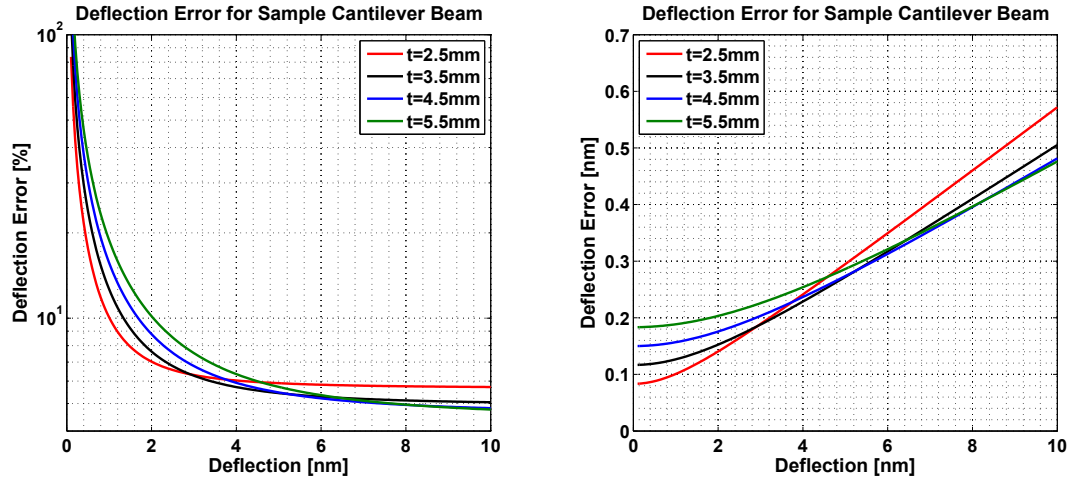
Figure 2.25: Example of an error estimation on the deflection channel.

erties (E, ρ). The material choice has been determined by a number of factors, as described in chapter 3. $E \simeq 110$GPa and $\rho = 2.7$ are therefore imposed by the choice of mica as a substrate. The geometric parameters are more flexible, and can be chosen in principle to optimise the quality of the measurements.

Spring Constant Choice

The various geometric parameters determine the resonance frequency and the spring constant of the cantilever, both at the free end, and at the tip. In order to assess the impact of the choice of cantilever spring constant, on the force measurement, it is instructive to think of the system as two springs in series, with respective spring constants k_t (cantilever at the tip position) and $k_j = S$ (tip-sample junction).

Assuming loading with force F, we apply an additional force dF by pushing the tip at one end by dz_{tip}. The force causes a deflection du_t at the other end such that: $dF = k_t du_t = S(dz_{\text{tip}} - du_t)$. This leads to the intuitive result:

$$du_t = dz_{\text{tip}} \left(1 + k_t/S\right)^{-1} \quad (2.81)$$

For a given resolution of z_{tip}, this implies that the maximum resolvable deflection is proportional to $(1 + k_t/S)^{-1}$. This shows that in cases when the cantilever spring constant is chosen much higher than the junction stiffness, the intrinsic resolution limit can worsen dramatically. One therefore has an advantage to make k_t as small as possible, provided that it remains higher than the strongest force gradients, to prevent snap-into-contact. Within a single approach curve, the junction stiffness will change greatly, but will remain below 400N/m in most cases. For these reasons, we choose spring constant k_t in the range of 50N/m to 300N/m.

Choice of Parameters

The previous point emphasised the physical basis for the choice of a low spring constant at the tip. However, previous discussions have also highlighted the importance of designing the cantilever to have a high resonance frequency. Rewriting equations (2.34), (2.75), and (2.77) as equations (2.82), (2.83), and (2.84).

$$f_1 = f_0 \; \frac{\tau}{L^2}, \qquad f_0 \simeq 1030\text{Hz (mica)} \quad (2.82)$$

$$f_1 = f_0' \; \frac{k_0^{1/3}}{L^2}, \qquad f_0' = f_0(4/Ew)^{1/3} \quad (2.83)$$

$$\frac{F}{\delta u} = \frac{2k_0 L^3}{2t^2 d_b} \quad (2.84)$$

Recalling the points made in section 2.5.1: in order to reduce the relative error on force measurement, and therefore improve accuracy, we require a large t, a large L, and a large k_0. To improve precision, on the other hand, we wish to minimise the ratio of F to δu, which implies a small k_0, a small L, and a large t (although it is constrained by L). In order to minimise vibration coupling issues, we would like to avoid specific system frequencies and have a large first mode frequency. According to equation (2.83), however, this is obtained for large values of k_0, and small values of L and w.

It becomes apparent that choosing values for design parameters involves significant trade-offs. In practical terms, the cantilever width, which is the most straightforward quantity to choose a value for, is set as small as possible without jeopardising the interferometer measurement, as the laser beams may miss a cantilever if it is too narrow. In practice, $w = 1.5$mm has proven to be a good choice. The next criterion that was favoured for the cantilevers used in this thesis was ensuring a high resonance frequency to avoid catastrophic interactions. Finally, the cantilever beam was chosen as small as possible with the tip as close to the free end as possible. As a rule of thumb, the aim of cantilever design is to obtain a spring constant on the order of 50N/m to 100N/m at a distance of 1mm from the free end.

Figure 2.26 illustrates the effect of different cantilever lengths and thicknesses on the resonance frequency, $F/\delta u$ and the spring constant. $F/\delta u$ provides a direct indication for the force precision when $\sqrt{\Delta \delta u^2}$ has been determined.
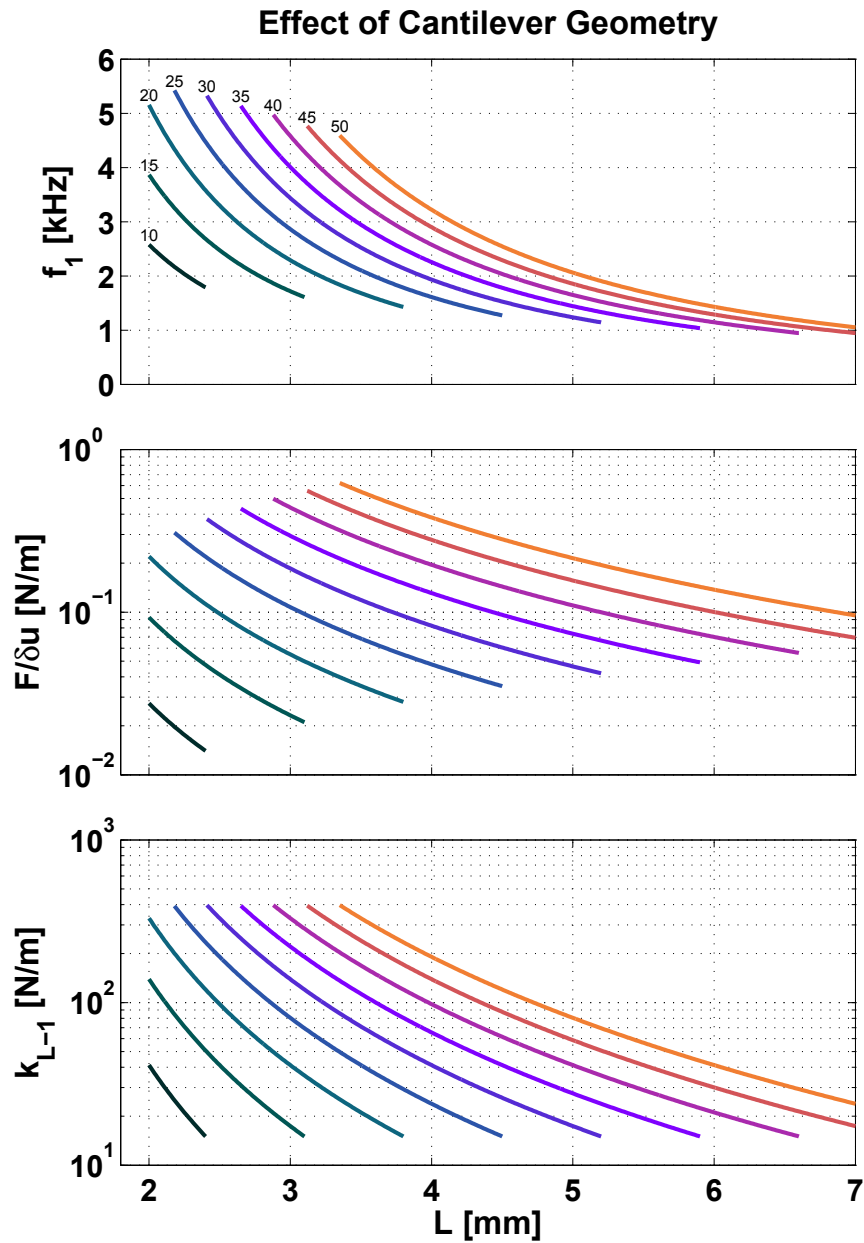
Figure 2.26: The effect of cantilever geometry on resonance frequency, $F/\delta u$, and spring constant at 1mm from the free end. A compromise needs to be made between high f_1, and low $F/\delta u$ and k_{L-1}. The different curves correspond to thickness between 10 and 50 microns. Assumptions: $w = 1.5$mm, $E = 110$GPa, and $\rho = 2.7$.

Chapter 3

Sample Preparation

The studies described in this thesis were performed using gold coated mica cantilevers. The cantilevers were cut from commercial sheets of mica and coated on both sides by thermal evaporation of gold. After anchoring to supporting clips and mechanical calibration, the samples were transferred in the FIM/STM/AFM system. They were then sputtered and annealed to produce large clean terraces of (111) oriented gold. The chemical cleanliness of the samples was then examined by Auger Electron Spectroscopy (AES). Finally, the samples were moved to the measurement chamber, where dynamic characterization was performed before combined STM/AFM.

3.1 Mica Preparation

3.1.1 Mica Structure

Mica was chosen as a substrate for gold deposition by thermal evaporation as it is widely documented that large terraces of monatomic Au(111) surfaces can be obtained in that way (see section 3.2.2). The term *mica* is used to refer to a group of naturally occurring sheet minerals that come in several forms, such as biotite, phlogopite, and muscovite. Principal constituents typically include silicon, aluminium, potassium and oxygen and hydrogen. The mica used in our investigations is Grade V-1 (highest quality) muscovite mica from SPI$^{\mathrm{TM}}$, and comes in 15mm × 15mm square sheets of 150μm thickness[1].

Muscovite mica is described by the chemical formula $KAl_2(Si_3Al)O_{10}(OH)_2$ [38], and is structured in negatively charged layers held together electrostatically by potas-

---

[1]Thinner muscovite mica from Axim$^{\mathrm{TM}}$ was also tested, and found to be of lesser quality. A large number of defects were found on the surface, which we attribute mostly to the cutting technique.

sium ions, as illustrated in figure 3.1. Mica can be readily cleaved at potassium layers, thus exposing a clean hexagonally structured top layer that can be atomically flat on a macroscopic scale.
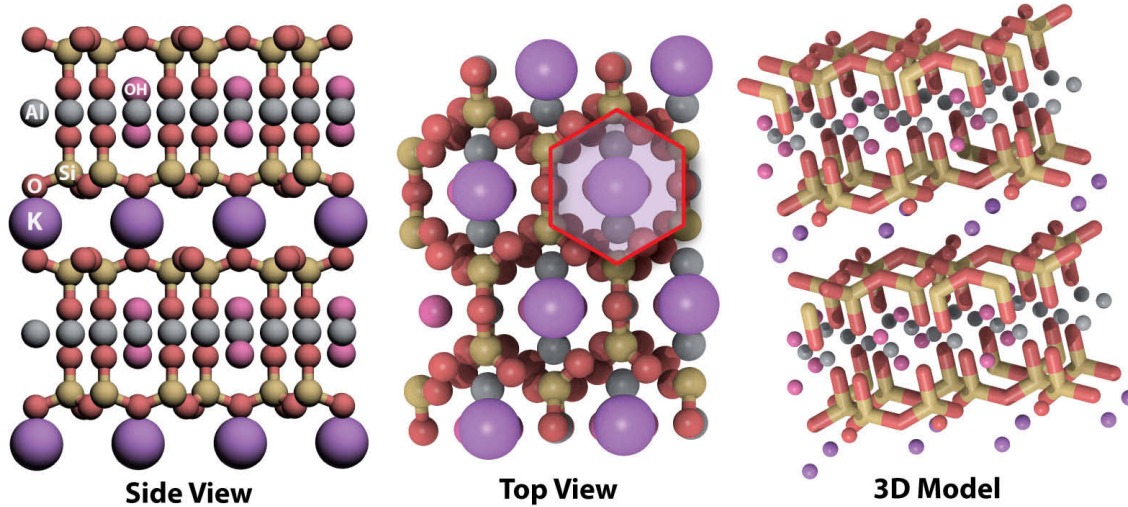
Figure 3.1: Three dimensional illustration of a muscovite mica structure, rendered by QuteMol [39]. In other forms of mica, K can be replaced by Na or Ca, and Al can be replaced by Mg or Fe among others. OH groups are sometimes replaced by F.

3.1.2   Cutting and Cleaving

The first step in preparing the mica substrate is the cleaving of the 150μm mica squares into 3-4 squares, each of which has a thickness of approximately 30μm-70μm. This is done using a razor blade cleaving technique recommended by SPI$^{\text{TM}}$, which consists of pressing the corner of a mica square laterally, with a razor blade held in parallel with the mica. Adding a drop of water in between the sheets as they dissociate makes the process easier, as the potassium ions in the surface dissolve, leaving two negatively charged surfaces that are easier to separate. The new mica squares are then conserved in ethanol until they are ready to be cut.

The cutting technique used is critical to ensure the success of our sample preparation. From the point of view of preparing a cantilever beam that can be used for force detection, we require that no inter-layer air bubbles be present. This is crucial, as it is necessary to have an optically flat backside for the reflection of the interferometer laser beams (see section 2.3). If the back surface is warped, the two reflected beams will not be parallel and will not recombine properly in the calcite crystal, thus causing a catastrophic loss in interferometer signal which can render the sample un-

usable. In addition, in order for the cantilever to behave as a proper Euler-Bernoulli beam and thus be accurately described by the equations developed in section 2.3, its geometry must be as close to a rectangular beam as possible. Imperfect geometry will also cause unwanted effects in the cantilever's dynamic behaviour. Lastly, large surface defects will affect the quality of the gold film and may also hamper STM operation.
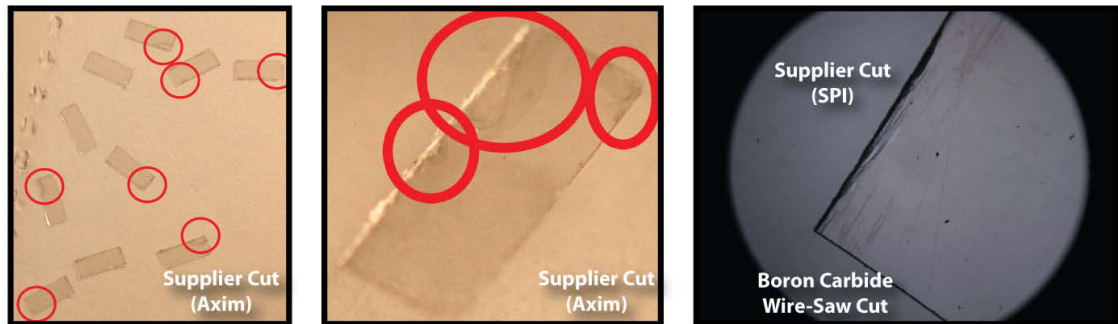
Figure 3.2: Illustration of the effect of various mica cutting techniques on the sample quality. On the left side and in the middle, small $2\text{mm} \times 5\text{mm} \times 25\mu\text{m}$ supplier (Axim$^{\text{TM}}$) cut mica samples. On the right side, SPI$^{\text{TM}}$ cutting on the top edge and diamond wire saw cutting on the bottom edge.

Several cutting techniques were explored. Cutting with a diamond disc saw was not successful, as samples were severely damaged during the procedure. Cutting with an office paper cutter tends to cause excessive bending of the samples. Cutting with a razor blade while pressing with a glass slide is better, and can work on thin samples of $25\mu\text{m}$ and less. Samples were also custom ordered from Axim$^{\text{TM}}$ at a predetermined optimal size for measurement, but the die cutting techniques employed were inadequate for our samples (Fig. 3.2). The undisclosed proprietary cutting techniques used by SPI$^{\text{TM}}$ are of better quality, but custom orders are relatively expensive. Other cutting techniques exist, such as hot platinum wire cutting [40], but they are relatively complicated to carry out and can produce large amounts of particulate contaminants [41].

The final, and most successful solution, was to use a wire saw using boron carbide powder (Model 810 from South Beach Technology Inc.). Before they are cut, the thin mica squares are attached with wax to a holder which is placed in the wire saw apparatus. Strips of $\sim 1-2mm$ width are then cut, and the wax is later dissolved in acetone, followed by a thorough cleaning in ethanol, and drying with nitrogen. The samples thus prepared exhibit very little edge deformation, as illustrated in figure 3.2.

Prior to gold evaporation, the mica strips were freshly cleaved using an adhesive tape cleaving technique. Adhesive tape is applied to the topmost layer of the mica strip and gentle removal will typically remove a 5μm slab of mica.

3.2 Gold Surface Preparation

The next step in the sample preparation is the deposition of a gold surface. Gold was chosen because of its highly inert behaviour and because it is widely documented, in terms of deposition techniques, surface science, and electronic and nanomechanical behaviour.

3.2.1 Au(111) Surface

The crystal structure of gold is Face Centred Cubic[2] (FCC), and the Au(111) surface is therefore close-packed, with a succession of layers in the order ABCABC. A diagram of an FCC(111) surface is shown in figure 3.3.
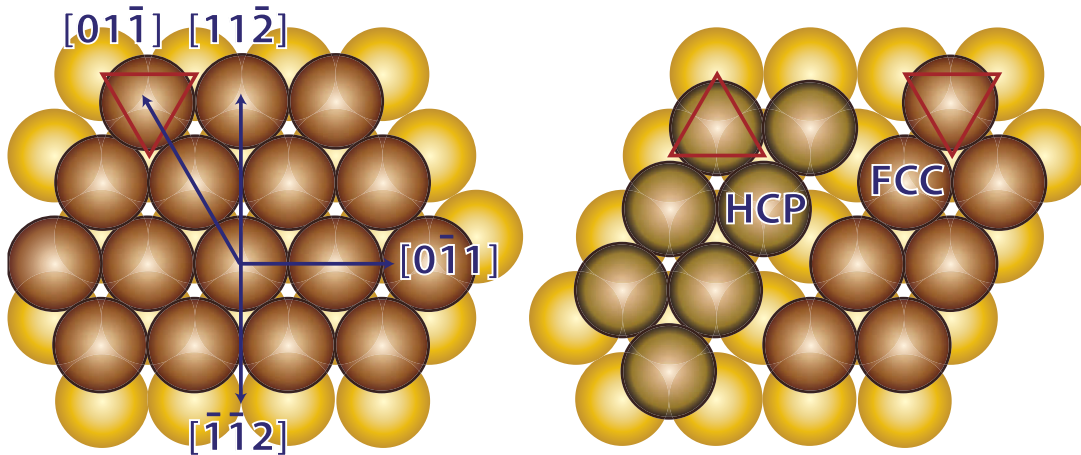
Figure 3.3: Schematics of the Au(111) surface of gold, with important crystal directions. FCC sites correspond to positions between three underlying atoms forming a downward pointing triangle. Hexagonal Close-Packed (HCP) sites correspond to underlying atoms forming an upward pointing triangle.

The previous description is accurate for gold below the topmost Au(111) layer. At the surface level however, gold atoms rearrange in a Au(111)$22 \times \sqrt{3}$ reconstruction in order to minimise energy. The effect is that in a very specific arrangement, certain atoms on the surface occupy Hexagonally Close-Packed (HCP) sites (see figures 3.3

---

[2]For details on common crystal structures, the reader is referred to the textbook by Ashcroft and Mermin [4].

and 3.4). Coupled with the three-fold symmetry of the substrate, this leads to a very recognisable herringbone pattern [42, 43].
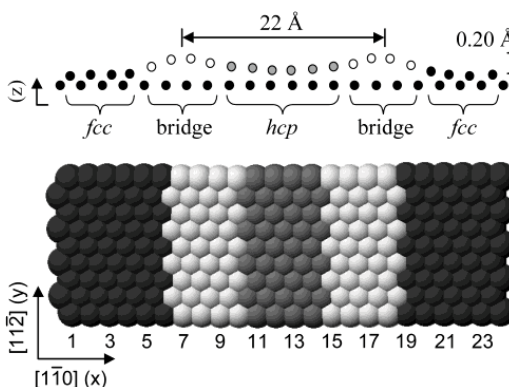
Figure 3.4: Model of the Au(111)$22 \times \sqrt{3}$ reconstruction. Reprinted with permission from [44]. Copyright American Chemical Society 2002.

3.2.2 Evaporation and Treatment

The method chosen for the deposition of a Au(111) surface is thermal evaporation. Numerous studies have been carried out by various authors to investigate the preparation of gold films on mica. Results published in the 1970's by Reichelt and Lutz found that while room temperature deposited gold on mica was polycrystalline, heating of the substrate favoured the growth of (111) oriented gold [45]. In 1987, Hallmark *et al.* observed atomic corrugation on Au(111) evaporated on mica preheated at $300^{o}C$ [46]. Long and flat Au(111) textured terraces of several hundreds of nanometres were also observed, separated by monatomic steps. Other studies have come to very similar conclusions [47–52]. These studies point to evaporation as a preferred technique to obtain these surfaces, and highlight the importance both of preheating the mica substrate and annealing the evaporated sample.

Evaporation Procedure

After fresh cleaving, rectangular mica strips are mounted on a sample heater and inserted in a dedicated thermal evaporator. The mica strips are then heated in high vacuum ($\sim 10^{-7}$mbar) at $400^{o}C$ for a minimum of 8 hours, after a temperature ramp on the order of 1-2 degree per minute. As the temperature of the mica strips is increased, when a sufficiently low pressure is reached, high purity gold pellets are pre-melted to allow potential contaminants to be degassed. The pellets are heated by running a very high current through an alumina coated molybdenum boat which

hosts the gold pellet. A second boat hosting titanium pellets allows the deposition of an adhesive layer when needed, for substrates such as glass. In the case of mica, an adhesive layer is not required, and titanium is simply used to enhance pumping by evaporating a layer of *getter* gas on the chamber surface while the samples are protected by a shutter.

After sufficient baking at 400^oC, gold is evaporated at a rate of $1 - 1.6 \mathring{A}/S$. The rate is monitored by a quartz crystal thickness monitor, and the evaporation is stopped when a 100nm layer has been deposited. The samples are then annealed at 400^oC for 1-3 hours and cooled by a temperature ramp down at a rate of 1-2 degree per minute. When the samples have reached room temperature, they are removed, cleaved, flipped, and reinserted in the evaporator, and the heating and evaporating procedures are repeated, thus leading to double sided gold coatings. This is done both to ensure a highly reflective layer on the backside which enhances the interferometer's performance, and to mitigate drift that could be caused by two parallel layers having different coefficients of thermal expansion, and subjected to inevitable temperature fluctuations.

Sputtering and Annealing

After evaporation, the samples are anchored onto molybdenum clips, as described in section 3.3.1. Their spring constant is then calibrated ex-situ (see section 3.3.1), and chosen samples are inserted in Ultra High Vacuum (UHV).

A quick analysis of the chemical composition of the surface by Auger Electron Spectroscopy (AES) (see section 3.2.3) confirms that untreated samples obtained as described earlier contain a substantial level of carbon contamination, in accordance with observations by Hallmark *et al.* [46].

In order to remove surface contaminants, the gold surface is bombarded by a 1keV beam of neon ions for 5-12mins. The beam is produced by an Ion Bombardment Gun after the preparation chamber has been filled by $\sim 10^{-5}$mbar of 99.95% purity neon, as monitored by a Residual Gas Analyser (RGA). Neon was chosen due to its low mass, which allows for gentle and controlled sputtering. The sputter yield of Ne^+ on gold (i.e. the number of substrate atoms sputtered off per incident ion) is indeed close to 1 [53]. For more information on theoretical and experimental aspects of sputtering, the reader is encouraged to refer to the comprehensive text by Behrish [53]. Finally, initial and advanced stages of sputtering on the Au(111) surface have been investigated by Lang *et al.* [54], and Michely *et al.* [55], respectively.

While the Au(111) surface is likely clean after sputtering, it is covered with sput-

tering pits and vacancy islands [54, 55], which are not conducive to the measurements we are seeking to perform. For this reason, the sample is annealed in UHV at a minimum temperature of 530^oC, for 45-60 minutes. While annealing will heal the surface [56], it also has a tendency to enhance the segregation of contaminants in the substrate towards the surface. For that reason, a new carbon peak on AES will often be observed on samples that had shown clean gold spectra after sputtering. When this is the case, an additional cycle of sputtering/annealing is required. Before the sample is moved to the measurement chamber, it is chemically characterized one last time by AES to ensure adequate cleanliness.

3.2.3 Chemical Characterization

In order to ensure that our surfaces are free of contamination, we characterize them by AES, which is based on the Auger process, first reported by Pierre Auger in the 1920's.

Auger Process
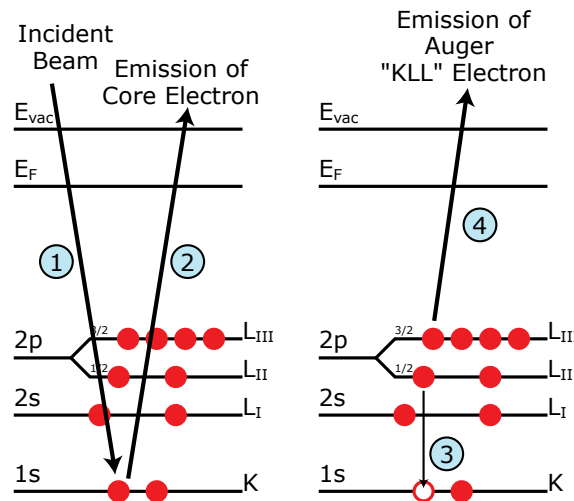
Figure 3.5: Auger process.

The Auger process is a four step process that can be described by the four following steps, as illustrated in figure 3.5. During the first step, an electron beam with an energy on the order of 1keV is directed at the sample. The incident beam, which can be replaced with an X-ray beam, excites and removes an electron from a core atomic level, thus leaving a hole behind. During step three, an electron from a higher level releases enough energy to fill the hole left after step two. This released energy is

enough to excite a third electron at a higher level, which leaves the atom with a very specific kinetic energy that depends only on the three levels involved and is therefore a signature of the atom in question. If the binding energies of the three levels involved are denoted with indices 1,2,3, with $E_1 > E_2 > E_3 > 0$, the kinetic energy of the Auger electron is given by equation (3.1). The particular Auger process that leads to a specific energy is labelled by concatenating the letters that correspond to the three levels involved. For instance, in the example of figure 3.5, the process illustrated is a KLL process. When one of the levels involved is a valence level, the letter V is used. Reference values of these energies are given in the *Handbook of Auger Electron Spectroscopy* [57], and shown in figure 3.6. Finally, it should be noted that the Auger energies can vary slightly when the atoms are chemically bound to other species.

$$E_{123} = E_1 - E_2 - E_3 \qquad (3.1)$$

Auger Electron Spectroscopy

AES is a ubiquitous surface analysis technique, and has been documented extensively. For a quick introduction to the topic, a number of surface science books can prove useful [59–61]. More in depth technique descriptions and reference material can be found in references [57, 62].

Figure 3.7 illustrates the interaction between the incident beam and the surface. The volume probed by the beam is on the order of $1\mu m^3$, a number of particles are emitted from the sample that can be detected by dedicated equipment. A number of electrons are elastically reflected and upon detection contribute to form an elastic peak used for instrument calibration and tuning. A large number of secondary electrons are also detected, which are extracted from the sample and have a broad range of energies which contribute to form a continuous background in the acquired spectra. Also contributing to the background are Auger electrons that have lost their original energy by scattering as they were leaving the samples. Upon differentiation with respect to incoming beam energy, this broad background vanishes, and true Auger peaks emerge that can be identified using figure 3.6 and reference spectra [57].

One important point to highlight is that AES is a surface sensitive technique. This is due to the fact that the inelastic mean free path for electrons for all elements follows a universal curve [63, 64]. As shown in figure 3.7,this curve shows that the mean free path has a minimum of a few monolayers when the electron energy is in the range of 10-500eV, which is precisely the range of energies primarily explored by
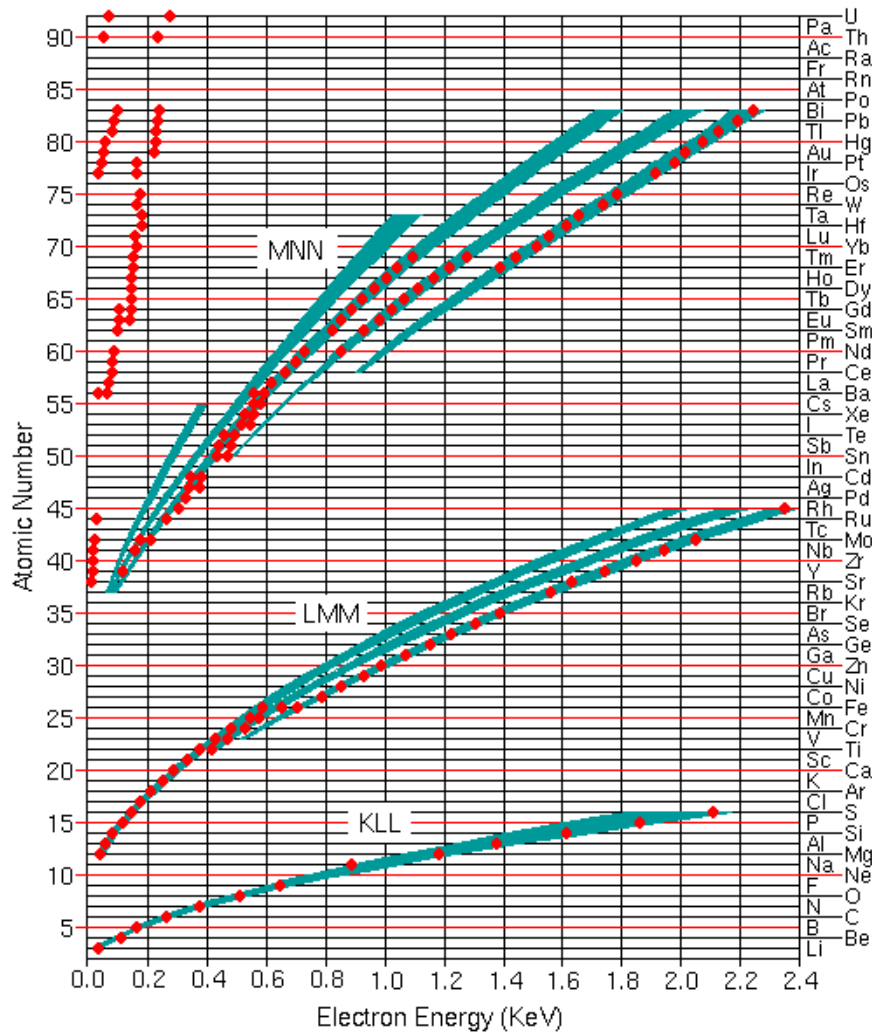
Figure 3.6: AES energies. Courtesy of [58].

AES. For an Auger electron to contribute to an Auger peak, it must have retained all of its original energy, and only electrons emitted within a mean free path distance of the surface are susceptible to fulfill that criterion. For gold, experimental results show that for energies between 80-140eV, the mean free path is essentially one monolayer [65]. Above 1keV, the mean free path for gold rises to close to $20\mathring{A}$ [59], and therefore, peaks that are occur at energies above 1keV will attest to the presence of gold within 10 monolayers or more of the surface.

Experimental Considerations

AES was performed with a commercial system (Varian 981-2707), which is based on a Cylindrical Mirror Analyser. Differential spectra are acquired through a lock-

(a) Auger Probing Volume

(b) Universal Curve of Inelastic Electron Mean Free Path. λ_m is in monolayers. Reprinted with permission from [63]. Copyright John Wiley and Sons 1979.
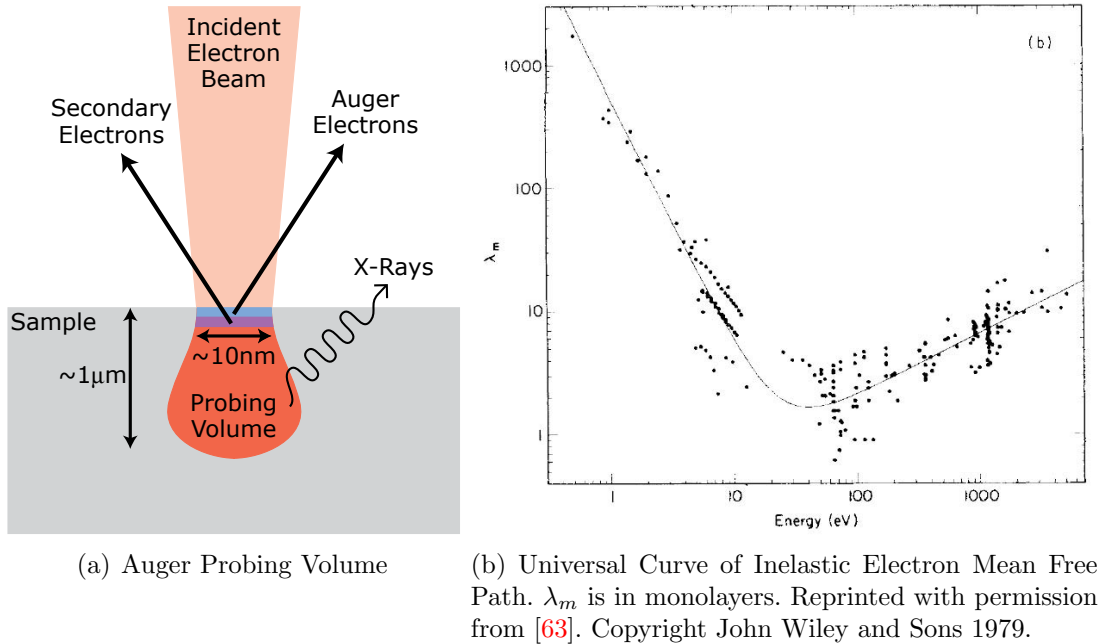
Figure 3.7: AES probing volume and electron mean free path.

in amplifier operated at 17kHz. Scanning AES was also implemented by using the STM/AFM acquisition system, by applying voltages to pairs of plates that deflect the incident beam in a raster fashion.

In order to facilitate comparisons between various samples and between several states of the same sample, spectrum acquisition parameters were standardized to acquire three types of spectra. E-type spectra are elastic scattering peaks and are obtained to calibrate and tune the system. A-type spectra correspond to a 2keV incident beam and acquire in the 100-600eV range, which is a wide enough range to identify gold on the topmost layer, as well as contaminant such as carbon, oxygen, and sulfur. B-type spectra are obtained with a 5keV incident beam and acquire in the 100-2100eV range. This allows identification of species that do not have a peak in the 100-600eV range, as well as the detection of gold under a potentially contaminated surface, which will extinguish the gold peaks in the low energy range, due to the mean free path considerations outlined earlier.

The most prominent contaminant in our investigations was carbon. When it is present, carbon is highly visible in the acquired spectra, not only because the low energy gold peaks are extinguished, but also because the relative sensitivity of AES to C is over 10 times higher than for Au [57]. Carbon free AES spectra are therefore extremely good indicators that the surface is clean.
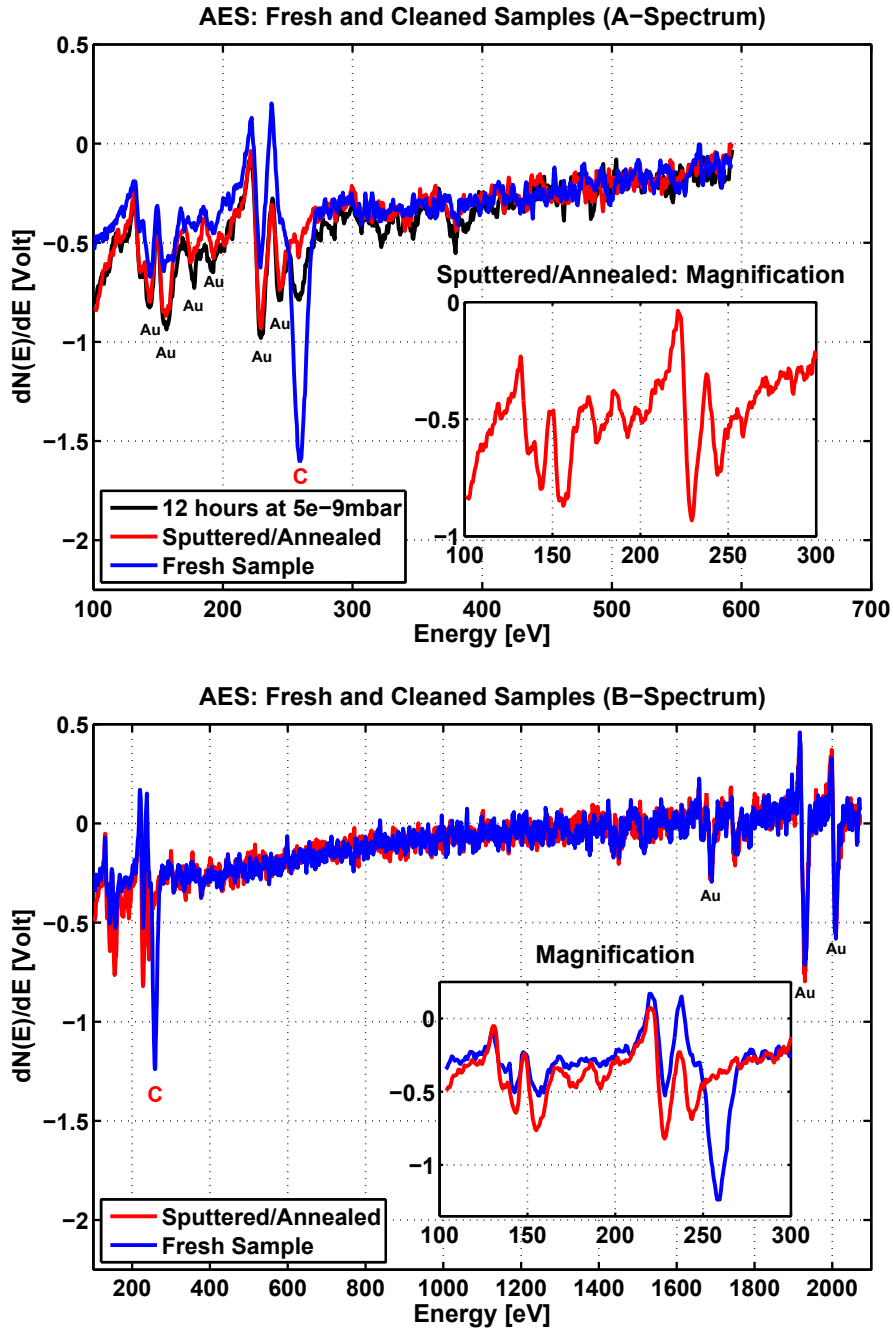
Figure 3.8: AES spectra of our gold surfaces. Fresh samples (i.e. directly after evaporation) present a carbon peak which disappears after a few cycles of sputtering/annealing. The top graph also shows a trace corresponding to a clean sample left in the preparation chamber at a pressure $\sim 5 \cdot 10^{-9}$mbar for 12 hours, which highlights the onset of carbon contamination. Finally, the bottom graph shows large gold peaks in the vicinity of 2keV for both the fresh (contaminated) and the clean samples, highlighting the presence of gold within the top 10-20 monolayers.

AES measurements allow us to verify the cleanliness of samples every time before they are transferred into the measurement chamber. The insights gathered after extended use of the system are very instructive. As illustrated in figure 3.8, and consistent with Hallmark's observations [46], freshly evaporated samples do tend to present significant carbon contamination which is easily removed by sputtering and annealing. Figure 3.8 also shows that a sample left in the preparation chamber after annealing (at which point it sustains a pressure above $5 \cdot 10^{-9}$mbar) will start to develop a noticeable carbon contamination layer after 12 hours. For this reason, after preparation, the samples are moved to the measurement chamber immediately where a titanium sublimation pump, a cryogenic pump, and a magnetically levitated turbomolecular pump allow a base pressure on the order of $5 \cdot 10^{-11}$mbar. Given that the impingement rate of contaminant is proportional to the pressure [62], this means that the sample can be expected to reach the same level of contamination after a duration 100 times longer. In practice, our samples show no sign of contamination after 2-3 days.

AES of surfaces prepared months prior and left in Petri dishes showed substantial contamination by carbon and oxygen, and more rarely by sulfur, at times when thiol molecules were used by colleagues in the laboratory. A few cycles of sputtering and annealing will easily restore the surface, showing that our samples have a very long shelf life.

A final note of interest is that peak energies measured by our AES system tend to differ from reference spectra by a few electronvolts. The energies measured are also different depending on the type of spectrum acquired, although they are consistent for each type of spectrum (A and B). Similar phenomena are reported in the literature, and attributed to changes in the vacuum energy level of the detector, which vary depending on the detector and its state [66]. In order to account for these differences, a reference database of measured peaks was collected for each spectrum type for common elements relevant to our investigations (C, O, Au, Si, S, etc.) and a linear fit was performed that successfully converted all measured values to standard reference values.

3.2.4   STM Characterization

The topography of prepared surfaces was characterized by Scanning Tunnelling Microscopy (STM), which showed results consistent with those reported in the literature as described previously. Figure 3.9 shows STM images acquired with a nanoscope STM in air. Large atomic terraces and monatomic steps are easily visible.

|2 micron x 2 micron | 330nm x 330 nm | 200nm x 200nm|
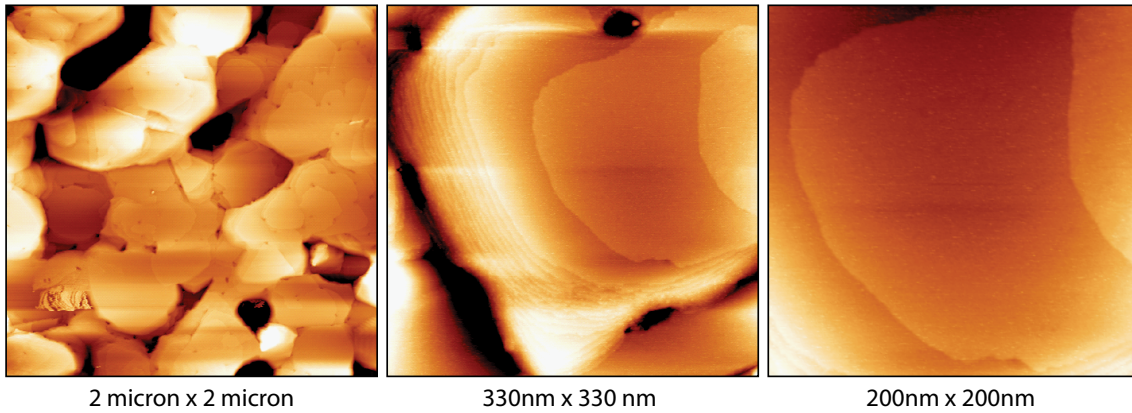
Figure 3.9: STM images of a freshly evaporated gold sample. The images were obtained in air with a nanoscope instrument.

| 150 nm x 150 nm | 100 nm x 100 nm | 100 nm x 100 nm |
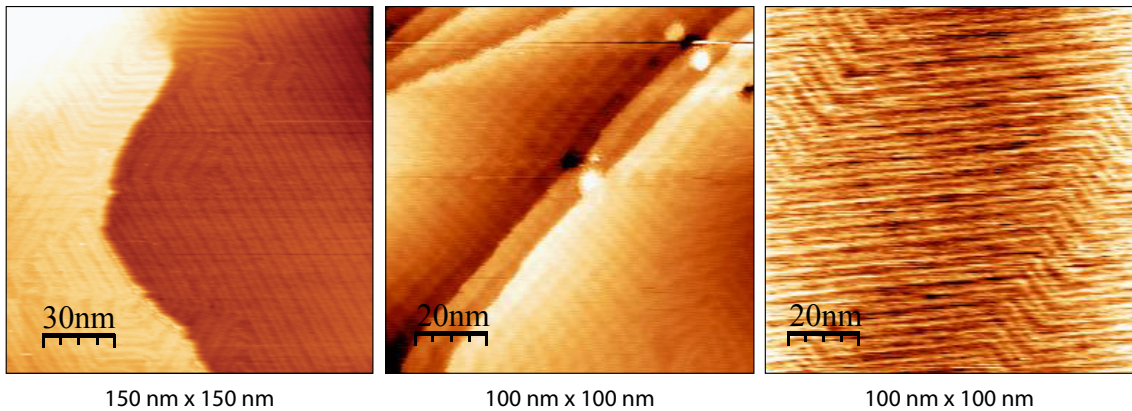
Figure 3.10: STM images of three different Au(111) samples in UHV, imaged with three different tips. The differences in image quality can be understood in terms of the tips employed. The holes and islands imaged in the middle scan were produced by nanoindentation.

After sputtering and annealing in UHV, large terraces can still be observed, and the herringbone reconstruction is ubiquitous. Figure 3.10 shows three different samples imaged with three different tips, which all present a herringbone pattern. The images confirm that we are imaging (111) oriented surfaces of gold, and also give an indication of the vertical sensitivity of our instrument, as the 20pm height difference between bridge and HCP sites (Fig. 3.4) can be resolved routinely. Imaging of the herringbone also provides a definite confirmation of our lateral calibration.

3.2.5   Other Gold Preparation Techniques

Several other gold preparation techniques have been documented in the literature which are reported to offer micron sized flat terraces: flame annealing [67], rapid

nitrogen annealing [68], and template-stripping [69, 70]. While they were not tested on our system, other group members have had success with some of these techniques.

3.3   Preparation and Characterization of Cantilever Beams

Obtaining a suitable surface for measurement is an essential step towards carrying out our experiments. However, a great level of care in preparing and characterising samples is also necessary to ensure both that the tip–sample junction is stable, and that quantitative force measurement is possible.

3.3.1   Cantilever Preparation and Calibration

Cantilever Support

After mica strips of a definite thickness and width have been prepared and coated with gold on both sides, they are cut to an appropriate length in order to optimise mechanical behaviour. Before they are usable, the strips need to be attached to our sample holder. The requirements on the support system are very stringent: the anchoring system needs to satisfy the assumptions made in section 2.2.1 that $u(0) = 0$ and $u'(0) = 0$. It further needs to be rigid and symmetric to ensure proper mechanical behaviour, to be UHV and high temperature compatible (our annealing procedures reach temperatures close to 550^oC), and to provide an electrical connection between the sample holder and the cantilever beam to allow STM operation.

Glue/Cement Anchoring

Excessive Mechanical Clamping

Mechanical Clamping

Asymmetric Mechanical Clamping
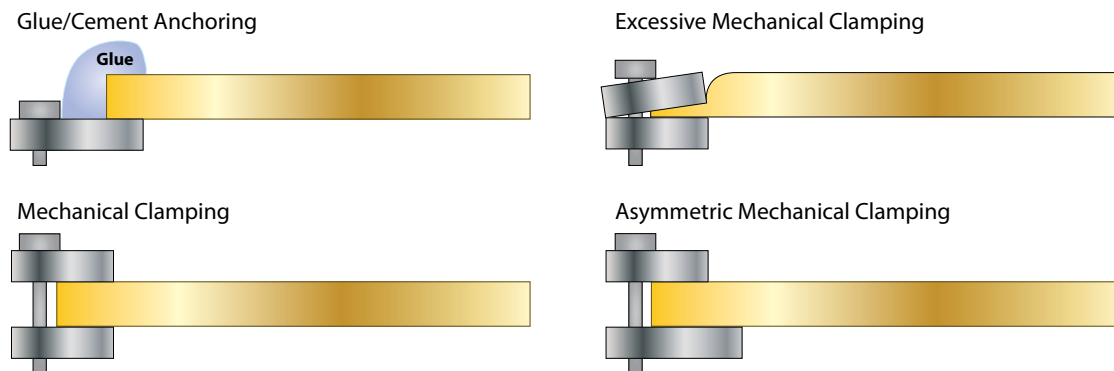
Figure 3.11: Mechanical and glue based anchoring.

In previous studies on the FIM/STM/AFM system, the cantilever was clamped mechanically to a sample holder. While this technique has the advantage of simplicity

and is correct in principle, it is prone to error. As highlighted in figure 3.11, two main issues can arise. The first issue is that of asymmetric clamping, and occurs when the top and bottom clips are not identical. In that case, the deflection equations derived in section 2.3 are no longer accurate, as the length of the cantilever is different at the top and the bottom. This leads to different spring constants for upward and downward bending, and makes the vibrations more complex, in a diving-board like fashion.

The second issue that can arise with mechanical clamping has a far more severe consequence as it can preclude force measurement altogether. It occurs when the clamping strength is excessive enough to deform the cantilever, often beyond the breaking point, as illustrated in figure 3.12.
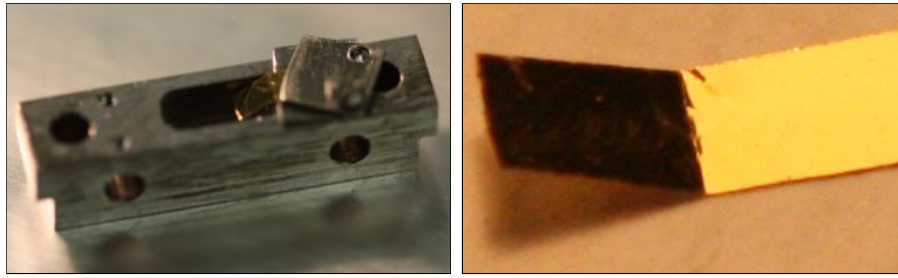
Figure 3.12: Release of a mechanically clamped cantilever beam showing the effect of excessively strong clamping.

Damage to cantilevers such as that depicted in figure 3.12 is not easily visible when the cantilever is mounted, as the latter may appear straight and even produce a reasonable value during the spring constant calibration procedure. This is likely due to the large scale $50 - 250 \mu$m downward bending imposed by our spring constant calibration procedure, which will eventually find an effective clamping point at the edge of the bottom clip. However, as the bending that we aim to measure during our experiments are at least four orders of magnitude smaller, they are insufficient to reach a consistent clamping point and therefore yield a spring constant measurement that is also valid at the nanoscale. When pushed down at that scale, the cantilever is more likely to rotate about its mechanical clamping point than to deform according to the *Static Deflection Model* described in section 2.2.1. In that case, its motion is more accurately described by a *Cantilever Rotation Model* (CRM), and can be approximated as remaining straight. In that case, using the notation of section 2.3.1, the slope of the beam is equal to u_t/t, or equivalently, $\delta u/d_b$, which leads to equation (3.2).

$$\frac{\delta u_{\text{CRM}}}{u_t} \quad = \quad \frac{d_b}{t} \quad (3.2)$$

In order to test whether the cantilever deflects or rotates, one can apply a very large force with a flat tip *in situ*. After the gold surface deforms entirely, there comes a point when the cantilever motion is expected to match the tip motion exactly. However, at the beginning of this project, the tip was found to move almost twice as much as the cantilever. A number of possible explanations (beam positions, torsional motion of the cantilever when the tip was not positioned at the centre, etc.) were explored and investigated, which allowed us to reduce the discrepancy to a factor of 1.5, but a one-to-one matching was never attained on a mechanically clamped cantilever. The explanation of the 1.5 factor comes immediately when looking at equations (3.2) and (2.46)[3]: If the cantilever rotates instead of deflecting according to the Static Deflection Model, δu^{CRM} as measured by the interferometer is 2/3 of the value that would be obtained if the clamping had been adequate. For that reason, the deflection measured when equation (2.46) is applied is false, and 1.5 times smaller than reality. It was therefore hypothesised that this factor would disappear when anchoring the cantilever properly with glue. This hypothesis has since been tested with 100% success for approximately a dozen (mica or silicon) cantilevers. Mechanical clamping was henceforth abandoned in favour of gluing.

The samples used for the investigations described in this thesis were attached by gluing the samples to molybdenum clips. Several types of glue were tested. An inorganic zirconia/silica based cement (Aron Ceramic E from Toagosei Co., LTD) that can sustain temperatures up to 1300^oC was tested. However, as the cement is non-conductive, an additional electrical contact is necessary. H20E from Epotek is a very convenient two component conducting epoxy. However, at temperatures higher than 300^oC, it begins to degrade dramatically, as confirmed by residual gas analysis in our preparation chamber. The solution that yielded the best results was a high purity silver paint from SPI$^{\text{TM}}$. Heating of samples at up to 550^oC in our preparation chamber and concurrent monitoring with a RGA showed that the glue degases very mildly and completely within 5-10mins. Subsequent sputtering and annealing, followed by Auger analysis showed contaminant-free gold spectra. For higher temperatures, the more expensive platinum paint from SPI$^{\text{TM}}$ may be used.

---

[3]Equation (2.46) states that $\dfrac{\delta u_{\text{SDM}}^{\max}}{u_t} = \dfrac{3}{2}\dfrac{d_b}{t}$

Spring Constant Calibration

After anchoring and before the sample is transferred to UHV, the spring constant of the cantilever is calibrated. As illustrated in figure 3.13, a rigid tip is pushed onto the sample, which rests on a microbalance. The tip is pushed down using a micrometric screw, which allows the tip to be moved by increments of 10μm. The force corresponding to that displacement is measured by converting the mass reading on the scale into a force. Taking 5-8 data points in a range of $10-80\mu$m or $50-350\mu$m in increments of 10 and 50 microns respectively, the slope of the force vs. deflection curve represents the spring constant at the tip position, k_t. The process is repeated 4-5 times and the average value is taken as the spring constant.
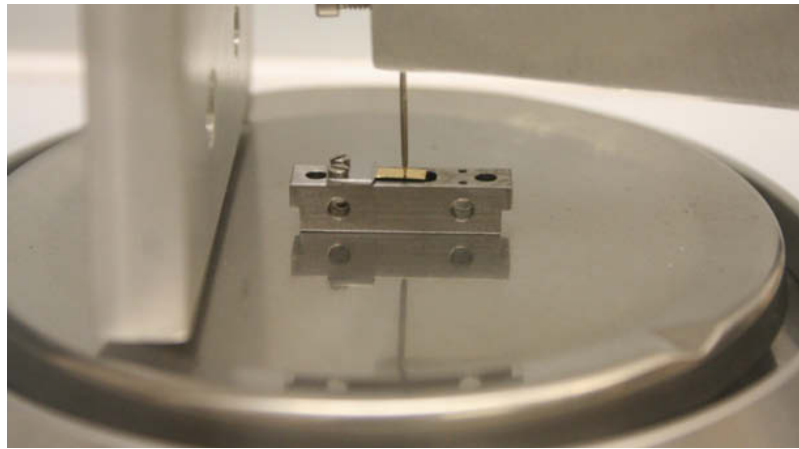
Figure 3.13: Microbalance based calibration

In figure 3.13, the spring constant is being measured in the middle of the cantilever. In the past, all positions of the tip during measurements would be recorded and spring constant calibrations would be carried out at each of these points after the sample was removed from UHV. For improved efficiency, we only measured the spring constant at the very end of the cantilever, k_0, and use equation (2.20), which states that $k_t = k_0 L^3/t^3$, to determine k_t from the tip position t and the full cantilever length L. This relation has tremendous practical value and is shown for reference in figure 3.14.

In UHV, a high resolution photograph with a 100mm macro lens attached to a Digital Single-Lens Reflex camera Canon EOS 450D, is acquired before each experiment. This allows the determination of t from L, which is measured prior to transferring the sample into UHV. This method allows quick and accurate determination of the spring constant immediately prior to measuring. It should also be noted that equation (2.20) was tested successfully by moving the tip/screw assem-

bly with an x–y stage to various locations for spring constant measurement. The results obtained confirmed the validity of equation (2.20), and provided the estimate $E_{\mathrm{mica}} \simeq 120\mathrm{GPa}$, also consistent with frequency measurements.
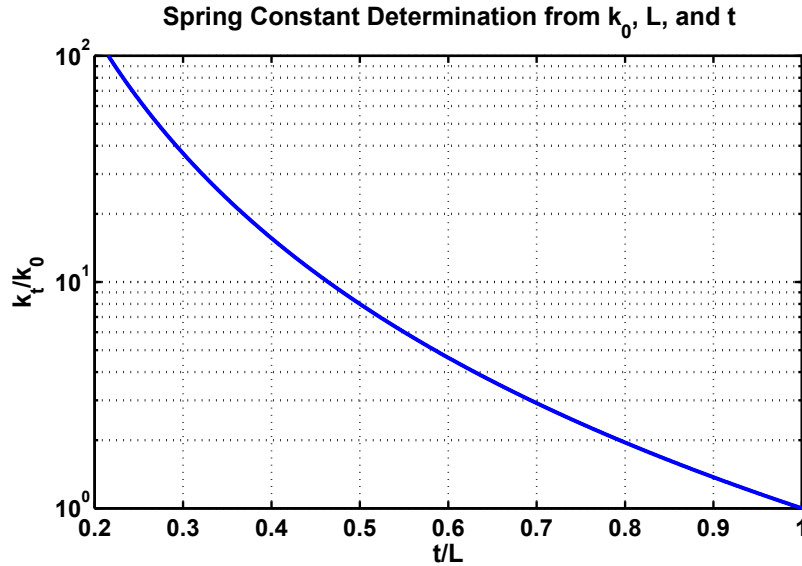
Figure 3.14: CB Calibration

3.3.2 Dynamic Characterization

The dynamic characterization of our cantilever beams relies primarily on the measurement of the dynamic response to both environmental vibrations and applied excitations. Our methodology and main results were introduced in sections 2.2.2 and 2.4. In the following, we outline some specific relevant issues of sample preparation and characterization.

Vibration Coupling and Changing Resonance

As mentioned in section 2.4.2, external vibrations can interact with a resonance of the cantilever and drive it in a manner that precludes all measurement and that can potentially cause damage to the tip. This makes quantitative assessment of the vibrations spectra particularly important prior to measurement in order to discard samples that are unsuitable. A close look at vibration spectra, such as that of figure 2.23, shows that the number of potential system frequencies to be avoided is surprisingly high. However, looking at spectra from the same cantilever on different days shows that specific vibrations can be present on particular days and undetectable on others, which means that a sample that appears unusable on a given day may

provide great results on a different day. This is particularly true when the resonance frequency of the sample changes a few tens of Hertz, which is observed regularly, particularly after annealing.
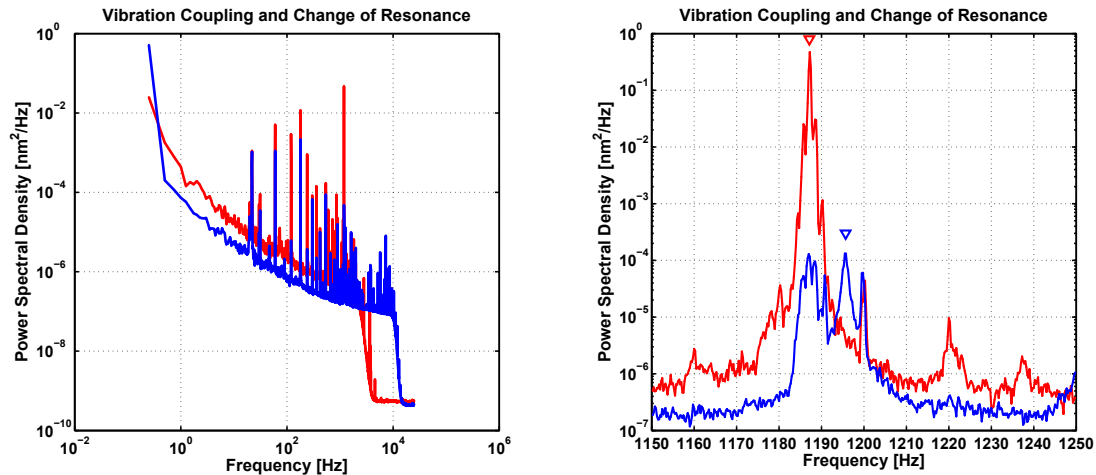
Figure 3.15: Vibration spectra of the same cantilever beam on different days. Vibrations are larger on the red traces. On the right, the first mode resonances are marked with triangles and differ by almost 10Hz, which is enough to make the red traces interact catastrophically with the excitations between 1180Hz and 1190Hz.

Figure 3.15 shows vibration spectra of the same cantilever beam taken on two different days, only three days apart. Although system vibrations are visible around 1185Hz on day 1, they do not interact catastrophically with the cantilever (blue trace). On day 2 (red trace), the cantilever resonance frequency has shifted down (after annealing) sufficiently to overlap substantially with the system frequency at 1185Hz, producing large amplitude excitation that make the obtention of a stable tunnelling gap impossible.

Evolution of Power Spectra

The previous paragraph highlighted the idea that a change in resonance frequency of a cantilever beam after annealing can have disastrous consequences on the tunnelling gap. However, a stable tunnelling gap is not a definitive criterion to characterize a cantilever as dynamically suitable. As the tip interacts with the sample, force gradients change, which changes the resonance frequency of the cantilever beam [20, 71]. These changes can be substantial in nanoindentation mode, when the tip is driven into the sample with loads upwards of 500nN, in which case the dynamic behaviour of the sample become closer to that of a doubly clamped cantilever, with

much higher resonance frequencies [72]. As the resonance frequency of the cantilever changes value, it can encounter large system vibrations which can cause it to enter into large scale oscillations that can preclude the resolution of subtle features, such as discrete plastic events.

In order to monitor changes in cantilever vibrations, the deflection signal can be acquired and cut into a number of segments, for which power spectra are computed and compared. The power spectra are then displayed in PSDE (Power Spectral Density Evolution) plots. One such plot is shown in figure 3.16.
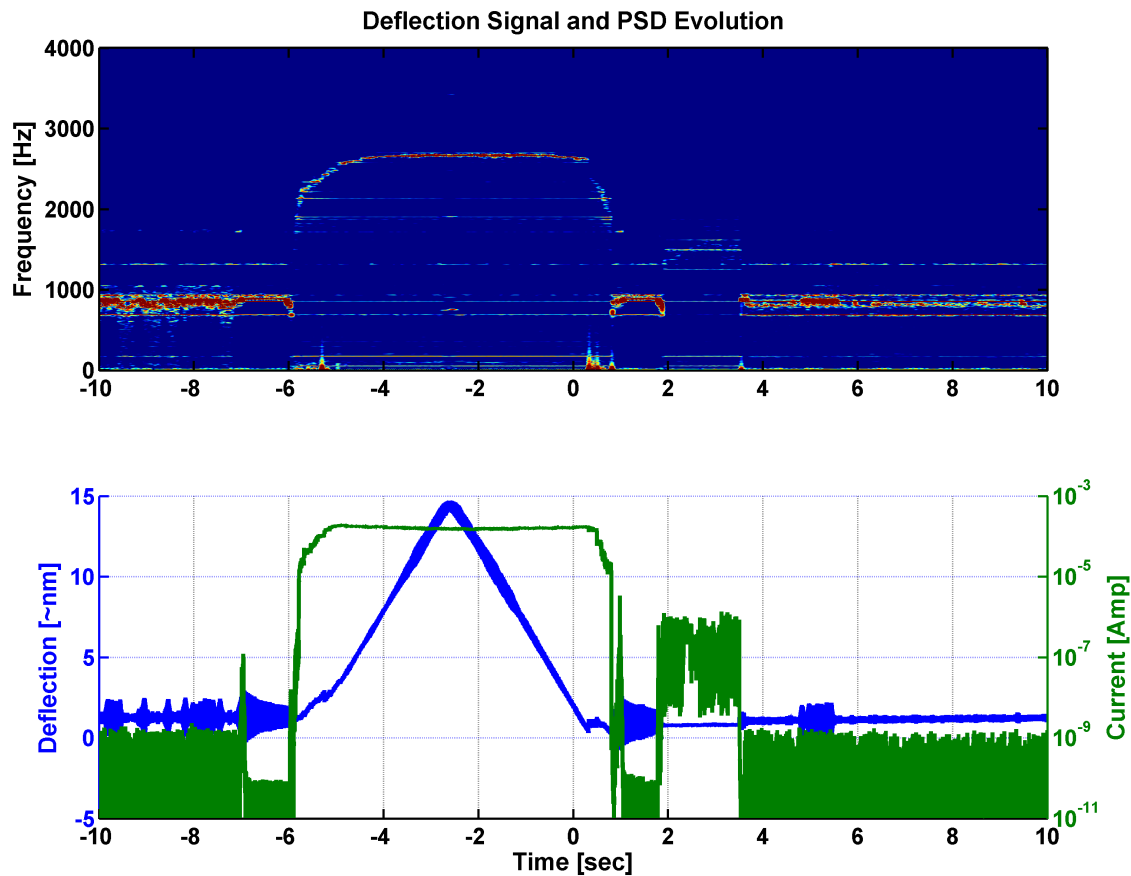
Figure 3.16: Power Spectral Density Evolution plot showing the evolution of the power spectrum of the deflection signal over time, during a nanoindentation experiment. The colour scale is logarithmic and adjusted to highlight features of interest. The current signal is acquired and displayed simultaneously.

In a PSDE plot, each vertical cross-section is a power spectrum displayed as a colour code. A number of insights can be extracted from the PSDE plot shown in figure 3.16. Before the indentation experiment starts, the resonance frequency of the cantilever is apparent just below 1kHz, with a number of system vibrations visible are horizontal lines. As the tip approaches the sample, an attractive interaction between

the tip and sample causes the sample resonance frequency to drop noticeably on the PSDE plot, whereas the attractive region is very difficult to observe in this case. As the tip is driven into the sample, the interaction becomes strongly repulsive and the resonance frequency increases, reaching system vibrations around 2kHz, after which point the cantilever remains mildly excited at those frequencies. During this nanoindentation cycle, several discrete events took place such as pop-ins and pop-outs (see section 5.1.2) which translate into noticeable $1/f$ noise in the low frequency region of the PSDE plot.

Another interesting feature of PSDE plots is that the variation of the resonance frequency looks strikingly similar to the current trace displayed below. While a detailed understanding of this phenomenon would require extensive investigations, the general trend can be intuitively understood in the following way: the resonance frequency is expected to vary like the junction stiffness, which is expected to be proportional to the contact radius, which is proportional to the square root of the contact area, itself being proportional to the current. It should be noted however that this reasoning is very approximate, particularly as the current is displayed in a logarithmic scale. Furthermore, it assumes that the deformation is strictly elastic, and that the electronic transport properties of the junction are not affected by the loading.

Q-Factor and Thermal Peak

In the example of figure 3.15, catastrophic vibrations of the cantilever were induced on a specific day which were very easy to identify. In other cases, the coupling of external vibrations to the system are more subtle effects. In order to establish whether the first mode of the cantilever is being excited by an external source other than the white excitation background caused by thermal fluctuations, one can perform a fit of the resonance peak, as suggested in section 2.2.2. This procedure allows the extraction of a Q-factor which can be used to plot the expected thermal peak.

Figure 3.17 shows a series of resonance peaks taken from the spectra of figure 2.23. The black bottom trace correspond to the expected thermal peak based on the Q-factor obtained through fits. The fits yielded a Q-factor of 2000, which is on the higher end of the range of values typically seen with the samples we use. The figure shows that when the vibration isolation systems are employed, for that cantilever, the vibration spectra are very close to the thermal limit. Also interesting to note is the peak slightly above 860Hz, which corresponds to the vibration of one of our turbomolecular pumps, which only shows up on the curves taken when the pump was in operation.
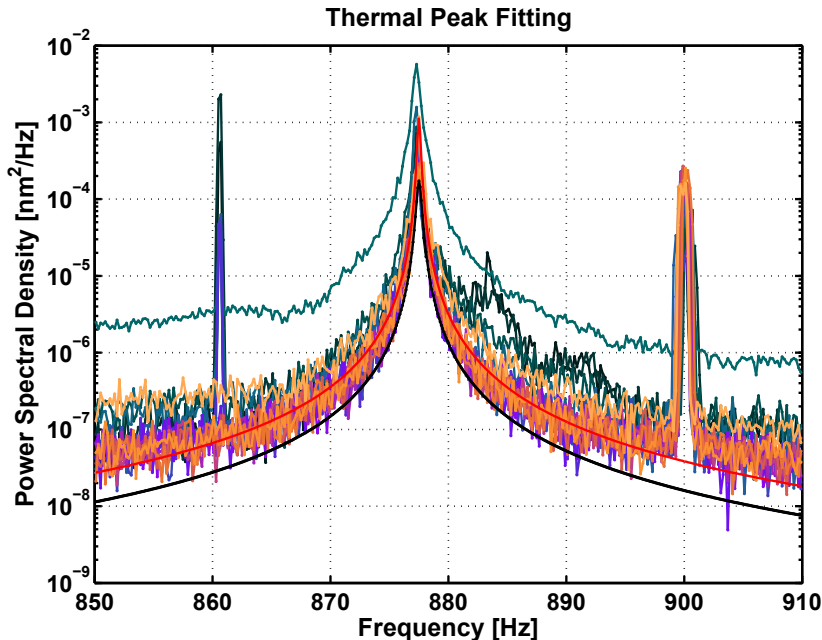
Figure 3.17: Thermal peak fitting of the first from the spectra displayed in figure 2.23. The bottom red trace is a fit to the last curve, and the black trace at the very bottom is the expected thermal peak.

3.4 Tip Preparation and Characterization

The preparation of our tips has been documented extensively by Lucier [16, 17] and Hagedorn [7]. The tips used during the experiments reported in this thesis are electrochemically etched from (111) oriented tungsten wire. Some tests were also carried out using similarly etched polycrystalline tungsten wire, which typically leads to (110) oriented tips. For the in-situ mechanical testing of cantilevers, cut Pt–Ir tips were also used at times.

The tips used for our experiments were annealed in UHV in our preparation chamber prior to being used, primarily to evaporate the tungsten oxide layer formed during etching. The tips were then characterized by Field Emission in order to obtain an estimate of their sharpness before they are transferred to the measurement chamber. This characterization by Field Emission is important as only very sharp tips with a tip radius below 10-15nm will be imageable in our Field Ion Microscopy (FIM) setup in its current form[4].

After the tip is transferred to the measurement chamber, it is imaged by FIM, as outlined briefly in section 2.1.2. In the present studies, FIM is used primarily to

---

[4]The limitation is due to the geometric constraints, combined with the fact that the coaxial wiring used for high voltage during FIM is the same used during tunnelling.

ensure that the structure observed is indeed that of W(111), to determine the overall structure of the tip, and to ensure that the tip has retained its general shape and integrity at the end of the experimental sequence.

The most important quantitative information relevant to our nanomechanical studies that is extracted from the FIM images is an approximate tip radius. In order to do so, a ring counting method is applied, as described in particular by Lucier [16].

Images of the tips are thus taken immediately before measurement, and when possible, after a sequence of experiments. In a minority of cases, the tip can suffer a crash that makes it impossible to image with our setup. When images can be obtained after a nanoindentation experiment (Fig. 3.18), tips are typically covered in adsorbates, which we interpret as gold adatoms taken from the sample. This is consistent with results from Fian *et al.* [73, 74].

As the voltage is increased, the adsorbates are field evaporated, eventually revealing the underlying W(111) surface. The radius of the tip measured with the ring counting method after experimentation remains the same, attesting to the integrity of the majority of our tips post-nanoindentation.
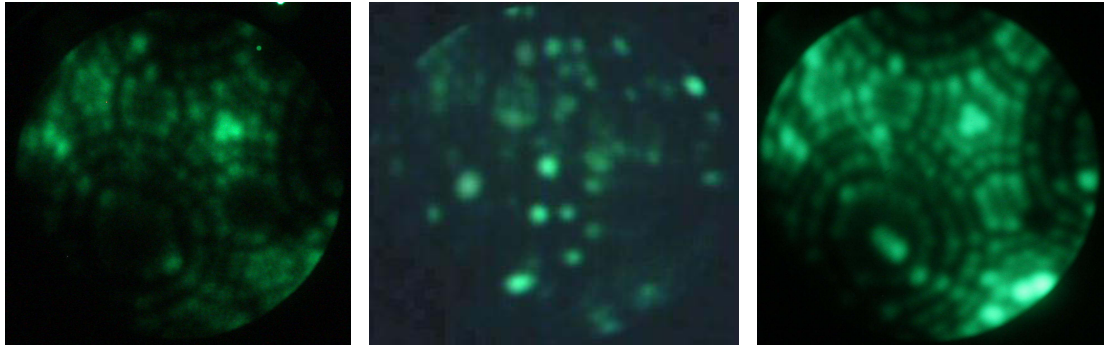
Figure 3.18: FIM images of the same tip immediately before nanoindentation, immediately after acquisition of a few hundred load-displacement curves, and after field evaporation of adsorbates. The image on the right shows that the tip structure was recovered. The lower quality of the middle image is due to the fact that a different camera was used to capture the dynamics of adsorbate field evaporation.

In rare cases, the tip may be damaged during a nanoindentation experiment. This is usually associated with noticeable and characteristic signatures in load-displacement curves. At times, tips remain imageable by FIM, in which case a substantial level of distortion is visible, as shown in the first image of figure 3.19. When the voltage is increased, substantial field evaporation occurs, and after a few minutes, a structure, such as that of the middle image, is exposed, which presents a visible crack. In some cases, continued field evaporation will eventually reveal a tip with a
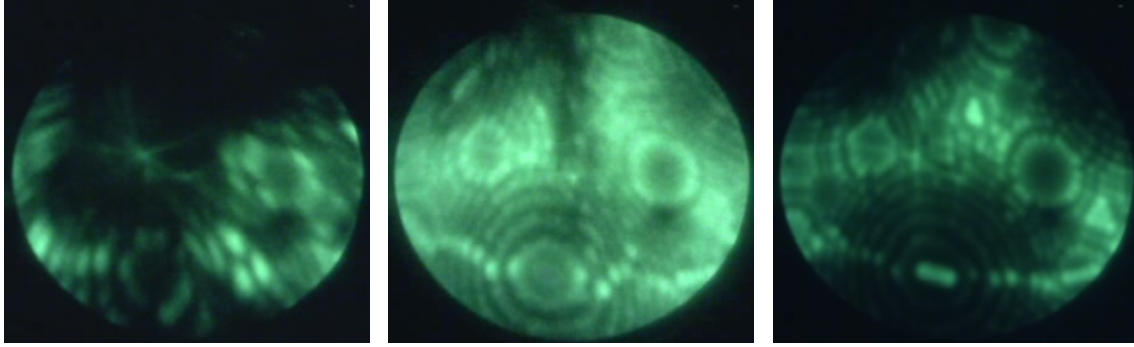
Figure 3.19: FIM imaging of a tip damaged during experiments.

much higher tip radius, and more subtle signs of damage.
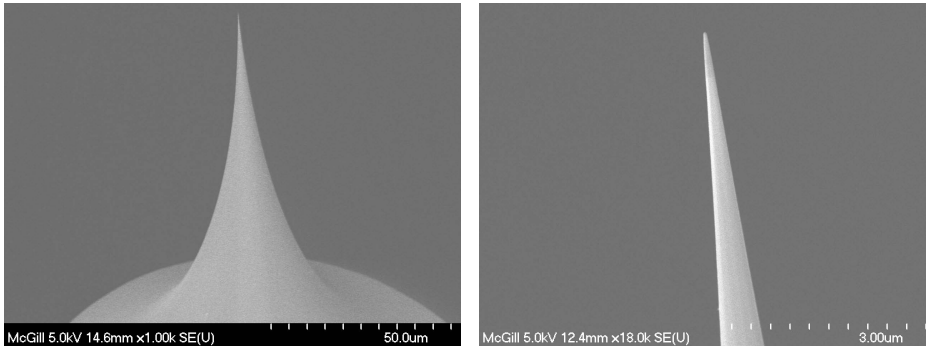
Figure 3.20: Scanning Electron Micrographs of (111) oriented tungsten tips. Images taken by T. Hagedorn.

The tungsten tips were also characterized by Scanning Electron Microscopy (SEM), as illustrated in figure 3.20. While SEM imaging for every tip is impractical in the absence of an in-situ SEM, results showed consistently that our tips can be described mesoscopically as rounded cones with a half-opening angle between 5-10 degrees, as illustrated in figure 3.21.
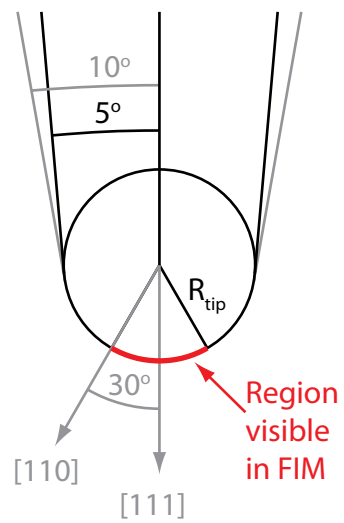
Figure 3.21: The mesoscopic shape of our tips can be described as a rounded cone of half opening angle between 5 and 10 degrees. Drawing courtesy of T. Hagedorn.

Chapter 4

Nanomechanics of Au(111): Investigations and Phenomenology

4.1   Nanoindentation of Au(111)

The nanomechanical properties of the Au(111) surface have been studied by a number of researchers, both experimentally and from a theoretical perspective. In doing so, an approach that has proven very powerful is that of nanoindentation. Important considerations relevant to the history, conduct, and interpretation of nanoindentation experiments are presented in chapter 5.

Simply put, indentation consists in the probing of the elastic and plastic properties of a sample by an indenter of known geometry and properties. This is generally done with dedicated equipment, designed especially to perform simultaneous measurement of the force applied on the indenter, and its penetration depth in the sample. The advent of Atomic Force Microscopy (AFM) in the past decades has accelerated the development of the field by allowing the imaging and probing of a wide variety of surfaces with submicron scale indenters.

Landmark experiments and simulations of nanoindentation on the Au(111) surface have been carried out by a number of researchers since the late 1990's. Some of the most important results from the literature are reviewed in section 6.1.2. A brief survey is nonetheless provided below to highlight the relevance of the investigations presently reported.

Corcoran *et al.* [75] performed nanoindentation experiments in 1997, with a load–controlled[1] feedback scheme. Load–displacement[2] curves were obtained that showed

---

[1]Load-control and displacement control are described in section 5.1.2.

[2]Displacement refers to the penetration depth of the indenter into the sample: it is the negative

several discontinuities attributed to discrete plastic events in the sample, such as dislocation generation and motion. Measured forces were on the μN scale, using an indenter with a tip radius of 205nm.

In 1998, Kiely, Houston, and colleagues [76, 77] reported a comprehensive study of gold surfaces that were passivated using thiols to suppress the effect of adhesion. The forces measured also range in the μN scale, and the indenters had tip radii between 70nm and 175nm. This work was conducted in parallel with atomistic simulation work by Kelchner *et al.* [78].

In 2006, Asenjo *et al.* [79], conducted nanoindentation experiments on Au(111) using an AFM in ambient conditions. The reported load–displacement curves suggest force resolutions on the order of 50nN, and the indenter used was poorly known with an assumed radius of 10nm.

Also in 2006, Cross *et al.* [80] published results previously obtained with our FIM/STM/AFM system. The data was obtained with a sharp indenter on with a tip radius on the order of a few nm. The tip was characterized by Field Ion Microscopy (FIM) prior to experiments carried out in Ultra High Vacuum (UHV). Force–distance curves were acquired with a simultaneous recording of the electrical current flowing between tip and sample. Several phenomena were observed and described such as discrete plastic events, reverse plasticity, and self healing of the surface upon indentation. Some of the studies limitations included poor force resolution ($\sim$ 15nN), limited current range ($I_t < 100$nA), and absence of Scanning Tunnelling Microscope (STM) imaging of the indented region.

In the present studies, nanomechanical properties of clean Au(111) surfaces in UHV are explored with very sharp W(111) indenters of tip radii between 3.6nm and 8.5nm. Great care was taken, and significant enhancements were implemented, to improve stability, reproducibility, and resolution. New insights were extracted from simultaneous measurement of the current over a wide range. In addition, STM/AFM images were successfully acquired before and after nanoindentation. Tips were also characterized by FIM before experiments, and in most cases, after as well.

---

of the tip sample separation.

4.2 Overview of Experiments

4.2.1 Data Sets

A number of experiments were conducted in different conditions, with different types of tips (polycrystalline, single crystal, and cut tips), and with different types of sample clamping. In the rest of this thesis, we report nanoindentation investigations using (111) oriented single crystal tungsten tips of varying radii, and samples anchored to molybdenum clips using a high purity silver paint supplied by SPI$^{\text{TM}}$. Ten data sets were obtained on ten separate days. On average, each data set contains approximately 100 static mode STM/AFM scans and 200 force-distance curves of varying range.

FIM imaging was performed prior to the acquisition of each data set, and in 8 out of 10 cases, after as well. The two data sets without tip imaging after nanoindentation correspond to very sharp tips ($R_{\text{tip}} < 4nm$), which were damaged by the application of large stresses during deep nanoindentation. In both cases, damage to the tip was immediately noticeable on force–distance curves showing abnormal features, such as the retraction trace appearing above the approach trace. Out of the eight tips that were imageable after nanoindentation, seven showed a mild coverage of adsorbate, which are interpreted as gold atoms extracted from the surface. Field evaporation of the adsorbates reveals an underlying W(111) tip with the same radius obtained prior to nanoindentation, which attests to the maintained integrity of the tip. In the remaining case, although damage signatures similar to those described earlier were observed on force–distance curves, the tip was imageable by FIM, which showed a large level of distortion and revealed a crack. Corresponding FIM images, shown in figure 3.19, are similar to images obtained with tips damaged during other experiments carried out on the FIM/STM/AFM system.

As outlined in section 4.2.2, two different cantilever beams were used in the acquisition of the ten data sets. However, the cantilever beam with the lowest resonance frequency (sample CB200905C, with $f_1 \simeq 715$Hz, and $k_0 = 4.5$N/m) did not allow the resolution of subtle events in the force channel. As illustrated in figure 4.1, this is explained by the fact that the sample becomes strongly excited after its rising resonance frequency reaches system frequencies around 1250Hz. The Power Spectral Density Evolution (PSDE) shows that the 1250Hz excitations are present initially, before the cantilever resonance frequency, starting around 715Hz starts rising, due to increased junction stiffness (see section 3.3.2). In these conditions, a combination of low–pass and band–pass filtering can reveal the general trend of the curve, but

any discrete event and subtle effects are impossible to discern.
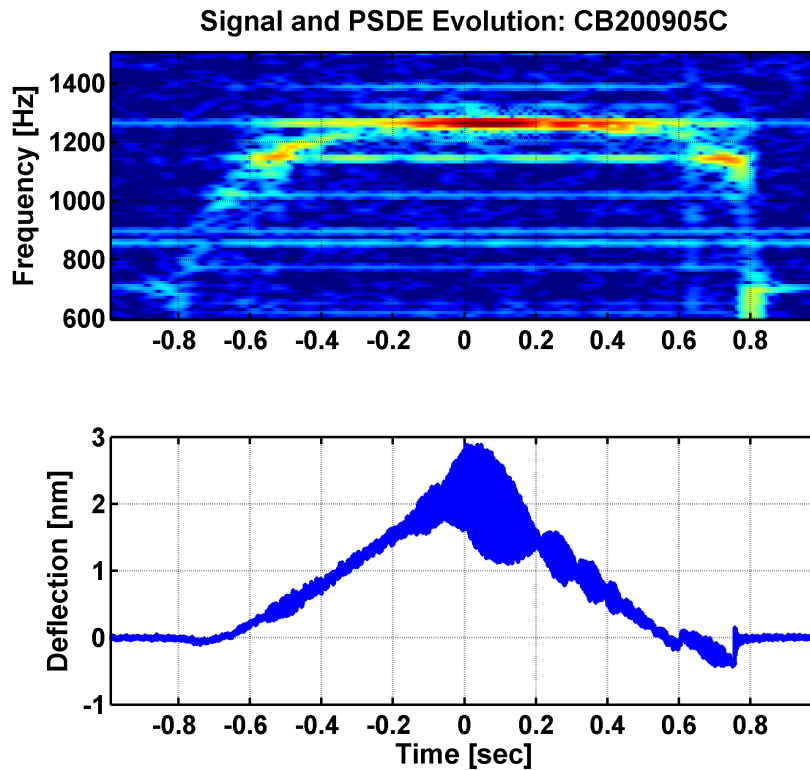
Figure 4.1: PSDE plot of the calibrated raw deflection channel during a nanoindentation curve acquired with sample CB200905C.

Therefore, sample CB200905C can produce useful results at low applied forces only. As we are interested in investigating the nanoindentation regime, the seven data sets obtained with it were therefore excluded from the analysis presented in the rest of this thesis.

The results and interpretations in this thesis are thus based on three experimental data sets, with three tips on a given sample (CB200905E) as described in section 4.2.2. These three data sets alone comprise over 250 STM scans and 750 force–distance curves, each containing tens of thousands of points on several channels. The tremendous wealth of information that they contain goes beyond the scope of this thesis. To facilitate their continued analysis in the future, the data sets have been labelled according to the day of measurement, and are therefore referred to henceforth as: DS20, DS28, and DS02.

4.2.2 Sample Used in the Present Study

The cantilever beam used to acquire the data is referred to as CB200905E. Its important characteristics are summarised in table 4.1.

Sample	CB200905E
k_0	$32 \pm 1.3 \text{N/m}$
L	$5.7 \pm .1 \text{mm}$
w	$1.4 \pm .1 \text{mm}$
τ	$\sim 50 \pm 5 \mu \text{m}$
f_1	$\sim 1500 - 1600 \text{Hz}$

Table 4.1: Geometric parameters of sample CB200905E.

Although the sample used for all three relevant data sets is the same, the tip position was different, which led to different spring constants, as summarised in table 4.2.

Data Set	t	k_t	z_1
DS20	$2.95 \pm .1 \text{mm}$	$230 \pm 2 \text{ N/m}$	$3 \pm .1 \text{mm}$
DS28	$3.40 \pm .1 \text{mm}$	$150 \pm 2 \text{ N/m}$	$3.5 \pm .1 \text{mm}$
DS02	$2.70 \pm .1 \text{mm}$	$300 \pm 3 \text{ N/m}$	$2.7 \pm .1 \text{mm}$

Table 4.2: Tip positions and spring constant values at the tip for all three data sets. z_1 is the position of the interferometer beam closest to the fixed end of the sample.

Noise spectra obtained through the acquisition system directly are shown in figure 4.2. The spectra were taken to estimate the random uncertainty on the measurement of the deflection. Table 4.3 shows a summary of the integrated noise over several bandwidths.

	5-50Hz	5-500Hz	5-1000Hz	5-2000Hz	5-5000Hz
DS20	1.0	1.0	1.1	15.2	15.4
DS20 (filt.)	0.4	0.5	0.5	1.6	1.6
DS28	0.8	1.2	1.3	8.9	10.9
DS28 (filt.)	0.2	0.2	0.3	0.9	0.9
DS02	1.0	1.4	1.5	7.0	8.9
DS02 (filt.)	0.2	0.3	0.4	0.8	0.8

Table 4.3: Quadratic uncertainty on the deflection measurement: $\Delta u_t{}^2$, in units 10^{-4}nm^2.

For comparison with CB200905C (Fig. 4.1), a PSDE of sample CB200905E during a nanoindentation experiment is shown in figure 4.3.
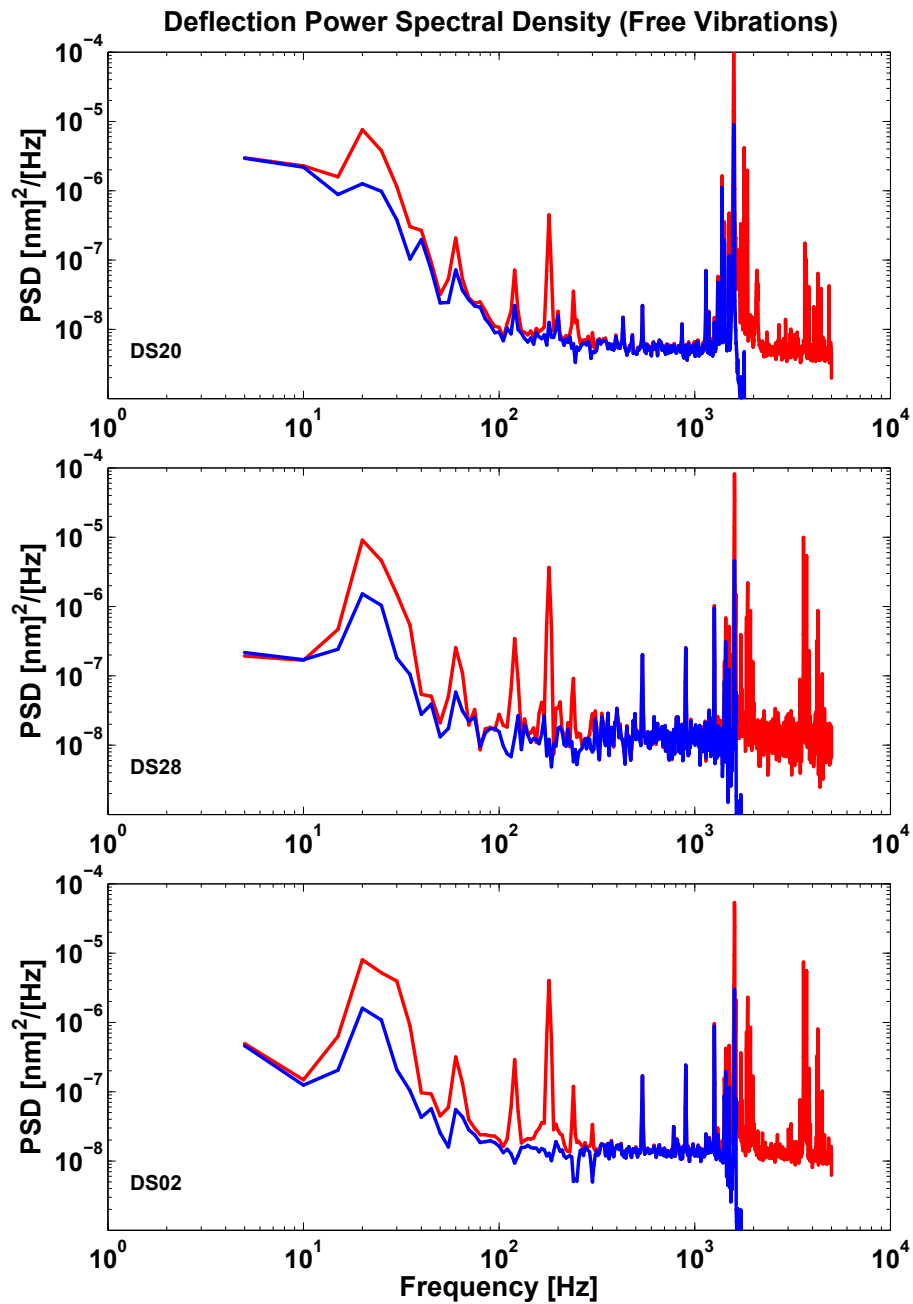
Figure 4.2: Power spectra of the deflection channel in the measurement conditions for data sets DS20, DS28, and DS02. Power spectra were obtained with the Welch method (see section A.2, using 400, 100, and 100 segments respectively, with a *hanning* window. Red traces correspond to raw data, and blue traces to the standard filters applied to our data sets.
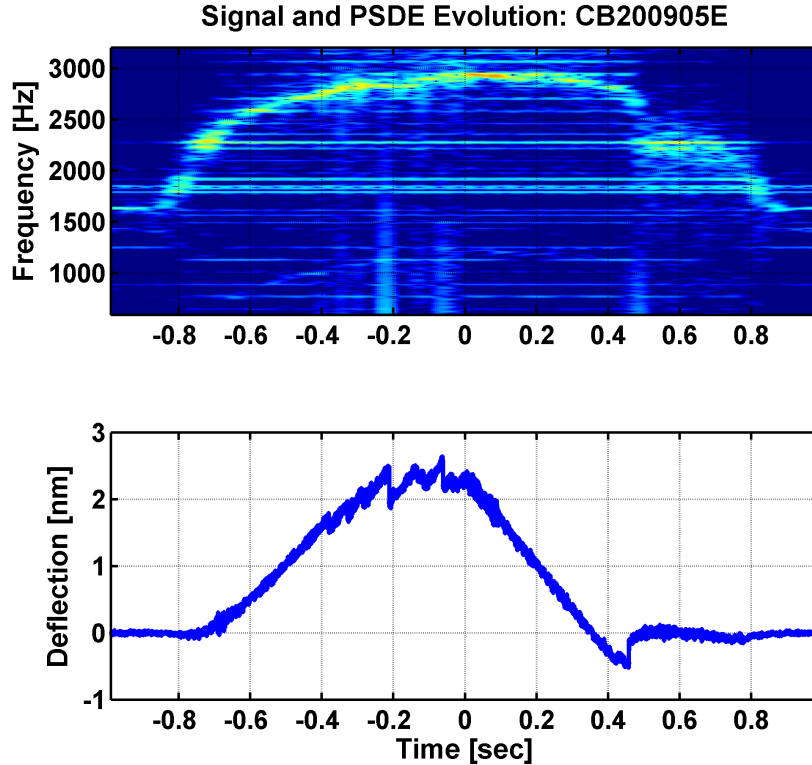
Figure 4.3: PSDE of sample CB200905E during a nanoindentation experiment. Contrary to sample CB200905C, in this case, vibrations remain moderate as the resonance frequency increases. Vertical lines in the PSDE plots correspond to $1/f$ noise caused by discrete plastic events observable in the deflection trace underneath.

4.2.3 Tips Used in the Present Study

The primary results reported here were obtained with three different tips, each used to produce different data sets, and labelled according to the name of the data sets in question. Characterization by FIM (Fig. 4.4) shows that the tips had three different radii: Tip20 ($R = 3.6 \pm 1.2$nm), Tip28 ($R = 6.1 \pm 1.2$nm) and Tip02 ($R = 8.5 \pm 1.2$nm). All tips were imaged prior to nanoindentation, and with the exception of Tip20, after nanoindentation as well. Tip20 was no longer imageable due to a fatal crash.

Three dimensional models of the tips can be built based on FIM images. One such reconstruction is presented in figure 4.5, courtesy of T. Hagedorn. In the model of Tip28, the top layer atoms are represented in red, the second layer atoms in orange, and the third layer in white. Other atoms visible in FIM are highlighted (also in orange). Atoms that are not visible in FIM due to insufficient local field

(a) Tip20 ($R = 3.6 \pm 1.2$nm)   (b) Tip28 ($R = 6.1 \pm 1.2$nm)   (c) Tip02 ($R = 8.5 \pm 1.2$nm)
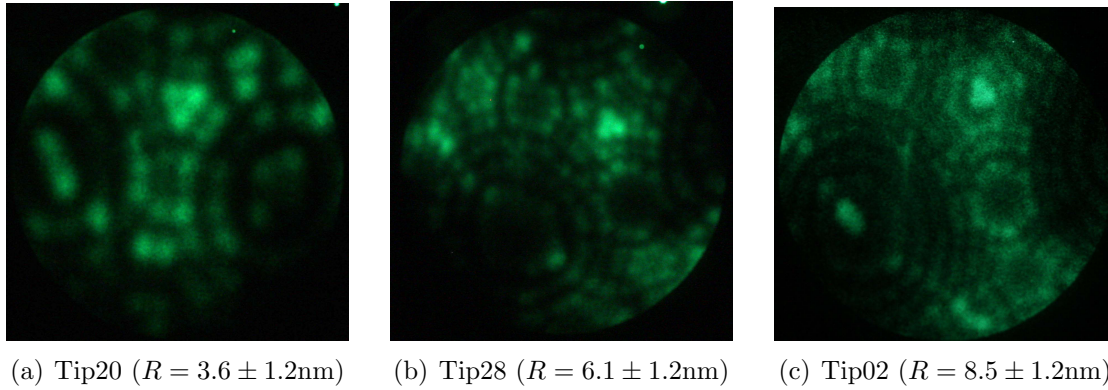
Figure 4.4: FIM images of the main three tips used in this study. The tip radii were determined by counting the number of rings between the 111 and 001 poles (3, 5, and 7 respectively).
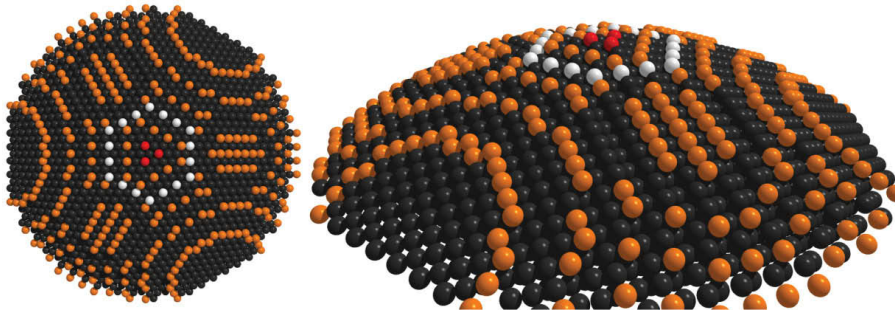
Figure 4.5: Three dimensional rendering of Tip28. Courtesy of T. Hagedorn.

enhancement are shown in dark grey.

   To gain an appreciation of the general size and shape of the different tips used, approximate models are easily built based solely on the tip radii, as depicted in figure 4.6. These differences lead to different contact mechanics behaviour and to significant size effects, as discussed in sections 6.2.3 and 6.3.2.

4.2.4   Experimental Conditions and Signal Processing

   Experiments were carried out in UHV and at room temperature. STM/AFM and indentation data was acquired using Scanita. Surface scans were taken before and after nanoindentation curves. These consisted in force–distance curves obtained by loading the surface through the extension of the piezotube according to a predetermined actuation profile. Forces and currents were recorded simultaneously during acquisition of nanoindentation curves. The rigid body displacement (or depth, defined as the negative of the tip–sample separation) was computed as outlined in section 2.3 and illustrated in figure 2.7. This allows the display of the data as load–displacement
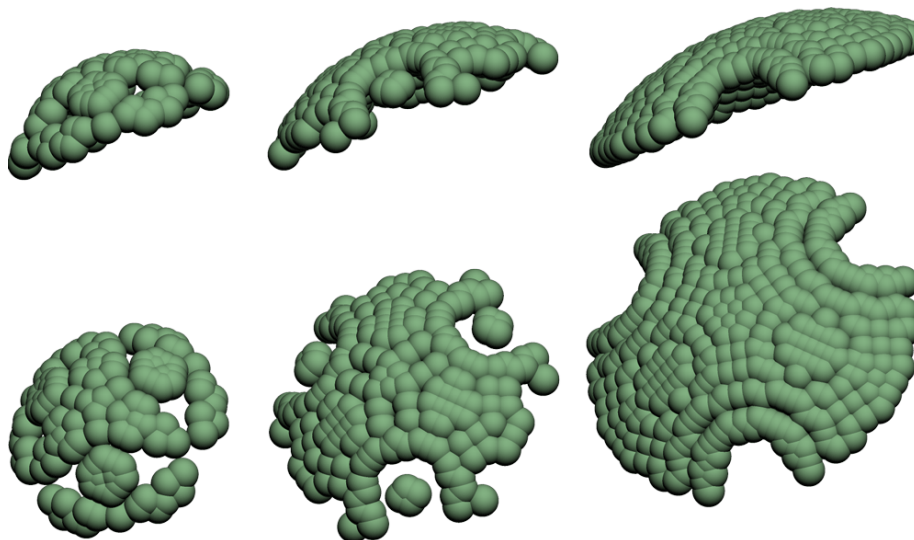
Figure 4.6: Approximate tip models based on tip radii of Tip20, Tip28, and Tip02 respectively. Only atoms of the outer shell are represented. Rendered using QuteMol [39].

curves, which are standard in the field of nanoindentation. Load–displacement curves were acquired at both pristine and previously loaded sample positions. In addition to single curves, sequences of 2-40 curves were acquired at times, to enable the investigation of cyclic loading.

Nanoindentation data was acquired at sampling rate of 10kHz, typically for a duration of 2 seconds for the full approach-retraction cycle. The raw data in volts was recorded, and later calibrated. The electronic current channel was acquired directly by Scanita, as described in section 2.1.3, with no antialiasing measure other than the low pass effect of sampling at 200kHz before downsampling to the control/data frequencies.

The interferometer signal channel (s in the notation of section 2.3), which is used to compute the deflection, the force, and the displacement, is acquired through a hardware order 8 low pass filter for antialiasing, in addition to 200kHz oversampling. In the data presented here, s was acquired in most cases through a 10kHz filter, which does not allow full suppression of the high frequency noise folding back in our measurement bandwidth[3]. For display and analysis, the s channel was post-processed and saved as an additional filtered channel which was used to compute filtered versions of the deflection, force and displacement. The post-processing filtering was performed in Matlab^TM, using customised processing tools based on Matlab^TM's *filter design*

---

[3]The 2.5kHz high order low pass filter normally used for this purpose was unfortunately out of order during these measurements.

toolbox (see section A.3). It consisted in applying an order 8 low pass butterworth filter with a cutoff frequency at 1500Hz, and 8 narrow 2Hz band-stop butterworth filter with centred frequencies at 20Hz, 30Hz, 60Hz, 120Hz, 180Hz, 240Hz, 250Hz, and 300Hz. This gentle and targeted filtering allows the suppression of line noise and spectral peaks at high frequencies[4] with minimal effect on the resolution of discrete mechanical events. The filters were applied using the Matlab™ *filtfilt* function which performs zero-phase filtering by processing the data in both forward and reverse directions, which reduces distortions of the data and effectively doubles the order of all applied filters. Comparisons between filtered and unfiltered data are shown in figure 4.2 in the frequency domain and in figure 4.8 in the time domain.

The data was logged and managed by implementing a data management, analysis, and exploration system developed using object oriented programming concepts in Matlab™. This system, described in appendix B, facilitates the tracking of data from a vast variety of sources along with contextual information such as user comments and experimental conditions. In particular, the generation of dynamic plots allows for the synthesis of large quantities of data into quantitative curve indicators displayed in simple interactive plots, an example of which is shown in figure 4.7. In these plots, each data point corresponds to a single load–displacement curve, and all points that correspond to a given curve sequence are connected and displayed in a different colour, with the first curve highlighted by a star in selected cases. The choice of colour is consistent throughout each data set for seamless comparison of indicators when examining multiple indicator plots.

In the case of figure 4.7, the two indicators displayed are the first piezo extension value and the experiment time. The latter is obtained by analysing a time stamp generated automatically when the acquisition of each curve starts, and is displayed here relative to the first curve. The former is the first value of the absolute extension applied to the piezotube at the very start of each nanoindentation experiment. Since this point is consistently chosen as 2nm above the point of tunnelling contact (100pA at 50mV), this provides an estimate of the correction performed by the controller to offset the vertical drift throughout the experimental session. Figure 4.7, which shows the relationship between these two indicators for the 349 curves that make up DS20, therefore allows the drift to be estimated, and yields an upper bound for the drift of about 2nm/min, corresponding to sub-angstrom errors for files acquired in less than

---

[4]High frequency mechanical peaks often correspond to vibrations of the cantilever beam on the free end side of the tip, where the interferometer beams are located. The tip effectively "clamps" the sample at the tip position, and therefore, high frequency peaks are likely to be detected at disproportionately high amplitudes relative to the actual level of vibration at the junction.

2 seconds. Drift was therefore considered negligible in our conditions and no drift correction measures were taken. No sign of appreciable piezotube nonlinearities were observed in the range of displacements. If present, these would produce characteristic image artefacts, such as distortion, and the elastic portions of loading and unloading would not overlap perfectly..
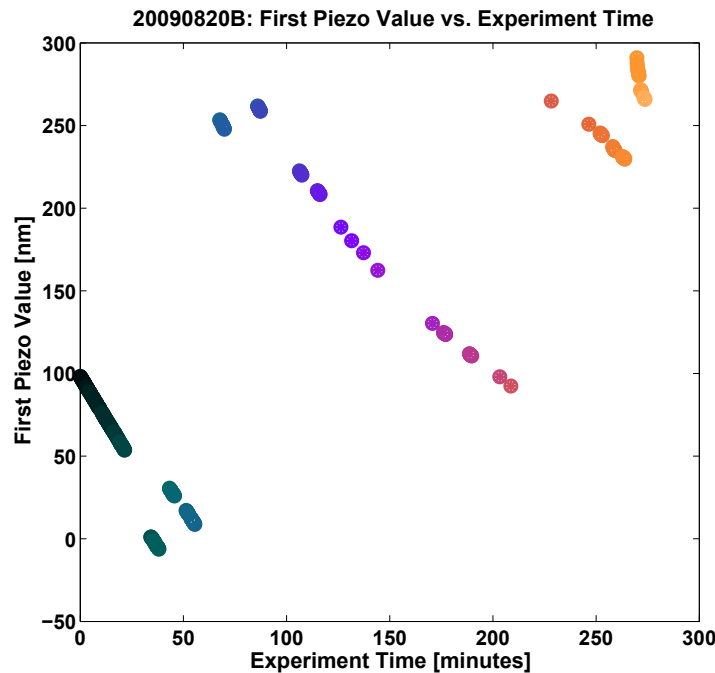
Figure 4.7: Time evolution of the first absolute position of the piezotube for DS20, providing an estimate of a vertical drift of 2nm/min, which is negligible in our conditions. Each point corresponds to a different curve, and connected curves correspond to a specific indentation sequence. The various starting points correspond to situations when the sample position was changed appreciably.

4.3 Overview of Results and Phenomenology

In the following section, a brief overview of the results and phenomenology will be introduced. We start with single loading, along with a description of the graphing conventions followed throughout this thesis. We then show various phenomenological aspects of the results, including loading with surface imaging and cyclic loading, particularly with respect to its effect on electronic transport.

4.3.1   Single Loading of Nanoscale Contacts

An example of a complete load–displacement curve obtained with our system is shown in figure 4.8. The colour code used here is maintained throughout, and summarised in table 4.4. Unless otherwise specified, depicted current traces are calibrated raw data from the log-IV converter without filtering or downsampling. To enhance the visibility of important features, the solid load traces correspond to a filtered version of the load according to the procedures described in section 4.2.4. Lighter background shades of the load curves correspond to calibrated and unfiltered raw data.

	Current	Load	Filtered Load
Loading	Green	Blue	Light Blue
Unloading	Red	Black	Grey

Table 4.4: Colour convention for load–displacement curves.
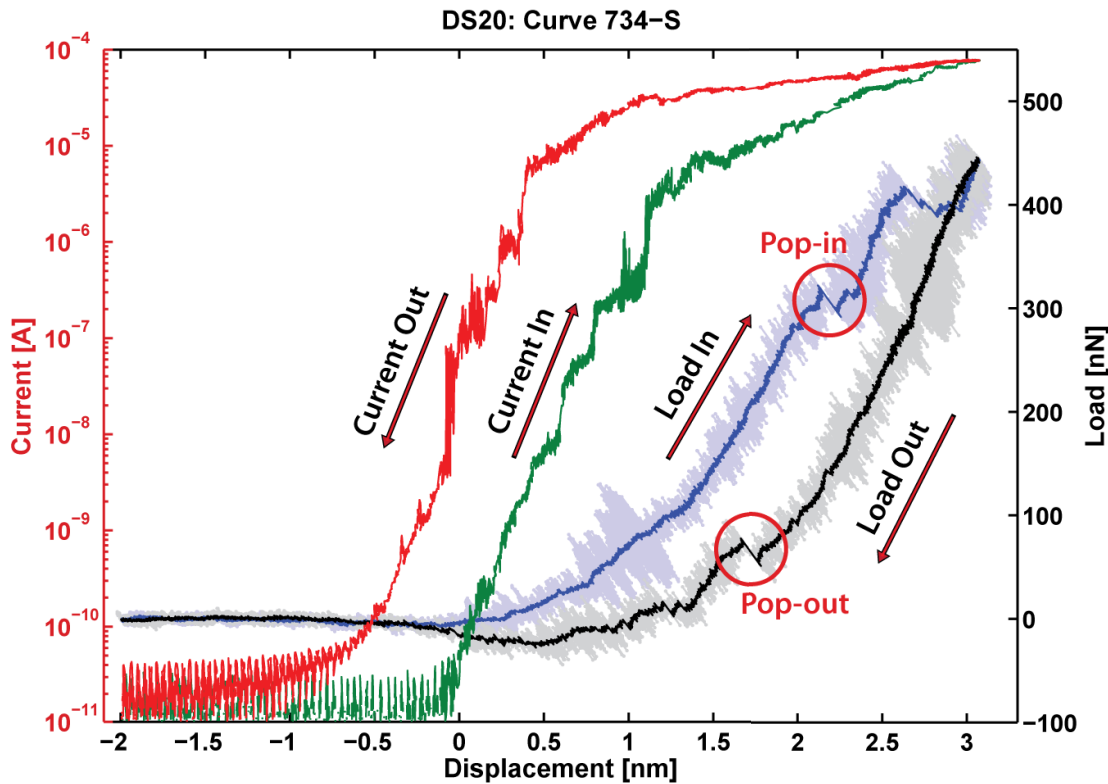
Figure 4.8: Example of a typical load-displacement curve. The colour code is consistent throughout the thesis. See text for details.

Several important insights can be extracted from the load–displacement curve shown in figure 4.8, which will be explored in detail in chapter 5. Briefly, as the

tip approaches the sample it eventually makes tunnelling contact, marked as zero[5]. Upon further extension of the piezotube, elastic–plastic loading of the contact occurs, with interspersed discrete plastic events having counterparts in the unloading curve (pop-ins and pop-outs respectively). Particularly noticeable, is a *yield* point, approximately 200-300nN in data set DS20, beyond which the curves change significantly. After the maximum load is applied, the tip is retracted immediately[6], starting with a quasi-linear unloading slope and eventually reaching negative loads, which is a signature of adhesion. Simultaneous current traces show jumps correlated with discrete events in the loading curve. In the example of figure 4.8, such correlations are difficult to see, due to the logarithmic scale in current, which masks subtle effects at high currents.

The curve displayed in figure 4.8 was acquired on a pristine area of the sample which was imaged before and after nanoindentation, as shown in figure 4.9. The figure represents a succession of four single indentation curves and associated images, which highlights the fact that the general look of both the curves and the associated surface damage is very reproducible. Such nanoindentation curves typically produce a hole and 1-3 hillocks that are typically one atom high islands with the threefold symmetry of the underlying substrate. The hexagonal structure of both craters and hillocks is also easy to see in figure 4.10, which shows image scans of the surface performed after the scans of figure 4.9. The hole–hillock structure caused by curve 734-S is further analysed in conjunction with the load–displacement curve in section 5.3.2.

The succession of images in figures 4.9 and 4.10 also provide interesting insights into the motion of the islands and craters. They show that craters are significantly less mobile than islands, which change shape and move during scanning. A large portion of the diffusion is tip assisted, as can be seen in the middle scan of figure 4.10. In that case, the scanning direction was down-up, and it is easy to see that the island near crater #1 is moved towards crater #4, along the $[11\bar{2}]$ direction, as evidenced by the herringbone reconstruction. Another island is *"collected"* at that point, and the assembly is moved towards the middle point between craters #2 and #4, as seen in the last image. After the last scan, several islands have completely disappeared.

---

[5]Unless otherwise specified, the zero-point in the curves presented are defined as the point at which tunnelling was stable at a tunnelling resistance of $\sim 500M\Omega$, immediately prior to the start of the curve acquisition.

[6]In typical nanoindentation experiments, a hold time is often added at maximum load to study time dependent effects. While Scanita does not allow the implementation of hold times, a Matlab^{TM}/Tcl/Tk TCP/IP system for remote acquisition with the ability to apply arbitrary actuation profiles was developed jointly with William Paul. Initial trial runs of the system in a realistic test environment have been performed successfully.
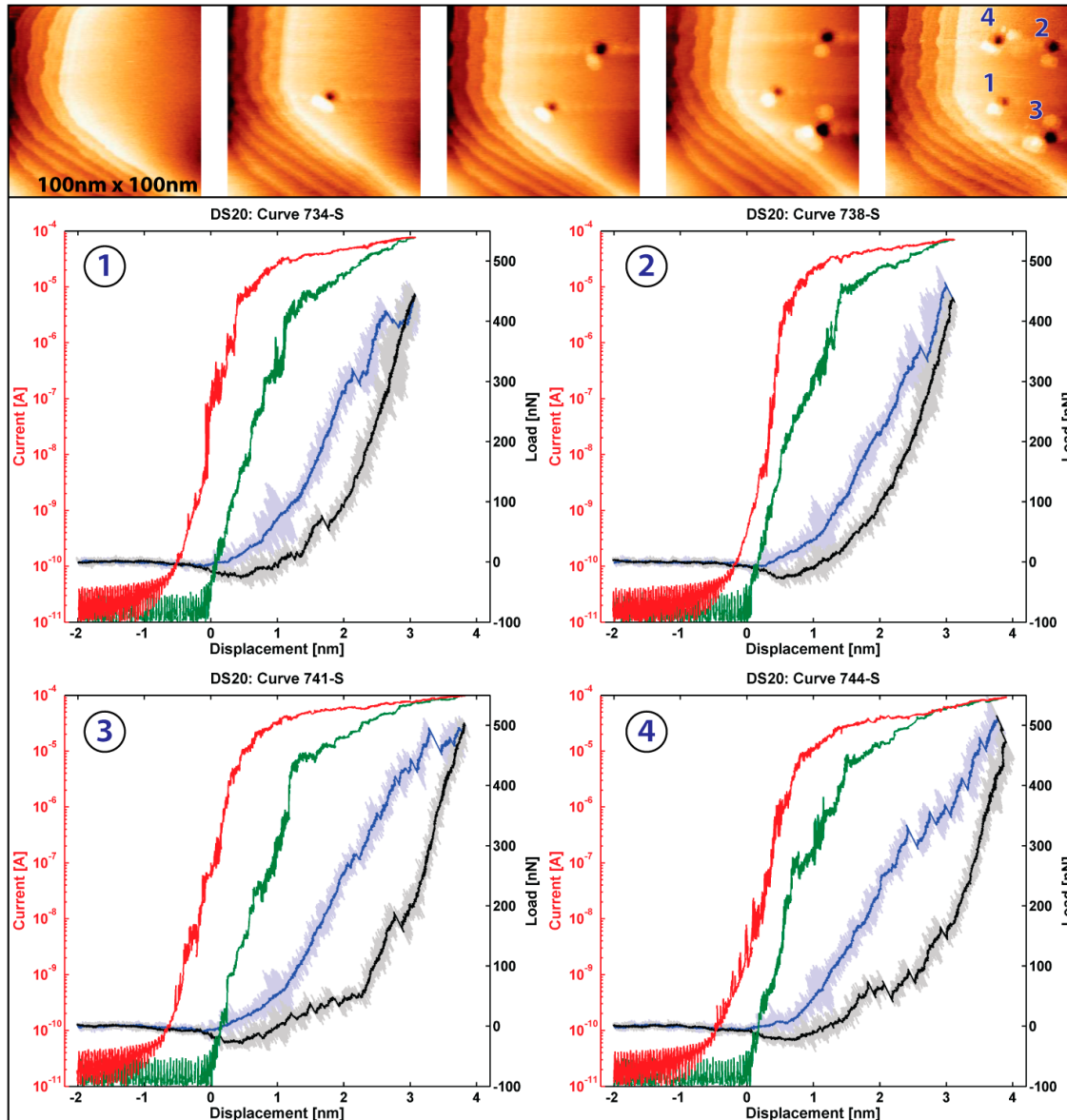
Figure 4.9: A series of five STM/AFM images, with a single nanoindentation curve acquired in-between. See text for details.

Although the craters have largely remained at their original location, a number of small vacancy clusters become visible on the surface. With the data at hand, it is difficult to separate out how much of this diffusion is intrinsic and how much is tip assisted, and whether the emerging vacancy clusters are formed directly from crater vacancies or whether they are produced by free vacancies being pinned by surface defects in the underlying substrate that are created during nanoindentation. These interesting questions could be answered by a systematic study, which would yield valuable insights of great relevance to fundamental surface physics and to technology,
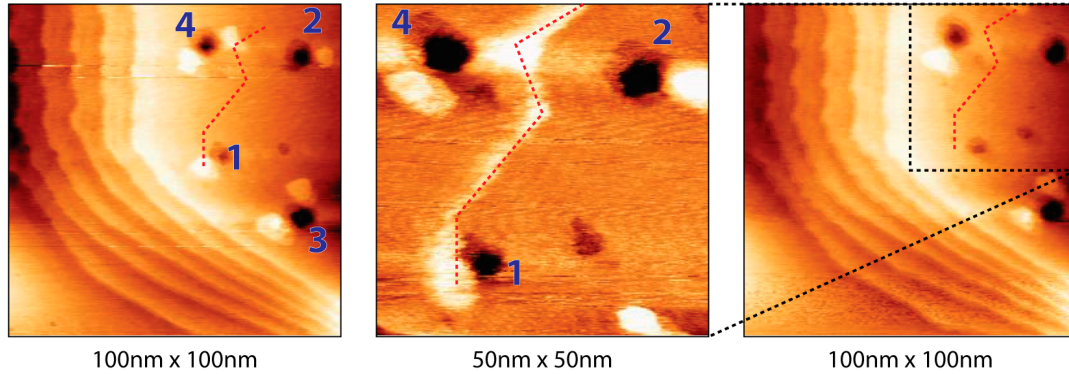
Figure 4.10: Three consecutive STM topography images. As shown by the square dotted outline (black), the middle image is a zoom into the top right corner of the other images. The scan for the middle image was in the upward direction, which assisted the island initially in the vicinity of hole #1 along the herringbone reconstruction towards one of the rightmost side islands of crater #4, subsequently moving it toward to top right corner of the next scan. The trajectory of the islands is highlighted by the red dotted line.

as they pertain to the resilience of structures machined at the nanoscale.

4.3.2 Cyclic Loading of Nanoscale Contacts

In addition to single indentation curves, it is interesting to investigate the effect of cyclic loading at a given location. For that reason, a number of sequences of 2-40 curves were taken with the same parameters. Such curves were typically obtained at 7 second intervals, which allows 5 seconds for the tunnelling feedback to be switched on and stabilised before a new indentation curve is taken.

A sequence of 20 curves from data set DS20 is shown in figures 4.11 and 4.12. The behaviour of the contact during this sequence shows very reproducible features. Typically, the first curve will show a large hysteresis which is a signature of plastic events, such as the formation of a crater. The hysteresis loop typically gets smaller within 2-3 curves before virtually disappearing, after which quasi-perfect elastic loading/unloading is seen at high loads, as the indenter and the crater conform appreciably. Smaller hysteresis loops are still visible at low loads, which are attributed to adsorbed gold on the tip or hillocks on the surface that need to be pushed away before full loading is in effect.

Plotting all curves from a given sequence in the same graph, as shown in figure 4.12, allows the extraction of interesting points and trends. For example, it is apparent that the maximum load decreases during the sequence. This behaviour is observed for a large number of other sequences as well. Several mechanisms can be
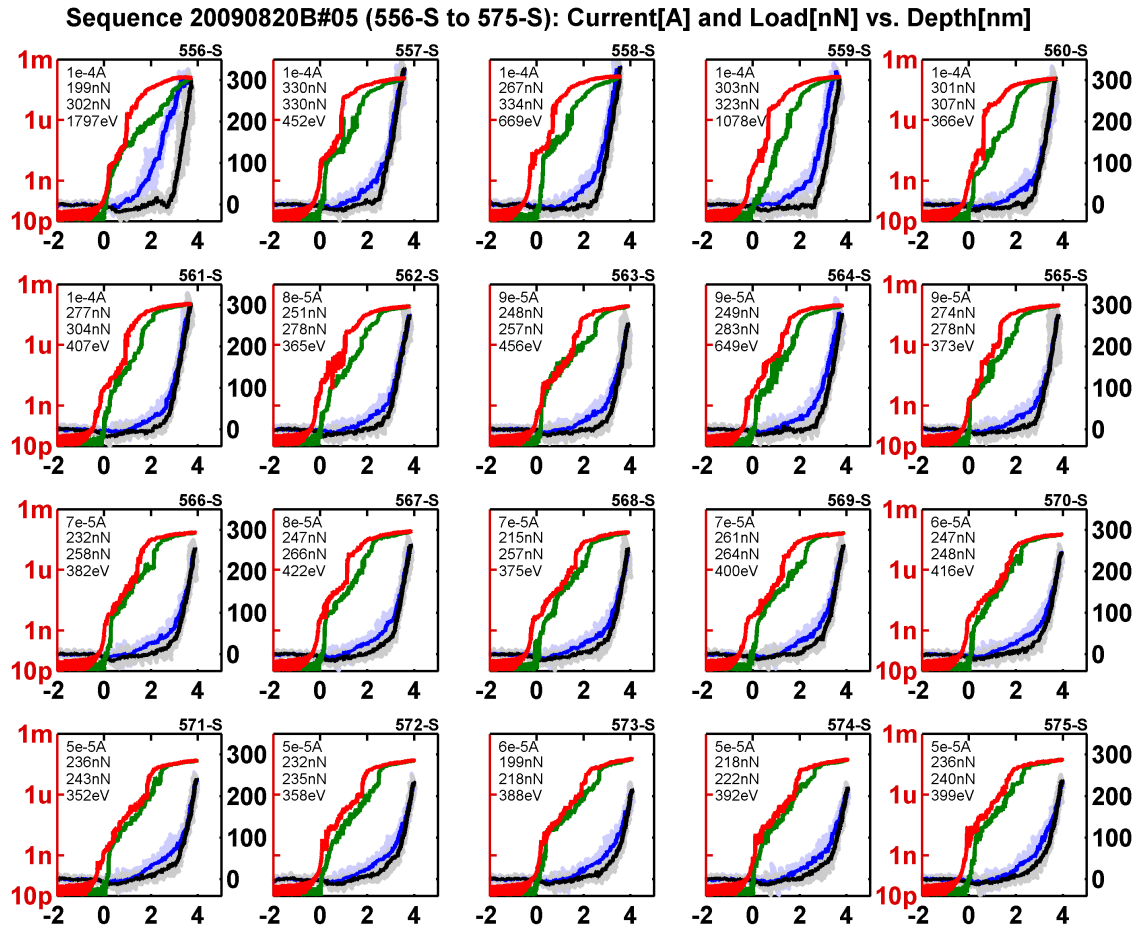
Figure 4.11: Cyclic loading of sequence 05, from data set DS20. The number in each plot correspond to maximum current, force at maximum current, maximum force, and total hysteresis energy respectively.

imagined in order to explain this phenomenon.

This is highlighted in figure 4.13, which proposes two simple mechanisms to account for these differences. During a given cyclic loading sequence, the zero point is defined differently for each data curve, as it corresponds to the point of tunnelling contact immediately prior to acquisition. When an island is present on the surface (due to prior nanoindentation, or when a gold atom is adsorbed on the tip) the tips starts off higher above the sample in absolute term. As the same piezotube extension profile is applied to all curves, the cases with the tip starting off higher will result in penetration that is not as deep into the sample as in the original case. As the hillocks and absorbed gold atoms are easily pushed away, the tip will face less resistance to penetration and the maximum loaded will be lowered. This has an important consequence on the computation of displacement: the displacement is computed using $h = z - u$, where h is the displacement, z the piezo extension,
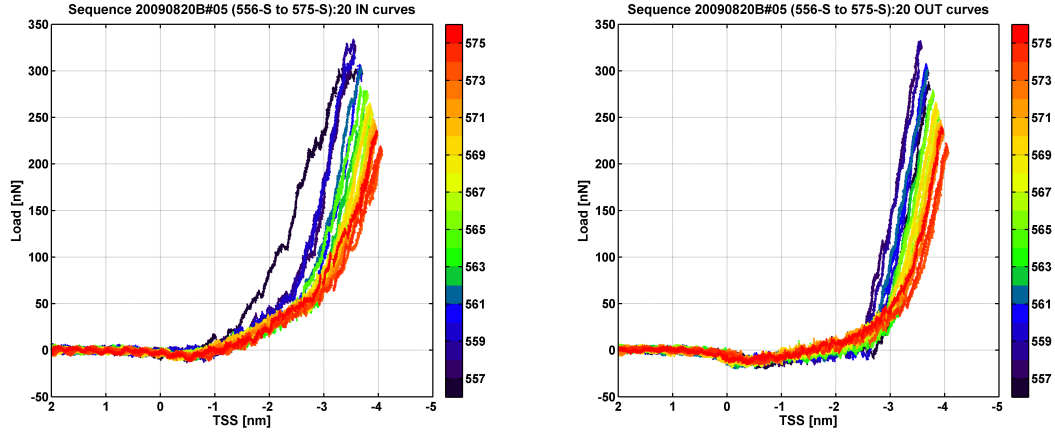
Figure 4.12: Cyclic loading of sequence 05, from data set DS20. All 20 curves are shown together to highlight their evolution. Dark blue colours correspond to early curves and red to later curves.
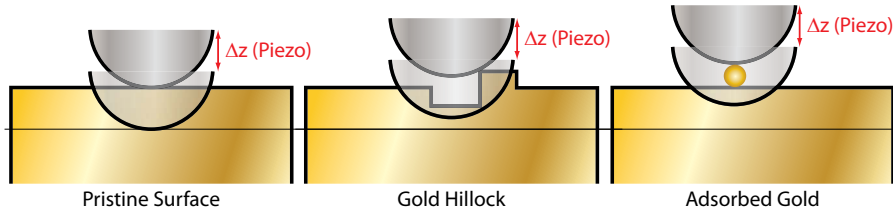
Figure 4.13: Proposed mechanisms to account for the reduced load upon cyclic loading. See text for details.

and u the cantilever beam deflection (see section 2.3.1 and figure 2.7). When a specific maximum piezo extension Δz is chosen for a given cyclic loading sequence, the maximum displacement is given by $h_{max} = \Delta z$ $u_{max} = \Delta z$ P_{max}/k_t. It is immediate to see that in cases where the maximum load is reduced ($\Delta P_{max} < 0$), the computed maximum depth will be increased by $\Delta h_{max} = \Delta P_{max}/k_t$, although the tip does not penetrate as deep into the surface.

The idea outlined in the previous paragraph is corroborated by the results shown in figure 4.14. The graph on the left is a plot showing the evolution of the maximum as a function of curve number, which shows that virtually all sequences with an initial peak load above ~ 200nN present the effect discussed earlier, albeit with disparate strength. The 200nN load limit is commensurate with the yield point in the indentation curves depicted above. A further look at a graph of all hysteresis energy against peak load (Fig. 4.15) unambiguously confirms that 200nN is the limit above which appreciable plastic deformation occurs, which is associated with the formation of hillocks of gold. Taken together with the graph on the right of figure
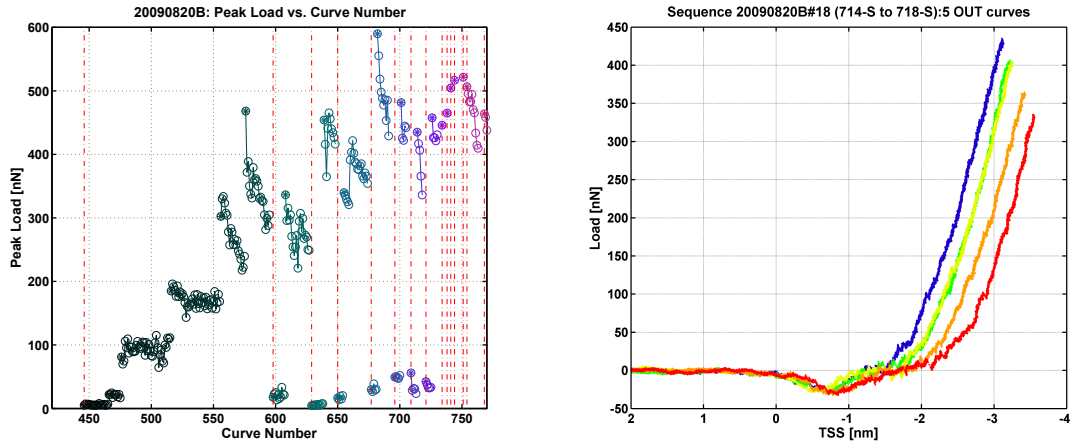
Figure 4.14: Reduced load upon cyclic loading. Left: above 200nN, the maximum load decreases appreciably upon cyclic loading. Vertical dotted lines correspond to a change of sample position to a pristine location. Right: Five retraction curves from a specific sequence showing differences in computed displacement that are multiples of the gold lattice constant.

4.14 which shows retraction curves separated by multiples of the gold lattice constant, these facts corroborate the validity of the proposed mechanisms.
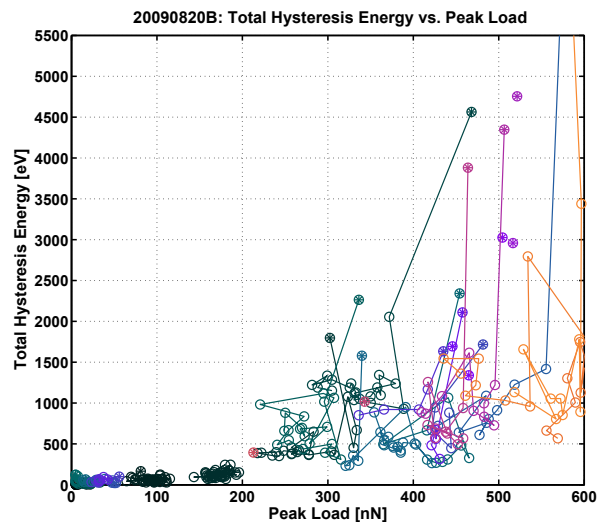
Figure 4.15: Plot of the hysteresis energy vs. peak load for DS20, suggesting a limit of elastic loading at $\sim$ 200nN. The graph also highlights that the hysteresis energy drops significantly after the first curve, as indicated by the starred circles corresponding to the first curve of each sequence.

In conclusion to the points made above, it becomes clear that undesirable offsets in the computed displacement can be introduced by gold adsorption and/or plastic loading. For this reason, in our analysis, we favour methods that rely on quantities

that are defined more robustly, such as load and stiffness, which is defined as the derivative of load with respect to displacement.

4.3.3 Electronic Transport Through Defects

Perhaps the most consistent effect of cyclic loading, which has not been previously reported in the literature to our knowledge, is the dramatic drop in conductance during loading. As we have seen, curve indicator plots, such as those presented earlier, can lead to interesting insights. In this section, we look more closely at the conductance of our contacts. To obtain a measure of the extent to which our nanomechanical contacts are good electrical contacts, we compute the maximum linear conductance by dividing the maximum current by the bias voltage. Further dividing by the conductance quantum, we obtain the graph of figure 4.16 with the 349 curves of DS20, which shows a clear decrease in conductance during cyclic loading. This result is consistent throughout all data sets acquired.
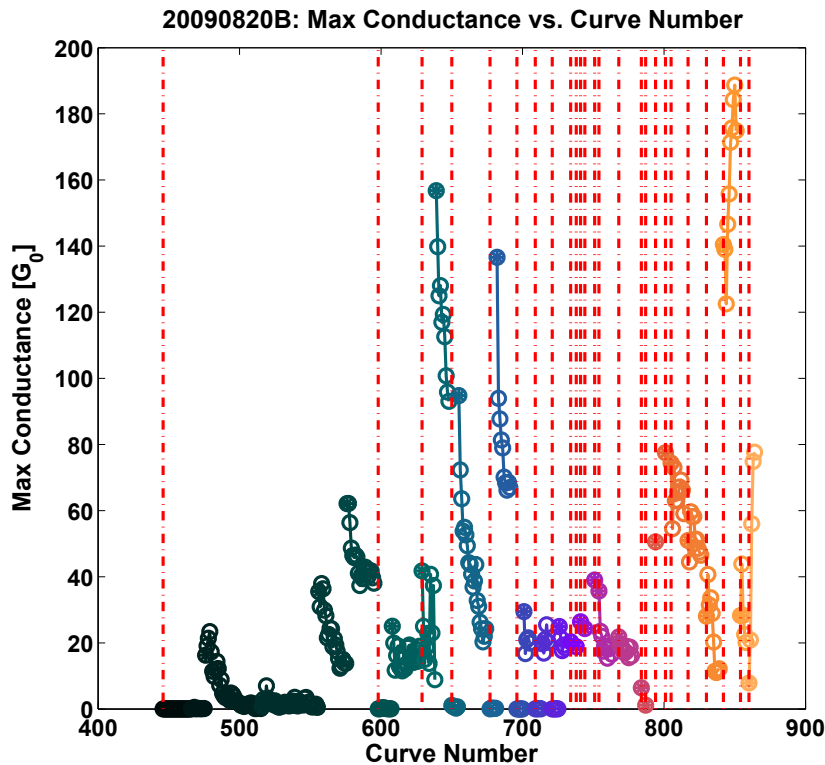
Figure 4.16: A plot of the maximum conductance vs. the curve number shows that the maximum conductance drops as cyclic loading is performed. Vertical dotted lines correspond to a change of sample position to a pristine location.

Several hypotheses were explored to account for this unexpected behaviour. One

such hypothesis relies on the observation made in section 4.3.2 that the maximum load diminishes during cyclic loading. A smaller load implies a smaller contact area, which implies a smaller conductance. However, a closer look at data shows that this hypothesis cannot account for the entire magnitude of the conductance drop. The relationship between load and current in elastic contact is expected to be roughly linear. By purely geometric arguments, after the onset of plastic deformation, the surface and indenter conform better, which implies an increased surface area and thus an increased conductance. In contrast, the data shows that in most cases, a drop effect on the order of 50% in the current will be accompanied by an effect on the order of 10-20% in load. In fact, in the case of sequence #3, displayed in figure 4.17, the load increases, the maximum load *increases* from 80nN to 110nN, while the maximum current drops two orders of magnitude.
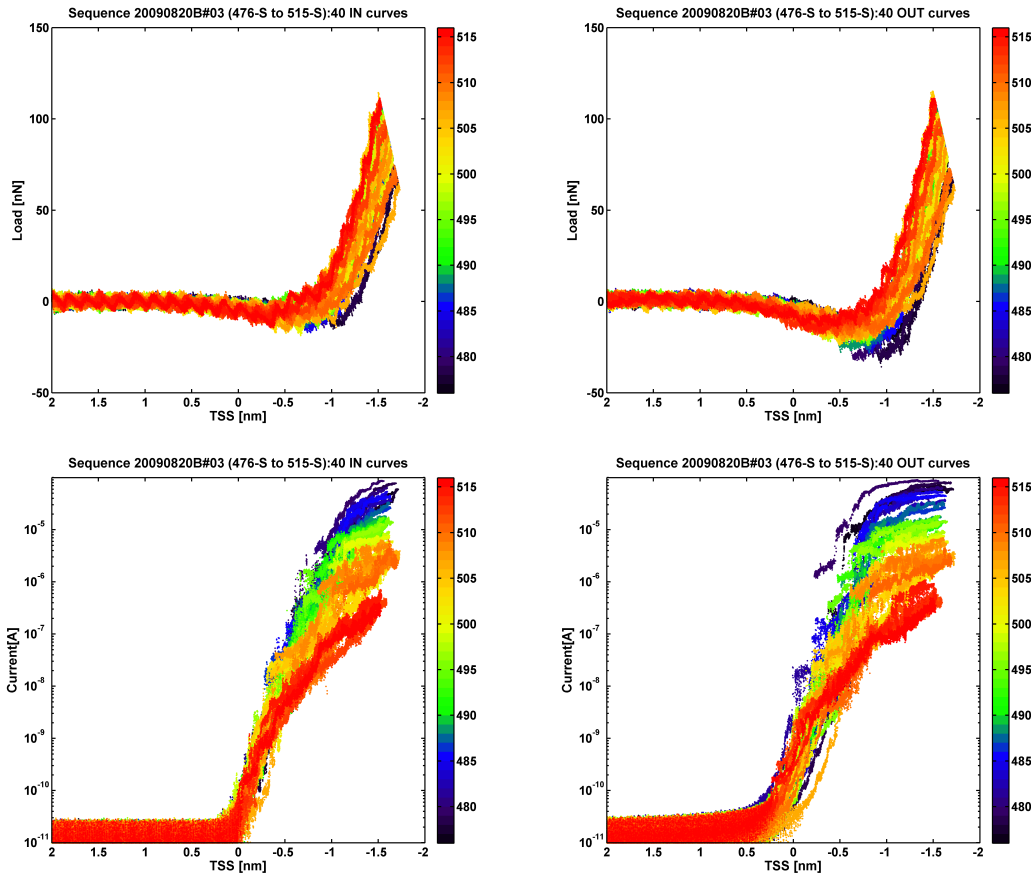
Figure 4.17: Sequence #3 from DS20 shows an drop in maximum conductance of two orders of magnitude, while the maximum load increases, suggesting a larger contact area.

Other hypotheses to explain this phenomenon involved damage to the tip. How-

ever, figure 4.16 clearly illustrates that moving to a new sample position (red vertical dotted lines) instantly increases the conductance ostensibly. Immediate increase in conductance pursuant to deeper indentation at a given location (such as is the case for the first few sequences of figure 4.16) also preclude the perspective of a contamination layer on the surface.

Having eliminated other likely possibilities, the preferred interpretation is that plastic deformation produces a large number of defects locally, which creates a volume below the surface that becomes progressively more disordered as the contact is cyclicly loaded, until the electronic transport is significantly suppressed. It would therefore be expected that moving to a new sample position would restore the conductance. Similarly, deeper indentation at a given location increases the probed volume, and therefore the conductance, until the cyclic loading eventually affects the entire volume, as observed experimentally. These results and interpretations have prompted collaboration with the group of Prof. Hong Guo at McGill University. The Guo group is able to conduct unprecedented *ab initio* simulations of electronic transport through defects, and preliminary tests and analyses performed by Jesse Maasseen have proven very promising. Figure 4.18 shows the effect of various degrees of disorder in a 4.1nm gold slab on the conductance. Preliminary results show that 25% disorder surprisingly leads to an 80% drop in conductance, which is on the same scale as the experimental results presented here. Further work is being done to investigate the effect of tungsten as one of the electrodes.

The dramatic drops in conductance that we observe have important fundamental and technological implications. In particular, as nanoscale electronic devices are built, ensuring robust contacts that are mechanically resilient and electronically reliable is a key requirement that will require in-depth knowledge into the contact mechanics and electronics at the nanometre scale.

4.3.4   Self Healing

A last phenomenological feature of a certain class of cyclic loading sequences should be highlighted: self healing. In rare cases, generally in highly stepped regions, a very unique type of cyclic loading sequence can be observed whereby matching inward and outward plastic events in the loading and unloading curves respectively will produce a compact hysteresis loop in the compressive region. As illustrated in figure 4.19, this is such that both above and below the loop, loading and unloading curves match perfectly. This type of behaviour was reported by Cross *et al.* [80] who interpreted the phenomenon as a recovery of the indented surface by reverse
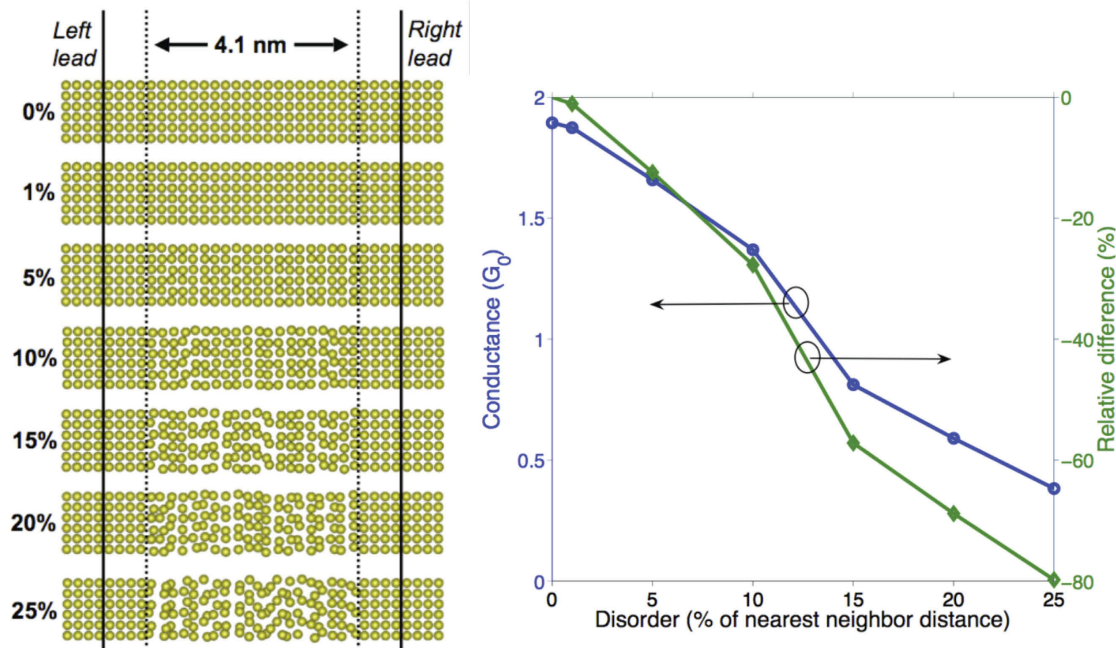
Figure 4.18: Preliminary *ab initio* simulations suggest that a 25% degree of disorder in a 4.1nm long gold slab can suppress 80% of the conductance. From J. Maassen

plasticity. It is also interesting to see that the loading curves and the unloading curves respectively are strikingly similar to one another. While sequences showing such behaviour are relatively infrequent, they have been observed observed for several combinations of tips and samples, and further investigation is warranted to determine their cause.

In this chapter, we have highlighted the wide variety of phenomena observed. In the following, the focus is placed on more quantitative analyses. First, traditional nanoindentation analysis methods will be outlined and applied to our context. And second, interesting emerging size effects will be detailed.
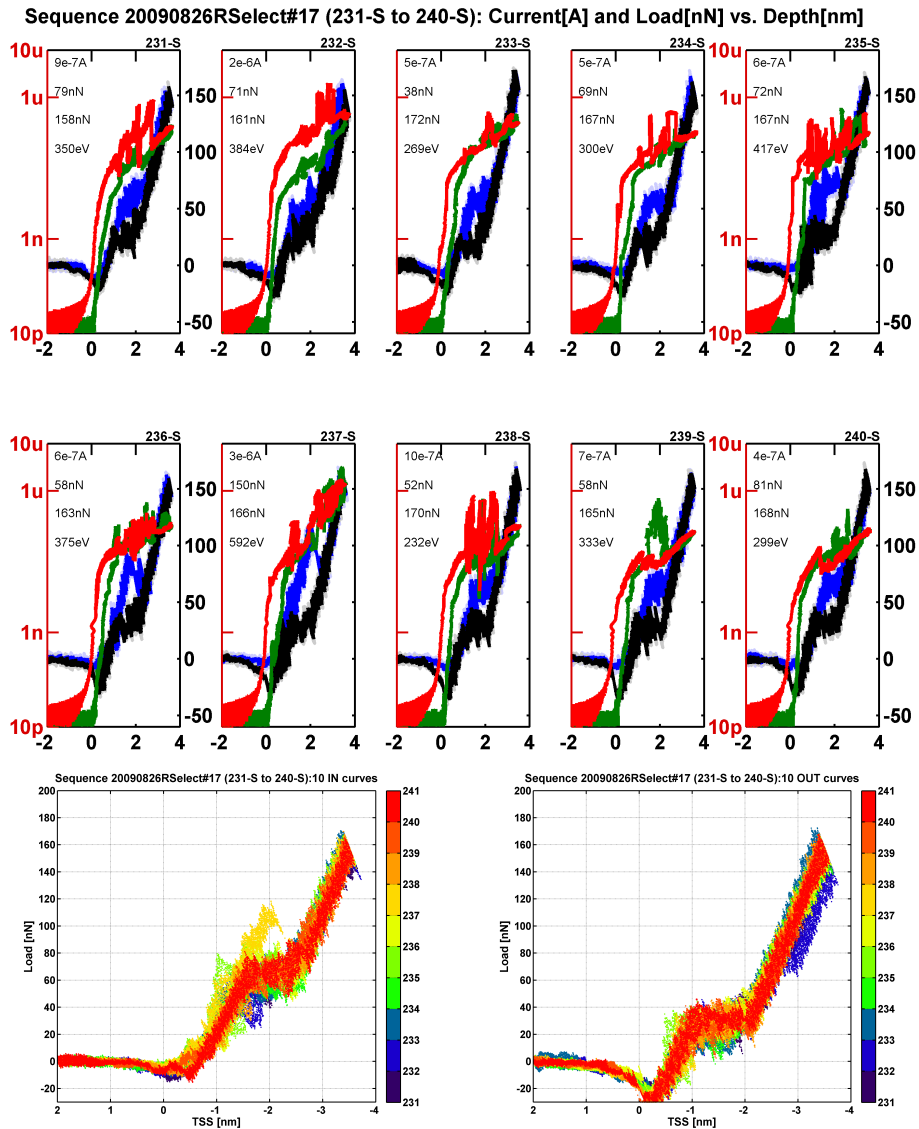
Figure 4.19: Self Healing and compact hysteresis loops. The data is from DS26, which is not otherwise analysed in this thesis, as it was obtained with sample CB200905C, which is the reason for the strong noise.

Chapter 5

Measurement of Elastic-Plastic Properties with Nanoindentation

5.1 From Hardness to Nanoindentation

Nanoindentation is a powerful tool that can be used to investigate the elastic and plastic behaviour of materials. It allows the extraction of important insights which are of relevance for both fundamental and applied science.

5.1.1 Hardness of Materials

Nanoindentation has its roots in the measurement of elastic and plastic properties of materials, and most notably their hardness. To this day, the concept of hardness has yet to be defined in an entirely satisfying manner. The term can be used to refer to a number of different concepts, such as scratch hardness, dynamic hardness and indentation hardness. It is the latter that is of interest in the present thesis. In general terms, it is a measure of a material's resistance to penetration and permanent deformation.

The hardness of materials has been widely studied for several centuries, with most advances taking place in the twentieth century, and stemming from the work of Brinell and, shortly thereafter, Meyer, in the early 1900's [81–85]. In an introductory review [86] to the *Focus Issue on Indentation Methods in Advanced Materials Research*, a special issue of the *Journal of Materials Research*, Hutchings provides an interesting historical perspective on indentation hardness, outlining in particular the valuable contributions of D. Tabor.

In an indentation hardness test, an indenter is pushed into a surface until plastic

deformation occurs. A quantity that typically has the dimensions of a pressure is then usually extracted by dividing the applied load by a characteristic surface area. In the Brinell Hardness Tests, the area used is the full curved surface area of the residual impression left by a ball of steel of specific weight and shape. The Meyer Hardness Test, which is of most interest to us, is based on the empirical observation that the mean pressure during indentation becomes largely constant when a plastic zone has fully developed. The Meyer Hardness is precisely that mean pressure, and is obtained by dividing the applied load by the projected area of the residual impression left by the indenter.

During an indentation test, a rigid indenter is driven into an elastic–plastic sample. As the compressive load is increased, three distinct regions can be observed (Fig. 5.1):

The Elastic Region: In this region, the deformation of the sample is completely reversible. This occurs when the mean applied pressure is below a certain threshold ($p_m \lesssim 1.1Y$) where Y is the Yield Stress.

The Elastic-Plastic Region: In this region, a plastic zone starts to develop. The response of the system in this region is very complex and at present, no analytical theory can fully describe this region.

The Plastic Region: In this region, the plastic zone is fully developed and the mean applied pressure stays largely constant. As the load is increased, the contact area increases correspondingly to keep the mean pressure relatively constant. It is observed both in experiments and simulations that this occurs when $p_m = H \simeq CY$ where C is a material dependent parameter and is usually close to three for metals [82, 83, 85].

5.1.2   Nanoindentation

Using concepts from Contact Mechanics, elastic–plastic properties of materials can be easily extracted from the measurement of the load and contact area during indentation. However, empowered by the advancement of science and technology, researchers became interested in probing the properties of ever smaller volumes of material, which gave rise to challenges in the accurate measurement of the area of residual impression. Hardness thus became increasingly difficult to measure, and new methods were developed.
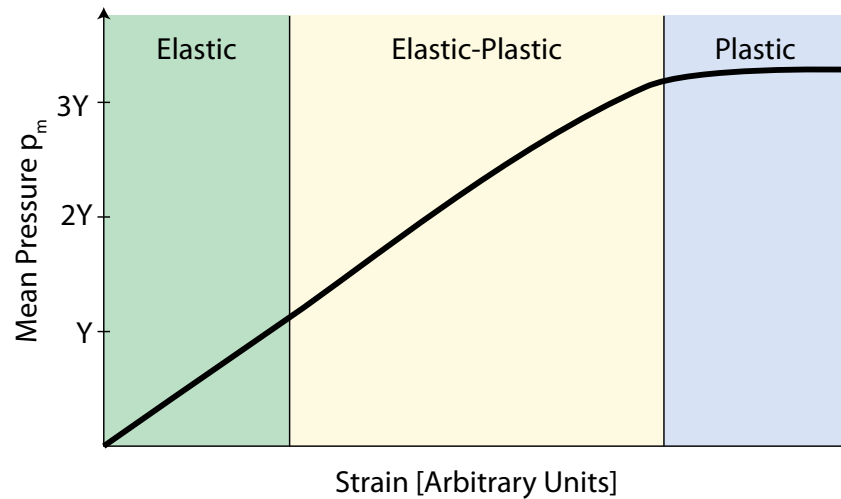
Figure 5.1: A typical indentation experiment can be divided in three regions. As the strain is increased, the mean applied pressure first increases reversibly (elastic), then irreversibly (elastic–plastic), and eventually reaches a limiting value, defined as the indentation hardness.

Depth sensing indentation, or instrumented indentation, is one such method. It is a simple extension of indentation testing based on the simultaneous measurement of the applied load (P) and the indentation displacement (or depth, h). Analysis of load–displacement curves, with knowledge of the tip geometry, allows for an estimation of the contact area and the measurement of important material properties such as elastic modulus, yield stress, and hardness.

Figure 5.2 illustrates typical load–displacement curves. When the load remains below a certain threshold, which depends on the tip geometry and the material properties of the sample, the deformation remains elastic (fig. 5.2 (b)), consistently with the first region in figure 5.1. Beyond that point, a plastic zone starts to develop as described previously.

Nanoindentation is the application of depth sensing indentation at the submicron scale. For an introduction to the experimental and analytical methods involved in nanoindentation, the reader may refer to the book by Fischer-Cripps [85] or reviews by the same author [87, 88]. Other interesting introductory references include the section on nanoindentation of the Handbook of Nanoscience and Nanotechnology [8], a review by Schuh [89], and the experimental technique paper by Hay [90]. A more detailed account of the analysis methods of load–displacement curves, along with other relevant references, will be given in section 5.3.
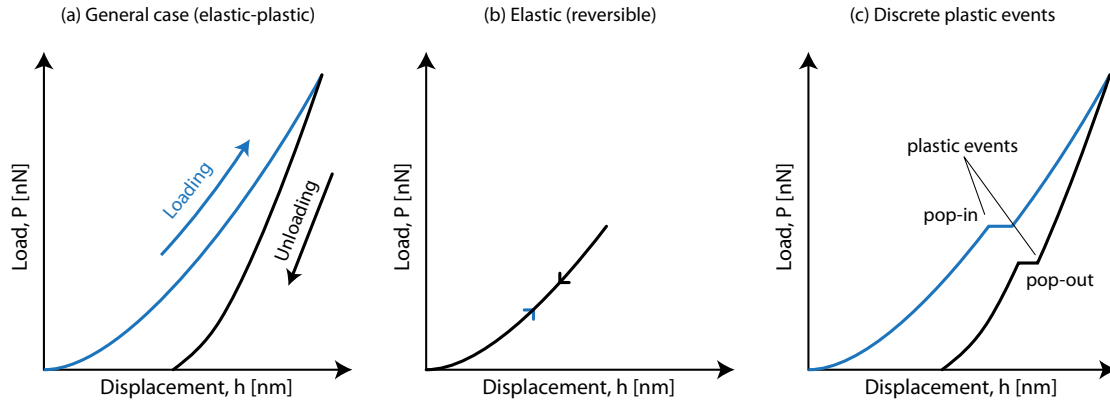
Figure 5.2: Typical Load–Displacement curves. The colours used here are consistent throughout the thesis (blue for approach/loading and black for retraction/unloading). In the general case of non-adhesive elastic–plastic behaviour (a), a hysteresis loop demonstrates the presence of irreversible (plastic) processes. In the elastic region (b), the loading and unloading curves overlap entirely. In some situations, discrete plastic events can be seen (c) which correspond to sudden rearrangements below the surface due to nucleation and motion of defects. Pop-in events (during loading) are far more frequent than pop-out events in the literature.

Acquisition of Nanoindentation Data

The acquisition of nanoindentation data is usually carried out using dedicated equipment capable of indenting a surface while simultaneously measuring the applied load and the indentation depth. Detailed descriptions of nanoindentation systems are available in chapter 11 of reference [85] and in [91].

Actuation Control   Most systems are load–controlled, which means that a chosen load profile is applied to the sample. This is due to the very nature of the loading systems that are employed, which typically rely on hardware that can be operated to impose a specific force through mechanical, electrostatic, or electromagnetic instrumentation. An independent measurement of the applied load is generally acquired, often by using a spring of very low stiffness in parallel with the indenter shaft. The chosen load profile is then maintained by using a feedback loop.

Another approach involves controlling the depth (or displacement), also through a feedback loop (Figure 5.3 highlights the difference between the two control methods). Whereas the overall shape of the load-displacement curve will largely remain the same, the details of the control method will greatly affect the way in which discrete plastic events are emphasised.

Displacement–control has been shown to be more sensitive to discrete plastic
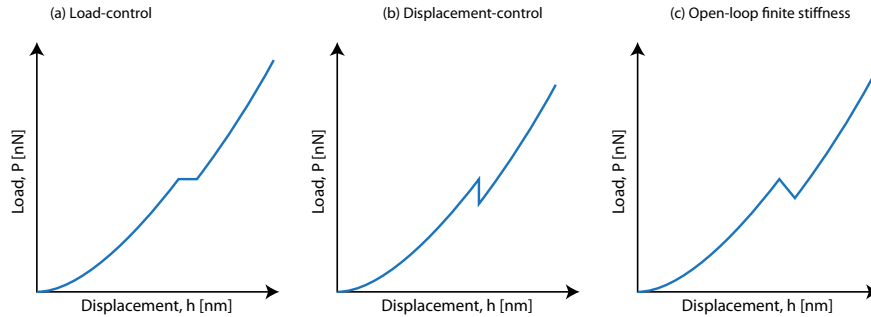
Figure 5.3: In load–control (a), a discrete plastic event appears as a horizontal jump in the P-h curve. In displacement–control (b), it appears as a vertical dip. In the absence of control and with a deflection measurement in series with the indenter (c), such as in the FIM/STM/AFM system, pops appear inclined along a negative slope that is equal to the spring constant of the deflection measurement system. See sections 5.1.2 and 5.1.3 for more details.

events [92–94]. This can be understood as follows: when a pop-in occurs, the depth increases suddenly, which would naturally be accompanied by a drop in load. In load–control, as the instrument attempts to maintain the load, a large amount of energy is added to the system in a poorly controlled fashion. This can cause a multitude of events that hinder the detection of the initial discrete event. In contrast, in displacement–control, the load drops immediately in a visible manner and the amount of energy added to the system remains controlled as the load is restored by the control system.

Soer *et al.* illustrate the difference between load and displacement–control for the detection of discrete plastic events in the case of Al (a) and Al-2.6%Mg (b) (Fig. 5.4). Their interpretation is that these yield events are associated with the nucleation of dislocation in the sample. The lower resolvability of yield events in the load–controlled experiment on sample (b) is then attributed to the fact that the burst is detectable only to the extent that the sample can generate a multitude of dislocations that can move freely, which is hampered in this case by solute pinning.

Actuation Profile   Whether load–controlled or displacement–controlled, the actuation profile (Fig. 5.5) that is applied to the indenter will typically contain the following features: a fixed rate approach to a specific target, an optional hold (or dwell) time, and a fixed rate retraction. A second dwell time is sometimes added close to the end of the retraction. The dwell times are added to investigate time dependent issues such as creep and thermal drift.
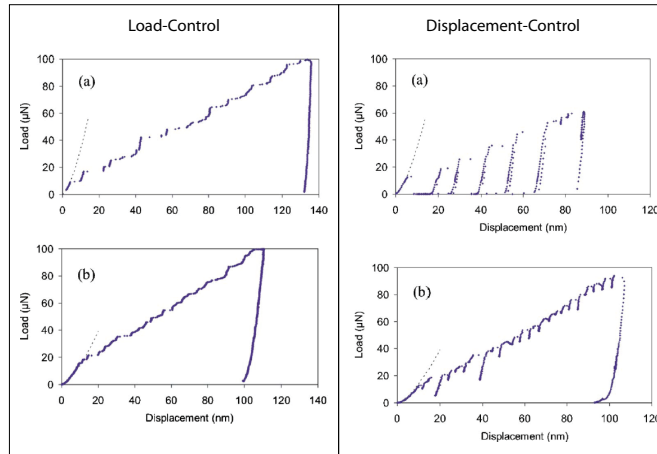
Figure 5.4: Load and displacement control P-h curves for Al (a) and Al-2.6%Mg (b). Reprinted with permission from [93]. Copyright 2007, American Institute of Physics. The higher resolvability of discrete plastic events in displacement–control is evident in the case of Al-2.6%Mg. See text for details.

5.1.3   Nanoindentation with the FIM/AFM/STM

Our system differs from traditional nanoindentation equipment in several ways, as we use a Scanning Probe Microscope that operates simultaneously in Scanning Tunnelling Microscopy (STM) and Atomic Force Microscopy (AFM) modes, in Ultra High Vacuum (UHV). The combination of these surface science tools allows, in particular, the imaging of the surface both before and after indentation, with force and displacement resolutions on the sub-nN and sub-angstrom scales respectively. In addition, we use a Field Ion Microscope (FIM) setup to characterize our indenter at the atomic scale, both before and after indentation.

Finite Stiffness Open–Loop control

Our AFM system (see section 2.1.1) allows for the direct measurement of the static deflection of the cantilever beam as the indenting tip in driven into the sample by the extension of the piezoelectric tube.

In the experimental setup as it currently stands, the measurement is carried out by applying a specific extension profile to the z-piezo (Fig. 5.5). A number of signals are acquired simultaneously, such as the applied voltage to the z-piezo, the tunnelling/contact current, and most importantly for the present discussion, the cantilever beam deflection. After calibration, this channel is used to compute both the displacement (the difference between the z-piezo travel and the deflection), and the load (by multiplying the deflection by the cantilever beam spring constant at
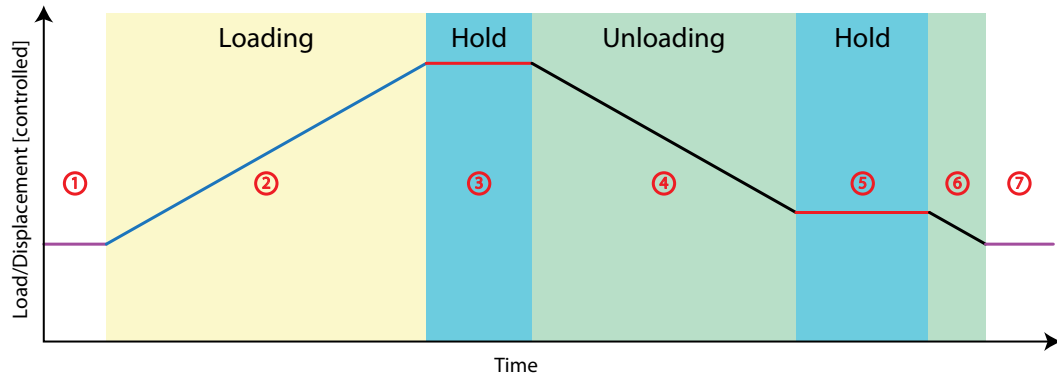
(a) Traditional nanoindentation: the load or displacement is controlled throughout the experiment. One or two optional hold/dwell times (3 and 5) can be implemented to allow the investigation of creep and thermal drift.
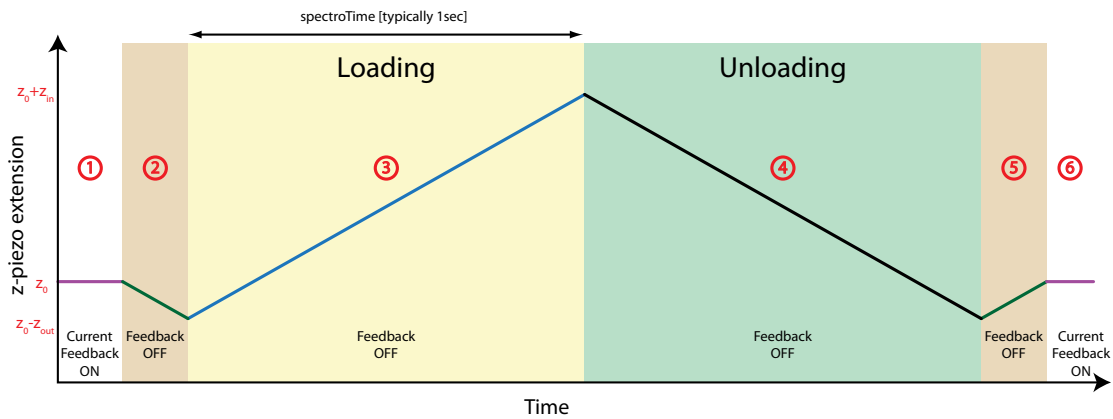
(b) Nanoindentation with the FIM/STM/AFM system: the actuation profile is applied to the z-piezo passively (without control) and does not incorporate hold times. The profile is imposed by the current software package, with z_{out}, z_{in}, and the duration of the approach (spectroTime) being the only user-defined parameter. The data acquired in sections 3 and 4 is used for analysis. Matlab^{TM}-based enhancements for the acquisition and the actuation profile are discussed in section 7.2.

Figure 5.5: Nanoindentation actuation profiles

the measurement location (Fig. 2.7)). Further details are provided in chapter 2, in particular in section 2.3.

No feedback mechanism is currently employed to control the extension of the piezo, once a specific extension profile shape and associated parameters have been selected and an indentation started. The applied load and measured displacement are acquired passively, according to a scheme that could be referred to as finite stiffness open–loop control. This type of control is such that our experiments fall into a regime that is somewhat between load–control and displacement–control. In load–control, a discrete plastic event would cause the displacement to burst forward. In
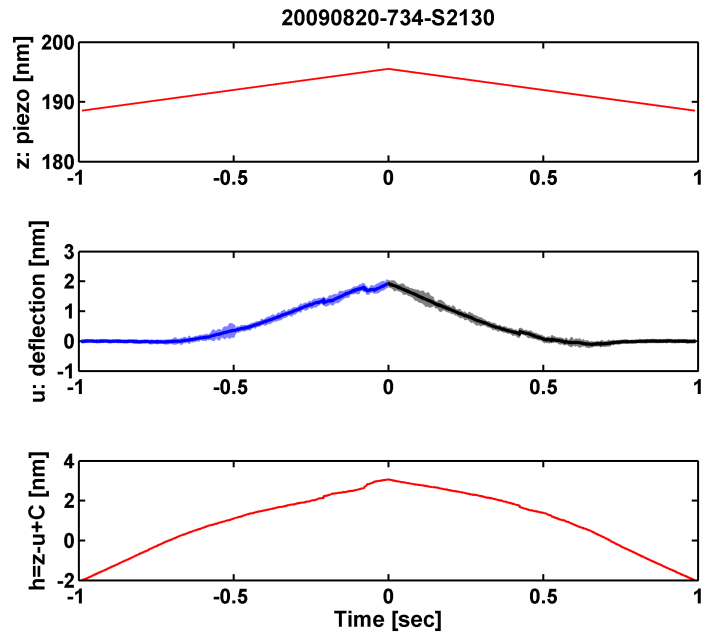
(a) Loading-unloading curve (3 and 4 in figure 5.5.b)
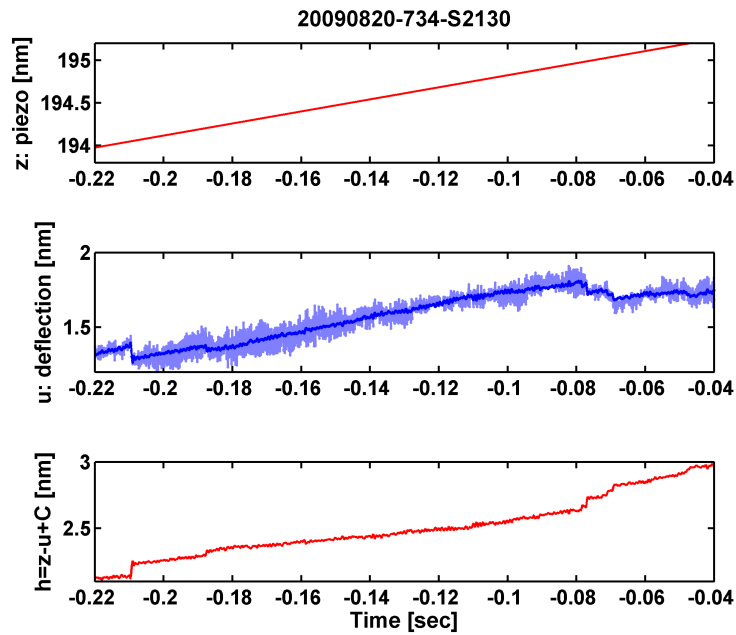
(b) Zoom on discrete events. Pops show up as sudden jumps in deflection (and thus force) and as sudden dips in depth.

Figure 5.6: Open-loop acquisition of nanoindentation data with the FIM/STM/AFM system. Real data with conventions outlined in section 4.3.1.

displacement–control, the load would burst downward while the displacement would remain constant[1]. In the case of a finite stiffness open–loop scheme with a specified z-piezo extension profile, a discrete plastic event would cause the displacement to increase suddenly, and nearly simultaneously, would cause the load to drop by a quantity equal to the product of the displacement burst by the stiffness of the cantilever beam at the contact point (Figs. 5.3 and 5.6).

Our system currently uses no active control of any form, as the load correction is provided automatically by the cantilever beam, which has a finite stiffness and a finite response time, determined by its geometry and material properties. The fast load drop allows for the sensitive detection of discrete plastic events, particularly when the current channel is used in concurrence with load data.
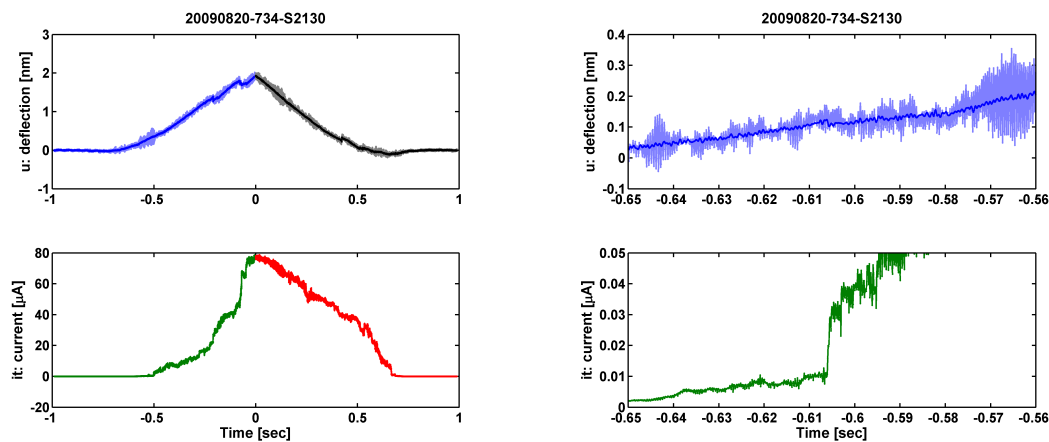
(a) Simultaneous observation of current and deflection over a wide current range (10pA to 80μA).

(b) Jumps in the current trace are very sensitive indicators of electronic/mechanical changes in the junction, and can facilitate the detection of sudden events in the deflection trace. In this case, the deflection jump is on the order of a tenth of an angstrom.

Figure 5.7: Current–based observation of discrete plastic events.

Simultaneous Current Measurement

As highlighted earlier, our STM/AFM acquires the current flowing between the tip and sample simultaneously with the force. In *imaging mode*, the current channel is used for STM feedback, and allows high quality imaging of the sample surface. In the case of the Au(111) surface, this translates to routine imaging of the herringbone reconstruction.

---

[1]This would happen in a time scale determined by the bandwidth of the control system.

During indentation, the current channel is acquired passively over a large range, which is made possible by a logarithmic current-to-voltage converter (see section 2.1.3). Although it is challenging to build a quantitative understanding from the current channel, it is an extraordinarily sensitive indicator of changes in the contact area and in the probed volume. This sensitivity facilitates the detection of plastic events, in addition to providing insights into the electronic transport behaviour of nanocontacts (Fig. 5.7).

Atomic Scale Contact Characterization

Another specificity of our system is the atomic scale characterization of the contact, in terms of both the indenter and the sample, through FIM and STM/AFM respectively. This ensures that the chosen sample location is suitable for testing, and that the indenter geometry is known, which is a requirement for analysis. After indentation, this further allows the characterization of the impression on the surface, as well as the confirmation that the indenter has not been damaged in the process.

5.2 Elastic Contact Models

The field of Contact Mechanics emerged in the late nineteenth century, with the development of a model for frictionless, non-adhesive, elastic contact by Hertz. Since then, much progress has been made in the field, and several approaches have been introduced to account for adhesion.

Common analysis techniques for nanoindentation data are largely based on the Hertz Model. While this is often appropriate for high load instrumented indentation, it can become problematic for metals at the nanometre scale, where adhesion forces can be significant.

Figure 5.8 illustrates the geometry and defines the relevant parameters that will be referred to in the following sections.

5.2.1 Non-Adhesive Elastic Contact

The Hertz model describes the elastic contact between two non-adhesive spheres, and is outlined in detail in the textbook by Johnson [83]. A more recent presentation is also given by Maugis [95].

The model assumes a circular contact area A, with radius a, and provides a
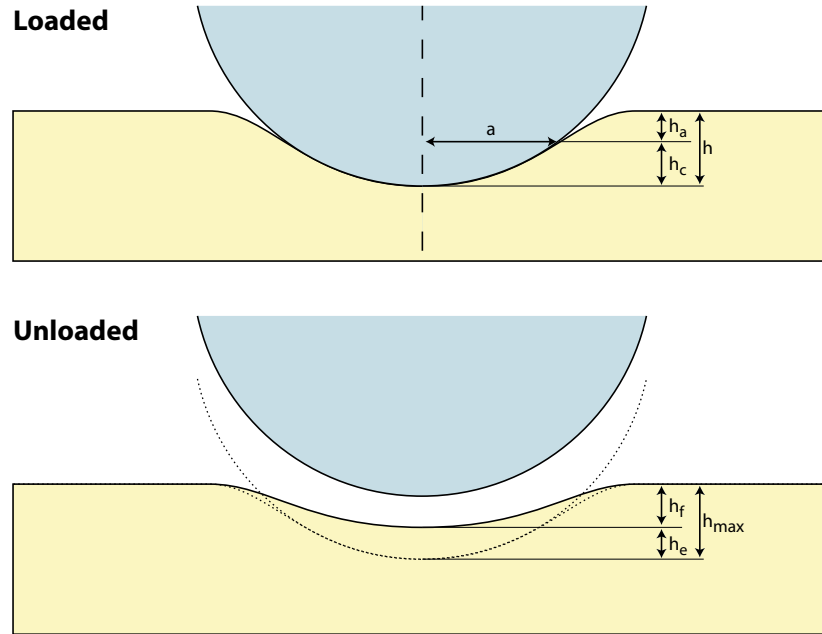
Figure 5.8: Contact geometry and important parameters: when the junction is under load a is the contact radius, h_{c} the height of the circle of contact (also known as the plastic or contact depth) with respect to the deformed surface, and h_{a} the displacement of the contact circle with respect to the original undeformed surface (also referred to as sink-in depth), such that $h = h_{\mathrm{a}} + h_{\mathrm{c}}$. When the maximum load is applied, $h = h_{\mathrm{max}}$. After the indenter is removed, a permanent indent remains at a final depth h_{f}. $h_{\mathrm{e}} = h_{\mathrm{max}} - h_{\mathrm{f}}$ is thus the elastic portion of the deformation during full load. See figure 5.11 for a visual illustration of these quantities on a P-h curve.

solution to the contact problem, as a function of the reduced elasticity modulus[2] E_r and the reduced radius of curvature R_r (or simply R) which are defined by:

$$E_r = \left(\frac{1 - \nu_i^2}{E_i} + \frac{1 - \nu_s^2}{E_s^2} \right)^{-1} \qquad (5.1)$$

$$R = \left(\frac{1}{R_i} + \frac{1}{R_s} \right)^{-1} \qquad (5.2)$$

In equations (5.1) and (5.2), the i index stands for indenter and s for sample. In equation (5.1), ν is the poisson ratio.

It follows from a non-adhesive elastic analysis that the pressure distribution is:

---

[2]In the literature, E_r is sometimes replaced by $\frac{3}{4}K$, as it simplifies the form of the Hertz equations. In the present thesis, $K = \frac{4E_r}{3}$ will be referred to as the modified reduced modulus.

$$p = p_0 \left\{ 1 - \left(\tfrac{r}{a} \right)^2 \right\}^{1/2} = \frac{3p_m}{2} \left\{ 1 - \left(\tfrac{r}{a} \right)^2 \right\}^{1/2} \qquad (5.3)$$

And the total load is therefore:

$$P = p_m \pi a^2 = \frac{2}{3} p_0 \pi a^2 \qquad (5.4)$$

It follows from the Hertz model that:

$$\text{Hertz:} \quad a = \left(\frac{3PR}{4E_r} \right)^{1/3} \qquad (5.5)$$

$$h = \frac{a^2}{R} = \left(\frac{9P^2}{16RE_r^2} \right)^{1/3} \qquad (5.6)$$

$$p_m = \frac{P}{\pi a^2} = \frac{2}{3} \left(\frac{6PE_r^2}{\pi^3 R^2} \right)^{1/3} \qquad (5.7)$$

$$P = \frac{4E_r a^3}{3R} = \frac{4E_r \sqrt{R}}{3} h^{3/2} \qquad (5.8)$$

Equation (5.8) is of great importance to indentation, as it describes readily the relationship between load and displacement in the case of the non-adhesive elastic contact of a sphere on a elastic half-space (in which case $R_s = \infty$ and thus $R = R_i$).

5.2.2   Adhesive Elastic Contact

Adhesion Map

Several approaches have been formulated to describe adhesion in the context of elastic contact of solids. The most well known approaches are the Bradley model, the Derjaguin-Müller-Toporov (DMT) model, and the Johnson-Kendall-Roberts (JKR) model. In short, the Bradley model [96] considers rigid spheres with a Lennard-Jones potential, the DMT model assumes a deformation with a hertzian profile [97] and the JKR model assumes that adhesion forces are active only in the contact area, which is larger than in the hertzian case, and leads to neck formation [98], as illustrated in figure 5.9.

In the presence of adhesion, a finite *pull-off* force is require to remove the tip from the sample, which varies between $2\pi\gamma R$ for the Bradley model and $\frac{3}{2}\pi\gamma R$ for JKR. The differences between the various models were described by Tabor [99] as limiting cases depending on the value of a specific parameter known as the Tabor parameter μ, which is a measure of the ratio of the elastic displacement of the surfaces at pull-off to the range of surface forces:
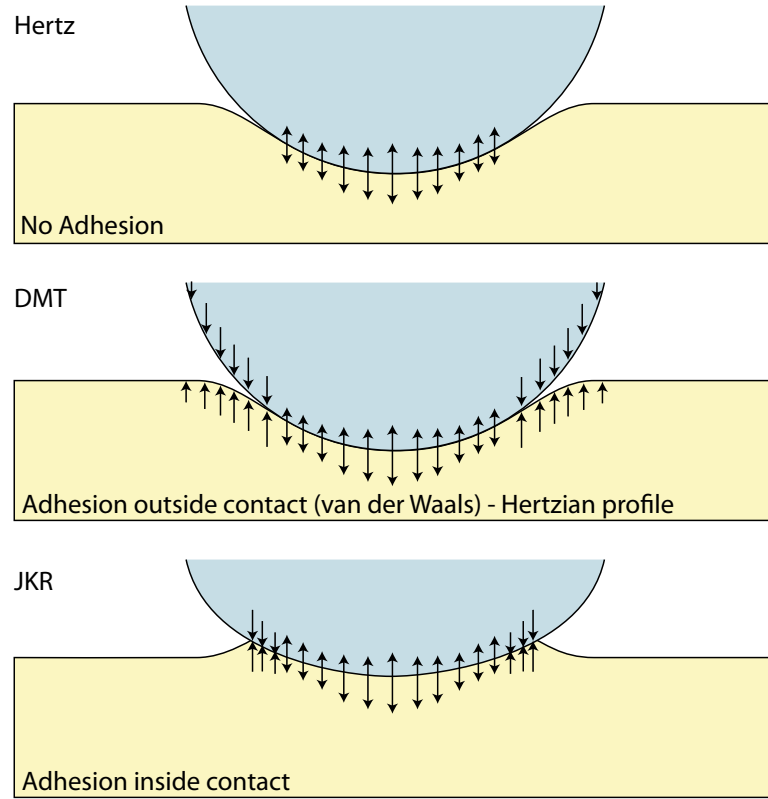
Figure 5.9: Profiles and description of assumptions of the Hertz, DMT, and JKR profiles with regards to adhesion.

$$\mu \;\; = \;\; \left(\frac{R\gamma^2}{E_r^2 \, z_0^3} \right)^{1/3} \quad (5.9)$$

In equation (5.9), z_0 is the equilibrium separation between atoms and w is the work of adhesion, also referred to as γ. The work of adhesion is given by $\gamma = \gamma_i + \gamma_s - \gamma_{is}$, where the first two terms are the surface energies of the indenter and sample respectively and where the third term is the interfacial energy.

Low values of the Tabor parameter (small, rigid spheres) imply a DMT behaviour, while high values (large, compliant spheres) imply a JKR behaviour. Maugis [100] later described the transition between the two by introducing a different formalism, the Maugis-Dugdale (MD) theory.

Johnson [101] conveniently summarised the validity of these various model into a two dimensional adhesion map (Fig. 5.10), which highlights in particular the fact that that all models reduce to the Hertz model when the load is sufficiently larger than the pull-off force.
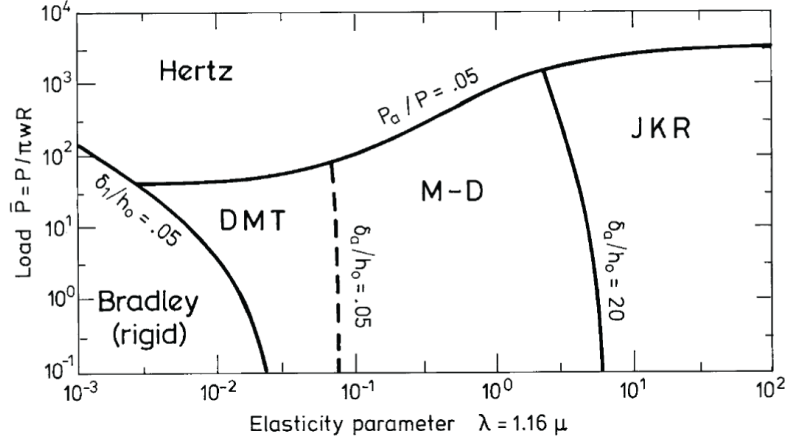
Figure 5.10: Adhesion map. Reprinted from [101]. Copyright 1997, with permission from Elsevier.

Maugis-Dugdale Approach

The Maugis-Dugdale (MD) model was introduced in 1992 [100], and is described in detail in references [95, 101]. In the MD formalism, a number of dimensionless parameters are introduced to described the system, as a function of γ, r, K, and σ_0, the stress (assumed constant) by the Dugdale model over a finite distance [100].

$$A = \frac{a}{(\pi \gamma R^2/K)^{1/3}} \qquad (5.10)$$

$$\overline{P} = \frac{P}{\pi \gamma R} \qquad (5.11)$$

$$\Delta = \frac{h}{(\pi^2 \gamma^2 R^2/K^2)^{1/3}} \qquad (5.12)$$

$$\lambda = \frac{2\sigma_0}{(\pi^2 \gamma^2 R^2/K^2)^{1/3}} \qquad (5.13)$$

In the previous equations, A is the dimensionless contact radius, $\overline{P}$ the dimensionless load, Δ the dimensionless displacement, and λ a transition parameter that can be related to the Tabor parameter by $\lambda = 1.16\mu$.

With this formalism, the adhesive contact is viewed as a crack [95], and the solutions can determined by fitting the data to the following equations, where $m = c/a$ is the ratio between c, which represents the radius within which adhesive force are active and a, and the contact radius.

$$\overline{P} = A^3 - \lambda A^2 \left(\sqrt{m^2 - 1} + m^2 \tan^{-1} \sqrt{m^2 - 1} \right) \quad (5.14)$$

$$\Delta = A^2 - \frac{4}{3} A\lambda \sqrt{m^2 - 1} \quad (5.15)$$

$$1 = \frac{\lambda A^2}{2} \left[\sqrt{m^2 - 1} + (m^2 - 1)\tan{-1}\sqrt{m^2 - 1} \right]$$
$$+ \frac{4\lambda^2 A}{3} \left[\sqrt{m^2 - 1} \tan^{-1} \sqrt{m^2 - 1} - m + 1 \right] \quad (5.16)$$

Hertz

With the MD formalism, the Hertz equations can be rewritten simply as:

$$\text{Hertz:} \quad A^3 = \overline{P} \quad (5.17)$$
$$\Delta = A^2 \quad (5.18)$$

DMT

The DMT model, first published in 1975 [97] and corrected in 1983 [102], provides for the following, in the case of a sphere-plane geometry:

$$\text{DMT:} \quad A^3 = \overline{P} + 2 \quad (5.19)$$
$$\Delta = A^2 \quad (5.20)$$

JKR

In the JKR [95, 98] case, one can write:

$$\text{JKR:} \quad A^3 = \overline{P} + A\sqrt{6A} \quad (5.21)$$
$$\Delta = \frac{A^3 + 2\overline{P}}{3A} = A^2 - \frac{2}{3}\sqrt{6A} \quad (5.22)$$

5.3   Analysis of Nanoindentation Data

In many ways, unloading curves are simpler to analyse and to understand than loading curves. A key assumption that is commonly made is that the behaviour of the system during unloading is entirely elastic. In addition, the vast majority of studies

ignore adhesion and treat the surfaces as smooth. These assumptions contribute to simplifying analysis by applying hertzian results.
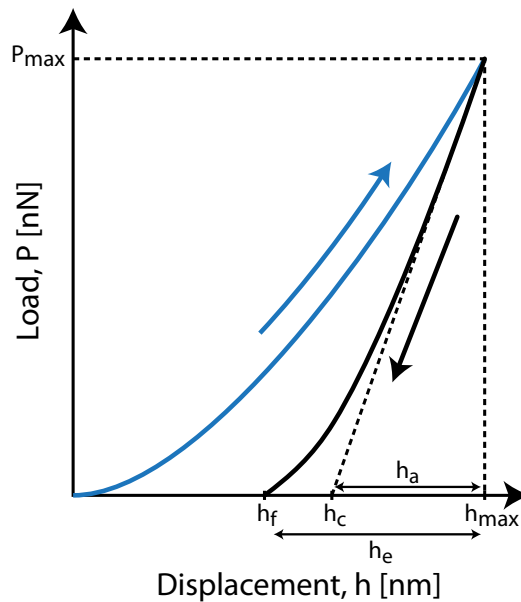
Figure 5.11: This P-h curve highlights important parameters that are used in nanoindentation analysis. See also figure 5.8. While the locations of h_f, h_e and h_max are unambiguous in this drawing, h_c here is extrapolated from the initial slope of the unloading, consistent with the approximation of an effective cylindrical flat punch by Doerner and Nix [103]. Oliver and Pharr [104] have shown that this is a lower bound for h_c, as the unloading curve appears distinctly nonlinear upon close examination. See text for details.

5.3.1 Analysis of Non-Adhesive Nanoindentation Data

For a general introduction to nanoindentation and curve analysis in particular, the reader may refer to the textbook and the two reviews by Fischer-Cripps [85, 87, 88].

From Sneddon to Doerner and Nix

Most of the widespread analysis methods have their foundation in the hertzian-like analysis of Sneddon [105], which provides solutions for the relationship between load, displacement, and contact area for frictionless non-adhesive elastic contact between a smooth profile axisymmetric indenter and a flat surface. In particular, he shows that common indenter geometries produce power law elastic loading curves of the form $P = \alpha h^m$, where $m = 1$ for a cylindrical flat punch, $m = 1.5$ for a

paraboloid of revolution (which a sphere approximates to for low displacements), and $m = 2$ for a cone.

In 1986, Doerner and Nix [103], using a pyramidal Vickers indenter, made the observation that, during the initial stages of unloading, the P-h curve appeared linear. They interpreted their observation by the fact that the contact area at initial unloading was constant, which was consistent with Sneddon's result of a linear P-h curve for a cylindrical flat punch. With an estimate of the contact area, material properties can then be extracted from the curve by using the following expressions for the contact stiffness S, and the hardness H:

$$S = \frac{\mathrm{d}P}{\mathrm{d}h} = 2\,a\,E_r = \frac{2}{\sqrt{\pi}}\sqrt{A}\,E_r \quad (5.23)$$

$$H = \frac{P_{\max}}{A} \quad (5.24)$$

The contact area at full load is then estimated by assuming that during contact, the surface conforms with the rigid indenter up to the plastic (or contact) depth h_{c}, and the area of contact can be obtained by assuming a specific area function for the indenter:

$$A = f(h_{\mathrm{c}}) \quad (5.25)$$

Doerner and Nix then estimate h_{c} by extrapolation, as the intercept of the initial unloading contact stiffness to zero load (Fig. 5.11).

The Oliver and Pharr Method

Oliver and Pharr later established the extended validity of the Sneddon equation[3] (5.23) to virtually all smooth axisymmetric geometries [106], and published an improved method in 1992 [104], which was further refined in 2004 [107]. The approach is an extension of the Doerner and Nix method which replaces the inaccurate approximation of the cylindrical flat punch by an *effective indenter shape* [108]. Oliver and Pharr show that unloading curves are distinctly nonlinear, and that they are best described by a power law (equation (5.26)), where the exponent m is typically between 1.2 and 1.6.

$$P = \alpha\,(h - h_{\mathrm{f}})^m \quad (5.26)$$

In a similar fashion to the Doerner and Nix method, the next step is to estimate

---

[3]It should be noted that in the literature, this formula is often presented as $\frac{\mathrm{d}P}{\mathrm{d}h} = 2\,\beta\,a\,E_r$, where *beta* $\simeq 1$ is a correction factor, in most cases geometric in nature. In this thesis, we will assume $\beta = 1$. For more details, see Refs. [106, 107].

A from h_c, using an indenter area function. However, h_c is determined differently, by using the value for the elastic sink-in h_a calculated by Sneddon, given by equation (5.27), where ε is a parameter that depends on the geometry (table 5.1).

$$h_a = \varepsilon \frac{P}{S} \quad (5.27)$$

The contact depth can then be estimated by subtracting the elastic sink-in depth from the displacement:

$$h_c = h - h_a = h - \varepsilon \frac{P}{S} \quad (5.28)$$

The area function is determined by the known geometry of the indenter, and can be refined by calibration against a sample of known material properties. For our purposes, the most relevant geometries are the sphere, the cone, the rounded tip cone and the cylindrical flat punch. The area functions for these geometries are given in table 5.1.

Geometry	Area function	m	ε
Flat punch	$A = \pi a^2$	1	1
Sphere	$A = 2\pi R h_c - \pi h_c^2$	$\simeq 1.5$	.75
Paraboloid	$A = 2\pi R h_c$	1.5	.75
Cone	$A = \pi \tan^2 \alpha \; h_c^2$	2	$\frac{2}{\pi}(\pi - 2) \simeq .72$
Rounded cone	$A = \pi \tan^2 \alpha \; h_c^2 + 2\pi R h_c$	-	-

Table 5.1: Area functions, m and ε for important geometries. a is the radius of the cylindrical flat punch, R the tip radius of curvature, and α the half-included angle for the cone.

The Oliver and Pharr approach relies on a number of important assumptions, such as the absence of adhesion, of excessive (non-elastic) sink-in, of pile-up and of reverse plasticity. For example, pile-up (the accumulation of material on the side of the indent) will typically lead to an underestimation of the contact area, which can cause E_r and H to be greatly overestimated.

One should note here that nanoindentation hardness is different from traditional hardness as the area used is the contact area, and not the (smaller) residual area.

5.3.2   Material Properties of Nanoscale Au/W Contacts

Before applying the Oliver and Pharr method to the data sets described previously, it should be highlighted that several difficulties arise due to the specifics of our

experimental system. As discussed in chapter 4, the definition of zero-point displacement is difficult and can be ambiguous in our system. This is further complicated by the adhesion observed, as evidenced by negative loads during retraction. Finally, the application of the Oliver and Pharr method required detailed knowledge of the geometry of the indenter, which is assumed to be smooth and axisymmetric. While these limitations are not entirely realistic, it is nonetheless instructive to attempt traditional nanoindentation analyses.

These difficulties in mind, however, alternative methods providing useful insights are explored first, starting with a classic residual impression analysis.

Analysis of unloading data with residual impression

The FIM/STM/AFM system allows us to characterize our specimens at the nanometre scale both before and after indentation. For that purpose, let us return to the P-h curve depicted in figure 4.8. As shown in figure 5.12, the curve was both preceded and followed by STM imaging.
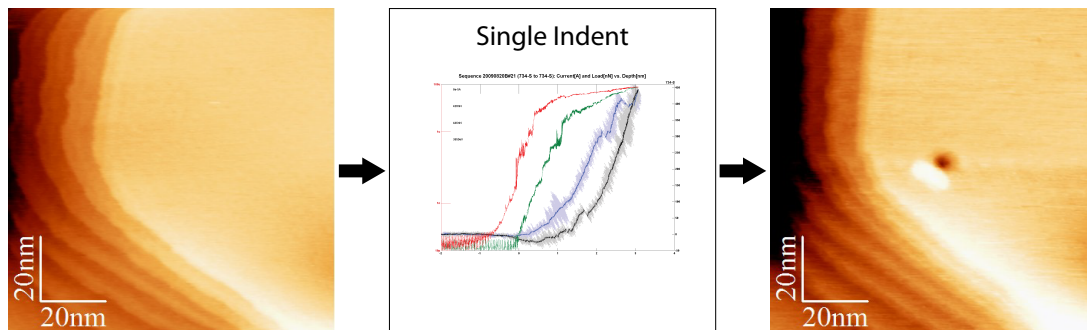
Figure 5.12: Scan/Indent/Scan sequence for curve 20-734. After indentation, an impression is visible, as well as a one-atom high island with the symmetry of the substrate.

The sequence in figure 5.12 can easily be analysed to extract the reduced elasticity modulus and the hardness. Analysis of the post-indentation scan produces a value for the projected area of the residual impression, which can be used to compute the hardness. In addition, the initial unloading slope can be used along with the Sneddon equation (5.23) to obtain the reduced elastic modulus.

The area and the volume of the residual impression and the side hillock were obtaining by *flooding analysis* with WSxM$^{\text{TM}}$[109] (Fig. 5.13), resulting in $A \simeq 21 \pm 2$nm^2. The initial unloading slope was then determined by a first order polynomial fit (Fig. 5.14) to be: $S = 370 \pm 20$N/m.

(a) 3D zoom on residual impression

(b) Residual impression

(c) Area analysis with WSxM™[109]
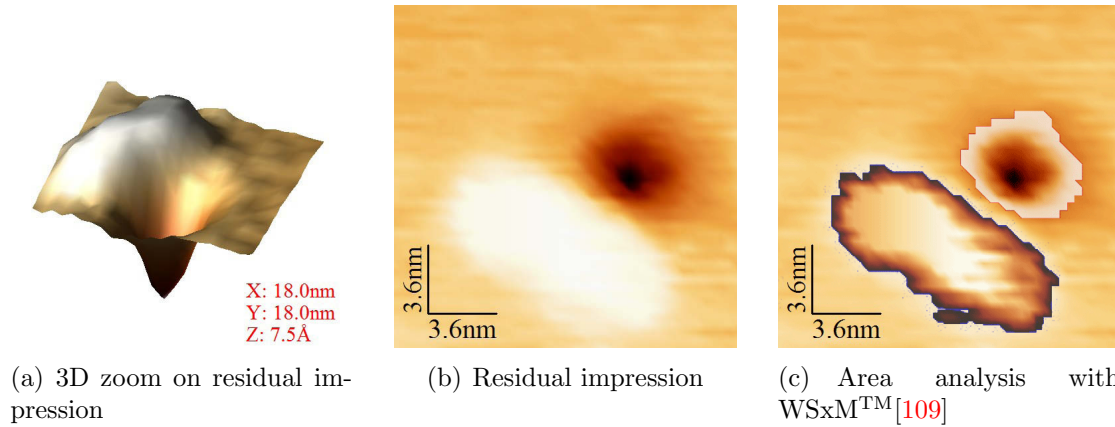
Figure 5.13: Analysis of the residual impression for indentation 20-734. The measured projected area of the residual impression was $A \simeq 21 \pm 2 \text{nm}^2$, with a volume $V \simeq 2.8 \pm 1 \text{nm}^3$. The area and volume measured for the hillock were respectively $53 \pm 5 \text{nm}^2$ and $3.0 \pm 1 \text{nm}^3$.
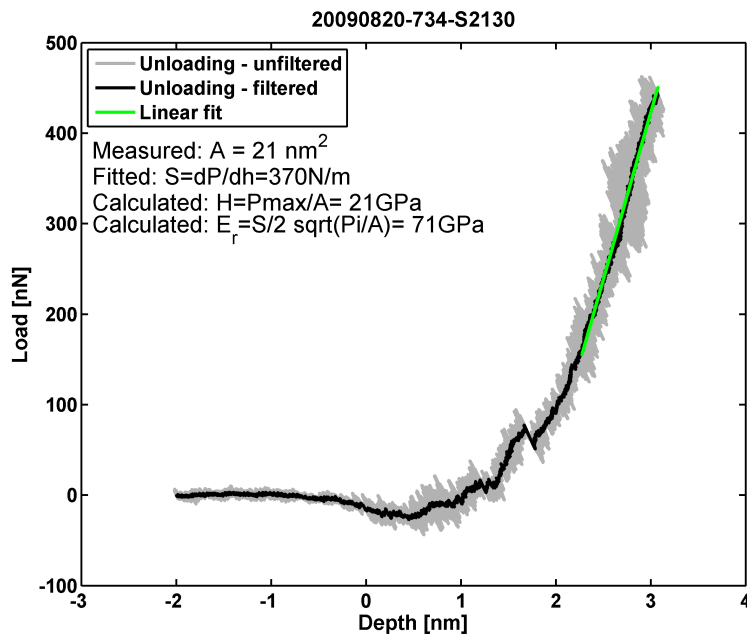
Figure 5.14: The unloading curve for indentation 20-734, leading to a measurement of an initial unloading contact stiffness $S = 370 \pm 20 \text{N/m}$, resulting in a estimation of $E_\text{r} = 71 \pm 9 \text{GPa}$ and $H = 21 \pm 3 \text{GPa}$.

The conclusion from this analysis is that the elastic modulus is very close to what one would expect based on literature values [110, 111]. Indeed, while the elastic modulus of gold in the (111) direction is 117GPa, the indentation modulus is smaller, as the shape of the stress distribution upon loading is such that several directions are probed, and the appropriate modulus turns out to be a weighted average of all

directions [112]. Values between 75-80GPa values are often quoted [76, 80].

In contrast, the hardness is over five times the largest values commonly found in the nanoindentation literature [75, 113–115]. Also worthy of note here, the fact that the surface was pristine, with clear observation of atomic steps and of the herringbone reconstruction.

A convenient feature of this analysis is that it was not necessary to have a precise knowledge of the tip geometry, or an accurate measure of the displacement. The latter is difficult to obtain due the challenges of zero-point determination, especially in the presence of adhesion.

Hardness over elastic modulus square

Another interesting parameter that can readily be obtained from the data, is H/E_r^2. Using equations (5.23) and (5.24), one can easily obtain an expression (equation (5.29)) for this parameter, which is independent of contact area and displacement, and depends only on the peak load and the unloading stiffness, both of which can be extracted directly from the P-h curve. This parameter was investigated by Joslin and Oliver [116], who concluded that it is a measure of the resistance to plastic penetration.

$$\frac{H}{E_r^2} = \frac{4}{\pi}\frac{P_{\text{max}}}{S^2} \quad (5.29)$$

H/E_r^2 is an easy parameter to compute for our nanoindentation data sets. For the primary data set, labelled DS20, from which curve 20-734 was taken, a very clear quadratic relationship appears when the peak load is plotted against the stiffness (Fig. 5.15).

A quadratic fit on the data shown in figure 5.15 yields $P_{\text{max}} \simeq .0027S^2$, where the units of P_{max} and S are respectively nm and N/m. Using equation (5.29), we obtain the following result:

$$\left(\frac{H}{E_r^2}\right)_{\text{DS20}} \simeq .0035 \text{ GPa}^{-1} \quad (5.30)$$

Equation (5.30) implies that for the expected value of E_r=75GPa, $H \simeq 20$GPa. Assuming E_r in the range of 50GPa to 80GPa (the range of most commonly reported values in the literature), the hardness is concluded to be in the range of 9GPa to 22GPaThese hardness values are far above the numbers that have been reported in the literature for gold, which lie in the vicinity of 0.5-4GPa. Looking at H/E_r^2 systematically can in principle provide key insights as to the individual behaviour of
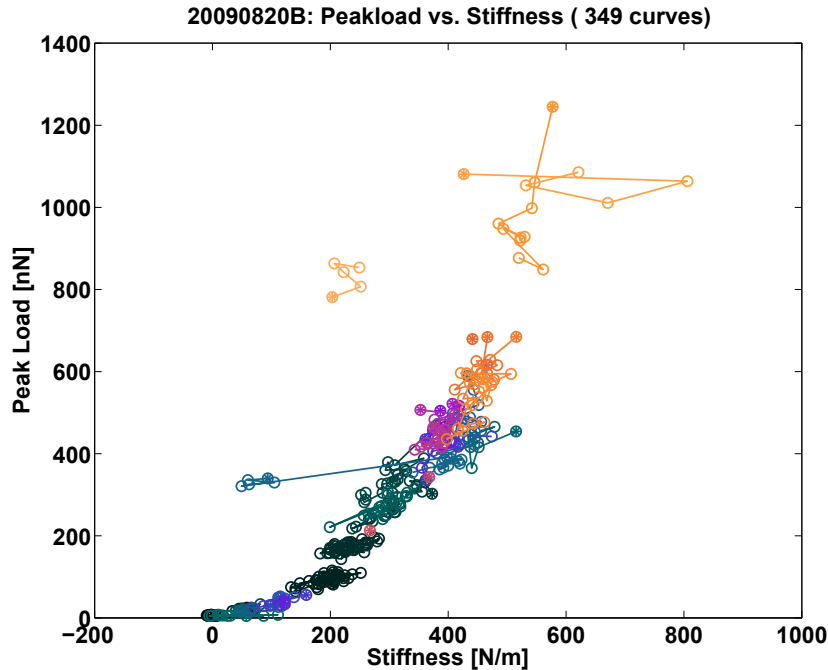
Figure 5.15: The relationship between peak load and contact stiffness is very clearly quadratic for data set 20, which corroborates the validity of equation (5.29). The points that do not fall on the parabola are due to errors in fitting and saturation during data collection.

indentation curves (Fig. 5.16). It can be seen for instance that aside from outliers, the overwhelming majority of points for our primary data set (DS20) fall very close to the .0035GPa^{-1} mark.

It is worth noting here that curve 20-734, which was the object of our previous analysis, behaves the same way as the group. Another interesting point on these curves is that for the majority of them, during cyclic loading (linked dots), the first (starred) point, which corresponds to the first curve in each indentation sequence, is higher in the plot. This suggests that for a pristine surface, the resistance to plastic penetration is higher, which can be understood in terms of motion of preexisting defects.

Although the data is straightforward to interpret in the case of DS20, the situation is very different for data set DS28. While the values for the $\dfrac{4}{\pi}\dfrac{P_{\max}}{S^2}$ parameter are on the same order, some very distinct differences can be seen, which suggests that the equation (5.29) is invalid in the context of DS28. The difference between these data sets is outlined in detail in chapter 6.
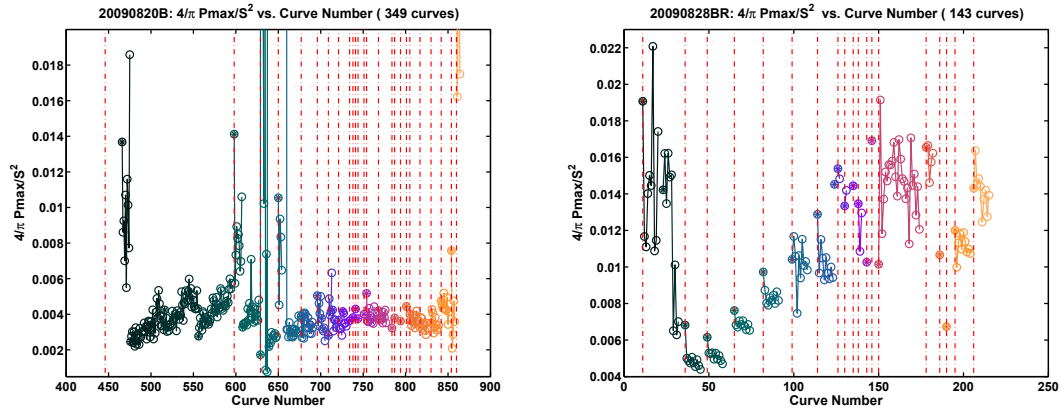
(a) Data set 20

(b) Data set 28
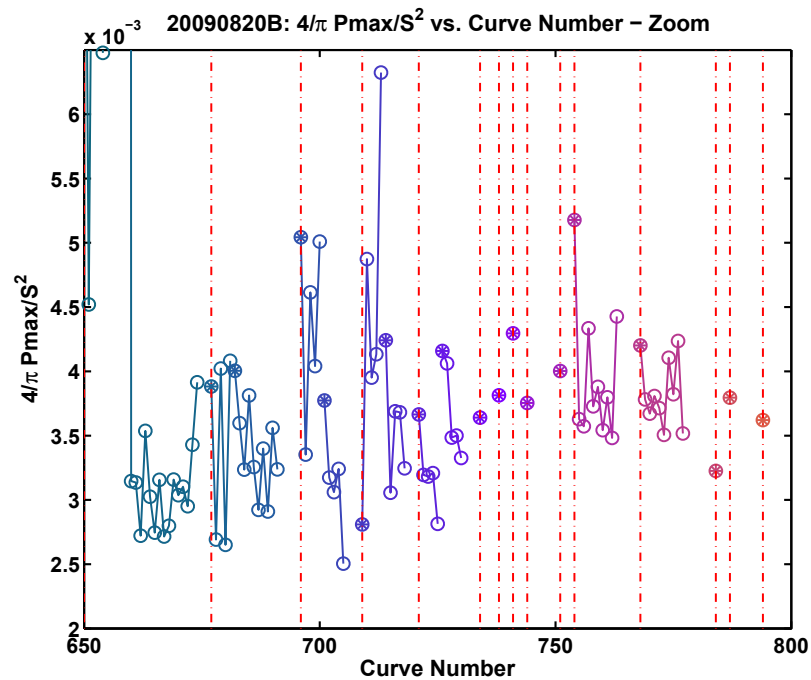
(c) Data set 20 (zoom)

Figure 5.16: H/$E_r$vs. curve number for data sets 20 and 28

Oliver and Pharr Analysis

In order to conduct an analysis of our data using the Oliver and Pharr method, we are faced with two main challenges: accounting for adhesion and estimating the zero-point displacement.

As discussed in section 6.2.2, data set DS20, obtained with a tip of radius 3.6nm, appears to behave in a manner consistent with the DMT model. This implies that

the unloading curves are simply shifted down by a constant force ($P_{\mathrm{po}} = 2\gamma\pi R$), which does not affect the measured value of the stiffness. In addition, the DMT model assumes that the adhesion does not alter the non-adhesive elastic profile, which implies that equation (5.28) for the contact depth should remain valid.

The estimation of the zero-point displacement is made difficult by the fact that gold sometimes adsorbed on the tip will allow electronic contact before the two surfaces start to deform elastically in the sense of contact mechanics. In traditional nanoindentation, the initial displacement upon contact is usually estimated by fitting a smooth curve to low positive values of the load, and extrapolating towards zero load [88]. In our case however, the presence of adhesion complicates matters as contact starts in the tensile zone.

Despite these challenges, an attempt to apply the Oliver and Pharr method was made by adjusting the zero-point displacement graphically. Fitting of the portion of the unloading data between 50nN and 100nN was found to show very little sensitivity
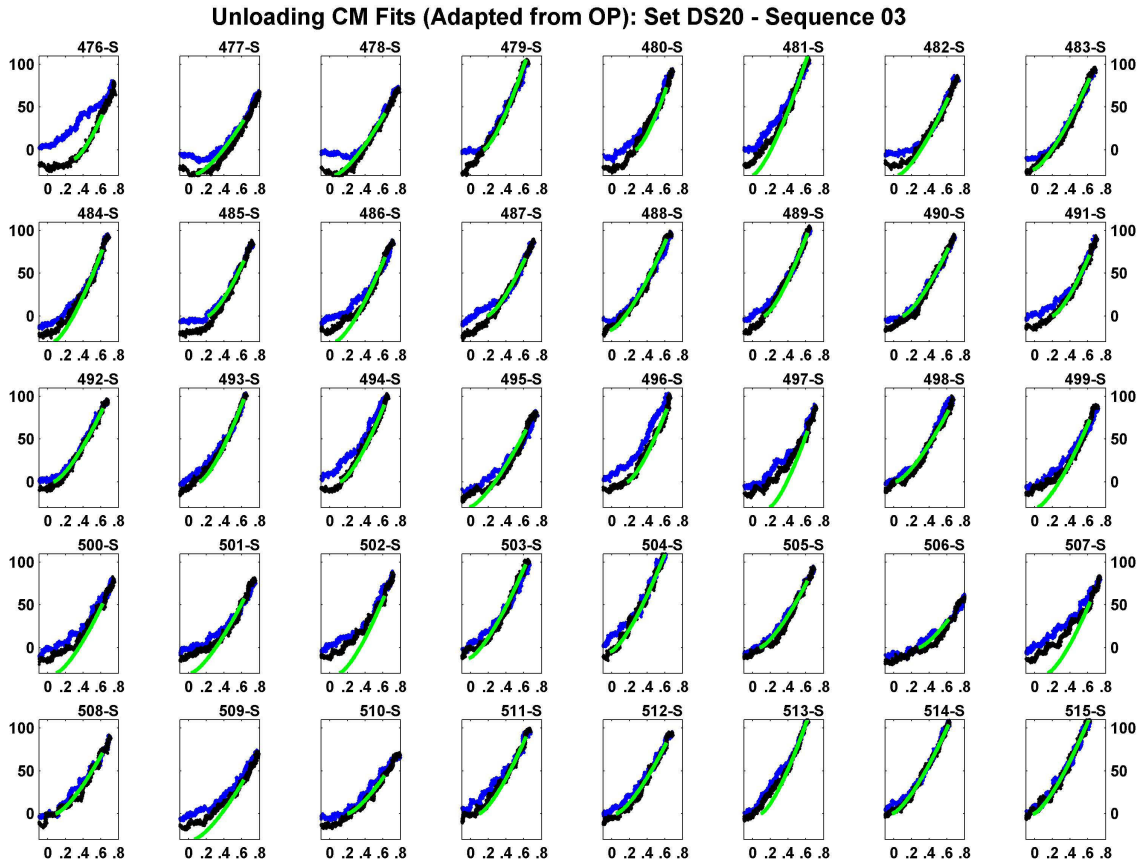
Figure 5.17: Application of the Oliver and Pharr method to curve sequence #03 of data set DS20. The zero-point displacement was adjusted graphically.

to the choice of exponent m. For that reason, m was fixed at the value expected for a spherical indenter in order to improve the quality of the fit, which is later differentiated to obtain a value for S, used to compute h_c and the contact radius based on equation (5.28), where h, S and P are all taken at peak displacement. The unloading fits are shown on figure 5.17, and the resulting values for the mean contact pressure p_m and the reduced elastic modulus are shown in figure 5.18. While the value for E_r is consistent with the literature, the mean contact pressure (which represents a lower bound for the hardness, as it becomes H when the plastic zone is fully developed) is far greater than typical reported results for the hardness. This is discussed further in section 6.3.2.

(a) Reduced Elastic Modulus
(b) Mean Contact Pressure
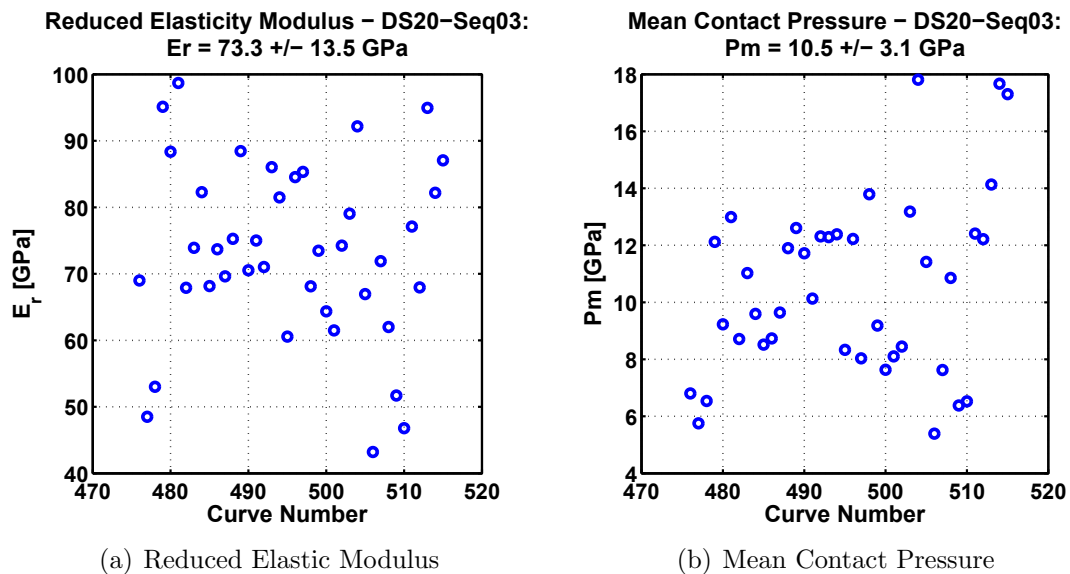
Figure 5.18: Reduced elastic modulus E_r and mean contact pressure p_m for the 40 data curves of sequence #03 in data set DS20. The mean value obtained for E_r is consistent with literature results. p_m however is much larger than reported results for hardness.

The results in this section confirm values of the indentation modulus that are consistent with predictions and literature results. In counterpart, they highlight very high values for the mean contact pressure for DS20, and a behaviour qualitatively different for DS28 and DS02. In order to explore the origin of these differences, the next chapter introduces and applies adhesion models in the context of nanoindentation.

Chapter 6

Plasticity and Adhesion: Effects of Very Sharp Indenters

6.1 Introduction

6.1.1 Loading Curves: Incipient Plasticity

In the previous chapter, we outlined and applied a number of methods for the analysis of unloading curves. Due to the high level of complexity and the variety of behaviours associated with the elastic–plastic transition, detailed analysis of loading curves is difficult. Several authors have presented models to describe its principal features, the most commonly accepted approach being that provided by Field and Swain [117]. Other references of interest include [118, 119].

We leave detailed analysis of loading curves for future work, and only briefly mention some relevant concepts here. As pointed out previously (section 5.1.1), loading curves start off in the elastic regime, following a specific elastic contact mechanics model. Upon loading, the mean contact pressure eventually reaches the yield stress, marking the beginning of the elastic–plastic region.

In classic dislocation theory, yield is attributed to the motion or generation of dislocations [4, 120, 121]. For a perfect crystal, the theoretical yield stress can be calculated by estimating the shear stress needed to move a plane of atoms with respect to another. In figure 6.1, the left drawing shows two planes of atoms. In moving the top row with respect to the bottom row by δx, the shear stress required is given by equation (6.1). In order to move the top row permanently, the *theoretical critical shear stress* is given by equation (6.2), where G is the shear modulus.

$$\tau = \frac{Gb}{2\pi a} \sin \frac{2\pi \delta x}{b} \quad (6.1)$$

$$\tau_c = \frac{b}{a}\frac{G}{2\pi} \simeq \frac{G}{2\pi} \quad (6.2)$$
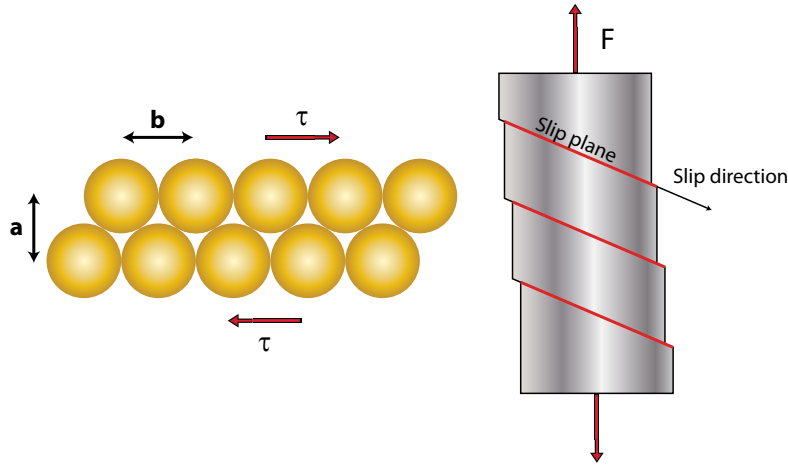
Figure 6.1: Shear Stress and Slip Planes

In a crystal under uniaxial loading, normal stresses become shear stresses when projected along a particular slip plane, and slip of a plane along a given direction can lead to macroscopic deformation, as shown on the drawing on the right of figure 6.1. The combination of slip plane and slip direction is referred to as a slip system. In order to quantify the normal stress required for yield, it is necessary to compute the shear stress *resolved* on the slip plane, in the plane direction. When the *maximum resolved shear stress* on a slip plane exceeds the critical shear stress for the slip system in question then slip will occur, in theory. In nanoindentation, the stress distribution is more complex than in the drawing presented above, and it is necessary to compute the stress distribution to find the maximum resolved shear stress. According to the Hertz model, one finds that the maximum resolved shear stress is $\tau_{\text{max-Hertz}} \simeq .47 p_m$ [122]. In the Au(111) surface, for accurate prediction, is it necessary to take into account the crystal structure, and resolve the shear stress on the crystal's primary slip systems, which are those associated with close-packed planes. In the case of the indentation of the Au(111) surface, the maximum resolved shear stress is equal to $\tau_{\text{max-Au(111)}} \simeq .228 p_m$ [76]. At the atomic level, atomistic simulations [78, 123] provided interesting insights and allow the visualization of events taking place.

As described in section 5.1.1, continued loading will cause the mean contact pressure to increase until it reaches a limiting value, the hardness. In the next section, a brief review of the literature on gold nanoindentation, particularly as it relates to
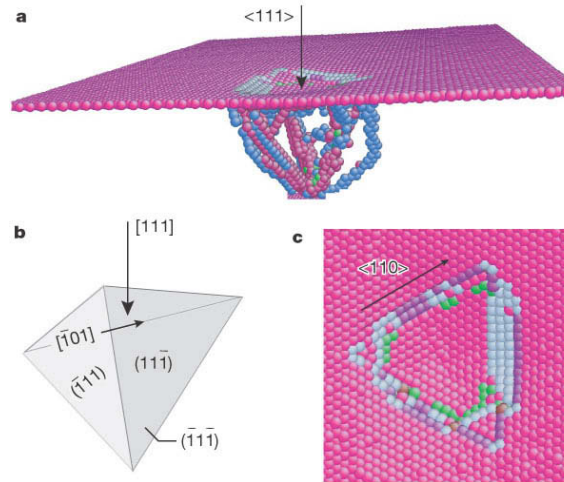
Figure 6.2: Atomistic simulation of incipient plasticity in Al(111). Reprinted by permission from Macmillan Publishers Ltd: Nature [123], copyright 2002.

the hardness of Au(111), will be presented.

6.1.2   Hardness of Gold in the Literature

The material properties of gold have been tested and studied for centuries. In the late 1960's, Gane *et al.* [124, 125] noticed that microindentation of gold featured a threshold load above which a permanent indent was left. Gane *et al.* also noticed a size effect according to which the microhardness of gold increased as indentation depth was reduced and as sharper indenters were used. The range of measured values reported were approximately between 200MPa and 1GPa. Pethica and Tabor [126] later noted that some of these results were partly due to the presence of an oxide layer on the surface, which, when reaching 5nm in thickness, can produce a tenfold increase in hardness. Nevertheless, Pethica *et al.* confirmed in 1983, in a seminal publication introducing results with depth sensing indentation [114], that a scaling effect does exist for reduced indentation depth measuring an increase of hardness to 2GPa for an indentation depth of 20nm. These studies were both carried out in air and at room temperature.

In 1989, Burnham and Colton [127] used a modified Atomic Force Microscope (AFM) in air to measure the hardness of a gold foil with a tungsten tip, which was imaged by Scanning Electron Microscopy (SEM). The tip had a radius of curvature of 100-200nm, with a cone half opening angle of approximately 10°. However, due to inexact knowledge of the tip geometry, the authors assume a cylindrical flat punch of 100nm radius. At a depth of 100nm, they obtain a hardness of 650MPa, which is

higher than the bulk value they quote as 500MPa.

In 1997, Corcoran *et al.* [75] obtained a set of load controlled load-displacement curves for Au(111), Au(100), and Au(110). These experiments were conducted in air under nitrogen flow, using a blunt-ended Berkovich diamond indenter (spherical radius of 205nm), and a commercial Hysitron Picoindenter with a Nanoscope III AFM/STM by Digital Instruments.

With displacements below 50nm, loads below 100μN, and a resolution on the order of 1μN and 1nm for load and displacement respectively[1], the authors note a succession of primarily elastic loading segments punctuated by discrete yielding events. Using a great deal of approximate assumptions, they obtain a value of the mean contact pressure at the first yield event of approximately 3GPa for Au(111) and 8.5GPa for Au(100).

Perhaps the most comprehensive experimental investigations of the indentation of gold single crystals, are the two studies from the group of J.E. Houston that were published at the end of the 1990s. In the first, Michalske and Houston [122] examined the nanoscale elastic and plastic responses of Au(111) single crystals using an Interfacial Force Microscope (IFM) [128]. To avoid adhesion between the tip and sample, the latter is passivated with a self-assembled layer of thiols [129]. The tips are electrochemically etched with radii of 50nm, 80nm, 125nm, and 250nm. A size effect is found, with yield stress increasing for sharper tips. The maximum resolved shear stress $\tau_m = .47p_{\mathrm{m}}$ is computed based on continuum elastic theory, and found to increase from 3GPa to 5GPa approximately as the tip radius is reduced. The effect is explained qualitatively using energetic arguments with respect to the generation of dislocation loops, but quantitative agreement is not met, as the critical resolved shear stress is expected to be smaller by a factor of 3.5.

The second study, by Kiely and Houston [76], also involves thiol passivated surfaces, this time of Au(111), Au(001), and Au(110). Using electrochemically etched tips of radii 70nm and 175nm, the elastic modulus and the yield stress is extracted. With a measured mean contact pressure at yield of 7.3GPa, a corresponding maximum resolved shear stress is calculated in the Au(111) orientation, for the $(1\bar{1}\bar{1})[\bar{1}0\bar{1}]$ slip system of $\tau_m = .228p_{\mathrm{m}} \simeq 1.7$GPa, which the authors conclude is the critical resolved stress. The same result is obtained for other orientations. Inspection of the load–displacement curves of the paper show no significant increase in mean contact pressure after numerous yield events.

A related theory paper by Kelchner *et al.* [78] published results of simulations of

---

[1]Estimated by visual inspection of the provided graphs.

indentation on Au(111) without adhesion, to mimic the passivated surface of Kiely *et al.* [76]. However, due to computer requirements, the tip radius was kept to 8nm, an order of magnitude sharper than in the experimental counterpart. In this care, a critical resolved shear stress of 7GPa was obtained.

In 2006, Asenjo *et al.* [79] used an AFM in ambient conditions, with a silicon cantilever having a tip radius of 10nm according to the supplier. After performing 16 indents of increasing depth in a 4x4 array with a separation of $\sim$ 500nm between neighbouring indents, the authors show resulting surface damage, including emerging screw dislocations. Similarly to Corcoran *et al.* [75] and Kiely *et al.* [76], the initial elastic response is fitted to the Hertz model, which allows the estimation of the mean contact pressure at yield. However, the elastic moduli that were computed were abnormally low, which the authors attributed to a combination of inaccuracy of the tip radius and lower elastic moduli near the surface. This led to a high degree of uncertainty with regard to the maximum resolved shear stress at yield, between 2 and 6.5GPa. It is also worth noting that the curves show a clear slope reduction after the yield event, which is a signature of elastic–plastic behaviour. This could be due to motion of generated or preexisting defects on the surface, as discussed in section 6.2.2. Interaction between the indents is also difficult to rule out, since damage as seen from the provided AFM image can extend hundreds of nanometres away from indents. Finally, a mild JKR type of adhesion between the tip and surface could also explain the lower elastic loading slope.

In 2008, Lucas, Gall, and Riedo [113] investigated the Au(111) surface with two tips of radii 70nm and 112nm, silicon and diamond coated silicon respectively. Hardness measurements were made by analysis of the residual impression and of the tip before and after each indentation using a silicon calibration grating. As illustrated in figure 6.3, the values of hardness obtained show a size effect between a value of 500MPa (bulk value) for indentation depths above 40nm and up to 4-5GPa for the sharper tip at small depths.

Finally, although the yield stress and hardness are not examined explicitly, the study by Cross *et al.* [80] is the first to explore the mechanics of the Au(111) surface in a clean environment, and with a sharp and atomically defined tip ($\sim$ 3nm). These investigations were carried out using our experimental system in the late 1990's, before several improvements were made in terms of stability, imaging, force resolution, and range of measurable currents.

A number of recent studies have been carried out on different gold surfaces. Particularly worthy of note is that by Smallwood and Unertl [130, 131], which also
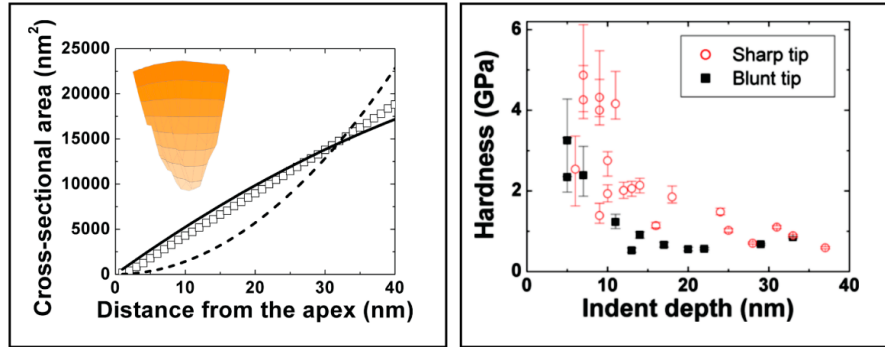
Figure 6.3: See text for details. Reprinted with permission from [113]. Copyright 2008, American Institute of Physics.

combined Field Ion Microscopy (FIM) and STM/AFM in Ultra High Vaccum (UHV). The surface of interest was Au(110), and the indenters were W(110) tips with radii between 5-25nm. Other studies were performed by Carrasco *et al.* [132–136] on the Au(100) surface, which detailed imaging of defects produced by indentation. Other studies of gold surfaces have also been carried out with an emphasis on thin films, and for example, the effect of the substrate and film thickness [115, 137–139].

6.2   Nanoindentation Data with Adhesion: Contact Models

Several approaches have been explored by researchers to account for adhesion in nanoindentation and related measurements. A number of them rely on fitting the elastic portion of the load–displacement data based on a Maugis-Dugdale model (MD) [140, 141].

Sirghi and Rossi [142, 143] modelled adhesion assuming that *"Sneddon's solutions [105] for the shape of the sample surface outside the contact area are not considerably affected by adhesion"*. Assuming further that the adhesion energy is the product of the work of adhesion γ by the total surface area at the interface between the indenter and the sample, they derive a theoretical expression for load–displacement curves for a conical geometry. While these assumptions are highly unrealistic for a JKR type of contact, this approach is useful in modelling data in cases when the Hertzian profile is not greatly affected by adhesion, such as DMT contacts.

The solutions put forth by Sirghi and Rossi are targeted towards AFM experiments with conical tips. Before attempting to apply this method to our data, it is useful to review some aspects related to our tip geometry.

6.2.1   Small Opening Angle Rounded Cone Tip Geometry

The W(111) tips used in our studies are characterized by FIM, which allows for the extraction of a radius of curvature for the tip apex. Further study of our tips by SEM reveals a conical shape with a half opening angle α between 5 and 10 degrees. In order to calculate the elastic and the adhesion energies, the first step is to obtain an expression for the axisymmetric tip profile: $z = z(\rho)$ in cylindrical coordinates.

Tip Geometry

The area function of the tip, given by the rounded cone expression in table 5.1, can serve as a starting point. Using $A = \pi(\tan^2 \alpha h_c^2 + 2Rh_c)$, one obtains equation (6.3).

$$z = \frac{R}{\tan^2 \alpha} \left(\sqrt{1 + \left(\frac{\tan \alpha}{R}\rho\right)^2} - 1 \right) \qquad (6.3)$$

In our experiments, $\rho \tan \alpha / R \ll 1$, since the contact radius is small compared to $R/\tan \alpha > R_{\min}/\tan \alpha_{\max} \simeq 20$nm: for our smallest tip ($R = 3.6$nm, and assuming a maximum value for α of 10 degrees). Equation (6.3) can then be re-written using a Taylor expansion[2]. Keeping only the first two terms, this leads to equation (6.4). In our case, the tips have a small enough cone opening angle that only the first term can be kept, effectively reducing the problem to a paraboloid (Fig. 6.4), described by equation (6.5).

$$z \simeq \frac{1}{2R}\rho^2 - \frac{\tan^2 \alpha}{8R^3}\rho^4 \quad (6.4)$$

$$z \simeq \frac{1}{2R}\rho^2 \qquad\qquad (6.5)$$

---

[2] $\sqrt{1+x} = 1 + x/2 - x^2/8 + ...$
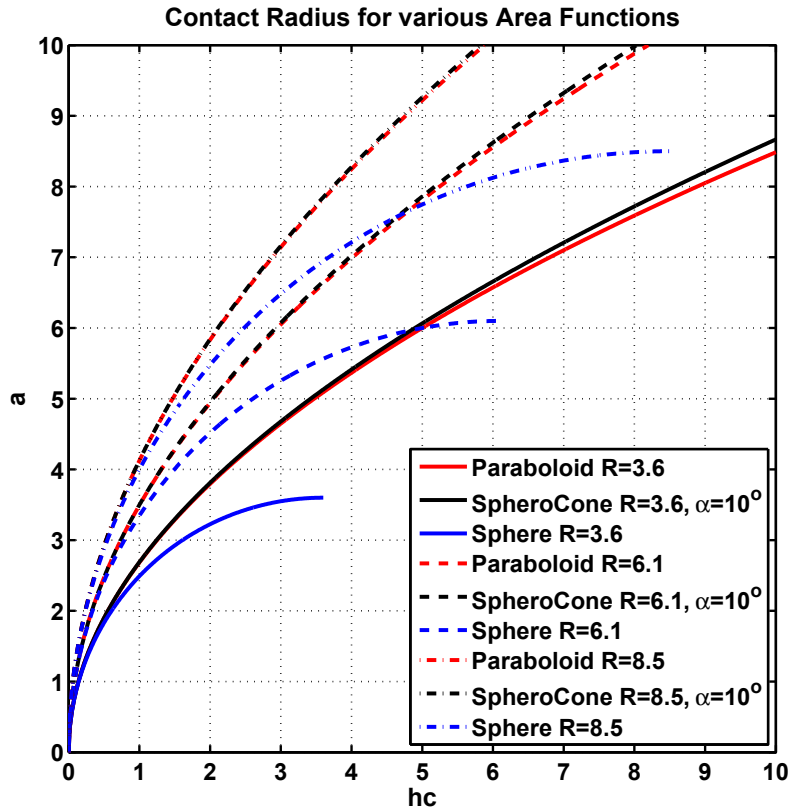
Figure 6.4: Contact radius for various area functions applicable to our tips. The three tip radii illustrated here (3.6nm, 6.1nm and 8.5nm) correspond to tips used in our most relevant data sets (respectively data sets DS20, DS28 and DS02. See section 4.2.3). All dimensions are in nm.

Elastic Deformation

The elastic response of a sample elastically deformed by an indenter with the shape specified earlier can be obtained by carrying out an analysis based on the results of Sneddon [105]. Writing $z(\rho)$ as a power series (equation (6.6)), one can derive expressions for the load P, displacement h and elastic sink-in h_a using Sneddon's results[3]. The expressions obtained (equations (6.6-6.10)) are given as a function of the contact radius a, and Γ is the gamma function.

$$z = \sum_{n=1}^{\infty} c_n \rho^n \tag{6.6}$$

$$P = 2aE_\mathrm{r}\sqrt{\pi} \sum_{n=1}^{\infty} \frac{n}{n+1} \frac{\Gamma\left(\dfrac{n}{2}+1\right)}{\Gamma\left(\dfrac{n}{2}+\dfrac{1}{2}\right)} c_n a^n \tag{6.7}$$

$$h = \sqrt{\pi} \sum_{n=1}^{\infty} \frac{\Gamma\left(\dfrac{n}{2}+1\right)}{\Gamma\left(\dfrac{n}{2}+\dfrac{1}{2}\right)} c_n a^n \tag{6.8}$$

$$h_\mathrm{a} = u_z(a,0) = \frac{2}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{\Gamma\left(\dfrac{n}{2}+1\right)}{\Gamma\left(\dfrac{n}{2}+\dfrac{1}{2}\right)} c_n a^n [\arcsin(1) - j_n] \tag{6.9}$$

$$j_n = \int_0^1 \frac{t^n}{\sqrt{1-t^2}} dt = \frac{\sqrt{\pi}}{n+1} \frac{\Gamma\left(\dfrac{n+1}{2}+1\right)}{\Gamma\left(\dfrac{n+1}{2}+\dfrac{1}{2}\right)} \tag{6.10}$$

Writing $z(\rho)$ in the form (6.6), with $c_2 = 1/2R$, $c_4 = -\tan^2\alpha/8R^3$, and $c_n = 0$ otherwise, one can use equations (6.6-6.10) to derive expressions for our tips' small opening rounded cone geometry (equations (6.11-6.13). Keeping only the first term in each of these equations, we recognise the expressions of P, h and h_a for a paraboloid punch with a radius of curvature R at the apex. These expressions are identical to those of the Hertz model, which are valid for spheres with $a \ll R$.

---

[3]The solution obtained in this case is referred to as the Segedin Formula in the literature. It should be noted however that the final expression for the load given by Sneddon (in reference [105]) for a power series is missing a factor of 2/(n+1). The correct formula is given in provided by Woirgard [144], and can be derived using Sneddon's general formula and equation (6.10).

$$P = \frac{4E_{\mathrm{r}}}{3}\frac{a^3}{R} - \frac{8E_{\mathrm{r}}}{15}\frac{a^5}{R^3}\tan^2\alpha \quad (6.11)$$

$$h = \frac{a^2}{R} - \frac{a^4\tan^2\alpha}{3R^3} \quad (6.12)$$

$$h_{\mathrm{a}} = \frac{a^2}{2R} - \frac{5a^4\tan^2\alpha}{24R^3} \quad (6.13)$$

Adhesion Model for a Paraboloid Tip with a Sneddon Elastic Profile

Having concluded from the previous analysis that our tip geometry is well approximated by a paraboloid of revolution (and for small displacements by a sphere with $a \ll R$), we can derive an expression for the adhesion force based on analogous assumptions to Sirghi and Rossi [142, 143], namely that the Sneddon solutions derived earlier are not affected by adhesion at first approximation. With that assumption, one can write the total adhesion energy $W_{\mathrm{adh}} = \gamma A_{\mathrm{lat}}$, where γ is the work of adhesion and A_{lat} the lateral surface area at the interface: $A_{\mathrm{lat}} = \frac{\pi a}{6h_{\mathrm{c}}^2}\left[(a^2 + 4h_{\mathrm{c}}^2)^{3/2} - a^3\right]$.

Using the fact that $h_{\mathrm{c}} = h - h_{\mathrm{a}} = \frac{a^2}{2R}$, the lateral surface area simplifies to $A_{\mathrm{lat}} = \frac{2\pi R^2}{3}\left[\left(1 + \frac{h}{R}\right)^{3/2} - 1\right]$, and the adhesion force P_{adh} can be written:

$$P_{\mathrm{adh}} = -\frac{\mathrm{d}W_{\mathrm{adh}}}{\mathrm{d}h} = -\gamma\pi R\sqrt{1 + \frac{h}{R}} \quad (6.14)$$

6.2.2   Choice of Adhesion Model

In the previous sections, we introduced a number of different contact models that account for adhesion. The choice of which model to use depends primarily on the tip geometry, the material properties of the sample being tested, and the range of possible interactions between tip and sample. Here, we investigate a clean surface of Au(111) characterized at the sub-nanometre scale in UHV with a clean W(111) tip atomically cleaned and characterized by FIM. Taking these two facts into account, the best adhesion model will be primarily determined by the tip geometry.

Inferring the tip geometry from FIM for our three main tips (see section 4.2.3 and figure 4.4), and assuming literature values for the surface energy for the Au-W interface $\gamma = 1J\ m^{-2}$, and $E_{\mathrm{r}} = 75\mathrm{GPa}$ [80], one can draw expected load–displacement curves for the different adhesion models for each data set.

Data Set DS20
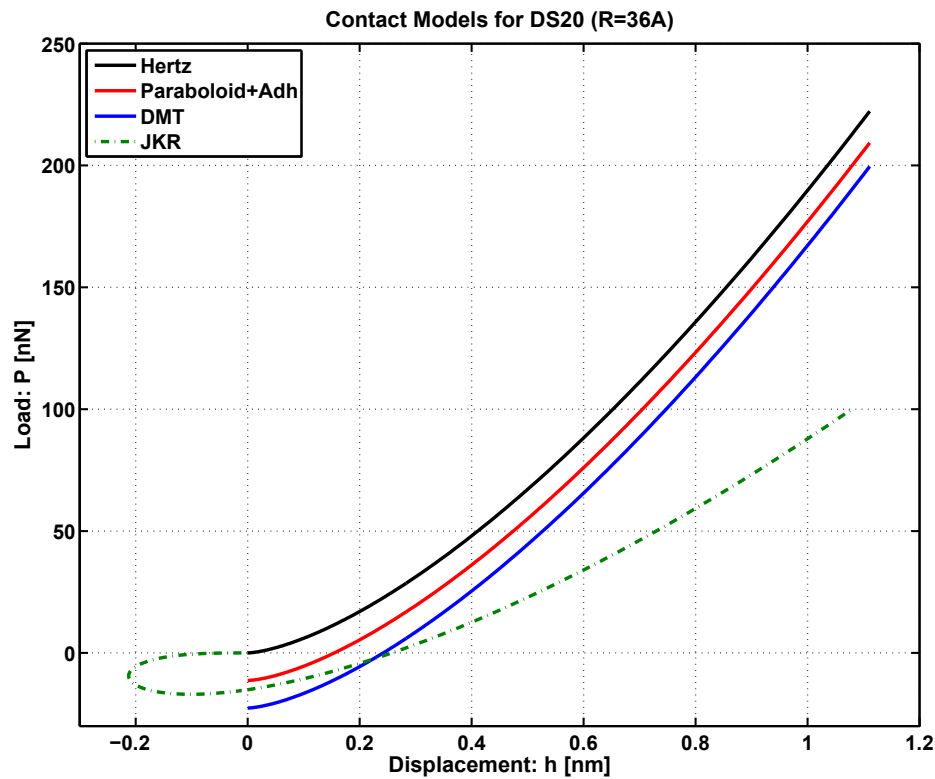
Figure 6.5: Various contact models with adhesion for data set DS20, for a tip of $36\mathring{A}$. The paraboloid with adhesion trace corresponds to the assumption explored in section 6.2.1, of adhesion in the contact area only, with a maintain Sneddon elastic profile (equations (6.11-(6.14)).

Data set DS20 was obtained with Tip20 which has a radius of $R = 3.6 \pm 1.2$nm as determined by interpretation of FIM images (section 4.2.3). Figure 6.5 shows the various adhesion models discussed earlier plotted in a load–displacement graph, which illustrates clearly the significant difference between the models that assume a non-adhesive elastic profile (Hertz, Sneddon Paraboloid with Adhesion and DMT, all represented in solid lines) and the JKR model (discontinued line), which assumes an altered profile with a larger contact radius due to adhesion (Fig. 5.9). In the JKR case, the larger contact radius upon initial contact happens in the negative displacement region as a neck is created.

In order to identify the most appropriate model for the curves in data set DS20, a selection of representative curves were plotted on a graph similar to that of figure 6.5. For that purpose, offsets were applied to the curves along the h-axis to account

for the zero-point in the contact models being different from the 100pA tunnelling setpoint used as a reference in our curves. The results, shown in figure 6.6, indicate that the contact models most appropriate to describe the data are those that assume little to no change to the non-adhesive elastic profile. These models differ mostly by small vertical offsets (a small and slowly varying difference in the case of equation (6.14)), and the effect on the curve slopes is very small. For this reason, stiffness based methods of analysis, such as the Oliver and Pharr method described in section 5.3.1, should still be largely applicable (see section 5.3.2).
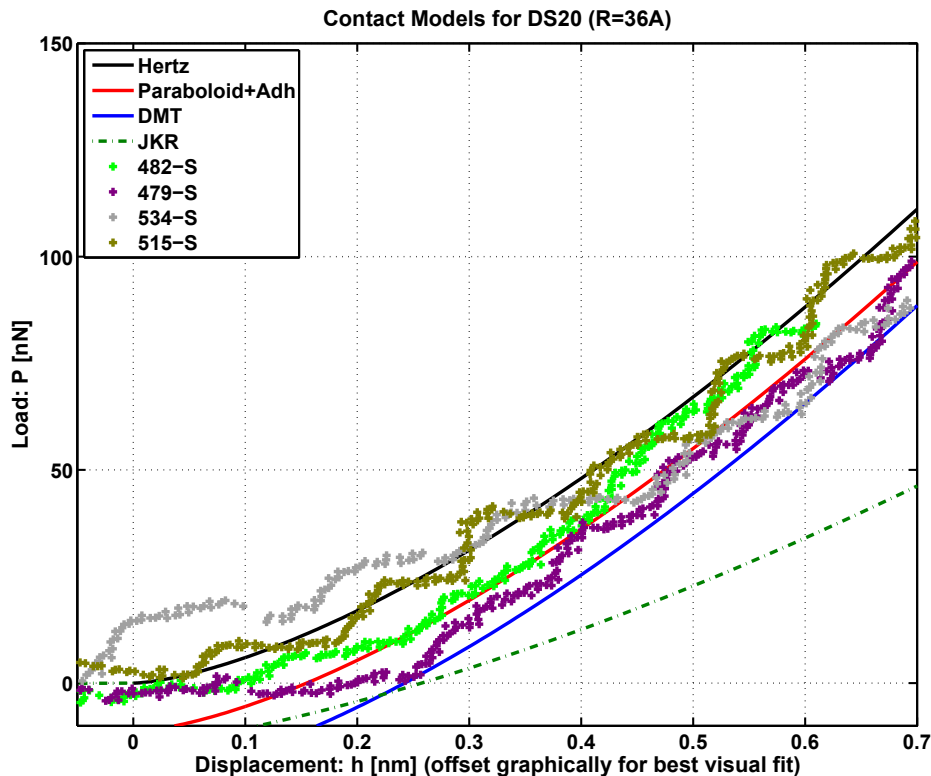
Figure 6.6: Visualisation of the elastic region of representative curves from DS20, and comparison with various contact mechanics models.

Data Set DS28

Data set DS28 was obtained with Tip28 of radius $R = 6.1 \pm 1.2$nm (see section 4.2.3). Figure 6.7 shows that the curves in DS28 appear to be far better represented with the JKR model than with the models that assume no changes to the non-adhesive elastic profile, which is in stark contrast to DS20. This shows that as the tip radius increases, we are transitioning from DMT behaviour towards JKR. This is

consistent with the adhesion map, as a larger radius implies a larger Tabor parameter (see equation (5.9) and figure 5.10).

Figure 6.7: Visualisation of the elastic region of representative curves from DS28, and comparison with various contact mechanics models.

A number of other interesting features can be observed on the selected curves. For example, some of them show discrete plastic events that are sometimes followed by a reduced slope (e.g. curve 023), while some continue with the same slope (e.g. curve 082). This could be explained by elastic–plastic behaviour vs. strictly elastic behaviour past the discrete plastic event. The former may be caused, for instance, by pre-existing defects in the probed region that could facilitate the generation and mobility of other defects. Other curves appear to have a lower slope initially, and only approach the JKR curve after a certain load (e.g. 049). This may be explained by the possible existence of a soft layer of gold on the tip that is pushed away before the actual sample is probed.

Data Set DS02

Data set DS02 was obtained with the largest of our three tips, Tip02, of radius $R = 8.5 \pm 1.2$nm. Based on the previous discussion, we would therefore expect a larger Tabor parameter, and clear JKR-like behaviour. Figure 6.8 confirms that this is indeed the case.
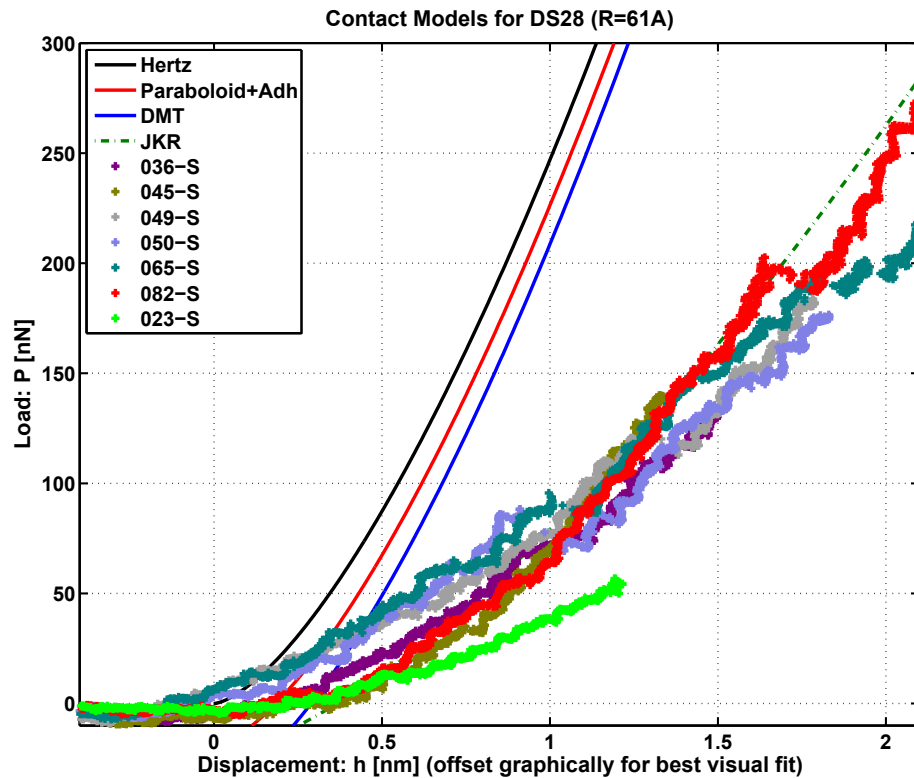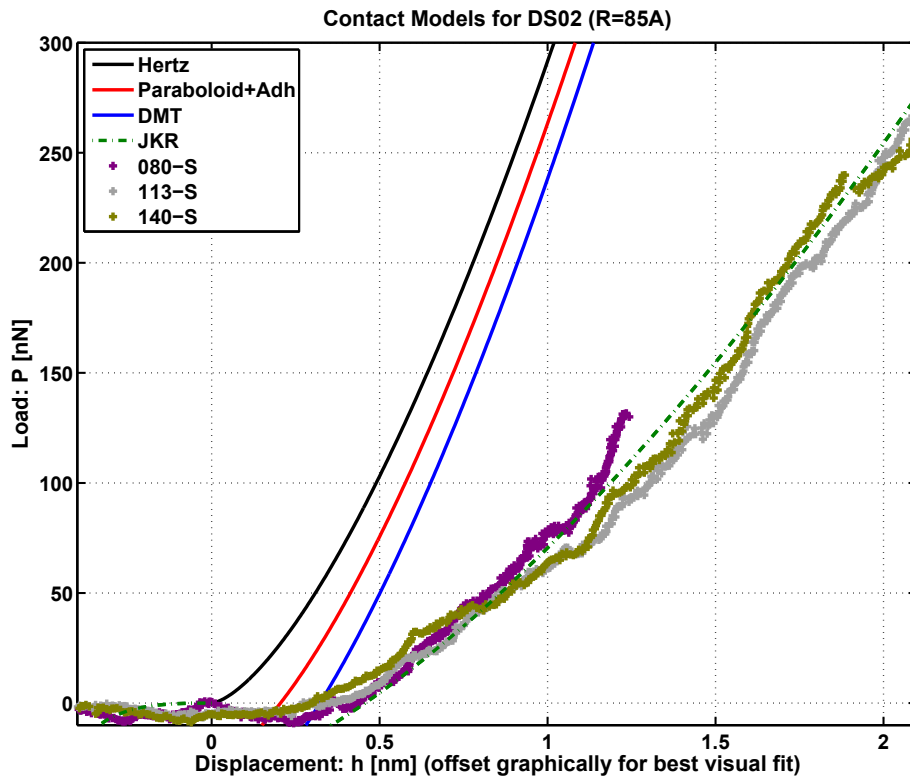

Figure 6.8: Visualization of the elastic region of representative curves from DS02, and comparison with various contact mechanics models.

6.2.3   Summary

In the previous discussion, we showed that the contact mechanics behaviour of our indenter–sample system could be described, to a certain extent in terms of adhesion models such as DMT and JKR. It is important to note that these models provide approximate descriptions only. Important deviations from these models are observed and expected.

◇ Contrary to assumptions made in developing common contact mechanics models, the small scale of the tips implies significant surface roughness, which could

account for changes of contact areas and stresses by a factor of two according to Luan and Robbins [145].

◇ The tip geometry is determined by FIM and in some cases by SEM. The former provides atomic scale detail of the apex and an estimate of the tip radius. The latter demonstrates that our tips are conical at a larger scale. While the two methods provide complementary information, the tip geometry in between those two scales can only be inferred. Assuming it is paraboloidal allows us to rely on the contact models described previously[4]. However, the exact details of the geometry at that scale can be of great importance, particularly when the system behaves in a JKR manner, as the contact radius increases rapidly and a larger portion of the tip surface becomes relevant even at low loads.

◇ The tip can become covered in gold from the surface, which needs to be pushed away upon indentation before the contact mechanics of our system are actually probed. This could explain lower values for stiffness in the initial stages of compression.

◇ Many of our curves show plasticity, whether it is due to pre-existing defects in the vicinity, or to the generation of defects in a pristine volume. This can lead to important differences when a crater has been formed, as the geometry is no longer sphere–plane.

In addition to the points made above, it should be noted that whereas DS20 shows signs of DMT behaviour, and DS28 and DS02 both show signs of JKR behaviour, it is likely that all of these data sets fall more accurately at various positions in the intermediate Maugis-Dugdale regime. This is corroborated by estimates for the μ and λ parameters (see table 6.1), assuming a W-Au flat–flat equilibrium interface separation $z_0 = .30$nm [80].

A more detailed analysis of the data sets based on a method developed by Piétrement and Troyon [146], and Carpick *et al.* [147] confirms that DS20 can be mostly treated as DMT while DS02 can be treated as JKR. DS28 is more difficult to fit, due to some discrete changes in loading stiffness which are probably due to the surface roughness of the tip. Therefore, for simplicity, the three data sets will be treated as DMT, JKR and JKR respectively, unless otherwise specified.

---

[4]For instance, assuming a paraboloidal tip allows us to use the Hertz equation (and related adhesion models) despite the fact that the Hertz assumption of small contact radius $a \ll R$ does not hold for a number of our measurements.

Data Set	Tip Radius	μ	λ	Observed behaviour
DS20	3.6nm	.29	.33	$\sim$DMT
DS28	6.1nm	.34	.40	$\sim$JKR
DS02	8.5nm	.38	.44	$\sim$JKR

Table 6.1: Tabor parameter μ and transition parameter λ estimates based on the measured tip radius by FIM, assuming literature values for $E_r = 75$GPa, $\gamma = 1Jm^{-2}$ and $z_0 = .3$nm [80]. These values would indicate that our data sets fall at various positions in the Maugis-Dugdale transition between DMT and JKR.

6.3 Nanoindentation Data with Adhesion: Material Properties

Having identified contact models that appear to describe our data sets successfully, our analysis can be carried forward with assumptions on elastic deformation profiles and stresses that could not be made based on analyses from section 5.3.

6.3.1 Stiffness Based Measurement of Material Properties

In our previous discussions, we concluded that the three main data sets of interest (DS20, DS28, and DS02) behave according to the DMT, JKR, and JKR models respectively. In addition, we showed that the zero-point displacement is difficult to obtain for a number of our curves, due in particular to adsorbed gold on the tip that produces electronic contact (and at times soft elastic–plastic contact) before complete mechanical contact between the indenter and the surface. For these reasons, we aim to devise DMT/JKR based methods that rely on quantities that can be measured with superior accuracy: maximum load P_{max} and initial unloading stiffness S.

Elastic Loading

Assuming elastic loading, we write the radius of contact and the expression of the stiffness given by Piétrement *et al.* [146] for DMT and JKR.

$$a_{\mathrm{DMT}} = \left(\frac{R}{K}(P + P_{\mathrm{po}})\right)^{1/3} \qquad (6.15)$$

$$a_{\mathrm{JKR}} = \left(\frac{R}{K}\left(\sqrt{P_{\mathrm{po}}} + \sqrt{P + P_{\mathrm{po}}}\right)^2\right)^{1/3} \qquad (6.16)$$

$$S_{\text{DMT}} \;=\; \frac{3}{2}Ka_{\text{DMT}} = \frac{3}{2}(K^2R)^{1/3}(P + P_{\text{po}})^{1/3} \quad (6.17)$$

$$S_{\text{JKR}} \;=\; \frac{3}{2}Ka_{\text{JKR}}\frac{1 - \frac{1}{2}\sqrt{\frac{4RP_{\text{po}}}{K}}\left(\frac{1}{a_{\text{JKR}}}\right)^{3/2}}{1 - \frac{1}{6}\sqrt{\frac{4RP_{\text{po}}}{K}}\left(\frac{1}{a_{\text{JKR}}}\right)^{3/2}} \quad (6.18)$$

We can then rewrite the stiffness equations in the more general form of equations (6.19)-(6.21), in order to highlight the dependence on P, R, and K, for high loads. The f functions in equations (6.19)–(6.21) are functions of the ratio of the load to the pull-off force, and become very close to one as the load increases.

$$S \;=\; \frac{3}{2}\left(K^2R\right)^{1/3}P^{1/3}\,f^{1/3}\left(\frac{P}{P_{\text{po}}}\right) \qquad (6.19)$$

$$f_{\text{DMT}}(x) \;=\; 1 + \frac{1}{x} \qquad (6.20)$$

$$f_{\text{JKR}}(x) \;=\; \left(\frac{1 + \sqrt{1+x}}{\sqrt{x}}\right)^{2/3}\frac{\sqrt{1+x}}{\sqrt{1+x+2/3}} \quad (6.21)$$

The contact radii equations (6.15) and (6.16) can also be rewritten in the simpler form of equation (6.22), using g functions which also become close to one as the load increases.

$$a^3 \;=\; \frac{PR}{K}\,g\left(\frac{P}{P_{\text{po}}}\right) \qquad (6.22)$$

$$g_{\text{DMT}}(x) \;=\; 1 + \frac{1}{x} \qquad (6.23)$$

$$g_{\text{JKR}}(x) \;=\; \left(\sqrt{\frac{1}{x}} + \sqrt{1 + \frac{1}{x}}\right)^2 \quad (6.24)$$

While the deformation remains elastic, one can use equation (6.19) and the tip radius obtained from FIM image analysis, to extract the modified elastic modulus K from a measurement of the stiffness at peak load. This is done using equation (6.25), where the argument of the f function was left out for the sake of readability.

$$\boxed{K \;=\; \sqrt{\frac{8S^3}{27PRf}} \quad (6.25)}$$

In order to obtain the mean contact pressure p_{m}, one can use equation (6.22) to eliminate the elastic modulus from equation (6.19), and rewrite the stiffness in the form of equation (6.26).

$$S = \frac{3Rg^{2/3}f^{1/3}}{2}\frac{P}{a^2} \quad (6.26)$$

The mean contact pressure $p_{\mathrm{m}} = P/\pi a^2$ can thus be readily computed by using equation (6.27).

$$p_{\mathrm{m}} = \frac{P}{\pi a^2} = \frac{2S}{3R\pi\left(g^2 f\right)^{1/3}} \quad (6.27)$$

Inelastic Loading

The analysis carried out previously assumed elastic loading of a sphere-flat contact. When plasticity sets in, the slope of the load–displacement curve (i.e. the stiffness) becomes smaller due to plastic events. This is no longer the case when unloading starts, as the system reverts to behaving elastically. However, the geometry during unloading is different, as the bottom surface is no longer flat. Assuming minimum departure from non-adhesive contact mechanics[5], the bottom surface can be approximated by a sphere of curvature $-1/\rho$, where the negative sign reflects the fact that the surface is now concave[83]. The new effective radius for R_{i}, defined by $1/R_{\mathrm{i}} = 1/R - 1/\rho$ is then expected to become larger as the surface deforms further. Equation (6.28) highlights the fact that, as the surface is plastically deformed, ρ goes from infinity to finite values, which causes R_{i} (initially equal to R for $\rho = \infty$) to become larger.

$$R_{\mathrm{i}} = R\frac{1}{1 - R/\rho} \quad (6.28)$$

Based on this reasoning, one can plug-in the initial unloading stiffness into equation (6.25), to compute an *apparent* modified Elastic modulus K_{app}. K_{app} is expected to be equal to the system's modified elastic modulus K when the loading is elastic (region 1 of 5.1) and to become progressively higher as high loads push the system into plastic deformation (region 2 of 5.1), at which point the effective radius R_{i} increases, and is underestimated by R, used in the denominator of equation (6.25) to determine K_{app}.

When the plastic zone has fully developed (region 3 of figure 5.1), the mean contact pressure becomes constant, and tends towards the hardness H of the material. Assuming that equation (6.27) is still valid in the plastic region, and noting that for

---

[5]This assumption is reasonable for data set DS20, which behaves in a DMT manner. While this approach is not mathematically rigorous in the general case of adhesive contacts, it can be very instructive and facilitate the extraction of a number of insights.

very high loads g and f are both very close to 1, one would therefore expect the stiffness to plateau at a value $S_\infty \simeq 3\pi RH/2$ when the load becomes high enough.

The effect of the stiffness plateau when the plastic zone has fully developed on K_{app} can be determined from equation (6.25), which shows that in that regime, K_{app} is dominated by the $1/\sqrt{P}$ factor. We would therefore expect K_{app} to decrease for high enough loads.

(a) DS28 (R=6.1nm) (b) DS02 (R=8.5nm)
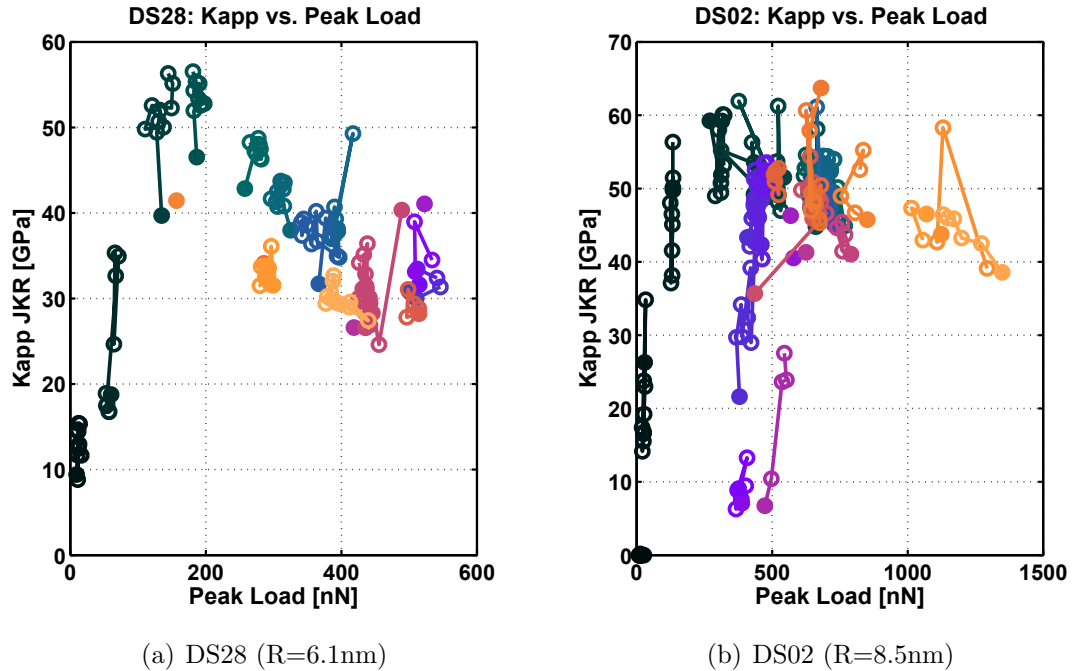
Figure 6.9: K_{app} vs. peak load for the data sets behaving in a JKR-like manner. K_{app} starts off small and quickly moves towards elastic values. As the plastic zone develops, the contact pressure plateaus, which causes K_{app} to drop as the load is further increased. For DS02, the blue and purple sequences with abnormally low K_{app} around 500nN look significantly different from other curves, which suggest that different physics is at play for them.

Figure 6.9 shows the evolution of K_{app} as the maximum load is increased for the JKR data sets (DS28 and DS02). As expected, K_{app} drops steadily for increasing peak load after a certain point. Figure 6.10 confirms that this is due to the mean contact pressure plateauing beyond a certain load.

In figure 6.9, the maximum value of K_{app} is still well below the expected value of $K_{\mathrm{Au\text{-}W}} \simeq 100GPa$ [80]. Similar issues have been reported in the literature [79, 148], and could be explained in our case by a number of factors, such as the increased sensitivity to indenter roughness and geometry by the JKR behaviour of the tip, or the wetting of the tip which could induced changes in the surface. The values

obtained for the mean contact pressure at a high load (between 5-7GPa), on the other hand, are on the same order of magnitude as previous results of nanoindentation on gold films and gold single crystals, which vary between .75-5GPa in the literature [75, 113–115, 127, 138, 139, 149]. In addition, the results reported in the literature are obtained with tips that are significantly blunter than ours, and indentation size effects are expected to lead to higher values of hardness for sharper tips.



(a) DS28 (R=6.1nm)   (b) DS02 (R=8.5nm)
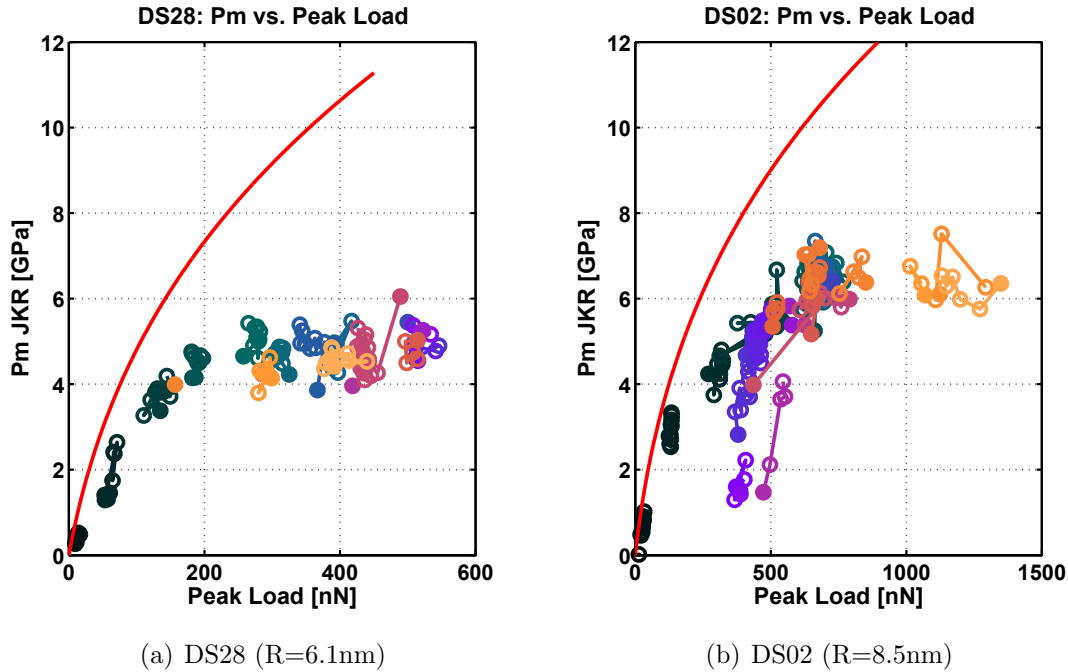
Figure 6.10: The mean contact pressure vs. peak load for JKR data sets. The red line corresponds to the expected JKR mean pressure for a fully elastic contact for tips of radii R=6.1nm and R=8.5nm respectively, with assumed values of E_r=75GPa (K=100GPa), and $\gamma = 1 Jm^{-2}$[80].

In figure 6.10, the estimate of the mean contact pressure was obtained using equation (6.27), and replacing S with the initial unloading stiffness, R by the appropriate tip radius as measured by FIM prior to the experiment, and evaluating the functions g and f at $x = P_{max}/P_{po}$, where the peak load P_{max}, and the pull-off force P_{po}were estimated directly from the P-h curves. The figure also illustrates that the system is unable to keep up with the theoretical JKR elastic pressure (red line), computed based on literature values for E_r and γ.

While data sets DS28 and DS02 behave largely as expected in terms of the arguments made in this section, it appears that data set DS20, obtained with a very sharp tip, shows a stark contrast in behaviour.

6.3.2   DS20: Effect of Very Sharp Tips

The analysis presented in the previous section, when applied to the tips with larger radii, appears to lead to a behaviour that is largely expected as described earlier. The situation is very different for data set DS20, which was obtain with a very sharp tip ($R = 3.6$nm).
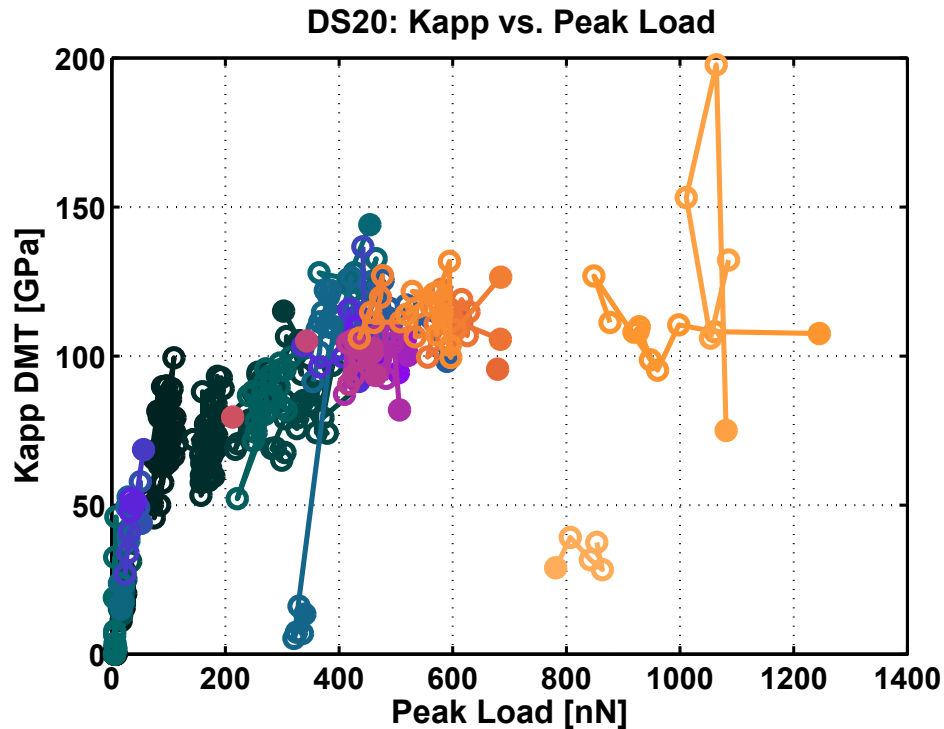
Figure 6.11: K_{app} vs. peak load for data set DS20, obtained with a very sharp tip ($R = 3.6$nm). In this case, K_{app} appears to increase continuously, for the full range of applied loads, which suggests that the mean contact pressure does not reach a limiting value. The points with abnormally low K_{app} (around 350nN) correspond to spurious effects due to a temporary saturation in our acquisition system and should be ignored. The points for loads above 700nN should also be taken with caution (see text).

The first observation that can be made from figure 6.11 is that K_{app} does not decrease as load is increased in the range of loads of the experiment. As confirmed by figure 6.12, this suggests that the mean contact pressure, which is a lower bound for hardness, does not reach a limiting value and continues to increase.

In figures 6.11 and 6.12, the outliers with a peak load around 350nN are due to spurious effects caused by a temporary saturation in the force channel and should be disregarded. In addition, points for loads above 700nN should be taken with great caution, as the corresponding curves exhibited highly unusual shapes, which
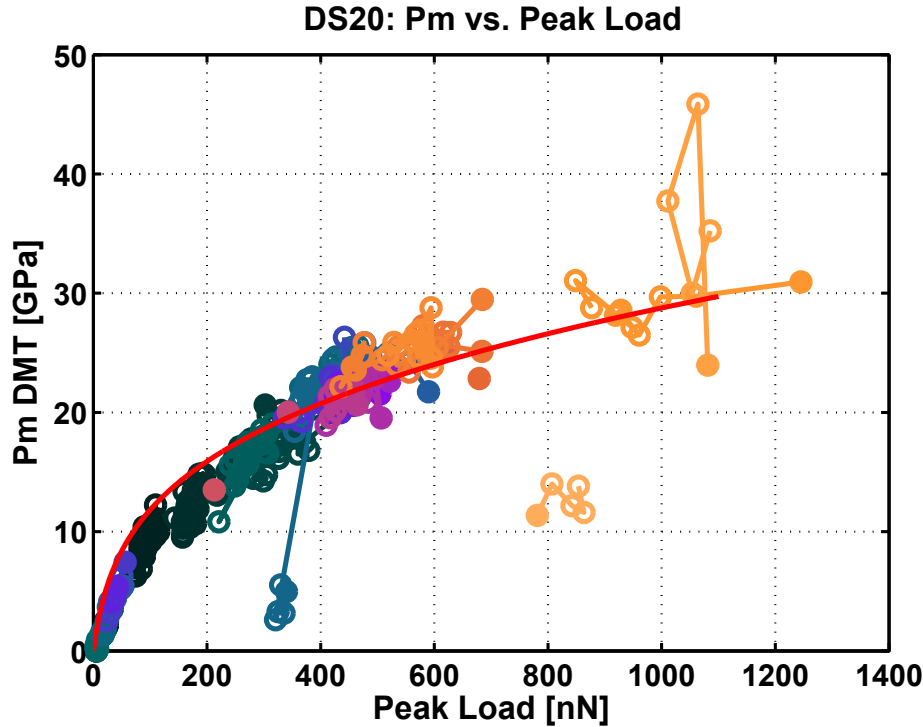
Figure 6.12: The mean contact pressure vs. peak load for DMT data set DS20. The red line corresponds to the expected DMT mean pressure for a fully elastic contact for a tip of radius R=3.6nm, with assumed values of E_{r}=75GPa (K=100GPa), and $\gamma = 1 Jm^{-2}$[80].

we attribute to catastrophic damaging of the tip: in the case of DS20, the tip was not imageable after the series of experiments, and the last few sets of curves show very peculiar behaviour with the unloading curve sometime appearing above the loading curve. A possible interpretation is that the applied load, as it was increased further to reach a limiting value of the mean contact pressure, became too high for our sharp tip to withstand.

It is interesting to note that the trend observed in figure 6.12 is in quantitative agreement with other methods used elsewhere in this thesis to estimate the mean contact pressure. For example, the Oliver and Pharr analysis carried out in section 5.3.2 for sequence #03 of this data set, with peak loads between 80nN–110nN, yielded a value of $p_{\mathrm{m}} = 10.5$GPa. Similarly, the residual impression analysis performed in section 5.3.2 yielded a value of p_{m} close to 20GPa. Within the limits of our error analysis, both of these points fall at the right place in figure 6.12.

It is reasonable to assume that the high contact pressures observed in this data set are due to the very sharp aspect of the indenter. Although it can be argued that in some cases indentation size effects result from instrumental artefacts, particularly

related to uncertainty in the estimation of the contact area at very small scales [150], genuine size effects have been well known and a wide range of theories has been proposed to account for them. In simple terms, it has been observed at a variety of scales that smaller typically means harder. Examples of experimental studies on the Au(111) surface include the work by Michalske and Houston [122], which explored tip radii down to 50nm, and more recently Lucas *et al.* [113], who used tips of 70nm and 112nm. Size effects were observed that led to higher hardness or shear stress values for smaller tip radii. A size effect by which smaller indentation depth increased hardness was also documented. In all cases, it appears that the determinant factor is the size of the volume being probed, which is a function of the tip radius and the indentation depth among others.

The underlying premise is that perfect crystals are harder than real annealed crystals because the latter contains a large number of dislocations that are mobile and facilitate deformation. However, as the number of dislocations increase during plastic deformation, they will eventually collide and stop one another or reach grain boundaries at which point they can become immobile, causing an increase in hardness (i.e. work hardening).

At a large scale for polycrystalline materials, grain size is therefore the determining factor, with the yield stress varying with the inverse square root of the grain size, which is known as the Hall-Petch relation [151, 152].

For a single crystal, theories based on strain gradients and the geometrically necessary dislocations, or GNDs [153], required to accommodate the higher strain gradients caused by smaller geometries were proposed by Fleck *et al.* [154], and later by Nix and Gao [155]. However, it was shown by Gerberich *et al.* that these effects cannot account for the Indentation Size Effects (ISE) observed for small depths, below 100nm.

A different mechanism to explain size effects on single crystal micro- and nano-pillars was introduced by Nix and coworkers [156–159], emphasising the concept of *dislocation starvation*. According to this theory, in a non geometrically confined volume attached to a larger crystal, such as a micropillar, dislocations will be generated and glide away from the volume leaving it effectively perfect, and thereby hardening it. Similar work by Shan and coworkers in a Transmission Electron Microscope (TEM) [160] shows that even as dislocations were initially present in a micropillar, a first compression test will make the pillar dislocation free, and a second test will require a much higher yield stress. The phenomenon of dislocation starvation was also observed in molecular dynamics simulations of nickel nanowires [161], where

dislocation nucleation, propagation, and escape events were separated by elastic behaviour.

Returning to the size effect observed in data set DS20, the fact that the mean contact pressure continues to increase well above the highest reported values of hardness for gold, and that it does not reach a plateau within the range of applied loads of the experiment is striking. In addition, a permanent and highly disordered region is created in the probed volume below the tip, as evidenced by the loss of conductivity (see section 4.3.3). Together, these elements suggest that a different mechanism of hardening may be at play. As pointed out by Kelchner *et al.* [78], a decrease of the indenter size to a very small radii will be accompanied by a decrease in the volume with a high enough shear stress to induce dislocation loops. One can hypothesise that below a critical volume, dislocation nucleation and propagation may not be the main mechanism of deformation. For instance, mechanisms involving point defects that would be attracted to the high stress fields could be imagined. In order to answer these questions, new theoretical developments are required, and atomistic simulations at the scale of radii below 5nm would most likely provide valuable insights that could be directly related to our observations.

Chapter 7

Conclusion

7.1 Summary

The object of the thesis was to investigate experimentally the nanomechanics of nanometre scale volumes under high compressive loads. This was done by performing a large number of nanometre scale indentations of the Au(111) surface, with tips of radii 3.6nm, 6.1nm, and 8.5nm. A few features distinguish this study from all prior nanoindentation studies of the Au(111) surface. These experiments were conducted in Ultra High Vacuum (UHV), with tips characterized at the atomically scale by Field Ion Microscopy (FIM). Load–displacement curves were obtained at high resolutions (on the order of 1nN and 10pm for load and displacement respectively) with simultaneous acquisition of tip–sample current spanning from 10pA to 1mA. The sample surface was imaged by a combined STM/AFM prior to and after single and cyclic loading experiments, and atomic characterization of tips was performed successfully after most indentation data sets. With the exception of the results by Cross *et al.* [80], obtained on the same system, the tips used in this study are of unprecedented sharpness in reported results of gold indentation, and as such, allow the probing of nanometre scale volumes and offer the possibility of direct comparison with atomistic simulations. Compared to reference [80], the results reported here are of higher resolution and larger scope, with a variety of phenomenological observations. These were enabled in particular by imaging of the surface after indentation, by the wider range of measurable currents, and by the variety of tip radii used.

Results reveal striking size effects, with mean contact pressure reaching 30GPa for the sharpest tip. Size effects were also observed with respect to adhesion, with the sharpest tip exhibiting a Derjaguin-Müller-Toporov (DMT) type of behaviour whereas the ones with larger radii behave according to the Johnson-Kendall-Roberts

([JKR](#)) model. Tip imaging after nanoindentation typically revealed moderate adsorption of gold on the tip surface, with the underlying tip surface largely unaffected. Dramatic suppression effects of cyclic loading on electronic transport properties of the probed volumes were observed, and confirmed by preliminary electronic transport calculations performed by the group of Professor Guo at McGill University.

A broad description of the phenomenology of nanometre scale indentation was also provided, comprising interesting phenomena such as self healing and surface dynamics involving island and vacancy cluster diffusion. Details of the unique instrument that enabled these measurements were provided, with an emphasis on enhancements recently implemented, particular as they relate to force measurement. Theoretical and experimental considerations relevant to force transducers used in the system were explored in detail, outlining the implications of the choice of design parameters on accuracy, precision, and the mitigation of external excitations. A comprehensive set of integrated software tools developed in the Matlab$^{\text{TM}}$ environment for data management, analysis, and exploration, were introduced and are transferable to a variety of contexts.

7.2   Future Work

Prospects for future work on this system are tremendous. One direction will involve continued analysis of the rich data sets described in this thesis. In particular, an analysis of discrete plastic events, their size distributions, and the loads at which they occur, will provide valuable insights into the plasticity and the yielding behaviour of our nanoscale volumes. This should be in conjunction with a detailed exploration of loading curves, including yield stress analysis using appropriate contact models, and a more quantitative investigation of the elastic–plastic transition. While it is difficult to predict in detail the outcome of continued analysis of this data, it will certainly inform the next experiments that should be performed.

On the instrumentation side, the system developed in conjunction with William Paul in the Matlab$^{\text{TM}}$ environment to control the acquisition system via TCP/IP is extremely promising. In the short term, it will allow the acquisition of signals at the higher bandwidth of 50kHz, with adequate antialiasing at 10kHz. Beyond improving the general quality of the signal, this will facilitate the resolution of discrete plastic events, as they occur on a timescale related to the cantilever resonance frequency, which is currently close to the maximum frequency detectable with our acquisition system. In addition, the application to the piezotube of arbitrary actuation profiles

will become possible, enabling drift tracking and correction, hold times to investigate time dependent issues such as creep, and the seamless integration of I-V spectroscopy, which will greatly expand the scope of our investigations into the electronic transport properties of the probed volumes. Finally, this system will allow the implementation of software feedback and feed-forward schemes, which could be used to implement the load or displacement control, specific loading rates and loading targets, as well as the ability to halt indentation at the first discrete plastic event. Subsequent imaging of the surface would allow the investigation of the conditions for the formation of permanent plastic indents. Imaging should also be used to study surface dynamics in detail, particularly to isolate tip mediated effects from spontaneous diffusion of islands and vacancy clusters. The studies may elucidate interesting surface decoration patterns (seemingly related to the herringbone reconstruction) which emerge in some of our data sets in regions that have been indented, as shown in figure 7.1.

(a) DS20: 100nm x 100nm   (b) DS26: 50nm x 50nm
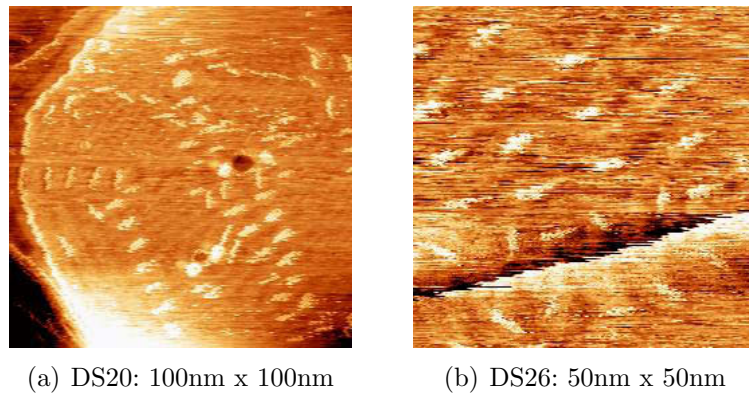
Figure 7.1: Surface decoration pattern which may become visible during nanometre scale indentation of Au(111)

Other studies can be imagined to explore incipient plasticity in Au(111). The study of plasticity at various distances from a step is one such study, which could be directly compared to the results of atomistic simulations by Zimmerman *et al.* [162]. In the simulation, the tip radius was 4nm, within the range of radii that can routinely be produced in our system, whereas existing literature employs tips over radii 25nm and higher [163]. Further investigations will also allow the determination of the specific conditions that lead to self healing.

Finally, key insights will likely come from comparison of the detailed results of our experiments with associated commensurate atomistic computations. Such comparisons can be expected to shed light on the very mechanisms responsible for stress induced mass transport at the nanometre scale.

Appendix A

Digital Signal Processing

The aim of the appendix is to highlight a few important points on Digital Signal Processing (DSP). For further information, the reader is invited to refer to a number of excellent resources available on the world wide web. In particular, *"The Scientist and Engineer's Guide to Digital Signal Processing"* [164] is an comprehensive reference freely available online[1], while the National Instruments® Developer zone[2] contains thorough and practical tutorials. In the following, we will briefly touch on three points of importance to us.

A.1 Antialiasing

A very important component of an acquisition system is the antialiasing filter, a hardware low-pass filter with a cutoff frequency that should be less than one half of the sampling frequency (the Nyquist frequency). The aim of the filter is to suppress all high frequencies before they are acquired. In the absence of antialiasing, frequencies above the Nyquist frequency appear as a lower frequency in the acquired signal, as highlighted by figure A.1, which illustrates a case of aliasing in time domain. Figure A.2 shows an experimental situation that highlights the same phenomenon, in the frequency domain. Aliasing increases the noise level as all levels above the Nyquist frequency fold back into the measurement bandwidth, and can lead to inaccurate interpretation of peaks at specific frequencies. One practical way to detect aliasing is simply changing the sampling frequency and keeping track of peaks that move.

---

[1] Available at: www.dspguide.com
[2] Available at zone.ni.com
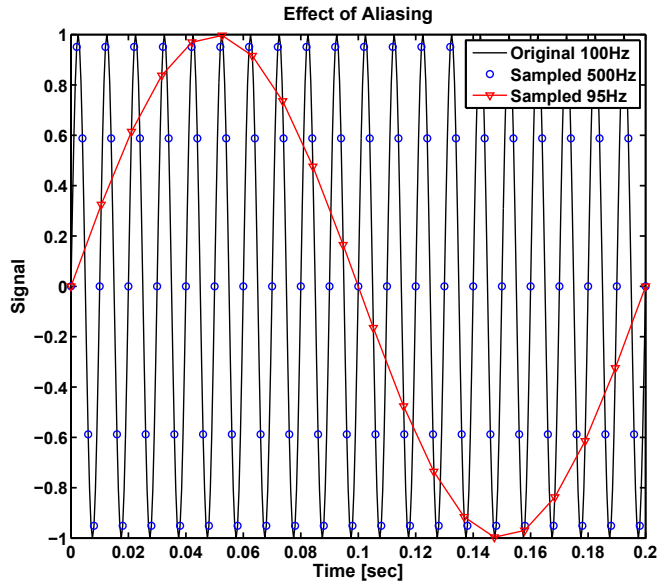
Figure A.1: Aliasing in the time domain. An original frequency of 100Hz (black curve) is sampled at 500Hz (blue circles), which allows for an accurate representation. However, when sampling at 95Hz (in red), the Nyquist frequency is 47.5Hz, which means that the apparent frequency will be —95-100—=5Hz, as confirmed by the red trace.
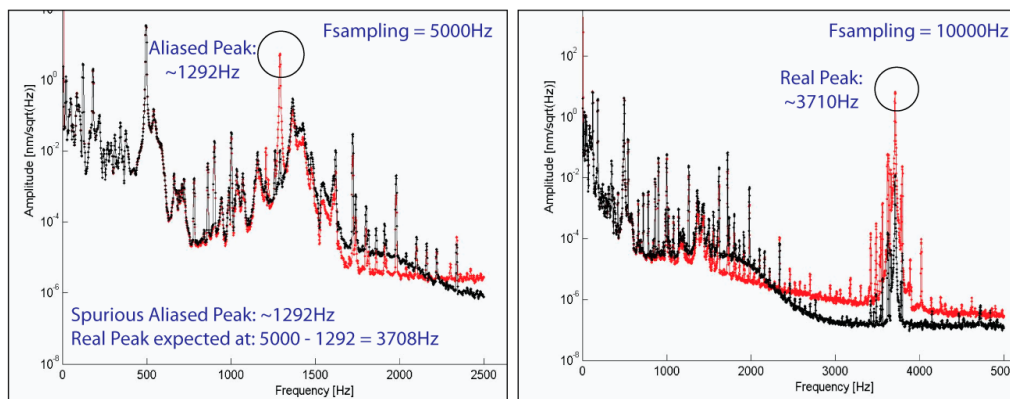
Figure A.2: Effect of aliasing in the frequency domain, during experimental tested. The red and black curves are amplitude spectral densities of the same signal. The black trace corresponds to a low order low-pass filter with a cut-off frequency at 2kHz, and performs mild antialiasing. The red curve is not acquired through an antialiasing filter, and therefore an excitation at 3710Hz shows up as an aliased peak at 1290Hz when acquired with a sampling rate of 5kHz. The Nyquist frequency is immediately visible from the graphs, as it corresponds to the last frequency bin.

A.2   Windowing and Welch Method

When computing the power spectral density of a signal, the signal is finite, and as such no longer satisfies the condition of periodicity in Fourier analysis. For that reason, when Discrete Fourier Transforms are computed, the power density corresponding to specific frequency will *leak* into other bins. Fast fourier transform algorithms can be understood as a bank of notch filters at regularly spaced frequencies. However, these filters have a finite bandwidth, and signal will leak into neighbouring bins. In order to mitigate this phenomenon and control the amount of spectral leakage from the general background vs. from neighbouring frequencies, *windowing* is performed. Windowing consists of multiplying the signal by a window that typically goes to zero at both ends of each segment that is analysed, so as to improve its pseudo-periodic character. Applying no window at all correspond to a rectangular window, and is characterized by large spectral leakage, with contributions from a broad range of bins. The flat top window is known to have large leakage near the centre frequency, and low leakage outside, which can be useful when the amplitude of peaks is sought. However, peaks will tend to appear wider in Power Spectral Density (PSD) plots, which can make Q-factor fits inaccurate. Hanning windows are a popular choice, as they provide a good compromise, as highlighted by figure A.3 For more information on windowing, the reader can refer to references [34, 164, 165], and on the National Instruments® website[3].
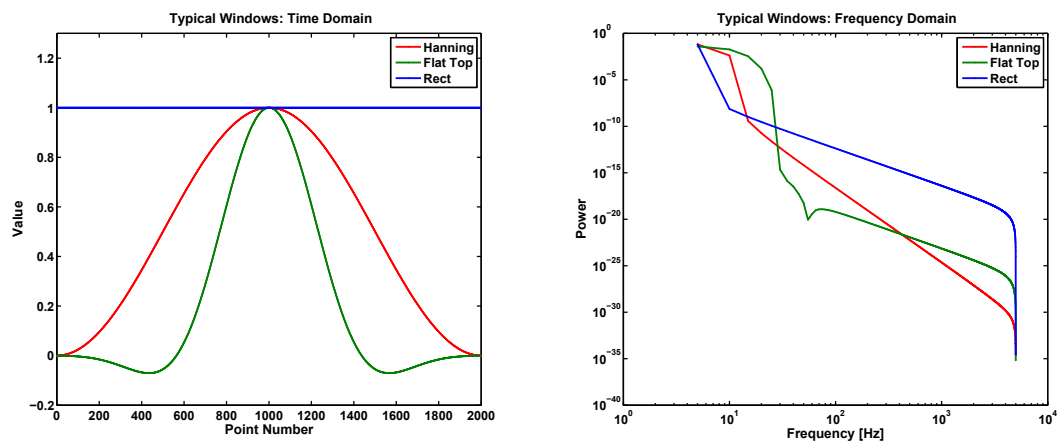

Figure A.3: Typical windows in time and frequency domains.

When computing a PSD plot, great care must be taken to normalise the spectrum properly, both with respect to the change from the 2-sided to the 1-sided PSD, and

---

[3]Available here: zone.ni.com/devzone/cda/tut/p/id/4844

to the window used. In this thesis, all PSD plots are computed using a Hanning window, using the Welch method, with is described in detail in [34].

A.3   Filtering

Finally, post-processing filtering techniques were often used in this thesis. Details are provided in section 4.2.4. A first version of our filtering techniques were developed manually using the techniques outlined in [164]. However, Matlab$^{\text{TM}}$ provides a powerful and convenient filter design toolbox, which was used as a basis for our filtering. As explained earlier, the filtering of curves in this thesis generally consists of an order 8 low-pass Butterworth filer, and narrow band-stop filters at specific frequencies. The amplitude transfer function of the low-pass filter used is shown in figure A.4, along with an example of one band-stop filter.
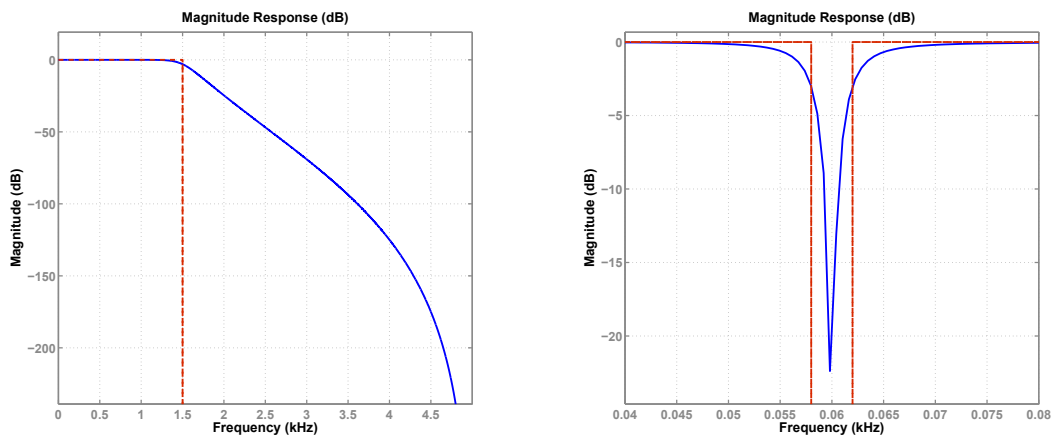

Figure A.4: Filter transfer functions.

Appendix B

Enhancing Data Acquisition, Management, and Analysis with Matlab$^{\text{TM}}$, Java$^{\text{TM}}$, and other Software Tools

B.1 Data Acquisition

The data presented in this thesis comes in many formats and from various sources: STM/AFM images and spectroscopy curves are obtained through Scanita, in an uncommon format loosely based on XML, FIM images are acquired with a variety of sources including SLR cameras, a customised webcam, video cameras, and CCDs. Information on cantilever and laser beam positions is gathered manually, based on manual calibration procedures and analysis of photographs. Other sources and types of data include AES spectra, sample signal for spectral analysis, bode plots, etc. Information on the specific conditions of measurements are typically recorded in lab books which become inconvenient when large volumes of data need to be managed, especially when they are shared among several users. In order to facilitate data management and collaboration, standardized software tools were developed in Tcl/Tk, Java$^{\text{TM}}$, and Matlab$^{\text{TM}}$. To the extent possible, data acquisition was performed directly in Matlab$^{\text{TM}}$, using the data acquisition and the image toolboxes. For instance, a software package was recently developed jointly with William Paul to control and bypass the STM/AFM acquisition system directly from Matlab$^{\text{TM}}$, using TCP/IP and Tcl/Tk. This was done to allow high bandwidth measurements with maximum

flexibility and control, in particular to apply arbitrary actuation profiles during spectroscopy. The system was successfully implemented in a testing environment.

In cases when data was acquired outside of the Matlab^{TM} environment, software was developed to collect all relevant information in electronic format, preprocess it, and centralise it in Matlab^{TM}. For instance, a Java^{TM} tool (SdfReader) was built to open and process Scanita data files, and convert to various useful format, such as ascii, binary, and WSxM^{TM}[109]. In addition, SdfReader is a Java^{TM} class that can be instantiated directly from Matlab^{TM}, and allow data and associated information to be imported readily into efficient and well structured objects developed in Matlab^{TM}, using Object Oriented Programming concepts.

B.2   Data Management

Data from most of our sources is centralised in an object based data management system in Matlab^{TM}, which allows the storage of all relevant associated information, such as STM/AFM data files, cantilever beam characteristics, tip positions, FIM images, and user comments.

The data management system is designed such that all relevant data are interconnected and accessible dynamically. Standardization and automation also allows immediate comparisons of raw data, processed data, and user generated indicators among very large data sets. Using useful Matlab^{TM} functionalities (such as 'ButtonDownFcn'), access to original data is instantaneous when clicking on data points in indicator plots. For instance, after entering one data loading command at the Matlab^{TM} prompt, relevant data sets are loaded, and a simple:

<div align="center">DS20.plot('peakload','stiffness')</div>

will instantly produce a plot of the stiffness vs. peak load graph for all 349 curves in data set DS20. Simply pressing 'p' on the keyboard will produce a 'pdf' rendering of the graph, shown in figure B.1. In this plot, each circle correspond a single load–displacement curve, while connected circles of a given colour correspond to a sequence of consecutive curves at a given position.

Investigating particular points of interest, such as outliers is readily done. Looking at figure B.1, a few points with loads around 350nN exhibit abnormally low stiffness values around 100N/m. The first immediate remark is that those points belong to the same sequence, and that other points in the same sequence follow the general trend. To further investigate the trend, a simple click on the cluster will produce an
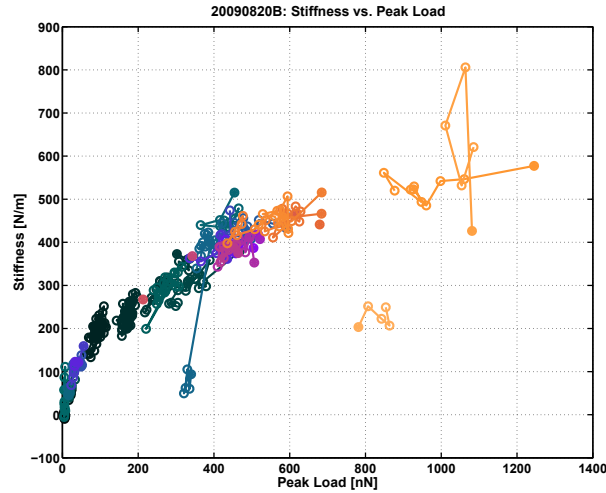
Figure B.1: Summary plot of the stiffness vs. peak load for all 349 curves in data set DS20.

additional figure, which includes all the plots in the sequence. Pressing 'g' this time (for 'pnG') will produce a png file of the figure[1], which is shown without further processing in figure B.2. This figure shows immediately the reason for the abnormal behaviour: the force channel was temporarily saturated during the acquisition of the first five plots, which affected the stiffness measurement. Each of the plots in the new figure are clickable, and various combinations of clicking and keyboard strokes will produce different customizable outcomes.

The data set structures also centralise information regarding relevant experimental parameters, electronic gains, etc. In addition, FIM images are also included in the structures, and clickable summaries of the chronology of a complete data set can be generated that include STM scans and FIM images. Such summaries are shown for DS20 on pages 161 and 162. These summaries contain standardized plot ranges so that sequences can be readily compared. The effect of curve 738-S in sequence #22 (top left corner of page 162) is immediately observed by looking at the surface scans before and after indentation. Sequence #22 contained a single curve. For multi-curve sequences, only the first curve is shown, along with the number of curves and the maximum piezo extension. FIM images and scans can also be obtained by clicking, which also provides all relevant information such as user comments taken manually during the acquisition.

The data management system is seamlessly integrated with a number of process-

---

[1]Pdf is avoided here because the figure contains millions of points that would make a vectorial pdf file very large.
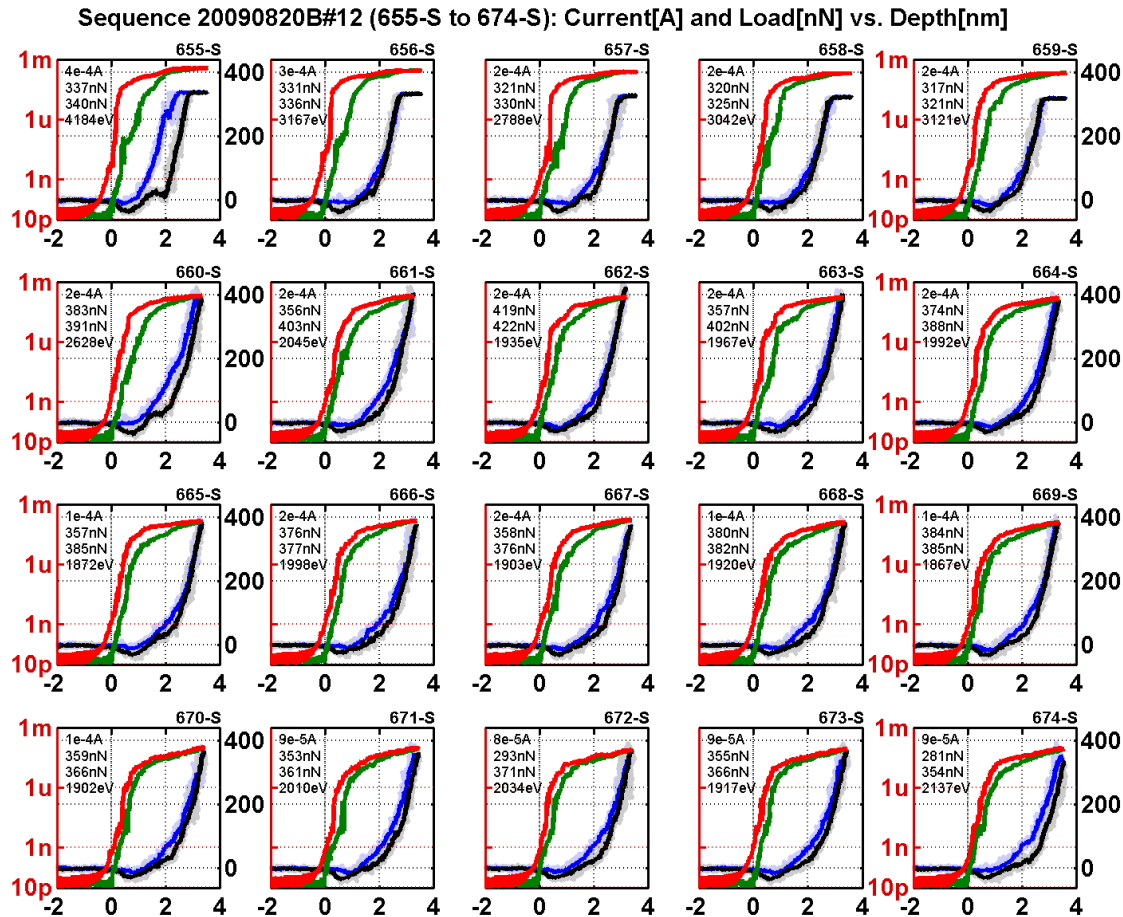
Figure B.2: DS20: Plots of the twelfth sequence.

ing and analysis tools that were developed directly in Matlab$^{\text{TM}}$. These comprise customizable power spectra generation tools based on the Welch method, Power Spectral Density Evolution (PSDE) (see section 3.3.2) generation tools, STM/AFM image processing and display tools, fitting tools (e.g. thermal peak fitting), etc.

This combination of software tools allows the efficient centralization of data and all associated information, which greatly facilitates data exploration, data processing, data analysis, archiving, and collaboration, which all are critical to successful investigations involving very large and complex data sets.

Full Summary 20090820B(Page 01)
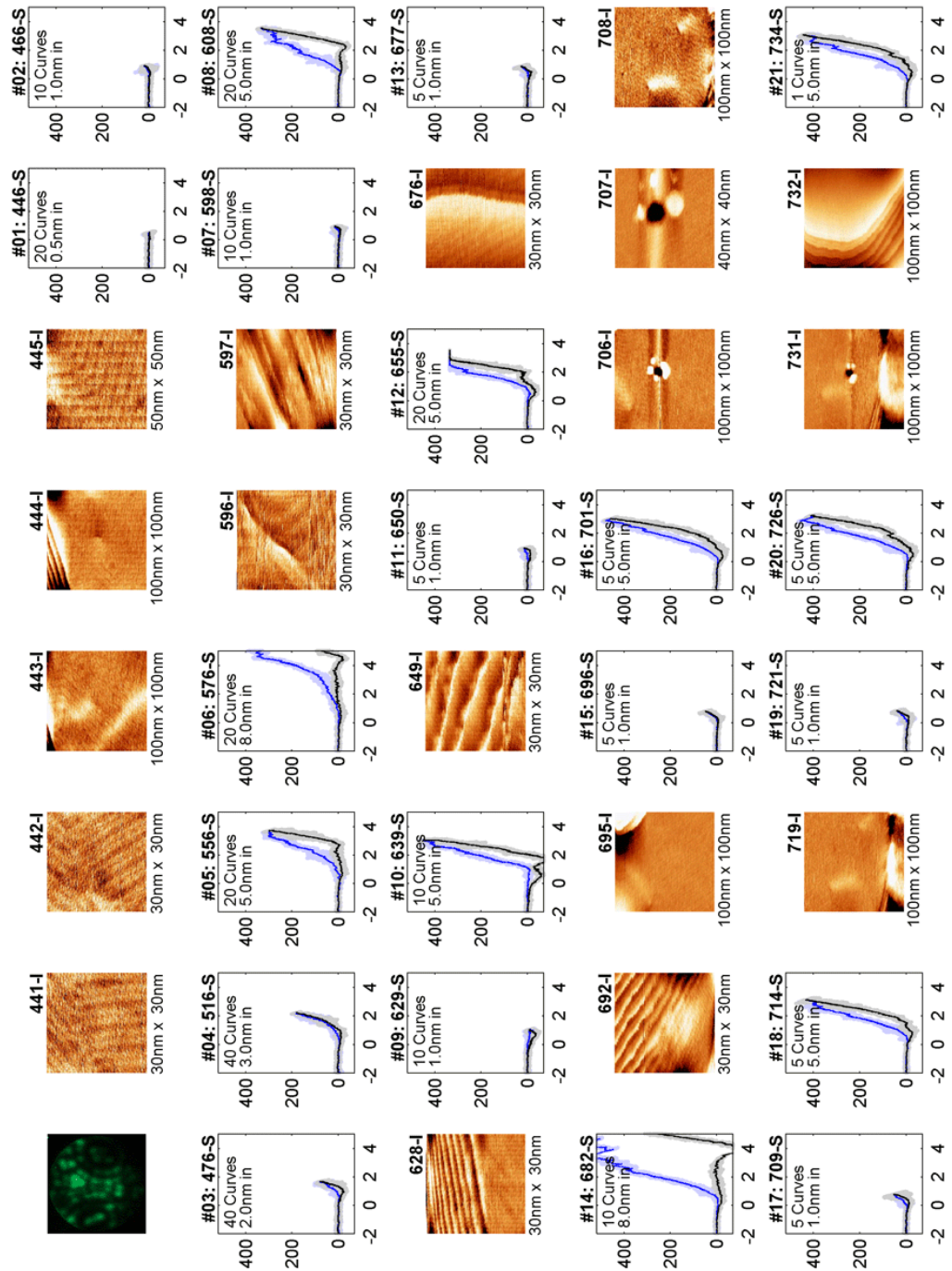
Full Summary 20090820B(Page 02)
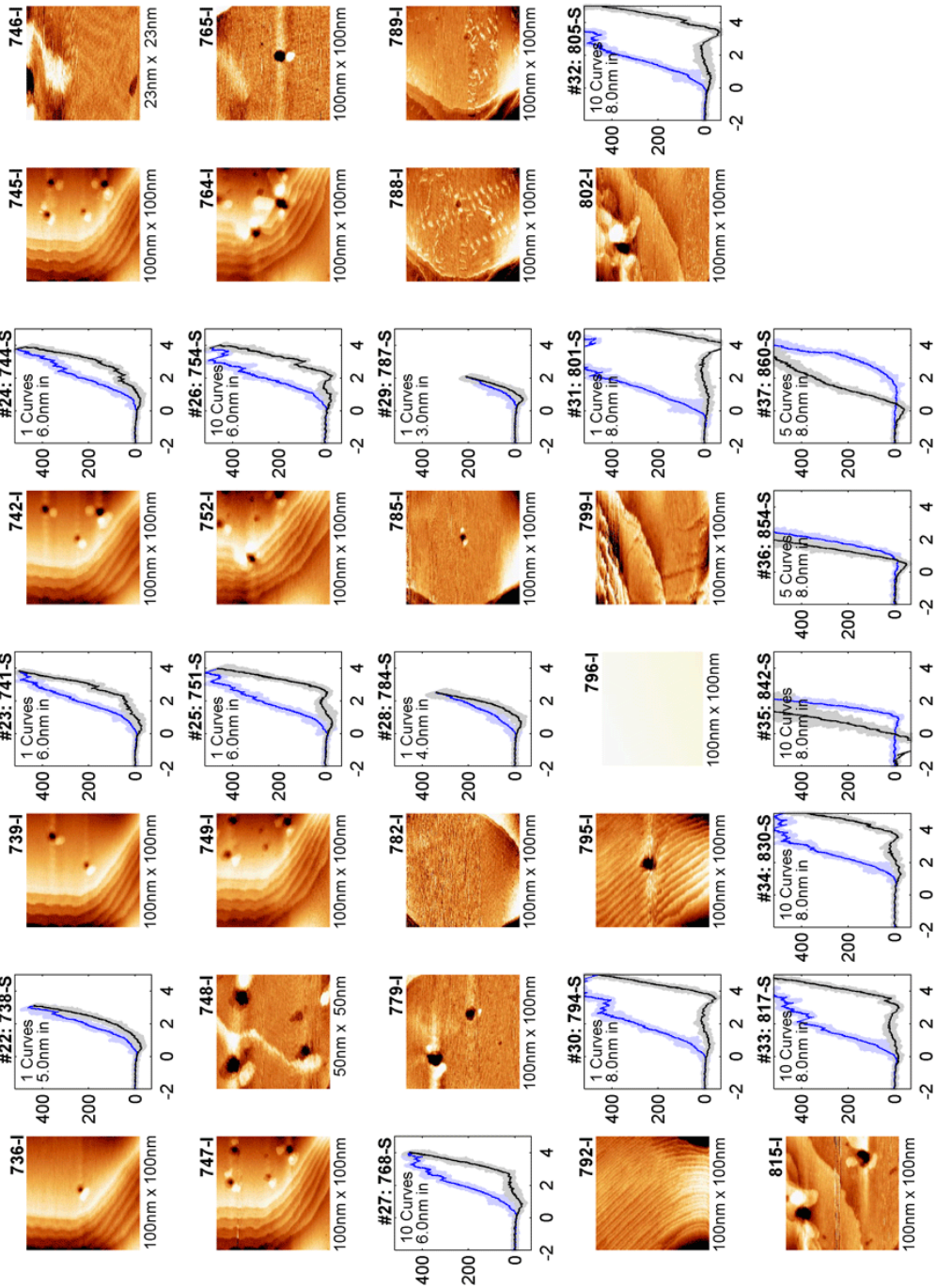
References

[1] Ted Sargent. *The dance of molecules : how nanotechnology is changing our lives.* Thunder's Mouth Press, New York, 2006.

[2] G Binnig, H Rohrer, Ch. Gerber, and E Weibel. Surface Studies by Scanning Tunneling Microscopy. *Physical Review Letters*, 49(1):57 LP – 61, July 1982.

[3] G Binnig, C F Quate, and Ch. Gerber. Atomic Force Microscope. *Physical Review Letters*, 56(9):930 LP – 933, March 1986.

[4] Neil W. Ashcroft and N. David Mermin. *Solid state physics.* Harcourt College Publishers, 1976.

[5] A Stalder. *Mechanics and Electronics on the Atomic Scale: A Study using Scanning Probe Methods.* Phd thesis, University of Fribourg, 1995.

[6] Yan Sun. *Study of Interactions at the Atomic Scale.* Phd thesis, McGill University, 2004.

[7] Till Hagedorn. *Atomic Contacts characterized by Force and Current.* Phd thesis, McGill University, 2010.

[8] Bharat Bhushan, editor. *Springer Handbook of Nanotechnology.* Springer, 2004.

[9] E Meyer, H J Hug, and R Bennewitz. *Scanning Probe Microscopy - The Lab on a Tip.* Springer-Verlag, Berlin, 2004.

[10] A S Foster and W A Hofer. *Scanning Probe Microscopy - Atomic Scale Engineering by Forces and Currents.* NanoScience and Technology. Springer, New York, 2006.

[11] D Sarid. *Exploring Scanning Probe Microscopy with Mathematica.* Wiley-VCH, 2007.

[12] W A Hofer, A S Foster, and A L Shluger. Theories of scanning probe microscopes at the atomic scale. *Reviews of Modern Physics*, 75(4):1287–1331, October 2003.

[13] C Julian Chen. *Introduction to Scanning Tunneling Microscopy*, volume 1. Oxford University Press, Oxford, England, 2nd edition, September 2007.

[14] D Sarid. *Scanning Force Microscopy - With Applications to Electric, Magnetic and Atomic Forces.* Oxfor University Press, New York, revised edition, 1994.

[15] S Morita, R Wiesendanger, and E Meyer. *Noncontact Atomic Force Microscopy.* Springer, Berlin, 2002.

[16] Anne-Sophie Lucier. *Preparation and Characterization of Tungsten Tips Suitable for Molecular Electronics Studies.* Msc thesis, McGill University, Montreal, 2004.

[17] Anne-Sophie Lucier, Henrik Mortensen, Yan Sun, and Peter Grütter. Determination of the atomic structure of scanning probe microscopy tungsten tips by field ion microscopy. *Physical Review B*, 72(23):1–9, December 2005.

[18] Tien T. Tsong. *Atom-probe field ion microscopy - Field ion emission and surfaces and interfaces at atomic resolution.* Cambridge University Press, Cambridge, 1990.

[19] M K Miller, A Cerezo, M G Hetherington, and G D W Smith. *Atom Probe Field Ion Microscopy.* Oxford University Press, Oxford, 1996.

[20] T R Albrecht, P Grutter, D Horne, and D Rugar. Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity. *Journal of Applied Physics*, 69(2):668–673, 1991.

[21] U Dürig, O Zuger, and A Stalder. Interaction force detection in scanning probe microscopy: Methods and applications. *Journal of Applied Physics*, 72(5):1778–1798, 1992.

[22] U Dürig, L Novotny, B Michel, and A Stalder. Logarithmic current-to-voltage converter for local probe microscopy. *Review of Scientific Instruments*, 68(10):3814–3816, 1997.

[23] N A Burnham, R J Colton, and H M Pollock. Interpretation of force curves in force microscopy. *Nanotechnology*, 4(2):64–80, 1993.

[24] B Cappella. Force-distance curves by atomic force microscopy. *Surface Science Reports*, 34(1-3):1–104, 1999.

[25] H Butt, B Cappella, and M Kappl. Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surface Science Reports*, 59(1-6):1–152, October 2005.

[26] M. Saint Jean, S. Hudlet, C. Guthmann, and J. Berger. Van der Waals and capacitive forces in atomic force microscopies. *Journal of Applied Physics*, 86(9):5245, 1999.

[27] S Schaer. *Dynamische Rasterkraftmikroskopie auf heterogenen Oberflachen*. Phd thesis, Uni Basel, 2004.

[28] Roy R. Jr. Craig. *Mechanics of Materials*. John Wiley & Sons, 2000.

[29] S Rast, C Wattinger, U Gysin, and E Meyer. Dynamics of damped cantilevers. *Review of Scientific Instruments*, 71(7):2772–2775, 2000.

[30] H J Butt and M Jaschke. Calculation of thermal noise in atomic force microscopy. *Nanotechnology*, 6(1):1–7, 1995.

[31] S Rast, C Wattinger, U Gysin, and E Meyer. The noise of cantilevers. *Nanotechnology*, 11(3):169–172, 2000.

[32] K Y Yasumura, T D Stowe, E M Chow, T Pfafman, T W Kenny, B C Stipe, and D Rugar. Quality factors in micron- and submicron-thick cantilevers. *Journal of Microelectromechanical Systems*, 9(1):117–125, 2000.

[33] C. Schönenberger and S. F. Alvarado. A differential interferometer for force microscopy. *Review of Scientific Instruments*, 60(10):3131–3134, October 1989.

[34] O M Solomon Jr. The effects of windowing and quantization error on the amplitude of frequency-domain functions. *IEEE Transactions on Instrumentation and Measurement*, 41(6):932–937, 1992.

[35] Dieter W Pohl. Some design criteria in Scanning Tunneling Microscopy. *Ibm Journal of Research and Development*, 30(4):417–427, 1985.

[36] M. Okano, K. Kajimura, S. Wakiyama, F. Sakai, W. Mizutani, and M. Ono. Vibration isolation for scanning tunneling microscopy. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 5(6):3313–3320, November 1987.

[37] A Stalder and U Dürig. Ultrahigh vacuum compatible cooling and vibration isolation stage. *Review of Scientific Instruments*, 64(12):3644–3646, 1993.

[38] F Ostendorf, C Schmitz, S Hirth, A Kuhnle, J J Kolodziej, and M Reichling. How flat is an air-cleaved mica surface? *Nanotechnology*, 19(30):305705, 2008.

[39] Marco Tarini, Paolo Cignoni, and Claudio Montani. Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1237–1244, 2006.

[40] Jacob N Israelachvili, Norma a Alcantar, Nobuo Maeda, Thomas E Mates, and Marina Ruths. Preparing contamination-free mica substrates for surface characterization, force measurements, and imaging. *Langmuir : the ACS journal of surfaces and colloids*, 20(9):3616–22, April 2004.

[41] M. Heuberger and M. Zäch. Nanofluidics: Structural Forces, Density Anomalies, and the Pivotal Role of Nanoparticles. *Langmuir*, 19(6):1943–1947, March 2003.

[42] JV Barth, H Brune, G Ertl, and RJ Behm. Scanning tunneling microscopy observations on the reconstructed Au(111) surface: Atomic structure, long-range superstructure, rotational domains, and surface defects. *Physical Review B*, 42(15):9307–9318, November 1990.

[43] Y Hasegawa and P Avouris. Manipulation of the reconstruction of the Au (111) surface with the STM. *Science*, 258(5089):1763–1765, December 1992.

[44] Kevin S. Schneider, Kenneth T. Nicholson, Daniel R. Fosnacht, Bradford G. Orr, and Mark M. Banaszak Holl. Effect of Surface Reconstruction on Molecular Chemisorption: A Scanning Tunneling Microscopy Study of H 8 Si 8 O 12 Clusters on Au(111) 233. *Langmuir*, 18(21):8116–8122, October 2002.

[45] K Reichelt and H O Lutz. Hetero-epitaxial growth of vacuum evaporated silver and gold. *Journal of Crystal Growth*, 10(1):103–107, June 1971.

[46] Vm Hallmark, S Chiang, Jf Rabolt, Jd Swalen, and Rj Wilson. Observation of atomic corrugation on Au(111) by scanning tunneling microscopy. *Physical review letters*, 59(25):2879–2882, December 1987.

[47] B Marchon, S Ferrer, DS Kaufman, M Salmeron, and W. The surface topography of pyrolitic carbons and of gold thin films by scanning tunneling microscopy: Grain boundaries and surface defects. *Thin Solid Films*, 154:65–73, 1987.

[48] C E D Chidsey, D N Loiacono, T Sleator, and S Nakahara. STM study of the surface morphology of gold on mica. *Surface Science*, 200(1):45–66, 1988.

[49] R. Emch, J. Nogami, M. M. Dovek, C. A. Lang, and C. F. Quate. Characterization of gold surfaces for use as substrates in scanning tunneling microscopy studies. *Journal of Applied Physics*, 65(1):79, 1989.

[50] J. Vancea, G. Reiss, F. Schneider, K. Bauer, and H. Hoffmann. Substrate effects on the surface topography of evaporated gold filmsa scanning tunnelling microscopy investigation. *Surface Science*, 218(1):108–126, August 1989.

[51] J A DeRose, T Thundat, L A Nagahara, and S M Lindsay. Gold grown epitaxially on mica: conditions for large area flat faces. *Surface Science*, 256(1-2):102–108, 1991.

[52] Y Golan, L Margulis, and I Rubinstein. Vacuum-deposited gold films. I. Factors affecting the film morphology. *Surface Science*, 264(3):312–326, 1992.

[53] Rainer Behrisch and Wolfgang Eckstein. *Sputtering by Particle Bombardment: Experiments and Computer Calculations from Threshold to MeV Energies. Topics in Applied Physics.* Springer, 2007.

[54] C A Lang, C F Quate, and J Nogami. Initial stages of sputtering on Au(111) as seen by scanning tunneling microscopy. *Applied Physics Letters*, 59(14):1696–1698, 1991.

[55] T Michely, K H Besocke, and G Comsa. Observation of sputtering damage on Au(111). *Surface Science*, 230(1-3):135–139, May 1990.

[56] D Porath, Y Goldstein, A Grayevsky, and O Millo. Scanning tunneling microscopy studies of annealing of gold films. *Surface Science*, 321(1-2):81–88, 1994.

[57] Lawrence E Davis. *Handbook of Auger electron spectroscopy : a reference book of standard data for identification and interpretation of Auger electron spectroscopy data.* Physical Electronics Industries, Eden Prairie, Minn., 1976.

[58] EAG. http://www.eaglabs.com/training/tutorials/aes_theory_tutorial/energies.php, retrieved on 2010/08/14.

[59] A Zangwill. *Physics at Surfaces.* Cambridge University Press, 1988.

[60] M Prutton. *Introduction to Surface Physics.* Oxford Science Publications, 1994.

[61] H Lüth. *Solid Surfaces, Interfaces and Thin Films.* Springer, 2001.

[62] Martin Prutton and Mohamed M El Gomati. *Scanning Auger Electron Microscopy.* John Wiley & Sons, 2006.

[63] M. P. Seah and W. a. Dench. Quantitative electron spectroscopy of surfaces: A standard data base for electron inelastic mean free paths in solids. *Surface and Interface Analysis*, 1(1):2–11, February 1979.

[64] David Penn. Electron mean free paths for free-electron-like materials. *Physical Review B*, 13(12):5248–5254, June 1976.

[65] D Norman and DP Woodruff. The energy dependence of the electron inelastic scattering mean-free-path in gold. *Solid State Communications*, 22(11):711–713, 1977.

[66] Daisuke Fujita and Kazuhiro Yoshihara. Practical energy scale calibration procedure for Auger electron spectrometers using a spectrometer offset function. *Surface and Interface Analysis*, 21(4):226–230, 1994.

[67] Mark H. Dishner, M M Ivey, S Gorer, J C Hemminger, and F J Feher. Preparation of gold thin films by epitaxial growth on mica and the effect of flame annealing. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 16(6):3295, November 1998.

[68] C Nogues and M Wanunu. A rapid approach to reproducible, atomically flat gold films on mica. *Surface Science*, 573(3):L383–L389, December 2004.

[69] M Hegner, P Wagner, and G Semenza. Ultralarge atomically flat template-stripped Au surfaces for scanning probe microscopy. *Surface Science*, 291(1-2):39–46, 1993.

[70] L Chai and J Klein. Large area, molecularly smooth (0.2 nm rms) gold films for surface forces and other studies. *Langmuir*, 23(14):7777–7783, 2007.

[71] F J Giessibl. A direct method to calculate tip-sample forces from frequency shifts in frequency-modulation atomic force microscopy. *Applied Physics Letters*, 78(1):123–125, 2001.

[72] U. Rabe, K. Janser, and W. Arnold. Vibrations of free and surface-coupled atomic force microscope cantilevers: Theory and experiment. *Review of Scientific Instruments*, 67(9):3281, 1996.

[73] a. Fian, Ch. Ernst, and M. Leisch. Combined atom probe and STM study of tip-substrate interactions. *Fresenius' Journal of Analytical Chemistry*, 365(1-3):38–42, September 1999.

[74] A Fian and M. Leisch. Study on tipsubstrate interactions by STM and APFIM. *Ultramicroscopy*, 95(1-4):189–197, June 2003.

[75] S G Corcoran, R J Colton, E T Lilleodden, and W Gerberich. Anomalous plastic deformation at surfaces: nanoindentation of gold single crystals. *Physical Review B (Condensed Matter)*, 55(24):16057–16060, 1997.

[76] J D Kiely and J E Houston. Nanomechanical properties of Au (111), (001), and (110) surfaces. *Physical Review B (Condensed Matter)*, 57(19):12588–12594, 1998.

[77] J D Kiely, K F Jarausch, J E Houston, and P E Russell. Initial stages of yield in nanoindentation. *Journal of Materials Research*, 14(6):2219–2227, 1999.

[78] C L Kelchner, S J Plimpton, and J C Hamilton. Dislocation nucleation and defect structure during surface indentation. *Physical Review B (Condensed Matter)*, 58(17):11085–11088, 1998.

[79] A Asenjo, M Jaafar, E Carrasco, and J M Rojo. Dislocation mechanisms in the first stage of plasticity of nanoindented Au(111) surfaces. *Physical Review B*, 73(7):75431, February 2006.

[80] Graham Cross, André Schirmeisen, Peter Grutter, and U Dürig. Plasticity, healing and shakedown in sharp-asperity nanoindentation. *Nature Materials*, 5(5):370–376, 2006.

[81] D Tabor. *The Hardness of Metals.* Oxford University Press, 1951.

[82] D Tabor. Indentation hardness: Fifty years on a personal view. *Philosophical Magazine A: Physics of Condensed Matter, Structure, Defects and Mechanical Properties*, 74(5):1207–1212, 1996.

[83] K L Johnson. *Contact Mechanics*. Cambridge University Press, Cambridge, 1985.

[84] A C Fischer-Cripps. *Introduction to Contact Mechanics*. Springer, second edition, 2007.

[85] Anthony C Fischer-Cripps. *Nanoindentation*. Mechanical engineering series. Springer, New York, 2nd edition, 2004.

[86] Ian M Hutchings. The contributions of David Tabor to the science of indentation hardness. *Journal of Materials Research*, 24(3):581–589, March 2009.

[87] A C Fischer-Cripps. A review of analysis methods for sub-micron indentation testing. *Vacuum*, 58(4):569–585, September 2000.

[88] Anthony C Fischer-Cripps. Critical review of analysis and interpretation of nanoindentation test data. *Surface and Coatings Technology*, 200(14-15):4153–4165, 2006.

[89] Christopher A Schuh. Nanoindentation studies of materials. *Materials Today*, 9(5):32–40, 2006.

[90] J. Hay. Introduction To Instrumented Indentation Testing. *Experimental Techniques*, 33(6):66–72, November 2009.

[91] R M VanLandingham. Review of instrumented indentation. *Journal of Research of the National Institute of Standards and Technology*, 108(4):249–265, July 2003.

[92] O L Warren, S A Downs, and T J Wyrobek. Challenges and interesting observations associated with feedback-controlled nanoindentation. *Zeitschrift fuer Metallkunde/Materials Research and Advanced Techniques*, 95(5):287–296, 2004.

[93] W A Soer, J Th. M De Hosson, A M Minor, Z Shan, S A Syed Asif, and O L Warren. Incipient plasticity in metallic thin films. *Applied Physics Letters*, 90(18):181924, 2007.

[94] Andrew M Minor, S A Syed Asif, Zhiwei Shan, Eric A Stach, Edward Cyrankowski, Thomas J Wyrobek, and Oden L Warren. A new view of the onset of plasticity during the nanoindentation of aluminium. *Nature Materials*, 5(9):697–702, 2006.

[95] D Maugis. *Contact, Adhesion and Rupture of Elastic Solids*. Springer, Berlin, 2000.

[96] RS Bradley. LXXIX. The cohesive force between solid surfaces and the surface energy of solids. *Philosophical Magazine Series 7*, 13(86):853–862, 1932.

[97] B V Derjaguin, V M Muller, and Yu P Toporov. Effect of contact deformations on the adhesion of particles. *Journal of Colloid and Interface Science*, 53(2):314–326, November 1975.

[98] KL Johnson, K. Kendall, and AD Roberts. Surface energy and the contact of elastic solids. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 324(1558):301–313, 1971.

[99] D Tabor. Surface forces and surface interactions. *Journal of Colloid and Interface Science*, 58(1):2–13, 1977.

[100] Daniel Maugis. Adhesion of spheres: The JKR-DMT transition using a dugdale model. *Journal of Colloid and Interface Science*, 150(1):243–269, 1992.

[101] K L Johnson and J A Greenwood. Adhesion map for the contact of elastic spheres. *Journal of Colloid and Interface Science*, 192(2):326, 1997.

[102] VM Muller, BV Derjaguin, and Y.P. Toporov. On two methods of calculation of the force of sticking of an elastic sphere to a rigid plane. *Colloids and Surfaces*, 7(3):251–259, August 1983.

[103] M F Doerner and W D Nix. A method for interpreting the data from depth-sensing indentation instruments. *Journal of Materials Research*, 1(4):601–609, July 1986.

[104] Warren C. Oliver and G M Pharr. An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *Journal of Materials Research*, 7(6):1564–1583, June 1992.

[105] Ian N Sneddon. The relation between load and penetration in the axisymmetric boussinesq problem for a punch of arbitrary profile. *International Journal of Engineering Science*, 3(1):47–57, May 1965.

[106] G M Pharr, Warren C. Oliver, and F R Brotzen. On the generality of the relationship among contact stiffness, contact area, and elastic modulus during indentation. *Journal of Materials Research*, 7(3):613–617, 1992.

[107] Warren C. Oliver and G M Pharr. Measurement of hardness and elastic modulus by instrumented indentation: Advances in understanding and refinements to methodology. *Journal of Materials Research*, 19(1):3–20, 2004.

[108] G M Pharr and A Bolshakov. Understanding nanoindentation unloading curves. *Journal of Materials Research*, 17(10):2660–2671, 2002.

[109] I Horcas, R Fernández, J M Gómez-Rodríguez, J Colchero, J Gómez-Herrero, and A M Baro. WSXM: a software for scanning probe microscopy and a tool for nanotechnology. *The Review of scientific instruments*, 78(1):013705, January 2007.

[110] F Milstein and S Chantasiriwan. Theoretical study of the response of 12 cubic metals to uniaxial loading. *Physical Review B - Condensed Matter and Materials Physics*, 58(10):6006–6018, 1998.

[111] J.-M. Zhang, Y Zhang, K.-W. Xu, and V Ji. Young's modulus surface and Poisson's ratio curve for cubic metals. *Journal of Physics and Chemistry of Solids*, 68(4):503–510, 2007.

[112] J.J. Vlassak and WD Nix. Measuring the elastic properties of anisotropic materials by means of indentation experiments. *Journal of the Mechanics and Physics of Solids*, 42(8):1223–1245, 1994.

[113] Marcel Lucas, Ken Gall, and Elisa Riedo. Tip size effects on atomic force microscopy nanoindentation of a gold single crystal. *Journal of Applied Physics*, 104(11), 2008.

[114] J B Pethica, R Hutchings, and Warren C. Oliver. Hardness measurement at penetration depths as small as 20 nm. *Philosophical Magazine A (Physics of Condensed Matter, Defects and Mechanical Properties)*, 48(4):593–606, 1983.

[115] Marianne Dietiker, Ralph D Nyilas, Christian Solenthaler, and Ralph Spolenak. Nanoindentation of single-crystalline gold thin films: Correlating hardness and the onset of plasticity. *Acta Materialia*, 56(15):3887–3899, 2008.

[116] D. L. Joslin and W. C. Oliver. A new method for analyzing data from continuous depth-sensing microindentation tests. *Journal of Materials Research*, 5(1):123–126, January 1990.

[117] J.S. Field and M.V. Swain. A simple predictive model for spherical indentation. *Journal of Materials Research*, 8(2):297–306, February 1993.

[118] D Maugis and H M Pollock. Surface forces, deformation and adherence at metal microcontacts. *Acta Metallurgica*, 32(9):1323–1334, 1984.

[119] SV Hainsworth, HW Chandler, and TF Page. Analysis of nanoindentation load-displacement loading curves. *Journal of Materials Research*, 11(8):1987–1995, 1996.

[120] John Price Hirth and Jens Lothe. *Theory of dislocations*. McGraw-Hill series in materials science and engineering. McGraw-Hill, New York, 1992.

[121] Derek Hull and D. J. Bacon. *Introduction to dislocations*. Butterworth-Heinemann, 2001.

[122] T A Michalske and J E Houston. Dislocation nucleation at nano-scale mechanical contacts. *Acta Materialia*, 46(2):391–396, 1998.

[123] Ju Li, Krystyn J Van Vliet, Ting Zhu, Sidney Yip, and Subra Suresh. Atomistic mechanisms governing elastic limit and incipient plasticity in crystals. *Nature*, 418(6895):307–310, 2002.

[124] N. Gane and F. P. Bowden. Microdeformation of Solids. *Journal of Applied Physics*, 39(3):1432, 1968.

[125] N. Gane and J. M. Cox. The micro-hardness of metals at very low loads. *Philosophical Magazine*, 22(179):881–891, November 1970.

[126] JB Pethica and D Tabor. Contact of characterised metal surfaces at very low loads: Deformation and adhesion. *Surface Science*, 89(1-3):182–190, 1979.

[127] N A Burnham and R J Colton. Measuring the nanomechanical properties and surface forces of materials using an atomic force microscope. *Journal of Vacuum Science & Technology A (Vacuum, Surfaces, and Films)*, 7(4):2906–2913, July 1989.

[128] J E Houston and T A Michalske. Interfacial-force microscope. *Nature*, 356(6366):266, 1992.

[129] R.C. Thomas, JE Houston, T.A. Michalske, and R.M. Crooks. The mechanical response of gold substrates passivated by self-assembling monolayer films. *Science*, 259(5103):1883, 1993.

[130] Steven A Smallwood. *Deformation studies of tungsten-gold contacts at the nanometer scale.* PhD thesis, The University of Maine, 2001.

[131] W N Unertl, R J Lad, and S A Smallwood. Nanometer-scale contact mechanics of tungsten-gold contacts in ultra-high vacuum. In *Materials Research Society Symposium Proceedings*, volume 1021 of *Surface and Interfacial Nanomechanics - 2007 MRS Spring Meeting*, pages 29–40, Laboratory for Surface Science and Technology, University of Maine, 5708 ESRB-Barrows Hall, Orono, ME 04469, United States, 2007.

[132] E Carrasco, O Rodríguez de la Fuente, and J M Rojo. Dislocation emission at the onset of plasticity during nanoindentation in gold. *Philosophical Magazine*, 88(3):281–296, 2008.

[133] E. Carrasco, M. A. González, O Rodríguez de la Fuente, and J M Rojo. Analysis at atomic level of dislocation emission and motion around nanoindentations in gold. *Surface Science*, 572(2-3):467–475, November 2004.

[134] E Carrasco, O Rodríguez de la Fuente, M A Gonzalez, and J M Rojo. Dislocation cross slip and formation of terraces around nanoindentations in Au(001). *Physical Review B*, 68(18):180102, 2003.

[135] E Carrasco, O R De La Fuente, and J M Rojo. Atomic events at the onset of plasticity in the Au(001) surface. In *Materials Research Society Symposium Proceedings*, volume 976 of *2006 MRS Fall Meeting*, pages 7–12, Física de Materiales, Universidad Complutense de Madrid, Av. Complutense s/n, Madrid, 28040, Spain, 2006.

[136] O R De La Fuente, E Carrasco, M A González, and J M Rojo. Dislocation-mediated mechanisms of mass transport around nanoindentations in Fcc metals. In Piqueras J., Zypman F.R., Bonnell D.A., and Shreve A.P., editors, *Materials Research Society Symposium - Proceedings*, volume 738 of *Spatially Resolved Characterization of Local Phenomena in Materials and Nanostructures*, pages 207–212, Depto. de Fís. de Materiales, Universidad Complutense, 28040 Madrid, Spain, 2003.

[137] Z Zong and W Soboyejo. Indentation size effects in face centered cubic single crystal thin films. *Materials Science and Engineering: A*, 404(1-2):281–290, September 2005.

[138] M. Göken, M. Kempf, M. Bordenet, and H. Vehoff. Nanomechanical characterizations of metals and thin films. *Surface and Interface Analysis*, 27(5-6):302–306, May 1999.

[139] M.J. Cordill, D.M. Hallman, N.R. Moody, D.P. Adams, and W.W. Gerberich. Thickness Effects on the Plasticity of Gold Films. *Metallurgical and Materials Transactions A*, 38(13):2154–2159, January 2007.

[140] Graham Cross. *Mechanics of Nanometer Scale Indentation of a Metal Surface*. Phd thesis, McGill University, Montréal, 1999.

[141] Jiunn-Jong Wu. Easy-to-Implement Equations for Determining Adhesive Contact Parameters with the Accuracy of Numerical Simulations. *Tribology Letters*, 30(2):99–105, March 2008.

[142] L. Sirghi and F. Rossi. Adhesion and elasticity in nanoscale indentation. *Applied Physics Letters*, 89(24):243118, 2006.

[143] L Sirghi and F Rossi. The effect of adhesion on the contact radius in atomic force microscopy indentation. *Nanotechnology*, 20(36):365702, 2009.

[144] J. Woirgard and J-C. Dargenton. An alternative method for penetration depth determination in nanoindentation measurements. *Journal of Materials Research*, 12(9):2455–2458, September 1997.

[145] Binquan Luan and Mark O Robbins. The breakdown of continuum models for mechanical contacts. *Nature*, 435(7044):929–932, 2005.

[146] O. Piétrement and M. Troyon. General Equations Describing Elastic Indentation Depth and Normal Contact Stiffness versus Load. *Journal of colloid and interface science*, 226(1):166–171, June 2000.

[147] R.W. Carpick, D.F. Ogletree, and Miquel Salmeron. A general equation for fitting contact area and friction vs load measurements. *Journal of colloid and interface science*, 211(2):395–400, 1999.

[148] B. Kracke and B. Damaschke. Measurement of nanohardness and nanoelasticity of thin gold films with scanning force microscope. *Applied Physics Letters*, 77(3):361, 2000.

[149] MM McCann and SG Corcoran. Nanoindentation behavior of gold single crystals. *Materials Research Society Symposium Proceedings*, 795:1–7, 2004.

[150] Michele Alderighi, Vincenzo Ierardi, Francesco Fuso, Maria Allegrini, and Roberto Solaro. Size effects in nanoindentation of hard and soft surfaces. *Nanotechnology*, 20(23):235703, June 2009.

[151] George Ellwood Dieter. *Mechanical metallurgy*. McGraw-Hill, 3rd edition, 1986.

[152] K Sieradzki, a Rinaldi, C Friesen, and P Peralta. Length scales in crystal plasticity. *Acta Materialia*, 54(17):4533–4538, October 2006.

[153] M. F. Ashby. The deformation of plastically non-homogeneous materials. *Philosophical Magazine*, 21(170):399–424, February 1970.

[154] NA Fleck, GM Muller, MF Ashby, and JW Hutchinson. Strain gradient plasticity: theory and experiment. *Acta Metallurgica et Materialia*, 42(2):475–487, 1994.

[155] W.D. Nix and Huajian Gao. Indentation size effects in crystalline materials: a law for strain gradient plasticity. *Journal of the Mechanics and Physics of Solids*, 46(3):411–425, 1998.

[156] Michael D Uchic, Dennis M Dimiduk, Jeffrey N Florando, and W D Nix. Sample dimensions influence strength and crystal plasticity. *Science (New York, N.Y.)*, 305(5686):986–9, August 2004.

[157] W D Nix, J Greer, G Feng, and E Lilleodden. Deformation at the nanometer and micrometer length scales: Effects of strain gradients and dislocation starvation. *Thin Solid Films*, 515(6):3152–3157, February 2007.

[158] J R Greer. Bridging the gap between computational and experimental length scales: A review on nano-scale plasticity. *Reviews on Advanced Materials Science*, 13(1):59–70, 2006.

[159] Julia Greer and William Nix. Nanoscale gold pillars strengthened through dislocation starvation. *Physical Review B*, 73(24):245410, June 2006.

[160] Z W Shan, Raja K Mishra, S A Syed Asif, Oden L Warren, and Andrew M Minor. Mechanical annealing and source-limited deformation in submicrometre-diameter Ni crystals. *Nature materials*, 7(2):115–9, February 2008.

[161] A Cao, Y Wei, and S X Mao. Alternating starvation of dislocations during plastic yielding in metallic nanowires. *Scripta Materialia*, 59(2):219–222, 2008.

[162] J A Zimmerman, C L Kelchner, P A Klein, J C Hamilton, and S M Foiles. Surface step effects on nanoindentation. *Physical Review Letters*, 87(16):0, 2001.

[163] J D Kiely, R Q Hwang, and J E Houston. Effect of surface steps on the plastic threshold in nanoindentation. *Physical Review Letters*, 81(20):4424–4427, 1998.

[164] Steven W. Smith. *The Scientist & Engineer's Guide to Digital Signal Processing*. California Technical Publishing, 1997.

[165] R Lyons. Windowing functions improve FFT results [spectral analysis]. *Test & Measurement World*, 18(7):37–38, 1998.

Glossary

A Contact area $= \pi a^2$. 108

E_r Reduced Modulus, also E^* in some references. 109

K Modified Reduced Modulus, $K = \dfrac{4E_r}{3}$. 109

R_r Reduced Radius of Curvature. 109

S Contact stiffness $= \dfrac{\mathrm{d}P}{\mathrm{d}h}$. 115

ν Poisson Ratio. 109

a Contact radius, also r_c in some references. 108

AES Auger Electron Spectroscopy. 49, 54

AFM Atomic Force Microscope/Microscopy. 6, 9, 75, 104, 127

CRM Cantilever Rotation Model. 63

DMT Derjaguin-Müller-Toporov adhesive contact model. 110, 130, 149

DSP Digital Signal Processing. 153

FCC Face Centred Cubic. 52

FIM Field Ion Microscope/Microscopy. 70, 76, 104, 130, 149

HCP Hexagonal Close-Packed. 52

IFM Interfacial Force Microscope. 128

ISE Indentation Size Effects. 147

JKR Johnson-Kendall-Roberts adhesive contact model. 110, 129, 150

MD Maugis-Dugdale adhesive contact model. 111, 112, 130

PSD Power Spectral Density. 155

PSDE Power Spectral Density Evolution. 68, 77, 160

RGA Residual Gas Analyser. 54

SEM Scanning Electron Microscopy. 72, 127

SPM Scanning Probe Microscope/Microscopy. 5, 6

STM Scanning Tunnelling Microscope/Microscopy. 6, 9, 76, 104

TEM Transmission Electron Microscope/Microscopy. 147

UHV Ultra High Vacuum. 5, 54, 76, 104, 130, 149

Index