Multiscale Approach for Simulations of Kelvin Probe Force Microscopy with Atomic Resolution

Inauguraldissertation
zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch–Naturwissenschaftlichen Fakultät
der Universität Basel

von

Ali Sadeghi
aus Iran

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von:

Prof. Dr. Stefan Goedecker
Prof. Dr. Pavel Jelinek

Basel, den 17. September 2013

Prof. Dr. Jörg Schibler
Dekan

Abstract

The distance dependence and atomic-scale contrast recently observed in nominal contact potential difference (CPD) signals simultaneously recorded by the Kelvin probe force microscopy (KPFM) using non-contact atomic force microscopy is addressed theoretically. In particular, we consider probing an insulating surface where the applied bias voltage affects electrostatic forces acting on the atomic scale. Our approach is a multiscale one. First, the electrostatics of the macroscopic tip-cantilever-sample system is treated, both analytically and numerically. Then the resulting electric field under the tip apex is inserted into a series of density functional theory calculations for a realistic neutral but reactive silicon nano-scale tip interacting with a NaCl(001) sample. Theoretical expressions for amplitude modulation (AM) and frequency modulation (FM) KPFM signals and for the corresponding local contact potential differences (LCPD) are obtained and evaluated for several tip oscillation amplitudes A up to 10 nm. For $A = 0.01$ nm, the computed LCPD contrast is proportional to the slope of the atomistic force versus bias in the AM mode and to its derivative with respect to the tip-sample separation in the FM mode. Being essentially constant over a few Volts, this slope is the basic quantity which determines variations of the atomic-scale LCPD contrast. Already above $A = 0.1$ nm, the LCPD contrasts in both modes exhibit almost the same spatial dependence as the slope. As the most basic quantity, the slope is shown to be approximately expressed in terms of intrinsic charge distribution and dipole moment and their variation due to the chemical interactions. The slope is also influenced by the macroscopic bodies.

As a second part, we introduce a method to measure the distances between atomic configurations which is useful when seeking the tip-apex structures. The broad application of this method includes conformational search and machine-learning based interatomic potentials.

Acknowledgments

The presented work here was performed during July 2009 to September 2013 at the University of Basel. That period of time was an enjoyable part of my life in which I obtained much more knowledge and experience both in physics and life.

First of all, I would like to greatly appreciate my PhD advisor, Prof. Stefan Goedecker, for all of his kindly and exceptionally support, scientific or not. With him, I was introduced to the field of computational physics, the amazing scientific mixture made out of physics, mathematics and computer science. As a member of his Computational Physics Group, my focus was mostly on developing efficient computational methods required in atomistic simulations, and contributing to the development of the electronic structure calculations code BigDFT which was also greatly used throughout of my research. I want also to especially appreciate Prof. Alexis Baratoff for frequent invaluable discussions on several topics. His broad knowledge of both theoretical and experimental details was essential in my work.

I gratefully acknowledge financial support from the Swiss National Science Foundation (SNF) and the Swiss National Center of Competence in Research (NCCR) on Nanoscale Science.

The Physics Department was not only an exceptional place to pursue my scientist research, but also offered a very warm and friendly atmosphere to me during my life in Basel. Many things would not have been possible if a lot of people had not helped me in various ways. I would like to thank all of them, and I apologize in advance for any omission. In alphabetical order, I thank Maximilan Amsler, Sandip De, Nazim Dughan, S. Alireza Ghasemi, Stefan Mohr, Michael Sicher, Bastain Shäfer, Huan Trau and Alexander Willand for their friendly supports.

Finally, I give my very special thanks to my lovley wife Akram, who was a patient supporter of my work. Words cannot express my gratitude and I would like to dedicate this thesis to her.

Contents

Introduction

Kelvin probe force microscopy (KPFM), which was introduced twenty years ago, [1, 2] has become an attractive, and indeed unique, non-contact technique to determine the electric surface characteristics of materials on the atomic scale. It has been successfully applied to the mapping of local variations of work function or surface potential along inhomogeneous surfaces of a broad range of materials. [3, 4, 5] KPFM is nowadays a popular tool with a wide variety of applications on semiconducting [6, 7] and insulating [8] surfaces, from quantum dots [9], molecules and solar cells [10, 11, 12] to charge states of an atom. [13]

A lot of work has been done both to improve the experimental technique and to theoretically explain the physics behind. Yet, the interpretation of the KPFM images, like images obtained by other scanning probe microscopy techniques, is not necessarily straightforward. The aim of this work is to shed some light on this complicated problem by a combined theoretical and computational approach.

The first and main part of this thesis, is devoted to explaining the multiscale approach developed for simulating KPFM experiments. In chapter 1, the contributions to the tip-sample interactions, and their relation to the detected signal in experiment are briefly explained. Chapter 2 presents an introduction to the fundamentals of the KPFM method. The electrostatic interactions are specially important in KPFM, and therefore we address them separately and in details in chapter 3. The results are generally applicable to any scanning probe microscopy where the bias-dependent interactions are present. The electrostatics of a conducting tip of a scanning probe microscope against a flat conductor coated with a thin or thick dielectric layer is treated analytically and numerically. In particular, exact and compact approximate expressions for the capacitance, force, force gradient, electric field profiles and their effective widths are derived for a spherical model tip by generalizing known solutions for the conducting sphere and

sample problem. These expressions allow convenient modelling of various measurements involving voltage-biased probes, estimation of lateral resolution and prediction of trends as a function of relevant parameters.

In chapter 4, by combining the macroscopic and atomistic contributions to the electrostatic force, expressions for the Kelvin signal both in AM- and FM-KPFM, and both for small and large amplitude oscillations are derived. Then, in chapter 5 an approximate variant of the multiscale method is presented. This also provides a microscopic insight into the physics of the phenomenon based on the polarization effects.

In the second part of the thesis, a method is introduced for measuring dissimilarities between molecular structures in computer simulations. In order to characterize molecular structures we introduce configuration fingerprint vectors which are counterparts of quantities used experimentally to identify structures. Components of such vectors can be associated to individual atoms and can then serve as an atomic fingerprint that identifies an atom within a structure. The Euclidean distance between the configuration fingerprint vectors satisfies the properties of a metric and can therefore safely be used to measure dissimilarities between configurations in the high dimensional configuration space. We show that these metrics correlate well with the root-mean-square distance (RMSD) between two configurations if this RMSD is obtained from a global minimization over all translations, rotations and permutations of atomic indices. We introduce a Monte Carlo approach to obtain this global minimum of the RMSD between configurations where atomic fingerprints are used to enhance the performance of the procedure.

Chapter 1

Probe-Sample Interactions

Three decades after the invention of the scanning tunneling microscope (STM), [14] diverse scanning probe microscopy (SPM) techniques have become available for the study of local chemical and physical surface properties of materials as well as for manipulating them down to the *atomic scale*. A probing tip scans the surface at distances comparable to atomic distances and *senses* the topmost atomic layers of the surface, providing a surface image. In STM, sensing the surface is done by measuring the current flow of the electrons that tunnel through the vacuum gap between the tip and a *conducting* sample. Contrary to STM, scanning force microscopy [15] (SFM) can be applied to both conducting and *non-conducting* samples, [16] because it *senses* the surface via the tip-sample *local interactions* rather than via the current.

Imaging a surface using an atomic force microscope (AFM) is performed by monitoring the deflection of a micro-cantilever interacting with the surface and then interpreting it as a signature of the interaction. The cantilever beam with a micro-tip of a few microns height attached to one end and driven at the other end, namely the probe, is therefore the heart of the AFM.

In the so-called static operation mode, the force acting on the probe is determined using the Hooke's law via measuring the bending of the cantilever beam of known stiffness. To prevent the tip to suddenly jump to the surface due to a strong attraction, the AFM tip is at contact (i.e. close proximity) with the surface where the probe is repelled from it. In this contact mode, feedback electronics keeps the deflection of the probe (and hence the force exerted on it) constant during the scan by adjusting the voltage fed into the piezoelectric which controls the height. The hight of the cantilever at each lateral position (x, y) is considered as the topography map of the

surface.

1.1 Interplay between interactions and dynamics

Much higher resolutions (down to atomic scale) can be achieved if the AFM works in the so-called dynamic mode where tip and sample are usually not in contact. In this case the cantilever is externally driven to oscillate at or close to its resonance frequency f_1 (or one of the harmonics) and the force between probe and sample is determined via measuring the perturbing effect of the sample on either the amplitude or the frequency of the natural oscillations of the tip. The technique is called amplitude modulation (AM) or frequency modulation (FM) depending on whether the modulation of the amplitude or the frequency according to the tip-sample interaction is exploited to extract information about the characteristics of the sample.

If not interacting with the sample surface, the free cantilever tends to oscillate at its resonance frequency f_1 if driven at this frequency. Upon the s-dependent interaction with the sample, s being the tip-sample separation, the oscillation frequency alters depending on the vertical gradient of the force F exerted on the probe. An extra stiffness

$$k_{\mathrm{ts}} \equiv -\frac{\partial F}{\partial s}$$

is added to the cantilever flexural stiffness k due to the tip-sample interactions and the resonance frequency is perturbed as

$$\frac{f}{f_1} = \sqrt{\frac{k + k_{\mathrm{ts}}}{k}} \simeq 1 + \frac{k_{\mathrm{ts}}}{2k} = 1 - \frac{1}{2k}\frac{\partial F}{\partial s}.$$

Then, the amplitude of the driven oscillation at f_1 is reduced rapidly. To retain the oscillation amplitude, one needs either to amplify the driving force or to readjust its frequency. In the FM-AFM, the amplitude is kept constant with a feedback mechanism which changes the frequency of the driving force to always match f, the shifted frequency. The frequency-shift $\Delta f = f - f_1$ is then recorded as a function of (x, y).

In the repulsion regime, the force-gradient $\partial_s F < 0$ and the cantilever seems stiffer ($\Delta f > 0$). If $\partial_s F > 0$, on the other hand, the effective stiffness is lowered and Δf is negative. For tip oscillation with ultrasmall amplitudes, the frequency shift is, to a very good approximation, given by [17]

$$\frac{\Delta f}{f_1} = -\frac{1}{2k}\frac{\partial F}{\partial s}. \tag{1.1}$$

In practice, however, the amplitude of the oscillation of the tip can be orders of magnitude larger than s. Then Δf is proportional to the force gradient averaged over a cycle, namely

$$\frac{\Delta f}{f_1} = -\frac{1}{2k}\left\langle \frac{\partial F}{\partial s} \right\rangle \tag{1.2}$$

which is given by [18]

$$kA\frac{\Delta f}{f_1} = -\frac{1}{2\pi}\int_0^{2\pi} F[d + A(1+\cos\phi)]\cos\phi d\phi, \tag{1.3}$$

where A is the oscillation amplitude and d the closest tip-sample distance. This issue is addressed in details in 4.3.2.

1.1.1 Contributions

The total force between the probe and sample has three main contributions from

- van der Waals (vdW)
- electrostatic
- and chemical

interactions. Accordingly, the frequency shift consists of three contributions $\Delta f = \Delta f_{\mathrm{vdW}} + \Delta f_{\mathrm{el}} + \Delta f_{\mathrm{chem}}$. The two first are long range and dominant at large separations where they cause an attractive force on the tip. The chemical interactions are dominant at very short separations and responsible to the atomic-scale resolved images.

1.2 Dispersion Interactions

The vdW forces are independent of the bias voltage and therefore are not important in the case of KPFM simulations. Therefore we address them only briefly here. The vdW interactions are long range and attractive (in the relevant separations here). The origin of the vdW interactions is the quantum fluctuation resulting in fluctuating electromagnetic fields from one material interacting with another.

The dispersion interaction between two atoms a distance r away can be described by an empirical potential based on quantum mechanics [19] as

$$V_{\mathrm{vdW}}(r) = -\frac{C_6}{r^6}, \tag{1.4}$$

where C_6 is a constant depending on the polarizability which in turn is related to the first ionizions energies of the atoms. The total vdW force between the macroscopic tip and sample within the Hamaker's approximation [20] is obtained from

$$\mathbf{F}_{\mathrm{vdW}} = -\rho_{\mathrm{tip}}\rho_{\mathrm{sample}}\int_{\mathrm{tip}}\int_{\mathrm{sample}} \nabla V_{\mathrm{vdW}}(|\mathbf{r}-\mathbf{r}'|)d\mathbf{r}d\mathbf{r}', \tag{1.5}$$

where the integrals are taken over the volumes of tip and sample and ρ denotes the number density of atoms in them. One can rewrite the latter expression as

$$\mathbf{F}_{\mathrm{vdW}} = -\frac{H}{\pi^2}\int_{\mathrm{tip}}\int_{\mathrm{sample}} \nabla \frac{1}{|\mathbf{r}-\mathbf{r}'|^6} d\mathbf{r} d\mathbf{r}', \tag{1.6}$$

where $H = \pi^2 C_6 \rho_{\mathrm{tip}} \rho_{\mathrm{sample}} \sim 10^{-19}$ J is the Hamaker constant and depends on the material properties independently from the geometries. For instance, for the setup used in chapter 4, namely a Si tip over NaCl surface, the geometric mean of Hamaker constants for Si and NaCl extracted from tabulated values, [21, 22] gives $H = 1.17 \times 10^{-19}$ J.

Apart from H, the vdW force depends only on the geometries. For many simple geometries the integral in Eq. (1.6) can be evaluated. [23] In particular for the sphere-plane (mimicking tip-sample) separated by s one can use the approximation

$$F_{\mathrm{vdw}} = -\frac{HR}{6s^2} \tag{1.7}$$

where R is the sphere radius. For a spherical Si tip of radius 20 nm atop a NaCl flat surface at separation $s = 1$ nm, the vdW interaction is 0.39 nN; with $s = 0.5$ nm the force is 1.56 nN. For more realistic geometries including the tip cone see Ref. [24].

1.3 Electrostatic Interactions

The electrostatic interactions between the macroscopic bodies of probe and sample, in the most general form, reads

$$F_{\mathrm{el}} = g(s)V^2 \tag{1.8}$$

where $V = V_b - V_{\mathrm{CPD}}$ (Eq. (2.2)) is the electric potential difference and

$$g(s) = \frac{1}{2}\frac{\partial C(s)}{\partial s} \tag{1.9}$$

is a function of the macroscopic geometries. In chapter 3 we evaluate the capacitance C for realistic tip-cantilever-sample geometries. In order to get a feeling, we exemplify here again the simplest relevant geometry, namely a conducting spherical tip of radius R over a perfect conducting plane a distance s away. For this geometry the commonly used approximate expression [25] is

$$g(s) = -\frac{\pi\epsilon_0 R^2}{s(R+s)}, \tag{1.10}$$

where $\epsilon_0 = 8.85$ pN/V^2 is the vacuum permittivity. Then, for typical values e.g. $R = 20$ nm and $s = 1$ nm, the electrostatic attraction is 0.53 nN for $V = 1$ Volt; if $s = 0.5$ nm the force is 1.09 nN.

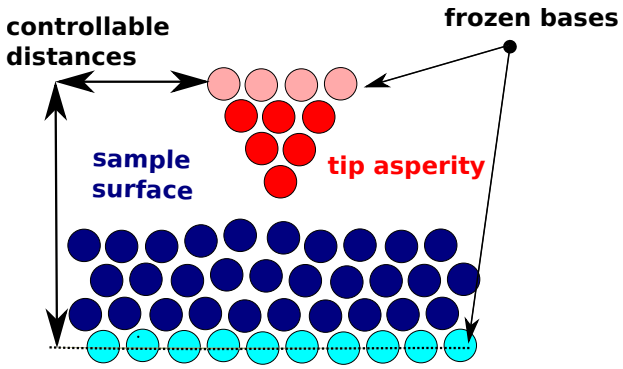

Figure 1.1: Illustration of the nano tip-sample system used to evaluate the short range interactions. Tip position is defined in terms of controllable distances between the frozen bases.

1.4 Chemical Interactions

The chemical forces, including covalent, ionic and/or metallic forces, can be repulsive or attractive in the relevant range of SPM. Indeed it is the chemical forces that provides the ability of high resolution imaging of the surface states. Because of their short-range nature, the interactions between only a few foremost atomic layers of tip and sample, as shown schematically in Fig. 1.1 need to considered in simulation. The base atoms of tip and of sample are kept frozen in their bulk positions and the remainder are allowed to relax in response to the chemical interactions upon changing the nominal relative position of the nano tip with respect to the sample. The latter position is defined based on the position of the frozen atoms.

The chemical force on the nano-tip is the vector sum of the forces on the individual atoms

$$\mathbf{F}_{\text{chem}} = \sum_{i \in \text{tip}} \mathbf{F}_i = \sum_{i \in \text{tip-base}} \mathbf{F}_i, \tag{1.11}$$

where $\mathbf{F}_i$ denotes the total force on atom i. Note that only the frozen atoms in the base contribute to the sum. The reason is that for any of the relaxed atoms $\mathbf{F}_{i,\text{relaxed}} = 0$ (at least up to a tolerance depending on the noise of the force evaluation).

Equivalently, one can use

$$\mathbf{F}_{\text{chem}} \equiv -\frac{\partial U_{\text{chem}}}{\partial \mathbf{R}_{\text{tip}}}, \tag{1.12}$$

where $\mathbf{R}_{\text{tip}}$ is the tip position with respect to the sample. Evaluating the potential energy of the chemical interactions U_{chem} is not easy in general. In particular, in case of KPFM, the electric field generated by the biased tip would polarize the atoms in the contact point and this way U_{chem} is a function of bias and macroscopic geometries.

1.4.1 Force fields

The simplest way to describe the chemical interactions is to approximate them with empirical force fields. Force field are sum of a few additive terms which describe the two or many particle interactions. If only pairs are considered, then

$$\mathbf{F}_{\text{chem}} = \sum_{i \in \text{tip}} \sum_{j \in \text{sample}} \mathbf{F}_{ij},$$

where $\mathbf{F}_{ij}$ is the force on atom i in tip from atom j in sample. For instance, the Lennard-Jones potential gives the interaction between two atoms as a function of their distance r_{ij}, namely

$$V_{\text{LJ},ij} = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}}. \tag{1.13}$$

(Then $U_{\text{chem}} = \sum_{i>j} V_{\text{LJ},ij}$.) The first term is repulsive but non-vanishing only at too short distances. It is supposed to describe the Pauli repulsion when the atoms are too close together such that the atomic orbitals overlap. The second term is attractive and like Eq. (1.4) describes the dispersion effects. If the free parameters A and B are fitted to experimental or quantum mechanical results, this simple potential can be a good approximation usable e.g. for inert gas molecules.

Another force field is the Morse potential [26]

$$V_{\text{Morse},ij} = D_0 \Big[1 - e^{-a(r_{ij} - r_{0,ij})}\Big]^2 \tag{1.14}$$

where $r_{0,ij}$ is the equilibrium distance at which the pairwise potential energy reaches its minimum D_0 and the constant a characterizes the widths of the potential. Again both repulsive and attractive terms exist.

There are more number of force fields that have been used in atomistic simulations of AFM. For instance, for ionic systems the shell model, which takes into account the polarization of the electron shells, has been implemented in the SciFi code. [27] Since we never used force fields in this work, we do not explore further this issue here, but as a final remark we mention that the interaction between the whole atomic cluster (mimicking the tip asperity) and the surface can proximately be described using e.g. the LJ or Morse potentials. The particle-particle distance is then replaced by the tip-sample distance and the free parameters are fitted to the more accurate calculations like those explained in the following.

1.4.2 Quantum mechanical methods

Empirical force fields are not able to describe the chemical interactions within the accuracy required for interpreting many of the atomically resolved images by the modern SPM techniques.

Indeed, physical and chemical properties of the systems in that range essentially depend on the details of the electron-electron and electron-core interactions. During the last century, the quantum mechanics theory has been shown to be pretty accurate to describe the material at this level. However, analytically solving the fundamental underlying equation, i.e. the Schrödinger's equation, is impossible except for trivial systems. The power of supercomputers along with considerable theoretical progress has been able to overcome, to some extend, this difficulty. By loosing some accuracy because of introducing some approximations, nowadays the numerical solutions to real-world problems are possible.

Density functional theory

Among a number of existing first-principles method for electronic structure calculations, the density functional theory (DFT) is applicable to extended systems as big as those required in simulations of AFM, i.e. containing few hundred atoms. Because of high computational demanding, one can do such calculations only for a limited number of tip positions. However, compared to other numerical methods to solve the Schrödinger's equation, the accuracy versus computational efficiency of DFT has made it a suitable tool for calculating accurately the chemical interactions in simulating the AFM. We also used DFT throughout this thesis.

DFT greatly simplifies the N-electron problem by mapping it into N one-electron Schrödinger's equations:

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{eff}}(\mathbf{r})\Big]\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}), \tag{1.15}$$

where $i = 1, 2, \cdots, N$. Each electron i feels a local effective potential V_{eff} generated by all electrons. In the Kohn-Sham [28] formalism, the effective potential is determined self-consistently through an iterative process, which is done numerically.

In principle, DFT is an exact theory. In practice, however, it is not exact because of required approximations for the exchange-correlations term. Nevertheless, in many applications including solid state physics, the accuracy is sufficient compared to the experiment. DFT is essentially based on two Hohenberg-Kohn theorems [29] showing that *electron density* $n(\mathbf{r}) = \sum |\phi_i(\mathbf{r})|^2$ can be considered as the basic quantity, hence the name of DFT. The theorems say that for a set of interacting electrons subject to an external potential V_{ext} there exists a ground-state electron density $n_0(\mathbf{r})$ which minimizes the functional

$$E[n(\mathbf{r})] = \mathcal{F}[n(\mathbf{r})] + \int n(\mathbf{r})V_{\text{ext}}(\mathbf{r})d\mathbf{r}. \tag{1.16}$$

Usually

$$V_{\text{ext}}(\mathbf{r}) = \sum_i^N \frac{q_i}{|\mathbf{r} - \mathbf{R}_i|}$$

i.e. identical to the electrostatic potential offered to the electrons by N nuclei of charges q_i and positions $\mathbf{R}_i$.

As a common practice, the universal functional

$$\mathcal{F}[n(\mathbf{r})] = T[n(\mathbf{r})] + \frac{1}{2}\iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{\mathrm{xc}}[n(\mathbf{r})]$$

includes contributions form the kinetic (T) and Hartree energies, as well as a contribution $E_{\mathrm{xc}}[n(\mathbf{r})]$ which represents the exchange and correlations (XC) between the electrons.

So far, there is no exact closed-form expression for the XC term. Different approximate forms have been suggested. The simplest, yet surprisingly good in many applications, is the local density approximation (LDA) in which the XC term depends merely on the electronic density in space

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n(\mathbf{r})] = \int n(\mathbf{r})\epsilon_{\mathrm{xc}}[n(\mathbf{r})]d\mathbf{r}$$

where ϵ_{xc} is the exchange-correlation density.

When doing calculations, more error sources come into play. One is the error due to modeling the electron-nucleus interaction by pseudopotentials. Another error is because of non-completeness of the basis set used to expand the wave function. In this work we used the BigDFT package [30] which uses a wavelet basis set. This makes it possible to reduce the mentioned error depending on how fine is the used grid.

Geometry optimization

In order to find the relaxed geometry of the atoms the total energy is minimized with respect to atomic positions $\mathbf{R}_i$. Within the Born-Oppenheimer approximation, the motion of nuclei and electrons are split to two independent problems; nuclei are frozen while when the electronic wave function is being calculated. Using methods such as the steepest descent or conjugate gradients, the atomic cores are moved according to the corresponding energy gradients (forces). Then, the electronic wave function has to be determined by solving the Schrödinger equation's for the new atomic positions; the forces are updated and the cores are accordingly moved and so forth until the forces become small enough. This makes the *ab initio* geometry relaxation computational demanding for large systems. Once the Hellmann-Feynman the free atoms are relaxed, the forces acting on the nano tip are summed up to get the chemical force exerted on the tip as function of the tip position (see Fig. 1.1).

Hellmann-Feynman forces

Fortunately, the Hellmann-Feynman theorem makes it possible to calculate all force components on all atomic cores at once from the converged ground state electronic wave function $\psi(\mathbf{r})$ at

each geometry relaxation step. According to this theorem, to get the gradient of the energy E with respect to an external parameter λ one calculates the expectation value of the gradient of the Hamiltonian $\hat{H}$ with respect to λ, namely

$$\frac{\partial E}{\partial \lambda} = \Big\langle \Psi(\lambda) \Big| \frac{\partial \hat{H}}{\partial \lambda} \Big| \psi(\lambda) \Big\rangle.$$

If the parameter λ is an atomic position, then the energy gradient is the force acting on the corresponding core. The forces acting on each atomic core has two sources, one from all other cores in the system and the other from the electrons which are determined by the electron density $n(\mathbf{r}) = \langle \psi | \psi \rangle$. Therefore, the force acting on atom i with core charge q_i and located at $\mathbf{R}_i$ is given by

$$\mathbf{F}_i = q_i \sum_{j \neq i}^{N} q_j \frac{\mathbf{R}_j - \mathbf{R}_i}{|\mathbf{R}_j - \mathbf{R}_i|^3} - q_i \int \frac{\mathbf{r} - \mathbf{R}_i}{|\mathbf{r} - \mathbf{R}_i|^3} n(\mathbf{r}) d\mathbf{r}.$$

Chapter 2

Principles of KPFM

2.1 Basic Concepts

2.1.1 Work function: metals versus insulators

For a conducting crystal, the work function corresponds to the energy difference between the vacuum level outside the surface at a distance large compared to the lattice spacing, yet small compared to the lateral dimensions of a homogeneous patch, and the bulk Fermi level. In this range, which is typical for conventional KPFM measurements, the potential acting on an electron outside the surface approaches the local vacuum level and becomes constant, except in the vicinity of surface steps or patch boundaries. Differences between local vacuum levels are solely due to electrostatic contributions which give rise to fringing electric fields around such boundaries.[31, 32] Below about a lattice spacing from the surface, the potential exhibits smooth 3D variations on the atomic scale.

If the sample is covered by a thin overlayer of foreign material, the work function can change owing to electron transfer and structural relaxation at the interface. [33] Similar changes can occur at the surface of a doped semiconductor, owing to band bending in a subsurface depletion layer. As long as electrochemical equilibrium occurs the Fermi level is aligned throughout the sample with the Fermi level of the back-electrode. However, if the sample is a wide-bandgap insulator, e.g. an alkali halide crystal, this equilibration may require very long times, so that the bulk Fermi level is not well-defined. Charge rearrangements and relaxation occur at the interface with the back electrode and cause an additive shift of the local vacuum level outside the surface with

respect to the Fermi level of the back electrode. In a real, thick enough insulator with charged impurities, such a shift will also be affected by the distribution of migrating charged defects at the interface, the surface and in the bulk of the sample. [8, 4]

2.1.2 Contact potential difference (CPD)

When two macroscopic objects characterized with different work functions, e.g. W_{tip} and W_{sample} of an AFM tip and the probed sample, are electrically connected, an electric current flows through the connection until the Fermi levels become aligned. The objects are then charged and a so-called *contact potential difference* (CPD) of

$$V_{\text{CPD}} = \frac{W_{\text{tip}} - W_{\text{sample}}}{e} \tag{2.1}$$

is developed between them, where $e = -1.60217657 \times 10^{-19}$ C is the elementary charge.[1] For an insulating sample W_{sample} must be referred to the Fermi level of the back-electrode but shifted due to the existence of the insulator (which is also affected by sample preparation), as explained above. What is interesting in practice is, however, the variation of the CPD throughout the surface, but not its absolute value.

The capacitor made out of the tip-sample combination has a capacitance $C(s)$ which depends on their geometries and relative positions, and in particular on their separation s. If the tip is biased at V_b with respect to the grounded sample, the *effective* potential difference between tip and sample is[2]

$$V = V_b - V_{\text{CPD}}. \tag{2.2}$$

The amount of the charge transfered because of electrical connection, namely

$$q(s, V_b, \text{CPD}) = C(s)(V_b - V_{\text{CPD}}) \tag{2.3}$$

would be zero if a bias voltage of $V_b = V_{\text{CPD}}$ is applied to compensate the CPD, as depicted in Fig. 2.1. Note that vanishing q (i.e. V) means also vanishing the electrostatic interaction between the capacitor electrodes (i.e. tip and sample) which reads $\frac{1}{2}\frac{\partial C}{\partial s}V^2$. This allows for measuring CPD and hence the work function of the sample with respect to the tip (whose work function is known).

KPFM is essentially based on the same concept; the electrostatic contribution to the oscillation of an atomically sharp tip scanned over the sample surface is minimized as much as

[1] Note that since $e < 0$, if the work function of a sample is higher than that of another sample, the CPD is also higher for the former.

[2] If the bias V_b is instead applied to the sample while the tip is grounded, then $V = V_b + V_{\text{CPD}}$.

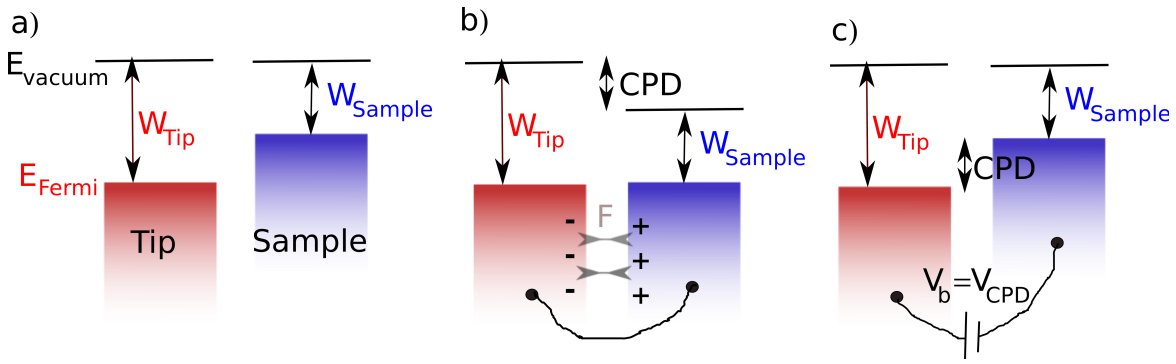

Figure 2.1: Illustration of CPD measurement. (a) Energy levels of electron in isolated tip and sample. The work function W is the difference between the vacuum and Fermi levels. (b) Upon electrical connection, the Fermi levels align via electron transfer towards the object with lower Fermi level (tip in this case). The amount of charge transfer and the resulting electrostatic interaction between the charged objects depends on CPD. (c) By compensating CPD with external bias, the charge and force are nullified. The applied bias therefore determines the CPD.

possible by tunning the DC bias, as detailed in the following section. This provides maps of the atomic-scale variation of the surface potential. When the resolved resolution is within atomic-scale [34, 35, 36, 37, 7] or less, [13, 12] one indeed measures the variation of the *local* CPD, known as the LCPD,

$$V_{\mathrm{LCPD}}(x,y) = \frac{W_{\mathrm{tip}} - W_{\mathrm{sample}}(x,y)}{e}, \tag{2.4}$$

where (x, y) denotes the lateral position of the tip over the sample surface. In practice, the detected LCPD also depends on the tip-height from the surface, as will be discussed in chapters 4 and 5.

2.2 Measurement

Since the capacitive electrostatic force is a quadratic function of V_b, so is also the corresponding contribution to the frequency shift. One way to measure CPD is therefore the direct method in which V_{CPD} is determined from the extremum of the parabolic curve $\Delta f_1(V_b)$ measured by slowly sweeping V_b at each measurement point.[38, 36, 39, 40]

Much faster and more sensitive measurement of LCPD is attained in the combined NCAFM-KPFM where both topography and LCPD are imaged simultaneously. [2] In most state-of-the-art NCAFM experiments a micro-fabricated cantilever with a tip at its free end (typically etched out of doped single-crystal silicon) oscillates with a constant amplitude at the frequency of a flexural

resonance, which is fundamental mode i.e. f_1. [41, 16] Distance-dependent tip-sample forces cause a frequency shift Δf_1 which can be very accurately measured [17] and used for distance control.

The applied bias consists of an AC voltage with angular frequency ω modulating the DC voltage:

$$V_b(t) = V_{\mathrm{DC}} + V_{\mathrm{AC}} \cos \omega t. \tag{2.5}$$

Assuming that the electric response is linear and in-phase with V_{AC}, the electrostatic force acting on the tip can be decomposed into three spectral components:

$$\begin{aligned} F(t) &= \frac{1}{2}\frac{\partial C}{\partial s} V_b^2 \\ &= F_{\mathrm{DC}} + F_\omega \cos \omega t + F_{2\omega} \cos 2\omega t \end{aligned} \tag{2.6}$$

where

$$F_{\mathrm{DC}} = \frac{\partial C}{\partial s}\Big(\frac{1}{2}(V_{\mathrm{DC}} - V_{\mathrm{CPD}})^2 + \frac{1}{4}V_{\mathrm{AC}}^2\Big), \tag{2.7}$$

$$F_\omega = \frac{\partial C}{\partial s}(V_{\mathrm{DC}} - V_{\mathrm{CPD}})V_{\mathrm{AC}}, \tag{2.8}$$

$$F_{2\omega} = -\frac{1}{4}\frac{\partial C}{\partial s}V_{\mathrm{AC}}^2. \tag{2.9}$$

In the Amplitude Modulation (AM) KPFM, [42] the second resonance mode of the cantilever is excited electrically (i.e. $\omega = 2\pi f_2$) while f_1 is used independently for the normal NCAFM operation. The KPFM-related signal is therefore the demodulated lever deflection at f_2. This signal is proportional to F_ω and thus to $(V_{\mathrm{DC}} - V_{\mathrm{CPD}})V_{\mathrm{AC}}$, as in Eq. (2.8). In the FM-KPFM, [43] on the other hand, the first resonance mode is used also for the KPFM operation and is electrically modulated at $\omega \ll 2\pi f_1$. Now, the detected signal corresponding to the KPFM operation is the modulation of the resonance frequency shift which is proportional to $\partial_s F_\omega$ and thus again proportional to $(V_{\mathrm{DC}} - V_{\mathrm{CPD}})V_{\mathrm{AC}}$.

Therefore, in either case the feedback circuit minimizes the corresponding detected signal by instantly adjusting $V_{\mathrm{DC}} = V_{\mathrm{CPD}}$ while scanning the tip parallel to the sample surface at a distance controlled by the (non-modulated) shift Δf_1. [16] V_{CPD} is recorded continuously and provides the CPD map of the scanned surface.

Because the scope of this thesis is theoretical, we do not further discuss experimental details e.g. on the signal-to-noise ratio. We only mention that both FM-KPFM, [34, 44, 45, 46, 36] and AM-KPFM [35, 37, 7] can detect lateral atomic-scale variations of V_{LCPD} in the range where Δf_1 exhibits similar variations on surfaces of semiconductors, as well as of ionic crystals.

2.2.1 Experimental evidence

When an AFM tip approaches a surface, short-range forces contribute to the tip-sample interaction and give rise to atomic-scale contrast in NCAFM. The short-range force component

perpendicular to the surface can be extracted from measurements of Δf_1 versus the closest tip approach distance d in an oscillation cycle.[47, 48] The contrast observed in V_{LCPD} in the same distance range cannot only be due to the long-range electrostatic force, but must be due to a short-range bias-dependent force.

Arai and Tomitori were the first to infer the existence of such a force from $\Delta f_1(V_b)$ curves recorded with a cleaned and sharpened silicon tip closer than 0.5 nm to a 7×7 reconstructed Si(111) surface. [49] In particular, above a Si adatom, they found a narrow peak growing with decreasing d superposed on the usual parabolic dependence around the plotted minimum of $-\Delta f_1(V_b)$ in their Fig. 1, i.e. for $V_b \simeq V_{\mathrm{CPD}}$. Later the same authors pointed out that an even sharper peak appeared at the same bias in the simultaneously recorded tunneling current. [50] This seemingly supported their original suggestion that the additional attractive force causing the peak in $-\Delta f_1(V_b)$ arose from the increased overlap due to the bias-induced energetic alignment of dangling bonds states localized at the tip apex and on Si surface adatoms. The formation of a covalent bond between those states has been shown to be responsible for the observed NCAFM contrast on the 7×7 reconstructed Si(111) surface. [51] In extensive recent measurements on the same system, however, Sadewasser *et al.* [36] reported parabolic $\Delta f_1(V_b)$ curves, but detected a rapid drop by about -1 V followed by a gradual increase in V_{LCPD} above a Si adatom with decreasing d in the range where the extracted short-range force showed a similar behavior. The apparent discrepancy with respect to Arai and Tomitori's observations is not so surprising because tunneling is seldom observed with clean silicon tips, although it is routinely measured in STM, as well as in NCAFM on conducting and even semiconducting samples when using metal-coated silicon tips. [52, 53]

An appreciable position- and distance-dependent DC tunneling current complicates the interpretation of LCPD measurements. This problem does not arise with insulating samples, but conversion to a DC transport current below the surface of a weakly doped semiconductor can cause a significant voltage drop within the sample owing to the finite bias required to compensate the LCPD. Especially in quasistatic measurements of $\Delta f(V_b)$ away from the compensation voltage, a strong DC electrostatic "phantom" force is generated which gives rise to atomic-scale contrast in NCAFM at separations where none is expected. [54, 55] Nevertheless, Arai and Tomitori's basic idea that bias-induced alignment of spatially localized surface states can lead to an enhanced site-dependent attractive force remains plausible even if a DC tunneling current cannot be sustained. Thus Krok and coworkers [46] suggested that the lower LCPD which they found across protruding In rows on the c(2×8) reconstructed InSb(001) surface was due to a bias-induced local electron transfer from a polar dangling bond on the electronegative Sb atom presumably picked by the Si tip to the nearest electropositive surface In atoms. The same authors also showed that the LCPD contrast between different lateral positions decays exponentially with increasing $d < 1$nm.

2.3 Theoretical Aspect

Understanding the connections between the observed contrast in V_{LCPD} and the atomic-scale variations of the electrostatic potential just outside the surface has been a challenging task, especially on unreconstructed cleavage faces of rocksalt-type crystals. [37] Above a flat homogeneous surface V_{LCPD} must, in principle, approach the corresponding V_{CPD} at somewhat larger tip-sample separations. In practice, however, this ideal behavior is often masked by a slow dependence caused by the finite lateral resolution of surface inhomogeneities, e.g. islands of materials with different work functions. This effect is less pronounced in FM- than in AM-KPFM. [56, 57, 46, 58]

2.3.1 Previous models

Several researchers developed models and computational schemes based on classical electrostatics which treated the tip and the sample (sometimes also the cantilever) as macroscopic bodies in order to interpret the resolution of KPFM images of inhomogeneous surfaces on lateral scales of several nanometers and above. [59, 60, 61, 62, 63, 64, 65, 66, 67] On the other hand, only few authors considered atomistic nano-scale tip-sample systems, either neglecting [36, 68, 69] or including the macroscopic contributions via simple approximations. In the first theoretical study of combined NCAFM-KPFM on an ionic crystal sample, [37, 70, 5] a formally correct partitioning was proposed between capacitive and short-range electrostatic forces induced by the effective macroscopic bias V. This analytic treatment also provided qualitative insights into the origin of atomic-scale LCPD contrast, although underestimating the capacitive force caused a quantitative disagreement with experimental results.

More reliable results were obtained for a NaCl(001) sample interacting with a model tip consisting of a conducting sphere terminated by a small charged NaCl cluster by allowing local atomic deformations. [71] These atomistic simulations were based on the SciFi code [27] which has provided detailed insights into NCAFM on ionic compounds. [72, 73]

The few simulations of KPFM based on DFT computations have been concerned with silicon model tips interacting with reconstructed Si surfaces, both clean and containing substitutional impurities. None of those purely microscopic calculations took into account the macroscopic capacitive contribution to the KPFM signal, however. Thus Sadewasser *et al.* [36] obtained qualitative agreement between variations of the perpendicular dipole moment and of the local chemical potential of their microscopic subsystem, both computed at zero bias, and the distance dependence of V_{LCPD} measured by FM-KPFM on the benchmark-like Si(111) 7×7 surface. Masago and coworkers [68] defined the V_{LCPD} within a tight-binding based DFT formalism as the difference between the Fermi levels (electrochemical potentials) of their tip and sample microscopic subsystems, which were forced to carry opposite charges determined so as to minimize the

force on their nanotip. Although overlap between tip and sample wave-functions was neglected, qualitatively correct V_{LCPD} images of charged surface and subsurface defects were obtained based on Coulomb interactions between Mulliken charges treated as point objects. Very recently the same authors included wave-function overlap to first order [69], and generated V_{LCPD} images showing partially occupied dangling bonds on the on the 5×5 analog of the Si(111) 7×7 surface at a smaller distance (0.4 nm) where a covalent bond begins to form between a dangling bond at the tip apex and a Si adatom.

Whereas bias-induced electron transfer is plausible for narrow-bandgap semiconductors like those previously mentioned, it is unlikely for overall neutral cleaved (001) surfaces of wide-bandgap insulators like alkali halides which neither have gap states, nor are reconstructed, but are only weakly rumpled. [74] In Ref. [37] the atomic-scale LCPD contrast observed on KBr(001) was attributed to opposite surface cation and anion displacements in response to local electric fields induced by the macroscopic (in accordance with our definition) field. However, the authors approximated the electric field E_z by V/R i.e. that at the surface of an isolated conducting spherical tip of radius R, the local unit cell polarizability by the bulk crystal (Clausius-Mossotti) expression, and neglected the macroscopic surface polarization. Although essentially constant on the scale of a nanometer-size nanotip, the latter, together with E_z is actually nonuniform on a lateral scale of order $\sqrt{Rs}$ for separations $s \ll R$. They evaluated the macroscopic and microscopic surface charges densities σ_m and σ_μ induced on a conducting model tip by their E_z and by the displaced surface ions, respectively. Using Eq. (3.30) they computed the modulation of the electrostatic force. After further justified approximations, they obtained opposite LCPDs above cations and anions which *increased exponentially* with d. In a subsequent article, [70] the same authors added a macroscopic force roughly representing the interaction of the cantilever with the back electrode, but still obtained a surprisingly large maximum in the absolute LCPD for $d \simeq 0.6$ nm. In a subsequent publication [71], more reliable results were obtained with the SciFi code [27] for a cubic NaCl cluster partly embedded into a conducting sphere interacting with a NaCl(001) sample similar to ours via empirical shell-model potentials. Cluster ions inside the sphere were fixed while the protruding cluster ions formed a small nanotip with a net charge +e at the apex. The protruding ions were allowed to relax and to induce image charges in the electrodes. The results obtained can be considered representative of what is expected for a small, charged nanotip interacting with an ionic crystal. A common justification for such a model is that real tips often pick up sample material and that simulations based on the same code produced reasonable results when compared to NCAFM measurements on ionic crystals, alkali halides in particular. [72, 73] However, in those and in previous SciFi computations [75] using a *larger protruding ionic nanotip* against an *overall neutral defect-free sample slab*, the inclusion of electrode polarization was deemed unnecessary. Indeed, contributions from images charges of close anion-

cation pairs tend to cancel out if $R \gg s$. More importantly, according to the Supplementary Material of Ref. [76], the distance dependence of V_{LCPD} calculated analytically for a single *charge or dipole fixed* below a conducting sphere facing a biased planar counter-electrode coincides with the results of full SciFi computations including image charges. This is observed down to a separation s of 0.7 nm for a small charged nanotip similar to that assumed in Ref. [71], whereas the much smaller V_{LCPD} obtained for the larger nanotip assumed in the above-mentioned SciFi computations is compatible with a smaller permanent dipole moment. The coincidence implies that at larger separations the electrode polarizations induced by the charge q and by the effective bias $V = V_b - V_{\mathrm{CPD}}$ are decoupled. Coupling presumably arises at separations approaching interatomic distances where ions (or atoms) inside the microsystem become appreciably polarized (electronically and/or owing to induced displacements) by local fields [77], thus leading to the site-dependent LCPD contrast superposed on the z-dependent long-ranged LCPD obtained in Ref. [71]. In the absence of localized, point-like net charges (or permanent dipoles) in the microsystem, the resulting force due to polarization of the microsystem and of the electrodes is proportional to V^2. It is, however, overwhelmed by the macroscopic capacitive contribution if the nanotip dimensions are small compared to R.

2.3.2 Present model

Earlier studies mentioned that the short-range tip-sample interaction is bias-dependent but provided no recipe to investigate it theoretically. Moreover, they did not clarify how long-range and short-range bias-dependent forces are connected and the role of each in the observed KPFM signals. We answer all of these questions and obtain and analyze in detail theoretical expressions for the site-dependent LCPD. Our approach is not limited to particular materials, but results are presented for the system described in chapter 4 which is representative of a neutral, but polarizable reactive clean Si tip interacting with an ionic crystal.

In the present work, which is based on separate classical electrostatics and *ab initio* calculations, we propose a more rigorous and accurate approach for coupling interactions acting on widely different length scales. The bias voltage is applied between the conducting probe and the grounded back electrode below the dielectric sample. The electric field generated between the macroscopic tip and sample, obtained from classical electrostatic treatment, is applied to the microscopic system consisting of a protruding nanotip a few Å away from a slab sample. Taking into account both interactions, an unambiguous definition of the Kelvin signal is provided. Based on highly accurate density functional calculations for nano-scale tip-sample systems for a realistic Si tip close to a NaCl(001) slab as an example of current interest, we conclude that the slope of the microscopic short-range force is the basic quantity that should be extracted from KPFM measurements. Expressions for V_{LCPD} in AM- and FM-KPFM are obtained and evaluated, first

for ultrasmall, then for finite tip oscillation amplitude A. Their magnitude and dependence on A are explained in detail. Experimental limitations and evidence for the predicted trends, as well as desirable measurements are also briefly discussed.

Chapter 3

Electrostatic Interactions

Among various interactions with the probing tip, the ubiquitous long-range electrostatic (ES) force is of special importance in SPM techniques involving voltage-biased conducting tips. [78] In electrostatic force microscopy (EFM) [79] the ES force is directly measured, whereas in Kelvin probe force microscopy (KPFM) [80] the contact potential difference (CPD) is mapped by compensating an AC signal related to the ES force. Those techniques, as well as scanning capacitance microscopy (SCM) can be used to determine the local charging properties of dielectric samples or surface layers and of semiconductor devices protected by insulating layers. In particular, two-dimensional electron gas (2DEG)-based structures can be controlled by voltages applied to a back-electrode (gate) and to top gates confining the 2DEG laterally. SCM is being extensively used to map lateral doping profiles, [81] but can also detect the *quantum capacitance* due to the occupation of confined electronic states e.g. in the quantum Hall effect. [82] Furthermore, scanning gate microscopy (SGM) studies, where a biased conducting tip locally perturbs electron waves or shifts the levels of confined states past the Fermi energy, have allowed one to map induced variations in the conductance of quantum constrictions, [83] quantum dots [84] and of increasingly complex structures of current interest.

In particular in the KPFM experiments, owing to electric field penetration into the dielectric sample, the tip shank and the cantilever significantly affect the capacitive force and its gradient even at sub-nanometer tip-surface separations where atomic-scale contrast appears. A main outcome is the electric field as function of the bias and probe-sample geometry which would be included into atomistic calculations.

In this chapter, we first investigate the electrostatics of a model spherical tip over dielectric samples from an analytic point of view. Then, a more realistic tip-lever geometry is treated numerically.

I. Analytic Approach

Typical insulator thickness h, tip radius R and average tip-sample distance s being of the same order (tens of nanometers) in such measurements, their lateral resolution has often been roughly assumed to be $\sim R$. Although 3D numerical solutions of the Poisson equation yield an accurate description of the tip-induced electrostatic potential for each particular probe-sample geometry, analytic expressions for the capacitance, the ES force and its vertical gradient as functions of R, s and h are highly desirable. The same holds also for the electric field profiles at the top and bottom surfaces of a uniform dielectric layer in contact with a flat back-electrode. Their widths provide useful estimates of the lateral resolution of local CPD or surface charge variations, but also of tip-induced conductance changes in buried semiconductor devices, at least if the field distribution at the interface is narrower than the structure lateral dimensions but exceeds the lateral screening length.

For flat conducting samples, Hudlet *et al.* [25] proposed an approximate analytic model which is surprisingly accurate, also for atomically thin insulating layers on metals. [76] Dielectric samples or layers on a conducting back-electrode, however, have been predominately simulated numerically because of the complexity introduced by partial field penetration (see, e.g., Refs. [85, 86] and references therein). In this chapter we obtain exact and novel approximate analytic results for a spherical model tip facing such a slab. The derived expressions can be used for further analysis of experiments on the above-mentioned types of samples. Our formalism can also be generalized to multilayer slabs.

3.1 Spherical tip atop a semi-infinite dielectric

The classical electric potential between a cylindrically symmetric conducting probe at potential V facing a dielectric slab grounded on the bottom can be calculated by means of the image charges method. In order to obtain an analytic solution, we model the tip as a sphere of the same radius R as the apex, see Fig. 3.1(a). Additional contributions to the capacitance C and the ES force F from the tip shank and the supporting sensor can be important for thick dielectric slabs. [86] However, they vary more slowly than the contribution from the tip apex, so that our approximation captures the main features of ES properties at tip-sample distances $s < R$.

The problem of a sphere facing a semi-infinite dielectric is solved by combining two textbook

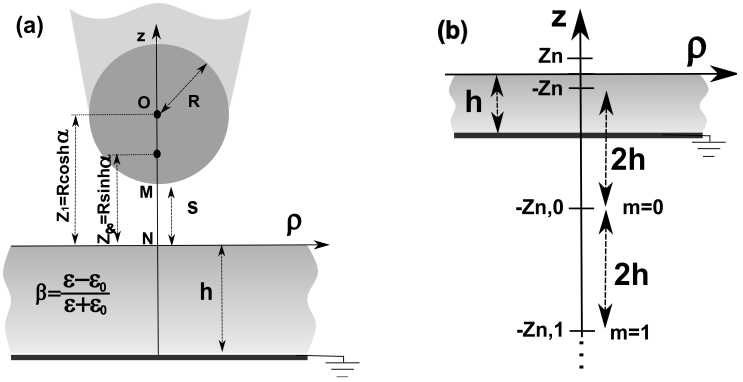

Figure 3.1: (a) The probe tip modelled as a conducting sphere of the same radius R as its apex at a separation s from a dielectric slab of thickness h. The tip is biased at V while the bottom back electrode is grounded. Image charges which sum up to the charge on the sphere are located between z_1 and z_∞. (b) Positions of image charges below the surface of the dielectric slab due to a point charge at z_n.

problems, [87] namely a point charge q at a distance r from the center of the conducting sphere or at a distance z_q from the surface of the dielectric. If an image charge $-qR/r$ is placed at a distance R^2/r from the sphere center on the same radial line as q, the sphere surface is an equipotential. In the second problem, the electric potential outside the dielectric can be obtained by adding the Coulomb potentials of q and of an image charge $-\beta q$ at $-z_q$ on the normal to the surface, where

$$\beta = \frac{\epsilon - \epsilon_0}{\epsilon + \epsilon_0},$$

ϵ and ϵ_0 being the permittivities of the dielectric and of vacuum (or that of the external medium), respectively. The potential inside the dielectric is that of single point charge $(1-\beta)q$ at z_q if that region were vacuum (or equivalently, of a point charge $(1+\beta)q$ at z_q if the whole space were filled with the dielectric). Physically, each image charge represents the effect of the polarization induced at the surface of the sphere or of the dielectric.

In the combined problem, i.e. sphere against dielectric, a charge $q_1 = 4\pi\epsilon_0 RV$ located at the center of the sphere ($z_1 = R + s$) tends to make the surface of the sphere an equipotential at V. The image $-\beta q_1$ at $-z_1$ below the dielectric surface, however, modifies the potential on the sphere surface. A second charge $q_2 = \beta q_1 R/2z_1$ is then placed at $z_2 = z_1 - R^2/2z_1$ to bring the sphere potential towards V, which induces in turn an image $-\beta q_2$ at $-z_2$ and so forth. The

resulting convergent series of point charges inside the sphere

$$q_{n+1} = \frac{\beta q_n R}{z_1 + z_n} \qquad (q_1 = 4\pi\epsilon_0 RV) \tag{3.1}$$

$$z_{n+1} = z_1 - \frac{R^2}{z_1 + z_n} \qquad (z_1 = R + s) \tag{3.2}$$

together with their corresponding images below the dielectric surface $\{-\beta q_n, -z_n\}$ satisfy the boundary conditions both on the sphere and dielectric surfaces. The attractive force on the sphere can be obtained by summing the Coulomb forces between the charges inside the sphere and their images inside the dielectric

$$F(s,V) = \frac{1}{4\pi\epsilon_0} \sum_{n,n'=1}^{\infty} \frac{-\beta q_n q_{n'}}{|z_n + z_{n'}|^2}. \tag{3.3}$$

The Green's function (GF) of the surface is

$$G_n^{\pm} = \frac{1}{\sqrt{\rho^2 + (z - z_n)^2}} - \frac{\beta}{\sqrt{\rho^2 + (z \pm z_n)^2}} \tag{3.4}$$

where G_n^+ and G_n^- refer to $z \geq 0$ and $z \leq 0$ regions, respectively. The electric potential $\Phi(\rho, z) = \frac{1}{4\pi\epsilon_0} \sum q_n G_n$ and the electric field $\mathbf{E} = \frac{-1}{4\pi\epsilon_0} \sum q_n \nabla G_n$ can be obtained outside the sphere, above or inside the dielectric slab.

3.1.1 Electric field profile

The z-component of the electric field just above the surface of the sample ($z = 0$)

$$E_z \;=\; \frac{1+\beta}{4\pi\epsilon_0} \sum_{n=1}^{\infty} \frac{q_n z_n}{(\rho^2 + z_n^2)^{3/2}} \tag{3.5}$$

is especially relevant in AFM and STM experiments because it polarizes atoms or ions and thus set up microscopic local fields which influence atomic-scale contrast. [77] For tip-sample separations s where such contrast appears, E_z approaches a uniform value $E_N \equiv E_z(\rho = 0)$ and can be inserted into atomistic model potential [70, 77] or *ab initio* simulations. [86]

Figure 3.2(a) shows how E_z gradually weakens as ρ increases. Its effective width can be characterized by $\rho_{1/2}$ at which $E_z = E_N/2$. When $\beta \simeq 0$ as well as for $s > R$, $\rho_{1/2}$ approaches the point-charge-like asymptotic linear relation $\sqrt{2^{2/3} - 1}(R + s) \simeq 0.766(R + s)$, as can be seen in Fig. 3.2(b). Over a perfect conductor ($\beta = 1$), on the other hand, $\rho_{1/2} \simeq \sqrt{2Rs}$ for sufficiently small s/R, as seen in the inset, hence formally vanishes upon contact because E_N then becomes infinite. Keeping in mind that for commonly used solid dielectrics [88] $\beta \geq 0.6$, the behavior highlighted in the inset indicates that $\rho_{1/2}$ *is considerably smaller than* R *if* $s \ll R$.

An alternative definition of the half-width directly related to the capacitance is based on the charge distribution at each surface. We define an effective area S^* such that

$$D_\perp^* S^* \equiv q^* = \int_S D_\perp dS,$$

where q^* is the total charge on the surface of area S, $D_\perp$ is the normal component of the electric displacement vector and $D_\perp^* \equiv D_{\perp,max}$. First we calculate the angular half-width θ_C for a conducting sphere separated by s from a semi-infinite dielectric. We have

$$\begin{aligned} q^* &= \sum_{n=1}^{\infty} q_n \\ S^* &= 2\pi R^2(1-\cos\theta_C) \\ D_\perp^* &= \epsilon_0 E_z(0,s) = \frac{1}{4\pi}\sum_{n=1}^{\infty} q_n\Big(\frac{1}{(z_n-s)^2}+\frac{\beta}{(z_n+s)^2}\Big). \end{aligned}$$

If $\beta = 0$, $z_n - s = R$ and $\theta_C = \pi$ as required for an isolated sphere. For a perfect conductor ($\beta = 1$), on the other hand, $\theta_C \simeq 45°$ at a separation $s = R/10$, as seen in Fig. 3.3 (a). These plots can be used to judge whether the approximation of the tip by a sphere is justified, e.g. θ_C should at least be smaller than 90° minus the cone half-angle for a conical tip terminated by a spherical cap.

Over the sample surface, we consider instead the polarization surface charge density $P_\perp = (\epsilon - \epsilon_0)E_z^-(\rho,0)$, [87] where $E_z^-(\rho,0) = \frac{\beta-1}{4\pi\epsilon_0}\sum q_n z_n/(\rho^2+z^2)^{3/2}$, thus

$$\begin{aligned} q^* &= -\beta\sum_{n=1}^{\infty} q_n \\ S^* &= \pi\rho_P^2 \\ P_\perp^* &= \frac{-\beta}{2\pi}\sum_{n=1}^{\infty}\frac{q_n}{z_n^2}. \end{aligned}$$

Therefore

$$\rho_P = \sqrt{2\sum_{n=1}^{\infty} q_n \Big/ \sum_{n=1}^{\infty}\frac{q_n}{z_n^2}}. \tag{3.6}$$

For a single point charge at z_n, we obtain $\rho_P = \sqrt{2}z_n$ (independent of β) which also coincides with the value at which the parallel component E_ρ is maximum. When $s \gg R$ or $\beta \ll 1$, ρ_P of the sphere approaches the point charge-like asymptotic linear behaviour, i.e. $\rho_P = \sqrt{2}(s+R)$, because then all charges vanish except q_1 at $z_1 = s + R$. As shown in Fig. 3.3 (b), for a perfect conductor ($\beta = 1$) $\rho_P \simeq \sqrt{2sR}$, i.e. like $\rho_{1/2}$, and is significantly smaller than R if $s \ll R$.

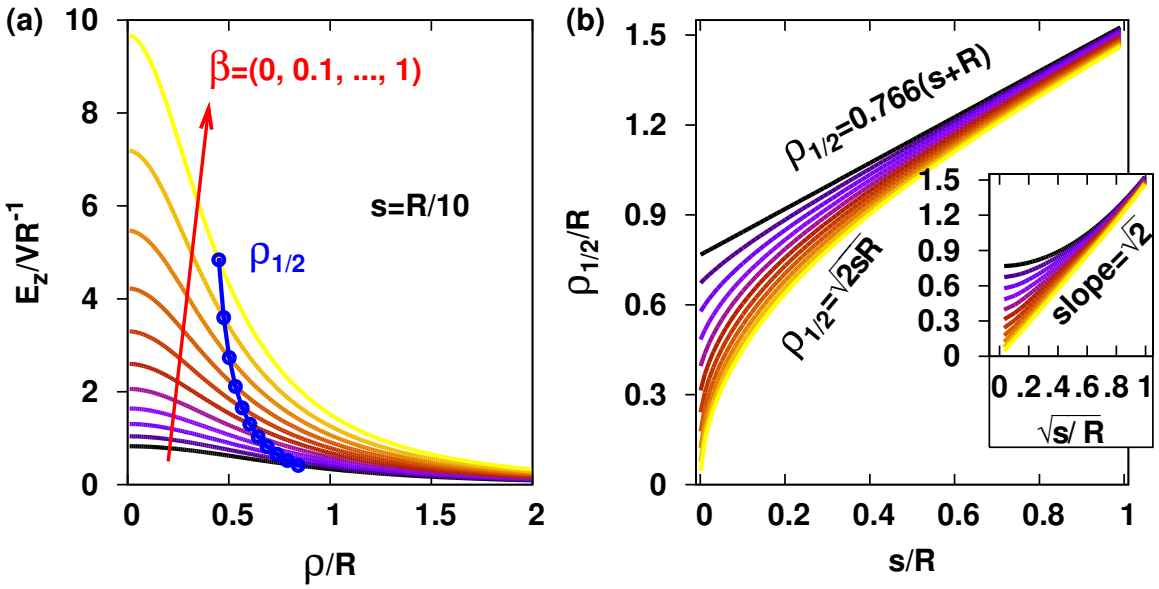

Figure 3.2: (a) Lateral profile of E_z just outside the sample and its half-width $\rho_{1/2}$ for a conducting sphere of radius R separated by $s = 0.1R$ from a semi-infinite dielectric. Blue points indicate $\rho_{1/2}$ at which E_z is half of $E_N = E_z(\rho = 0)$. (b) Variation of $\rho_{1/2}\ /R$ against s/R and $\sqrt{s/R}$ (inset); the curves are bounded by $\rho_{1/2} = \sqrt{2Rs}$ at small s/R for $\beta = 1$ and $\rho_{1/2} = 0.766(s + R)$ for $\beta = 0$ or at large s/R for all β.

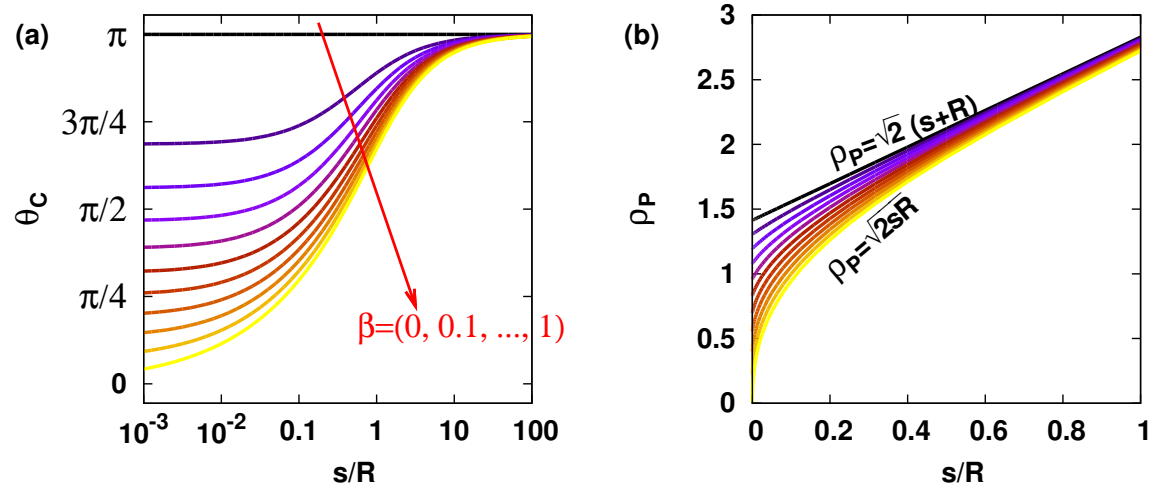

Figure 3.3: θ_C (a) and ρ_P (b) as a function of s/R for a sphere facing a semi-infinite dielectric. In (b) the curves are bounded by $\rho_P = \sqrt{2}(s + R)$ for $\beta = 0$ and $\rho_P \simeq \sqrt{2sR}$ for $\beta = 1$ and $s \ll R$.

3.1.2 Closed-form expressions for capacitance, force, force gradient and field

Following Ref. [89], one can write Eq. (3.1) as a second order homogeneous difference equation with constant coefficients

$$\frac{1}{q_n} = \frac{2\cosh\alpha}{\beta}\frac{1}{q_{n-1}} - \frac{1}{\beta^2}\frac{1}{q_{n-2}} \ (n \geq 3)$$

the solution of which is a linear combination of $\exp(\pm n\alpha)$ where

$$\cosh\alpha = \frac{z_1}{R} = 1 + \frac{s}{R}.$$

Substituting the known expressions for q_1 and q_2, we obtain

$$q_n = 4\pi\epsilon_0 RV \sinh\alpha\Big(\frac{\beta^{n-1}}{\sinh n\alpha}\Big), \tag{3.7}$$

$$z_n = R\sinh\alpha\coth n\alpha \quad (n \geq 1), \tag{3.8}$$

which are a simple generalization of the solution for a semi-infinite conductor ($\beta = 1$) [89] for arbitrary β and provide a convenient expression for the capacitance

$$C = \frac{1}{V}\sum_{n=1}^{\infty} q_n = 4\pi\epsilon_0 R\sinh\alpha \times \sum_{n=1}^{\infty}\frac{\beta^{n-1}}{\sinh n\alpha}. \tag{3.9}$$

Being the capacitance $C_{sph} = 4\pi\epsilon_0 R$ of the sphere alone in vacuum, the constant leading term in C can be ignored because only variations of C (i.e. of $C - C_{sph}$) with tip position are of interest.

Corresponding expression for the attractive electric force on the tip ($F = C'V^2/2$ where $C' = dC/ds$) is given by

$$F = 2\pi\epsilon_0 V^2 \sum_{n=2}^{\infty}\frac{\beta^{n-1}}{\sinh n\alpha}\big(\coth\alpha - n\coth n\alpha\big). \tag{3.10}$$

Similarly, the vertical force gradient (dF/ds) is given by

$$F' = \frac{2\pi\epsilon_0 V^2}{R\sinh\alpha}\sum_{n=2}^{\infty}\frac{\beta^{n-1}}{\sinh n\alpha}\Big[\frac{n^2}{\sinh^2 n\alpha} - \frac{1}{\sinh^2\alpha} + n\coth n\alpha(n\coth n\alpha - \coth\alpha)\Big]. \tag{3.11}$$

In dynamic EFM or SCM experiments with stiff deflection sensors F' is proportional to the resonance frequency shift which is used to control the tip-sample separation s. [16]

Finally, the maximum electric field outside the surface, i.e. $E_N \equiv E_z(\rho, z = 0)$, reads

$$E_N = \frac{V}{R}\Big(\frac{1+\beta}{\sinh\alpha}\Big)\sum_{n=1}^{\infty}\frac{\beta^{n-1}\sinh n\alpha}{\cosh^2 n\alpha}. \tag{3.12}$$

3.1.3 Limiting values

Except for α=0 and β=1, q_n decays exponentially towards zero. For an ideal conductor ($\beta = 1$) these expressions diverge in the limit $s \to 0$ (i.e. $\alpha \to 0$). For dielectrics, the resulting series for all quantities converge. We obtain $q_n(s = 0) = q_1\beta^{n-1}/n$, $z_n(s = 0) = R/n$ and

$$C_{max} = -4\pi\epsilon_0 R\Big(\frac{\ln(1-\beta)}{\beta}\Big). \tag{3.13}$$

Like C, $-F$ and F' are monotonically decreasing functions of s. Their upper bounds attained at $s = 0$ ($\alpha = 0$), namely

$$\begin{aligned} F_{max} &= -\frac{2}{3}\pi\epsilon_0 V^2\Big[\frac{\ln(1-\beta)}{\beta} + \frac{1}{(1-\beta)^2}\Big], \\ F'_{max} &= \frac{4\pi\epsilon_0 V^2}{45R}\Big[\frac{\ln(1-\beta)}{\beta} + \frac{1}{(1-\beta)^2} + \frac{21\beta}{(1-\beta)^4}\Big], \end{aligned}$$

are finite if $\beta < 1$ as shown in Fig. 3.4. The result for $F(s{=}0)$ is stated without proof in Ref. [90]. Finally,

$$E_{N,max} = \frac{V}{R}\frac{1+\beta}{(1-\beta)^2}.$$

They all provide useful upper bounds on the corresponding quantities in case of a realistic probe tip of apex radius R. [86] Their dependences on β are plotted in Fig. 3.4. As an example, the limiting values for NaCl ($\epsilon_r = 5.9$, $\beta = 0.71$) are $C_{max}/\pi\epsilon_0 R = 6.98$, $F_{max}/\pi\epsilon_0 V^2 = -6.77$ (i.e. $F_{max} = -0.188$ nN/V^2 independent of the sphere radius), $F'_{max}/\pi\epsilon_0 V^2 R^{-1} = 188.7$ and $E_{N,max}/VR^{-1} = 20.4$.

3.1.4 Approximate expressions

For $s/R < 0.1$ many terms in Eqs. (3.9-3.12) are required to get a reasonable accuracy. However, the truncation error is dramatically reduced, as illustrated in Table I, by adapting a procedure proposed for the sphere-conductor problem. [27]. Equation (3.8) shows that z_n approaches $z_\infty = R\sinh\alpha$ as $\exp(-2n\alpha)$ whereas q_n decays as $\beta^{n-1}\exp(-n\alpha)$. If the charge series is truncated at some q_k for which $z_{n>k} \simeq z_\infty$, the remainder can be summed up analytically as a correction term

$$q^{(k)}_{corr} \equiv \sum_{n=k+1}^{\infty} q_n \simeq \frac{q_{k+1}}{1-\beta e^{-\alpha}} = \frac{q_1\beta^k/(1-\beta e^{-\alpha})}{\sinh((k+1)\alpha)/\sinh\alpha} \tag{3.14}$$

lumped at z_∞ where we have used $q_{n+1}/q_{n>k} \simeq \beta R/(z_1 + z_\infty) = \beta e^{-\alpha}$. Moreover, Eq. (3.14) leads to compact, accurate analytical expressions. For example

$$C^{(1)} - C_{sph} \simeq 2\pi\epsilon_0 R\Big(\frac{\beta/(1-\beta e^{-\alpha})}{\cosh\alpha}\Big), \tag{3.15}$$

$$C^{(2)} - C_{sph} \simeq 2\pi\epsilon_0 R\Big(\frac{\beta}{\cosh\alpha} + \frac{\beta^2/(1-\beta e^{-\alpha})}{4\cosh^2\alpha - 1}\Big) \tag{3.16}$$

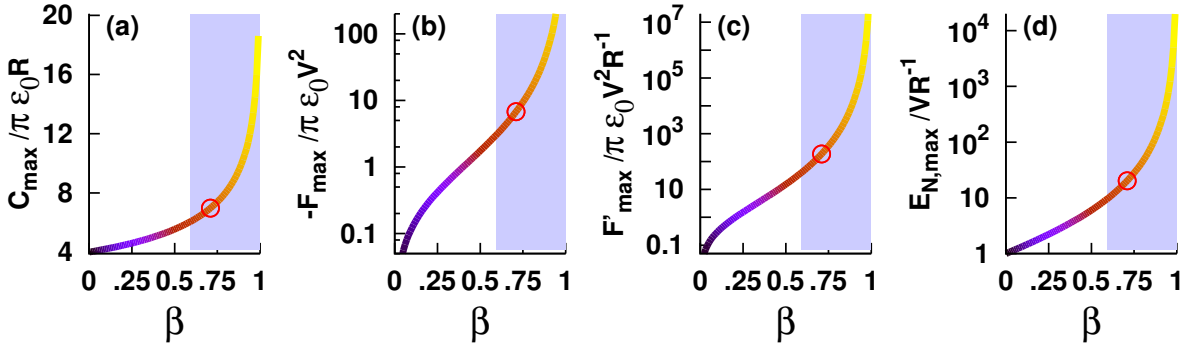

Figure 3.4: The upper bounds of C, $-F$, F' and E_N (in units $\pi\epsilon_0 R$, $\pi\epsilon_0 V^2$, $\pi\epsilon_0 V^2 R^{-1}$ and VR^{-1}, respectively) as a function of β, attained for a conducting sphere touching ($s = 0$) the semi-infinite dielectric surface. The relevant range for typical values [88] $\beta_{\rm PMMA} \simeq \beta_{\rm SiO_2} = 0.59$, $\beta_{\rm NaCl} = 0.71$ (circles), $\beta_{\rm Al_2O_3} = 0.80$, $\beta_{\rm Si} = 0.86$, $\beta_{\rm HfO_2} = 0.92$ and $\beta_{\rm LiNbO_3} = 0.93 - 0.98$, is shaded.

Table 3.1: Relative error in calculating C using k point charges in addition to $q^{(k)}_{corr}$.

	$\beta = 0.71$ (NaCl)				$\beta = 1$ (conductor)				
s/R	k=1	2	10	20	1	2	10	20	Eq. (3.17)
0.1	.02	.004	10^{-7}	10^{-15}	.05	.01	10^{-7}	10^{-13}	0.02
0.2	.008	.001	10^{-11}	$<10^{-16}$	.02	.002	10^{-8}	$<10^{-16}$	0.02
0.5	.001	10^{-5}	10^{-16}	$<10^{-16}$	.003	10^{-4}	10^{-14}	$<10^{-16}$	0.01

approximate Eq. (3.9) within 5% and 1%, respectively, as seen in Table 3.1. Our novel formulas should be useful in theoretical modeling and data interpretation. Without the correction, Eq. (3.16) becomes $2\pi\epsilon_0 R\beta/(1 + s/R)$ with an error larger than 33% at $s = R/10$. However, it gives the correct asymptotic behavior $C - C_{sph} = 2\pi\epsilon_0\beta R^2/s$ for $s \gg R$. The prefactor of the approximation proposed by Hudlet *et al* [25] for a conducting sample, i.e.

$$C - C_{sph} = 2\pi\epsilon_0 R \ln\left(1 + \frac{R}{s}\right), \tag{3.17}$$

was adjusted to match this asymptotic behavior, although their basic assumption (constant field along field lines perpendicular to the sphere and sample surfaces) is reasonable only for small s/R. Nevertheless, Eq. (3.17) is remarkably accurate at moderate separations, as demonstrated in the last column of Table I.

3.2 Point charge above a dielectric slab

Next we consider the intermediate problem of a point charge at ($\rho = 0, z_n \geq 0$) against a dielectric slab of thickness h; see Fig. 1(b). When $h \to \infty$, the GF is given by Eq. (3.4). For finite h, G_n is, however, modified because the field lines become perpendicular to the surface of the back-electrode. Appropriate expressions are derived and plots of the resulting field profiles and of their half-widths are presented below. Compared to the treatment in Ref. [91], the GF approach is more convenient, especially for extending to the case of a multilayer slab for which similar boundary conditions are applied at each interface.

3.2.1 Green's function

In cylindrical coordinates, the Green's function (GF) above and below the slab surface has the form [87]

$$G_n^+ = \int_0^\infty \Big(e^{-k|z-z_n|} + A_1 e^{-kz}\Big) J_0(k\rho) dk, \tag{3.18}$$

$$G_n^- = \int_0^\infty \Big(A_2 e^{-kz} + A_3 e^{+kz}\Big) J_0(k\rho) dk, \tag{3.19}$$

J_0 being the zero order Bessel function of the first kind. Recall that

$$\int_0^\infty e^{-k|z-z_n|} J_0(k\rho) dk = \frac{1}{\sqrt{\rho^2 + (z-z_n)^2}}.$$

The boundary conditions [1]

$$\begin{aligned} G_n^-(\rho, -h) &= 0 \\ G_n^+(\rho, 0) &= G_n^-(\rho, 0) \\ \frac{\partial G_n^+}{\partial z}\Big|_{z=0} &= \epsilon_r \frac{\partial G_n^-}{\partial z}\Big|_{z=0} \end{aligned}$$

determine the coefficients

$$\begin{aligned} A_1 &= -\Big(\frac{\beta + e^{-2kh}}{1 + \beta e^{-2kh}}\Big) e^{-kz_n}, \\ A_2 &= -e^{-2kh} A_3, \\ A_3 &= \Big(\frac{1-\beta}{1+\beta e^{-2kh}}\Big) e^{-kz_n}. \end{aligned}$$

Using

$$\frac{1}{1+\beta e^{-2kh}} = \sum_{m=0}^{\infty} (-\beta)^m e^{-2mkh}$$

[1] These conditions are not satisfied by the GF proposed for $z > 0$ in Ref. [85].

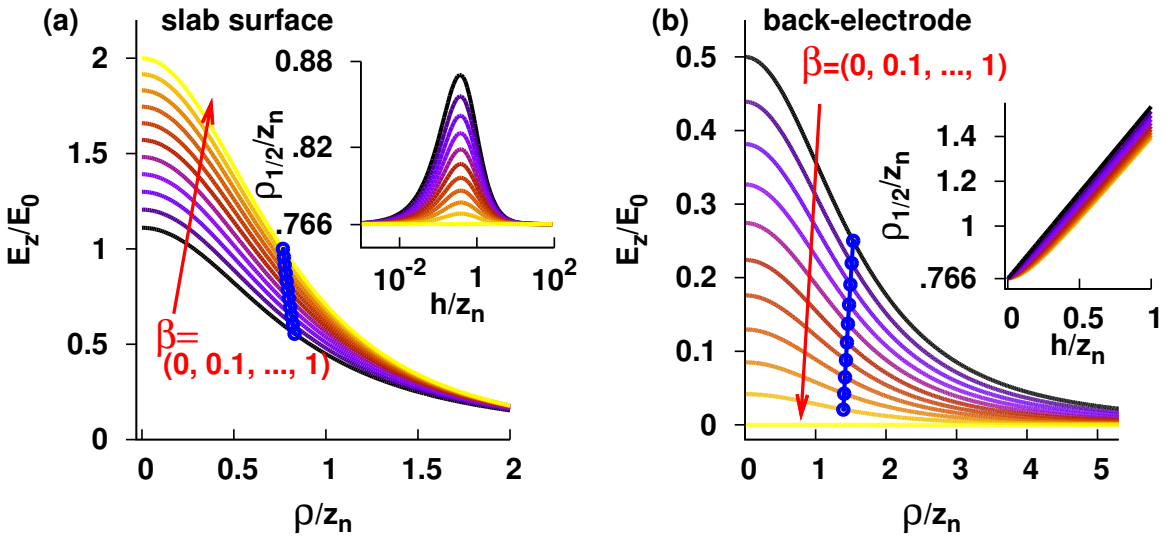

Figure 3.5: (a) E_z (in units of $E_0 = q_n/4\pi\epsilon_0 z_n^2$) vs ρ/z_n just outside the sample for a point charge at z_n above a dielectric slab of thickness $h = z_n$ for several values of β. Blue points indicate the half-width $\rho_{1/2}$ while its variation with h is depicted in the inset; $\rho_{1/2}$ =0.766z_n for $\beta = 1$. (b) Same as (a) but just outside the coated back-electrode; $\rho_{1/2}$ =0.766$(h + z_n)$ for $\beta = 0$.

we obtain expressions for GF for any field point (ρ, z) above the surface

$$G_n^+ = \frac{1}{\sqrt{\rho^2 + (z - z_n)^2}} - \frac{\beta}{\sqrt{\rho^2 + (z + z_n)^2}} - (1 - \beta^2) \sum_{m=0}^{\infty} \frac{(-\beta)^m}{\sqrt{\rho^2 + (z + z_{nm})^2}} \tag{3.20}$$

and inside the slab

$$G_n^- = (1 - \beta) \sum_{m=0}^{\infty} (-\beta)^m \Big(\frac{-1}{\sqrt{\rho^2 + (z + z_{nm})^2}} + \frac{1}{\sqrt{\rho^2 + (z + 2h - z_{nm})^2}} \Big) \tag{3.21}$$

where

$$z_{nm} = z_n + 2(m + 1)h. \tag{3.22}$$

Eqs. (3.20,3.21) reduce to Eq. (3.4) for $h \to \infty$, and the GF of a perfect conductor is obtained for $\beta = 1$. Above a dielectric slab of finite h, however, G_n^+ includes additional terms because q_n induces, in addition to the first image $-\beta q_n$ at $-z_n$, an infinite sequence of alternating image charges $-(1 - \beta^2)(-\beta)^m q_n$ at equidistant positions $-z_{nm}$, as depicted in Fig. 3.1(b). All those image charges sum up to $-q_n$, as required.

Inside the slab, G_n^- simply corresponds to a series of point charges $(1 - \beta)(-\beta)^m q_n$, $m \geq 0$, located at $z_{nm} - 2h$ and their mirror images with respect to the back-electrode plane.

3.2.2 Electric field profiles

In Fig. 3.5, we show E_z profiles induced just outside the slab surface and its interface with the back-electrode by a point charge. With increasing ϵ, i.e. β, the field above the surface is enhanced, whereas E_z above the back-electrode drops because of reduced penetration into the dielectric; concomitantly, the respective half-widths slightly decrease. With increasing h/z_n, $\rho_{1/2}$ increases monotonically at the back-electrode interface and stays only slightly below the linear dependence $\rho_{1/2} = \sqrt{2^{2/3}-1}(h+z_n) \simeq 0.766(h+z_n)$ obtained when $\beta = 0$. On the sample surface, however, $\rho_{1/2} \simeq 0.766 z_n$, apart from a small peak rising to $0.88 z_n$ around $h/z_n = 0.7$ upon decreasing β. This behaviour can be related to the spreading of field lines emanating from q_n at a given angle which end perpendicular to the sample surface if $\beta = 1$, but perpendicular to the back-electrode otherwise and remain between those for β=1 and $\beta = 0$ (vacuum instead of dielectric layer).

For a biased sphere centered at $z_1 = s + R$, Fig. 3.2(a) shows a more pronounced field enhancement caused by the image charges $q_{n>1}$ closer to the surface, and $\rho_{1/2}$ is somewhat smaller, except when $\beta \to 0$, but this limit corresponds to a missing sample.

3.3 Spherical tip atop dielectric slab

In the problem of a biased conducting sphere against a finite dielectric slab, each q_n inside the sphere generates an infinite series of images on the slab side at positions $-z_{nm}$, and each of those induces an image closer to the sphere center to bring it towards an equipotential. The images within the sphere are recursively given by

$$Q(m, q_n, z_n) = \begin{cases} q_n \beta R/(z_1 + z_n), & m = -1 \\ q_n(1-\beta^2)(-\beta)^m R/(z_1 + z_{nm}) & m \geq 0 \end{cases} \tag{3.23}$$

$$Z(m, z_n) = z_1 - \frac{R^2}{z_1 + z_{nm}}. \tag{3.24}$$

When $h \to \infty$ or $\beta = 1$, all images vanish except $Q(-1, q_n, z_n) = q_{n+1}$ and Eqs. (3.1,3.2) are ecovered. Now, to solve the combined sphere-slab problem, one puts the first point charge $q_1 = 4\pi\epsilon_0 RV$ at the sphere center $z_1 = R + s$. An infinite series of images $Q(m, q_1, z_1)$ is then induced inside the sphere, each of which has in turn an infinite images $Q[l, Q(m, q_1, z_1), Z(m, z_1)]$ and so on. In a numerical treatment, the infinite series can be truncated as soon as Q becomes small enough. For relevant parameters, Q is at least 10^{-16} times smaller than q_1 when $m > 10$; for the same reason only a limited number of nested sums must be considered. In Fig. 3.6 we illustrate the convergence of our procedure for a particular example. The normalized capacitance of the system is $C/\pi\epsilon_0 R = 7.22$ compared to 5.86 and 7.46 in the case of a semi-infinite slab and a perfect conductor, respectively, as given by Eq. (3.9).

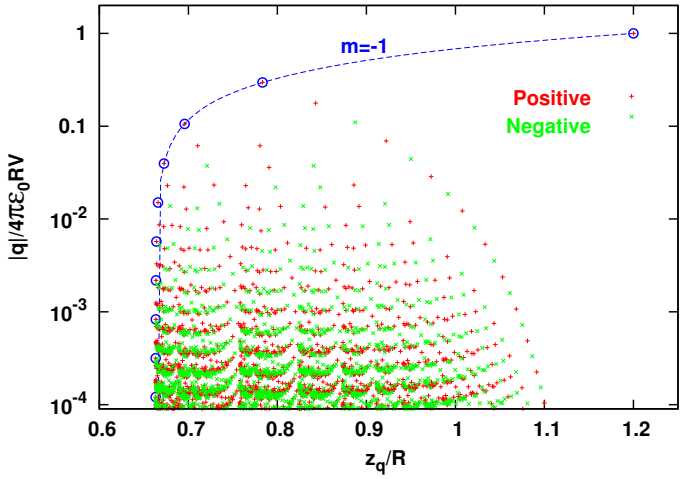

Figure 3.6: The normalized magnitudes of image charges vs. their positions generated within a conducting sphere of radius R at a distance $s = 0.2R$ from an NaCl slab ($\beta = 0.71$) of thickness $0.2R$. The largest point charge of magnitude one is at the sphere center ($z_1 = 1.2R$), while no charges appear below $z_\infty = 0.66R$, z being measured from the slab surface. The first series of image charges given by Eq. (3.23) with $m = -1$ are shown with circles and define an h-independent upper bound. Further charges with magnitudes larger than 10^{-4} are shown in red (positive) and green (negative).

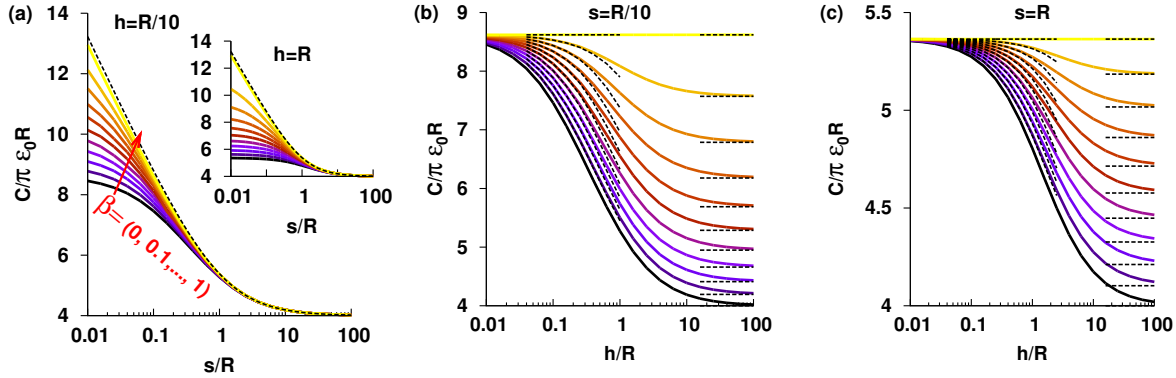

Figure 3.7: Computed normalized capacitance (a) vs s/R for $h = R/10$ and $h = R$ (inset) and vs h/R for $s = R/10$ (b) and $s = R$ (c) of a conducting sphere of radius R separated by s from a dielectric slab of thickness h. In (a) the dashed lines correspond to Eq. (3.17). In (b) and (c) the dashed lines on the left correspond to a sphere at an effective separation $s + h/\epsilon_r$ from the back-electrode (see text) and those on the right to the sphere at separation s from a semi-infinite dielectric.

3.3.1 Capacitance

The capacitance C is obtained from the total charge on the sphere , namely

$$CV = q_1 + \sum_{m=-1}^{\infty} Q(m, q_1, z_1) + \sum_{l,m=-1}^{\infty} Q[l, Q(m, q_1, z_1), Z(m, z_1)] + \cdots . \quad (3.25)$$

A systematic summation of the resulting nested series with a prescribed precision is possible as described above. The resulting dependence of the capacitance of the sphere-slab system as a function of s/R and h/R is depicted in Fig. 3.7. First of all, the dashed lines in Fig. 3.7(a) show that for $\beta = 1$, Eq. (3.17) agrees within 2% over the whole examined range of s/R. With increasing s/R, C first becomes almost independent of β at a value which grows with h/R, then approaches the capacitance C_{sph} of the isolated sphere as $2\pi\epsilon_0 R^2/s$. Similarly, at large h/R, C approaches values given by Eq. (3.9) shown by dashed lines on the right side in Fig. 3.7(b). The slow approach reflects the influence of the additional image charges. Note that $\beta = 1$ and 0 correspond to a biased sphere at respective separations s and $s + h$ from a perfect conductor.

A smooth interpolation between those two limits is obtained by *replacing s with $s + h/\epsilon_r$, where $\epsilon_r = \epsilon/\epsilon_0$, while keeping $\beta = 1$ in Eq. (3.9).* As shown by dashed lines on the left side in Fig. 3.7(b), the resulting approximation is within 1% of the exact $C(s/R, h/R)$ for $h \leq R/3$ if $s = R/10$. For larger separations $s \sim R$ the deviation remains within 1% as long as $h \leq R$, see Fig. 3.7(c). In view of its remarkable agreement with Eq. (3.9) for $\beta = 1$, *Eq. (3.17) together with the same substitution provides an almost as good but simple approximation to the exact C for small enough h/R.* The same combination $s + h/\epsilon_r$ appears in the denominator of C in the case of a parallel-plate capacitor of thickness $s + h$ partially filled with a dielectric slab of thickness h. However, the proposed approximation remains valid when the field profiles at the top and bottom surfaces of the dielectric slab are far from uniform, e.g. if $s \ll h < R$. This is evidenced by the $E_z(\rho)$ profiles and by their respective half-widths as discussed in the following.

3.3.2 Electric field profiles

The electric field profiles for a conducting sphere atop a dielectric slab are shown in Figs. 3.8 and 3.9. To compute $E_z(\rho, z) = -\sum q_k \partial_z G_k^{\pm}/4\pi\epsilon_0$, where the GF is given by Eqs. (3.20-3.21), we used the same point charges q_k which are kept in evaluating Eq. (3.25) as shown in Fig. 3.7. With increasing β, the field just outside the surface is enhanced, whereas E_z just above the back-electrode drops because of reduced penetration into the dielectric. Concomitantly, the respective half-widths slightly decrease.

At the sample surface (insets of Figs. 3.8(a) and 3.9(a)), $\rho_{1/2}$ decreases towards values matching those in Fig. 3.2(b) at the same s if $h \to \infty$, but towards a common value which agrees with that in Fig. 3.2(b) for $\beta = 1$ if $h \to 0$. A weak maximum consequently appears,

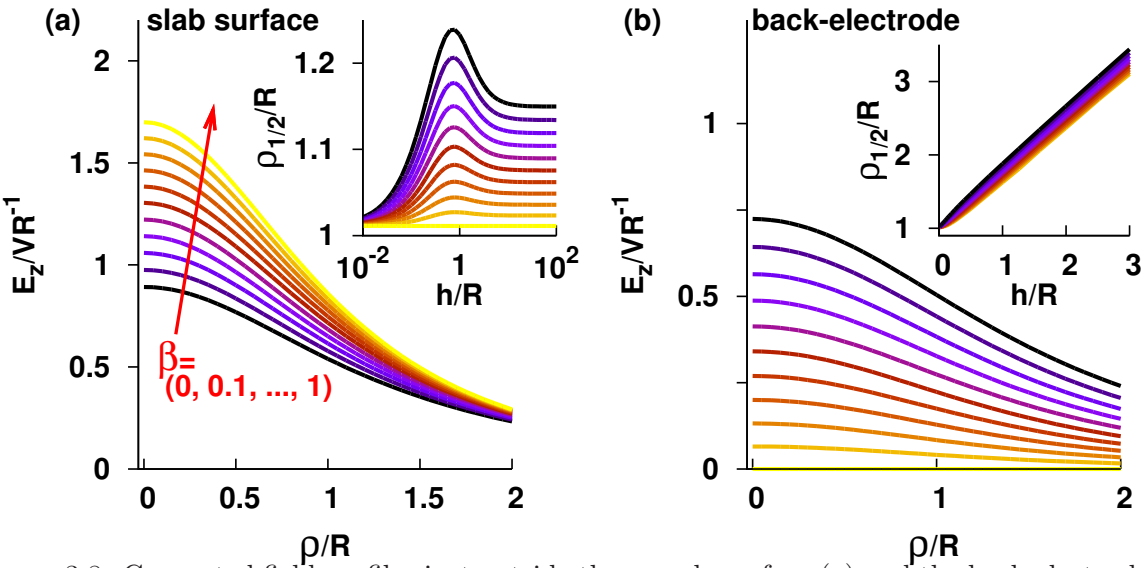

Figure 3.8: Computed field profiles just outside the sample surface (a) and the back-electrode (b) for a conducting sphere of radius R at a separation $s = R/2$ from a dielectric slab of thickness $h = R/2$. The corresponding half-widths $\rho_{1/2}$ as functions of h/R are shown in the insets.

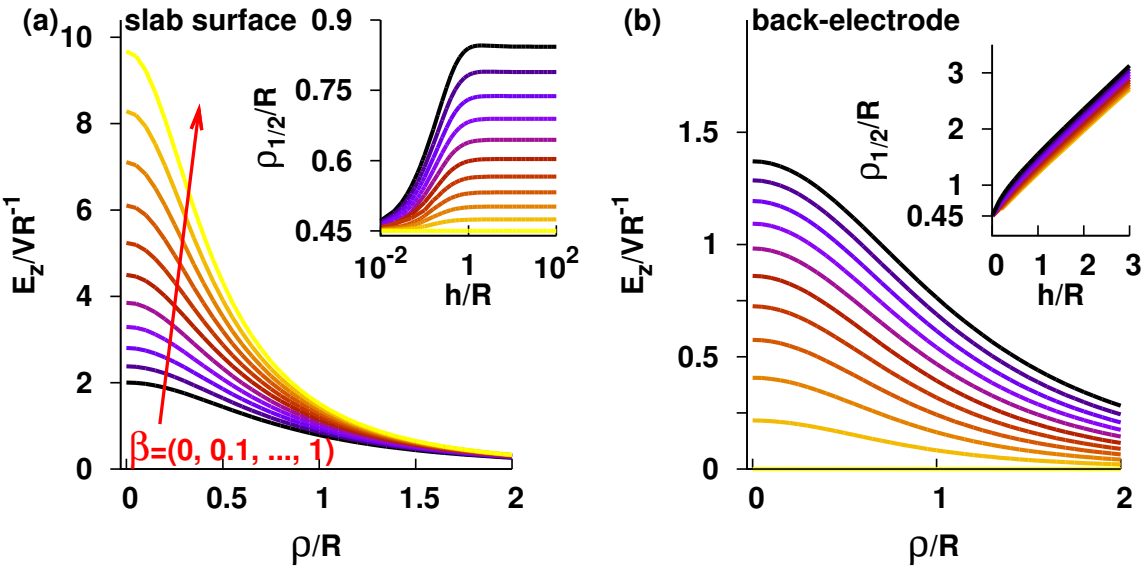

Figure 3.9: Same as Fig. 3.8, but for $s = R/10$ (cf. Fig. 3.2(a)).

around $h \sim R$ which, however, is absent if $s = 0.1R$. If β is close to 1 (undoped or depleted semiconductor capping layer), $\rho_{1/2}$ remains close to the $\beta = 1$ value for all h. Thus, like in Fig. 3.2(b), the half-widths at the surface are significantly below R, while their spread increases, if s/R is small. At the back-electrode interface (insets of Figs. 3.8(b) and 3.9(b)), $\rho_{1/2}$ increases monotonically with increasing h/R and stays below the line $\rho_{1/2} = 0.766(s + R + h)$ obtained when $\beta = 0$ for a lumped charge at the sphere center, as illustrated in the inset of Fig. 3.5(b), albeit at the separation $s + h$ from the back-electrode. Therefore, this line is approached only if $s + h$ becomes comparable to R. In the opposite limit, $\rho_{1/2} \simeq \sqrt{2(s + h)R}$, in accordance with Fig. 3.2(b) When $h \to 0$ (missing dielectric layer) the half-widths on the back-electrode and on the surface coincide with the $\beta = 1$ values in Fig. 3.2(b), namely $\rho_{1/2}/R = 1.02$ if $s = R/2$ and 0.45 if $s = R/10$. Nevertheless $\rho_{1/2}$ at the back-electrode considerably exceeds the half-width at the surface in the common experimental situation when the closest approach distance $s \ll h \leq R$.

II. Numerical Approach

In the following we first discuss previous numerical approaches, then present our own computationally simple, yet flexible finite-difference (FD) scheme with controlled accuracy to treat electrostatic tip-sample interactions on macro- and mesoscopic scales. Calculating the cantilever-tip-sample electrostatic interaction is, in fact, an intricate electrostatic boundary-value problem. One difficulty is due to the distance-dependent redistribution of the surface charge density on an electrode at constant bias voltage.

For the model macroscopic system of a semi-infinite sample and spherical tip, we obtained in Sec. 3 exact expressions for the interaction energy and forces by the analytic method of images. For more realistic geometries including tip and cantilever, even without extrinsic charges, the problem is still nontrivial. The main difficulty is due to the presence of several length scales determined by the nontrivial shape of the AFM probe. For a conducting tip represented as a cone with a spherical end cap above a conducting plane, a simple assumption (constant electric field along each field line approximated by a circular arc normal to the surfaces) led to an analytic expression for the force on the tip. [25] Recent numerical calculations [92, 76] showed that Hudlet's expression is surprisingly accurate. Somewhat different analytical expressions and estimates for the lateral resolution in AM- and FM-KPFM were obtained for similar probes, also including a tilted cantilever. [62]

Evidently, the difficulty mentioned above is present also for numerical methods which directly deal with Poisson's equation. We first overview a few existing numerical approaches and then present our new method.

3.4 Previous approaches

3.4.1 Numerical Images Charges Method

For cylindrical geometries, many authors proposed numerical schemes based on the image charge method which is applicable to simple geometries involving spherical and planar surfaces. [87] Thus Belaidi *et al* [93] placed N point charges on the symmetry axis and determined their positions and strengths by forcing the potential on the tip surface to be V by a nonlinear least squares fit. The previously mentioned authors also described how contributions of the spherical cap, the tip shank and the cantilever to the macroscopic force lead to characteristic distance dependencies on scales determined by the geometry and dimensions of those parts. A *linearized* version of the *numerical image charge method* where the positions of axial point and line charges were fixed was applied to study tip-shape effects for conductive and *dielectric samples* [63, 90] *and thin films on conducting substrates* [85], also including the influence of the cantilever[94]. It is not known to what extent the boundary conditions must be satisfied for a given accuracy in the numerical image method, unlike in the analytic method where the positions and strengths of the image charges change with tip-sample separation and the boundary conditions are fully satisfied.

3.4.2 Poisson Solvers

A more systematic approach to multi-length-scale problems is the *boundary element method* (BEM) [64, 65, 92]. In this method the 3D (2D) differential Poisson's equation is transformed into 2D (1D) integral (Green's functions) equations on the surfaces of conductive or dielectric components, including CPD discontinuities and surface charges if desired. [66] The accuracy of BEM is controlled by the mesh resolution and is applicable to complex probe-sample systems, e.g. including a realistic cantilever [95]. The size of the resulting linear system of equations is small compared to volumetric discretization methods. However, because of the memory requirement of $\mathcal{O}(N^2)$ to store the fully populated matrix and complexity of $\mathcal{O}(N^3)$ to solve the linear equations, BEM has mostly been applied to systems with a relatively small number N of grid points, e.g. problems of high symmetry and homogeneity for which it is feasible to derive the Green's function analytically. Somewhat earlier a few authors adapted Green's function methods developed for more complex near-field optics problems to investigate lateral resolution in KPFM on inhomogeneous samples [60, 63]. One advantage of BEM is that the LCPD of such samples can be expressed as a 2D convolution of the CPD and/or of a fixed surface charge distribution with a point-spread function which depends only on the relative position of the scanning probe. [64, 92, 95] The distance-dependent lateral resolution can be quantified by the width of that function. Moreover, if one assumes that only one of those distribution is present, its can be determined by inversion of the BEM matrix upon discretization on the adjustable BEM mesh. [66]

Conceptually more straightforward approaches involving surface elements have been applied to conductive probe and sample systems. In the simplest one, the tip surface is approximated as a regular staircase (or, equivalently, as an *array of capacitors in parallel*), [60, 96, 57]. More accurate methods rely on adjustable meshes. Thus the *finite element method* (FEM) was used to calculate the electrostatic force acting on a conical tip, [61] while a commercial FEM software was recently applied to simulate a realistic cantilever and tip of actual shape and dimensions over a *conducting* flat sample with a CPD discontinuity. [97] More sophisticated software packages have been used to solve the Poisson's equation in the presence of space charges, e.g. for structured samples involving doped semiconductors [59, 98]. Numerical methods which involve 3D discretization require a very large number of grid points even if the mesh is carefully adjusted; the computational box must therefore be truncated at some finite extent.

3.5 Finite-difference method

As an alternative we present a finite-difference method (FDM) on a 3D non-uniform grid which is capable of dealing with realistic sizes of the cantilever, tip and sample. Inhomogeneous metallic and dielectric samples as well as thin dielectric films on metal substrates, can be straightforwardly treated with this method. The most attractive feature of our FDM compared to FEM or BEM computations is its ease of implementation. Since the electrostatic potential varies smoothly and slowly at distances far from the tip apex, we use a grid spacing which increases exponentially away from this region. Consequently, the number of grid points depends logarithmically on the truncation lengths, and an extension of the computational box costs relatively few additional grid points. It allows us to simulate the cantilever as well as thick dielectric samples according to their actual sizes in experiments.

The capacitance $C(s)$ between the probe and the sample back-electrode depends only on the tip-sample separation s, provided that their geometries are fixed. [89] The macroscopic electrostatic energy due to the effective voltage difference $V = V_b - V_{CPD}$ between the conducting tip and back-electrode is given by $U_c(s, V) = \frac{1}{2}C(s)V^2$. The electrostatic force exerted on the tip is proportional to the capacitance-gradient $C'(s) = \partial C/\partial s < 0$

$$F_M(s, V) = -\frac{\partial U}{\partial s} = +\left(\frac{\partial U_c}{\partial s}\right)_V = +\frac{1}{2}C'(s)V^2 < 0. \tag{3.26}$$

Similarly, the force-gradient is proportional to $C''(s) = \partial^2 C/\partial s^2$. We emphasize the difference between the total electrostatic energy U of the macroscopic system and the capacitive energy U_c which leads to the positive sign on the RHS of Eq. (3.26). Because this a very common mistake by many authors, the reason is restated here.

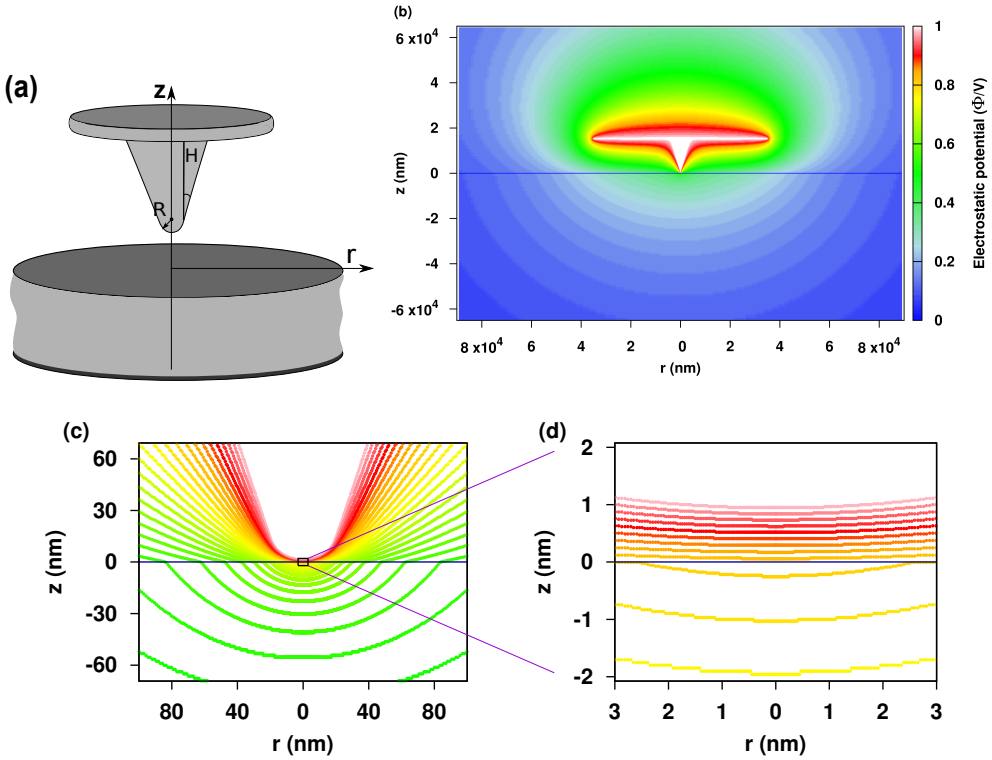

Figure 3.10: (a) Schematic of the macroscopic model AFM probe-sample system with cylindrical symmetry: a 15 nm high cone with 15^o half-angle terminated by a spherical cap of radius $R = 20$ nm is attached to a disk of thickness 0.5 μm. The radius of the disk is 35 μm which matches the area of a typical cantilever. The sample is a 1-mm thick dielectric slab with the relative permittivity $\epsilon/\epsilon_0 = 5.9$ of NaCl. An effective bias of $V = 1$ Volt is applied to the conducting probe while the back electrode at the bottom and the surrounding enclosure of height and radius $10^6 R = 20$ mm (not shown) are grounded. (b) 2D (r, z) map of the macroscopic electrostatic potential Φ normalized to V for the model system in (a). The white region corresponding to $\Phi = 1$ reflects the probe geometry; successive contours differ by 0.01. The sample-vacuum interface is indicated by the horizontal line at $z = 0$ and the tip-sample separation is 1 nm. (c), (d) Zooms into the apex-surface proximity region in (b) with $\times 10^3$ and 3×10^4 magnifications, respectively. The staircase shape of the contours reveals the resolution of the nonuniform mesh at different locations.

3.5.1 Sign of the macroscopic electrostatic force

Using the virtual work method, the macroscopic electrostatic tip-sample interaction can be calculated from the potential energy stored in the capacitor formed between the tip and the back-electrode. The (real) force acting on the tip F_s, which is considered constant during a virtual arbitrary infinitesimal tip displacement δs, performs a virtual work $\delta w = F_s \cdot \delta s = -\delta U$, where $U = U_c + U_b$ is the total energy of the system including contributions from both the capacitor and

the biasing battery which maintains a fixed potential difference V between the both electrodes. In response to this displacement, the battery transfers a charge δQ between the electrodes in order to keep their potential difference fixed. It costs a change of $\delta U_b = -\delta Q \cdot V$ in the energy of the battery. Whereas the energy of the capacitor changes by $\delta U_c = \frac{1}{2}\delta Q \cdot V$, which implies $U_b = -2U_c$, i.e.

$$\delta U = \delta U_c + \delta U_b = -\delta U_c.$$

The electrostatic force is therefore

$$F_s = -\frac{\delta U}{\delta s} = +\frac{\delta U_c}{\delta s} = +\frac{1}{2}\frac{\delta C}{\delta s}V^2$$

and is always attractive because $\delta C/\delta s < 0$.

3.5.2 Discretization

The electrostatic energy

$$U_c(s, V) = \frac{1}{2}\int \epsilon(\mathbf{r})|\nabla\Phi|^2 \, d\mathbf{r}$$

can be determined once the electrostatic potential $\Phi(\mathbf{r}; s, V)$ is known at any point $\mathbf{r}$ in space. In general, when the dielectric constant $\epsilon(\mathbf{r})$ varies in space, Φ satisfies the generalized form of Poisson's equation

$$\nabla \cdot [\epsilon(\mathbf{r})\nabla\Phi(\mathbf{r})] = -\rho(\mathbf{r}), \tag{3.27}$$

ρ being the charge density. Minimization of the energy-like functional

$$I\left[\Psi(\mathbf{r})\right] = \frac{1}{2}\int \epsilon(\mathbf{r})\left|\nabla\Psi\right|^2 d\mathbf{r} - \int \rho\Psi d\mathbf{r}. \tag{3.28}$$

subject to Dirichlet boundary conditions leads to Φ, the solution of the Poisson's equation Eq.(3.27) with the same boundary conditions. [87] Using a discretized variational approach, we therefore minimize the functional

$$I\left(\{\Psi_{\mathbf{n}}\}\right) = \sum_{\mathbf{n}}\left(\frac{1}{2}\epsilon_{\mathbf{n}}\left|\nabla\Psi\right|^2_{\mathbf{n}} - \rho_{\mathbf{n}}\Psi_{\mathbf{n}}\right)v_{\mathbf{n}}. \tag{3.29}$$

3.5.3 Implementation

On a non-uniform grid, we delimit the volume $v_{\mathbf{n}}$ of the volume element assigned to node $\mathbf{n}$ by neighboring nodes. Then, $\Psi_{\mathbf{n}}$, $\rho_{\mathbf{n}}$, $\epsilon_{\mathbf{n}}$ and the electric field $-\nabla\Psi_{\mathbf{n}}$ are *evaluated at the center of the volume element* by linear interpolation between the nodes adjacent to $\mathbf{n}$ in orthogonal directions. This ensures that the field is effectively evaluated to second order in the product of

grid spacings and that discontinuities in $\nabla\Psi_{\mathbf{n}}$ and $\epsilon_{\mathbf{n}}$ at material interfaces are correctly treated. Although the formalism is general and can be applied to any 3D system on a judiciously chosen nonuniform 3D orthogonal grid, in the following examples we *consider a cylindrically symmetric setup without free charges* in order to allow comparison with most previous computations. In cylindrical coordinates, each volume element is a truncated tube of height $h_k^{(z)}$ with inner and outer radii r_i, r_{i+1}, respectively, and $v_{\mathbf{n}} = \pi(r_{i+1}+r_i)h_i^{(r)}h_k^{(z)}$, $h_i^{(r)} = r_{i+1}-r_i$ and $h_k^{(z)} = z_{k+1}-z_k$ being respectively the radial and vertical spacings of the appropriate nonuniform grid. The radial and vertical components of $\nabla\Psi$ are approximated on the circle of radius $r_i + 0.5h_i^{(r)}$ at $z_k + 0.5h_k^{(z)}$ as $(\Psi_{i+1,k} - \Psi_{i,k})/h_i^{(r)}$ and $(\Psi_{i,k+1} - \Psi_{i,k})/h_k^{(z)}$. Since the FD approximation of the electric field is a linear combination of the potential values on nearest neighbor nodes, the functional in Eq. (3.29) is quadratic and the minimization condition $\partial I/\partial\Psi_{\mathbf{n}} = 0$ yields a system of *linear* equations $A\Phi = \mathbf{b}$ where the vector $\mathbf{b}$ describes imposed boundary values and charge distributions. Because $A_{\mathbf{mn}} = \partial^2 I/\partial\Psi_{\mathbf{m}}\partial\Psi_{\mathbf{n}}$ is a sparse, symmetric and band matrix, the system can be solved efficiently by an iterative procedure, which may, however, suffer from conditioning problems due to the nonuniformity of the grid. For an accurate solution, a mesh with high enough resolution is required in regions where $\Phi(\mathbf{r}; s, V)$ varies strongly. We used the PARDISO package [99, 100] to solve the resulting huge system of equations. An implementation of our FDM is distributed under GNU-GPL license as the CapSol code [101].

Once $\Phi(\mathbf{r}, s, V{=}1)$ is determined for several separations s, the system capacitance is obtained as $C(s) = \int \epsilon(\mathbf{r})|\nabla\Phi|^2\, d\mathbf{r} \simeq \sum_{\mathbf{n}} \epsilon_{\mathbf{n}} |\nabla\Phi|^2_{\mathbf{n}}\, v_{\mathbf{n}}$. Then a simple second order FD approximation is used to evaluate $C'(s)$ and $C''(s)$ from $C(s)$. The electrostatic force acting on an arbitrary area S of a conducting part can also be evaluated as

$$\mathbf{F}_S = \frac{1}{2\epsilon_0}\int_S \sigma(s)^2 \hat{n} dS, \tag{3.30}$$

where $\sigma(s) = -\epsilon\partial\Phi/\partial n$ is the surface charge density guaranteeing that the tip surface is an equipotential, and $\hat{n}$ is the unit vector normal to the surface element dS. For a system with cylindrical symmetry the net force on a part of the probe delimited by two cylinders of radii $r_1 < r_2$ is vertical and given by $F = \pi\epsilon_0 \int_{r_1}^{r_2} |\nabla\Phi|^2 r dr$, however we prefer to use Eq. (3.26) to calculate the total macrosocopic force on the probe. In the following subsections we validate the performance of our FDM by comparisons with previous results obtained by other methods for cylindrically symmetric systems. We mainly consider the macroscopic model system which is shown schematically in Fig. 3.10(a) and is described in the caption. The conducting probe consists of a conical tip terminated by a spherical cap of radius R attached to a cantilever modelled as a disk of the same area as a typical cantilever,[60] and the sample by a thick dielectric slab. Dirichlet boundary conditions are applied on a very large cylindrical box. Figures 3.10(b-d) show a typical computed 2D (r, z) map of the electrostatic potential normalized to effective potentital V at three

magnifications. Note that the grid spacing changes by six orders of magnitude (hundredths of nm around the tip apex to tens of μm near the box walls). The indented contours reveal the resolution of the grid at different locations, e.g. $R/400$=0.05 nm in the gap between tip and sample in this case. Figure 3.10(d) clearly shows that for a separation of 1 nm a large fraction of the voltage drop occurs within the thick dielectric sample. Whereas the contour spacing between the tip apex and the surface is constant to a good approximation, it gradually increases inside the dielectric, in contrast to what occurs in a parallel plate capacitor. Actually the capacitance remains finite for an infinitely thick sample even in the (macroscopic) contact limit $s \to 0$.

3.5.4 Convergence and Accuracy

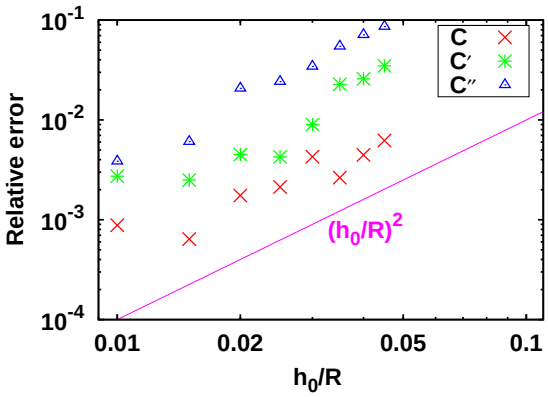

Figure 3.11: Convergence analysis with respect to the finest grid spacing h_0 for a conducting sphere of radius R in front of a thick dielectric of relative permittivity $\epsilon/\epsilon_0 = 5.9$. Points computed by our FDM for the macroscopic capacitance C, the force $\propto C'$ and force gradient $\propto C''$ are compared to the analytic solution for a semi-infinite dielectric The sphere-surface separation is $s = R/20$ and the computational box extends to $10^6 R$ in the radial and vertical directions. The straight line in the log-log plot indicates the expected quadratic error scaling (see text).

Grid spacing

We first test our implementation for the problem of a conducting sphere of radius R separated by s from a semi-infinite dielectric surface for which an analytic solution of controllable accuracy is available. Such a convergence analysis also yields the parameters needed to achieve a desired accuracy. Compared to the analytic solution of the sphere-dielectric system, the convergence with respect to the finest grid spacing h_0 shows a nearly quadratic error scaling (Fig. 3.11) as is expected for a second order FDM. In order to consistently preserve the shape of the tip

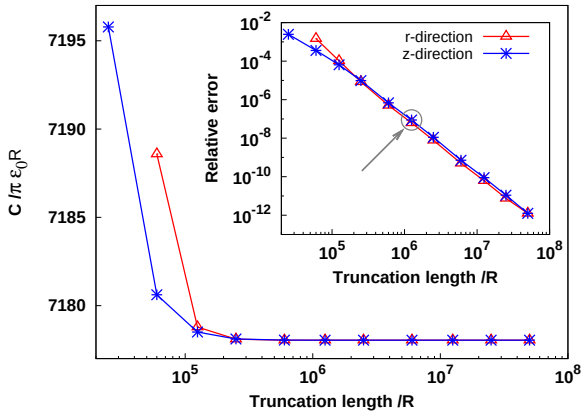

Figure 3.12: Convergence analysis with respect to the radial and vertical extents of the FDM computational box for the macroscopic system described in the caption of Fig. 3.10, the tip-sample separation and finest mesh size being $s = R/20$ and $h_0 = R/100$, respectively. The normalized capacitance of the system approaches the same asymptotic value upon increasing the truncation length in one direction while the other one is sufficiently large and fixed. Relative deviations with respect to the asymptotic value are shown in the inset. The arrow indicates the truncation length adopted in subsequent FDM computations.

approximated by the orthogonal mesh, the tip-sample separation must be changed in steps of h_0. Then the errors of the second order FD approximations of C' and C'' are quadratic versus h_0, even if C is known exactly. Once these errors are added to those of C in the Poisson's solver, the overall errors in C' and C'' are larger than the error in C, although they remain quadratic versus h_0, as seen in Fig. 3.11. The accuracy could be improved by using higher order FD approximations for the electric field by using further neighboring points. However, a corresponding improvement of the approximation of curved surfaces on the orthogonal FD-mesh is then also required. Note that, for consistency, the surface of the sphere must be approximated as a staircase with variable step heights and widths which also change when the grid-spacing is changed. At larger separations, the error scaling deviates from quadratic towards linear behaviour. Then the absolute value of the error is small and a larger grid-spacing can be used. The capacitance, force and force-gradient of our test system at a rather small separation of $s = R/20$ can be calculated within a relative error of 0.005 compared to the analytic solution if $h_0 = R/100$. For the cantilever-tip-sample system [Fig. 3.10(a)] a uniform grid with $h^{(r)} = h^{(z)} = h_0 = R/100$ is used around the tip apex up to a distance of twice the tip apex radius in both radial and vertical directions. Outside this range the grid becomes gradually coarser with a growth factor of 1.01.

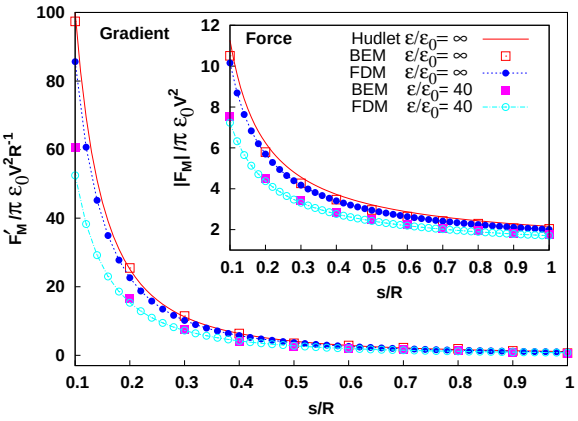

Figure 3.13: Normalized macroscopic electrostatic force (inset) and force-gradient computed by our FDM versus the normalized tip separation s/R from a dielectric ($\epsilon/\epsilon_0 = 40.0$) and a conducting ($\epsilon/\epsilon_0 = \infty$) sample compared to BEM computations (Ref.[92]), as well as to Hudlet's approximation (Ref. [25]) in the second case (see text). The cantilever is absent, as assumed in those two treatments, but the remaining parameters are as described in the caption of Fig. 3.10(a).

Space truncation

A convergence analysis with respect to the size of the computational cylinder is shown in Fig. 3.12 for the model system described in Fig. 3.10. The capacitance approaches the same asymptotic value when the truncation length in a particular direction is increased while the other one is kept fixed and sufficiently large. If the computational box extends to $10^6 R$ in the radial and vertical directions, the relative deviation of the capacitance from its asymptotic value is only 10^{-7} (as indicated by the arrow in Fig. 3.12). We use these cutoff parameters in all subsequent FDM computations reported here.

Comparison

In Fig. 3.13 we compare results obtained by our FDM with previous accurate BEM computations [92] for a system like in Fig. 3.10(a) but *without the cantilever* for a conducting and a dielectric ($\epsilon/\epsilon_0 = 40$) sample. The force and the force-gradient evaluated by the two methods are in very good agreement for both kinds of samples. For the conducting sample, Hudlet's analytic approximation [25] deviates by only a few percent from the numerical results. In the following we show that the contribution of the cantilever can be quite appreciable for a dielectric sample.

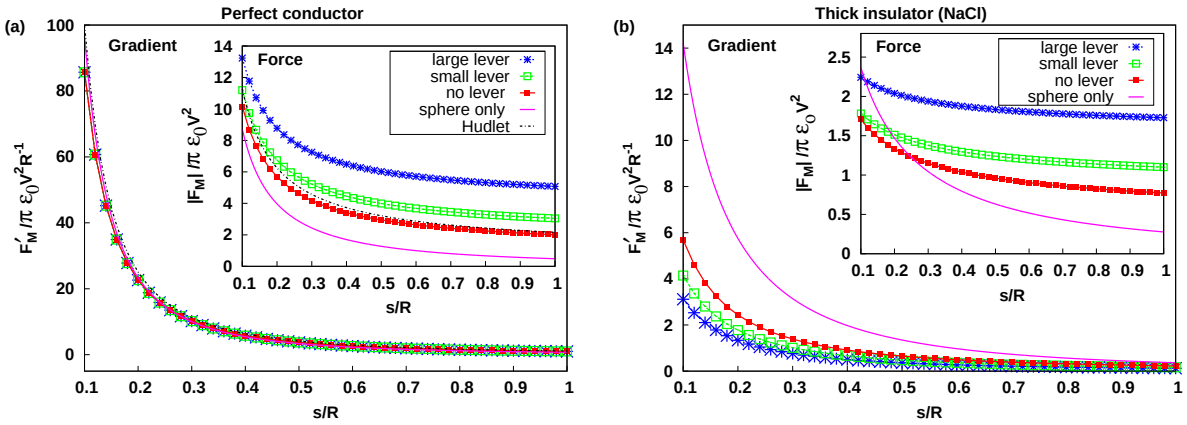

Figure 3.14: Effect of the cantilever (size) on the macroscopic electrostatic force (inset) and force-gradient at different normalized tip separations from a conducting (a) and dielectric (b) sample. The cantilever is modelled as either a small or a large disk with radii of 20 and 35 μm, respectively. Other parameters are as in caption of Fig. 3.10(a). The solid lines show corresponding results for a tip approximated by a conducting sphere with radius $R = 20$ nm obtained by summing the analytic series for semi-infinite samples of both kinds.

3.6 Results

3.6.1 Force and force-gradient

The macroscopic electrostatic force and force-gradient versus the normalized tip-surface separation s/R for the system in Fig. 3.10 are shown in Fig. 3.14 for three different geometries: without, with a small and a large cantilever modelled as disks of thickness 0.5 μm. The small disk radius is equal to the width of a typical rectangular AFM cantilever (20μm) while the total area of the large disk (of radius 35 μm) matches the area of the rectangular cantilever. The presence of the cantilever increases the capacitance and the electrostatic force. Because the cantilever is more than 10 μm away from the surface, its contribution to the force is often considered constant for tip-sample separations smaller than R, and therefore does not contribute to the force gradient. Our calculations [Fig. 3.14(a)] confirm that this is in fact true for a conductive sample. In this case, the main contribution to the force-gradient comes from the spherical cap, as can be seen from the solid line which corresponds to the analytic solution for a conducting spherical tip. However, the conical shank of the tip and the cantilever affect the force at large separations, as shown in the inset and noticed earlier. [59, 93, 25, 60] On the other hand, if s/R is small, as shown in Fig. 3.14(b) and also emphasized in previous work [90, 94, 85], *over a thick dielectric sample both the force and the force-gradient are significantly decreased, owing to field penetration into the sample.*

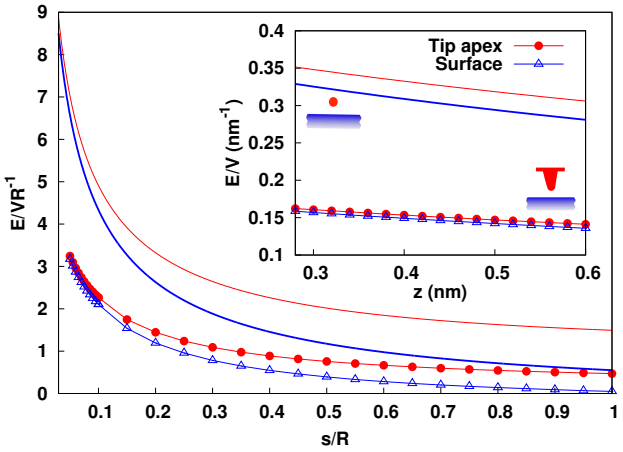

Figure 3.15: Normalized macroscopic electric field in the vacuum on the tip surface and on the dielectric sample surface (ϵ/ϵ_0=5.9) versus their normalized separation for the probe described in the caption of Fig. 3.10(a) (curves with symbols) and for a tip approximated by a conducting sphere of the same radius (continuous curves). Inset: zoom into the range $z = s - h$ where atomic-scale contrast appears for $R = 20$ nm, $h = 0.72$ nm; the electric field between the tip and the surface changes by only a few percent and is hence nearly uniform.

3.6.2 Electric field

A quantity of particular relevance in our multi-scale approach is the macroscopic electric field in the vacuum gap between the spherical tip end and the sample surface which polarizes the microscopic system. The variation of the electric field normalized to V/R at two points on the symmetry axis in the vacuum gap just below the tip and just above the surface versus their normalized separation is shown in Fig. 3.15. The same quantities are shown magnified in the inset for nanotip separations relevant for atomic-scale contrast, i.e. $z = s - h <\sim 0.6$ nm, differ little and drop only weakly with increasing z. In the same distance range the z-component of the electric field is two orders of magnitude stronger than the radial component parallel to the surface. These features are also clearly illustrated by the essentially equispaced horizontal equipotential contour lines in the vacuum region shown in Fig. 3.10(d). This important observation greatly simplifies the desired coupling to atomistic calculations: we *can consider the electric field E_z at the midpoint of the macroscopic tip-surface distance $s = z + h$ as a uniform external field acting on the isolated microscopic tip-sample system.* The connection between those two scales is schematically illustrated in Fig. 4.1.

Figure 3.14 shows that for a conducting sample the force gradient can be accurately described by a spherical tip if $s < R$, although the force itself is increasingly underestimated at larger separations [25, 38]. In contrast, *for a thick dielectric sample*, the same description only

provides the order of magnitude of F_M at small s/R, but exhibits a faster decrease with increasing separation and overestimates F'_M. Figure 3.15 reveals that a spherical model tip overestimates the electric field E_z under the tip at all separations, which then approaches V/R on the sphere (and zero on the surface) when $s \gg R$. This occurs because the induced surface charges can spread to the conical shank and the cantilever in the more realistic model. The contributions of those parts to the force F_M become nevertheless stronger than that of the sphere alone already at small s/R. In general, if $s/R \to 0$, the *electric field under the tip, hence the force and the force gradient are enhanced* owing to an increasingly *localized surface polarization* of both tip and sample, but remain finite if the sample is a dielectric, as explicitly demonstrated by the solution for a spherical tip. Comparison with that solution (the solid curves in Fig. 3.14) shows that *even at small separations* contributions from *both the conical shank and the cantilever contribute to the force*, whereas *mainly the conical shank affects the force gradient*. Hence, ignoring those contributions causes an overestimation of the force-gradient if the sample is an insulator.

In the original analytic work of Bocquet, Nony and collaborators the nonuniform macroscopic surface polarization of the dielectric sample and also the resulting electric field and force enhancements at small s/R were not taken into account. More importantly, only the polarization of the back-electrode was considered, leading to a capacitive force on the assumed spherical tip [37] and also on the cantilever [70] underestimated by orders of magnitude (as discussed in the first part of this chapter). In a subsequent publication which assumed a similar setup [71], the macroscopic surface polarization was presumably correctly included, although details were not provided. Fig. 3.15(b) shows that contrary to conducting surface, where the electric field (and force-gradient) can be precisely described by a single sphere, for case of a thick dielectric slab it can give only the order of magnitude of the electric field. Like force-gradient, the sphere-slab model overestimation of the electric field by a factor of two in this especial case. It worths to tone that replacing the sphere-slab system with a sphere within an effective dielectric $\epsilon_0 \leq \tilde{\epsilon} \leq \epsilon$ [37] which gives the electric field on the sphere surface as $E/VR^{-1} = 2\tilde{\epsilon}/(\tilde{\epsilon} + \epsilon_0) < 2$ on the sample surface as $E(s)/VR^{-1} = 2\tilde{\epsilon}/(\tilde{\epsilon} + \epsilon_0)(1 + s/R)^2 < 2$ underestimates the electric field strength (cf Fig.3.15). When using a sphere model, one should notice this over- and underestimation of the field intensity which influences the strength of the polarization effects in the atomic-scale. It plays the most important role in interpreting the LCPD as will be discussed in more details in the next Section. (Also, note that a plane-capacitor approximationpredicts a very small $E = V/(s + h/\epsilon)$ for a dielectric slab with large thickness h.)

Conclusions

In summary, the electrostatic interactions of a conducting model with different types of sample surfaces was addressed from two approaches, namely analytically and numerically. First, using the method of image charges for a biased model spherical tip facing a semi-infinite dielectric, we found a simple generalization of the solution for the sphere-planar conductor problem. Approximate, but accurate compact formulas were obtained for the capacitance and related quantities of current interest in scanning gate or scanning capacitance experiments on doped semiconductors or 2DEGs capped by insulating layers, besides electrostatic force microscopy of insulating thin films on metal substrates. Note that the tip is typically oscillated at or close to a resonance frequency of the force sensor used to control the closest approach distance; probed quantities must therefore be averaged over the tip trajectory. [16, 86] Green's functions for field points above and inside a dielectric slab with finite thickness grounded at the bottom were used to setup a systematic numerical solution. For experimentally relevant situations, where the tip apex radius exceeds the tip-sample separation s and the slab thickness h, our numerical results are within 1% of the total capacitance for the sphere-planar conductor problem at an effective separation $s + h/\epsilon_r$. The computed field profile widths at the slab surface and at the back-electrode interface indicate that the common assumption of a tip-surface capacitor in series with a sample capacitor of effective radius comparable to the tip apex radius R is seldom justified. We recommend instead to use the above-mentioned approximation. A worthwhile next step would be to include screening by non-ideal conductors, e.g. buried 2DEGs or surface layers, e.g. graphene or metallic surface states.

In the second approach, we addressed the same problem from a numerical approach. The electrostatic problem of the macroscopic bodies of the voltage-biased AFM probe (including the tip and the cantilever) against the grounded sample, treated as macroscopic perfect conductors or insulators, is solved by a finite-difference method with controlled accuracy on a non-uniform mesh by minimizing an energy-like functional which leads to solution of the Poisson's equation. The method is capable of treating complex geometries with widely different dimensions, but is illustrated here for systems with cylindrical symmetry. The solution yields the electric potential and field distributions and the capacitance $C(s)$ of the system from which the electrostatic force F_M acting on the probe and its gradient are calculated as functions of the macroscopic tip-sample separation s. By comparing results obtained with and without the cantilever, as well as with the analytic solution for a tip approximated by a conducting sphere, the *contributions of the cantilever, the conical tip shank and of its spherical end can be recognized. If the sample is a thick insulator, all three affect the macroscopic force, whereas the last two affect the force gradient even at sub-nanometer separations relevant for atomic-scale contrast.*

Chapter 4

Multiscale Modeling of KPFM

In this chapter, the multiscale approach for simulations of KPFM with atomic resolution is described in details. Based on the results of previous chapter and evaluating the chemical interactions subject to external bias in this chapter, expressions for V_{LCPD} in AM- and FM-KPFM are obtained and evaluated, first for ultrasmall, then for finite tip oscillation amplitude, and their magnitude and dependence on are explained. In chapter 3 we showed that both tip and cantilever contribute to the electrostatics force and its gradient over a nonconducting surface where the electric field penetrates into the sample. Here we show that how this electric field affects the atomic-scale contrast in KPFM images. We also explain how the influence of the effective bias V can be included into atomistic calculations, as well as shortcomings of previous attempts to do so. We critically discuss previous atomistic calculations, as well as experimental evidence for short-range electrostatic interactions. Highly accurate density functional calculations for nano-scale tip-sample systems are then discussed and illustrated for a realistic Si tip close to a NaCl(001) slab as an example of current interest.

Experimental limitations and evidence for the predicted trends, as well as desirable measurements are also briefly discussed. The same framework is used in the next chapter where the multiscale is simplified by introducing some approximations.

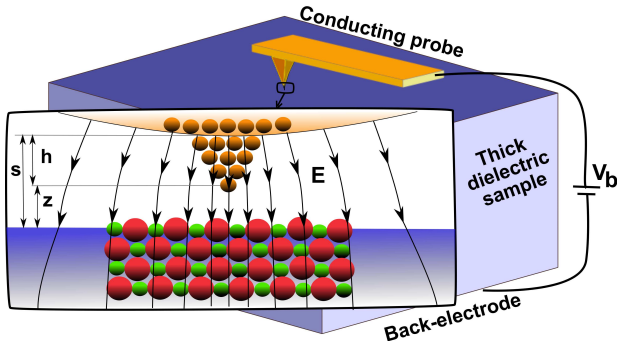

Figure 4.1: Sketch of the AFM setup showing its macroscopic and microscopic parts on two very different scales. The macroscopic tip-sample separation is $s = z + h$, where h is the nanotip height and z is the distance between the apex atom and the sample surface. (Both h and z are nominal values with atomic relaxation due to chemical interactions excluded.) Zoom window: the macroscopic electric field $\mathbf{E}$ depicted by the black field lines is applied as an external field to the atomistic subsystem.

4.1 Multiscale Approach

The method is multiscale in the sense that it couples the interactions of different types acting on widely different length scales, i.e. from a fraction of nm in the contact point to about a mm in the macroscopic bodies, as schematically illustrated in Fig. 4.1. The macroscopic system treated by classical electrostatics consists of the probe (cantilever plus tip) and of a sample described by its bulk dielectric constant. The bias voltage V_b is applied between the probe and the grounded back electrode, considered as perfect conductors. The microscopic system consists of a protruding nanotip less than 1 nm away from a slab of a few sample layers, both treated atomistically. Applying the electric field generated by the macroscopic tip and sample to the microscopic system leads to an unambiguous definition of V_{LCPD} on defect-free, overall neutral surfaces of crystalline materials. This provides the desired well-defined relationship between the bias-voltage and short-range forces which was lacking in previous approaches to LCPD contrast based on DFT computations. [36, 68, 69]

4.2 Bias-Dependent Chemical Interactions

4.2.1 Model and Method

Computations are performed within the local-density approximation to density functional theory (DFT) using norm-conserving HGH pseudopotentials [102] and the BigDFT package. [30] Relying on a wavelet basis set with locally adjustable resolution, this package calculates the

self-consistent electron density, the total energy and its electrostatic component with selectable boundary conditions [103] i.e. periodic in two directions and free in the third in our case. This allows us apply an external field perpendicular to the surface without artifacts which can arise from periodic images in the z direction when using plane-wave of mixed basis sets. The voltage biased macroscopic system determines the uniform electric field $E_z \propto V = V_b - V_{CPD}$ applied to the microscopic part.

As illustrated in the zoom window of Fig. 4.1, our microscopic system consists of a nanotip of height h protruding from the spherical end of the macroscopic tip and of a wider two-layer slab of sample atoms. Figure 4.2 illustrates the microscopic system used in the DFT computations reported here. The nanotip at the very end of a silicon tip is modelled as a cluster with a fixed (001) base of eight Si atoms with all dangling bonds passivated by H atoms in order to mimic the connection to the rest of the tip. The remaining Si atoms were pre-relaxed using the Minima Hopping Method [104] previously employed to generate low-energy structures of silicon clusters and of similar model tips. [105, 106] The free Si atoms adopted a disordered configuration with several exposed under-coordinated atoms. In particular the protruding apex atom is threefold coordinated and hence has a dangling bond with a small dipole moment pointing towards the surface. As will be discussed later, the relatively large size of this model nano-tip makes it possible to capture most of the charge redistribution induced in the tip apex due to the tip-sample short range interactions.

As we verified, a distance five times the lattice constant of NaCl is large enough to get rid of the electrostatic interaction between this nano-tip and its images in the main in-plane symmetry directions along which periodic boundary conditions applied. Therefore our sample slab consists of two 10×10 NaCl(001) layers containing 200 ions in total. For such a large system, it is sufficient to perform calculations only at one single k-point, namely center of the surface Brillouin zone. Pre-relaxation of the sample only caused a small rumpling which preserved the basic periodicity of the truncated (001) surface. Although the silicon model tip and the sample were initially individually pre-relaxed, all tip and sample atoms were subsequently frozen in some of our KPFM simulations. In this way we could assess pure electronic polarization effects without effects due to the interaction-induced displacements of ion cores.

The silicon model tip was positioned so that its foremost atom was 6.5 Å above a sodium or chlorine surface ion, then moved towards the sample in steps of 0.2 Å. At each step the Kohn-Sham equations are solved iteratively. The topmost layer of the Si tip and the bottom layer of the slab are kept fixed while other ions are free to relax until the Hellman-Feynman force exerted on each ion becomes smaller than 1 pN. This extremely tight tolerance is, however, required only when detecting the variation of the force upon small changes in the bias is required. As explained in chapter 1, the force F_μ exerted on the model tip is obtained by summing the z-components of

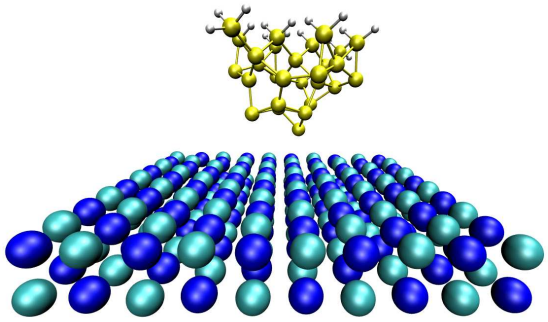

Figure 4.2: The microscopic model system used in the DFT calculations. The apex of a silicon AFM tip is modelled as a pre-relaxed $Si_{29}H_{18}$ cluster. All eight atoms in the top layer are passivated by hydrogen atoms and kept fixed at their bulk positions. The model sample consists of two NaCl(001) layers each containing 10×10 ions with the bottom layer kept frozen. Periodic boundary conditions are applied only along the lateral directions.

the forces over atoms of the tip. Since the free atoms are well relaxed, their contribution to that force is not significant and was used to ensure that the error is small enough.

4.2.2 Force spectroscopy

Figure 4.3 shows the microscopic force versus the tip-apex separation from Cl and Na surface sites. The same procedure is repeated at each tip-sample separation for a few field strengths E_z corresponding to $-2 \leq V = V_b - V_{\mathrm{CPD}} \leq 2$ Volts. For such biases and distances where F_μ becomes site-dependent, a nearly uniform macroscopic electric field of ~ 0.15 V/nm occurs in the vacuum gap per unit bias voltage, as shown in the inset of Fig. 3.15. No instabilities caused by electronic and/or atomic rearrangements appeared in that range of parameters.

4.2.3 Force varsus bias

The variation of the microscopic force as a function of bias voltage at the particular separation $z =$ 0.3 nm is shown in the insets in Fig. 4.3. In contrast to the macroscopic capacitive force, *the short-range force depends linearly on the applied bias voltage*. Theoretical considerations which explain this linear behavior will be presented in the next chapter where we show that this linear term is remarkably close to the interaction between distance-dependent but bias-independent charge densities on the tip and sample with the macroscopic electric field. Earlier studies obtained such a term by treating native ions or charged atoms adsorbed on the sample surface and/or the tip apex as point charges. [37, 76, 77] Deviations from the linear behavior could occur for larger biases, especially near instabilities, as observed in computations for a charged nanotip. [71]

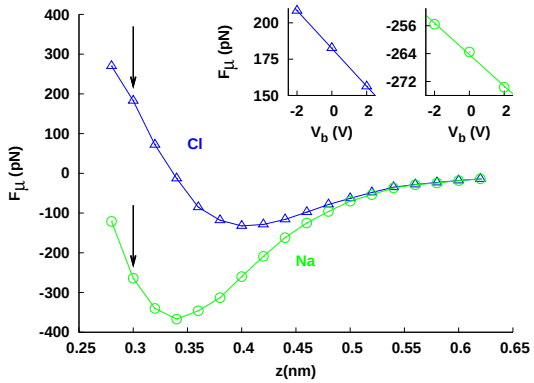

Figure 4.3: Microscopic force on the Si nanotip above Na and Cl surface ions from *ab initio* calculations without an applied electric field. Insets: variation of the force as a function of the macroscopic bias voltage at a tip-surface separation of 0.30 nm indicated by arrows.

Slope a

The *basic quantity which determines the deviation of the LCPD from the background CPD* is the voltage-independent slope of the short-range force with respect to the applied voltage

$$a(\mathbf{R}) = \frac{\partial}{\partial V} F_\mu(\mathbf{R}; \mathbf{E}(V)), \tag{4.1}$$

where $\mathbf{R} = (x, y, z)$ denotes the tip position. As discussed in chapter 2, the background CPD is not a well-defined quantity for an insulator. For a real doped silicon tip-NaCl(001) sample, it would be different from the CPD of our microscopic system if charge equilibrium is achieved, as enforced by the self-consistency of the computations. Besides, no CPD is explicitly included in the description of the macroscopic system. Thus the effective bias $V = V_b - V_{\mathrm{CPD}}$ would differ from that in a real system. Nevertheless, as long as this bias is in the Volt range, the slope a is unaffected.

The slope a, shown in Fig. 4.4(a), exhibits characteristic site-dependent distance dependence at separations less than 5 Å, and is larger above the more polarizable Cl ion. The underlying physics will be explored in the next chapter. The microscopic force-gradient F'_μ is also a linear function of bias voltage; The variation of its slope

$$a'(\mathbf{R}) \equiv \frac{\partial F'_\mu}{\partial V} = \frac{\partial}{\partial z}\frac{\partial F_\mu}{\partial V} = \frac{\partial a}{\partial z} \tag{4.2}$$

with distance, calculated by a second order finite difference approximation, is shown in Fig. 4.4(c). Figures 4.4(b) and (d) show that a and a' are stronger if relaxation due to chemical interactions is not performed; but the contrast appears below nearly the same distance and exhibits almost the same distance dependence. Thus, for the assumed neutral Si nanotip, the contrast is mainly due to electronic polarization rather than to bias-induced ion displacements.

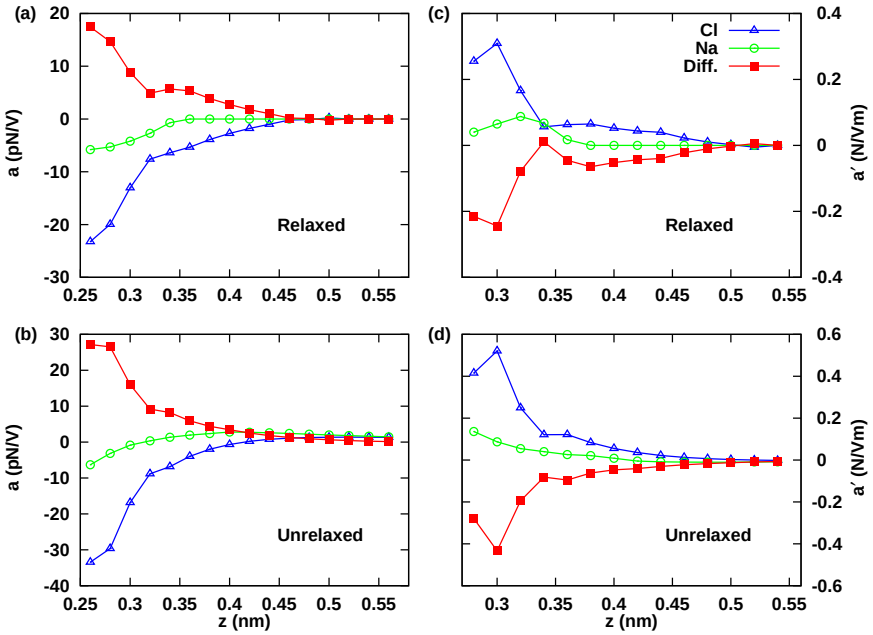

Figure 4.4: Distance-dependence of a and a' above Na and Cl surface ions with (a,c) or without(b,d) relaxation of the free atoms and ions during the tip approach. The difference between Na and Cl sites (i.e. the contrast) is shown by red (filled) symbols.

4.2.4 Contribution to the total force

In the approximation that the macro- and microscopic systems are coupled only through the macroscopic electric field (the required corrections are discussed in the following), the z-component of the V-dependent total force exerted on the tip is

$$F - F_{\mathrm{vdW}} = F_M(s;V) + F_\mu(\mathbf{R};\mathbf{E}(V)) \tag{4.3}$$

where, as before, $s = z + h$ and $V = V_b - V_{\mathrm{CPD}}$. However, some corrections might need to be considered beyond this approximation as well be discussed in the following. The macroscopic electrostatic force is quadratic $F_M = \frac{1}{2}C'(s)V^2$ contrary to the microscopic force F_μ which was shown to be linear in V.

Note that the vdW force, being only a function of the mesoscopic geometries, is bias- and site-independent, and is therefore henceforth ignored, although it affects the overall resonance frequency shift Δf in a NCAFM measurement.

4.2.5 Corrections

Microscopic capacitance: The first correction is due to an additional capacitive contribution $\frac{1}{2}\delta C' V^2$ caused by the presence of a polarizable nanoscale material (nanptip) in the gap between the macroscopic bodies. Owing to the small lateral dimensions of the nanotip compared to the radius of the macroscopic tip end, this correction is small, [77] although it can become noticeable and site-dependent if the nanotip is strongly polarizable and nearly contacts the sample. [71, 77] In chapter 5 we show that this term is small for our model system.

Constant charge/dipole shift: The second correction arises if the microsystem contains a localized net charge [27] or permanent dipole moment [107]. This leads to a site-independent LCPD with an approximate power-law approach towards a background CPD of several Volts. The interaction of the nanotip charge distribution with the macroscopic field $\mathbf{E}$ could in principle be included in our description at separations s where $\mathbf{E}$ can no longer be considered uniform. In that range, however, the charge or dipole might be approximated as point objects, as justified in the case of a conducting sample in the Supplementary Material of Ref. [76]. Because the charge or dipole are intrinsic (i.e. V-independent), their interaction with external field is proportional to V, so that this correction would give rise to long-range contributions to the slopes a and a' (see next chapter). [76, 77] In the case of our neutral Si nanotip and sample slab, this correction is small.

Nanotip size: The third correction arises because in reality the nanotip is in electrical contact with the macroscopic tip, and the electron density at the interface differs from that near the top of the isolated cluster used as the nanotip if it is small. Then the charge distribution near the apex, which dominates F_μ could also be affected. We used a rather large model tip in order to get rid of this shortcoming.

4.3 Computed KPFM signals

The force gradient is more sensitive than the force to short-range interactions which are responsible for atomic-scale contrast in NCAFM and KPFM. Direct detection of the gradient is in principle possible if the variation of F_μ over the peak-to-peak oscillation amplitude remains linear, e.g. if it is comparable to the spacing 0.2 Å of the computed points in Fig. 4.3. We first consider this simple limit which is commonly assumed in the KPFM literature, but is seldom achieved in NCAFM experiments. This is useful to explain the idea behind connecting our simulations results to experimentally measured signals. Finally, we address the finite amplitude oscillations which is relevant to the experiments.

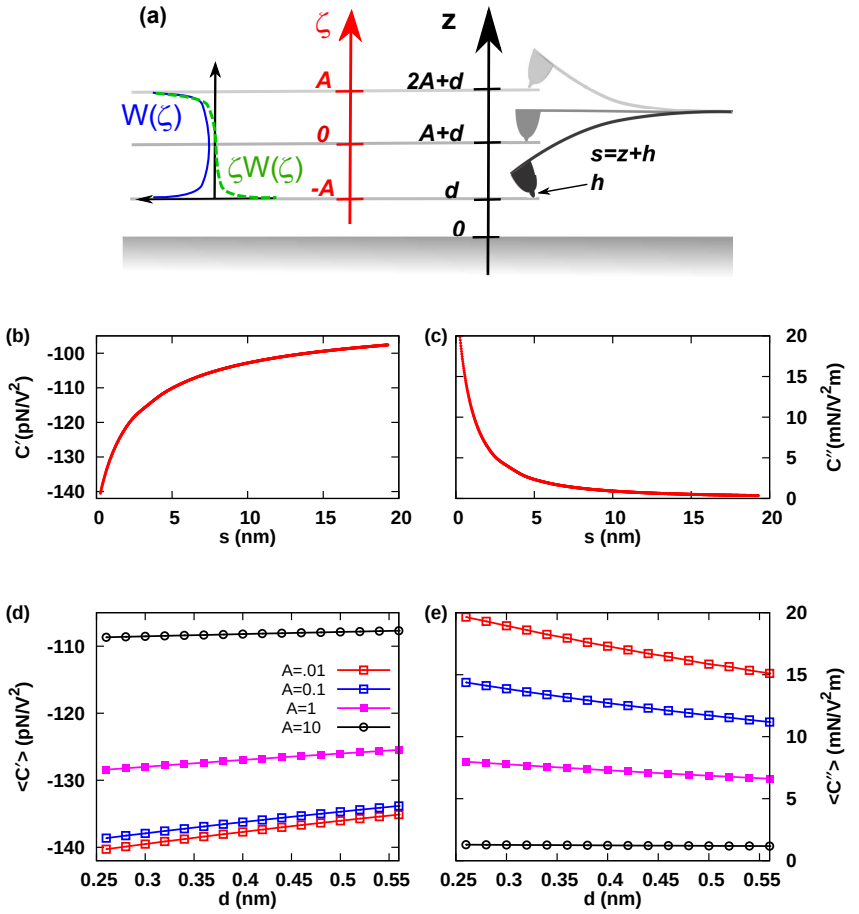

Figure 4.5: (a) Sketch of the cantilever-tip probe oscillating in its fundamental mode with a finite amplitude A. Dependencies of the first (b) and second (c) gradients of the capacitance on the macroscopic separation $s = z + h$ calculated for the setup shown in Fig. 3.10(a), and of their cycle averages (d) and (e) tip oscillation amplitudes A =0.01, 0.1, 1 and 10 nm as a function of the closest approach distance of the nanotip apex $d = z_{\min}$. The averages are calculated using Eqs. (4.14) and (4.15). The weight functions $w(\zeta)$ and $\zeta w(\zeta)$ are plotted as a function of $\zeta = z - d - A$ in (a).

4.3.1 Ultrasmall amplitude limit

As explained in chapter 2, the electrostatic contribution to the tip-sample force is minimized in the KPFM technique by applying a bias voltage which consists of a modulating AC voltage with angular frequency ω in addition to the DC voltage

$$V_b(t) = V_{\rm DC} + V_{\rm AC}\cos\omega t \tag{4.4}$$

resulting in an effective potential difference

$$V = (V_{\rm DC} - V_{\rm CPD}) + V_{\rm AC}\cos\omega t. \tag{4.5}$$

Ignoring the vdW force in Eq. (4.3), the total force is then decomposed into three spectral components

$$F(t) = F_{\rm DC} + F_\omega\cos\omega t + F_{2\omega}\cos 2\omega t \tag{4.6}$$

which is different from Eq. (2.6) in the sense that F_μ is also now contributing

$$\begin{aligned} F_\omega &= \Big[\frac{d(F_M + F_\mu)}{dV_b}\Big]V_{\rm AC} \\ &= \Big[C'(s)\,(V_{\rm DC} - V_{\rm CPD}) + a(\mathbf{R})\Big]V_{\rm AC}, \end{aligned} \tag{4.7}$$

where we have assumed that the response to $V_{\rm AC}$ is linear and instantaneous. Note that F_ω is the only relevant component because $V_{\rm LCPD}$ is operationally defined by nulling the KPFM signal generated by this force component. This fact is discussed in the following for the two common KPFM methods, i.e. AM and FM.

AM-KPFM

In the ultrasmall amplitude limit, the deflection signal detected in AM-KPFM is proportional to $F_\omega\cos\omega t$ which is nulled if

$$V_{\rm DC} = V_{\rm LCPD}^{\rm AM} = V_{\rm CPD} - \frac{a(\mathbf{R})}{C'(s)}. \tag{4.8}$$

Because the background $V_{\rm CPD}$ is not well-defined as discussed before, and only $a(\mathbf{R})$ is site-dependent, we consider only the deviation of $V_{\rm LCPD}$ from $V_{\rm CPD}$ which is responsible for atomic-scale contrast, i.e.

$$L^{\rm AM} \equiv V_{\rm LCPD}^{\rm AM} - V_{\rm CPD} = -\frac{a(\mathbf{R})}{C'(s)}. \tag{4.9}$$

As illustrated by the points for $A = 0.01$ nm in Fig. 4.5(d) for a dielectric sample, the z-dependence of C' is weak over the range ($s = z + h < 1$ nm) where $a(\mathbf{R})$ is appreciable (see

Fig. 4.4(a)). Therefore $V_{\mathrm{LCPD}}^{\mathrm{AM}} - V_{\mathrm{CPD}}$ differs from $a(\mathbf{R})$ by an essentially z-independent factor. Depending on the nanotip height h, this may no longer hold in the case of a conductive sample or thin dielectric film on a conductive substrate. Note that the ultrasmall amplitude $A = 0.01$ nm is equivalent to $A \to 0$ in our calculations, as will be discussedin this section.

FM-KPFM

In FM operation mode, the frequency shift is the detected signal. We assume that the total frequency shift is obtained from the sum of the different contributions present in Eq. (4.3), namely

$$\Delta f - \Delta f_{\mathrm{vdW}}(s) = \Delta f_M(s; V) + \Delta f_\mu(\mathbf{R}; \mathbf{E}(V)). \tag{4.10}$$

In FM-KPFM the contribution of the modulated electrostatic force component F_ω to the frequency shift of the first resonant mode Δf_1 is detected and nulled. In the ultrasmall amplitude limit Δf_1 is proportional to the force-gradient [17] (see Eq. (1.1)), and would therefore be nulled if

$$F'_\omega \quad = \quad \left[C''(s)\,(V_{DC} - V_{CPD}) + a'(\mathbf{R})\right] V_{AC}$$

is nullified. The FM-counterpart of Eq. (4.9) is therefore

$$L^{\mathrm{FM}} \equiv V_{\mathrm{LCPD}}^{\mathrm{FM}} - V_{\mathrm{CPD}} = -\frac{a'(\mathbf{R})}{C''(s)}. \tag{4.11}$$

The site- and distance dependence of L^{FM} is again mainly determined by $a'(\mathbf{R})$ because in the range $s < 1$ nm where a' is appreciable C'' the denominator of Eq. (4.11) is almost constant, cf. Fig. 4.4(c) and the points for $A = 0.01$ nm in Fig. 4.5(e). The calculated LCPD deviations for $A = 0.01$ nm in the AM and FM modes are plotted in Figs. 4.6(a) and (e).

For the ultrasmall amplitude A=0.01 nm, which would likely not provide an adequate signal-to-noise ratio in practice, the calculated L^{FM} is about hundred times stronger than L^{AM} and exceeds the range of validity (± 2 V) of our DFT computations (see the horizontal lines in Figs. 4.6(e,f)), as well as the range of experimentally measured values. Hence this result cannot be trusted and it is important to consider averaging over the range covered by the finite tip oscillation, as is explained in the following.

4.3.2 Finite amplitude oscillations

In NCAFM with cantilevers the oscillation amplitude A is between several and a few tens of nanometers, so that the macroscopic capacitive electrostatic force can change by several orders

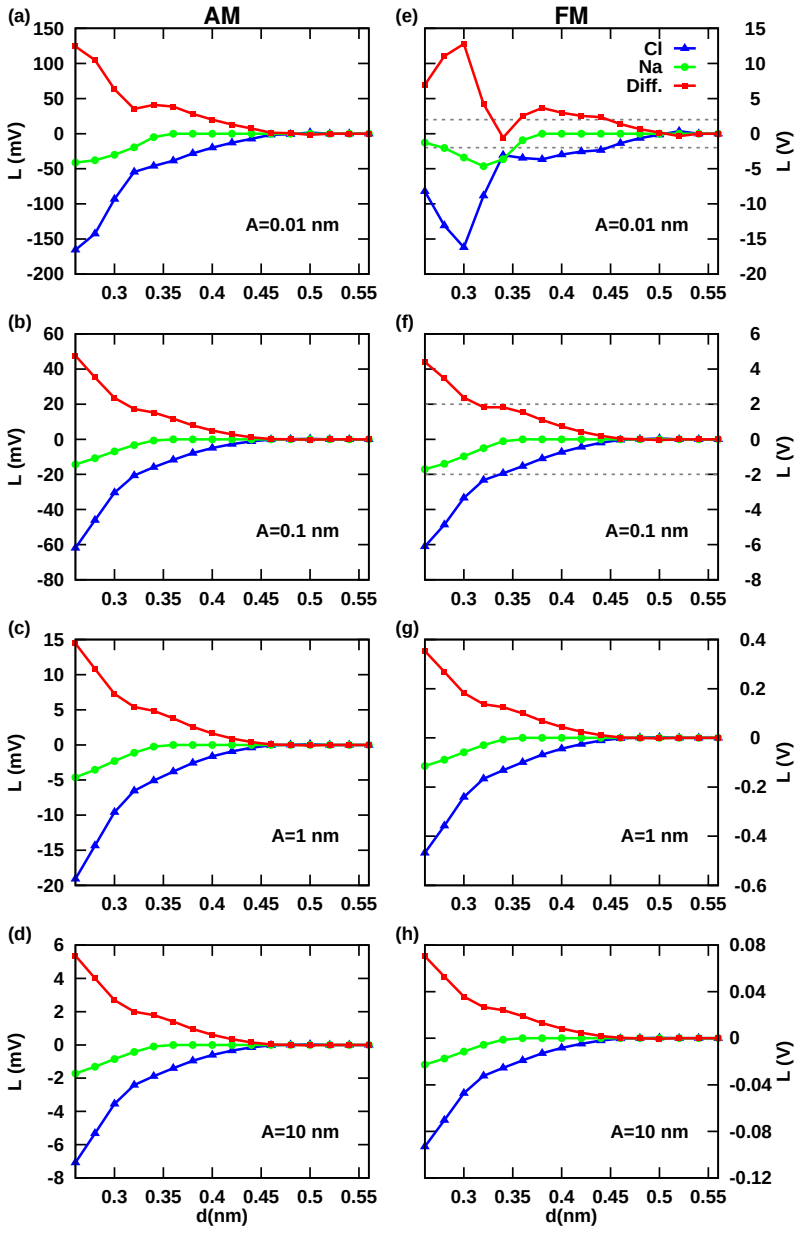

Figure 4.6: Calculated deviations $L = V_{\mathrm{LCPD}} - V_{\mathrm{CPD}}$ for AM- (left column) and FM-KPFM (right column) versus closest tip apex-sample distance for tip oscillation amplitudes $A = 0.01$ nm (a,e),$A = 0.1$ nm (b,f), $A = 1$ nm (c,g), and $A = 10$ nm (d,h). In (e,f) the dashed horizontal lines indicate the range of validity of our DFT calculations (± 2 V).

of magnitude over an oscillation cycle. In practice, the detected AM and FM KPFM signals are given by differently weighted averages, namely [7]

$$\langle F_\omega \rangle = \frac{1}{2\pi}\int_0^{2\pi} F_\omega[d + A(1+\cos\phi)]d\phi$$

and [18]

$$kA\frac{\Delta f_\omega}{f_\omega} = -\frac{1}{2\pi}\int_0^{2\pi} F_\omega[d + A(1+\cos\phi)]\cos\phi d\phi$$

where k is the flexural stiffness of the cantilever and $d = z_{min}$ is the closest tip apex-sample separation. Substituting the force from Eq. (4.7) and setting these averages to zero, one obtains

$$L^{\mathrm{AM}} = -\frac{\langle a(\mathbf{R})\rangle_w}{\langle C'(s)\rangle_w}, \tag{4.12}$$

$$L^{\mathrm{FM}} = -\frac{\langle a'(\mathbf{R})\rangle_{1/w}}{\langle C''(s)\rangle_{1/w}}, \tag{4.13}$$

where the cycle averages depend both on d and A and are defined as

$$\langle g\rangle_w \equiv \frac{1}{\pi}\int_{-A}^{A} w(\zeta)g(d+A+\zeta)d\zeta, \tag{4.14}$$

$$\begin{aligned}\langle g'\rangle_{1/w} &\equiv \frac{1}{\pi A}\int_{-A}^{A}\zeta w(\zeta)g(d+A+\zeta)d\zeta \\ &= \frac{1}{\pi A^2}\int_{-A}^{A}\frac{1}{w(\zeta)}g'(d+A+\zeta)d\zeta.\end{aligned} \tag{4.15}$$

As depicted in Fig. 4.5(a),

$$\zeta = z - (A+d)$$

whereas the weight functions

$$w(\zeta) = 1/\sqrt{A^2-\zeta^2}$$

and $\zeta w(\zeta)$ have square root singularities at the turning points of the oscillation. Note that if $A \to 0$ then $\langle g\rangle_w$ tends to $g(d+A)$. Similarly, the expression on the second line of Eq. (4.15) justifies the notation $\langle g'\rangle_{1/w}$ and shows that this quantity tends to $g'(d+A)$ when $A \to 0$, besides helping to relate the distance dependence of L^{FM} to those of a' and $C''(s)$. However, because a is computed with high precision, whereas a' is obtained by interpolation, we use the expression on the first line for numerical purposes. Furthermore, since a is known only at equispaced separations z_i where the DFT computations have been performed, the integrals in Eqs. (4.14) and (4.15) must be discretized. The adopted procedure, which deals with the singularities of the weight function $w(\zeta)$ at the integration limits, [108] is presented below. An important result is that the discretized version of the expression in the first line of Eq. (4.15) reduces to the second order FD approximation of $g'(d+A)$ when $2A$ matches the spacing between adjacent z_i values, in accordance with the expression on the second line.

4.3.3 Discretized integrals for finite tip oscillation amplitudes

Assuming that $N+1$ equispaced data points $\{z_i\}$ are sufficiently close together such that $g(z)$ remains almost constant within an interval length $\delta = 2A/N$, the integration in Eq. (4.14) can be approximated by a finite sum

$$\langle g(z)\rangle_w \simeq \frac{1}{\pi}\sum_{i=0}^{N} w_i g_i$$

where $g_i \equiv g(z_i)$ is either $g_i = a(z_i)$ or $g_i = C(z_i + h)$; Since $w(\zeta) = 1/\sqrt{A^2-\zeta^2}$ we obtain

$$w_i = \int_{\zeta_i^-}^{\zeta_i^+} w(\zeta)d\zeta = \arcsin(\frac{\zeta_i^+}{A}) - \arcsin(\frac{\zeta_i^-}{A})$$

where

$$\zeta_i^{\pm} = (i \pm \frac{1}{2})\delta - A$$

are the midpoints between ζ_i and $\zeta_{i\pm1}$. Taking into account the rapid variation of $w(\zeta)$ near the integration limits defined as $\zeta_0^- = -A$ and $\zeta_N^+ = A$, the square root singularities of $w(\zeta)$ at those turning points are approximately included with this modified trapezoid integration method. Sufficiently far from those points $w_i \simeq w(\zeta_i)\delta$ so that the standard trapezoid approximation is recovered. The analogous approximation for Eq.(4.15) namely

$$\langle g'(z)\rangle_{1/w} = \langle g(z)\rangle_{\zeta w} \simeq \frac{1}{\pi}\sum_{i=0}^{N} w_i^* g_i$$

involves [108]

$$w_i^* = \frac{1}{A}\int_{\zeta_i^-}^{\zeta_i^+} \zeta w(\zeta)d\zeta = \sqrt{1-\left(\frac{\zeta_i^-}{A}\right)^2} - \sqrt{1-\left(\frac{\zeta_i^+}{A}\right)^2}.$$

Note that in the $A \to 0$ limit only the data points at the two limits are taken into account. Indeed, if $N = 1$, $A = \delta/2$ and $W_0 = W_1$, hence $\langle g\rangle_w = (g_0 + g_N)/2$, and $W_0^* = -W_1^*$, hence $\langle g'\rangle_{1/w} = (g_N - g_0)/2A$, so that Eqs. (4.12) and (4.13) consistently approximate the corresponding zero-amplitude equations, Eqs. (4.9) and (4.11). Similarly, if N=2, $A = \delta$ and one obtains $w_0 = w_2, w_1 = 0$ and $w_0^* = -w_2^*, w_1^* = 0$ and Eqs. (4.9,4.11) are again recovered.

4.4 Results

Owing to the very different z-dependencies of $a(z)$ and $C'(s)$, shown respectively in Figs. 4.4(a) and 4.5(b), their cycle averages depend in different ways on d and A. The same holds for $a'(z)$

and $C''(s)$, shown respectively in Figs. 4.4(c) and 4.5(c). Figures 4.5(d) and 4.5(e) show the cycle averages of C' and C'' versus the closest tip-apex approach distance d for oscillation amplitudes $A = 0.01$, 0.1, 1 and 10 nm, whereas the cycle-averages of V_{LCPD} calculated from Eqs. (4.12) and (4.13) are plotted in Fig. 4.6 for AM-KPFM (left column) and FM-KPFM (right column) for the same amplitudes in the range where $a(z)$ is finite. In that range, the cycle averages for $A = 0.01$ nm agree with the non-averaged quantities. Since the primary quantities were calculated at points spaced by 0.02 nm, this is not surprising in view of the remarks at the end of the preceding subsection. Thus, apart from small deviations introduced by the discretization procedure, the points in Figs. 4.6(a) and 4.6(e) which were actually calculated for $A = 0.01$ nm coincide with those given by Eqs.(4.9) and (4.11), and exhibit essentially the same distance dependencies as $a(d)$ and $a'(d)$, as already discussed in 4.3.1.

Already above $A = 0.1$ nm, however, the LCPD contrasts in both modes exhibit almost the same spatial dependence as $a(d)$, although their respective magnitudes decrease if A is increased. Nevertheless, L^{FM} significantly exceeds L^{AM}; this can be understood as follows. As seen in Figs. 4.5(d) and 4.5(e), $\langle C''\rangle_{1/w}$ drops much faster than -$\langle C'\rangle_w$ if A is increased. As explained in the discussion of Fig. 3.14(b) this behavior reflects the increasing influence of the relative contributions of the tip shank and of the cantilever to $C'(s)$ in the range covered by the peak-to-peak oscillation. Especially $\langle C'\rangle_w$ is affected by the cantilever contribution which causes the very gradual levelling of $C'(s)$ apparent in Fig. 4.5(b). As seen in Fig. 4.5(c), this slowly varying contribution tends to cancel out in $C''(s)$, and, according to the second line in Eq. (4.15), in $\langle C''\rangle_{1/w}$ as well.

On the other hand, $\langle a\rangle_w$ and $A\langle a'\rangle_{1/w}$ essentially coincide once a exceeds the range where a is noticeable. Indeed, the main contributions to those averages come from the vicinity of $z = d$ where the integrands in Eqs. (4.14) and (4.15) (first line) match. Expanding $w(\zeta)$ about this turning point, one finds that $\langle a\rangle_w \sim A^{-1/2}$ whereas $\langle a'\rangle_{1/w} \sim A^{-3/2}$, just like Δf_1 behaves in NCAFM. [18] According to Fig. 4.5(b,c) the same argument cannot be applied to $\langle C''\rangle_{1/w}$ for $A \leq 10$ nm, and not at all to $\langle C'\rangle_w$ because $C'(s)$ varies only slowly up to $s = R = 20$ nm. Fig. 4.7 shows how the finite oscillation amplitude affects the relevant cycle averages, as well as ΔV_{LCPD} in the AM mode (left column) and in the FM mode (right column) at the closest tip apex-sample separation $d = 0.30$ nm indicated by arrows in Fig. 4.3.

The same trends persist at all separations $d < 0.5$ nm where LCPD contrast appears. $\langle a\rangle_w$ drops as $A^{-1/2}$, and $\langle a'\rangle_{1/w}$ drops as $A^{-3/2}$ already beyond $A = 0.1$ nm, while $\langle C'\rangle_w$ varies only little and $\langle C''\rangle_{1/w}$ begins to drop somewhat slower than A^{-1} only above $A = 1$ nm. The resulting amplitude dependencies in both modes reflect the different dependencies of the numerators and denominators in Eqs. (4.12) and (4.13).

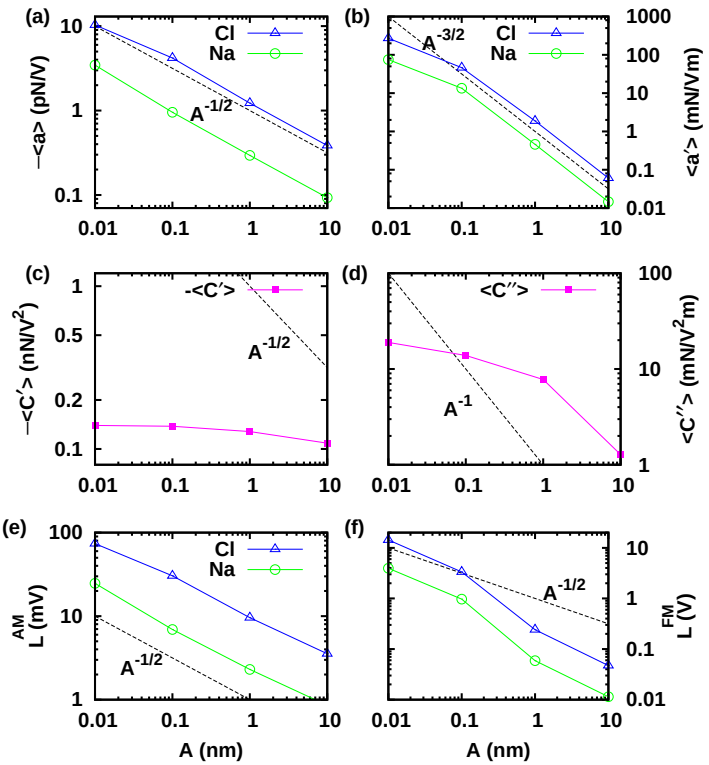

Figure 4.7: Amplitude dependencies of the cycle averages $\langle a \rangle_w$ and $\langle a' \rangle_{1/w}$ (a,b), $\langle C' \rangle_w$ and $\langle C'' \rangle_{1/w}$ (c,d) and of the resulting deviations L^{AM} and L^{FM} (e,f) at a closest tip apex separation of d=0.3 nm above Cl and Na surface sites. Dashed lines show the corresponding powers of A.

4.4.1 Discussion

Expressions formally similar to Eqs. (4.9) and (4.11) were previously suggested. There are essential differences, however. In some works, the denominators came from a *short-range polarization contribution* $\propto V^2$ (see next chapter) to the *microscopic force* F_μ rather than from the much larger capacitive force F_M which is seriously underestimated [70] or completely ignored. [68, 69] Barth *et al* [76] used the correct capacitive term in the denominators but the numerators are obtained from a classical methods, constant point charges or fixed dipoles which were applicable only to separations larger than the onset of the chemical interactions. They showed that a point charge or fixed dipole attached to the tip apex would shift the KPFM signal by a constant value [76] but then not the contrast.

Nony *et al* [70] also noticed that $\langle a \rangle_w$ and $A\langle a' \rangle_{1/w}$ almost coincide when A exceeds a few nanometers. This results in a comparable V_{LCPD} for AM and FM modes if A exceeds a few nanometers. However, by including the correct F_M and taking into account the different

amplitude dependencies of the denominators in Eqs. (4.12) and (4.13), we conclude that the contrast should remain larger in the FM than in the AM mode for a given closest approach distance d and oscillation amplitude A. This prediction is independent of the particular system considered, but the mode-dependent signal to noise ratio must also be considered. Thus Kawai *et al.* [7] calculated the minimum detectable CPD as a function of A and showed that it is smaller in the AM mode. Taking into account the discussions of Figs. 3.14 and 4.5, $\langle C' \rangle_w$ would be larger if the cantilever area is larger whereas $\langle C'' \rangle_{1/w}$ would be unaffected, whereas both quantities would be larger if the cone angle is broader or if the sample is a metal rather than an insulator, but $\langle C'' \rangle_{1/w}$ would be more strongly affected. On the other hand $\langle a \rangle_w$ and $\langle a' \rangle_{1/w}$ would be larger if the tip apex is charged [71] rather than neutral, or if the sample is a semiconductor with a reconstructed surface which exposes partially charged species like Si(111) 7×7 [36, 7]. From this point of view the system studied here is especially challenging. Furthermore, the contrast ratio slowly decreases if A is increased, e.g. by a factor which drops from about 100 to 10 for oscillation amplitudes between 0.01 and 10 nm in our example.

4.4.2 Experimental Limitations

For a meaningful comparison with NCAFM-KPFM measurements it is important to take *experimental limitations* into account. In view of the long-range LCPD variations due to surface and bulk inhomogeneities on real samples, one should compare computed atomic-scale LCPD variations with the difference between the LCPD measured at sub-nanometer separations d in the middle of a flat homogeneous island or terrace and the extrapolated long-range, essentially site-independent LCPD. This procedure would also suppress most of the long-range contributions to $\langle a \rangle_w$ and $\langle a' \rangle_{1/w}$ which would arise in the case of a charged or strongly polar tip [76]. Moreover, the comparison should be done with the same tip at constant d (slow distance control) because atomic-scale variations of d at constant $\Delta f_1(x, y, d)$ would induce such variations in the LCPD even if the latter is site-independent but has a different distance dependence as Δf_1.

For the distance controller to function properly, Δf_1 must be chosen on the branch where this frequency shift becomes more negative if d is decreased. Furthermore, the maximum restoring force kA must be much larger than the maximum tip-sample attraction [18]. For measurements with standard NCAFM cantilevers ($k \sim$ 20-40 N/m) this criterion is typically satisfied by using oscillation amplitudes $A > 5$ nm, and atomically resolved imaging is typically performed at distances $d \sim$ 0.4-0.5 nm. According to Fig. 4.6 the LCPD contrast which is then predicted to be 20-100 mV in the FM mode and a few mV in the AM mode approaches the experimental limits in both modes. Even if the AM-KPFM signal is enhanced by setting the modulation frequency at the second flexural resonance of the cantilever, the LCPD contrast predicted by our model would remain the same. This contrast would be stronger if the tip were charged. Unfortunately,

available data showing atomic-scale contrast on (001) surfaces of NaCl and KBr is insufficient for a meaningful comparison between AM and FM KPFM. However, LCPD maps obtained with sputter-cleaned Si tips and similar measurement parameters on Si(111) 7×7 surfaces show that the contrast between Si adatoms and corner holes in the FM-mode [36] is about ten times stronger than in the AM-mode [7]. Moreover, data obtained from a direct determination of the maximum of Δf_1 versus bias voltage V_b agreed well with those obtained by nulling the FM-KPFM signal at the modulation frequency [36].

The sizable LCPD contrast of several Volts predicted in the FM mode for amplitudes $A <$ 0.1 nm should, however, be readily observable when using a tuning fork instead of a cantilever. Owing to the much higher stiffness $k \simeq 1800$ N/m of this deflection sensor, the above-mentioned criterion can be satisfied with such amplitudes close to the ultrasmall limit [16]. Combined NCAFM-KPFM measurements using such tuning forks with PtIr tips have only been done at low temperature by the time-consuming direct method mentioned before. [40, 39] Unfortunately, no FM-KPFM measurements showing atomic-scale LCPD contrast on alkali halide (001) surfaces have so far been reported.

4.5 Summary and Outlook

A general multiscale approach was proposed to compute electrostatic forces responsible for atomic-scale contrast in KPFM performed simultaneously with NCAFM. The approach is not restricted to particular sample or tip materials and can be used for conductors and semiconductors thin/thick films and samples. The problem is split into two parts coupled in a remarkably simple but novel fashion. The electrostatic problem of the voltage-biased AFM probe over a grounded sample, is solved first. The solution provides not only the electrostatic interaction but also the electric field distribution in the contact point. Instead of the bias voltage V_b, the nearly uniform electric field obtained in that range is then applied as an external field to the microscopic part which can be treated by empirical atomistic or first principles methods. The *ab initio* BigDFT wavelet code employed here enabled us to compute the short-range bias-dependent force on the tip apex represented by a cluster with an unprecedented accuracy of 1 pN. For the Si-nanotip-NaCl(001) system considered here, this microscopic force F_μ is a linear function of the bias in the investigated range $-2 \leq V_b - V_{\mathrm{CPD}} \leq 2$ Volts. We argue that this is a general result, except close to atomic-scale instabilities caused by strong enough forces which could arise at very small separations and/or very large effective biases.

Adding the macroscopic and microscopic bias-dependent forces, expressions are obtained for the KPFM signals in the AM and the FM modes. The atomic-scale deviation of LCPD from its common asymptotic value CPD at large separations is the ratio of the derivatives $a = \frac{dF_\mu}{dV_b}\Big|_{V_b=V_{\mathrm{CPD}}}$

and $C' = \frac{dC}{dz}$ averaged over the tip oscillation amplitude with different weights in AM- and FM-KPFM. We explain the amplitude dependence of the atomic-scale LCPD contrast in both modes and predict that for typical amplitudes used in measurements with standard NCAFM cantilevers, this contrast should be much stronger in the FM mode. This is a consequence of the contributions of the cantilever and the tip shank to the KPFM signal in the AM mode, which are stronger on insulating samples. The same conclusion has previously been reached in comparisons of AM- and FM-KPFM measurements of long-range LCPD variations; such variations are caused by interactions of the biased probe with CPD inhomogeneities and surface charges on scales of several nanometers and above on conducting samples partly covered with ultrathin overlayers of different materials [56, 57]. However, the strong mode-dependent influence of distant contributions to C' on the atomic-scale LCPD contrast has, to our knowledge, not been recognized because previous work on this topic assumed that only the tip apex mattered at sub-nanometer separations.

Because V_{LCPD} depends on measurement parameters, it is desirable to *extract the more fundamental quantity* a from combined KPFM measurements, just like the microscopic force F_μ is extracted from NCAFM measurements using, e.g. a widely accepted inversion algorithm [48] or one based on the direct inversion of the discretized version of the first line of Eq.(4.13) described in 4.3.3 by back-substitution [108]. In the next chapter, we show that a is approximately given by the gradient of the normal component of the dipole moment multiplied by the electric field normalized to $V_b - V_{\mathrm{CPD}}$. Since $L = V_{\mathrm{LCPD}} - V_{\mathrm{CPD}}$ is predicted to be stronger in FM-KPFM, whereas its distance dependence is governed by the weighted average $\langle a' \rangle_{1/w}$ modes, the most appealing way to obtain $a(d)$ would be to extract a' then integrate it from the range where L vanishes down to the desired separation d. The averages $\langle a' \rangle_{1/w}$ and $\langle C'' \rangle_{1/w}$ can be separately obtained from direct measurements of the frequency shift Δf_1 as a function of bias [36], namely from the shift of the maximum and the curvature of parabolic fits at several (x, y, d) positions. The signal/noise ratio of those averages can be improved by using AC modulation and lock-in detection at the modulation frequency. The averages could then be determined from the zero intercept L^{FM} and the slope of the FM-KPFM signal $\langle \Delta f_\omega \rangle_w$ versus DC bias. An analogous procedure could be applied to determine $\langle a \rangle_w$ and $\langle C' \rangle_w$ from the AM-KPFM signal $\langle F_\omega \rangle_w$, then a itself by inversion, using suitably modified algorithms [108, 109]. Because the AM-KPFM signal/ratio is much superior if the modulation frequency f is at the second cantilever resonance [7], L^{AM} could be determined more accurately even if it is smaller than in FM-KPFM. In any case, note that the slope a reflects variations of the electrostatic potential outside the sample surface which are, however, locally enhanced by the proximity of the tip apex. Since the latter is in turn also polarized and deformed [110], a cannot simply be described as the convolution of the unperturbed electrostatic potential with a merely distance-dependent tip point-spread function, as in macroscopic electrostatics. [64]

Complications due to averaging over the tip oscillation amplitude are to a certain extent avoided with tuning fork deflection sensors which enable direct measurements of $\langle \Delta f_\omega \rangle_w$ vs. bias, using amplitudes approaching the ultrasmall limit [40, 39]. Spectacular results have thus been obtained on isolated molecules adsorbed on a thin epitaxial NaCl(001) film by using tips with well-defined apex species stable at low temperature [111]. Most recently, L^{FM} contrast reflecting changes in the intramolecular charge distribution has been observed upon a configurational switch triggered by a judiciously applied pulse [12]. Our results shown in Figs. 4.6(e), (f) and (d) show that L^{FM} and a' still have a significant amplitude dependence between $A = 0.1$ and 0.01 nm, so that inversion is still necessary to obtain accurate results for typical amplitudes used with tuning fork sensors.

Since such measurements use hard metal tips, while metal-coated tips are also used in NCAFM and/or KPFM measurements with cantilevers it would desirable to develop appropriate nanotip models and to perform simulations like those described here. In particular, the recently fabricated sharp and stable W and Cr coated silicon tips [107, 53] and the stable atomic-scale resolution achieved with Cr-coated cantilevers at separations exceeding the usual range $d < 0.5$ nm merit further attention. Atoms or molecules intentionally picked by the apex and/or adsorbed on the sample [40, 111, 77, 12] would be worth studying using our approach in order to take into account modifications of their electronic and geometric structure due to bonding and charge transfer. Another class of systems which merit further investigations involve silicon nanotips with a picked-up cluster of foreign material, NaCl in particular, which have so far been studied by DFT in the absence of a sample [112] or represented by a cluster of the same material as the sample using empirical interaction potentials [72, 71].

Note finally that all macroscopic probe models, including ours, provide a better description of metallic or metal-coated tips than of real silicon tips. Indeed, even if the native oxide is removed by sputtering, a silicon layer of few nanometers depleted of charge carriers still separates the tip surface from the highly doped conducting tip interior. Although it was taken into account in previous treatments of KPFM of semiconductor devices [3], this depletion layer remains to be included when modelling Si tips, e.g. by allowing a smaller effective radius R of the equipotential at the applied bias voltage and a larger effective separation s from the sample surface.

Chapter 5

Microscopic Description of Atomic Scale Variation of CPD

Describing the underlying mechanism of the atomic variation in the KPFM images has been challenging and has attracted great interest. Since the development of KPFM techniques, the KPFM signals have been considered as *local* contact potential difference, because it is essentially the macroscopic CPD which in addition shows local variations if a sharp tip is used. The equations derived in the multiscale model in the preceding chapter demonstared, however, that the long range electrostatic forces, which determine the curvature of the frequency-shift parabola versus bias, play a crucial role in the quantity which is measured as KPFM signal. Furthermore, the experimentally measured KPFM signals are averaged quantities and do not represent the instantaneous variation of some local quantity unless the cantilever is driven with an ultra-small amplitude. As a consequence, one should be careful in ascribing the measured quantity to a local potential difference. In fact our model suggests that the KPFM signals represent a quantity which arises from a tight coupling between the macroscopic capacitive force and the slope of short range bias-dependent forces with respect to the applied voltage. Working in the same multiscale framework of the previous chapter, we present in this chapter a theory for atomic-scale variation of CPD.

How to define LCPD?

Several ways to attribute the calculated quantities to the LCPD have been proposed in the literature. Perhaps the most trivial definition, inspired also by its name *surface potential*, is to simply consider the electrostatic potential evaluated on a horizontal plane at a given distance from the surface. A Gaussian convolution (with a width of the same order as the tip radius) of this quantity was successfully used to reproduce the LCPD images of a random distribution of surface charged impurities. [113] The spatial resolution is, however, far from a nm scale and only slightly smaller than the tip radius. Mohn *et al.* [12] showed that the component normal to the surface of the electric field evaluated on such a plane a few Å above a free naphthalocyanine molecule looks very similar to the high resolution LCPD map of the molecule adsorbed on a NaCl thin film imaged by exploiting a CO molecule-terminated tip-apex. The reason why instead of the electrostatic potential its gradient is connected to the LCPD in this picture is that the tip-apex is indeed polarized by the electric field arising directly from the charge distribution of the molecule.

On the other hand, the nano tip could explicitly be included in the model. Model tip-sample systems considered simply as fixed charges/dipoles [76] and/or polarizable atoms, [37, 77] have previously been used to predict correctly the contrast on a nanometer scale. To get atomic resolution for more sophisticated surfaces, more realistic models and atomistic calculations are required. In a joint experimental-theoretical study, Sadewasser *et al.* [36] showed that the changes in the surface dipole moment and consequently in the chemical potential induced by bond formation upon the approach of a nano-tip towards a semiconducting surface appear in the same tip-sample separations where the atomic scale contrast in the KPFM signal is detected. The reason is that the CPD depends on a variety of parameters, including on the work function which is shifted due to the surface dipole. [114] Masago *et al.* [68] defined *quantitatively* the LCPD as the difference between the Fermi levels of the tip and sample subsystems in a combined system. This definition is essentially applicable for large separations where the overlap of the wave functions of the tip and surface atoms is negligible. Using a perturbation approach, it was also applied to smaller distances about 4 Å where the chemical interactions matter. [69]

Almost all proposed quantitative models [115, 37, 71, 68, 69, 4, 5, 86] are commonly based on the experimental definition of the LCPD, namely finding a macroscopic bias which minimizes the bias-dependent tip-sample interactions for each tip position. In all of the mentioned references, this interaction is expressed as a quadratic function of the bias. The extremum of this parabola deviates from the macroscopic V_{CPD} and this deviation which is determined by the ratio of the coefficients of the linear to the quadratic terms, as seen in Eq. (5.3). Different authors have, however, tried to obtain this ratio from completely different approaches. In chapter 4, we critically discussed the shortcomings of some approaches, and presented a multiscale framework

which takes into account all contributions, macroscopic and atomistic. Here, we introduce some simplifications to obtain general expressions for the signals detected in the KPFM technique in terms of induced changes in the dipole moment. Short- and intermediate-range contributions are split. Based on this, we provide some microscopic insights into the problem of the atomic origin of the LCPD contrast. In particular, it turns out that the distance dependence of the LCPD contrast follows the dependence of the dipole moment. Comparable results are obtained for the same setup used in the previous chapter, i.e. a realistic silicon nano-scale tip interacting with a NaCl(001) sample treated by density functional theory (DFT).

5.1 Multiscale definition of LCPD

When the electrostatic force is to be compensated, the bias voltage V_{DC} is close to the CPD and thus the effective potential difference $V = V_{\mathrm{DC}} - V_{\mathrm{CPD}}$ is small, and the short range force can therefore be approximated to the second order by

$$F_\mu(\mathbf{R}; V) - F_\mu(\mathbf{R}; V = 0) \simeq a(\mathbf{R})V + b(\mathbf{R})V^2 + \cdots . \tag{5.1}$$

This in addition to the macroscopic capacitive force $F_M = \frac{1}{2}C'V^2$ gives the total V-dependent force[1]

$$F(\mathbf{R}, V) = a(\mathbf{R})V + \Big[\frac{1}{2}C'(s) + b(\mathbf{R})\Big]V^2, \tag{5.2}$$

where $C' = \partial_s C < 0$ is the capacitance gradient, s equals z plus a constant and $\mathbf{R} \equiv (x, y, z)$ is nominal position of the tip-apex with respect to surface (see Fig. 4.1).

The vertex of the parabola in Eq. (5.2) occurs at

$$V_{\mathrm{DC}} = -\frac{a}{2b + C'} + V_{\mathrm{CPD}}$$

rather than at V_{CPD}. In the AM mode with ultra-small amplitude oscillations (where the compensated signal is proportional to Eq. (5.2), see chapter 2) the position of the vertex is operationally obtained either by sweeping V_{DC} or via a feed-back circuit, and recorded as the LCPD map of the scanned surface. In this case, the local deviation of LCPD from the macroscopic CPD,

$$L \equiv V_{\mathrm{LCPD}} - V_{\mathrm{CPD}},$$

at some scan point $\mathbf{R}$ is

$$L^{\mathrm{AM}}(\mathbf{R}) = -\frac{a}{2b + C'}. \tag{5.3}$$

[1] Although we showed in chapter 4 that F_μ is almost linear when V changes in the range of few Volts, we keep the quadratic term for the sake of completeness. Finally, it will be shown that the microscopic contribution to the quadratic term originated from $b(\mathbf{R})$ is dominated by the macroscopic counterpart.

The deviation is clearly both site- and capacitance dependent.

In the FM mode, on the other hand, where the minimized signal is the frequency shift and thus proportional to the gradient of Eq. (5.2), one sees that the vertex occurs as

$$V_{\mathrm{DC}} = -\frac{a'}{2b' + C''} + V_{\mathrm{CPD}},$$

resulting in

$$L^{\mathrm{FM}}(\mathbf{R}) = -\frac{a'}{2b' + C''}. \tag{5.4}$$

Owing to their atomic origin, a and b decays beyond the short-range interactions while for intermediate separations L approaches a constant value. Apart from b and b', expressions obtained for L are identical to those in chapter 4. For finite-amplitude oscillations, averaging over the oscillation cycles is performed as explained in that chapter. However, the amplitude should be of order of 1 Å or less to achieve atomic resolution.

5.1.1 Computational details

Our test *microsystem* consists of a model tip facing a flat slab illustrated in Fig. 5.1 and described in more detail in chapter 4. The model tip is a 29-atom Si cluster while the (001) surface of the NaCl sample is modeled with a two-layer slab, containing in total 200 ions. Periodic boundary conditions are applied along the lateral directions while free boundary conditions are applied in the vertical direction along which the electric field is also applied. The bottom layer of the sample and the base (top) layer of the tip are frozen while other atoms are fully relaxed upon approaching the tip to the sample in steps of 0.2 Å. The nominal tip-sample separation s is defined as the distance between the tip apex atom and the sample surface ignoring the relaxations induced by mutual interactions, and is therefore equal to the distance between the frozen layers minus a constant.

As justified in chapter 3, the uniform vertical electric field E_z imposed on the microsystem for a specified macroscopic bias voltage V is obtained from the solution of the classical electrostatic problem for a realistic *macroscopic* AFM probe-sample model. The macroscopic sample is a 1-mm thick dielectric slab ($\epsilon_r = 5.9$ for NaCl) grounded at the bottom. The probe is a 15 μm high conducting cone with an opening angle of 15° terminated by a spherical cap of radius 20 nm. The cone is attached to a cantilever approximated by a disk of radius 10 μm.

DFT calculations were performed within the local density approximation (LDA) using norm-conserving HGH pseudopotentials [102] to represent the effect of the atomic cores on the self-consistently determined *valence electron density* n. The BigDFT package, which relies on a wavelet basis set to expand wave-functions on an adjustable grid in real space [30], was employed

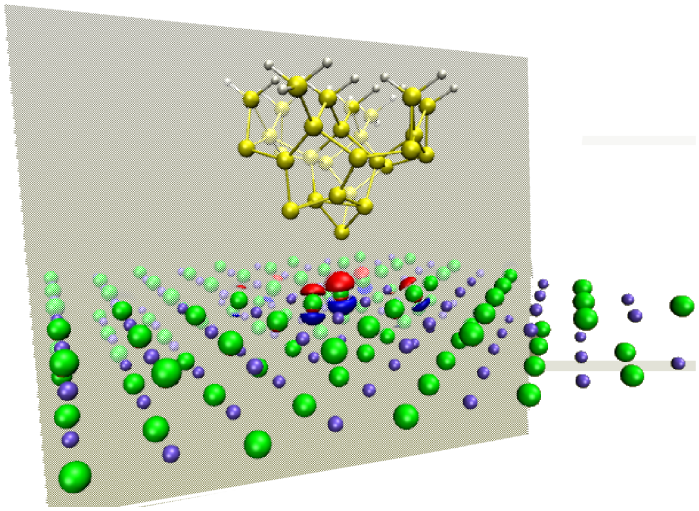

Figure 5.1: Perspective view of the microsystem used in the DFT calculations showing Cl (green), Na (violet), Si (yellow) and H (white) atoms. Red (lighter gray) and blue (darker gray) isosurfaces of $\Delta n = \pm 0.002$ e/Å^3 depict, respectively, electron excess and depletion generated due to tip-sample chemical interactions at a nominal separation 3.2 Å from a Cl ion. The plane of the 2D maps in the following figures is shown in gray and is perpendicular to the surface and passes through a row of alternating Na and Cl ions and the tip apex.

to treat the combined as well as the isolated tip and sample. In order to obtain the *difference density*

$$\Delta n = n - n_0^{\text{tip}} - n_0^{\text{sample}} \tag{5.5}$$

due to wave-function overlap and to the relaxation of the atomic cores in response to Hellman-Feynman forces, computations for each separation were performed in the same grid for the combined and the isolated subsystems. Contrary to common practice, the densities n_0^{tip} and n_0^{sample} of the isolated (but relaxed separately) tip and sample were subtracted from the total density n computed at the relaxed positions of the atomic cores in the *interacting* microsystem. Similarly, $n(\mathbf{E})$ is computed at core positions shifted by the imposed electric field $\mathbf{E}$, and a fixed computational grid was used when computing the difference

$$n_{\text{ind}} = n(\mathbf{E}) - n(\mathbf{E} = 0) \tag{5.6}$$

induced by the electric field.

On the sample side, the valence electron density (which integrates to eight electrons per NaCl pair) is almost entirely concentrated around the closed-shell Cl anions. As a consequence, the most pronounced differences Δn or n_{ind} have the appearance of the $3p_z$ orbitals on Cl ions

with their electron excess lobes pointing along the displacement vectors of the corresponding Cl^{+7} cores in the surface layer. Figure 5.1 shows the atomic arrangement of the microsystem together with isosurfaces of Δn. The mentioned orbitals appear on a few anions under the tip pointing towards the dangling bond on the foremost Si atom.

5.1.2 Microsystem and charge density

Consider the microscopic region of the contact point containing all the atoms of tip and all the atoms of the sample that contribute to the short range interactions.[2] The charge density

$$\rho \equiv -n(\mathbf{r}) + \sum_k q_k \delta(\mathbf{r} - \mathbf{r}_k), \tag{5.7}$$

where n is the electron density while q_k denotes the core charge of atom k at position $\mathbf{r}_k$, has three contributions

$$\rho(\mathbf{r}; \mathbf{R}, \mathbf{E}) = \rho_0(\mathbf{r}) + \Delta\rho(\mathbf{r}; \mathbf{R}) + \rho_{\mathrm{ind}}(\mathbf{r}; \mathbf{R}, \mathbf{E}). \tag{5.8}$$

Here

$$\rho_0(\mathbf{r}) \equiv \rho_0^{\mathrm{tip}}(\mathbf{r}) + \rho_0^{\mathrm{sample}}(\mathbf{r})$$

is the charge density of a *fictitious prototype system* composed of the *non-interacting* nano-tip and nano-sample with charge densities ρ_0^{tip} and ρ_0^{sample}, respectively. Due to chemical interaction of the nano-tip and nano-sample, a charge density $\Delta\rho(\mathbf{r}; \mathbf{R})$ has been induced in the system as a function of tip position $\mathbf{R}$. If the system is subject to an external electric field $\mathbf{E}(\mathbf{r})$ (proportional to the bias voltage), a charge density is also induced which is approximated to the first order by

$$\rho_{\mathrm{ind}}(\mathbf{r}; \mathbf{R}, \mathbf{E}) = \boldsymbol{\gamma} \cdot \mathbf{E}.$$

The vector field

$$\boldsymbol{\gamma}(\mathbf{r}; \mathbf{R}) \equiv \frac{\partial \rho}{\partial \mathbf{E}}\Big|_{\mathbf{E}=0} \tag{5.9}$$

describes the *local* electric polarizability as a function of tip position $\mathbf{R}$.

The electric field developed by the macroscopic bodies inside the microscopic region connects the two length scales. This multi-scale feature is absent, for instance in the DFT-based tight binding simulation method for KPFM [68]. In this work a $\mathbf{R}$-dependent charge distribution similar to Eq. (5.8) is used but the external electric field due to the macroscopic parts is missing.

[2] V-independent forces, e.g. vdW interactions, do not enter into the play in this context.

Chemically exited changes in electron density and atom core positions

Figure 5.2 shows 2D maps of the total valence electron densities n (left) and the difference Δn (center), as well as integral of Δn (right) on planes parallel to the surface for the tip apex facing a Na (top row) and a Cl (bottom row) surface sites at a short representative separation, i.e. 3.2 Å.

For this separation, the microscopic force on the tip is close to maximum attraction above a Na surface site, but slightly repulsive above a Cl surface site (see Fig. 4.3) in consistence with the shown displacement patterns of the relaxed cores. At this close separation, density changes due to incipient bonding of the tip apex atom to the nearest Cl anions are visible in the n maps, but are quite prominent in Δn. This strong change is responsible for the decrease of the system dipole moment upon tip approach above a Cl site apparent in Fig. 5.4(b). The contribution of each Cl $3p_z$-like polarization to the dipole moment is to large extent cancelled by the displacement of each Cl^{7+} core. Nevertheless, DFT results show that a net contribution about a factor seven smaller is still left. Above the Na site, displacements of surface ions just aside the tip apex atom reveal that they have been pushed away by non-electrostatic (Pauli) repulsion. This also happens when the tip apex is above a Cl site for adjacent Cl anions along $\langle 110 \rangle$ directions which are not visible in the cutting plane. Less pronounced and more distorted Si $3p_z$-like polarization clouds appear around under-coordinated Si atoms at the apex and on the sides of the model tip. In the apex vicinity, electron accumulation above the Na site switches to electron deficiency above the Cl site. Further away, the Si^{4+} core displacements and the concomitant density changes have appreciable components orthogonal to the surface normal, hence contribute less to the total dipole moment than the apex region.

Electrically induced changes in electron density and atom core positions

Figure 5.3 shows 2D maps and contours of n_{ind} induced by an uniform external electric field above a Na (left) and a Cl (right) surface sites. The maps clearly reveal the roughly three times stronger electron accumulation lobes pointing towards the tip apex on the nearest Cl anions compared to Na cations also seen in Fig. 5.1. In contrast to Fig. 5.2, more distant anions are also polarized. Even those in the fixed bottom layer are polarized, albeit less by almost an order of magnitude compared to those in the top layer, which show a pattern of core displacements consistent with the field direction. In a thicker slab, internal layers would be even less polarized owing to dielectric screening, which is however, underestimated by the LDA functional adopted here.

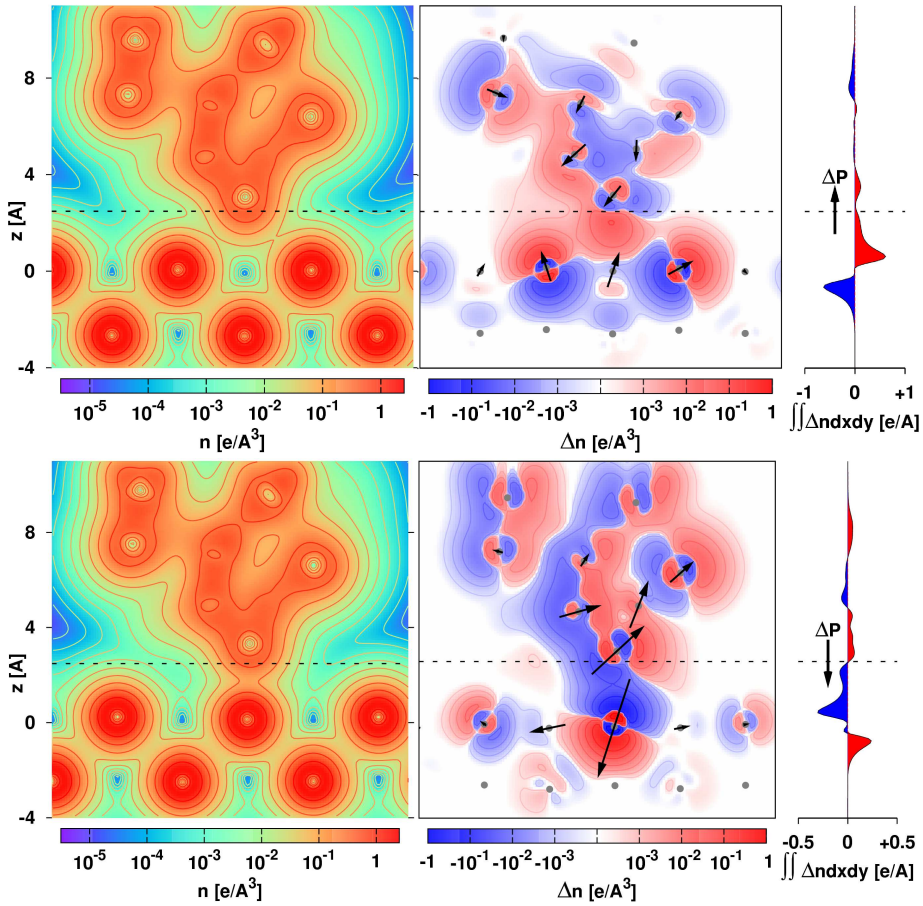

Figure 5.2: 2D maps of the valence electron density n (left) and its deviation $\Delta n = n - n_0$ from the superposed densities n_0 of the isolated subsystems (center) and plane integrals $\int \Delta n dx dy$ as a function of z (right) at a nominal separation of 3.2 Å of the tip apex Si atom from a Na (top) and Cl (bottom) ion. The map plane is normal to the surface and passes through the foremost atom of the tip as shown in Fig. 5.1. Δp illustrates the induced dipole moment. The horizontal dashed lines indicate z with the minimum plane-integrated Δn. In the difference maps dots show the positions of atomic cores, while core displacements induced upon tip approach are depicted by centered arrows magnified by suitable factors for better visualization. Contours of constant n and Δn values differing by a constant factor $\sqrt{10}$ allow one to recognize regions where those quantities decay exponentially. The region with $|\Delta n| < 10^{-4}$ e/Å^3 is white, while red and blue regions depict electron excess and depletion, respectively; the two first contour sets correspond to $\Delta n = \pm 10^{-3.5}, \pm 10^{-3}$ e/Å^3.

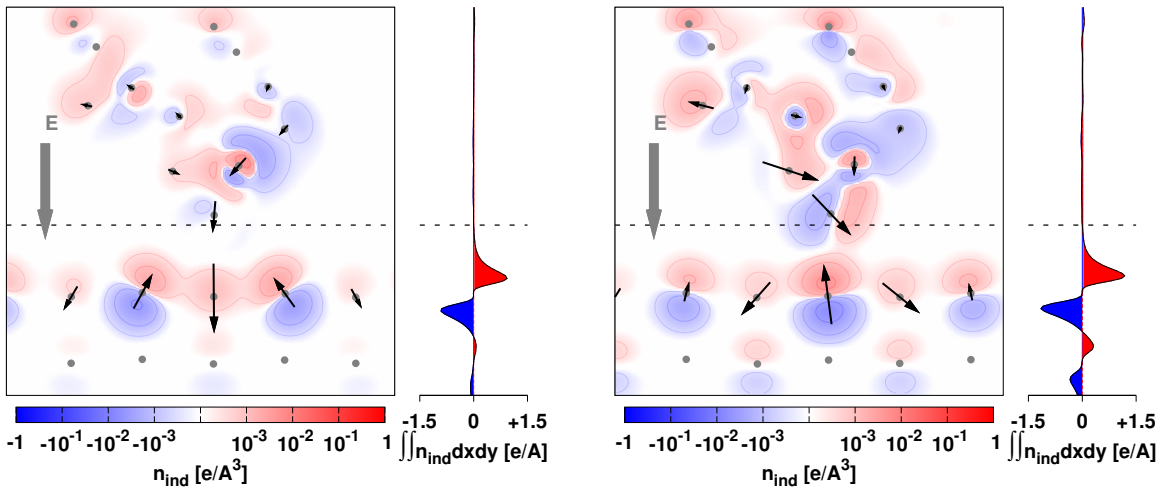

Figure 5.3: 2D maps of the induced electron density n_{ind} by uniform external electric field (gray arrows) corresponding to a macroscopic tip bias of 2 Volts for a nominal separation of 3.2 Å (for which $E_z = 0.316$ V/nm) above a Na (left) and a Cl (right) surface site. Side curves show plane integrals $\int n_{\text{ind}} dxdy$ as a function of z. The horizontal dashed lines indicate z with the minimum plane-integrated n_{ind}. Dots show the positions of atomic cores, while centered arrows scaled by a factor of 50 show their field-induced displacements. The region with $|\Delta n| < 10^{-4}$ e/Å^3 is white, while red and blue regions depict electron excess and depletion, respectively. The two first contour sets correspond to $\Delta n = \pm 10^{-3.5}$ and $\pm 10^{-3}$ e/Å^3.

5.1.3 Dipole Moment

The dipole moment of the system changes in response to chemical interactions and electric field in the same way as the charge density

$$\begin{aligned} \mathbf{p} &\equiv \int \mathbf{r}\rho(\mathbf{r})d\mathbf{r} \\ &= \mathbf{p}_0 + \Delta\mathbf{p}(\mathbf{R}) + \alpha\mathbf{E}. \end{aligned} \tag{5.10}$$

$\mathbf{p}_0 \equiv \int \mathbf{r}\rho_0(\mathbf{r})d\mathbf{r} = \mathbf{p}_0^{\text{tip}} + \mathbf{p}_0^{\text{sample}}$ is the permanent dipole moment of the prototype system[3] and $\Delta\mathbf{p}(\mathbf{R}) = -\int \mathbf{r}\Delta n(\mathbf{r})d\mathbf{r} + \sum_k q_k \Delta\mathbf{r}_k$ arises from chemical interactions. The electrically induced dipole moment is $\alpha\mathbf{E}$ where the elements of the polarizability tensor are

$$\alpha_{ij}(\mathbf{R}) \equiv \frac{\partial p_i}{\partial E_j}\Big|_{\mathbf{E}=0} = \int \mathbf{r}_i \boldsymbol{\gamma}_j d\mathbf{r} \quad ; \quad i,j \in \{x,y,z\} \tag{5.11}$$

where $\boldsymbol{\gamma}$ is defined by Eq. (5.9).

[3] If the Fermi levels are not aligned in the isolated tip and sample subsystems, an offset correction is also required.

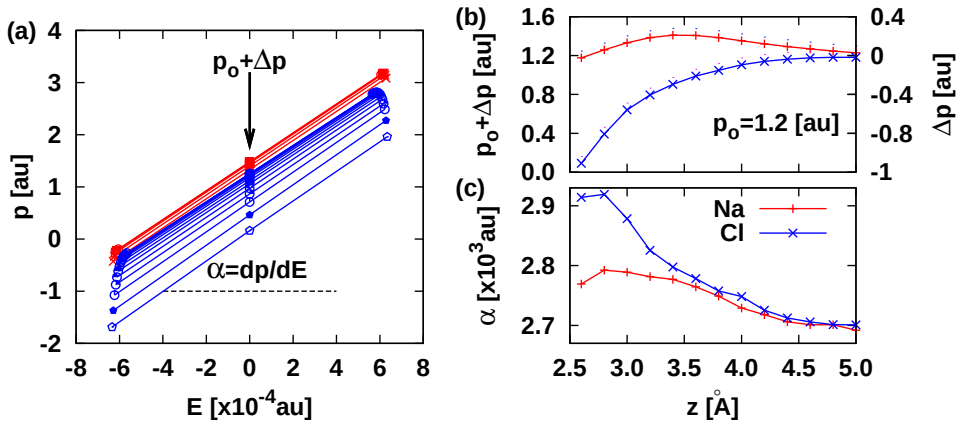

Figure 5.4: (a) Calculated electric dipole moment normal to sample surface as a function of electric field (note that 10^{-4} au ≈ 0.0514 V/nm). Symbols show the calculated values at field strengths corresponding to $V = 0, \pm 2$ Volts, while lines are drawn to help eye. Red color is used for the tip on top of a Na ion and blue for on Cl. Each line corresponds to one tip position which could easier be followed from (b). (b) $p_0 + \Delta p = p(E = 0)$ as a function of tip-height, calculated as the intercept of lines in (a), where $p_0 = p_0^{\text{tip}} + p_0^{\text{sample}}$ and Δp is the interaction-induced dipole moment. 1 au ≈ 2.54 Debye. (c) Polarizability $\alpha = \partial p / \partial E$, calculated as the slope of lines in (a). 10^3 au $= 16.5$ pN.nm^3/V^2.

The dipole moment in the z direction for the model system calculated with DFT is shown in Fig. 5.4(a). The moment is almost linear versus the electric field strength when the bias voltage applied to macroscopic electrodes varies from -2 to +2 V. First of all, this linearity verifies our assumption that higher than the first order terms which are already omitted Eqs. (5.8) and (5.10) are not significant.

The intercepts of these lines gives the dipole moment at zero bias (i.e. $p_z(E_z = 0)$) for different tip positions, as plotted in Fig. 5.4(b) as a function of separation from Na and Cl sites. Indeed, this characterizes the chemically excited dipole moment which strongly depends on tip position.

On the other hand, the slopes of the lines in Fig. 5.4(a) give the polarizability α, as plotted in Fig. 5.4(c), which depends only weakly on the tip position. In this case α varies within 10%, and hence we ignore the $\mathbf{R}$-dependence of α.

5.1.4 Energy

The electrostatic energy of the microscopic system subject to an external electric field $\mathbf{E} = -\nabla\phi_{\text{ext}}$ is

$$u_\mu = \frac{1}{2}\iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + \int \rho(\mathbf{r})\phi_{\text{ext}}(\mathbf{r})d\mathbf{r}. \tag{5.12}$$

Using Eq. (5.8), terms with different V-dependencies are split:

$$\begin{aligned} u_\mu &= \frac{1}{2}\iint \frac{[\rho_0(\mathbf{r})+\Delta\rho(\mathbf{r})][\rho_0(\mathbf{r}')+\Delta\rho(\mathbf{r}')]}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' \\ &+ \int \big(\rho_0+\Delta\rho\big)\underbrace{\big(\phi_{\text{ext}}+\phi_{\text{ind}}\big)}_{\propto V}d\mathbf{r} + \int \underbrace{\rho_{\text{ind}}\big(\phi_{\text{ext}}+\frac{1}{2}\phi_{\text{ind}}\big)}_{\propto V^2}d\mathbf{r}. \end{aligned} \tag{5.13}$$

Finally, we obtain the quantities required to evaluate Eq. (5.3), i.e. a and b. Assuming that $\phi_{\text{ind}} = \int \frac{\rho_{\text{ind}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'$ modulates ϕ_{ext} only slightly, recalling that $F_\mu = -\partial_s u_\mu$ and using Eqs. (5.1) and (5.13), we obtain

$$a(\mathbf{R}) = -\frac{1}{V}\frac{\partial}{\partial s}\int(\rho_0(\mathbf{r})+\Delta\rho(\mathbf{r}))\phi_{\text{ext}}(\mathbf{r})d\mathbf{r}, \tag{5.14}$$

$$b(\mathbf{R}) = -\frac{1}{V^2}\frac{\partial}{\partial s}\int\rho_{\text{ind}}(\mathbf{r})\phi_{\text{ext}}(\mathbf{r})d\mathbf{r}. \tag{5.15}$$

Therefore the linear V-dependence is merely due to the interaction of the intrinsic (i.e. noninduced) charge density with the external electric field of the macroscopic tip, while the quadratic term is caused by the induced charges with this field. If one considers the involved sizes, and as will also been shown later on, the latter is much smaller than the macroscopic capacitive interactions.

5.1.5 Approximate expressions for local deviation from CPD

Atomic resolution is achieved only at small separations. In this case, the electric field underneath the tip-apex is nearly uniform and parallel to the tip axis. This means that ϕ_{ext} can be simply approximated as a linearly decreasing function from V on the tip surface towards the sample surface (our interested region is the contact point). The integrals in Eqs. (5.14) and (5.15) can then be expressed in term of dipole moments:

$$a \simeq -\frac{1}{V}\frac{\partial}{\partial s}\Big[(p_0+\Delta p)E\Big], \tag{5.16}$$

$$b \simeq -\frac{1}{V^2}\frac{\partial}{\partial s}(p_{\text{ind}}E). \tag{5.17}$$

$p_0 + \Delta p = p(\mathbf{E} = 0)$ is the dipole moment at zero bias while p_{ind} is the induced dipole moment due to applying an electric field $\mathbf{E}$. We emphasize that a and b are V-independent, because $p_{\text{ind}}, E \propto V$. Therefore, for convenience, we introduce a purely geometrical function

$$\mathcal{E} = \frac{E}{V}$$

which is $\mathcal{E}(s) = 1/s$ in the simplest case, i.e. for a parallel-plane capacitor. In general, $\mathcal{E}$ is the solution of the Poisson's equation[4] for the given macroscopic tip-sample system, as plotted in Fig. 3.15 for our test system as well as for a spherical tip above a dielectric. Recalling that $p_{\text{ind}} = \alpha E$, where $\alpha = \alpha_{zz}$ is the polarizability of the system, we have

$$a \simeq -\partial_s(p_0\mathcal{E}) - \partial_s(\Delta p\mathcal{E}), \tag{5.18}$$

$$b \simeq -\partial_s(\alpha\mathcal{E}^2). \tag{5.19}$$

5.2 Discussion

5.2.1 Linear terms: short- and intermediate range contributions

The two terms in Equation (5.18) can be considered as intermediate- and short range terms, respectively. Accordingly, it is useful to have the corresponding contributions to L separated as

$$V_{\text{LCPD}}^{\text{AM/FM}} = \text{L}_{\text{ir}}^{\text{AM/FM}} + \text{L}_{\text{sr}}^{\text{AM/FM}} + V_{\text{CPD}}. \tag{5.20}$$

This formula is valid for an ultrasmall amplitude oscillation; extension to finite amplitude is straightforward. The intermediate-range term L_{ir} is caused by the interaction of p_0 with the external field, while the short-range term L_{sr} by the interaction of the chemically induced dipole Δp_0 with the field.

Intermediate range shift: L_{ir}

The permanent dipole moment p_0 is not necessarily due a dipole, but also exists if a nonzero net charge, either extended or point-like, exists for which $\mathbf{p}_0 = \int \mathbf{r}\rho_0 d\mathbf{r}$ (e.g. for a net point charge q it gives $\mathbf{p}_0 = q\mathbf{r}_q$, Caused either by fixed point charges or dipoles, the shift introduced to the background CPD is given by

$$\text{L}_{\text{ir}}^{\text{AM}} = \frac{\partial_s(p_0\mathcal{E})}{2b + C'}, \quad \text{L}_{\text{ir}}^{\text{FM}} = \frac{\partial_s^2(p_0\mathcal{E})}{2b' + C''}. \tag{5.21}$$

First of all, the shift depends strongly on the macroscopic parts via $\mathcal{E}$ and C. The shift also depends on whether the charge/dipole is fixed on the sample surface or carried by the tip.

[4] The macroscopic capacitance $C(s)$ is also obtained by solving the Poisson's equation.

If the charge/dipole is carried by the tip the shift is site-independent[5] and hence does not feel the electric field variations during the tip lateral motion, but it is distance-dependent because it feels the change in the field due to vertical motion of tip. Such a shift which is determined by the macroscopic parameters via $\mathcal{E}$ and C, both should be noticed when the LCPD maps are interpreted and could be exploited to characterize a charged or polarized tip.[6] [76]

In contrast, if the charge/dipole is on the surface ($\partial_s p_0 = 0$), the intermediate shift changes with the tip lateral position. The imaging resolution of such an object is strictly determined by the tip geometry and its separation from the surface. If the tip is sharp and close, the surface charge/dipole senses well the variation of $\mathcal{E}$ developed by the tip during its lateral motion. The size of the spot appearing in the LCPD map is therefore of the same order as the effective width of the field. In 3.1.1 we showed that the latter, if considered to be the field half-width, is considerably smaller than the tip radius. Even if the spot size is an order of magnitude smaller than the tip radius id is still somehow larger than the atomic resolution.

Short range shift: $L_{\mathbf{sr}}$

The chemical interaction (and bond formation) between the foremost atoms of the tip and the sample change the dipole moment as Δp. This short range term, which is also affected by the electric field and macroscopic capacitance, is responsible for the atomic-scale contrast. Again, a sharp tip over a conducting surface generates rapidly spreading field lines and the contrast in the LCPD is therefore enhanced. The short range shifting away from $L_{\mathrm{LR}} + V_{\mathrm{CPD}}$ is given by

$$\mathrm{L}_{\mathrm{sr}}^{\mathrm{AM}} = \frac{\partial_s(\Delta p \mathcal{E})}{2b + C'}, \quad \mathrm{L}_{\mathrm{sr}}^{\mathrm{FM}} = \frac{\partial_s^2(\Delta p \mathcal{E})}{2b' + C''}. \tag{5.22}$$

5.2.2 Capacitive terms: microscopic versus macroscopic contributions

Equation (5.19) can be rewritten as $b \simeq -2\alpha\mathcal{E}\mathcal{E}'$ provided that the polarizability is constant, as seen in Fig. 5.4(c) for our test system for which it is about 45 pN.nm^3/V^2. From the inset of Fig. 3.15 for a realistic macroscopic probe-sample we know $\mathcal{E} \simeq 0.15$ nm^{-1} and $\mathcal{E}' \sim -0.049$ nm^{-2}. Equation (5.19) then gives $b \simeq -2\alpha\mathcal{E}\mathcal{E}' = 0.66$ pN/V^2, namely more than two orders of magnitude smaller than $-C' \simeq 135 - 140$ pN/V^2 at the same separations (see Fig. 4.5(b)). We conclude that $2b + C' \approx C'$. Similarly, for the FM mode where one needs $2b' + C''$, we see that for our test system $C'' \simeq 15 - 20$ mN/m.V^2 (see Fig. 4.5(c)) dominates $b' \simeq 0.87$ mN/m.V^2 by more than an order of magnitude. Therefore, both in AM and FM modes, the macroscopic quadratic term

[5] For a point charge q fixed to the tip $\partial p_0/\partial s = q$, because $\partial z_q/\partial s = 1$. Then $a = q\mathcal{E} + qz_q\mathcal{E}'$ is site-independent but s-dependent.

[6] Equation (5.21) becomes identical to Eqs. (1) and (2) in Ref. [76] where polarizability and chemical interactions are absent ($\alpha, \Delta p = 0$), $\mathcal{E} = 1/s$ and conducting sphere-plane approximation is used.

well dominates its microscopic counterpart. This result is in contrast to the finding/assumption of Refs. [37, 68, 69]. For materials different from our test case, α could be larger than that of our Si-NaCl system, but it is unlikely that it can compensate this two-order of magnitude difference. However, to make it applicable to potential problems where this is not the case, we still keep both quadratic terms in our formalism. Even if $b(\mathbf{R})$ (or its gradient in the FM mode) is very large, since α is almost constant, the *atomic-scale variation* of the LCPD is still essentially determined only by $a(\mathbf{R})$.

In fact, the last term in Eq. (5.13) can be considered as the energy of a *microscopic capacitor* made of the polarizable atomic contents in the gap of the *macroscopic tip-sample capacitor*. The capacitive energy is quadratic versus the bias and is comes from the interaction between the electric field with the charges it induces on the macroscopic bodies and microscopic contact region. The much larger physical size of the former suggests that it dominates the latter. However, this is not necessarily also valid for the vertical gradients.

5.3 Case Study

5.3.1 NaCl(001) surface

The characteristic variation of Δp when $s < 5$ Å in Fig. 5.4(b) suggests that the atomistic origin of the LCPD contrast comes from the chemically induced dipole moments. We plot in Fig. 5.5

$$L^{\mathrm{AM}} = \frac{(p\mathcal{E})'}{C'},$$

where the dipole moment p is calculated when $E = 0$. Evaluating the latter expression is much easier and less computationally demanding compared to minimizing $F_M + F_\mu$ as explained in chapter 4 (see Eq. (4.9) and plotted in Fig. 4.6(a) (it is also plotted here for comparison). All other panels of Fig. 4.6 could be reproduced with such a very good agreement because the basic quantity i.e. the slope a is already approximated very well using Eq. (5.18).

Figure 5.5 resembles Fig. 5.4(b) with a very similar general trend. In other words, tracking the variation of $\partial_s \Delta p$ with tip position can already provide a qualitative picture of the LCPD variation. But to obtain a quantitative value, it should be combined with $\mathcal{E}$ and C' in the proposed way.

5.3.2 Copper phthalocyanine (CuPc) molecule

As the next case study, we compare the calculated LCPD with the experimental result for a copper phthalocyanine (CuPc) molecule. The molecule is adsorbed either on the Cu(111) surface or on a two-layer NaCl film coating it. These are modeled in simulation as slabs of four Cu

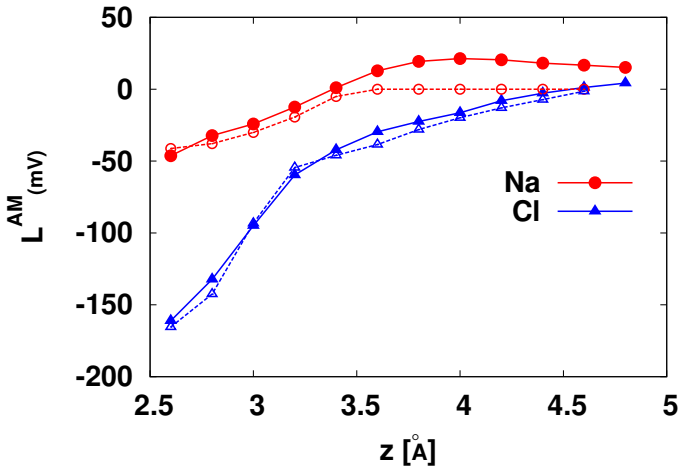

Figure 5.5: Shift in CPD calculated from $L^{\mathrm{AM}} = (p\mathcal{E})'/C'$ (filled symbols) as a function of tip-height from Na and Cl sites on NaCl(001) surface. The same quantity calculated using Eq. (4.9) and plotted in Fig. 4.6(a) is also shown by blank symbols for comparison. Since C' is identical in both cases, the excellent agreement demonstrates that the basic quantity a is approximated very well by Eq. (5.18); Consequently, all panels of Figs. 4.6 and 4.7 are reproduced (not shown) using this approximated a.

layers and two NaCl layers, respectively. The tip apex is modelled with a 5-atom pyramid Cu cluster (see Fig. 5.6) mimicking a tip which has picked up Cu atoms. Tip is laterally moved along lines parallel to the symmetry axes of the molecule at a constant distance of 2.0 Å from the molecular plane. Since the relaxed geometry of the molecule on either surface is curved with lobes approaching the surface, the tip height is measured from the central Cu atom of the molecule. The experiment [116] was performed in the FM mode with an ultrasamll amplitude oscillations (tip vertical moment is always witnin 0.2 Å) and hence we do not need to average over a cycle; see 4.3.3.

The shift from CPD is calculated by

$$L^{\mathrm{FM}} = \frac{(p\mathcal{E})''}{C''}.$$

Finite difference is used to calculate vertical gradients from the values on two adjacent separations ± 0.1 Å away. Equation (3.17) was used[7] with the appropriate approximation for thin films as explained in 3.3.1, namely $s \to s + h/\epsilon_r$ where $\epsilon_r = 5.9$ for two-monolayer NaCl film of thickness $h = 0.56$ nm. We used $R = 20$ nm as the tip radius and $s = 0.7$ nm for its separation from the substrate.

[7] Within the approximation that Hudelt *et al.* [25] used, $C'' = 2\pi\epsilon_0 R[1/s^2 - 1/(R+s)^2]$ and $\mathcal{E} = 1/s$.

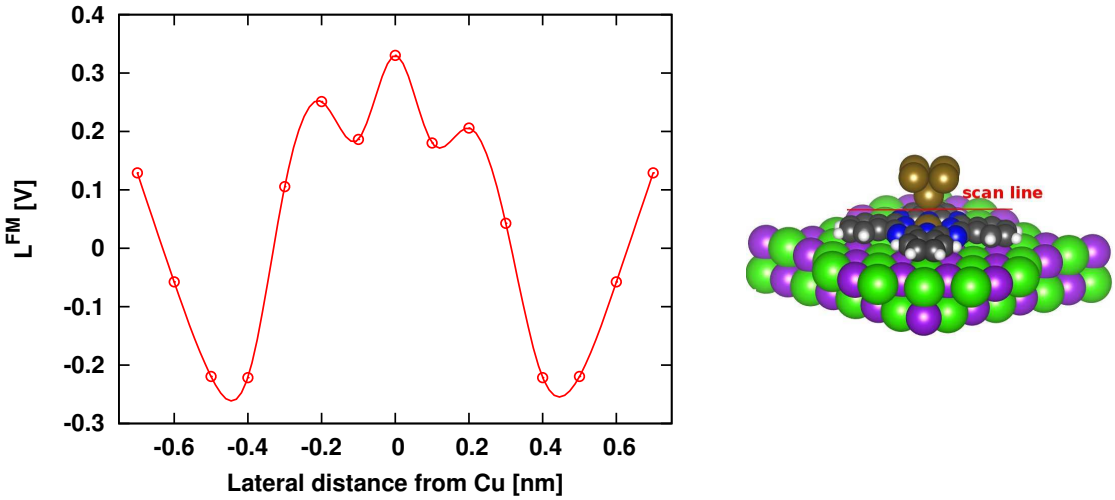

Figure 5.6: Shift in CPD calculated from $L^{\mathrm{FM}} = (p\mathcal{E})''/C''$ above a CuPc molecule absorbed on NaCl two-layer film. The tip is modeled by th pyramid Cu_5 cluster and is moved along a line parallel to the symmetry axis of the molecule as indicated in the side view shown on the right.

The calculated shift from CPD, L^{FM}, for the CuPc on the NaCl layers is plotted in Fig. 5.6. As in the experiment, it reveals a peak on top of the central Cu atom and two valleys aside, showing an excellent agreement within sub-nm lateral resolution.

For the case of CuPc adsorbed on the Cu(111) surface, L^{FM} is shown in Fig. 5.7. The relaxed geometry of the molecule over the surface, shown in Figure, no longer shows fourfold symmetry. Two wings of the molecule get aligned with the $[11\bar{2}]$ direction of the Cu(111) surface such that the center of their hexagons, like the central Cu atom, lie on the bridge sites of the surface. The two other are aligned with the $[1\bar{1}0]$ direction and the hexagons are located on the surface hollow sites. Note that one of the hollow sites is fcc while the other is hcp. Such asymmetry, reflected also in imaged LCPD, is well seen in the calculated LCPD plots. The range of the variation of the LCPD throughout the molecule is a few times smaller compared to the NaCl case, in agreement to the experiment.

Conclusions

In summary, the atomic-scale deviation from the macroscopic CPD was studied by splitting the electrostatic energy into contributions based on their dependences on the bias voltage and on their short- or intermediate range character. The curvature of the force-bias parabola is determined by the bias-induced charges with the electric field E caused by the biased tip, whereas the linear term which shifts the vertex and thus the LCPD is mainly determined by the interaction of

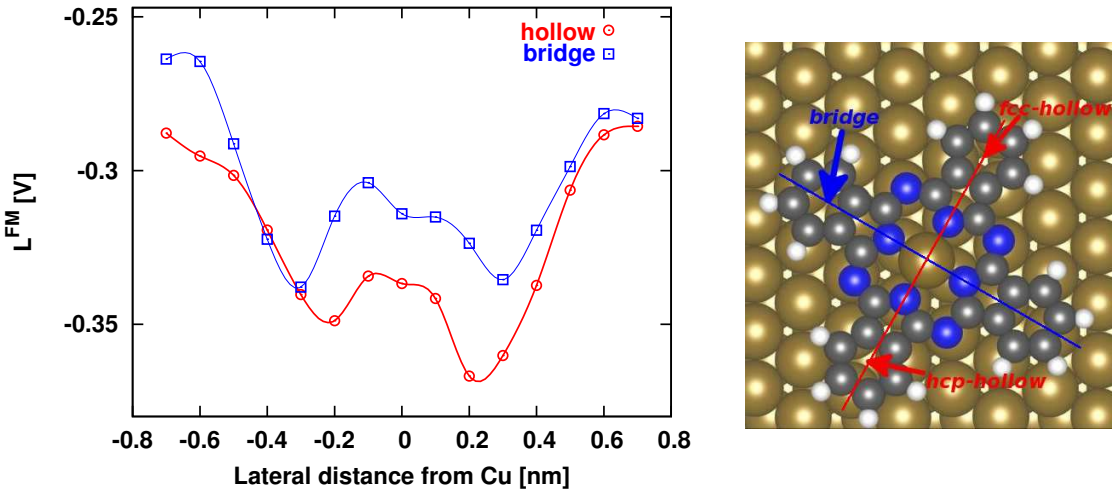

Figure 5.7: Same as Fig. 5.6 but for the CuPc molecule absorbed on Cu(111) surface. The scan lines are parallel to the molecular axes as shown on the top view of the relaxed geometry on the right.

chemically excited charges with the field. The latter is simply given by pE where p is the total dipole moment in absence of E. We showed how the LCPD image, which is believed to be the local contact potential difference between the tip and the probed sample, is a related to both the charge distribution at the contact point and the tip-sample geometry. The values predicted by the proposed simple expressions show excellent qualitative agreement with experiment.

Chapter 6

Metrics for measuring distances in configuration spaces

Quantifying dissimilarities between molecular structures is an essential problem encountered in physics and chemistry. Comparisons based on structural data obtained either from experiments or computer simulations can help identifying or synthesising new molecules and crystals. A broad diversity of structures can only be obtained if identical configurations are eliminated. It is therefore highly desirable to have numerically affordable fingerprints that allow in a reliable way to detect identical configurations in the presence of noise which can either arise from experimental measurements or from structural relaxations in numerical simulations. Maintaining a broad diversity of structures is also a prerequisite for efficiency in any structure prediction method in material science and solid state physics [117, 118, 119, 120, 121] and conformer search in structural biology and drug discovery. [122, 123, 124, 125, 126, 127, 128] In the latter case, most of the proposed approaches [129, 130, 131, 132] use approximate methods that reduce the structure description information, e.g. by excluding the side chains in a protein or a two dimensional representations of the molecule, [133] to speed up the searching procedure. [134] In the case of solid state physics, fairly accurate dissimilarity measures are required. Within the structure prediction methods based on the evolutionary algorithms, [117] the required diversity of populations can only be maintained if strongly similar configuration are eliminated. Within the Minima Hopping structure prediction method, [118] an identification of identical configurations is required as well to prevent trapping in funnels that do not contain the global minimum. Some

machine learning approaches [135] are also based on similarity measures.

It is natural to characterize the dissimilarity between two structures p and q by a real number $d(p, q) \geq 0$. In order to give meaningful results $d(p, q)$ should satisfy the properties of a metric, namely

- coincidence axiom: $d(p, q) = 0$ if and only if $p \equiv q$,
- symmetry: $d(p, q) = d(q, p)$,
- triangle inequality: $d(p, q) + d(q, r) \geq d(p, r)$.

The coincidence axiom ensures that two configurations p and q are identical if their distance is zero, and vice versa. The triangle inequality is essential for clustering algorithms. If it is not satisfied, then it could happen that a configuration that belongs to one cluster in configuration space is also part of another cluster even though the distance between the two clusters is very large in the configuration space.

Since measuring distances between configurations is required in many applications, a considerable effort has been made to find cheap, yet reliable, distance measures that are not affected by the alignment of the two structures whose distance is being measured and by the indexing of the atoms in the structures. In the field of chemoinformatics a large number of different descriptors have been proposed to establish relations between structure and functionality. [136] For example, a structure can be represented by a binary string whose elements are set depending on whether some specific patterns exist in the structure. Then the similarity between structures is described by the Tanimoto coefficient. [132, 137] Another class of approaches is based on a generalizations of standard physical descriptors such as coordination numbers. Cheng *et al.* [138] used for instance the statistical properties (average, variance and bounds) of the coordination numbers while Lee *et al.* [139] used their weighted histograms in order to characterize the structures. Histogram-based methods were also used for the identification of crystalline structures. [140] All these methods have several tuning parameters such as the width of histogram bins or cutoff radii for the determination of coordination numbers [139] and their performance can critically depend on the choice of these parameters.

In this work we will introduce a family of parameter free metrics for measuring distances in configuration spaces. We show that these metrics fulfil all the mathematical requirements and demonstrate their excellent performance for a representative set of benchmark systems including covalent, metallic (simple or transition), ionic and organic structures. In the case of periodic systems, additional complexity comes into play because of non-uniqueness of the elementary cell. In the present work, our focus is on isolated molecules. The configurations in our test set are metastable low energy configurations obtained during a structure search using the Minima Hop-

ping Method [118] on the density functional theory (DFT) level as implemented in the BigDFT code. [30]

6.1 RMSD

A configuration of n alike atoms is uniquely represented by $\mathbf{R} \equiv (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) \in \mathbb{R}^{3\times n}$, where the column vector $\mathbf{r}_i$ represents the Cartesian coordinates of atom i. A distance based on the naive Frobenius norm

$$\|\mathbf{R}^p - \mathbf{R}^q\| = \Big(\sum_{i=1}^{n} \|\mathbf{r}_i^p - \mathbf{r}_i^q\|^2\Big)^{1/2} \tag{6.1}$$

can not be used to compare two configurations p and q, because it is not invariant with respect to translations or rotations of one configuration relative to the other. For this reason the commonly used root-mean-square distance (RMSD) is defined as the minimum Frobenius distance over all translations and rotations. By minimizing $\sum_i^n \|\mathbf{r}_i^p + \mathbf{d} - \mathbf{r}_i^q\|^2$ with respect to the translation vector $\mathbf{d}$ one obtains $\sum_i^n (\mathbf{r}_i^p + \mathbf{d} - \mathbf{r}_i^q) = 0$, i.e. the required translation is the difference between the centroids $\mathbf{d} = \frac{1}{n}\sum_i^n \mathbf{r}_i^q - \frac{1}{n}\sum_i^n \mathbf{r}_i^p$. Therefore we will assume in the following that all $\mathbf{r}_i$ are measured with respect to the centroids of the corresponding configuration which allows us to drop the minimization with respect to the translation $\mathbf{d}$. Then, finding the rotation $\mathbf{U}$ around the common centroid which minimizes

$$RMSD_l(p,q) = \frac{1}{\sqrt{n}} \min_{\mathbf{U}} \|\mathbf{R}^p - \mathbf{U}\mathbf{R}^q\| \tag{6.2}$$

is a local minimization problem and hence we denote this version of the RMSD by RMSD$_l$. The Kabsch algorithm [141] provides the solution to this problem based on the Euler angles. We perform the local minimization by an alternative method based on quaternions [142] which is more stable and numerically very cheap. [143, 144]

The RMSD$_l$ is however not invariant under index permutations of chemically identical atoms. If the configuration p and q are identical, Eq. (6.2) will be different from zero if we permute for instance in $\mathbf{R}^q$ the positions $\mathbf{r}_i^q$ and $\mathbf{r}_j^q$ of atoms i and j. The minimum Frobenius distance obtained by considering all possible index permutations for an arbitrary rotation $\mathbf{U}$ is

$$RMSD_{\mathbf{P}}(p,q) = \frac{1}{\sqrt{n}} \min_{\mathbf{P}} \|\mathbf{R}^p - \mathbf{U}\mathbf{R}^q\mathbf{P}\|, \tag{6.3}$$

$\mathbf{P}$ being an $n \times n$ permutation matrix. This assignment problem is solved in polynomial time using the Hungarian algorithm. [145] However, what is really needed is a solution of the combined problem of the global minimization over all rotations and permutations, namely

$$RMSD(p,q) = \frac{1}{\sqrt{n}} \min_{\mathbf{P},\,\mathbf{U}} \|\mathbf{R}^p - \mathbf{U}\mathbf{R}^q\mathbf{P}\|. \tag{6.4}$$

The global minimum RMSD fulfills all the properties of a metric. The coincidence and symmetry properties are easy to see. Using the standard triangle inequality, the proof of the triangle property is as follows:

$$
\begin{aligned}
& RMSD(p,q) + RMSD(q,r) \\
&= \frac{1}{\sqrt{n}} \min_{\mathbf{P},\mathbf{U}} \|\mathbf{U}\mathbf{R}^p\mathbf{P} - \mathbf{R}^q\| + \frac{1}{\sqrt{n}} \min_{\mathbf{P},\mathbf{U}} \|\mathbf{R}^q - \mathbf{U}\mathbf{R}^r\mathbf{P}\| \\
&= \frac{1}{\sqrt{n}} \|\mathbf{U}_{pq}\mathbf{R}^p\mathbf{P}_{pq} - \mathbf{R}^q\| + \frac{1}{\sqrt{n}} \|\mathbf{R}^q - \mathbf{U}_{rq}\mathbf{R}^r\mathbf{P}_{rq}\| \\
&\geq \frac{1}{\sqrt{n}} \|\mathbf{U}_{pq}\mathbf{R}^p\mathbf{P}_{pq} - \mathbf{R}^q + \mathbf{R}^q - \mathbf{U}_{rq}\mathbf{R}^r\mathbf{P}_{rq}\| \\
&\geq \frac{1}{\sqrt{n}} \|\mathbf{R}^p - \mathbf{U}_{rp}\mathbf{R}^r\mathbf{P}_{rp}\| \\
&= RMSD(p,r)
\end{aligned}
$$

where $\min_{\mathbf{P},\mathbf{U}} \|\mathbf{U}\mathbf{R}^p\mathbf{P} - \mathbf{R}^q\|$ is shown by $\|\mathbf{U}_{pq}\mathbf{R}^p\mathbf{P}_{pq} - \mathbf{R}^q\|$ for convenience.

6.1.1 Iterative global minimization of RMSD

Since $\mathbf{U}$ and $\mathbf{P}$ are not independent, no algorithm exists which can find the global RMSD within polynomial time. Just doing a search by alternating rotation and permutation steps using local minimizations and the Hungarian algorithm, respectively, is not guaranteed to converge to the global minimum with a finite number of steps. Trying out all possible permutations would lead to a factorial increase of the computing time with respect to n and this approach is therefore not feasible except for very small systems. We use a two-stage method for finding the global RMSD with moderate computational effort. The flowchart of the algorithm is depicted in Fig. 6.1 with the two different stages shown on the left and right sides. In the first stage we try to find the optimal global alignment of the two structures being compared. We first align two of the three principal axes of inertia of one configurations with the corresponding axes of the other one. A trial alignment is always followed by the application of the Hungarian algorithm to find the index permutation that gives the smallest RMSD. [146] The index matching in the Hungarian algorithm is done in the Cartesian space by associating to each atom $i \in$ p the closest atom $j \in$ q such that $\sum_i^n \|\mathbf{r}_i^p - \mathbf{r}_j^q\|$ is minimal. In other words, the columns of the $n \times n$ matrix made by $\|\mathbf{r}_i^p - \mathbf{r}_j^q\|$ are reordered such that its trace is minimal. The implementation of the Hungarian algorithm based on Ref. [147] finds the optimal index permutation within polynomial time and with a small prefactor. After this initial index matching, a rotation using quaternions is applied to refine the molecular alignment. If the required rotation is significant, the atomic index assignment should be repeated. This whole procedure is iterated until the atomic indices remain fixed after applying

Table 6.1: Number of remaining distinct configurations, average RMSD and average CPU-time (on single 2.4 GHz Intel core) for superimposing one pair of configurations at different steps of the two-stage RMSD global minimization. In the axes alignment (AA) stage, the principal axes of inertia as well as three molecular sets of axes obtained from vectorial atomic fingerprints are used. Every molecular alignment is always followed by the application of the Hungarian algorithm to find the optimal index permutation. In the Monte Carlo (MC) stage, random permutations are tried out which are followed by local minimization to get the optimal rotation. Because of the stochastic nature of the MC part, the reported values might change in different runs.

		Si_{32}			Mg_{26}			$C_{22}H_{24}N_2O_3$		
		remaining distinct	$\overline{\text{RMSD}}$ [Å]	$\bar{t}_{\text{CPU}}$ [s]	remaining distinct	$\overline{\text{RMSD}}$ [Å]	$\bar{t}_{\text{CPU}}$ [s]	remaining distinct	$\overline{\text{RMSD}}$ [Å]	$\bar{t}_{\text{CPU}}$ [s]
Unanalyzed		317	1.40		111	3.44		60	2.75	
AA	axes of inertia	184	1.16		60	1.08		42	1.93	
	$(\mathbf{W}, \mathbf{W}')_{\alpha_1}$	184	1.06		59	1.06		42	1.89	
	$(\mathbf{W}, \mathbf{W}')_{\alpha_2}$	184	1.04		59	1.03		42	1.81	
	$(\mathbf{W}, \mathbf{W}')_{\alpha_3}$	184	1.02	< 0.001	59	1.01	< 0.001	42	1.78	< 0.001
MC	iter.$=10^3$	184	.978	0.03	59	.985	0.02	42	1.52	0.05
	iter.$=10^4$	184	.910	0.13	59	.864	0.11	42	1.51	0.13
	iter.$=10^5$	184	.852	1.1	59	.852	1.0	42	1.51	1.6
	iter.$=10^6$	184	.792	12.1	59	.824	10	42	1.51	15
	iter.$=10^7$	184	.791	132	59	.824	119	42	1.51	163

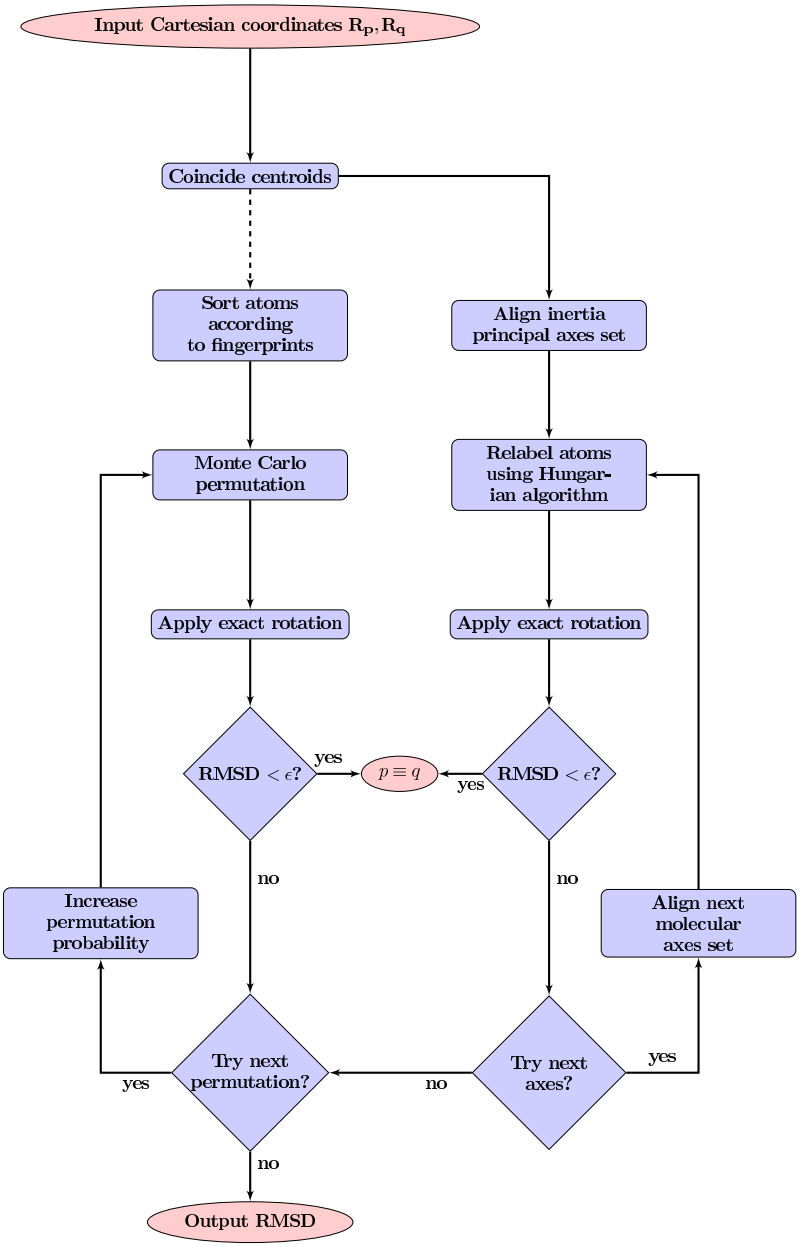

Figure 6.1: Flowchart of the algorithm of global minimization of RMSD in two major steps. The loop on the right runs over several sets of axes and matches atoms of a pair of configurations via aligning their molecular axes. The left loop shows the Monte Carlo (MC) permutation of identical particles while the parameters are dynamically tuned to obtain an acceptance rate close to 50%. The dashed line means that the right loop can be excluded.

the rotation. This procedure has allowed us to detect all identical configuration in this first stage, as seen in Table 6.1.

Since all the global alignment methods are empirical and can fail we apply several of them successively. After the first global alignment based on the principal axes of inertia we apply some more alignments steps based on axes which are derived from local atomic fingerprints (see next section). We set up an overlap matrix with s and p type Gaussian orbitals (see Appendix A) and find its principal eigenvector (i.e. the eigenvector with the largest eigenvalue; see Fig. A.1). Defining $\mathbf{w}_i = s_i \mathbf{p}_i$, where s_i and $\mathbf{p}_i$ are respectively s- and p-type components of the principal eigenvector belonging to atom i we can form two axes $\mathbf{W}$ and $\mathbf{W}'$

$$\mathbf{W} = \sum_i^n \mathbf{w}_i, \tag{6.5}$$

$$\mathbf{W}' = \sum_i^n \mathbf{w}_i \times \mathbf{r}_i \tag{6.6}$$

where the sum runs over the atoms, $\mathbf{r}_i$ represents the positions of atoms with respect to the center of mass and $\times$ denotes the cross product. First, we align $\mathbf{W}^q$ with $\mathbf{W}^p$ and then rotate q around it such that the plane made by ($\mathbf{W}^q$,$\mathbf{W}'^q$) coincides with the plane made by ($\mathbf{W}^p$,$\mathbf{W}'^p$). Depending on the width of the Gaussian used to construct the overlap matrix, several sets of axis may be constructed and tried one-by-one in this stage. If the alignment according to a new set of axes results in a smaller RMSD, we accept it. In Table 6.1 we show the results of the alignment of the principal axes of inertia as well as three sets of ($\mathbf{W}, \mathbf{W}'$) axes obtained by three different Gaussian widths α.

If a small enough RMSD is not found, we enter into an iterative stage (see left side of Fig. 6.1) where randomly chosen atoms are permuted within a thresholding Monte Carlo (MC) approach followed by applying the optimal rotation. In the thresholding MC step, two chemically identical atoms are selected according to a uniform random distribution. If by swapping them the RMSD is reduced, the permutation is accepted. To exclude the possibility of getting stuck in a local minimum, the permutation is also accepted if it causes the RMSD to increase by less than an adjustable parameter ξ. This parameter is dynamically updated at each step: if the acceptance rate so far is less/greater than 50%, then ξ is increased/decreased by a factor of 1.1. In this way, the average acceptance rate approaches 50% during the minimization. The iteration stops when the global minimum RMSD does not decrease any more for a large number of iterations. As seen in Table 6.1, the number of required MC iterations depends on the system size. For instance, for the biomolecule 10^4 MC iterations (which take on average 0.13 second on a single 2.4 GHz Intel core) are sufficient to find the global minimum RMSD between two configurations of this molecule. For a more systematic investigation of the scaling, we take the global minima of the Lenard-Jones (LJ) clusters with different sizes and apply random displacements of the unit

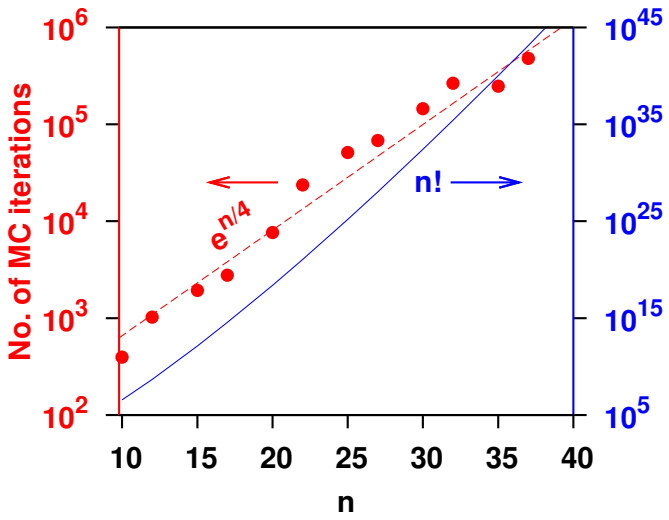

Figure 6.2: Number of MC iterations (averaged over 10 pairs of LJ clusters) to obtain the permutations corresponding to the global minimum of RMSD a function of number of particles n. The dashed line shows $41\exp(n/4)$ which is obtained by least square fit. For comparison, $n!$ is plotted with solid line.

magnitude to every atom (i.e. the RMSD between the randomized structures is almost one in the LJ length units). The averaged number of required MC iterations to get the asymptotic value of the RMSD (as obtained by 10^7 iterations), as a function of the cluster size n is shown in Fig. 6.2. Even though the number of iterations increases exponentially it is several orders of magnitude smaller than the number of possible permutations, i.e. $n!$.

6.2 Fingerprint Distances as Metrics

While the RMSD can be considered as the most basic quantity to measure the dissimilarities, finding the global minimum RMSD is numerically costly. Only in case that two structures are nearly identical the global minimum of RMSD is calculated with a polynomial computational time because no MC permutation is then required. Otherwise, even if the above described algorithm is used, the computational time increases exponentially with the number of permutable particles. In the following we will therefore introduce a family of metrics which are cheaper to calculate than the global RMSD yet in good agreement with it. We consider symmetric $N \times N$ matrices whose elements depend only on the interatomic distances $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$ of an n-atom configuration. Vectors $\mathbf{V}$ containing eigenvalues of such a matrix form a configurational fingerprint which allows

to identify a structure. The normalized Euclidean distance

$$\Delta_{\mathbf{V}}(p,q) = \frac{1}{\sqrt{N}}\|\mathbf{V}^p - \mathbf{V}^q\| \tag{6.7}$$

measures the dissimilarly between p and q with no need to superimpose them. Although we use the eigenvalues to form the vector $\mathbf{V}$ for describing entire structures throughout this work, one can also fill the vector $\mathbf{V}$ by the elements of selected eigenvectors. Then each element of $\mathbf{V}$ corresponds to an atom and the ensemble belonging to one atom forms an atomic fingerprint or descriptor of the local environment of the atom. For instance, if the principal eigenvector of the overlap matrix of one s and one p-type GTO per atom is used (as in Eqs. (6.5) and (6.6)), each individual atom is accordingly described by four numbers, as depicted in Fig. A.1

Since the matrix depends only on interatomic distances, the same holds true for the eigenvalues and eigenvectors, and $\mathbf{V}$ is thus invariant under translations, rotations and reflections of the configuration. In order to make $\Delta_{\mathbf{V}}$ also independent of the atomic indices, the elements of each $\mathbf{V}$ are sorted in an ascending order. This sorting can introduce discontinuities in the first derivative of the fingerprint distance with respect to changes in the atomic coordinates (e.g. when there is a crossing of eigenvalues) but does not destroy the important continuity of the fingerprint distance itself.

The coincidence axiom for a configurational fingerprint is satisfied if the dimension N of the matrix is sufficiently large and if therefore the resulting fingerprint vector is sufficiently long. We show in Appendix C that how a hypersurface of constant fingerprint can be constructed if the length of the fingerprint is short. What we would like to show however is the opposite, namely that no distinct configurations with identical fingerprints exist if the fingerprint is long enough. Since the fingerprint distance is a non-linear function, it can in principle not be excluded that two distinct configurations with identical fingerprints exist even if the fingerprint vector is longer than the threshold value. Since we recommend for a unique identification fingerprints which are considerably longer than the threshold value, namely fingerprints of length $3n$ or even $4n$ it is however extremely unlikely that such configurations exist and the coincidence axiom can be taken to be fulfilled. To confirm this assumption numerically as well, we did extensive numerical searches where we tried to find a second configuration which has a fingerprint which is identical to the fingerprint of a reference configuration. The initial guess for the second configuration was random and then this second configuration was moved in such a way as to minimize the difference between the fingerprints. All these numerical minimizations lead to non-zero local minima, i.e. we were not able to find numerically any violation of the coincidence axiom for vectors of length $3n-3$ based on the Hessian matrix and vectors of length $4n$ based on an overlap matrix with s and p orbitals.

Even though the eigenvalue vector is much shorter than the vector containing all matrix

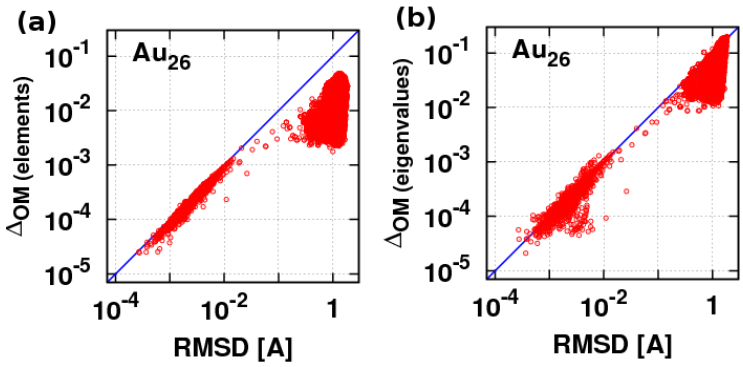

Figure 6.3: Correlation of the pairwise Euclidean distances based on vectors consisting either of all of the sorted elements of the overlap matrix (a) or eigenvalues of this matrix (b) and the RMSD for 1000 metastable configurations of a 26 atom gold cluster. The gap in the fingerprint distances between identical and distinct configuration is larger if eigenvalues are used (panel a).

elements, the fingerprint distances based on the eigenvalues are better than those obtained by sorting all the matrix elements depending on interatomic distances into a vector. One can in some cases construct distinct so-called homometric configurations [148] for which the fingerprint vectors of the sorted matrix elements are identical whereas the eigenvalue vectors are not identical and allow thus to distinguish between them. In addition, our empirical results of Fig. 6.3 show that the gap between identical and distinct pairs is larger for the eigenvalues than for the sorted matrix elements. Because the geometry relaxations were stopped when the force on each atom is within 0.01 eV/Å, identical configurations are in practice identical only up to some finite precision,i.e the atomic positions of the configurations are contaminated by noise. Two configurations are considered to be identical (i.e. belong to the same cluster) if their distance is below a certain threshold. An unambiguous threshold for distinguishing between distinct and non-distinct configurations can only be found if a well detectable gap exists in the distance space. Hence the existence of a large gap is an important benefit of a fingerprint method.

In an application to Ni clusters Grigoryan *et al.* [149] used the sorted interatomic distances to find the similarities between an $(n-1)$-atom cluster and $(n-1)$-atom parts of an n-atom cluster. This similarity measure also leads to a gap which is smaller than the one obtained from eigenvalue based fingerprints of either the corresponding r_{ij} matrix or the matrices proposed in this article (cf. Figs. 6.4 and 6.5). So it seems to be a general feature that fingerprints based on the eigenvalues are better than those based on sorted matrix elements.

In the following we will describe several matrix constructions which can be used for fingerprinting. These matrices are closely related to measurable quantities that are traditionally used by experimentalists to identify structures.

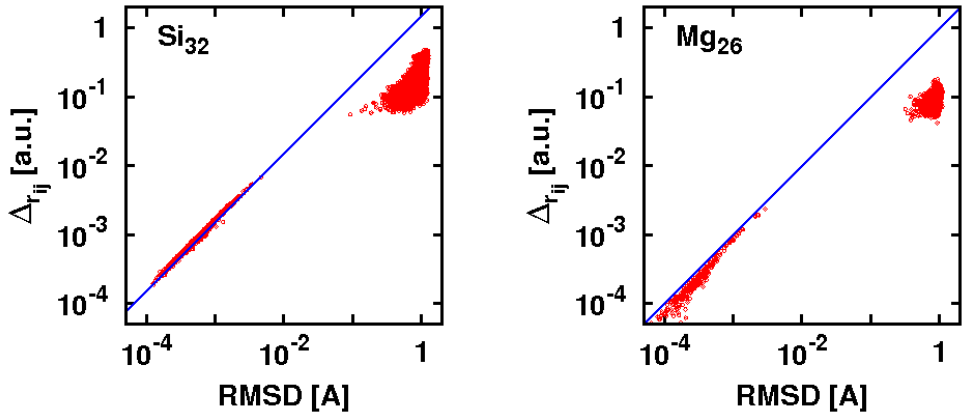

Figure 6.4: Correlation of Euclidean distance of the sorted interatomic distances with RMSD for the metastable configurations of the Si_{32} and Mg_{26} clusters. The gap that allows to discriminate distinct from non-distinct configurations is smaller in both cases compared to the fingerprints based on eigenvalues.

6.2.1 Hamiltonian Matrix

Emission and absorption spectra arise from transitions between discrete electronic energy levels. Each element has its characteristic energetic levels and therefore atomic spectra can be used as elemental fingerprints. When atoms are assembled into structures the electronic states of the constituent atoms are modified depending on the arrangement of the atoms. A computational analogue to electronic energy levels probed by various spectroscopic experiments are the Kohn-Sham (KS) energy eigenvalues, even though they do not represent the physical excitation energies. Since the Hamiltonian matrix depends only on the interatomic distances, the sorted KS eigenvalues are invariant to translations, rotations, reflections and permutations of atoms.

We examine fingerprints that are based on the occupied KS eigenvalues only as well as fingerprints that are based both on the occupied and unoccupied eigenvalues. The former were obtained from the self-consistent eigenvalues calculated in a large wavelet basis, [30] whereas, for simplicity, the latter were obtained from the non-self-consistent input guess eigenvalues calculated in a minimal Gaussian type atomic orbitals (GTO) basis set for a charge density which is a superposition of atomic charge densities. The configuration distances obtained from the occupied KS eigenvalues, denoted by $\Delta_{KS}(p,q)$, show an excellent correlation with the RMSD for all three test sets, see Fig. 6.5. Even though the vector $\mathbf{V}_{GTO}$ is in all cases longer than the vector $\mathbf{V}_{KS}$ (e.g. two times longer in case of the Si cluster), fingerprint distances based on the former do not better correlate with the RMSD than fingerprint distances based on the latter, although, as will be discussed in Sec. 6.3, the coincidence theorem is not satisfied.

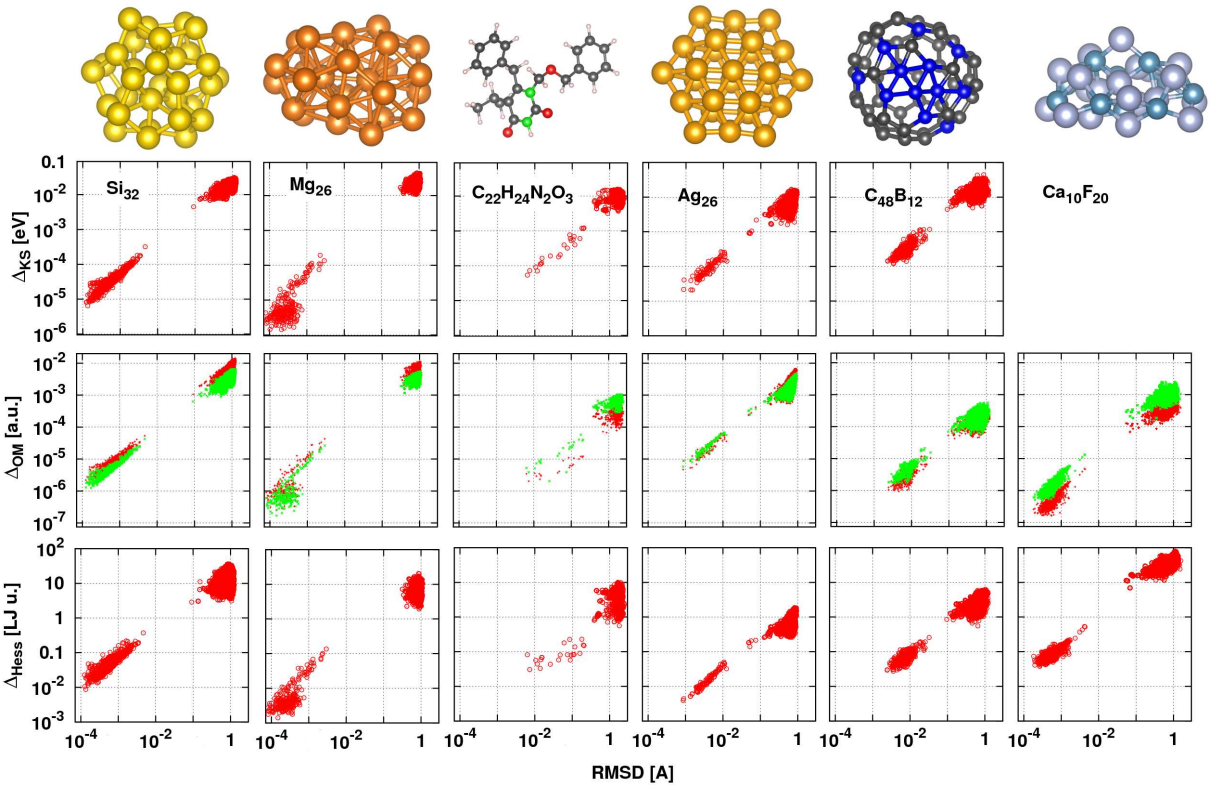

Figure 6.5: Comparison of RMSD based distances between configurations with fingerprint based distances. The fingerprints are based on the eigenvalues of the Kohn-Sham Hamiltonian matrix (first row), the overlap matrix (second row) and the Lennard-Jones Hessian matrix with the RMSD for sets of semiconductor (silicon), simple metal (magnesium), organic (6-benzyl-1-benzyloxymethyl-5-isopropyl uracil), transition metal (silver), covalent fullerene-type ($C_{48}B_{12}$) and ionic (calcium fluoride) clusters. Shown on top are representative configurations. Each set consists of a few hundred configurations, all being low-energy local minima within DFT, except those of $Ca_{10}F_{20}$ which are local minima of the Tosi-Fumi potential (parameters from Ref. [150]). For the latter system the Kohn-Sham eigenvalues are obviously not calculated. For the five sets where Kohn-Sham eigenvalues can be calculated their number is determined by the occupied valence states and is given, respectively from left to right, by $64 = 2n$, $26 = n$, $n < 70 < 2n$, $26 = n$ and $n < 114 < 2n$, n being the number of atoms. For the overlap matrix, results for both s-only (red) and s-and-p (green) overlap matrices are shown, leading to fingerprint vectors of lengths n and $4n$, respectively. For the Hessian matrix $3n - 3$ eigenvalues are non-zero. Even in the cases where the length of the fingerprint vector is shorter than $3n - 6$ the agreement with the RMSD is good and allows always to identify distinct and non-distinct configurations.

6.2.2 Overlap Matrix

A matrix which has similar properties as the Hamiltonian matrix is the overlap matrix (OM) expressed in terms of GTOs. Contrary to the Hamiltonian, all elements of the OM can easily be calculated analytically (Appendix A). In the simplest case where only uncontracted s-type GTO's are used, the resulting fingerprint consists of n scalars. Information about the radial distribution can be incorporated in the OM by adding p and d type GTO's. In this way the configurational fingerprint vector becomes also longer than $3n-6$ and the coincidence axiom is expected to be satisfied. By including both s and p type GTO's in the overlap matrix, the number of eigenvalues becomes $4n$ and all of our attempts to reduce $\Delta_{\mathbf{V}}(p,q)$ to very small values failed when p and q are distinct.

If the fingerprint is used to calculate distances between our test set of local minima configurations, it turns out that adding p-type orbitals gives only a marginal improvement, in the sense that the distance gap separating identical and distinct configurations gets larger. Adding additional d-type orbitals has virtually no effect. This is related to the fact that it is very unlikely that two local minima lie on the hypersurface that leaves the fingerprint invariant. The width of the GTO's was in all our tests given by the covalent radius of the atom on which the GTO was centered.

6.2.3 Hessian Matrix

The vibrational properties, which are frequently used experimentally to identify structures, are closely related to the Hessian matrix which consists of the second order derivatives of the energy with respect to the atomic positions. The vibrational frequencies are up to a scaling factor related to the mass of the atoms equal to the square root of the eigenvalues of the Hessian matrix. This matrix also belongs to the class of matrices with the desired properties. Unfortunately the calculation of the Hessian is rather expensive in the context of a DFT calculation and can also be cumbersome with sophisticated force fields. We will therefore not further pursue approaches based on an Hessian which is calculated within the same high level method as the energy and forces. It however turns out that eigenvalues or eigenvectors of the Hessian matrices which are derived from another cheaper potential such as the LJ potential give also good fingerprints. This is shown in Fig. 6.5 for our six test systems after the lengths were scaled to the equilibrium bond-length of the LJ potential.

6.3 Discussion

Various $n \times n$ matrices, have been used previously to characterize molecular configurations. The definition of a molecular descriptor can be based on either eigenvalues, spectral moments (defined

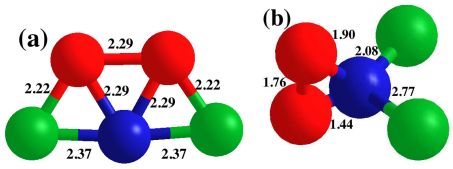

Figure 6.6: Two distinct configurations of the Si_5 cluster with an identical set of SPRINT coordinates, i.e. 3.59 (green), 4.37 (red), 4.85 (blue), using the parameters given in the Supplementary Material of Ref. [151]. The planar structure shown in (a) is a local minimum in LDA-DFT. The numbers show the bond-lengths in Å.

as the kth power of the eigenvalues, where the natural number $k \leq n$ is then the order of the moment) or even the elements of the eigenvector associated with the largest eigenvalue (i.e. the principal eigenvector) of many matrices e.g. adjacency, Laplacian, distance and reciprocal distance, distance-path, etc.; for review see Ref. [136]. The contact matrix from the graph theory exhibits discontinuities when the atomic distances cross the cutoff radius. By introducing a smooth cutoff these discontinuities disappear and the resulting matrix has been used as a fingerprinting tool in the SPRINT method. [151] Presumably not only the contact matrix but also other matrices from spectral graph theory such as the Laplace matrix could be used in a similar way. We did for instance not find significant differences in performance between the contact and Laplacian matrices. We found however that fingerprints based on either of them are rather sensitive to the form of the smooth cutoff function. Tuning of the parameters of this cutoff function is therefore required to obtain good results. In both cases, the resulting atomic fingerprints are real scalars which mostly contain information about the number of nearest neighbours of each atom and might be insufficient to characterize the chemical environment of an atom. Better chemical environment descriptors can however be obtained by adding information about the radial distribution of the neighbours. [152, 153] The Coulomb matrix is another matrix whose eigenvalues have been used to characterize configurations. [135] The off-diagonal elements of this n-by-n matrix are the pairwise Coulomb repulsions $q_i q_j / r_{ij}$, while the diagonal is filled with $q_i^{2.4}/2$, q_i being the core charge of atom i.

As discussed before, such a fingerprint of length n is not long enough to satisfy the coincidence axiom and can thus fail to detect structural differences. This has already been shown for the Coulomb matrix. [154] We show in Fig. 6.6 two distinct configurations of a Si_5 cluster which have identical sets of SPRINT coordinates. Note that the Si atoms with identical SPRINT coordinates in the configuration shown in Fig. 6.6(b), have very different environments. This shows that SPRINT, like any other $n \times n$ matrix-based fingerprint, fails to describe uniquely the entire structure and/or the chemical environment of an atom.

6.4 Summary

In summary, we have shown that the RMSD, the most natural measure of dissimilarity between two configurations, satisfies the properties of a metric when it is obtained by a global minimization over all rotations and index permutations. We have presented a Monte Carlo method to calculate the global minimal RMSD which does not require to try out all possible index permutations and which is thus computationally feasible. At the same time we have introduced fingerprints which are much cheaper to calculate because they do not require a structural superposition. Nevertheless the fingerprint based distances correlate in all our test cases with the RMSD, in the sense that small RMSD distances correspond to small fingerprint distance and vice versa. In contrast to numerous previously proposed fingerprints they satisfy the coincidence axiom and allow therefore to distinguish distinct from non-distinct configurations in a unique way. Within a DFT calculation the metric based on the Kohn-Sham eigenvalues is a good choice since the eigenvalues are a byproduct of any DFT calculation and thus no extra effort is required to obtain them. For the coincidence axiom to be satisfied, the number of bound eigenstates whose Kohn-Sham eigenvalues can be included in the fingerprint vector has however to be larger than $3n - 6$. If Kohn-Sham eigenvalues are not available, the method based on the eigenvalues of the overlap matrix constructed from s and p orbitals is recommended, since it leads to matrices whose elements can be calculated analytically and because the fingerprint vector is long enough ($4n$) to make the probability of a violation of the coincidence axiom vanishingly small. Even if the coincidence axiom is violated, it turns out in practice that it is very rare that different physically reasonable metastable configurations give rise to identical fingerprints. For our test sets of low energy local minima configurations metrics which violated the coincidence axiom therefore allowed nevertheless in all cases to distinguish between distinct and non-distinct configurations. In other applications where small movements away from metastable configurations lead to a change of physical properties, such as in force fields based on machine learning, a violation of the coincidence theorem can however not be tolerated. All the proposed variants of our approach are parameter free and no parameter tuning is therefore required.

Appendix A

Overlaps between GTO's

The normalized Gaussian type orbitals (GTO) centered at the atomic positions $\mathbf{r}_i$ in Cartesian coordinates are given by

$$\phi_i^{\mathbf{l}}(\mathbf{r}) = N_{\mathbf{l}}(x - x_i)^{l_x}(y - y_j)^{l_y}(z - z_i)^{l_z} e^{-\alpha_i \|\mathbf{r}-\mathbf{r}_i\|^2}$$

where $\mathbf{l} = (l_x, l_y, l_z)$ and $N_{\mathbf{l}}$ is the normalization factor. Depending on the angular moment $L = l_x + l_y + l_z$ the functions are labeled as as s-type (L=0), p-type (L=1), d-type (L=2) and so on. We take the Gaussian width α_i inversely proportional to the square of the covalent radius of atom i throughout this work.

The Gaussian product theorem says that the product of two Gaussian functions is again a Gaussian function. Therefore the overlap integrals between a pair of GTO's, namely

$$\langle \phi_i^{\mathbf{l}} | \phi_j^{\mathbf{l}'} \rangle = \int d\mathbf{r} \phi_i^{\mathbf{l}}(\mathbf{r}) \phi_j^{\mathbf{l}'}(\mathbf{r}) \tag{A.1}$$

can be evaluated analytically. This gives the normalization factors as

$$N_{\mathbf{l}}(\alpha_i) = \frac{1}{\sqrt{\langle \phi_i^{\mathbf{l}} | \phi_i^{\mathbf{l}} \rangle}} = (2\alpha_i/\pi)^{3/4} \sqrt{n_{l_x} n_{l_y} n_{l_z}}, \quad n_k = \frac{(4\alpha_i)^k}{(2k-1)!!}. \tag{A.2}$$

All GTO's are recursively obtained by differentiating

$$\phi_i^s(\mathbf{r}) = \left(\frac{2\alpha_i}{\pi}\right)^{3/4} e^{-\alpha_i \|\mathbf{r}-\mathbf{r}_i\|^2}$$

with respect to the Cartesian components of $\mathbf{r}_i$. For instance

$$\phi_i^{p_x}(\mathbf{r}) = 2\sqrt{\alpha_i}(x - x_i)\phi_i^s(\mathbf{r})$$

can also be expressed as

$$\phi_i^{p_x}(\mathbf{r}) = \frac{1}{\sqrt{\alpha_i}}\frac{\partial\phi_i^s(\mathbf{r})}{\partial x_i}. \tag{A.3}$$

The general formula for the overlap integrals, i.e. the elements of the overlap matrix, is given, e.g., by Eq. (3.5) in Ref. [155] and can also be calculated from recursion relations. [156] For convenience, we restate the simplified relations for the special cases involving s and p-type GTO's all in terms of the basic quantity

$$S_{ij} = S_{ji} = \Big(\frac{2\sqrt{\alpha_i\alpha_j}}{\alpha_i + \alpha_j}\Big)^{3/2} \exp\Big[\frac{-\alpha_i\alpha_j}{\alpha_i + \alpha_j} r_{ij}^2\Big] \tag{A.4}$$

where $r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|$, which is indeed the s-s overlap integral

$$\langle\phi_i^s|\phi_j^s\rangle = S_{ij}$$

Using Eq. (A.3) we obtain

$$\begin{aligned}\langle\phi_i^{p_x}|\phi_j^s\rangle &= \frac{1}{\sqrt{\alpha_i}}\frac{\partial S_{ij}}{\partial x_i} \\ &= -\Big(\frac{2\sqrt{\alpha_i}\alpha_j}{\alpha_i + \alpha_j}\Big)(x_i - x_j)S_{ij}\end{aligned} \tag{A.5}$$

and

$$\langle\phi_i^{p_x}|\phi_j^{p_{x'}}\rangle = \Big(\frac{2\sqrt{\alpha_i\alpha_j}}{\alpha_i + \alpha_j}\Big)\Big[\delta_{x,x'} - \frac{2\alpha_i\alpha_j}{\alpha_i + \alpha_j}(x_i - x_j)(x_i' - x_j')\Big]S_{ij} \tag{A.6}$$

where $x, x' \in \{x, y, z\}$ and δ denotes the Kronecker delta. The derivative of the basic quantity S_{ij} with respect to the atomic positions

$$\frac{\partial S_{ij}}{\partial x_k} = (\delta_{ik} - \delta_{jk})\Big(\frac{-2\alpha_i\alpha_j}{\alpha_i + \alpha_j}\Big)(x_i - x_j)S_{ij} \tag{A.7}$$

is required to calculate the derivative of the overlap matrix elements, which in turn determine the derivative of its eigenvalues

$$D_{\nu,x_k} \equiv \frac{\partial V_\nu}{\partial x_k} = \langle\nu|\frac{\partial O}{\partial x_k}|\nu\rangle, \tag{A.8}$$

where the eigenvector $|\nu\rangle$ corresponds to the eigenvalue V_ν of the overlap matrix O.

Eigenvectors associated to small eigenvalues seem not to contain any useful information. We therefore use the principal eigenvector of the overlap matrix as an atomic fingerprint, see Fig. A.1. This vector gives the coefficients required to construct the pseudo-orbital with the largest pseudo charge density. This charge density has similarities to a true charge density since it is large in regions between neighboring atoms where covalent bonding can occur (Fig A.2).

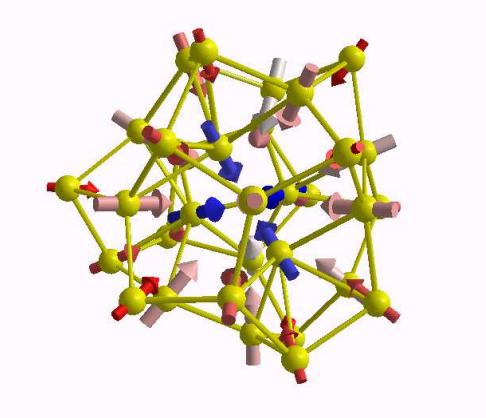

Figure A.1: Description of atomic environments for a Si_{32} cluster using the combined scalar and vectorial atomic fingerprints. Each atomic fingerprint consists of a scalar and a vector which are the corresponding s and (p_x, p_y, p_z) components of the principal eigenvector of the $4n \times 4n$ overlap matrix. The color of the vectors indicates the value (red corresponds to small values and blue to large values) of the scalar (s-type) fingerprint.

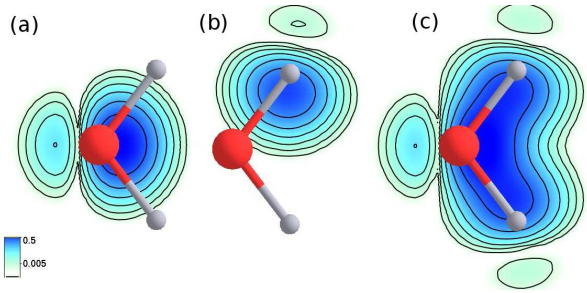

Figure A.2: Contributions from an oxygen (a) or hydrogen atom (b) to the total (c) pseudo-charge density $|\psi(\mathbf{r})|^2$ on the molecular plane for a water molecule. The coefficients of the orbitals $\phi_i^{\mathbf{l}}$ from which the pseudo-wavefunction ψ is made, are the elements of the principal eigenvector of the overlap matrix constructed from s and p-type GTO's.

Appendix B

Closed-form of superimposing rotation

A quaternion $\mathcal{Q} = (\mathcal{Q}_0, \mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3)$ is an extension of the idea of complex numbers to one real ($\mathcal{Q}_0$) and three imaginary parts. According to the Euler's rotation theorem, a rotation in space which keeps one point on the rigid body (centroid in our case) fixed, can be represented by four real numbers: one for the rotation angle and three for the rotation axis (we assume that the center of rotation is on the origin). A unit quaternion, i.e. $\|\mathcal{Q}\|^2 = \mathcal{Q}_0^2 + \mathcal{Q}_1^2 + \mathcal{Q}_3^2 + \mathcal{Q}_4^2 = 1$, can represent conveniently this axis-angle couple as

$$\mathcal{Q} = \Big(\cos\big(\frac{\theta}{2}\big), \hat{\mathbf{u}} \sin\big(\frac{\theta}{2}\big) \Big)$$

where θ is the rotation angle around the unit axis $\hat{\mathbf{u}} = a\hat{\mathbf{i}} + b\hat{\mathbf{j}} + c\hat{\mathbf{k}}$. The corresponding orthogonal rotation matrix is

$$\mathbf{U} = \begin{bmatrix} \mathcal{Q}_0^2 + \mathcal{Q}_1^2 - \mathcal{Q}_2^2 - \mathcal{Q}_3^2 & 2\mathcal{Q}_1\mathcal{Q}_2 - 2\mathcal{Q}_0\mathcal{Q}_3 & 2\mathcal{Q}_1\mathcal{Q}_3 + 2\mathcal{Q}_0\mathcal{Q}_2 \\ 2\mathcal{Q}_1\mathcal{Q}_2 + 2\mathcal{Q}_0\mathcal{Q}_3 & \mathcal{Q}_0^2 - \mathcal{Q}_1^2 + \mathcal{Q}_2^2 - \mathcal{Q}_3^2 & 2\mathcal{Q}_2\mathcal{Q}_3 - 2\mathcal{Q}_0\mathcal{Q}_1 \\ 2\mathcal{Q}_1\mathcal{Q}_3 - 2\mathcal{Q}_0\mathcal{Q}_2 & 2\mathcal{Q}_2\mathcal{Q}_3 + 2\mathcal{Q}_0\mathcal{Q}_1 & \mathcal{Q}_0^2 - \mathcal{Q}_1^2 - \mathcal{Q}_2^2 + \mathcal{Q}_3^2 \end{bmatrix}. \quad \text{(B.1)}$$

The optimum rotation U which minimizes RMSD, indeed maximizes the correlation between $\mathbf{R}^p$ and $\mathbf{R}^q$, i.e. the atomic Cartesian coordinates with respect to the common center of mass. Based on quaternions, [142] the optimum U is given by $\mathcal{Q}$ which is identical to the principal

eigenvector of the 4×4 symmetric, traceless matrix

$$\mathcal{F} = \begin{bmatrix} \mathcal{R}_{xx} + \mathcal{R}_{yy} + \mathcal{R}_{zz} & \mathcal{R}_{yz} - \mathcal{R}_{zy} & \mathcal{R}_{zx} - \mathcal{R}_{xz} & \mathcal{R}_{xy} - \mathcal{R}_{yx} \\ \mathcal{R}_{yz} - \mathcal{R}_{zy} & \mathcal{R}_{xx} - \mathcal{R}_{yy} - \mathcal{R}_{zz} & \mathcal{R}_{xy} + \mathcal{R}_{yx} & \mathcal{R}_{xz} + \mathcal{R}_{zx} \\ \mathcal{R}_{zx} - \mathcal{R}_{xz} & \mathcal{R}_{xy} + \mathcal{R}_{yx} & -\mathcal{R}_{xx} + \mathcal{R}_{yy} - \mathcal{R}_{zz} & \mathcal{R}_{yz} + \mathcal{R}_{zy} \\ \mathcal{R}_{xy} - \mathcal{R}_{yx} & \mathcal{R}_{xz} + \mathcal{R}_{zx} & \mathcal{R}_{yz} + \mathcal{R}_{zy} & -\mathcal{R}_{xx} - \mathcal{R}_{yy} + \mathcal{R}_{zz} \end{bmatrix} \tag{B.2}$$

where $\mathcal{R}$ is the correlation matrix whose elements are $\mathcal{R}_{xy} = \sum_i^n x_i^p y_i^q$ and so. Note that, Eq. (6.2) is then given by

$$RMSD(p,q) = \sqrt{\frac{1}{n}\Big(\|\mathbf{R}^p\|^2 + \|\mathbf{R}^q\|^2 - 2\lambda^*\Big)} \tag{B.3}$$

where λ^* is the largest eigenvalue of $\mathcal{F}$.

Appendix C

Constant-fingerprint hypersurfaces

Using a constructive iterative procedure, we show in the following that the coincidence axiom for a configurational fingerprint is not satisfied if the dimension of the matrix is not sufficiently large and if therefore the resulting fingerprint vector is not sufficiently long. Consider two configurations p and q which are close. The difference of the fingerprint vectors is then given by a first order Taylor expansion

$$\mathbf{V}^p - \mathbf{V}^q \simeq D(q)(\mathbf{R}^p - \mathbf{R}^q). \tag{C.1}$$

Note that, instead of the $3 \times n$ matrix notation here we use a column vector $\mathbf{R} \in \mathbb{R}^{3n}$ for representing the atomic coordinates. Since $\mathbf{V}$ is a column vector of length N, the first derivative $D(q) \equiv \frac{\partial \mathbf{V}}{\partial \mathbf{R}}\big|_{\mathbf{R}=\mathbf{R}^q}$ is a $N \times 3n$ matrix. We assume that D has always the largest possible rank for the three types of matrices discussed in more detail in this Section. For the Hamiltonian matrix this maximal rank $r_{\max}$ equals $\min(N, 3n-6)$ if all N eigenstates included in the fingerprint vector are bound. For the overlap matrix $r_{\max}$ equals $\min(N-1, 3n-6)$ because the diagonal elements are independent of the configuration. For the Hessian matrix $r_{\max} = 3n-6$ for configurations that are local minima with respect to the interaction potential and $r_{\max} = 3n-3$ for all other cases. [157]

If $r_{\max}$ is less than $3n-6$ one can find on a hypersurface of dimension $3n-6-r_{\max}$ (i.e. the nullity of D) configurations with identical fingerprint vectors, which are given as a solution of the equation

$$D\delta\mathbf{R} = \mathbf{0}. \tag{C.2}$$

Formulated in words, configurational displacement vectors $\delta\mathbf{R}$ which are in the null space of D

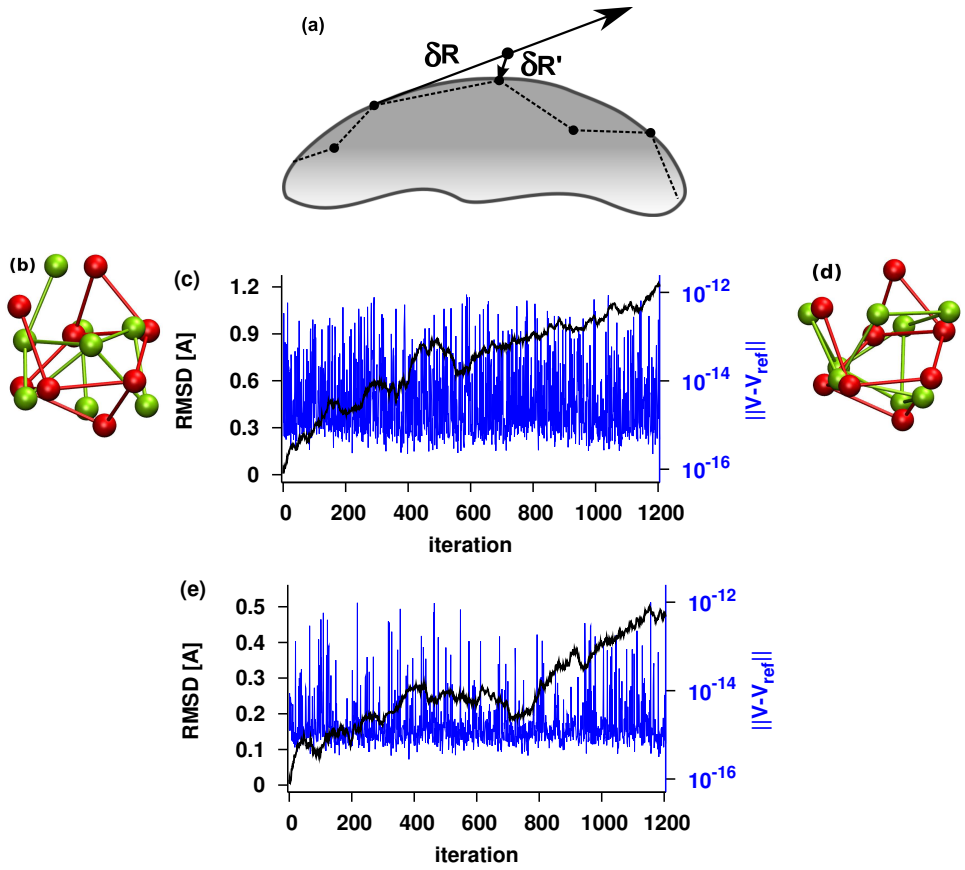

Figure C.1: (a) Schematic illustration of the exploration of the hypersurface defined by $\mathbf{V} = \mathbf{V}_{\text{ref}}$ consisting of iterative movements along $\delta\mathbf{R}$ (in the null space of D) followed by Newton step(s) $\delta\mathbf{R}'$ to come back to the hypersurface. Panel (b) shows two configurations (in red and green) of a Si_8 cluster whose fingerprint vectors of length n, obtained from an overlap matrix with one set of s-type GTO's, are identical. Panel (c) shows the evolution of the RMSD during the exploration of the hypersurface leading from the red structure to the green structure. Panels (d) and (e) contain the some information as panels (b) and (c) but for a fingerprint of length $2n$ obtained from an overlap matrix with two sets of s-type GTO's. In both cases $\|\mathbf{V} - \mathbf{V}_{\text{ref}}\|$ is vanishingly small.

leave the fingerprint invariant to first order. For configurations which are further apart the first order approximation breaks down but Eq. (C.2) can still be used as a starting point for mapping out such a hypersurface iteratively. We perform a move with a small amplitude along a vector $\delta\mathbf{R}$ in the null space of D. To correct for the small second and higher order deviations of the eigenvalues away from the hypersurface of constant eigenvalues defined as $\mathbf{V} = \mathbf{V}_{\text{ref}}$ we then solve

$$D\delta\mathbf{R}' = \mathbf{V}_{\text{ref}} - \mathbf{V} \tag{C.3}$$

for the required displacement $\delta\mathbf{R}'$. Like Eq. (C.1), the latter equation does not have a unique solution and we can therefore choose an arbitrary set of r_{max} coordinates which we want to modify in order to go back onto the hypersurface of constant eigenvalues. If the corresponding $r_{\text{max}} \times r_{\text{max}}$ matrix made out of D was ill-conditioned, we select another set of r_{max} atomic modification coordinates to ensure that Eq. (C.3) is solved accurately. Since this moving back to the hypersurface requires only tiny displacements a single solution of the linear system is sufficient. If this was not the case it could be repeated which would correspond to a Newton iteration. By iterating this procedure of moves along the null space followed by moves that bring us exactly back on the hypersurface we can obtain clearly distinct configurations whose fingerprints are identical up to machine precision. Such examples are shown in Fig. C.1 where the procedure is also illustrated schematically. Note that at each iteration we orthogonalize $\delta\mathbf{R}$ of the previous iteration to the row space of current D. This reduces the probability of moving backwards to the starting point.

Bibliography

[1] Weaver, J., Abraham, D. J. Vac. Sci. Technol. B **9** (1991) 1559

[2] Nonnenmacher, M., O'Boyle, M., Wickramasinghe, H. Appl. Phys. Lett. **58** (1991) 2921

[3] Sadewasser, S., Glatzel, T., Rusu, M., Jager-Waldau, A., Lux-Steiner, M.C. Appl. Phys. Lett. **80** (2002) 2979

[4] Barth, C., Foster, A.S., Henry, C.R., Shluger, A.L.: Recent trends in surface characterization and chemistry with high-resolution scanning force methods. Advanced Materials **23** (2011) 477–501

[5] Sadewasser, S., Glatzel, T.: Kelvin Probe Force Microscopy: Measuring and Compensating Electrostatic Forces. Springer, Heidelberg Dordrecht London New York (2011)

[6] Rosenwaks, Y., Shikler, R., Glatzel, T., Sadewasser, S.: Kelvin probe force microscopy of semiconductor surface defects. Phys. Rev. B **70** (2004) 085320

[7] Kawai, S., Glatzel, T., Hug, H.J., Meyer, E.: Atomic contact potential variations of si(111)-7 7 analyzed by Kelvin probe force microscopy. Nanotechnology **21** (2010) 245704

[8] Barth, C., Henry, C.R.: Surface double layer on (001) surfaces of alkali halide crystals: A scanning force microscopy study. Phys. Rev. Lett. **98** (2007) 136804

[9] Yamauchi, T., Tabuchi, M., Nakamura, A.: Size dependence of the work function in inas quantum dots on gaas(001) as studied by Kelvin force probe microscopy. Appl. Phys. Lett. **84** (2004) 3834–3836

[10] Glatzel, T., Marrón, D.F., Schedel-Niedrig, T., Sadewasser, S., Lux-Steiner, M.C.: Cugase[sub 2] solar cell cross section studied by Kelvin probe force microscopy in ultrahigh vacuum. Appl. Phys. Lett. **81** (2002) 2017–2019

[11] Hoppe, H., Glatzel, T., Niggemann, M., Hinsch, A., Lux-Steiner, M.C., Sariciftci, N.S.: Kelvin probe force microscopy study on conjugated polymer/fullerene bulk heterojunction organic solar cells. Nano Lett. **5** (2005) 269–274

[12] Mohn, F., Gross, L., Moll, N., Meyer, G.: Imaging the charge distribution within a single molecule. Nature Nanotechnology **7** (2012) 227

[13] Gross, L., Mohn, F., Liljeroth, P., Repp, J., Giessibl, F.J., Meyer, G.: Measuring the charge state of an adatom with noncontact atomic force microscopy. Science **324** (2009) 1428–1431

[14] Binnig, G., Rohrer, H., Gerber, C., Weibel, E.: Tunneling through a controllable vacuum gap. Appl. Phys. Lett. **40** (1982) 178–180

[15] Binnig, G., Quate, C.F., Gerber, C.: Atomic force microscope. Phys. Rev. Lett. **56** (1986) 930–933

[16] Giessibl, F.J.: Advances in atomic force microscopy. Rev. Mod. Phys. **75** (2003) 949–983

[17] Albrecht, T.R., Grütter, P., Horne, D., Rugar, D.: Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity. J. App. Phys. **69** (1991) 668–673

[18] Giessibl, F.J.: Forces and frequency shifts in atomic-resolution dynamic-force microscopy. Phys. Rev. B **56** (1997) 16010–16015

[19] London, F.: The general theory of molecular forces. Trans. Faraday Soc. **33** (1937) 8–26

[20] Hamaker, H.: The londonvan der waals attraction between spherical particles. physica **4** (1937) 1058–1072

[21] Vaa, T., Bergstrom, L.: Hamaker constants of inorganic materials. Adv. Colloid Interface Sci. **70** (1997) 125–169

[22] French, R., Cannon, R., DeNoyer, L., Chiang, Y.M.: Full spectral calculation of non-retarded hamaker constants for ceramic systems from interband transition strengths. Solid State Ionics **75** (1995) 13 – 33

[23] Israelachvili, J.N.: Intermolecular and surface forces. revised 3rd edn. Academic press (2011)

[24] Argento, C., French, R.: Parametric tip model and force–distance relation for hamaker constant determination from atomic force microscopy. J. App. Phys. **80** (1996) 6081–6090

[25] Hudlet, S., Saint Jean, M., Guthmann, C., Berger, J.: Evaluation of the capacitive force between an atomic force microscopy tip and a metallic surface. Eur. Phys. J. B **2** (1998) 5–10

[26] Morse, P.M.: Diatomic molecules according to the wave mechanics. ii. vibrational levels. Phys. Rev. **34** (1929) 57–64

[27] Kantorovich, L., Foster, A., Shluger, A., Stoneham, A.: Role of image forces in non-contact scanning force microscope images of ionic surfaces. Surf. Sci. **445** (2000) 283 – 299

[28] Kohn, W., Sham, L.J.: Self-consistent equations including exchange and correlation effects. Phys. Rev. **140** (1965) A1133–A1138

[29] Hohenberg, P., Kohn, W.: Inhomogeneous electron gas. Phys. Rev. **136** (1964) B864

[30] Genovese, L., Neelov, A., Goedecker, S., Deutsch, T., Ghasemi, S.A., Willand, A., Caliste, D., Zilberberg, O., Rayson, M., Bergman, A., Schneider, R.: Daubechies wavelets as a basis set for density functional pseudopotential calculations. J. Chem. Phys. **129** (2008) 014109

[31] Landau, L.D., Lifshitz, E.M., Pitaevskii, L.P.: Electrodynamics Of Continuous Media. 2nd edn. Pergamo, Oxford (1993)

[32] Woodruff, D.P., Denchar, T.A.: Modern Techniques of Surface Science. 2nd edn. Cambridge University Press, Oxford (1994)

[33] E.Olsson, F., Persson, M. Surf. Sci. **540** (2003) 172

[34] Kitamura, S., Suzuki, K., Iwatsuki, M., Mooney, C.: Atomic-scale variations in contact potential difference on au/si(111) 77 surface in ultrahigh vacuum. Appl. Surf. Sci. **157** (2000) 222 – 227

[35] Enevoldsen, G.H., Glatzel, T., Christensen, M.C., Lauritsen, J.V., Besenbacher, F.: Atomic scale Kelvin probe force microscopy studies of the surface potential variations on the tio_2(110) surface. Phys. Rev. Lett. **100** (2008) 236104

[36] Sadewasser, S., Jelinek, P., Fang, C.K., Custance, O., Yamada, Y., Sugimoto, Y., Abe, M., Morita, S.: New insights on atomic-resolution frequency-modulation Kelvin-probe force-microscopy imaging of semiconductors. Phys. Rev. Lett. **103** (2009) 266103

[37] Bocquet, F., Nony, L., Loppacher, C., Glatzel, T.: Analytical approach to the local contact potential difference on (001) ionic surfaces: Implications for Kelvin probe force microscopy. Phys. Rev. B **78** (2008) 035410

[38] Guggisberg, M., Bammerlin, M., Loppacher, C., Pfeiffer, O., Abdurixit, A., Barwich, V., Bennewitz, R., Baratoff, A., Meyer, E., Güntherodt, H.J.: Separation of interactions by noncontact force microscopy. Phys. Rev. B **61** (2000) 11151–11155

[39] König, T., Simon, G.H., Rust, H.P., Heyde, M.: Work function measurements of thin oxide films on metalsmgo on ag(001). J. Phys. Chem. C **113** (2009) 11301–11305

[40] Gross, L., Mohn, F., Liljeroth, P., Repp, J., Giessibl, F.J., Meyer, G.: Measuring the charge state of an adatom with noncontact atomic force microscopy. Science **324** (2009) 1428–1431

[41] Giessibl, F.J.: Atomic resolution of the silicon (111)-(7x7) surface by atomic force microscopy. Science **267** (1995) 68–71

[42] Kikukawa, A., Hosaka, S., Imura, R.: Vacuum compatible high-sensitive Kelvin probe force microscopy. Rev. Sci. Instrum. **67** (1996) 1463–1467

[43] Kitamura, S., M, I.: High-resolution imaging of contact potential difference with ultrahigh vacuum noncontact atomic force microscope. Appl. Phys. Lett. **72** (1998) 3154–3156

[44] Okamoto, K., Sugawara, Y., Morita, S.: The elimination of the artifact in the electrostatic force measurement using a novel noncontact atomic force microscope/electrostatic force microscope. Appl. Surf. Sci. **188** (2002) 381 – 385

[45] Okamoto, K., Yoshimoto, K., Sugawara, Y., Morita, S. Appl. Surf. Sci. **210** (2003) 128

[46] Krok, F., Sajewicz, K., Konior, J., Goryl, M., Piatkowski, P., Szymonski, M.: Lateral resolution and potential sensitivity in Kelvin probe force microscopy: Towards understanding of the sub-nanometer resolution. Phys. Rev. B **77** (2008) 235427

[47] Dürig, U.: Relations between interaction force and frequency shift in large-amplitude dynamic force microscopy. Appl. Phys. Lett. **75** (1999) 433–435

[48] Sader, J.E., Jarvis, S.P.: Accurate formulas for interaction force and energy in frequency modulation force spectroscopy. Appl. Phys. Lett. **84** (2004) 1801–1803

[49] Arai, T., Tomitori, M.: Observation of electronic states on Si(111)-(7 × 7) through short-range attractive force with noncontact atomic force spectroscopy. Phys. Rev. Lett. **93** (2004) 256101

[50] Arai, T., Tomitori, M.: Electric conductance through chemical bonding states being formed between a Si tip and a Si(111)-(7 × 7) surface by bias-voltage noncontact atomic force spectroscopy. Phys. Rev. B **73** (2006) 073307

[51] Lantz, M.A., Hug, H.J., Hoffmann, R., van Schendel, P.J.A., Kappenberger, P., Martin, S., Baratoff, A., Güntherodt, H.J.: Quantitative measurement of short-range chemical bonding forces. Science **291** (2001) 2580–2583

[52] Sugimoto, Y., Nakajima, Y., Sawada, D., Morita, K., Abe, M., Morita, S.: Simultaneous AFM and STM measurements on the Si(111)-(7 × 7) surface. Phys. Rev. B **81** (2010) 245322

[53] Kinoshita, Y., Naitoh, Y., Li, Y.J., Sugawara, Y.: Fabrication of sharp tungsten-coated tip for atomic force microscopy by ion-beam sputter deposition. Rev. Sci. Instr. **82** (2011) 113707

[54] Guggisberg, M., Bammerlin, M., Baratoff, A., Lthi, R., Loppacher, C., Battiston, F., L, J., Bennewitz, R., Meyer, E., Gntherodt, H.J.: Dynamic force microscopy across steps on the si(111)-(77) surface. Surf. Sci. **461** (2000) 255 – 265

[55] Weymouth, A.J., Wutscher, T., Welker, J., Hofmann, T., Giessibl, F.J.: Phantom force induced by tunneling current: A characterization on si(111). Phys. Rev. Lett. **106** (2011) 226801

[56] Glatzel, T., Sadewasser, S., Lux-Steiner, M.: Amplitude or frequency modulation-detection in Kelvin probe force microscopy. Appl. Surf. Sci. **210** (2003) 84 – 89

[57] Zerweck, U., Loppacher, C., Otto, T., Grafström, S., Eng, L.M. Phys. Rev. B **71** (2005) 125424

[58] Glatzel, T., Zimmerli, L., Koch, S., Such, B., Kawai, S., Meyer, E.: Determination of effective tip geometries in Kelvin probe force microscopy on thin insulating films on metals. Nanotechnology **20** (2009) 264016

[59] Hochwitz, T., Henning, A. K. ND Levey, C., Daghlian, C., Slinkman, J. J. Vac. Sci. Technol. B **14** (1996) 457

[60] Jacobs, H.O., Leuchtmann, P., Homan, O.J., Stemmer, A. J. Appl. Phys. **84** (1998) 1168

[61] Belaidi, S., Lebon, F., Girard, P., Leveque, G., Pagano, S. Appl. Phys. A **66** (1998) S239

[62] Colchero, J., Gil, A., Baró, A.M.: Resolution enhancement and improved data interpretation in electrostatic force microscopy. Phys. Rev. B **64** (2001) 245403

[63] Gómez-Moñivas, S., Froufe, L.S., Carminati, R., Greffet, J.J., Sáenz, J.J.: Tip-shape effects on electrostatic force microscopy resolution. Nanotechnology **12** (2001) 496

[64] Strassburg, E., Boag, A., Rosenwaks, Y.: Reconstruction of electrostatic force microscopy images. Rev. Sci. Instrum. **76** (2005) 083705

[65] Konior, J.: Vertical and lateral electrostatic forces in a tip-plane system studied with a green function plus surface charge method. J. Appl. Phys. **101** (2007) 084907

[66] Shen, Y., Barnett, D.M., Pinsky, P.M. Rev. Sci. Instrum. **79** (2008) 023711

[67] Baier, R., Leendertz, C., Lux-Steiner, M.C., Sadewasser, S.: Toward quantitative Kelvin probe force microscopy of nanoscale potential distributions. Phys. Rev. B **85** (2012) 165436

[68] Masago, A., Tsukada, M., Shimizu, M.: Simulation method of Kelvin probe force microscopy at nanometer range and its application. Phys. Rev. B **82** (2010) 195433

[69] Tsukada, M., Masago, A., Shimizu, M.: Theoretical simulation of Kelvin probe force microscopy for Si surfaces by taking account of chemical forces. J. Phys.: Cond. Matt. **24** (2012) 084002

[70] Nony, L., Bocquet, F., Loppacher, C., Glatzel, T.: On the relevance of the atomic-scale contact potential difference by amplitude-modulation and frequency-modulation Kelvin probe force microscopy. Nanotechnology **20** (2009) 264014

[71] Nony, L., Foster, A.S., Bocquet, F., Loppacher, C.: Understanding the atomic-scale contrast in Kelvin probe force microscopy. Phys. Rev. Lett. **103** (2009) 036802

[72] Hoffmann, R., Kantorovich, L.N., Baratoff, A., Hug, H.J., Güntherodt, H.J.: Sublattice identification in scanning force microscopy on alkali halide surfaces. Phys. Rev. Lett. **92** (2004) 146103

[73] Ruschmeier, K., Schirmeisen, A., Hoffmann, R.: Atomic-scale force-vector fields. Phys. Rev. Lett. **101** (2008) 156102

[74] de Wette, F.W., Kress, W., Schröder, U.: Relaxation of the rocksalt (001) surface: Alkali halides, MgO, and PbS. Phys. Rev. B **32** (1985) 4143–4157

[75] Bennewitz, R., Foster, A.S., Kantorovich, L.N., Bammerlin, M., Loppacher, C., Schär, S., Guggisberg, M., Meyer, E., Shluger, A.L.: Atomically resolved edges and kinks of nacl islands on cu(111): Experiment and theory. Phys. Rev. B **62** (2000) 2074–2084

[76] Barth, C., Hynninen, T., Bieletzki, M., Henry, C.R., Foster, A.S., Esch, F., Heiz, U.: AFM tip characterization by Kelvin probe force microscopy. New J. Phys. **12** (2010) 093024

[77] Bocquet, F., Nony, L., Loppacher, C.: Polarization effects in noncontact atomic force microscopy: A key to model the tip-sample interaction above charged adatoms. Phys. Rev. B **83** (2011) 035411

[78] Kalinin, S., Gruverman, A.: Scanning probe microscopy: electrical and electromechanical phenomena at the nanoscale. Springer, New York (2007) Vols. I and II.

[79] Weaver, J.M.R., Abraham, D.W.: High resolution atomic force microscopy potentiometry. J. Vacuum Sci. Technol. B **9** (1991) 1559–1561

[80] Nonnenmacher, M., O'Boyle, M.P., Wickramasinghe, H.K.: Kelvin probe force microscopy. Appl. Phys. Lett. **58** (1991) 2921–2923

[81] Kopanski, J., Marchiando, J., Lowney, J.: Scanning capacitance microscopy measurements and modeling: Progress towards dopant profiling of silicon. J. Vac. Sci. Tech. B **14** (1996) 242–247

[82] Suddards, M., Baumgartner, A., Henini, M., Mellor, C.: Scanning capacitance imaging of compressible and incompressible quantum hall effect edge strips. New J. Phys. **14** (2012) 083015

[83] Topinka, M., LeRoy, B., Shaw, S., Heller, E., Westervelt, R., Maranowski, K., Gossard, A.: Imaging coherent electron flow from a quantum point contact. Science **289** (2000) 2323–2326

[84] Pioda, A., Kičin, S., Ihn, T., Sigrist, M., Fuhrer, A., Ensslin, K., Weichselbaum, A., Ulloa, S., Reinwald, M., Wegscheider, W.: Spatially resolved manipulation of single electrons in quantum dots using a scanned probe. Phys. Rev. Lett. **93** (2004) 216801

[85] Sacha, G.M., Sahagún, E., Sáenz, J.J.: A method for calculating capacitances and electrostatic forces in atomic force microscopy. J. Appl. Phys. **101** (2007) 024310

[86] Sadeghi, A., Baratoff, A., Ghasemi, S.A., Goedecker, S., Glatzel, T., Kawai, S., Meyer, E.: Multiscale approach for simulations of Kelvin probe force microscopy with atomic resolution. Phys. Rev. B **86** (2012) 075407

[87] Jackson, J.D.: Classical Electrodynamics. Wiley, New York (2001)

[88] Wilk, G.D., Wallace, R.M., Anthony, J.M.: High-kappa gate dielectrics: Current status and materials properties considerations. J. Appl. Phys. **89** (2001) 5243–5275

[89] Smythe, W.R.: Static and dynamic electricity. 2nd edn. McGraw-Hill, New York (1950)

[90] Gómez-Moñivas, S., Froufe-Pérez, L.S., Caamaño, A.J., Sáenz, J.J.: Electrostatic forces between sharp tips and metallic and dielectric samples. App. Phys. Lett. **79** (2001) 4048–4050

[91] Lyuksyutov, S.F., Sharipov, R.A., Sigalov, G., Paramonov, P.B. arXiv:cond-mat/0408247v2 (2004)

[92] Shen, Y., Barnett, D.M., Pinsky, P.M.: Modeling electrostatic force microscopy for conductive and dielectric samples using the boundary element method. Eng. Anal. Bound. Elem. **32** (2008) 682 – 691

[93] Belaidi, S., Girard, P., Leveque, G.: Electrostatic forces acting on the tip in atomic force microscopy: Modelization and comparison with analytic expressions. J. Appl. Phys. **81** (1997) 1023–1030

[94] Sacha, G.M., Sáenz, J.J.: Cantilever effects on electrostatic force gradient microscopy. Appl. Phys. Lett. **85** (2004) 2610–2612

[95] Elias, G., Glatzel, T., Meyer, E., Schwarzman, A., Boag, A., Rosenwaks, Y.: The role of the cantilever in Kelvin probe force microscopy measurements. Beilstein J. Nanotechnol. **2** (2011) 252–260

[96] Sadewasser, S., Glatzel, T., Shikler, R., Rosenwaks, Y., Lux-Steiner, M.: Resolution of Kelvin probe force microscopy in ultrahigh vacuum: comparison of experiment and simulation. Appl. Surf. Sci. **210** (2003) 32 – 36

[97] Valdrè, G., Moro, D.: 3d finite element analysis of electrostatic deflection and shielding of commercial and fib-modified cantilevers for electric and Kelvin force microscopy: Ii. rectangular shaped cantilevers with asymmetric pyramidal tips. Nanotechnology **19** (2008) 405502

[98] Charrier, D.S.H., Kemerink, M., Smalbrugge, B.E., de Vries, T., Janssen, R.A.J.: Real versus measured surface potentials in scanning Kelvin probe microscopy. ACS Nano **2** (2008) 622–626

[99] Schenk, O., Bollhöfer, M., Römer, R.A.: On large scale diagonalization techniques for the Anderson model of localization. SIAM Review **50** (2008) 91–112

[100] Schenk, O., Wächter, A., Hagemann, M.: Matching-based preprocessing algorithms to the solution of saddle-point problems in large-scale nonconvex interior-point optimization. Comput. Optim. Appl. **36** (2007) 321–341

[101] http://pages.unibas.ch/comphys/comphys/software.

[102] Hartwigsen, C., Goedecker, S., Hutter, J.: Relativistic separable dual-space gaussian pseudopotentials from h to rn. Phys. Rev. B **58** (1998) 3641–3662

[103] Genovese, L., Deutsch, T., Neelov, A., Goedecker, S., Beylkin, G.: Efficient solution of poisson's equation with free boundary conditions. J. Chem. Phys. **125** (2006) 074105

[104] Goedecker, S.: Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. J. Chem. Phys. **120** (2004) 9911–9917

[105] Pou, P., Ghasemi, S.A., Jelinek, P., Lenosky, T., Goedecker, S., Perez, R.: Structure and stability of semiconductor tip apexes for atomic force microscopy. Nanotechnology **20** (2009) 264015

[106] Ghasemi, S.A.: Atomistic Simulations of Atomic Force Microscopy. PhD thesis, University of Basel, Basel, Switzerland (2010)

[107] Teobaldi, G., Lämmle, K., Trevethan, T., Watkins, M., Schwarz, A., Wiesendanger, R., Shluger, A.L.: Chemical resolution at ionic crystal surfaces using dynamic atomic force microscopy with metallic tips. Phys. Rev. Lett. **106** (2011) 216102

[108] Pfeiffer, O.: Quantitative dynamische Kraft- und Dissipationsmikroskopie auf molekularer Skala. PhD thesis, University of Basel, Basel, Switzerland (2004)

[109] Sader, J.E., Sugimoto, Y.: Accurate formula for conversion of tunneling current in dynamic atomic force spectroscopy. Appl. Phys. Lett. **97** (2010) 043502

[110] Kawai, S., Glatzel, T., Koch, S., Baratoff, A., Meyer, E.: Interaction-induced atomic displacements revealed by drift-corrected dynamic force spectroscopy. Phys. Rev. B **83** (2011) 035421

[111] Gross, L., Mohn, F., Moll, N., Liljeroth, P., Meyer, G.: The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy. Science **325** (2009) 1110–1114

[112] Amsler, M., Ghasemi, S.A., Goedecker, S., Neelov, A., Genovese, L.: Adsorption of small nacl clusters on surfaces of silicon nanostructures. Nanotechnology **20** (2009) 445301

[113] Burson, K.M., Cullen, W.G., Shaffique, A., Dean, C., Watanabe, K., Taniguchi, T., Kim, P., Fuhrer, M.S.: Direct imaging of charged impurity density in common graphene substrates. Nano lett. (2013)

[114] Wigner, E., Bardeen, J.: Theory of the work functions of monovalent metals. Phys. Rev. **48** (1935) 84–87

[115] Krok, F., Sajewicz, K., Konior, J., Goryl, M., Piatkowski, P., Szymonski, M.: Lateral resolution and potential sensitivity in Kelvin probe force microscopy: Towards understanding of the sub-nanometer resolution. Phys. Rev. B **77** (2008) 235427

[116] Fremy, S., *et al.* (unpublished)

[117] Oganov, A.R.: Modern Methods of Crystal Structure Prediction. 1st edn. Wiley-VCH, Berlin (2010)

[118] Goedecker, S.: Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. J. Chem. Phys. **120** (2004) 9911–9917

[119] Amsler, M., Goedecker, S.: Crystal structure prediction using the minima hopping method. J. Chem. Phys. **133** (2010) 224104

[120] Neumann, M., Leusen, F., Kendrick, J.: A major advance in crystal structure prediction. Angew. Chem. Inter. Ed. **47** (2008) 2427–2430

[121] Oganov, A.R., Valle, M.: How to quantify energy landscapes of solids. J. Chem. Phys. **130** (2009) 104504

[122] Downs, G.M., Willett, P.: Similarity searching in databases of chemical structures. Rev. Comput. Chem. **7** (1996) 1–66

[123] Velasquez, E., Yera, E.R., Singh, R.: Determining molecular similarity for drug discovery using the wavelet riemannian metric. In: IEEE Symposium on BionInformatics and BioEngineering. BIBE (2006) 261–268

[124] Karakoc, E., Cherkasov, A., Sahinalp, S.C.: Distance based algorithms for small biomolecule classification and structural similarity search. Bioinfo. **22** (2006) e243–e251

[125] Kuntz, I.D., Meng, E.C., Shoichet, B.K.: Stru. based mol. design. Accounts of Chemical Research **27** (1994) 117–123

[126] Downs, G.M., Willett, P., Fisanick, W.: Similarity searching and clustering of chemical-structure databases using molecular property data. J. Chem. Info. Comp. Sci. **34** (1994) 1094–1102

[127] Zhang, Y.: Progress and challenges in protein structure prediction. Curr. Opin. Stru. Bio. **18** (2008) 342 – 348

[128] Gillet, V.J., Wild, D.J., Willett, P., Bradshaw, J.: Similarity and dissimilarity methods for processing chemical structure databases. Comp. J. **41** (1998) 547–558

[129] Sheridan, R.P., Kearsley, S.K.: Why do we need so many chemical similarity search methods? Drug disc. today **7** (2002) 903–911

[130] Ponec, R., Amat, L., Carb-dorca, R.: Molecular basis of quantitative structure-properties relationships (qspr): A quantum similarity approach. J. Comp. Aided Mol. Design **13** (1999) 259–270

[131] Lemmen, C., Lengauer, T.: Computational methods for the structural alignment of molecules. J. Comp. Aided Mol. Design **14** (2000) 215–232

[132] Flower, D.R.: On the properties of bit string-based measures of chemical similarity. J. Chem. Inf. Comp. Sci. **38** (1998) 379–386

[133] Allen, B.C.P., Grant, G.H., Richards, W.G.: Similarity calculations using two-dimensional molecular representations. J. Chem. Info. Comp. Sci. **41** (2001) 330–337

[134] Schwarzer, F., Lotan, I.: Approximation of protein structure for fast similarity measures. RECOMB '03, New York, NY, USA, ACM (2003) 267–276

[135] Rupp, M., Tkatchenko, A., Müller, K.R., von Lilienfeld, O.A.: Fast and accurate modeling of molecular atomization energies with machine learning. Phys. Rev. Lett. **108** (2012) 058301

[136] Todeschini, R., Consonni, V.: Molecular descriptors for chemoinformatics. Wiley-VCH Weinheim (2009)

[137] Martin, R.L., Smit, B., Haranczyk, M.: Addressing challenges of identifying geometrically diverse sets of crystalline porous materials. J. Chem. Inf. Model. **52** (2012) 308–318

[138] Cheng, J., Fournier, R.: Structural optimization of atomic clusters by tabu search in descriptor space. Theor. Chem. Acc. **112** (2004) 7–15

[139] Lee, J., Lee, I.H., Lee, J.: Unbiased global optimization of lennard-jones clusters for $n \leq 201$ using the conformational space annealing method. Phys. Rev. Lett. **91** (2003) 080201

[140] Valle, M., Oganov, A.R.: Crystal fingerprint space a novel paradigm for studying crystal-structure sets. Acta Crystallographica Section A **66** (2010) 507–517

[141] Kabsch, W.: A discussion of the solution for the best rotation to relate two sets of vectors. Acta Cryst. **A34** (1978) 827–828

[142] Horn, B.K.P., Hilden, H., Negahdaripour, S.: Closed-form solution of absolute orientation using orthonormal matrices. J. Optic. Soc. Am. **5** (1988) 1127–1135

[143] Coutsias, E.A., Seok, C., Dill, K.A.: Using quaternions to calculate rmsd. J. Comp. Chem. **25** (2004) 1849–1857

[144] Theobald, D.L.: Rapid calculation of RMSDs using a quaternion-based characteristic polynomial. Acta Crystallographica Section A **61** (2005) 478–480

[145] Kuhn, H.W.: The hungarian method for the assignment problem. Naval Res. Logistics Quart. **2** (1955) 83–97

[146] Helmich, B., Sierka, M.: Similarity recognition of molecular structures by optimal atomic matching and rotational superposition. J. Comp. Chem. **33** (2012) 134–140

[147] Carpaneto, G., Martello, S., Toth, P.: Algorithms and codes for the assignment problem. Anna. Oper. Research **13** (1988) 191–223

[148] Patterson, A.: Homometric structures. Nature (London) **143** (1939) 939–940

[149] Grigoryan, V.G., Springborg, M.: Structure and energetics of ni clusters with up to 150 atoms. Chem. Phys. Lett. **375** (2003) 219–226

[150] Benson, G., Dempsey, E.: The cohesive and surface energies of some crystals possessing the fluorite structure. Proc. R. Soc. A. **266** (1962) 344–358

[151] Pietrucci, F., Andreoni, W. Phys. Rev. Lett. **107** (2011) 085504

[152] Steinhardt, P.J., Nelson, D.R., Ronchetti, M.: Bond-orientational order in liquids and glasses. Phys. Rev. B **28** (1983) 784–805

[153] Bartók, A.P., Kondor, R., Csányi, G.: On representing chemical environments. Phys. Rev. B **87** (2013) 184115

[154] Moussa, J.E.: Comment on fast and accurate modeling of molecular atomization energies with machine learning. Phys. Rev. Lett. **109** (2012) 059801

[155] Clementi, E., Davis, D.: Electronic structure of large molecular systems. J. Comp. Phys. **1** (1966) 223 – 244

[156] Obara, S., Saika, A.: General recurrence formulas for molecular integrals over cartesian gaussian functions. J. Chem. Phys. **89** (1988) 1540–1559

[157] Field, M.J.: A practical introduction to the simulation of molecular systems. Cambridge University Press (1999)

Publications

- Sadeghi, A., Baratoff, A., Ghasemi, S.A., Goedecker, S., Glatzel, T., Kawai, S., Meyer, E., *Multiscale approach for simulations of Kelvin probe force microscopy with atomic resolution.* Phys. Rev. B **86**, 075407 (2012)

- Willand, A., Kvashnin, Y.O., Genovese, L., Vázquez-Mayagoitia, Á., Deb, A.K., Sadeghi, A., Deutsch, T., Goedecker, S., *Norm-conserving pseudopotentials with chemical accuracy compared to all-electron calculations.* J. Chem. Phys. **138**, 104109 (2013)

- Sadeghi, A., Baratoff, A., Goedecker, S., *Electrostatic interactions with dielectric samples in scanning probe microscopies.* Phys. Rev. B **88**, 035436 (2013)

- Kawai, S., Sadeghi, A., Feng, X. Lifen, P., Pawlak, R., Glatzel, T., Willand, A., Orita, A., Otera, J., Goedecker, S., Meyer, E., *Obtaining detailed structural information about supramolecular systems on surfaces by combining high-resolution force microscopy with ab initio calculations.* ACS Nano **7**, 9098 (2013)

- Mohr, S., Pochet, P., Amsler, M., Schaefer, B., Sadeghi, A., Genovese, L. Goedecker, S., *Boron aggregation in the ground states of boron-carbon fullerenes.* arXiv:1305.2302 (2013)

- Sadeghi, A., Ghasemi, S.A., Schaefer, B., Mohr, S., Lill, M., Goedecker, S. *Metrics for measuring distances in configuration spaces* J. Chem. Phys. **139**, 184118 (2013)