



UNIVERSITÉ CATHOLIQUE DE LOUVAIN

Faculté des Sciences Appliquées
Unité de Physico-Chimie et de Physique des Matériaux

First-Principles Molecular Dynamics Study of SiO₂: Surface and Interface with Si

Dissertation présentée en vue de l'obtention du grade de
Docteur en Sciences Appliquées

par

Gian-Marco RIGNANESE

Promoteurs: Prof. J.-P. Michenaud
Dr. X. Gonze

Membres du jury:
Prof. R. Car
Dr. J.-C. Charlier
Prof. J.-P. Issi
Prof. Ph. Lambin
Prof. J.-P. Vigneron

Juin 1998

Acknowledgments – Remerciements

Ces quelques lignes sont dédiées à tous ceux et celles qui, de près ou de loin, ont contribué à l'aboutissement de cette thèse. Je voudrais leur témoigner toute ma gratitude pour le soutien qu'ils m'ont apporté tout au long de ces quatre années tant d'un point de vue scientifique que d'un point de vue humain.

Je tiens tout d'abord à remercier le Professeur J.-P. Michenaud qui m'a accueilli au sein du PCPM (laboratoire de Physico-Chimie et Physique des Matériaux). La confiance qu'il m'a témoignée en acceptant d'être mon promoteur de thèse et qu'il m'a renouvelée au quotidien a joué un rôle important dans l'aboutissement de mon travail de recherche. L'ambiance qu'il a contribué à créer au laboratoire, alliant rigueur scientifique et convivialité, fut également un des ingrédients de la réussite.

Ma reconnaissance s'adresse particulièrement au Docteur X. Gonze, également promoteur de cette thèse, qui m'a introduit dans le groupe *ab initio* du PCPM. Je le remercie pour l'amitié qu'il m'a portée, la disponibilité et la patience dont il a fait preuve à mon égard. Sa compétence, ses conseils, ses critiques ont largement contribué à ma formation scientifique. Très présent tout au long de mes recherches, il a su me laisser la liberté voulue pour m'amener à voler de mes propres ailes.

Ma gratitude va également au Professeur R. Car qui m'a accueilli dans son groupe à l'IRRMA (Institut Romand de Recherche Numérique en Physique de Matériaux) pendant près de neuf mois. Ce séjour a été un tournant décisif pour l'avancement de mes recherches. Je lui suis très reconnaissant de m'avoir donné accès tant au Cray T3D du projet PATP (collaboration entre CRAY Research Suisse et l'Ecole Polytechnique Fédérale de Lausanne), qu'à la NEC-SX4 du Centre Suisse pour le Calcul Scientifique (CSCS) de Manno, et ce même après mon retour en Belgique. Je voudrais également le remercier pour sa présence dans ce jury et ces conseils précieux pour la rédaction de ce manuscrit.

Mes remerciements s'adressent particulièrement au Docteur J.-C. Charlier. En me parlant pour la première fois du calcul *ab initio* alors que je n'étais encore qu'en seconde candidature, il a orienté le cours de mon existence. Lorsque le doute naissait en moi, il m'a écouté et conseillé me permettant ainsi d'aller de l'avant. Boute-en-train sans pareil, il a égayé le quotidien de notre travail tant à Louvain-la-Neuve qu'à Lausanne pour

mon plus grand plaisir. Je lui suis d'ailleurs très reconnaissant pour l'accueil chaleureux que lui et sa petite famille m'ont réservé lors de mon séjour en Suisse. La liste des choses pour lesquelles je voudrais le remercier est trop longue pour pouvoir être exhaustif. Merci Jean-Christophe, Mireille, Ilona et Elise...

Je tiens à remercier le Prof. J.-P. Issi qui a accepté de présider ce jury. Grâce à lui, ainsi qu'au Professeur J.-P. Michenaud, le laboratoire PCPM est reconnu non seulement pour sa recherche scientifique, mais également pour ses qualités humaines.

Ma gratitude s'adresse également au Prof. Ph. Lambin pour sa disponibilité et les conseils avisés qu'il m'a prodigués au cours de ces quatre années. Je voudrais aussi le remercier pour sa présence dans ce jury.

Ma reconnaissance va aussi au Prof. J.-P. Vigneron pour le soutien qu'il m'a apporté tout au long de l'élaboration de cette thèse. Je le remercie en outre d'avoir accepté de faire partie de mon comité d'encadrement et de ce jury.

Le présent travail n'aurait pu s'effectuer sans l'appui financier du Fonds National de la Recherche Scientifique, auquel j'exprime ma gratitude. Mes remerciements vont également au Ministère de l'Education, de la Recherche et de la Formation de la Communauté Française de Belgique pour la bourse de voyage qui m'a été octroyée en vue d'un séjour de 3 mois en Suisse. Je voudrais aussi remercier CRAY Research Suisse et l'EPFL qui ont financé le reste de ce séjour.

Durant ces quatre années, le plaisir de travailler au sein de l'équipe *ab initio* du PCPM fut un puissant moteur pour ma recherche. Je tiens à remercier François, Jean-Michel et Philippe pour tous ces moments passés ensemble, pour leur disponibilité et leur ouverture d'esprit. Plus particulièrement, je remercie Jean-Michel pour ses connaissances informatiques qui m'ont été d'une grande aide et dont il m'a transmis une infime partie.

Le laboratoire PCPM est une grande famille en permanente évolution, qui fut un cadre d'épanouissement scientifique et humain incomparable pour la maturation de ce travail. C'est avec plaisir que je remercie Andrei, Arnaud, Aurel, Bernard, Brigitte, Cédric, Christophe, Claude, Didier, Eusebiu, Fabienne, Farida, Guy, Jean-Benoît, Jean-Luc, Laurent, Luc L., Luc M., Luc P., Lucian, Michel, Patrick, Paul-André, Rita, Sadia, Sorin, Stoyka, Sylvain, Véronique, Vincent, Virginie, Xavier.

Ma reconnaissance va également aux docteurs X. Blase, A. Canning, A. Dal Corso, A. De Vita, G. Galli, F. Gygi, A. Pasquarello pour leur accueil au sein de l'IRRMA lors de mon séjour en Suisse et les très bonnes relations que nous avons conservées après mon retour.

Je voudrais aussi remercier tous mes amis et amies pour les bons moments passés ensemble et le soutien qu'il m'ont apporté dans les moments difficiles. Merci à Barbara, Bruno, Cécile, Corinne, Emile, Denis, Florence, Fred B., Fred P., Hélène, Jean-Louis, Jean-Philippe, Luc, Luca, Martine, Pascal, Valérie, Virginie, Walter.

La présentation de cette thèse doit également beaucoup à mes parents. Je voudrais leur exprimer ici toute ma reconnaissance pour leur amour et leur compréhension. Ils m'ont aidé à trouver ma voie et m'ont encouragé à m'y investir pour me réaliser pleinement. Je remercie aussi mon frère, ma soeur et mon beau-frère pour leurs attentions et leur soutien quotidiens. J'aimerais pour eux dédier cette thèse à Chiara...

Contents

Acknowledgments – Remerciements	1
Contents	5
Introduction	7
1 First-principles molecular dynamics	13
1.1 Introduction	13
1.2 Molecular Dynamics	15
1.2.1 The equations of motion	16
1.2.2 Algorithms for molecular dynamics	17
1.2.3 Modified equations of motion	19
1.2.4 Thermodynamic constraints	20
1.3 First-principles Interatomic potential	28
1.3.1 Conditions and assumptions	28
1.3.2 Density Functional Theory	29
1.3.3 Solution of the Kohn-Sham equations	32
1.4 First-principles Molecular Dynamics	35
1.4.1 Introduction	35
1.4.2 The Car-Parrinello method	36
1.5 Practical implementation	39
1.5.1 Electron-electron interaction	40
1.5.2 Periodic cells and Plane waves	45
1.5.3 Electron-ion interactions	51
1.5.4 Ion-ion interactions	63
1.6 Conclusion	64
2 The nitrated Si(001)-SiO₂ interface	67
2.1 Introduction	67
2.2 Core-level shifts calculation	73

2.2.1	Theoretical approach	74
2.2.2	Small molecules	76
2.3	Core-level shifts at the Si(001)-SiO ₂ nitrided interface	79
2.3.1	Auxiliary test molecules	80
2.3.2	Structural model	80
2.3.3	N 1s core-level shifts at the interface	91
2.3.4	Si 2p core-level shifts at the interface	94
2.4	Discussion and Conclusions	98
3	The SiO₂ α-Quartz (0001) Surface	101
3.1	Introduction	101
3.2	The dry surface	104
3.2.1	The bulk structure of α -quartz	104
3.2.2	The fracture of α -quartz	107
3.2.3	The cleaved surface	108
3.2.4	The dimer surface	109
3.2.5	The VAP surface	115
3.2.6	The dense surface	116
3.3	The wet surface	128
3.3.1	Dehydration of wet surfaces	130
3.3.2	Hydration of the dense surface	131
3.3.3	Wetting of partially hydroxylated surfaces	135
3.4	Conclusion	137
	Conclusion	141
	Appendix A: Beyond the LDA	145
	Bibliography	151

Introduction

SiO_2 , also referred to as silicon dioxide or silica, is one of the most abundant material on earth. It exists all around the world in amorphous or glassy phases as well as in a variety of crystalline polymorphs. The crystalline forms of SiO_2 occur mostly as quartz (Fig. 0.1), sand, and flint, but also in igneous, metamorphic and sedimentary rocks. Moreover, the earth's mantle is composed predominantly of silica. The amorphous silica can be found in nature as well: in the skeletons of radiolaria, diatoms, and sponges, for instance. Most glasses are primarily amorphous SiO_2 , with other substances in small quantities imparting characteristic properties and color.

SiO_2 is also produced industrially, and plays a critical role in nowadays technology. For instance, silica and silicon are the heart and soul of the semiconductor industry which has given us the myriad of electronic devices that have become so much part of our lives. In fact, the quality and properties of thermal SiO_2 , as compared with those of the oxides that form on other semiconductors, are largely responsible for the dominance of Si in semiconductor technology. The cost-effectiveness of this silicon based technology completes the ingredient list that made it become one of the most spectacular industrial successes of all times. Even today there are no tangible signs of decline of this technology and its potential is expected to remain unmatched for at least another decade. Similarly, the success of glass technology can be attributed to the high transmission of optical wavelengths in fibers and their low production cost. The rapidity of information transfers, which has become crucial in modern society, requires the massive installation of optical fiber networks, making of glass manufacturing a highly competitive industrial field.

Because of its use in technology and its abundance in nature, silica has been the object of extensive studies and research in many areas of science: physics, chemistry, crystallography, geology, geophysics, mineralogy and others. In all these areas, important contributions to the understanding of the microscopic structure and dynamics of silica have been made through



Figure 0.1: Quartz crystal.

molecular dynamics computer simulations. Indeed, these techniques, which have been developed in the last four decades, permit detailed analysis at the atomic scale.

The methods for atomic simulation that have been used in the framework of the study of silica fall into two categories: classical interatomic potential methods, which do not include the explicit treatment of the electrons, and direct quantum simulations. However, so far most investigations relied on empirical classical potentials, while first-principles approaches were limited to the crystalline forms of SiO_2 . Indeed, the elevated computational cost of these schemes has prevented their application to non-periodic silica systems (amorphous phases, surfaces, and interfaces). Although classical approaches have provided an adequate description of structural properties, their accuracy in treating properties in which the electronic structure plays a more subtle role is limited. Therefore, despite the numerous investigations on the topic, a series of problems in these systems remain to be understood.

In the present work, we used a first-principles approach to study two peculiar aspects of silica: its interface with silicon and its surface.

The first-principles approach used through out this study is the Car-Parrinello method. It is based on the molecular dynamics technique and thus involves solving the equations of motion for a system of interacting atoms. Each atom moves according to the force on it caused by all the other atoms. In the Car-Parrinello scheme, these forces are determined directly from the electronic structure of the system independently of any empirical parameter by means of the Density Functional Theory, and are thus very accurate over a wide range of bonding situations. This method has been applied successfully to a large number of systems, providing detailed information on

electronic as well as structural properties.

The SiO₂ interface with silicon is particularly relevant in the framework of very large scale integration (VLSI) of metal-oxide-semiconductor field effect transistors (MOSFET's). Indeed, given the small size of the devices (currently less than 100 Å) and the trend towards further miniaturization, the understanding of the microscopical structure of the interface has become critical to improve the performances. In this regard, the incorporation of a low concentration of N atoms near the Si(001)-SiO₂ interface (which is the technologically relevant SiO₂ interface with silicon) has recently drawn considerable attention. It appears as one of the most promising ways to match industrial programs requiring high-quality ultrathin gate dielectrics for VLSI. While the electrical properties are essentially as good as for conventional oxides, silicon nitrided oxides suppress the dopant diffusion across the oxide-silicon interface, and are therefore particularly suited for thin SiO₂ gates. The ultimate goal is to incorporate the nitrogen-rich layer at any chosen position within the dielectric, and this relies on a good understanding of the bonding chemistry of the N atoms. Core-level photoelectron spectroscopy has been one of the principal experimental tools for the investigation of the interface. When applied to the N 1s core-level, this technique is sensitive to the concentration of N atoms and can provide distribution profiles when resolved as a function of take-off angle or used in conjunction with chemical etching. These profiles depend on the growth process, but an accumulation of N atoms is often observed at the interface. The N 1s photoemission spectra also provide information on the bonding environment of the incorporated N atoms. The experimental spectra show a broad principal peak approximately at the same energy as in bulk Si₃N₄, suggesting the occurrence of N-Si₃ configurations. However, the asymmetric shape of this peak as well as its anomalous shift with oxide thickness are still poorly understood. The presence of a second, well separated, peak is also debated at this stage, together with possible N bonding configurations involving oxygen atoms.

We study the incorporation of nitrogen at the Si(001)-SiO₂ interface using a first-principle approach for the calculation of core-level shifts. This scheme allows to take core-hole relaxation effects into account despite the use of pseudopotentials. The calculated N 1s shifts for a series of small molecules are found to be in good agreement with the corresponding experimental data, demonstrating the validity of this computational method. The approach that is used in this study is the following. We first discriminate between the possible nitrogen bonding configurations by comparing the calculated N 1s core-level shifts for a set of carefully chosen molecules with

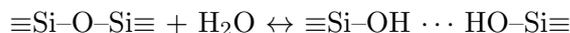
the available experimental informations about the interface. With all the pre-selected bonding configurations, we generate several model interfaces by varying the distance of the incorporated N atoms from the interface plane and by relaxing the atomic positions within First-Principles Molecular Dynamics. The N $1s$ and Si $2p$ core-level shifts are calculated for the different models and compared with experimental data, allowing for the following conclusions. First, our models give a good representation of the local structure for the various nitrogen-bonding configurations. Second, adopting the single first-neighbor configuration N-Si₃, we can explain the shift with oxide thickness of the principal XPS line and the appearance of two components, which are the main features of the N $1s$ experimental spectra. Our interpretation relies on core-hole relaxation and second nearest neighbor effects. These combine to give larger binding energies in the oxide than at the interface, in accord with experimental observations. On one side, core-hole relaxation affects N $1s$ shifts differently depending on the distance of the N atoms to the screening Si substrate. On the other side, a second nearest neighbor environment rich in Si and N atoms as at the interface induces opposite shifts compared to an O rich environment as found in the oxide. Third, our investigation of other configurations, in which the N atom is bonded to O atoms, tends to rule out their existence at the interface. Indeed, the N $1s$ shifts for N-Si₂O configurations are found at $\Delta=1.5$ eV, in a region of the spectrum where generally no peaks are observed in experiments. On the contrary, for N-Si₂H configurations, the calculated shifts are right in the range of experimental data ($\Delta=0.4$ eV). However, we believe that these configurations appear only in a very limited amount in nitrated interface, due to the high concentration of hydrogen that would be required. Finally, the calculated Si $2p$ core-level shifts are found to be in good agreement with experiments and other theoretical studies. We propose that both Si-O and Si-N bonds contribute to the suboxide peaks, contrarily to what is generally believed.

The surface of SiO₂ plays an important role in a number of technological applications. Obvious examples include adsorption, surface diffusion, nucleation, cluster growth, and thin film growth in areas such as catalysis, microelectronics, coatings, and sensor devices. Amongst the different SiO₂ surfaces, the (0001) α -quartz surface can be considered as a model surface for the different crystallographic forms of SiO₂. Besides, its study can also be profitable for the understanding of amorphous silica surface. Unfortunately, whereas the bulk structure of quartz has been widely studied, the experimental data on the structure of this surface are rather scarce. To our knowledge, there are to date only few studies dedicated to experimen-

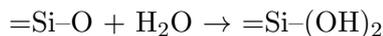
tal determination of the crystallographic structure of the (0001) α -quartz surface.

We investigate this surface using Car-Parrinello Molecular Dynamics. We start our study with two different initial configurations: the cleaved surface with non-bridging oxygens at the top and a 2×1 reconstruction with 2-membered rings. We obtain two other reconstructions by performing constant-temperature molecular dynamics simulations: the valence alternation pair surface, and the dense surface. The former presents an intimate pair of over- and under-coordinated oxygen atoms near the surface with 3-membered rings. The latter, which is the most stable of all investigated structure, presents an interesting densification of the two uppermost layers of SiO_2 tetrahedral units, with the formation of 3-membered and 6-membered rings that do not exist in bulk α -quartz. Contrarily to valence alternation pairs, whose existence in amorphous SiO_2 is well-know, the structure of the dense surface is newly proposed in this study and we suggest experiments to confirm this prediction.

Another important topic in regard to the SiO_2 surface lies in its interactions with water. These play an important role in microelectronics (wafer-bonding techniques), catalytic support, optical wave guides, and sol-gel processing. At the atomic level, the dynamics of the hydrolysis of siloxane bonds:



and non-bridging oxygens:



still lack of clarity.

We try to bring some light on the hydration/dehydration mechanism using First-Principles Molecular Dynamics. Unfortunately, at the moment of writing, many calculations are still in progress so that we present only a limited number of results. The simulation of the dehydration of the wet quartz surface implies large energy barriers (about 3 eV/surface unit cell at least) which are too high in regard to the time scale accessible to our simulation. So, we turn to the simulation of the hydration of both dry and partially hydroxylated surfaces. Our approach to the problem consists in throwing water molecules on various quartz surface using Car-Parrinello Molecular Dynamics. On one hand, the study confirms the hydrophobic behavior of siloxane bonds. Indeed, for the dense surface, we see that the water molecules rebound whatever their kinetic energies. On the other hand,

under-coordinated silicon atoms (such as those with non-bridging oxygens) are shown to be very reactive, the hydrolysis reaction being facilitated by adjacent silanol groups (that can be found on top of partially hydroxylated surfaces). We suggest that at the atomic level the formation of hydrogen-bond chains plays an important role in the hydration mechanism.

This work is organized as follows. In Chapter 1, we provide a theoretical description of the First-Principles Molecular Dynamics technique used in this study. In Chapter 2, we investigate the incorporation of nitrogen at the Si(001)-SiO₂ interface. The first-principles calculation of N 1s and Si 2p core-level shifts at this interface allows the unambiguous interpretation of photoemission results, providing microscopic information concerning the bonding structure of N at this interface. In chapter 3, we study the (0001) α -quartz surface, as model surface for other crystalline forms of SiO₂. We consider both the dry surface, as obtained in vacuum conditions, and the wet surface, as formed in normal ambient. The calculations provide informations on the structure of both crystalline and amorphous SiO₂ surfaces. Moreover, the models that are generated can be used in further studies. Finally, the simulations allow to gain insight on the mechanisms of hydration/dehydration at the microscopic level.

Chapter 1

First-principles molecular dynamics

1.1 Introduction

In recent years, considerable progress has been achieved in the scale and scope of Molecular Dynamics (MD) simulations with interatomic potentials derived from First Principles (FP). There is no doubt that much of this progress must be attributed to the original approach proposed by Car and Parrinello (1985), to the study of electronic and thermodynamic properties of condensed matter systems. This technique combines the MD scheme (see for instance Ciccotti and Hoover 1986) for the computation of statistical properties of classical systems, with the FP treatment of interatomic forces due to the quantum electronic system, as derived from Density Functional Theory (DFT). Originally developed in the solid state community, this scheme is having an ever-increasing impact on chemistry and even biochemistry.

MD is a well-established methodology which allows to determine the time evolution of any physical system for which quantum effects are not relevant. Based on Newton mechanics, such a calculation requires the simultaneous solution of a very large number of differential equations. This is obviously impossible analytically, and iterative techniques had to be developed to solve these equations numerically. Computer simulation has thus become an invaluable tool to gain insight into system microscopic behavior. The application range of Molecular Dynamics encompasses statistical mechanics, solid state physics, materials science, chemistry and biology.

In the course of the development of MD, simulations became increasingly dependent on the availability of realistic interaction potentials. In principle,

explicit knowledge of the electronic ground state at each atomic position is required, in order to have a correct description of the interatomic forces. In classical MD simulations, however no attempt is made to solve this complex many-body problem, and interactions between atoms are modeled with empirical or semi-empirical potentials.

DFT deals with the ground-state of a quantum-mechanical many-body system of electrons. In principle, it can provide the exact charge density and forces on atoms, for a chosen ionic configuration. In practice or within tractable approximations, it has been shown to give very accurate energies and atomic forces for many systems of interest in material science. Thus, there is a well-established way to calculate forces from a fundamental electronic theory, in an *ab initio* manner. However until mid-eighties, DFT energies and forces have been considered prohibitively complex, from a computational point of view, for direct applications to statistical mechanics simulations which usually require the study of a number of atoms ranging from several tens up to a few thousands, evolving through configurations.

The unified approach to DFT and MD has made it possible to use forces derived within first-principles in MD simulations. This has been accomplished by devising a way to generate at the same time ionic trajectories and the corresponding electronic ground state, and developing efficient techniques for the solution of the Schrödinger equation for the single-particle orbitals.

After this brief introduction, the body of this Chapter is organized as follows. In section 1.2, we introduce the Molecular Dynamics technique to study the evolution of a system of interacting particles. The corresponding equations of motion are developed and algorithms for their practical implementation are discussed. Modified equations of motion are proposed to find the energy minimum of the systems and also to perform dynamics under thermodynamic constraints (constant temperature, volume or pressure). In section 1.3, we outline how an interatomic potential can be derived from First Principles. On this purpose, we briefly describe Density Functional Theory and introduce the most commonly used methods to solve the self-consistent equations that derive from this scheme. In section 1.4, we present the Car-Parrinello method to perform First-Principles Molecular Dynamics (FP-MD). We analyze the reasons and conditions which make it a practical approach for MD simulations. In section 1.5, we discuss the practical implementation of FP-MD scheme. We introduce different energy functionals to describe electron-electron interactions (more specifically exchange and correlation energies) starting from the Local Density Approximation (LDA) and then going beyond it. We outline how the use of periodic cells and a

plane-wave basis set simplify the problem in the framework of Bloch's theorem by transferring it from the real to the reciprocal space. We discuss how the electron-ion interactions can be modeled by the frozen-core approximation and the use of pseudopotentials, and present the various forms (local, non-local, semi-local or separable) and the most widely used types (norm-conserving and ultrasoft) for the latter. At the end of the section, we briefly discuss the ion-ion interactions.

1.2 Molecular Dynamics

Newton mechanics allows in principle to predict the evolution of any interacting-particles system as far as quantum effects are negligible. However, such calculation requires the simultaneous solution of a huge number of differential equations even for a tiny fragment of matter: an obviously impossible task for traditional analytic techniques. Early computers allowed for the first time in history to approach these typical problems by brute force, at least for systems with hundred to thousand degrees of freedom: Molecular Dynamics was born.

The first MD calculations were mainly concerned with statistical mechanics: their goal was to evaluate the average values of physical quantities of complex systems. Alder and Wainwright (1957) studied the motion of 100 hard spheres with equal speed and different velocities. They were able to prove that momentum distribution for such a system converges rapidly to equilibrium.

The first attempt to investigate materials dates to the early work by Gibson *et al.* (1960): MD was used to investigate the dynamics of radiation damage. Their system was based on a short-range repulsive interatomic potential, plus a constant inward force responsible for the cohesion of the crystal. A single atom was endowed with a high kinetic energy, and the effect on the crystal was studied. Simulation was used as a very powerful microscope, in the paper several "photographs" of the atomic trajectories are illustrated: for the first time a window over the microscopic world was open. Through that window many researcher looked and still look, to gain qualitative insight on microscopic processes besides quantitative evaluation of thermodynamic properties.

Rahman (1964) initiated the study of systems with continuous potentials: he used 864 particles in periodic boundary conditions to simulate liquid argon. It was probably surprising to find that few particles (a very small collection, if compared with Avogadro number) can reproduce ther-

modynamic properties of real systems. It is exactly this fortunate fact which makes computer simulation of condensed matter possible. Since then many researchers followed the route opened by these vanguards, and many reports based on these techniques have been published.

In this section, we introduce the Lagrange, Newton and Hamilton equations of motion (EOM) describing the evolution of a system of interacting particles, and that are actually used in classical Molecular Dynamics. The Verlet (1967) algorithm is presented for the practical implementation of the EOM, and compared to other existing schemes. The equations of motion are then modified to provide schemes to find the energy minimum of the systems leading to steepest descent (SD) and the damped dynamics (DD). Finally, we discuss how thermodynamic constraints can be imposed in MD by means of Nosé-Hoover thermostat (constant temperature), and Parrinello-Rahman method (constant pressure).

1.2.1 The equations of motion

The molecular dynamics technique solves the classical equations of motion (EOM) of N particles interacting through a potential \mathcal{V} . The Euler-Lagrange equations of motion are of the form:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{R}}_I} \right) - \frac{\partial \mathcal{L}}{\partial \mathbf{R}_I} = 0 \quad (1.1)$$

where the classical lagrangian \mathcal{L} is defined in terms of the kinetic \mathcal{K} and potential \mathcal{V} :

$$\mathcal{L} = \mathcal{K} - \mathcal{V}. \quad (1.2)$$

The normal definitions of \mathcal{K} and \mathcal{V} give:

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^N M_I \dot{\mathbf{R}}_I^2 - \mathcal{V}(\{\mathbf{R}_I\}) \quad (1.3)$$

where M_I and \mathbf{R}_I are the mass and the position of the i^{th} particle and the interaction potential is a function of the positions of the particles: $\mathcal{V} = \mathcal{V}(\{\mathbf{R}_I\})$. By inserting Eq. (1.3) in the Euler-Lagrange EOM given by Eq. (1.1), one obtains the Newton equations of motion of the system:

$$M_I \ddot{\mathbf{R}}_I = \mathbf{F}_I \quad (1.4)$$

where \mathbf{F}_I is the force acting on the i^{th} particle. These $3N$ equations of motion are coupled since \mathbf{F}_I depends on the positions of all the particles in the system through the gradient of the potential:

$$\mathbf{F}_I = -\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I}. \quad (1.5)$$

The hamiltonian, $\mathcal{H} = \mathcal{K} + \mathcal{V}$, is a conserved variable of the motion. The conservation of \mathcal{H} is a useful test of accuracy of the numerical algorithm used to solve the EOM. By introducing the momenta \mathbf{p}_I conjugate to \mathbf{R}_I , the Hamilton equations of motion can also be derived:

$$\dot{\mathbf{R}}_I = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_I} = \frac{\mathbf{p}_I}{M_I} \quad (1.6)$$

$$\dot{\mathbf{p}}_I = -\frac{\partial \mathcal{H}}{\partial \mathbf{R}_I} = -\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I}. \quad (1.7)$$

and the hamiltonian of the system is written:

$$\mathcal{H} = \sum_{i=1}^N \frac{\mathbf{p}_I^2}{2M_I} + \mathcal{V}(\{\mathbf{R}_I\}). \quad (1.8)$$

1.2.2 Algorithms for molecular dynamics

In its most straightforward realization, molecular dynamics is a very simple technique. Given an initial configuration $\{\mathbf{R}_I\}$ of N particles at time t and their interaction potential, the resulting forces \mathbf{F}_I acting on each atom are calculated. Newton equations of motion are then solved numerically for a small time interval Δt under the assumption of constant force, to obtain the system configuration at time $t + \Delta t$. In the limit $\Delta t \rightarrow 0$ the solution is exact. The procedure can be repeated indefinitely, and the evolution of the system can therefore be followed.

A simple discretization of the corresponding equations of motion, first used by Verlet (Verlet 1967), is obtained by expanding in Taylor series the position at time $t + \Delta t$ and $t - \Delta t$:

$$\mathbf{R}_I(t + \Delta t) = \mathbf{R}_I(t) + \dot{\mathbf{R}}_I(t)\Delta t + \frac{1}{2}\ddot{\mathbf{R}}_I(t)\Delta t^2 + \mathcal{O}(\Delta t^3) \quad (1.9)$$

$$\mathbf{R}_I(t - \Delta t) = \mathbf{R}_I(t) - \dot{\mathbf{R}}_I(t)\Delta t + \frac{1}{2}\ddot{\mathbf{R}}_I(t)\Delta t^2 - \mathcal{O}(\Delta t^3). \quad (1.10)$$

By summing the two terms and rearranging, one gets

$$\mathbf{R}_I(t + \Delta t) = 2\mathbf{R}_I(t) - \mathbf{R}_I(t - \Delta t) + \ddot{\mathbf{R}}_I(t)\Delta t^2 + \mathcal{O}(\Delta t^4) \quad (1.11)$$

and by subtracting:

$$\dot{\mathbf{R}}_I(t) = \frac{\mathbf{R}_I(t + \Delta t) - \mathbf{R}_I(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^3). \quad (1.12)$$

Therefore, given the coordinates \mathbf{R}_I at time t and at time $t - \Delta t$, it is possible to calculate the configuration at time $t + \Delta t$, and the velocities at time t .

A number of simple modifications to the basic Verlet scheme such as the ‘leap-frog’ and ‘velocity-Verlet’ algorithms have been suggested to try to improve the method of handling the velocities (Allen and Tildesley 1987). There also exists another type of algorithms to solve the EOM: the Gear predictor-corrector algorithms (Gear 1966). Basically, a significant number of time derivatives of the particles positions are stored at a particular step. Taylor series expansions are then used to advance each derivative forward from time t to $t + \Delta t$, producing a series of predicted positions and their derivatives. The predicted positions can then be used to calculate the forces at $t + \Delta t$ and, through Eq. (1.4), the ‘correct’ acceleration. The difference between the correct and predicted accelerations is used with a set of Gear constants to correct all of the predicted derivatives of the motion. The Gear coefficients are chosen to optimize the stability and the accuracy of the trajectories.

The Verlet-like algorithms are simpler than the Gear algorithms, and a fifth-order predictor-corrector scheme requires at least three times more memory than the Verlet algorithm. More importantly, the Gear algorithms become rapidly unstable and inaccurate with increasing time step whereas the root-mean-square fluctuations in \mathcal{H} is almost a linear function of Δt^2 for the Verlet algorithm. Nowadays, the Verlet-like algorithms are the most widely used. Whichever algorithms is used, it can be assumed that the problem of solving Newton equations can be resolved in a more or less straightforward manner and that long trajectories can be generated in this way, i.e. that the positions and velocities are known as a function of time.

However, even in the most powerful computer the amount of computational power available is limited and the number of particles that can be simulated is finite (at most a few million particles at the time of writing), as is the length of the trajectories (up to a nanosecond). One standard way of alleviating the finite size problem is to eliminate surface effects by imposing periodic boundary conditions (PBC), i.e. a finite-size simulation

box is repeated to fill the whole space. This of course fixes an upper limit to the length scale that can be simulated. In general, only phenomena with a characteristic length of 10-50 Å can be reproduced. There are, however, no good techniques to extend the time scale limitation and although important progress has been made in recent years, the short time scale restriction is one of the most severe limitations of MD.

1.2.3 Modified equations of motion

Molecular dynamics can also be used to find local minima in the potential energy by slightly modifying the Newton EOM.

The simplest scheme consists in replacing Eq. (1.4) by first-order EOM:

$$M_I \dot{\mathbf{R}}_I = - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I}. \quad (1.13)$$

This technique is called steepest descent (SD), since the particles follow the steepest-descent direction at each time step. Although each iteration of SD algorithm moves the particles towards the minimum of the potential energy, there is no guarantee that this minimum will be reached in a finite number of iterations. In many cases, a very large number of SD steps is needed to get close to the minimum of E . This is especially the case when this minimum lies in a long narrow valley. The rate of convergence of the SD method is limited by the fact that, after the particles have moved along a given gradient direction, a subsequent displacement along the new gradient reintroduces errors proportional to the previous gradient. The conjugate-gradient (CG) scheme provides an efficient method to solve this problem. By using informations obtained from all the sampling points along the path followed by the particles, the new direction (called conjugate gradient) is generated so that each movement of the particles is independent of the previous ones in terms of minimization of the energy potential.

An alternative scheme consists in adding a dissipative term (damping factor) to the equations of motion:

$$M_I \ddot{\mathbf{R}}_I = - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} - \gamma M_I \dot{\mathbf{R}}_I. \quad (1.14)$$

This damped dynamics (DD) algorithm is quite efficient, it will be commented more deeply in the case of electrons dynamics.

1.2.4 Thermodynamic constraints

Molecular dynamics equations of motion, generated by Eq. (1.4), correspond to constant- NVE conditions (constant number of particles, volume, and total energy). In statistical mechanics, this is referred to the microcanonical ensemble. The EOM can also be modified in order to produce either constant- NVT (canonical ensemble), or constant- NPT (isobaric-isothermal ensemble) conditions.

From statistical mechanics we know that to produce a new ensemble one has simply to couple weakly the system of interest with a suitable reservoir. The way this is achieved theoretically, however, is not applicable to molecular dynamics simulation. The reservoir is usually a virtually infinite system with many more degrees of freedom than the system of interest, and the coupling is done through surfaces, becoming infinitely small in the thermodynamic limit. In simulation we really do not want to spend time integrating non-interesting degrees of freedom, and periodic boundary conditions are usually employed just to get rid of unwanted surfaces when studying bulk samples.

The key idea, proposed by Andersen (1980) is that MD reservoirs can be represented by only one or a few degrees of freedom, and the coupling can be applied uniformly to all particles. This is done by inventing new equations of motion for the extended system, composed of the system of interest plus the required reservoirs. These equations are derived by a lagrangian, or in Hamilton's canonical form, only in a virtual system of coordinates and become non-canonical when expressed in term of the real variables of the system of interest. The particular form of these equations is chosen in such a way that the microcanonical equilibrium distribution produced by a constant-energy trajectory of the extended system reduces to the desired distribution when the coordinates and momenta of the system of interest are considered, that is after integration over the extra reservoir variables.

We shall focus our attention on the method proposed by Nosé and Hoover for the constant-temperature molecular dynamics, and on the so-called Andersen and Parrinello-Rahman method for constant-pressure molecular dynamics.

Nosé-Hoover thermostat

The constant-temperature molecular dynamics method was first proposed by Nosé (1984), and later clarified by Hoover (1985) and Nosé himself (1986). The interested reader can refer to his latest review article for a more detailed

and fully comprehensive treatment (Nosé 1991). In this approach, the thermal reservoir is described by a single degree of freedom, which however does not have a simple physical meaning (Brañka and Parrinello 1986), and the coupling contains a scaling term which applies uniformly to all particle velocities and is in fact a scaling on simulation time. We will first concentrate on Hoover's approach which is simpler since it does not require the use of virtual variables and scaling concepts.

The temperature of the system can be obtained by using the energy equipartition theorem:

$$T = \frac{1}{3Nk_B} \sum_{i=1}^N M_I \dot{\mathbf{R}}_I^2. \quad (1.15)$$

It can be seen that it is directly related to the velocities and thus it can simply be adjusted to a fixed external temperature T_{ext} by introducing a variable damping in the Newton equation of motion:

$$M_I \ddot{\mathbf{R}}_I = -\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} - \dot{\chi} M_I \dot{\mathbf{R}}_I. \quad (1.16)$$

The thermodynamic friction coefficient which appears in this equation in turn evolves in time according to the following equation:

$$W_T \dot{\chi} = \sum_{i=1}^N M_I \dot{\mathbf{R}}_I^2 - 3Nk_B T_{ext} \quad (1.17)$$

where W_T is the inertia factor associated to this coefficient. It can easily be understood that the system of particles behaves like if there were a thermostat. Indeed, the driving force for the damping factor dynamics is the difference between the effective temperature of the system T (Eq. (1.15)) and the imposed external temperature T_{ext} . When the temperature of the system is bigger than T_{ext} , the friction is increased and thus the velocities tend to be reduced inducing a diminution of T , and conversely.

The corresponding Hamilton equations of motion are obtained by introducing a momentum π_χ conjugate to χ , in addition to those associated to the particles positions:

$$\dot{\mathbf{R}}_I = \frac{\mathbf{p}_I}{M_I} \quad (1.18)$$

$$\dot{\mathbf{p}}_I = -\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} - \frac{\pi_\chi}{W_T} \mathbf{p}_I \quad (1.19)$$

$$\dot{\chi} = \frac{\pi_\chi}{W_T} \quad (1.20)$$

$$\dot{\pi}_\chi = \sum_{i=1}^N \frac{\mathbf{p}_I^2}{M_I} - 3Nk_B T_{ext}. \quad (1.21)$$

Assuming ergodicity of the system, it can be shown (Hoover 1985, Nosé 1991) that this set of equations samples the canonical ensemble in the large- N limit, but may present troubles for small or stiff systems. The proof uses the conservation of probability in the phase space (generalized Liouville equation) to show that the distribution:

$$\rho(\{\mathbf{R}_I\}, \{\mathbf{p}_I\}, \chi, \pi_\chi) \propto \exp\left(\frac{\mathcal{H}_T}{k_B T_{ext}}\right) \quad (1.22)$$

is stationary.

The hamiltonian \mathcal{H}_T associated to these EOM can be written:

$$\mathcal{H}_T = \sum_{i=1}^N \frac{\mathbf{p}_I^2}{2M_I} + \mathcal{V}(\{\mathbf{R}_I\}) + \frac{\pi_\chi^2}{2W_T} + 3Nk_B T_{ext}\chi. \quad (1.23)$$

In this formulation, first proposed by Hoover (1985), the hamiltonian is a constant of motion but it is not canonical. Indeed, Eqs. (1.18-1.21) have additional force terms with respect to the canonical equations deriving from this hamiltonian.

By contrast, the Nosé formulation (1984) leads to a canonical Hamiltonian and associated canonical equations of motion. This is done by introducing a scaled time τ which is related to the real one through a virtual variable ξ in the following way:

$$d\tau = \xi dt. \quad (1.24)$$

This scaling also applies uniformly to all particles velocities since these imply time derivatives. The Nosé hamiltonian is written:

$$\mathcal{H}'_T = \sum_{i=1}^N \frac{\mathbf{p}'_I{}^2}{2M_I \xi^2} + \mathcal{V}(\{\mathbf{R}'_I\}) + \frac{\pi_\xi^2}{2W_T} + 3Nk_B T_{ext} \ln \xi. \quad (1.25)$$

It is a constant of motion $d\mathcal{H}'_T/d\tau = 0$, and it is canonical in the virtual variables \mathbf{R}'_I , \mathbf{p}'_I , ξ and π_ξ . These are related to the real variables by the following transformation:

$$\mathbf{R}'_I = \mathbf{R}_I \quad (1.26)$$

$$\mathbf{p}'_I = \xi \mathbf{p}_I \quad (1.27)$$

$$\xi = \exp \chi \quad (1.28)$$

$$\pi_\xi = \pi_\chi. \quad (1.29)$$

It can be shown this is not a canonical transformation, this is the reason why in Hoover formulation the hamiltonian is no longer canonical.

Andersen method

When considering constant-pressure molecular dynamics, two cases must be distinguished: solid systems (cell periodically to fill all the space) and molecular systems. The treatment of the latter will not be developed here. Note however that the key point lies in the application of the coupling of the volume fluctuations to the molecular center of mass rather than to each individual atomic positions (Melchionna *et al.* 1993).

In the case of solid systems, the restriction that the MD cell be kept constant in volume and in shape severely restricts the applicability of the method to problems involving structure transformations, in which these parameters play an essential role. In order to overcome this difficulty, Andersen (1980) proposed a method so as to allow for changes in volume of the MD cell containing a system of particles under external hydrostatic pressure P_{ext} .

The pressure of the system can be obtained by using the virial theorem:

$$P = \frac{1}{3V} \sum_{I=1}^N \left(M_I \dot{\mathbf{R}}_I^2 - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} \cdot \mathbf{R}_I \right). \quad (1.30)$$

The volume, considered as an additional dynamical variable, evolves as a function of the difference between the imposed external pressure P_{ext} and the internal pressure, according to the following equation:

$$W_P \ddot{V} = \frac{1}{3V} \sum_{I=1}^N \left(M_I \dot{\mathbf{R}}_I^2 - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} \cdot \mathbf{R}_I \right) - P_{ext} \quad (1.31)$$

where W_P is the inertia factor associated to V .

The Newton equation of motion of the particles is then:

$$M_I \ddot{\mathbf{R}}_I = - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} + \frac{1}{3V} \left(\ddot{V} - \frac{2\dot{V}}{3V} \right) M_I \mathbf{R}_I. \quad (1.32)$$

The corresponding Hamilton equations of motion are obtained by introducing a momentum π_V conjugate to V , in addition to those associated to the particles positions:

$$\dot{\mathbf{R}}_I = \frac{\mathbf{p}_I}{M_I} + \frac{\dot{V}}{3V}\mathbf{R}_I \quad (1.33)$$

$$\dot{\mathbf{p}}_I = -\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} - \frac{\dot{V}}{3V}\mathbf{p}_I \quad (1.34)$$

$$\dot{V} = \frac{\pi_V}{W_P} \quad (1.35)$$

$$\dot{\pi}_V = \frac{1}{3V} \sum_{I=1}^N \left(M_I \dot{\mathbf{R}}_I^2 - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} \cdot \mathbf{R}_I \right) - P_{ext}. \quad (1.36)$$

Assuming ergodicity of the system, it can be shown (Andersen 1980) that this set of equations samples the isobaric-isoshape ensemble to an accuracy of $\mathcal{O}(N^{-2})$ when calculating ensemble averages of intensive parameters ($\mathcal{O}(N^{-1})$ for extensive parameters).

The hamiltonian \mathcal{H}_P associated to these EOM can be written:

$$\mathcal{H}_P = \sum_{i=1}^N \frac{\mathbf{p}_I^2}{2M_I} + \mathcal{V}(\{\mathbf{R}_I\}) + \frac{\pi_V^2}{2W_P} + P_{ext}V. \quad (1.37)$$

In this formulation, the hamiltonian is a constant of motion but it is not canonical.

By introducing the virtual variables \mathbf{R}'_I , \mathbf{p}'_I , V' and π'_V , related to the real variables by the following transformation:

$$\mathbf{R}'_I = V^{-1/3}\mathbf{R}_I \quad (1.38)$$

$$\mathbf{p}'_I = V^{1/3}\mathbf{p}_I \quad (1.39)$$

$$V' = V \quad (1.40)$$

$$\pi'_V = \pi_V \quad (1.41)$$

the hamiltonian, that can be written:

$$\mathcal{H}'_P = \sum_{i=1}^N \frac{\mathbf{p}'_I{}^2}{2M_I V^{2/3}} + \mathcal{V}(\{V^{1/3}\mathbf{R}'_I\}) + \frac{\pi'^2_V}{2W_P} + P_{ext}V' \quad (1.42)$$

becomes canonical in the virtual variables. Besides, it remains a constant of motion $d\mathcal{H}'_P/dt = 0$.

Parrinello-Rahman method

Shortly after, Parrinello and Rhaman (1980) extended the method proposed by Andersen (1980) to allow the simulation cell to change also its shape. In this scheme, the volume variable is replaced by the 3×3 matrix variable \mathbf{h} whose columns are, in order, the components of the three lattice vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 . The volume is simply given by:

$$V = \det \mathbf{h} = \mathbf{a}_1 \cdot (\mathbf{a}_2 \wedge \mathbf{a}_3) \quad (1.43)$$

where \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 , in that order, are assumed to be a right-handed triad. The positions of the particles $\{\mathbf{R}_i\}$ are written in terms of \mathbf{h} and of their relative coordinates $\{\mathbf{S}_i\}$:

$$\mathbf{R}_i = \mathbf{h}\mathbf{S}_i = S_{i,1}\mathbf{a}_1 + S_{i,2}\mathbf{a}_2 + S_{i,3}\mathbf{a}_3. \quad (1.44)$$

The whole \mathbf{h} matrix evolves so that the pressure of the system adapts to the imposed external pressure P_{ext} :

$$W_P \ddot{\mathbf{h}} = [\sigma - P_{ext}\delta] \mathbf{h}^{-1} \det \mathbf{h} \quad (1.45)$$

where W_P is the inertia factor associated to the dynamics of \mathbf{h} and σ is the stress tensor which is obtained from the stress theorem:

$$\sigma = \frac{1}{\det \mathbf{h}} \sum_{I=1}^N \left(M_I (\mathbf{h}\dot{\mathbf{S}}_I) (\mathbf{h}\dot{\mathbf{S}}_I)^T - \frac{\partial \mathcal{V}(\{\mathbf{h}\mathbf{S}_I\})}{\partial \mathbf{S}_I} \cdot \mathbf{S}_I^T \right). \quad (1.46)$$

Note that the pressure of system can be evaluated from the stress tensor σ :

$$P = \frac{1}{3} \text{Tr} \sigma. \quad (1.47)$$

The Newton equation of motion of the particles expressed in terms of the relative coordinates $\{\mathbf{S}_I\}$ are written:

$$M_I \ddot{\mathbf{S}}_I = -\mathbf{g}^{-1} \frac{\partial \mathcal{V}(\{\mathbf{h}\mathbf{S}_I\})}{\partial \mathbf{S}_I} - M_I \mathbf{g}^{-1} \dot{\mathbf{g}} \mathbf{S}_I \quad (1.48)$$

where \mathbf{g} is the metric tensor defined by $\mathbf{g} = \mathbf{h}^T \mathbf{h}$. These can also be written in terms of the real positions $\{\mathbf{R}_I\}$:

$$\begin{aligned} M_I \ddot{\mathbf{R}}_I = & -\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} + M_I \ddot{\mathbf{h}} \mathbf{h}^{-1} \mathbf{R}_I \\ & + M_I \left[\dot{\mathbf{h}} \mathbf{h}^{-1} - (\dot{\mathbf{h}} \mathbf{h}^{-1})^T \right] (\dot{\mathbf{R}}_I - \dot{\mathbf{h}} \mathbf{h}^{-1} \mathbf{R}_I) \end{aligned} \quad (1.49)$$

The corresponding Hamilton equations of motion are obtained by introducing momenta $\pi_{\mathbf{h}}$ conjugate to \mathbf{h} , in addition to those associated to the particles positions $\{\mathbf{R}_I\}$:

$$\dot{\mathbf{R}}_I = \frac{\mathbf{p}_I}{M_I} + \dot{\mathbf{h}}\mathbf{h}^{-1}\mathbf{R}_I \quad (1.50)$$

$$\dot{\mathbf{p}}_I = -\frac{\partial\mathcal{V}(\{\mathbf{R}_I\})}{\partial\mathbf{R}_I} - \left(\dot{\mathbf{h}}\mathbf{h}^{-1}\right)^T \mathbf{p}_I \quad (1.51)$$

$$\dot{\mathbf{h}} = \frac{\pi_{\mathbf{h}}}{W_P} \quad (1.52)$$

$$\dot{\pi}_{\mathbf{h}} = [\sigma - P_{ext}\delta] \mathbf{h}^{-1} \det \mathbf{h}. \quad (1.53)$$

Assuming ergodicity of the system, it can be shown that this set of equations samples the isobaric-isothermal ensemble. Indeed, in the large- N limit, the kinetic energy associated to the variables describing the cell is small compared to that of the system of N particles.

The hamiltonian \mathcal{H}_P associated to these EOM can be written:

$$\mathcal{H}_P = \sum_{i=1}^N \frac{\mathbf{p}_I^2}{2M_I} + \mathcal{V}(\{\mathbf{R}_I\}) + \frac{\pi_{\mathbf{h}}^2}{2W_P} + P_{ext} \det \mathbf{h}. \quad (1.54)$$

In this formulation, the hamiltonian is a constant of motion but it is not canonical.

By introducing the virtual variables \mathbf{R}'_I , \mathbf{p}'_I , \mathbf{h}' and $\pi'_{\mathbf{h}}$, related to the real variables by the following transformation:

$$\mathbf{R}'_I = \mathbf{S}_I = \mathbf{h}^{-1}\mathbf{R}_I \quad (1.55)$$

$$\mathbf{p}'_I = \mathbf{h}\mathbf{p}_I \quad (1.56)$$

$$\mathbf{h}' = \mathbf{h}^T \quad (1.57)$$

$$\pi'_{\mathbf{h}} = \pi_{\mathbf{h}} \quad (1.58)$$

the hamiltonian, that can be written:

$$\mathcal{H}'_P = \sum_{i=1}^N \frac{\mathbf{p}'_I{}^T \mathbf{g}^{-1} \mathbf{p}'_I}{2M_I} + \mathcal{V}(\{\mathbf{h}\mathbf{R}'_I\}) + \frac{\pi_{\mathbf{h}}^2}{2W_P} + P_{ext} \det \mathbf{h}' \quad (1.59)$$

becomes canonical. Besides, it remains a constant of motion $d\mathcal{H}'_P/dt = 0$.

Parrinello and Rahman (1981) further extended the method to include external stress. However several authors have pointed out some shortcomings of the original method.

First, it has found (Nosé and Klein 1983) that only six of the nine variables of the matrix \mathbf{h} have an interesting meaning, when the three degrees of freedom describing the overall rotation of the MD box are extracted out. In order to avoid this spurious rotation, which would simply confuse the analysis of the trajectories of the system, two options have been considered both of which reduce the number of extra variables to six (geometrical constraints). In one case (Nosé and Klein 1983) the matrix h is kept symmetric by disregarding the asymmetric contributions in the equations of motion, and in the other one (Ferrario and Ryckaert 1985) the variables are reduced to six by fixing the MD box orientation with respect to a laboratory frame, and specifically by choosing a_1 parallel to the x -axis and a_2 to lie in the xy plane. This choice results in an upper triangular matrix h . More recently, an alternative method to get rid of the rotation of the MD cell was proposed by Lill and Broughton (1994) which consists in symmetrizing the infinitesimal strain at each time step (dynamical constraint).

Second, Wentzcovitch (1991) showed that Parrinello-Rahman equations of motion are not invariant with respect to the interchange between equivalent cells, referred to as modular transformations and defined by $\mathbf{h}' = \mathbf{M}\mathbf{h}$ with $\det \mathbf{M} = 1$). A method was also suggested to avoid this problem where the shape of the cell is described in terms of the strain tensor ϵ and a reference structure: $\mathbf{h} = \epsilon \mathbf{h}_0$.

Finally, Cleveland (1988) pointed out that the consistency between the condition of mechanical equilibrium and the virial theorem is only verified in the large- N limit for the Parrinello-Rahman dynamics. Souza and Martins (1997) proposed the metric tensor \mathbf{g} as variable to use in molecular dynamics simulations with variable-cell shape. They showed this choice is very convenient since the new equations of motion have several properties that were not present in early expressions, namely, absence of rigid rotation and invariance with respect to modular transformations. This scheme also solves the problem of consistency raised by Cleveland. Moreover, the generalization to anisotropic stress leads to much simpler equations than in the original formulation.

To summarize, we have presented the classical Molecular Dynamics technique to describe the evolution of a system of interacting particles. The underlying equations of motion have been introduced, and Verlet (1967) algorithm has been discussed with respect to other schemes for solving Newton equations numerically. Changes have been brought to the equations of

motion providing methods to locate the energy minimum of the systems namely steepest descent (SD) and damped dynamics (DD). Finally, we have formulated how MD simulations can be performed under thermodynamic constraints such as constant temperature (Nosé-Hoover thermostat), or constant pressure (Andersen and Parrinello-Rahman methods).

1.3 First-principles Interatomic potential

In conventional molecular dynamics simulations, the interactions between atoms are modeled with empirical potentials. Although many systems have successfully been investigated with model potentials, it is difficult to find empirical interactions which work for different states of matter and for a wide class of materials. In complex cases, such as those involving covalent bonds, there is no general agreement about such basic question as whether or not the interactions need to encompass 2-, 3-, 4- or higher n -body terms. Furthermore, the empirical molecular dynamics approach suffers from an important conceptual limitation: the correlation between local atomic structure, e.g. bonding properties, and atomic dynamics is missing, as well as the effect of the atomic dynamics on the electronic properties. A more accurate approach would be to derive the interatomic potential directly from first principles, i.e. from explicit knowledge of the electronic ground state at each atomic position.

In this section, we report the basic assumptions underlying the achievement of an interatomic potential from First Principles, and emphasize the conditions that make it really useful in actual MD simulations. The Density Functional Theory, which meets all these requirements, is presented along the lines of Hohenberg and Kohn (1964) for the energy functional and Kohn and Sham (1965) for single-particles orbital formulation. Finally, we deal with the resolution of the self-consistent equations that derive from this scheme by means of steepest descent, conjugate gradients or damped dynamics.

1.3.1 Conditions and assumptions

The derivation of interatomic forces from first principles relies upon two basic assumptions. We suppose that ions can be regarded as classical particles. Furthermore, we restrict ourselves to systems for which separation between the classical motion of the ions and the quantum motion of electrons can be achieved, i.e. systems satisfying the so-called Born-Oppenheimer approximation. This “adiabatic principle” is in general satisfied for semiconductors

and insulators, where the electronic excitation spectrum has a gap between the ground and the first excited state much larger than the energy associated with the ionic motion. The issue of the validity of the BO approximation for metals is much more subtle (Chester 1961). However it is generally believed that it holds also in this case provided that the plasma frequency is much larger than the ionic thermal energies.

For systems for which the BO approximation holds, the interaction potential among the ions, $\mathcal{V}(\{\mathbf{R}_I\})$ can be derived *ab initio*:

$$\mathcal{V}(\{\mathbf{R}_I\}) = \langle \Psi_0 | \mathcal{H}(\{\mathbf{R}_I\}) | \Psi_0 \rangle. \quad (1.60)$$

Here $\mathcal{H}(\{\mathbf{R}_I\})$ is the hamiltonian of the system at fixed ionic positions $\{\mathbf{R}_I\}$ (which includes the contribution due to the interactions amongst ions) and $\Psi_0 = \Psi_0(\{\mathbf{R}_I\})$ the corresponding instantaneous ground state. For any configuration $\{\mathbf{R}_I\}$, the Schrödinger equation has to be solved to get Ψ_0 then the forces on the ions can be evaluated by the use of the Hellman-Feynman theorem:

$$\frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} = \left\langle \Psi_0 \left| \frac{\partial \mathcal{H}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I} \right| \Psi_0 \right\rangle. \quad (1.61)$$

There are two basic requirements to be fulfilled in order that the use of Eqs. (1.60) and (1.61) in MD simulations is meaningful and practical. An accurate theory for the many-body hamiltonian $\mathcal{H}(\{\mathbf{R}_I\})$ is needed, capable of yielding with the same accuracy the ground state $\Psi_0(\{\mathbf{R}_I\})$ for different ionic configurations. Indeed, rather different atomic arrangements are sampled during a MD simulation. Furthermore, the theory should lead to a computational scheme which is affordable in practice, with modern computer capabilities.

A viable framework is offered by the Density Functional Theory proposed by Hohenberg and Kohn (1964) which provides an accurate and practical mean of calculating $\mathcal{V}(\{\mathbf{R}_I\})$. It scales acceptably as the function of the system size ($\mathcal{O}(N^3)$ where N is the number of electrons or less with newly developed techniques (Mauri *et al.* 1993, Ordejón *et al.* 1993, Li *et al.* 1993, Daw 1993) and has proved accurate for the description of many properties of both clusters and bulk systems.

1.3.2 Density Functional Theory

Density-functional theory calculations are based on the theorems introduced by Hohenberg and Kohn (1964) and on an independent-particle orbital formulation provided by Kohn and Sham (1965).

Hohenberg and Kohn proved that the total energy of an electron gas (even in the presence of a static external potential) is a unique functional of the electron density. The minimum value of the total-energy functional is the ground-state energy of the system, and the density that yields this minimum value is the exact single-particle ground-state density. Kohn and Sham then showed how it is possible, formally, to replace the many-electron problem by an exactly equivalent set of self-consistent one-electron equations. The central assertion used in establishing this single-particle orbital scheme is that, for any *interacting* system of electrons, we assume that there exists an auxiliary system of *non-interacting* particles such that the exact ground-state density of the interacting system equals the ground-state density of the auxiliary non-interacting system. For more details about density-functional theory see von Barth (1984), Dreizler and da Providencia (1985), Jones and Gunnarsson (1989), Kryachko and Ludena (1990).

The Kohn-Sham energy functional

The Kohn-Sham total-energy functional for a set of N_e electronic states ψ_i can be written

$$E_{tot}[\{\mathbf{R}_I\}, \{\psi_i\}] = \sum_i^{N_e} \langle \psi_i | -\frac{1}{2}\nabla^2 | \psi_i \rangle + \int V_{ion}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \\ + \int \frac{1}{2} \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{XC}[n(\mathbf{r})] + E_{ion}[\{\mathbf{R}_I\}] \quad (1.62)$$

where $V_{ion}(\mathbf{r})$ is the static total electron-ion potential, and $n(\mathbf{r})$ is the electronic density given by

$$n(\mathbf{r}) = \sum_i^{N_e} |\psi_i(\mathbf{r})|^2. \quad (1.63)$$

Eq. (1.62) formally defines the exchange and correlation energy functional $E_{XC}[n(\mathbf{r})]$. E_{tot} has been split into five contributions: the kinetic energy of the corresponding noninteracting system, the interaction energy with the external potential E_{el-ion} , the classical electrostatic energy for the electrons E_H , the exchange-correlation energy E_{XC} which contains all the remaining quantum interactions between the electrons, and the Coulomb energy E_{ion} associated with interactions among the ions at positions $\{\mathbf{R}_I\}$.

Only the minimum value of the Kohn-Sham energy functional has a physical meaning. It is equal to the ground-state energy of the system of electrons with the ions in positions \mathbf{R}_I .

However, in its present form, the Kohn-Sham formulation is not useful for practical computations, since the functional E_{XC} , which includes all the intricacies of the many-body electronic problem, is not known explicitly. The most widely used approximations for the exchange-correlation energy functional will be described in section 1.5.1.

Kohn-Sham equations

It is necessary to determine the set of wavefunctions ψ_i that minimize the Kohn-Sham energy functional. These are given by the self-consistent solutions to the Kohn-Sham equations (Kohn and Sham 1965):

$$\left[-\frac{1}{2}\nabla^2 + V_{KS}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (1.64)$$

where ψ_i is the wave function of electronic state i , ϵ_i is the Kohn-Sham eigenvalue, and V_{KS} is the Kohn-Sham potential. The latter is defined by:

$$V_{KS}(\mathbf{r}) = V_{ion}(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r}) \quad (1.65)$$

where V_H is the Hartree potential of the electrons given by

$$V_H(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (1.66)$$

and, the exchange-correlation potential, V_{XC} , is given formally by the functional derivative

$$V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \quad (1.67)$$

The Kohn-Sham equations represent a mapping of the interacting many-electron system onto a system of non-interacting electrons moving in an effective potential due to all the other electrons. If the exchange correlation functional were known exactly, then taking the functional derivative with respect to the density would produce an exchange-correlation potential that included the effects of exchange and correlation exactly.

The Kohn-Sham equations must be solved self-consistently so that the occupied electronic states generate a charge density that produces the electronic potential that was used to construct the equations.

1.3.3 Solution of the Kohn-Sham equations

In conventional DFT electronic structure calculations for solids and molecules, the KS self-consistent equations are solved as follows. An initial value for the electronic density $n(\mathbf{r})$ is guessed: for example it is taken to be the superposition of atomic charge densities. The KS potential is then calculated for this density, and the KS equation is solved by diagonalization of the KS Hamiltonian matrix \mathcal{H}_{KS} : this is set up according to a chosen basis set for the expansion of ψ_i . From the eigenvectors of \mathcal{H}_{KS} a new $n(\mathbf{r})$ is calculated, this output is mixed with the input in an adequate fashion to provide a new starting density, and the whole process is repeated till self consistency. Since the cost of standard diagonalization grows as $\mathcal{O}(M^3)$, where M is the number of basis functions used to expand the Kohn-Sham orbitals, this procedure becomes very costly for large systems and one is forced to use iterative techniques (Haydock 1972, Davidson 1975, Wood and Zunger 1985) to find the lowest occupied levels, instead of a direct diagonalization of \mathcal{H}_{KS} .

The KS equations can be solved in a different fashion using the variational principle that holds for the energy functional. Its minimum can be directly searched with respect to the single-particle orbitals ψ_i . Minimization can be achieved by introducing a fictitious dynamics in the space of electronic degrees of freedom ψ_i 's. The simplest dynamics that one can conceive is provided by steepest descent (SD), and can be formulated in terms of first-order equations in the fictitious variable t :

$$\mu\dot{\psi}_i = -\frac{\delta E}{\delta\psi_i^*} \quad (1.68)$$

Here the dot indicates time derivative, the functional derivative is $\frac{\delta E}{\delta\psi_i^*} = \mathcal{H}_{KS}\psi_i$ and \mathcal{H}_{KS} depends nonlinearly on the ψ_i . The parameter μ is a fictitious electronic mass. It is used to tune the speed of the electronic dynamics and does not describe any other physical property. When the ions are held fixed, this mass can be included in the definition of the time step and it is irrelevant. However, as we shall see, the ions are allowed to move, the ratio between μ and the physical ionic masses is important since it defines the relative speed of the ionic and of the fictitious electronic motion.

When solving Eq. (1.68) numerically, orthonormality constraints must be explicitly imposed and the fictitious dynamics becomes:

$$\mu\dot{\psi}_i = -\frac{\delta E}{\delta\psi_i^*} + \sum_j \Lambda_{ij}\psi_j \quad (1.69)$$

where the sum extends over the N occupied states. In practice, after an unconstrained step is performed, the predicted ψ_i are corrected by an appropriate orthonormalization procedure which gives the coefficients Λ_{ij} . Once the system has reached the minimum, $\dot{\psi}_i = 0$, and thus $\mathcal{H}_{KS}\psi_i = \sum_j \Lambda_{ij}\psi_j$, so by diagonalization of the constraint matrix $\mathbf{\Lambda}$ one recovers the KS equations. The unitary transform which diagonalizes $\mathbf{\Lambda}$ can be used to obtain the KS eigenstates and eigenvalues. Experience has shown that, at fixed nuclear configuration, E has only one minimum, and therefore the SD procedure leads to the absolute minimum. Note that in this dynamical way one achieves the objective of self-consistently minimizing E without ever performing explicit diagonalization. The efficiency of the SD scheme is controlled by the number of steps that are necessary to achieve convergence. This can be quite large, especially in metallic situations. Various attempts to improve the SD approach have been made.

Teter *et al.* (1989) obtained a very tractable scheme by adapting the well-known method of conjugate gradient (CG) minimization, and introducing an interesting preconditioning of the gradients. Moreover, they showed that by working band-by-band for the updating and the orthonormalization of the ψ_i the convergences is further improved with respect to solving for all the bands simultaneously.

Tassone *et al.* (1994) proposed an alternative second-order dynamics with a damping term:

$$\mu\ddot{\psi}_i = -\frac{\delta E}{\delta\psi_i^*} - \gamma\mu\dot{\psi}_i + \sum_j \Lambda_{ij}\psi_j \quad (1.70)$$

They showed the rate of convergence of this damped dynamics (DD) is much faster than that of SD dynamics provided the damping factor γ is adjusted during the dynamics. Indeed, it is convenient to use steepest descent in the first steps of the minimization when the highest frequency components dominate the deviation of the energy from the minimum. Subsequently, when only the slowest frequencies are left, damped dynamics becomes much more convenient, especially in those case of extremely slow convergence rate (e.g. metals). They also developed a preconditioning of the dynamics to improve the numerical efficiency of all the fictitious dynamical methods previously introduced. This is achieved by replacing the constant fictitious mass parameter μ in Eqs. (1.68), (1.69), and (1.70) with an arbitrary positive definite operator $\hat{\mu}$. The resulting increased arbitrariness in the choice of $\hat{\mu}$ is then exploited to compress the highest components of the spectrum of the fictitious electron dynamics. Recalling that these are due basically to

the high energy unoccupied states which are free-particle-like, they choose an operator $\hat{\mu}$ which is diagonal in q space with eigenvalues $\mu(q)$ given by

$$\begin{aligned} \mu(q) &= \mu_0 & \text{if } \frac{1}{2}q^2 < E_p, \\ \mu(q) &= \mu_0 \frac{q^2}{2E_p} & \text{if } \frac{1}{2}q^2 > E_p. \end{aligned} \tag{1.71}$$

Below a certain cutoff energy E_p , it is worth considering a constant mass μ_0 , because the low energy eigenstates have a relevant potential energy contribution and are not free-particle-like. The preconditioning cutoff E_p therefore represents the threshold above which the states are dominated by the kinetic energy.

As already mentioned, the basic step of iterative schemes such as SD, CG or DD consists in evaluating the action of \mathcal{H}_{KS} on the ψ_i , followed by the orthonormalization of the single-particle orbitals. If the ψ_i are expanded in plane waves (with M plane waves for each of the N electronic states), the computation for all the states of $\mathcal{H}_{KS}\psi_i$ by matrix-vector multiplication requires, in principle, NM^2 floating point operation, whereas the orthonormalization requires N^2M . Car and Parrinello (1985) out that when using local pseudopotentials $\mathcal{H}_{KS}\psi_i$ can be conveniently calculated in real space by means of fast Fourier transform (FFT) techniques and requires $\mathcal{O}(NM \log M)$ operations. As apparent, in these iterative scheme the large number M (usually $M \gg N$) enters only via $M \log M$ in the number of operations, and this gives a significant improvement over standard diagonalization techniques where M enters as M^3 . An additional advantage of SD, CG and DD iterative techniques is that the matrix \mathcal{H}_{KS} is never explicitly required, thus permitting a significant save in memory. The storage scales as N^2 . Of course, for sufficiently large systems, the orthogonalization workload, growing like N^2M dominates, and the numerical cost of the algorithm still grows as the cube of the system size.

New schemes based on an orbital picture and on new energy functionals, which is proven to have the KS ground-state energy as its absolute minimum value, have been recently developed (Mauri *et al.* 1993, Ordejón *et al.* 1993, Li *et al.* 1993, Daw 1993). The crucial feature of these functionals is that their minimization implies neither explicit orthogonalization of the orbitals nor inversion of the overlap matrix. The use of this approach within a localized orbital formulation leads straightforwardly to the so called $\mathcal{O}(N)$ methods whose workload grows linearly with the system size.

In summary, we have outlined the basic assumptions underlying the derivation of interatomic forces from First Principles. We have emphasized that to be really useful for MD simulations, conditions of accuracy

and tractability should be met. We have presented the Density Functional Theory to solve the many-electron problem providing an effective mean to obtain accurate interatomic potentials. The Hohenberg-Kohn for the energy functional has been introduced, as well as the Kohn-Sham formulation of the many-body problem in terms of single-particles orbitals. Finally, we have presented the most widely used methods to solve the KS self-consistent equations namely steepest descent, conjugate gradients or damped dynamics.

1.4 First-principles Molecular Dynamics

1.4.1 Introduction

In principle, any of the methods described in the previous section can be used to solve the KS equations and then derive *ab initio* the interaction potential amongst the ions $\mathcal{V}(\{\mathbf{R}_I\})$, at each atomic configuration $\{\mathbf{R}_I\}$. One could then perform *ab initio* MD simulations by following three separate steps, at each MD move: (i) solve self-consistently the KS equations for a given ionic configuration $\{\mathbf{R}_I\}$; (ii) compute forces acting on ions (\mathbf{F}_I), according to the Hellman-Feynman theorem ($\mathbf{F}_I = -\nabla_{\mathbf{R}_I}\mathcal{V}$); (iii) solve Newton equations of motions: $M_I\ddot{\mathbf{R}}_I = -\nabla_{\mathbf{R}_I}\mathcal{V}$.

An alternative methodology was proposed some years ago by Car and Parrinello (Car and Parrinello 1985), in which the electrons and the ions are treated on the same footing. In this scheme, the ionic trajectories and the corresponding electronic ground state are generated at the same time, without repeating the full self-consistent procedure to solve the KS equations at each MD step. This approach, together with global optimization techniques will be discussed in detail in the next section.

Before doing so, it is useful to see how to generalize the minimization procedure of the energy functional E with respect to the electronic degrees of freedom to include ionic degrees of freedom. This idea of simultaneously relaxing both ions and electrons was first discussed by Bendt and Zunger (1983). Within SD, this is easily done by considering the two following sets of equations:

$$\mu\dot{\psi}_i = -\frac{\delta E}{\delta\psi_i^*} + \sum_j \Lambda_{ij}\psi_j \quad (1.68)$$

$$M_I\dot{\mathbf{R}}_I = -\frac{\partial E}{\partial\mathbf{R}_I} \quad (1.71)$$

where the masses μ and M_I have been introduced in order to account for the difference in time scales associated with the two sets of parameters $\{\psi_i\}$ and $\{\mathbf{R}_I\}$. Eqs. (1.69) and (1.69) define mass weighted SD trajectory on the E surface that moves downhill from the starting point to a nearby local minimum. By varying the ψ_i at fixed ionic positions, a single minimum on the E surface is found, which coincides with a point of the BO potential energy surface \mathcal{V} defined by

$$\mathcal{V}(\{\mathbf{R}_I\}) = \min_{\{\psi_i\}} E[\{\psi_i\}, \{\mathbf{R}_I\}]. \quad (1.72)$$

As a consequence, the local minima of the E surface are also local minima of the \mathcal{V} surface. Eqs. (1.69) and (1.69) can be used to optimize complex atomic structures, by relaxing simultaneously electronic and nuclear degrees of freedom. In conventional gradient methods this is accomplished by following a trajectory on the \mathcal{V} surface, which means that perfect electronic self-consistency is required for any nuclear configuration. By using Eqs. (1.69) and (1.69) one follows instead a trajectory on the fictitious E surface, and the optimal relative relaxation rate for electrons and ions can be achieved by adjusting the respective mass parameters. The scheme may be generalized to include additional degrees of freedom like the volume or the shape of the unit cell.

The scheme just presented has a severe limitation, allowing only for local and not global optimizations. Indeed, it is well known that the potential energy surface \mathcal{V} can have several minima. To deal with this problem one needs a global optimization method. We show in the next section that a satisfactory methodology to the global optimization problem¹ is based on statistical mechanics, and it is intimately connected to the calculation of finite-temperature properties and the possibility of performing DFT-based MD simulations.

1.4.2 The Car-Parrinello method

The key-point in the FP-MD method proposed by Car and Parrinello (1985) is the definition of a fictitious dynamical system, associated with the physical system, whose potential energy surface E is an appropriate functional of both ionic and electronic degrees of freedom; the fictitious system is devised in such a way that the trajectories generated by its dynamics reproduce very closely those of the physical system with potential energy surface \mathcal{V} . The

¹The global optimization problem is *NP*-complete, and thus “the solution” is not available

classical lagrangian of the physical system is given by the sum of the ionic kinetic energy and the ionic potential energy with reversed sign (Eq. (1.3)). The generalized lagrangian of the fictitious system is defined as:

$$\begin{aligned} \mathcal{L} = & \sum_i \mu_i \int |\dot{\psi}_i(\mathbf{r})| d\mathbf{r} + \frac{1}{2} \sum_{i=1}^N M_I \dot{\mathbf{R}}_I^2 - E[\{\psi_i\}, \{\mathbf{R}_I\}] \\ & + \sum_{ij} \Lambda_{ij} \left(\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} - \delta_{ij} \right) \end{aligned} \quad (1.73)$$

\mathcal{L} does not depend explicitly on time, and is a functional of two sets of classical degrees of freedom, the ψ_i and the \mathbf{R}_I , which depend on time. The parameter μ , expressed in units (mass) \times (length)², plays the role of a generalized mass for the electronic degrees of freedom; how to choose this parameter remains to be specified. For simplicity we have consider a unique μ for the ψ_i , independent from the electronic state, even though this is by no means necessary. The first and second term in Eq. (1.73) are the kinetic energy of the electronic and ionic degrees of freedom, K_e and K_I , respectively. E is the potential energy of the coupled electron-ion fictitious system. The lagrangian multipliers Λ_{ij} are used to impose orthonormality conditions on the ψ_i ; in the language of classical mechanics, they are just simple holonomic constraints, that is they are expressible as $f(\psi) = 0$ where the function f does not depend explicitly on time and therefore the constraints do not do any work. The Euler equations associated with the lagrangian of Eq. (1.73) are:

$$\mu \ddot{\psi}_i = - \frac{\delta E}{\delta \psi_i^*} + \sum_j \Lambda_{ij} \psi_j \quad (1.74)$$

$$M_I \ddot{\mathbf{R}}_I = - \frac{\partial E}{\partial \mathbf{R}_I} \quad (1.75)$$

At first sight there is no relation between the ionic dynamics generated from Eq. (1.75) and that obtained from the correct equation of motion (EOM) for the nuclei, derived from the lagrangian of Eq. (1.3):

$$M_I \ddot{\mathbf{R}}_I = - \frac{\partial \mathcal{V}(\{\mathbf{R}_I\})}{\partial \mathbf{R}_I}. \quad (1.76)$$

Indeed, in general the nuclear trajectories generated by Eq. (1.75) and those obtained from Eq. (1.76) do not coincide, unless $E[\{\psi_i\}, \{\mathbf{R}_I\}]$ is at the instantaneous minimum. However, the parameter μ and the initial conditions

$\{\psi_i\}_0, \{\dot{\psi}_i\}_0$ can be chosen in such a way that the time-scale for the electronic degrees of freedom is much shorter than that of the nuclei. In this case nuclear trajectories, initially lying on the BO surface, will deviate from it only after times that are significantly longer than the MD time step and, in some cases, even longer than the MD observation time. In other words, if μ and $\{\psi_i\}_0, \{\dot{\psi}_i\}_0$ are chosen so that the two sets of classical degrees of freedom, ions and electrons, are only weakly coupled, the transfer of energy between them is small enough, to allow the electrons to follow adiabatically the ionic motion, remaining close to the BO surface. In such a metastable situation, meaningful temporal averages can be computed. This dynamics, which we call classical adiabatic dynamics, is meant to reproduce in a computationally effective way what indeed occurs in real life, that is electrons following adiabatically the ionic motion. The role of Eq. (1.74) is that of keeping the electrons on the BO surface, without solving self-consistently the KS equations at each step, but for the initial configuration.

We note that thermodynamic constraints can also be imposed like in the case of classical molecular dynamics. Constant-pressure FPMD has been introduced recently by Bernasconi *et al.* (1995) in which the shape of the MD cell is allowed to vary along the lines of Parrinello-Rahman method. It is thus possible to study crystal structure transformations from first-principles (Focher *et al.* 1995, Scandolo *et al.* 1995). Constant-temperature FPMD are a bit more particular since two different Nosé-Hoover thermostats can be applied, one to the ionic and the other to the electronic degrees of freedom. The role of these thermostats is to maintain the ions at the desired temperature while at the same time keeping the electrons “cold”, namely very close to the ground state. This opportunity has been exploited to avoid transfer of energy between electrons and ions in metals (Blöchl and Parrinello 1992).

To summarize, we have presented first-principles molecular dynamics as proposed by Car and Parrinello (1985). In this approach, the system of electrons and ions that evolve according to the BO approximation is regarded as a two-component system with two different energy scales. While ions move at the preassigned temperature, the electrons that have to follow adiabatically, remaining in the instantaneous ground state, are essentially at zero temperature. With this in mind, the BO system which has a quantum mechanical component, namely the electrons, is mapped into a system in which ions and electrons are treated classically.

1.5 Practical implementation

The FP-MD scheme just described was originally based on a pseudopotential description of the interaction between ionic cores and valence electrons, in conjunction with plane-wave (PW) basis sets and the Local Density Approximation (LDA) for the exchange-correlation energy functional (Car and Parrinello 1985). To date, the method has been mostly applied in its original version, for reasons of computational efficiency, however, a variety of modifications have also been tested.

The choice of LDA to describe exchange-correlation effects in the electronic system is quite simple to implement in the framework of pseudopotential-plane-wave formalism, since the E_{XC} is a functional of the density. Though it is a really drastic approximation, the method works surprisingly well for many physical properties. However, the LDA has some deficiencies and some functionals have been proposed to try to solve them. These new approximations to E_{XC} have been successfully introduced in CP scheme (White and Bird 1994).

The choice of PW expansions for the single-particle orbitals has several advantages. It permits the use of FFT techniques which are computationally very efficient; PW's do not depend on atomic positions and therefore forces acting on atoms can be easily computed via a straightforward application of the Hellman-Feynman theorem. PW's are free of basis-set superposition errors and allow computation of the total energy of different atomic arrangements with the same accuracy. This is particularly relevant in MD simulations, where rather different atomic configurations are explored. Finally the convergence of PW calculations can be controlled in a very simple manner, since it depends only upon the number of Fourier components included in the expansion of the ψ_i . The disadvantage of PW expansions is the very large number (M) of basis functions needed to represent the electronic orbitals, compared, for example, to basis sets of gaussian or atomic-like orbitals. PW basis sets call, of course, for the use of pseudopotentials, since the number of PW's needed to describe localized core states is too large and cannot be afforded from a computational point of view. Other basis sets have also been tested in the framework of CP method. For instance, an all-electron formulation within an augmented plane-waves scheme has been used by Blöchl (1994).

The rest of this section is organized as follows. In section 1.5.1, we concentrate first on the electron-electron interactions, introducing the physical origin of exchange and correlation. We describe various approximations to the E_{XC} functional starting from LDA and going beyond emphasizing their

advantages and inconveniences. In section 1.5.2, we explain how the resolution of the problem can be simplified by the use of periodic cells and a plane-wave basis set in the framework of Bloch's theorem, and present the reciprocal space formulation of the Schrödinger equation. In section 1.5.3, we discuss the modeling of electron-ion interactions with the frozen-core approximation and the use of pseudopotentials. We illustrate the various possible forms for the latter (local, non-local, semi-local or separable) and introduce Kleinman-Bylander procedure. We outline the differences between the two most widely used types of pseudopotentials, namely norm-conserving and ultrasoft. Finally, in section 1.5.4, we briefly present the Ewald summation technique for the calculation of ion-ion interaction energy.

1.5.1 Electron-electron interaction

Exchange-correlation energy

Before addressing the issue of the approximations most widely used for the exchange-correlation energy functional, we will try to introduce briefly the physical origin of this contribution to the total energy. On this purpose, we move back to the many-body problem and introduce the Hartree and Hartree-Fock approximations.

In many-electron systems, the most difficult problem is posed by the need to take account of the effects of the electron-electron interaction. Fortunately, a good insight in the properties of such systems can already be obtained from the study of effective non-interacting particles models. In these models, the interacting electron problem is replaced by a non-interacting system of particles in which the external field is replaced by an effective external field which incorporates to some extent the interparticle interactions.

When the fictitious electrons are non-interacting, the different variables can be separated in the many-body wavefunction, that basically takes the form of a "Hartree product":

$$\Psi_H(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) = \psi_1(\mathbf{r}_1) \cdot \psi_2(\mathbf{r}_2) \cdot \dots \cdot \psi_{N_e}(\mathbf{r}_{N_e}) \quad (1.77)$$

In this case, the electron-electron interaction energy E_{el} reduces to the classical electrostatic energy (Coulomb energy) between the electrons, which is called the Hartree energy E_H .

This simple model can be improved by taking into account the antisymmetric character of the many-body wavefunction under exchange of any two electrons because the electrons are fermions. The resulting wavefunction

can be written in terms of a ‘‘Slater determinant’’:

$$\Psi_S(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N_e}) = (1/\sqrt{N!}) \times \det \begin{bmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \cdots & \psi_1(\mathbf{r}_{N_e}) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \cdots & \psi_2(\mathbf{r}_{N_e}) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{N_e}(\mathbf{r}_1) & \psi_{N_e}(\mathbf{r}_2) & \cdots & \psi_{N_e}(\mathbf{r}_{N_e}) \end{bmatrix} \quad (1.78)$$

More generally, the wavefunction should also include a dependence on the spin of the electrons. This is obtained by a simple product of two Slater determinants (one for spin-up electrons and one for spin-down electrons).

The antisymmetry of the wavefunction produces a spatial separation between electrons that have the same spin and thus reduces the Coulomb energy of the electronic system. The reduction in the electron-electron interaction energy due to the antisymmetry of the wavefunction is called the exchange energy E_X . This is generally referred to as the Hartree-Fock approximation.

The real many-body wavefunction has extra dimensionalities with respect to a Slater determinant, so that the electron-electron interaction energy can be further reduced below its Hartree-Fock value, if electrons that have opposite spins are also spatially separated. The difference between the electron-electron interaction energy in the many-body case and in the Hartree-Fock approximation is called the correlation energy E_C .

We now turn to Kohn-Sham single-orbital formulation in which electron-electron energy is written as a functional of the density:

$$E_{el}[n(\mathbf{r})] = E_H[n(\mathbf{r})] + E_{XC}[n(\mathbf{r})] \quad (1.79)$$

where the problem remains of approximating the exchange-correlation energy functional in order to obtain a useful scheme for practical computations. It should be noted that the exchange-correlation energy of the Kohn-Sham formalism is not equivalent to that of the many-electron system. The difference originates in a transfer of part of the many-body kinetic energy to the exchange-correlation term within KS formulation. (Levy and Perdew 1985).

The Hohenberg-Kohn theorem provides some motivation for using a function of the electron density to describe the exchange-correlation energy. That is what is done in the local-density approximation (LDA) and in some scheme going beyond LDA.

Local-density approximation

The simplest method of describing the exchange-correlation energy of an electronic system is to use the local-density approximation (LDA) proposed by Kohn and Sham (1965). Due to its simplicity and effectiveness, this approximation is almost universally used in total-energy pseudopotential calculations.

The exchange-correlation energy functional can be written:

$$E_{XC}[n(\mathbf{r})] = \int \epsilon_{XC}(\mathbf{r})n(\mathbf{r})d\mathbf{r} \quad (1.80)$$

that is an integral of the exchange-correlation energy per electron at a point \mathbf{r} times the local density of particles $n(\mathbf{r})$. In the local-density approximation, it is assumed that $\epsilon_{XC}(\mathbf{r})$ only depends on the local density $n(\mathbf{r})$ and is equal to the exchange-correlation energy per electron in a homogeneous electron gas that has the same density as the electron gas at point \mathbf{r} :

$$\epsilon_{XC}^{LDA}(\mathbf{r}) = \epsilon_{XC}^{hom}[n(\mathbf{r})]. \quad (1.81)$$

The exchange-correlation energy functional is thus purely local.

The exchange part can be obtained analytically using Hartree-Fock technique:

$$\epsilon_X^{LDA}(\mathbf{r}) = \frac{\alpha}{r_s} \quad (1.82)$$

where $\alpha = -\frac{3}{4}\frac{3}{2\pi}^{2/3}$ and the local Seitz radius $r_s = r_s(\mathbf{r})$ is defined by:

$$r_s = \left(\frac{3}{4\pi n(\mathbf{r})} \right)^{1/3}. \quad (1.83)$$

It is the radius of the sphere that would contain exactly one electron in the homogeneous electron gas of density $n(\mathbf{r})$.

The correlation part is derived by interpolating homogeneous electron-gas data obtained by quantum Monte-Carlo calculations (Ceperley and Alder 1980). Several parameterizations exist for this functional. Throughout the present study, we have mainly used that of Perdew and Zunger (1981) that is written:

$$\epsilon_C^{LDA}(\mathbf{r}) = \begin{cases} \frac{\alpha_1}{1+\alpha_2 r_s^{1/2} + \alpha_3 r_s} & \text{if } r_s > 1 \\ \beta_1 + \beta_2 \ln r_s + \beta_3 r_s + \beta_4 r_s \ln r_s & \text{if } r_s < 1 \end{cases} \quad (1.84)$$

where $\alpha_1 = -0.1423$, $\alpha_2 = 1.9529$, $\alpha_3 = 0.3334$, $\beta_1 = -0.0480$, $\beta_2 = 0.0311$, $\beta_3 = -0.0116$, and $\beta_4 = 0.0020$. We will also introduce the parameterization proposed by Perdew and Wang (1992) (that will be needed in the framework of generalized gradient approximation):

$$\epsilon_C^{LDA}(\mathbf{r}) = -2a(1 + \alpha_1 r_s) \times \ln \left[1 + \frac{1}{2a(\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_4 r_s^2)} \right] \quad (1.85)$$

where $a = 0.0310907$, $\alpha_1 = 0.21370$, $\beta_1 = 7.5957$, $\beta_2 = 3.5876$, $\beta_3 = 1.6382$, and $\beta_4 = 0.49294$.

The local-density approximation ignores corrections to the exchange-correlation energy at a point \mathbf{r} due to nearby inhomogeneities in the electron density. Considering the inexact nature of this approximation, it is remarkable that calculations performed using the LDA have been so successful. This can be partially attributed to the fact that the exchange-correlation hole in the local-density approximation fulfills the sum rule according which this hole should integrate to -1 (Harris and Jones 1974, Gunnarsson and Lundqvist 1976, Langreth and Perdew 1977).

Note that an extension to spin-polarized systems has also been proposed (Gunnarsson and Lundqvist 1976). In this approximation, referred to as the local-spin-density LSD approximation, the exchange-correlation energy per electron is a functional of the local “up” and “down” electron spin densities $n_\uparrow(\mathbf{r})$ and $n_\downarrow(\mathbf{r})$.

Beyond the LDA

During the last 20 years, a variety of recipes have been proposed to improve local-density approximation. We refer to the abundant literature for a systematic discussion and to Appendix A where a brief description of the most commonly used approximations is proposed.

In the present study, when hydrogen bonds had to be simulated (see Chapter 3), we used the functionals proposed by Perdew and Wang (1991), that fall into the class of generalized gradient approximations (GGA) for which $\epsilon_{XC}(\mathbf{r})$ is expressed in terms of the density n and its gradient.

The exchange part is written:

$$\epsilon_X^{PW'91}(\mathbf{r}) = \epsilon_X^{LDA}(\mathbf{r}) \left(\frac{1 + a_1 s \sinh^{-1}(a_2 s) + (a_3 + a_4 e^{-100s^2})}{1 + a_1 s \sinh^{-1}(a_2 s) + a_5 s^4} \right) \quad (1.86)$$

where $a_1 = 0.19645$, $a_2 = 7.7956$, $a_3 = 0.2743$, $a_4 = -0.1508$, $a_5 = 0.004$. The scaled density gradient $s = s(\mathbf{r})$ is defined by:

$$s = \frac{|\nabla n|}{2k_F n} \quad (1.87)$$

where the local Fermi wave vector $k_F = k_F(\mathbf{r})$ is

$$k_F = (3\pi^2 n)^{1/3}. \quad (1.88)$$

The correlation part is written:

$$\epsilon_C^{PW'91}(\mathbf{r}) = \epsilon_C^{LDA}(\mathbf{r}) + H(n, s, t) \quad (1.89)$$

where $t = t(\mathbf{r})$ is another scaled density gradient defined by:

$$t = \frac{|\nabla n|}{2k_s n} \quad (1.90)$$

with the local screening wave vector $k_s = k_s(\mathbf{r})$ written as:

$$k_s = (4k_F/\pi)^{1/2}. \quad (1.91)$$

The correction factor to the LDA correlation functional is given by:

$$H = \frac{\beta^2}{2\alpha} \ln \left(1 + \frac{2\alpha}{\beta} \frac{t^2 + At^4}{1 + At^2 + A^2t^4} \right) + C_{c0}[C_c(n) - C_{c1}]t^2 e^{-100s^2} \quad (1.92)$$

where

$$A = \frac{2\alpha}{\beta} [e^{-2\alpha\epsilon_C^{LDA}(\mathbf{r})/\beta^2} - 1]^{-1} \quad (1.93)$$

and $\alpha = 0.09$, $\beta = 0.0667263212$, $C_{c0} = 15.7559$, $C_{c1} = 0.003521$. The function $C_c(n)$ is defined by:

$$C_c(n) = C_1 + \frac{C_2 + C_3 r_s + C_4 r_s^2}{1 + C_5 r_s + C_6 r_s^2 + C_7 r_s^3}, \quad (1.94)$$

with $C_1 = 0.001667$, $C_2 = 0.002568$, $C_3 = 0.023266$, $C_4 = 7.389 \times 10^{-6}$, $C_5 = 8.723$, $C_6 = 0.472$, $C_7 = 7.389 \times 10^{-2}$. The combination of these two functionals to obtain $\epsilon_{XC}(\mathbf{r})$ is referred to as PW'91 approximation.

Although experience in GGA is not as consolidated as that for LDA, the emerging picture is that in general GGA is an improvement over LDA,

especially for bonding energies (Johnson *et al.* 1993) and the description of hydrogen bonds (Sprik 1991, Sim *et al.* 1992).

Note that the GGA energy density is a smoothly varying function of the electron charge density n and of $|\nabla n(\mathbf{r})|/n(\mathbf{r})$, and present no computational problems. The corresponding exchange-correlation potential, however, involves the gradient of a function with cusps, and presents extreme convergence difficulties in plane-wave calculations. White and Bird (1994) introduced a method which deals efficiently with these problems.

In summary, we have demonstrated the physical origin of the splitting of the electron-electron interaction energy into Hartree, exchange and correlation energies. We have described various approximations to the E_{XC} functional, starting from the Local Density Approximation which, due to its simplicity and effectiveness, is almost universally used in total-energy pseudopotential calculations. Then, we have emphasized its limitations and introduced new recipes that have been proposed to go beyond LDA. We have stressed that the Generalized Gradient Approximation is an improvement over LDA, especially for bonding energies and the description of hydrogen bonds.

1.5.2 Periodic cells and Plane waves

In the framework of DFT, it was demonstrated that certain observables of the many-body problem can be mapped into equivalent observables in an effective single-particle problem. However, there still remains the formidable task of handling an infinite number of non-interacting electrons moving in the static potential of an infinite number of nuclei or ions. Two difficulties must be overcome: a wavefunction must be calculated for each of the infinite number of electrons in the system, and, since each electronic wavefunction extends over the entire solid, the basis set required to expand each wavefunction is infinite. Both problems can be surmounted by performing calculations on periodic systems and applying Bloch's theorem to the electronic wavefunctions. This naturally leads to the use of plane-wave basis sets and to k-point sampling techniques.

Supercell approximation

In simulations of bulk systems, techniques have been developed to take advantage of the translational symmetry of the corresponding hamiltonian to reduce the complexity of the problem (see Bloch's theorem, hereafter). To this end, the systems are represented by a MD box which is periodically

repeated to infinity.

This is a natural choice for perfect crystals, but this is no more the case for aperiodic configurations of atoms such as disordered systems, systems with a point-defect, surfaces or even molecules. The supercell approximation allows one to adopt periodic boundary conditions (PBC) even in these latter cases.

One simply constructs a large unit-cell (called supercell) containing the configuration in question and repeats it periodically throughout the space. By studying the properties of the system for larger and larger supercells, one can gauge the importance of the induced periodicity and systematically filter it out.

For disordered systems, such as amorphous solids or liquids, the supercell is chosen to be large enough so that the imposed periodicity does not affect the dynamical properties of the system.

For systems with a point-defect, the supercell contains the defect surrounded by a region of bulk crystal. It is essential to include enough bulk solid in the supercell to prevent the defects in neighboring cells from interacting. The independence of defects in neighboring cells can be checked by increasing the volume of the supercell until the computed defect energy has converged. It can then be assumed that defects in neighboring unit cells no longer interact.

For surfaces, the supercell contains a crystal slab and a vacuum region. To ensure that the results of the calculation accurately represent an isolated surface, the vacuum regions must be wide enough so that faces of adjacent crystal slabs do not interact across the vacuum region, and the crystal slab must be thick enough so that the two surfaces of each crystal slab do not interact through the bulk crystal. Typically, the vacuum region should be at least 5 Å and the slab should be 10-12 Å thick.

Finally, for molecules, again the supercell needs to be large enough so that the interactions between the molecules are negligible. In general, a 10-12 Å cube (depending on the size of the molecule) is sufficient.

Therefore, the techniques developed in electronic structure calculations within DFT, which take advantage of the translational symmetry of the hamiltonian, can also be used for DFT-MD simulations for mostly all ‘real’ systems.

Bloch’s theorem

Bloch’s theorem (see Ashcroft and Mermin 1976) states that in a periodic system each electronic wave function can be written as the product of a

wavelike part and a cell-periodic part:

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{i,\mathbf{k}}(\mathbf{r}) \quad (1.95)$$

where k is a wave vector of the BZ and i is the band index.

The cell-periodic part of the wave function can be expanded using a basis set consisting of a discrete set of plane waves whose wavevectors are reciprocal lattice vectors of the crystal:

$$u_{i,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{i,\mathbf{k}}(\mathbf{G})e^{i\mathbf{G}\cdot\mathbf{r}} \quad (1.96)$$

where \mathbf{G} are the reciprocal lattice vectors. Therefore each electronic wave function can be written as a sum of plane waves:

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{i,\mathbf{k}}(\mathbf{G})e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \quad (1.97)$$

where $c_{i,\mathbf{k}}(\mathbf{G})$ are the Fourier components (FC) of the single-particle wavefunction, which are treated as time-dependent degrees of freedom in the MD calculation.

Plane-wave basis sets

Using Bloch's theorem, the electronic wave functions at each \mathbf{k} point can be expanded in terms of a discrete plane-wave basis set (Eq. (1.97)). In principle, an infinite plane-wave basis set is required to expand the electronic wave functions in Eq. (1.97). However, the coefficients $c_{i,\mathbf{k}}(\mathbf{G})$ for the plane waves with small kinetic energy $\frac{1}{2}|\mathbf{k} + \mathbf{G}|^2$ are typically more important than those with large kinetic energy. Thus the plane-wave basis set can be truncated to include only plane waves that have kinetic energies less than some particular cutoff energy E_{cut} . If a continuum of plane-wave basis states were required to expand each electronic wave function, the basis set would be infinitely large no matter how small the cutoff energy. Application of the Bloch theorem allows the electronic wave functions to be expanded in terms of a discrete set of plane waves. Introduction of an energy cutoff to the discrete plane-wave basis set produces a finite basis set.

The truncation of the plane-wave basis set at a finite cutoff energy will lead to an error in the computed total energy. However, it is possible to reduce the magnitude of the error systematically by increasing the value of E_{cut} . The latter depends on the specific system and in particular upon the choice of the pseudopotential for the description of the core-valence interaction (see section 1.5.3). For a given pseudopotential, in principle,

the cutoff energy should be increased until the calculated total energy has converged. The choice of E_{cut} determines the accuracy of the calculation.

One of the difficulties associated with the use of plane wave basis sets is that the number of basis states, that will be noted $^2 \bar{N}_{PW}^d$ (the subscript “d” stands for “discontinuous”), changes discontinuously with cutoff energy, as illustrated in Fig. 1.1. In general these discontinuities will occur at different cutoffs for different k points. In addition, at a fixed-energy cutoff, a change in the size or shape of the unit cell will cause discontinuous variations of the number of plane waves in the basis set, so that the total energy and the stress curves obtained from calculations at constant E_{cut} are ragged.

Froyen and Cohen (1986) proposed a method to handle this problem for stress calculations. It consists in adding a correcting factor, called Pulay stress by analogy with the Pulay force (Pulay 1969), which approximately accounts for the difference between the number of plane waves used in the calculation and the fictitious number of states in the basis set if it were obtained by a continuous distribution in the reciprocal space \bar{N}_{PW}^c :

$$\bar{N}_{PW}^c = \left(\frac{V}{8\pi^3} \right) \frac{4}{3} \pi \left((2E_{cut})^{1/2} \right)^3 = \frac{V}{6\pi^2} (2E_{cut})^{3/2} \quad (1.98)$$

where the subscript “c” stands for “continuous”.

Francis and Payne (1990) introduced a correction to the energy by integrating the Pulay stress expression of Froyen and Cohen. However, we showed (Rignanese *et al.* 1995) that the latter definition of the Pulay stress was inaccurate and proposed a more rigorous technique for both total energy and stress calculations which applies also in the case of anisotropic deformations. This scheme relies on the interpolation of the energy as a function of the number of plane waves $E_{tot}[\bar{N}_{PW}]$ and on a scaling hypothesis that allows us to perform this interpolation for a unique reference configuration \mathbf{h}^0 (where \mathbf{h}^0 describes the cell shape cfr. section 1.2, the volume of the cell being simply $V_0 = \det \mathbf{h}^0$). For any configuration \mathbf{h}^1 with volume $V_1 = \det \mathbf{h}^1$, the correction to the total energy is given by:

$$\Delta E = E_{tot} \left[\frac{V_0}{V_1} \bar{N}_{PW}^c \right] - E_{tot} \left[\frac{V_0}{V_1} \bar{N}_{PW}^d \right] \quad (1.99)$$

whereas the correction to the stress tensor is:

$$\Delta \sigma_{\alpha\beta} = \delta_{\alpha\beta} \frac{V_0}{V_1^2} \bar{N}_{PW}^d \frac{\partial E_{tot}[\bar{N}_{PW}]}{\partial \bar{N}_{PW}} \bigg|_{\bar{N}_{PW} = \frac{V_0}{V_1} \bar{N}_{PW}^d} \quad (1.100)$$

²Since the number of plane waves varies with the electronic wavevector k , when we note \bar{N}_{PW} we refer to its average on the special k -point set.

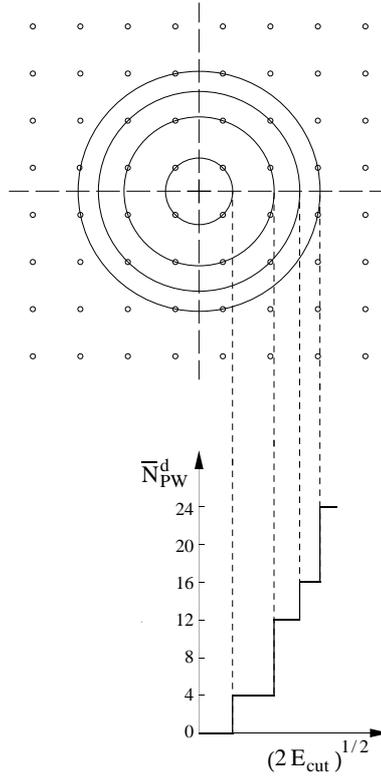


Figure 1.1: Number of plane waves \bar{N}_{PW}^d in a basis set defined by $\frac{1}{2}|\mathbf{k} + \mathbf{G}|^2 < E_{cut}$ in a 2D case. The upper graph illustrates how these k-points are enclosed in a circle of radius $(2E_{cut})^{1/2}$. The lower graph illustrates the stair-like evolution of \bar{N}_{PW}^d as a function of E_{cut} .

When performing Parrinello-Rahman (variable cell-shape) molecular dynamics within the framework of DFT, these discontinuities constitute a serious source of nuisance. Bernasconi *et al.* (1995) proposed a modified kinetic term for the energy functional that allows to work as close as possible to the constant E_{cut} , while avoiding non-physical discontinuities. These scheme has proved very successful (Focher *et al.* 1995, Scandolo *et al.* 1995).

k-point sampling

The computation of the electronic density $n(\mathbf{r})$, then of the electronic potential $v_{KS}(\mathbf{r})$ and of the total energy E_{tot} in principle requires an integral over the whole BZ. This would necessitate the knowledge of the electronic wavefunctions at \mathbf{k} points in the entire Brillouin zone. However, the electronic wave functions at \mathbf{k} points that are very close together will be almost identical. Hence it is possible to represent the electronic wavefunctions over a region of \mathbf{k} space by the wavefunctions at a single \mathbf{k} point. With this averaging operation, the electronic density can be written as a weighted sum over only a finite number of \mathbf{k} points, instead of being a integral over the whole BZ:

$$n(\mathbf{r}) = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_i^{N_e} |\psi_{i,\mathbf{k}}(\mathbf{r})|^2 \quad (1.101)$$

with $\sum_{\mathbf{k}} w_{\mathbf{k}} = 1$.

Chadi and Cohen (1973), Monkhorst and Pack (1976), and Evarestov and Smirnov (1983) have devised methods for obtaining special sets of k points in the Brillouin zone which give good average of the wavefunctions.

The magnitude of any error in the total energy due to inadequacy of the k -point sampling can always be reduced by using a denser set of \mathbf{k} points. The computed total energy will converge as the density of \mathbf{k} points increases.

Note that the number of \mathbf{k} points needed can be further decreased by taking into account symmetry considerations. This is an important source of computational saving in BZ sampling with special points. However, a general lattice distortion breaks the point symmetry of the lattice, and in this case BZ sampling with special points is not much help.

In general, the computation of finite-temperature properties of crystals and the simulation of disordered systems is most appropriately performed with large supercells and the use of the Γ point to sample the BZ. Indeed, there is a computational advantage in the use of the Γ point: at $\mathbf{k} = (0, 0, 0)$, one can choose the single-particle orbitals $\psi_i(\mathbf{r})$ to be real, since the phase factor of the wavefunction is arbitrary. Thus, for each reciprocal lattice

vector \mathbf{G} , the Fourier components $c_i(\mathbf{G})$ of the orbitals satisfy the symmetry relation:

$$c_i(-\mathbf{G}) = c_i^*(\mathbf{G}). \quad (1.102)$$

One can take advantage of this property to reduce the independent FC (e.g. the basis set) by a factor of two, substantially reducing the computational cost of a MD simulation.

When plane waves are used as a basis set for the electronic wavefunctions, the Kohn-Sham equations assume a particularly simple form. Substitution of Eq. (1.97) into Eq. (1.64 and integration over \mathbf{r} gives the secular equation:

$$\sum_{\mathbf{G}'} \left[\frac{(\mathbf{k} + \mathbf{G})^2}{2} \delta_{\mathbf{G},\mathbf{G}'} + V_{KS}(\mathbf{G} - \mathbf{G}') \right] c_{i,\mathbf{k}}(\mathbf{G}) = \epsilon_{i,\mathbf{k}} c_{i,\mathbf{k}}(\mathbf{G}') \quad (1.103)$$

where

$$V_{KS}(\mathbf{G} - \mathbf{G}') = V_{ion}(\mathbf{G} - \mathbf{G}') + V_H(\mathbf{G} - \mathbf{G}') + V_{XC}(\mathbf{G} - \mathbf{G}'). \quad (1.104)$$

In this form, the kinetic energy is diagonal, and the various potentials are described in terms of their Fourier transforms. Note that special care must be taken for the $\mathbf{G} = 0$ component of the Hartree and ionic potential to avoid electrostatic divergences (see for instance Gonze 1993-1994). Similar divergences appear at the energy level for the Hartree term, E_{el-ion} and E_{ion} (see section 1.5.4).

To summarize, we have discussed the use of periodic cells and plane-wave basis sets to expand the wavefunctions by introducing Bloch's theorem in the frame work of DFT. We have presented some difficulties associated with the use of plane waves and the solutions that have been brought. We have introduced the special k-points technique to sample the Brillouin zone, allowing for the problem to be reformulated in the reciprocal space and leading to simpler equations.

1.5.3 Electron-ion interactions

Two major inconveniences are associated with the use of plane waves for expanding the electronic wavefunctions. First, an extremely large plane-wave basis set is required to describe correctly the all-electron wavefunctions. Indeed, these include tightly bound core orbitals, which are localized in the core region, implying plane-wave components up to quite large wave vectors. Second, a very large number of plane waves are needed to follow the rapid

oscillations of the valence wavefunctions in the core region. These oscillations maintain the orthogonality with the core wavefunctions, as required by the exclusion principle.

We now briefly explain how the first problem may be avoided within the frozen-core approximation and how the second can be limited by the use of pseudopotentials. We present the various possible forms for the latter (local, non-local, semi-local or separable) and introduce Kleinman-Bylander procedure to transform any semi-local pseudopotential into a separable form. We compare the norm-conserving and ultrasoft pseudopotentials and show how they modify Kohn-Sham formalism.

Frozen-core approximation

The frozen-core approximation is based on the following observations. It is well known that most physical and chemical properties of solids are much more dependent on the valence electrons than on the core electrons. On the other hand, as the core electrons do not directly participate to the chemical bonding, one expects that they are only slightly affected by modifications of the atomic environment. It may therefore reasonably be approximated that the configuration of the core electrons within the solid is equivalent to that of the isolated atoms. In terms of the density, the frozen-core approximation corresponds to assume that:

$$n(\mathbf{r}) = n_{core}^{atom}(\mathbf{r}) + n_{val}(\mathbf{r}) \quad (1.105)$$

Within this approximation, the problem of treating the core electrons is considered as being solved (i.e. it has been solved at the atomic level), while the study restricts to the investigation of the behavior of the valence electrons within the ionic potential, partly screened by the core electrons. Note that the separation between valence and core states is not strict, like the one usually considered in undergraduate chemistry lectures, since electrons from deep energy levels can always be treated as valence electrons. In practice, the number of valence states changes according to the environment or to the degree of approximation allowed, even for the same atomic species.

Pseudopotential approximation

The pseudopotential approximation allows the valence wavefunctions to be expanded using a much smaller number of plane-wave basis states. Basically, it consists in a mathematical transformation by which the ionic potential V_{ion} screened by the core electrons is replaced by a weaker pseudopotential

V_{ps} that acts on a set of pseudo-wavefunctions rather than the true valence wavefunctions, and leading to the same eigenenergy in the Schrödinger equation. Ideally, the pseudopotential is constructed so that the effect of the core states on the valence states is effectively reproduced and in such way that the pseudo-wavefunctions have no radial nodes in the core region (defined by a cut-off radius r_{cut}). Outside the core region, the pseudopotential must reduce to the ionic potential, in order for the wavefunction and the pseudo-wavefunction to be identical:

$$V_{ps}(\mathbf{r}) = V_{ion}(\mathbf{r}) \text{ for } r > r_{cut} \quad (1.106)$$

and

$$\psi_{ps}(\mathbf{r}) = \psi(\mathbf{r}) \text{ for } r > r_{cut} \quad (1.107)$$

with the following spherical boundary conditions:

$$\psi_{ps}(\mathbf{r}) = \psi(\mathbf{r}) \text{ for } r = r_{cut} \quad (1.108)$$

and

$$\frac{d\psi_{ps}(\mathbf{r})}{d\mathbf{r}} = \frac{d\psi(\mathbf{r})}{d\mathbf{r}} \text{ for } r = r_{cut} \quad (1.109)$$

In order to be inserted in Schrödinger equations and give real eigenvalues, pseudopotentials should be linear, hermitian, operators. The most general form of such operators acts upon wavefunctions as follows:

$$\langle \mathbf{r} | V_{ps} | \psi \rangle = \int v_{ps}(\mathbf{r}, \mathbf{r}') \psi(\mathbf{r}') d\mathbf{r}' \quad (1.110)$$

where $v_{ps}(\mathbf{r}, \mathbf{r}')$ is the kernel of operator V_{ps} .

As we want pseudopotentials to replace atomic potentials, these operators should also be invariant under rotations. It is possible to show that this condition leads to the following development:

$$V_{ps} = \sum_{l,m} |Y_{l,m}\rangle V_l \langle Y_{l,m}| \quad (1.111)$$

where $|Y_{l,m}\rangle$ are the spherical harmonics and V_l is the pseudopotential operator for angular momentum l . Thus, the operator V_{ps} decomposes the wavefunction into spherical harmonics, the relevant pseudopotential operator V_l is then applied to each of these. In terms of the kernel, Eq. (1.111) can be rewritten:

$$v_{ps}(\mathbf{r}, \mathbf{r}') = \sum_{l,m} Y_{l,m}^*(\theta, \phi) v_l(r, r') Y_{l,m}(\theta', \phi') \quad (1.112)$$

In this expression, there is one kernel for each angular momentum channel, and these kernels depend on the radial distances r and r' . These conditions define the *non-local potentials*.

By contrast, the simplest form for a pseudopotential is called local pseudopotential. In this case, the kernel is a function only of the distance from the nucleus (meaning also that there is no angular momentum dependency):

$$v_{ps}(\mathbf{r}, \mathbf{r}') = v_{LOC}(r)\delta(r - r'), \quad (1.113)$$

and thus, the corresponding operator, noted $V_{LOC}(r)$, acts on the wavefunction as follows:

$$\langle \mathbf{r} | V_{LOC}(r) | \psi \rangle = v_{LOC}(r)\psi(\mathbf{r}). \quad (1.114)$$

It is quite an interesting form for the simplicity of the calculations. Unfortunately, in general, such a pseudopotential cannot reproduce accurately the properties of the all-electron wavefunction. However, it should be noted that the original ionic potential is a local potential. Thus, the pseudopotential is also local outside the cutoff radius sphere. Indeed, above this limit, the pseudopotential must reduce to the original potential, as already mentioned above.

For most elements of the periodic table, it is a good approximation to assume that $v_l(r, r') = v_{\bar{l}}(r)\delta(r - r')$ for $l \geq \bar{l}$, where \bar{l} can be one of the few lowest value (s, p, and sometimes d). Indeed, high angular momentum wave functions have vanishing presence probability within the cut-off radius, due to the centrifugal potential. This latter observation is an interesting starting point to build a more satisfactory form for the pseudopotential:

$$v_{ps}(\mathbf{r}, \mathbf{r}') = v_{LOC}(r)\delta(r - r') + \sum_{l=0}^{\bar{l}-1} \sum_m Y_{l,m}^*(\theta, \phi) v_{NL,l}(r, r') Y_{l,m}(\theta', \phi') \quad (1.115)$$

where $v_{NL,l}(r, r') = v_l(r, r') - v_{LOC}(r)\delta(r - r')$ is a short-ranged kernel. The first term on the right hand side $v_{LOC}(r) = v_{\bar{l}}(r)$ is called the local part of the pseudopotential, whereas the second term is referred to as the non-local part. The corresponding operator is thus:

$$V_{ps} = V_{LOC}(r) + V_{NL} = V_{LOC}(r) + \sum_{l,m} |Y_{l,m}\rangle V_{NL,l} \langle Y_{l,m}| \quad (1.116)$$

This form is much more practical for computations than Eq. (1.111) however further simplifications are introduced in actual calculations for the non-local part of the pseudopotential.

It should be mentioned the Kohn-Sham energy functional is now:

$$\begin{aligned}
E_{tot}[\{\mathbf{R}_I\}, \{\psi_i\}] &= \sum_i^{N_e} \langle \psi_i | -\frac{1}{2}\nabla^2 + \sum_I V_{NL,I} | \psi_i \rangle \\
&+ \int \sum_I V_{LOC,I}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \int \frac{1}{2} \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\
&+ E_{res} + E_{XC}[n(\mathbf{r})] + E_{ion}[\{\mathbf{R}_I\}]
\end{aligned} \tag{1.117}$$

where E_{res} is a correction to account for the deviation from the purely Coulombic behavior of the local part of the pseudopotentials (see for example Gonze 1993-1994):

$$E_{res} = \frac{1}{V} \left(\sum_J Z_J \right) \sum_I \int \left[V_{LOC,I}(\mathbf{r}) + \frac{Z_I}{|\mathbf{r}|} \right] d\mathbf{r}. \tag{1.118}$$

Note also that the Fourier transform of Kohn-Sham potential is now non-local in Eq. (1.103) and hence it is written:

$$\begin{aligned}
V_{KS}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}') &= V_H(\mathbf{G} - \mathbf{G}') + V_{XC}(\mathbf{G} - \mathbf{G}') \\
&+ V_{LOC}(\mathbf{G} - \mathbf{G}') + V_{NL}(\mathbf{k} + \mathbf{G}, \mathbf{k} + \mathbf{G}').
\end{aligned} \tag{1.119}$$

The most widely used pseudopotentials belong to two different classes. There are the *semi-local* potentials, in which each of the $V_{NL,I}$ is diagonal in r :

$$v_{NL,I}(r, r') = v_{SL,I}(r)\delta(r - r') \tag{1.120}$$

and the *separable* potentials, in which the kernels are a simple product

$$v_{NL,I}(r, r') = \xi_l^*(r) f_l \xi_l(r') \tag{1.121}$$

Semi-local pseudopotential are easily visualized, but separable pseudopotentials are definitely more powerful for numerical techniques.

Kleinman and Bylander (1982) introduced a procedure to transform any semi-local pseudopotential into a separable form:

$$\xi_l(r) = v_{SL,I}(r)R_l(r) \tag{1.122}$$

and

$$f_l = \int v_{SL,I}(r)R_l^2(r)4\pi r^2 dr \tag{1.123}$$

where $R_l(r)$ is the radial part of the atomic reference pseudo-wavefunction for which the pseudopotential was calculated. Hence the operator corresponding to the non-local part of the pseudopotential is written:

$$V_{NL}^{KB} = \sum_{l,m} \frac{|V_{SL,l}(r)\Phi_{l,m}(\mathbf{r})\rangle \langle \Phi_{l,m}(\mathbf{r})V_{SL,l}(r)|}{\langle \Phi_{l,m}(\mathbf{r})|V_{SL,l}(r)|\Phi_{l,m}(\mathbf{r})\rangle} \quad (1.124)$$

where $\Phi_{l,m}(\mathbf{r}) = Y_{l,m}(\theta, \phi)R_l(r)$.

Note that separable pseudopotentials have to be handled with some caution, since in some cases they can give rise to undesired spurious solutions (called ghost states) at energies lower than the ground state energy. The appearance of these ghost states has been traced back to the choice of the local pseudopotential (Gonze *et al.* 1990, Gonze *et al.* 1991).

Norm-conserving pseudopotentials

Along the sixties and seventies, pseudopotentials have been widely used, though in a quite empirical manner, providing a great number of remarkable results in solid-state physics (Philips and Kleinman 1959, Abarenkov and Heine 1965, Cohen and Bergstresser 1966, Cohen and Heine 1970). However, at the end of the seventies, the pseudopotential concept was formalized mathematically and cautiously defined as a transformation of the frozen-core Hamiltonian. This led to a new generation of pseudopotentials being essentially exact in a wide energy range, and capable of mimicking the original potential in a large variety of atomic environments, a property referred to *transferability*.

The key concept making this step possible is *norm conservation* brought in condensed matter physics by Hamann, Schlüter and Chiang (1979). The underlying idea is the crucial role of the density in total-energy calculations. Indeed, the energy of the electronic system can be written as a functional of the electron density. If that energy is to be desired accurately, it is necessary that outside the core region the pseudo-wavefunctions and real wavefunctions be identical, not just in their spatial dependences but also in their absolute magnitudes, so that the two wavefunctions generate identical charge densities. It is possible to show that if the pseudopotential is adjusted to ensure that the integrals of the squared amplitudes of the real and the pseudo-wavefunctions inside the core regions are identical, the equality of the wavefunction and pseudo-wavefunction is guaranteed outside the core region.

$$\int_{r < r_c} |\psi_{ps}(\mathbf{r})|^2 d\mathbf{r} = \int_{r < r_c} |\psi(\mathbf{r})|^2 d\mathbf{r} \quad (1.125)$$

This condition that defines the norm-conservation guarantees that the norm of the pseudo-wavefunctions and real wavefunctions are identical outside as well inside the core region.

The spherical boundary conditions Eqs. (1.108) and (1.109) can be reformulated in terms of the logarithmic derivative:

$$D(\epsilon, r) = D_{ps}(\epsilon, r) \text{ for } r = r_{cut} \quad (1.126)$$

where by definition:

$$D(\epsilon, r) = \frac{r}{\psi(\epsilon, r)} \frac{\partial \psi(\epsilon, r)}{\partial r} = \frac{\partial \ln \psi(\epsilon, r)}{\partial \ln r} \quad (1.127)$$

Hamann, Schlüter and Chiang (1979) showed that norm-conservation condition (Eq. (1.125)) implies the equality between the first-order energy dependence of the logarithmic derivatives:

$$\frac{dD(\epsilon, r)}{d\epsilon} = \frac{dD_{ps}(\epsilon, r)}{d\epsilon} \text{ for } r = r_{cut}. \quad (1.128)$$

Based on this considerations, Bachelet, Hamann, and Schlüter (1982) proposed a simple analytical form of norm conserving pseudo-potentials (PP's). This kind of PP was used in the present study for Si and H atoms to investigate the Si(001)/SiO₂ interface. Note that in their original formulation, these PP's are semi-local but they can easily be transformed into a separable form using Kleinman-Bylander scheme.

After the success of norm-conserving pseudopotential in many applications, the transferability has often been associated with the correct variation of logarithmic derivatives of the pseudo-wavefunction with energy. Thus, many attempts to improve norm-conserving pseudopotentials have centered around the logarithmic derivatives of the atomic and pseudo-atomic wavefunctions in order to improve the energy range over which they match. A method for the construction of pseudopotentials that corrects even the higher-order energy dependence of the logarithmic derivative (known as *extended-norm-conserving* condition) has been established by Shirley *et al.* (1989).

Concentration on the logarithmic derivatives has also led to methods generating *soft* pseudopotentials whose wavefunctions require a minimum number of plane waves which facilitates the convergence to the solution (Rappe *et al.* 1990, Troullier and Martins 1991). In the present study, we used Troullier-Martins PP's for Si, O, and H atoms to investigate SiO₂ surface.

A major contribution to transferability in pseudopotentials came from Louie, Froyen and Cohen (Louie *et al.* 1982), who realized that the deletion of the electron density of the core states removed by the pseudopotential construction gave a non-linear error in the exchange-correlation potential in the LDA. By restoring the core density n_{core} , they were able to show that atoms would respond correctly over a wider range of conditions (*core correction*).

An additional condition for transferability (*chemical hardness conservation*) has also been recently proposed by Teter (1993). The requirement that the total energies of atom and pseudo-atom match to second order with arbitrary changes in the valence-state occupancy yields major decreases in the errors made using pseudopotentials.

The typical method for generating a pseudopotential for an atom proceeds as follows. All-electron calculations are performed for an isolated atom in its ground state and some excited states, using a given form for the exchange-correlation density functional. This provides valence electron eigenvalues and valence electron wavefunctions for the atom. A parametrized form for the pseudopotential is chosen. The parameters are then adjusted, so that a pseudoatom calculation using the same form for exchange-correlation as in the all-electron atom gives both pseudowave functions that match the valence wavefunctions outside some cutoff radius r_{cut} , pseudo-eigenvalues that are equal to the valence eigenvalues. The pseudopotential obtained in this fashion is then used, without further modification, for any environment of the atom. The electronic density in any new environment of the atom is then determined using both the pseudopotential obtained in this way and the same form of exchange-correlation functional as employed in the construction of the pseudopotential. Finally, it should be noted that pseudopotentials are constructed with r_{cut} ranging typically from one to two times the value of the core radius. Note also that the easiest approach to increase the pseudopotential transferability is to simply decrease the cutoff radius r_{cut} used to generate the pseudopotential and the pseudo-wavefunctions, thereby reducing the difference between the all-electron and pseudopotential results. However, there are practical limits on how far we can decrease r_{cut} , for example it must be larger than the outermost node of the all-electron wavefunction if we insist on having nodeless pseudo-wavefunctions.

Ultrasoft pseudopotentials

A more radical approach has been suggested by Vanderbilt (1990), which involves to relax the norm-conserving condition on the pseudo-wavefunction.

The associated pseudopotentials are termed *ultrasoft*. The rationale behind this technique is that in most cases a high cutoff energy is required for the plane-wave basis-set only when there are tightly bound orbitals that have a substantial fraction of their weight inside the core region of the atom. In the framework of norm-conserving pseudopotentials, the cutoff energy cannot be substantially reduced in this case because there must be planewave components up to a large enough wave vector to allow the majority of the weight of the wavefunction to be kept within the core. On the contrary, in the context of ultrasoft pseudopotentials, the pseudo-wavefunctions are not required to have the same charge as the all-electron wavefunctions, and can therefore be designed to be as smooth as possible in the core region, so that the cut-off energy for the planewave basis set is reduced considerably. All that is required is to rebuild the correct electron density to account for the part of the charge which is not described by the pseudo wavefunctions ψ_i . This is done by augmenting the pseudo-density with appropriate functions (noted $Q_{nm}^I(\mathbf{r})$) localized in the core region and defined as follows:

$$n(\mathbf{r}) = \sum_i \left[\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) + \sum_{nm,I} Q_{nm}^I \langle \psi_i | \beta_n^I \rangle \langle \beta_m^I | \psi_i \rangle \right] \quad (1.129)$$

where the functions β_n^I are strictly localized in the core region and are also used to define the non-local pseudopotential:

$$V_{NL} = \sum_{nm,I} D_{nm}^0 |\beta_n^I\rangle \langle \beta_m^I| \quad (1.130)$$

The functions β_n^I and Q_{nm}^I are related to the atomic functions β_n and Q_{nm} by

$$\beta_n^I(\mathbf{r}) = \beta_n(\mathbf{r} - \mathbf{R}_I), \quad (1.131)$$

$$Q_{nm}^I(\mathbf{r}) = Q_{nm}(\mathbf{r} - \mathbf{R}_I). \quad (1.132)$$

The $Q_{nm}(\mathbf{r})$ are constructed in the atomic pseudization procedure in such a way that, at the reference energies, the electron density of the pseudo-wavefunctions as defined by Eq. (1.129) be the same as the all-electron density. The functions β_n and Q_{nm} and the parameters D_{nm}^0 are obtained from first principles in the ultrasoft pseudopotential scheme, and characterize the atomic species. Here, for simplicity, we consider a single atomic species.

Technically, the possibility of separating the description of the wavefunctions from that of the electron density is acquired by making use of a generalized orthonormality condition for the wave functions:

$$\langle \psi_i | S | \psi_j \rangle = \delta_{ij}, \quad (1.133)$$

where the operator S is defined by

$$S = 1 + \sum_{nm,I} q_{nm} |\beta_n^I\rangle \langle \beta_m^I| \quad (1.134)$$

Here $q_{nm} = \int Q_{nm}(\mathbf{r}) d\mathbf{r}$. It can easily be checked that the condition given by Eq. (1.133) is consistent with the definition of the electron density (see Eq. (1.129)) leading to

$$\int n(\mathbf{r}) d\mathbf{r} = N_e \quad (1.135)$$

where N_e is the number of valence electrons.

In Vanderbilt's ultrasoft pseudopotential scheme the total energy is given by the same expression as in the norm-conserving case, where the non-local contribution to the pseudopotential is now V_{NL} , of Eq. (1.130). However, the functional derivative entering Eq. (1.74) is modified, with respect to the norm-conserving scheme, since $\delta n(\mathbf{r})/\delta \psi_i^*(\mathbf{r}')$ is not simply given by $\psi_i(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}')$, but is:

$$\frac{\delta n(\mathbf{r})}{\delta \psi_i^*(\mathbf{r}')} = \psi_i(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}') + \sum_{nm,I} Q_{nm}^I \langle \beta_m^I | \psi_i \rangle \beta_n^I(\mathbf{r}'), \quad (1.136)$$

This leads to

$$\frac{\delta E}{\delta \psi_i^*} = \left[-\nabla^2 + \sum_{nm,I} D_{nm}^I |\beta_n^I\rangle \langle \beta_m^I| + V_{KS} \right] \psi_i = \epsilon_i S \psi_i \quad (1.137)$$

where ϵ_i are the eigenvalues of the hamiltonian defined by the terms between squared brackets, V_{KS} is given by Eq. (1.65) in which V_{ion} is the local part of the pseudopotential, and

$$D_{mn}^I = D_{mn}^0 + \int V_{KS}(\mathbf{r}) Q_{mn}^I(\mathbf{r}) d\mathbf{r}. \quad (1.138)$$

Note, that D_{mn}^0 are just parameters, whereas the D_{nm}^I depend through V_{KS} on the wavefunctions and have to be updated in the self-consistent procedure.

In norm-conserving pseudopotential schemes the energy cutoff E_{cn} required to describe the electron density is four times the energy cutoff E_{cut} of the wave functions. The simple relationship $E_{cn} = 4 \cdot E_{cut}$ does not always hold in the ultrasoft pseudopotential scheme, because of the augmentation functions Q_{nm} entering the expression of the electron density Eq. (1.129). It is therefore appropriate to introduce two independent energy cutoffs. The soft part of the electron density, i.e. the first term on the right hand side of Eq. (1.129) can again be described with a cutoff which is four times E_{cut} . In contrast, the part of Eq. (1.129) which depends on the augmentation functions Q_{nm} requires a much higher cutoff, E_{cn} . The cutoff E_{cn} can be reduced by smoothing the Q_{nm} . An inner core region (determined by an inner radius which is smaller than the cutoff radius of the core region) is defined in which the Q_{nm} are modified in such a way that all the electrostatic properties outside the inner core region are preserved. In general, the smoothing of the Q_{nm} allows the use of a cutoff E_{cn} which is approximately equal to $4 \cdot E_{cut}$. Hence, the calculation can be carried out similarly to the norm-conserving case. For some elements, such as transition metals, the Q_{nm} require $E_{cn} > 4 \cdot E_{cut}$, even after smoothing. In this case the problem of having a large number of plane waves can still be overcome by recognizing that the Q_{nm} are localized in real space.

In the ultrasoft pseudopotential scheme, the inclusion of orthonormality constraints in the lagrangian Eq. (1.73) has to be generalized and the last term entering Eq. (1.74) becomes:

$$\sum_{ij} \Lambda_{ij} (\langle \psi_i | S | \psi_j \rangle - \delta_{ij}) \quad (1.139)$$

The Euler equations associated with the lagrangian of Eq. (1.73) including the constraints of Eq. (1.139) are:

$$\mu \ddot{\psi}_i = -\frac{\delta E}{\delta \psi_i^*} + \sum_j \Lambda_{ij} S \psi_j \quad (1.140)$$

$$M_I \ddot{\mathbf{R}}_I = -\frac{\partial E}{\partial \mathbf{R}_I} \sum_{ij} \Lambda_{ij} \left\langle \psi_i \left| \frac{\partial S}{\partial \mathbf{R}_I} \right| \psi_j \right\rangle \quad (1.141)$$

It should be noted that a contribution to the ionic forces from the orthonormality constraints (Eq. (1.133)) appears because the operator S depends on the ionic positions through the functions β_n^I (Eq. (1.131)). The expression for the ionic forces can however easily be obtained (Laasonen,

Pasquarello, Car, Lee and Vanderbilt 1993), though it produces some complications for the practical implementation. This contribution of orthogonality constraints to the ionic forces is absent in the case of norm-conserving schemes.

The main feature of Vanderbilt's ultrasoft pseudopotential scheme is to reduce the required plane-wave cutoff for the valence wavefunctions. This is achieved by allowing the pseudo-wavefunction to be as soft as possible outside the core region, which is determined by a given cutoff radius r_{cut} . The larger the cutoff radius, the lower the energy cutoff for the plane waves. The upper physical limit for the cutoff radius is given by half a characteristic bond length r_{bond} . It turns out that for large cutoff radii, $r_{cut} \simeq r_{bond}$, it is necessary to use more than one reference energy for every angular momentum channel in order to have pseudopotentials which are as accurate as conventional norm-conserving pseudopotentials. Therefore, besides the complication of having a generalized eigenvalue problem (Eq. (1.137)) and wavefunction dependent D_{nm}^I (Eq. (1.138)), the number of β_n functions has generally to be doubled, increasing the related computational cost correspondingly. In spite of these complications, the number of required plane waves is so reduced with respect to the norm-conserving case that the present scheme is still advantageous compared to conventional methods for a variety of systems.

Vanderbilt PP's were used in this study for N and O atoms to investigate Si(001)/SiO₂ interface.

In summary, the replacement of the true ionic potential by a weaker pseudopotential has a number of advantages. In addition to the reduction of the number of plane-wave basis states needed to expand the electronic wavefunctions, the pseudopotential approximation removes the core electrons, leading to fewer electronic wavefunction calculations and to the disappearance of the rapid oscillations of the valence wavefunctions in the cores of the atoms.

More importantly, the total energy of the valence electron system is typically a thousand times smaller than the energy of the all-electron system. The differences between the electronic energies of different ionic configurations appear almost entirely in the energy of the valence electrons, so that the accuracy required to determine these differences in a pseudopotential calculation is much smaller than in an all-electron calculation. However the total energy of the system is no longer meaningful, only differences are.

1.5.4 Ion-ion interactions

Besides of electron-electron and electron-ion interaction energies (see sections 1.5.1 and 1.5.3, respectively), the total energy of the system also includes the contribution of the ion-ion repulsion $E_{ion}(\mathbf{R}_I)$ for a given configuration of the nuclei:

$$E_{ion}(\{\mathbf{R}_I\}) = \sum_{1 \leq I < J \leq N_p} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} = \frac{1}{2} \sum_{I \neq J}^{N_p} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \quad (1.142)$$

where the factor $\frac{1}{2}$ accounts for double counting of pairs of ions.

As long-ranged in both real and reciprocal spaces, the Coulomb energy of the ionic system using is not easily computed using summation.

Ewald (Ewald 1917a, Ewald 1917b, Ewald 1921) developed a rapidly convergent method for performing the Coulomb summations over periodic lattices. Ewald's summation is based on the following identity :

$$E_{Ewald} = \sum_{\mathbf{l}} \frac{1}{|\mathbf{l} - \tau|} = \frac{1}{\sqrt{\pi}} \sum_{\mathbf{l}} \int_{\eta}^{\infty} \frac{e^{-|\mathbf{l} - \tau|^2 \rho}}{\sqrt{\rho}} d\rho + \frac{\pi}{V} \sum_{\mathbf{G}} \int_0^{\eta} e^{-\frac{|\mathbf{G}|^2}{4\rho}} e^{-i\mathbf{G} \cdot \tau} \frac{1}{\rho^2} d\rho \quad (1.143)$$

where \mathbf{l} are the real-lattice vectors, \mathbf{G} are reciprocal-lattice vectors, τ is a real-space vector. The identity holds for all positive value of η and is invariant in η .

Such identity provides a method for rewriting the lattice summation for the Coulomb energy due to the interaction between an ion positioned at τ_1 in the unit cell and an array of atoms positioned at the set of points $\tau_2 + \mathbf{l}$. Indeed, it suffices to take $\tau = \tau_1 - \tau_2$ in Eq. (1.143). At first sight, the infinite Coulomb summation on the left-hand side of this equation is replaced by two infinite summations, one over real-lattice vectors and the other over reciprocal-lattice vectors. However, for an appropriate value of η , the two summations become rapidly convergent in their respective spaces. These two summations can then be computed with only a few respective lattice vectors if η is adjusted to optimize the convergence of the separate sums over \mathbf{G} 's and \mathbf{l} 's.

It should be noted that a divergence occurs in the reciprocal-lattice summation for $\mathbf{G} = 0$. This has to be related to similar problems in E_{el} and E_{el-ion} . As there is no Coulomb potential at $\mathbf{G} = 0$ in a charge neutral system, there cannot be any contribution to the total energy from this $\mathbf{G} = 0$

component of the Coulomb potential (Ihm *et al.* 1979). In fact, these contributions to the total energy from the electron-electron, electron-ion and ion-ion interactions cancel exactly, so that, in practice, the $\mathbf{G} = 0$ contribution can be omitted in each of these three terms.

Finally, to obtain the correct ionic energy two terms must be added to the Ewald summation from which the $\mathbf{G} = 0$ term has been removed (Yin and Cohen 1982):

$$E_{ion}(\{\mathbf{R}_I\}) = \frac{1}{2} \sum_{I,J} Z_I Z_J \left\{ \sum_{\mathbf{l} \neq 0} \frac{\text{erfc}(\sqrt{\eta} |\mathbf{l} - \tau_I + \tau_J|)}{|\mathbf{l} - \tau_I + \tau_J|} - \frac{\pi}{\eta V} - 2\sqrt{\frac{\eta}{\pi}} \delta_{IJ} + \frac{4\pi}{V} \sum_{\mathbf{G} \neq 0} \frac{1}{|\mathbf{G}|^2} e^{-\frac{|\mathbf{G}|^2}{4\eta}} \cos[\mathbf{G} \cdot (\tau_J - \tau_I)] \right\} \quad (1.144)$$

where Z_I , Z_J and τ_I , τ_J are the respective valences and locations of the ions I and J in the unit cell, and $\text{erfc}(x) = \int_x^\infty e^{-y^2} dy$ is the complementary error function. The ionic interaction of a nuclei with its own charge is omitted by imposing $\mathbf{l} \neq 0$ in the real-space summation when $I = J$.

This completes the development of the different contributions to the total energy of the system.

1.6 Conclusion

In this chapter, we have presented First-Principles Molecular Dynamic methodology which, by combining Density Functional Theory and Molecular Dynamics technique, made it possible to use forces derived within first-principles in MD simulations. Originally developed in the solid state community by Car and Parrinello (1985), this scheme rapidly crossed the frontier of chemistry and even biochemistry.

Section 1.2 has been dedicated to the classical Molecular Dynamics, that had been developed to study the evolution of a system of interacting particles. We have presented the corresponding equations of motion discussing the Verlet algorithms for their practical implementation. First-order (steepest descent) and damped dynamics equations of motion have been proposed providing a technique to find the energy minimum of the system. The imposition of thermodynamic constraints (constant temperature or pressure) has also been discussed by means of Nosé-Hoover thermostats and Parrinello-Rahman method. We have also emphasized the need to have a more and

more accurate description of the interatomic potential, which naturally lead to First-Principles technique.

In section 1.3, we have demonstrated the usefulness of DFT to deal with the ground-state of a quantum-mechanical system of electrons, by obtaining a single-orbital picture (Kohn-Sham formalism) of the many-body problem. We have also commented the most widely used methods to solve the self-consistent Kohn-Sham equations to derive finally a First-Principles interatomic potential.

In section 1.4, we have presented the Car-Parrinello method to perform First-Principles Molecular Dynamics (FP-MD), which is the origin of numerous achievements in the field. We have analyzed the key concepts that make it a really practical approach for MD simulations.

Finally section 1.5 was dedicated the practical implementation of FP-MD scheme. Different energy functionals have been proposed to describe the exchange-correlation part of electron-electron interactions energy, which is not known exactly in the framework of DFT. We have started from the Local Density Approximation (LDA) and then introduced several schemes (GGA, SIC, ...) going beyond it. We have demonstrated the advantages of using plane-wave basis sets to expand the wavefunctions in the case of periodic cells. We have introduced the reciprocal-space formulation of Kohn-Sham equations which is particularly suited for practical computations. We have discussed how the electron-ion interactions can be modeled by the frozen-core approximation and the use of pseudopotentials, reducing the plane waves in the basis set. We have distinguished local, non-local, semi-local and separable forms for the pseudopotentials, outlining their respective advantages. We have introduced the norm-conservation condition in the context of the transferability of the PP's and illustrated how this requirement could be relaxed in the case of ultrasoft pseudopotentials. In the end, we have briefly discussed the ion-ion interactions to complete the description of total energy components.

Chapter 2

The nitrated Si(001)-SiO₂ interface

2.1 Introduction

Thermally grown silicon dioxide films have many obvious applications in silicon devices and technology, including gate dielectrics of metal-oxide-semiconductor field effect transistors (MOSFET's). In the framework of very large scale integration (VLSI), the relevant scale for these devices has been decreasing continuously and is currently smaller than 100 Å. Because of this trend, the understanding of the structural properties of the Si(001)-SiO₂ interface at a microscopic level has acquired an enhanced technological importance.

Many experimental techniques (Pantelides 1978, Helms and Deal 1988, Helms and Deal 1993) have been used to investigate the detailed atomic structure of this interface. However, various difficulties have to be faced among which the problem of accessing a buried interface, the amorphous nature of the SiO₂ component and the dependence on sample preparation techniques. All these factors make it really difficult to obtain a complete description of the structural properties at the Si(001)-SiO₂ interface.

Among the large variety of experimental methods of investigation of this interface, X-ray photoemission spectroscopy (XPS) on Si 2*p* core-levels stands out as one of the most successful tools (Grunthaner *et al.* 1987, Himpsel *et al.* 1988, Lu *et al.* 1993). By complementing these studies with a theoretical approach that allows to calculate core-level shifts based on first-principles, a quite satisfactory picture of the microscopic structure at the interface is reached (Pasquarello *et al.* 1995, Pasquarello *et al.* 1996).

Despite of these invaluable advances in the knowledge of the atomic details of the Si(001)-SiO₂ interface, there remain several technological and reliability problems with silicon dioxide in the deep submicron regime. It is known, for instance, that very thin SiO₂ films are not sufficiently good masks against impurity diffusion. Moreover, the growth of high quality very thin layer of SiO₂ is rather difficult due to the defect density, integrity and yield problems. It is also known that high energy radiation can generate high density of interface states in the oxides, resulting in degradation of device performance. Many efforts have been made to overcome these difficulties.

In this framework, the incorporation of a low concentration of N atoms near the Si(001)-SiO₂ interface has recently drawn considerable attention (Hao *et al.* 1995, Kang *et al.* 1994, Carr and Buhrman 1993, Yao *et al.* 1994, Green *et al.* 1994). It appears as one of the most promising ways to match industrial programs requiring high-quality ultrathin gate dielectrics for VLSI. Indeed, silicon nitrided oxides (SiO_xN_y) exhibit several advantages over conventional oxides. Carr and Buhrman (1993) showed that the inclusion of nitrogen reduces interface state generation. Also, higher dielectric breakdown values have been reported (Hao *et al.* 1995, Kang *et al.* 1994) for these oxynitrides compared to silicon oxide. Moreover, the device obtained with these new gate dielectrics display improved *I-V* and *C-V* characteristics (Yao *et al.* 1994) and increased resistance to ionizing radiation (Chang *et al.* 1984). It was suggested that these improvements are due to a strain relaxation at the interface. Indeed, it is known that there is a large mismatch in the Si atomic density across the interface from 2.2×10^{22} atoms/cm⁻³ in the SiO₂ to 5×10^{22} atoms/cm⁻³ in bulk Si. As the Si density in Si₃N₄ is 4×10^{22} atoms/cm⁻³, between that of bulk Si and SiO₂, the incorporation of nitrogen creates a buffer Si₃N₄-like layer which reduces the mismatch-induced strain at the Si(001)-SiO₂ interface. Finally, the most interesting feature of oxynitrides is that they show an augmented resistance to boron diffusion from doped-polysilicon gates (Green *et al.* 1994).

X-ray photoemission spectroscopy on N 1s core-levels has widely been applied to silicon nitrided oxides (Carr and Buhrman 1993, Bhat, Ahn, Kwong, Arendt and White 1994, Hegde *et al.* 1995, Lu *et al.* 1995, Sutherland *et al.* 1995, Kaluri and Hess 1996, Kamath *et al.* 1997). This technique is sensitive to the concentration of N atoms and can provide distribution profiles when resolved as a function of take-off angle or used in conjunction with chemical etching. These profiles depend on the growth process. In fact, quite a large variety of preparation techniques have been used for nitrogen incorporation in gate oxides: furnace (Kang *et al.* 1994, Carr and Buhrman 1993, Hegde *et al.* 1995, Hussey *et al.* 1996), rapid thermal pro-

cessing (Carr and Buhrman 1993, Yao *et al.* 1994, Green *et al.* 1994, Ting *et al.* 1991, Lu *et al.* 1995, Bhat, Yoon, Kim and Kwong 1994), and downstream plasma (Saito 1996, Chen *et al.* 1995). Various oxidation agents have also been tested: nitrous oxide N_2O (Carr and Buhrman 1993, Bhat, Yoon, Kim and Kwong 1994, Lu *et al.* 1995, Saito 1996), nitric oxide NO (Yao *et al.* 1994, Hegde *et al.* 1995), ammonia NH_3 (Bhat, Ahn, Kwong, Arendt and White 1994), or mixtures of N_2 and O_2 (Kang *et al.* 1994, Saito 1996). There are some differences in the nitrogen distribution resulting from the combination of a given preparation technique and a given oxidation species, but an accumulation of N atoms is often observed at the interface (Carr and Buhrman 1993, Green *et al.* 1994, Bhat, Ahn, Kwong, Arendt and White 1994, Hegde *et al.* 1995, Lu *et al.* 1995, Sutherland *et al.* 1995, Kaluri and Hess 1996, Lu *et al.* 1996, Kamath *et al.* 1997).

XPS N 1s spectra also provide information on the bonding environment of the incorporated N atoms. The experimental spectra (Fig. 2.1) show a broad principal peak (FWHM \approx 1.5 eV), approximately at the same energy as in bulk Si_3N_4 (397.0 eV), which appears to shift to larger binding energies for samples of increasing oxide thickness (Carr and Buhrman 1993, Lu *et al.* 1995, Sutherland *et al.* 1995, Kaluri and Hess 1996, Kamath *et al.* 1997). Such shifts are generally observed at Si(001)- SiO_2 interfaces for Si 2p and O 1s levels, and are attributed to core-hole relaxation or charging effects (Pasquarello *et al.* 1995). However, the N 1s shift with oxide thickness (Fig. 2.2) is found to shift by 0.4 eV more than other oxide peaks (Carr and Buhrman 1993). This fact together with the asymmetric shape of the XPS peak led to the assumption that two different components contributed to this peak and that the N atoms at the interface and in the oxide would have different bonding properties (Carr and Buhrman 1993, Hegde *et al.* 1995, Lu *et al.* 1995, Kaluri and Hess 1996).

The component arising from the interfacial region has invariably been assigned to N atoms bonded to three Si atoms (N-Si₃) because of its proximity to the bulk Si_3N_4 line. On the other hand the oxide component, which is shifted by $\Delta=0.85$ eV to larger binding energies (Lu *et al.* 1995), has given rise to conflicting interpretations (Carr and Buhrman 1993, Hegde *et al.* 1995, Lu *et al.* 1995, Kaluri and Hess 1996), which all rely on a chemical change in the first-neighbor shell of the N atoms. Whereas the occurrence of N-Si₃ configurations is generally accepted, the underlying reasons for the shift with oxide thickness and the asymmetric shape of the N 1s XPS line are still poorly understood.

The presence of a second, well separated, XPS peak is also debated at this stage. Some XPS spectra (as illustrated in main part of Fig. 2.1) show

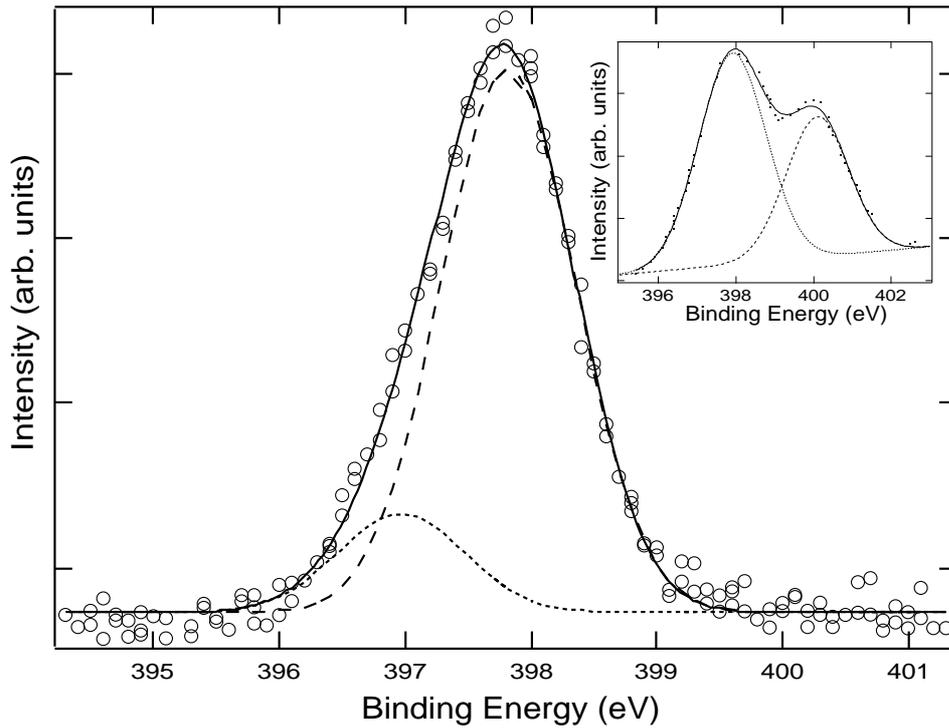


Figure 2.1: N 1s XPS spectrum obtained by Lu *et al.* (1995). The dotted and dashed lines are curve-fitted peaks corresponding to binding energies of 396.97 eV and 387.82 eV, respectively. The solid line is the sum of these two peaks. Inset, N 1s XPS spectrum obtained by Bhat, Ahn, Kwong, Arendt and White (1994). The dotted and dashed lines are curve-fitted peaks corresponding to binding energies of 397.6 eV and 399.9 eV, respectively. The solid line is the sum of these two peaks.

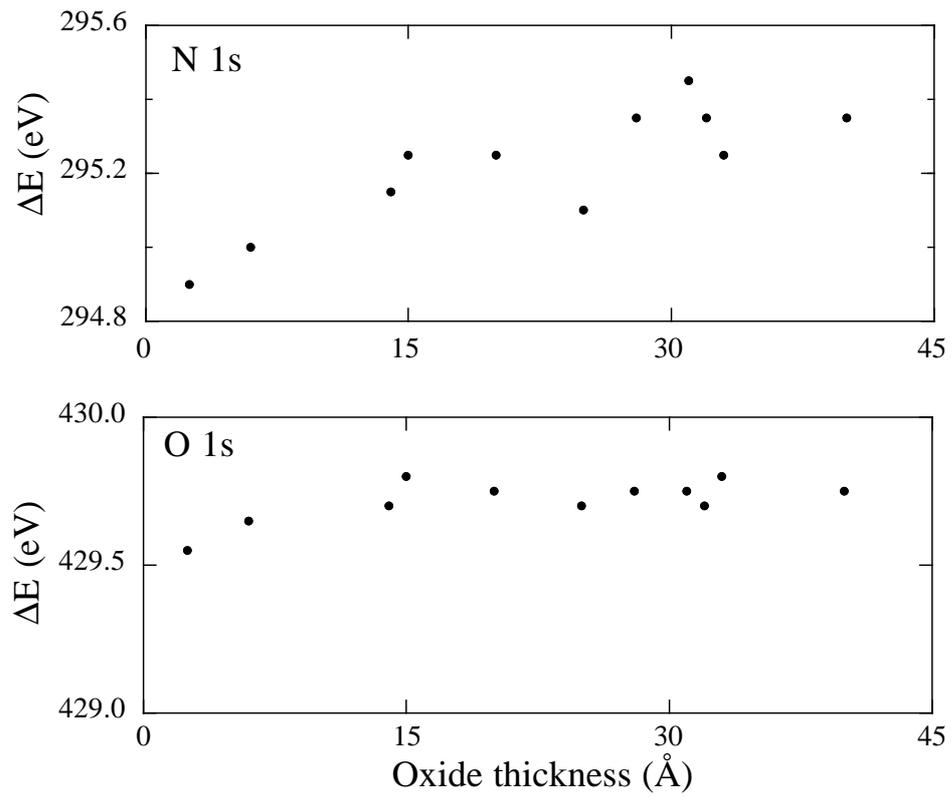


Figure 2.2: Separation of the N 1s and O 1s peaks from the Si 2p peak as a function of the oxide thickness, as obtained by Carr and Buhrman (1993).

no evidence for any other peak in a range of ± 3 eV from the N-Si₃ peak (Lu *et al.* 1995, Sutherland *et al.* 1995), whereas in other experiments (such as the one represented in the inset of Fig. 2.1) show a second distinct peak is observed at $\Delta=2.2$ eV (Bhat, Ahn, Kwong, Arendt and White 1994) or $\Delta=3.0$ eV (Hegde *et al.* 1995), and attributed to N-Si₂O configurations. Therefore, the identification of N bonding configurations involving oxygen atoms remains uncertain.

The purpose of our study is to characterize N atoms incorporated at the Si(001)-SiO₂ interface by establishing a correspondence between their bonding environment and the N 1s core-level shifts measured in photoemission experiments. The remainder of the chapter is organized as follows. In section 2.2, we introduce the different approximations that are currently used to calculate the core-level shifts. We propose two methods that have already been used successfully in the framework of Density Functional Theory, and that allow to determine the initial-state shifts and the full-shifts respectively. Before considering the Si(001)-SiO₂ interface, we assess the status of these theoretical methods by computing of N 1s core-level shifts for small molecules and by comparing the results with experimental data. It is found that it is crucial to include core-hole relaxation effects to obtain a good agreement with experiment. In section 2.3, we first try to gain insight on the possible nitrogen bonding configurations at the interface, by computing the N 1s core-level shifts for a series of selected molecules and comparing with experimental data. With the configurations that fit the experiments, we generate different nitrated interface models by introducing N atoms in various locations with respect to the interface plane. The interface models are relaxed within first principles, and the quality of resulting local structure is analyzed. The N 1s and Si 2p core-level shifts are computed and discussed with respect to experimental photoemission spectra.

As a conclusion, we propose an explanation of the features of N 1s experimental spectra (namely, the shift with oxide thickness of the principal XPS line and of the appearance of two components) in terms of the single first-neighbor configuration N-Si₃. Indeed, we show that the combination of core-hole relaxation and second nearest neighbor effects leads to larger binding energies in the oxide than at the interface, in accord with experimental observations. On one hand, core-hole relaxation affects N 1s shifts differently depending on the distance of the N atoms to the screening Si substrate. On the other hand, a second nearest neighbor environment rich in Si and N atoms as at the interface induces opposite shifts compared to an O rich environment as found in the oxide. For what concerns other possible bonding environments, our results essentially rule out the existence

of N-Si₂O configurations at the interface, since the calculated N 1s shifts ($\Delta=1.5$ eV) is found in a region of the spectrum where generally no peaks are observed in experiments. On the contrary, N-Si₂H configurations can not be excluded since the calculated shift ($\Delta=0.4$ eV) falls right in the range of experimental data. However, we suggest that these appear only in a very small amount in nitrified interface, since these configurations require higher hydrogen concentrations than those actually found in experiments. Ultimately, based on the calculated Si 2p core-level shifts, we propose a new interpretation of the suboxide experimental peaks in terms of both Si-O and Si-N bonds.

2.2 Core-level shifts calculation

X-ray photoemission spectroscopy is routinely used as an experimental tool to investigate unknown microscopic structures at surfaces and interfaces. This technique allows to determine the binding energy of some core states of given atoms (C 1s, O 1s, N 1s, or Si 2p for instance) by measuring the kinetic energy of photoexcited electrons. Hence, it acts as a probe which is mainly sensitive to the local potential, providing a invaluable, though indirect, measure of the chemical environment or oxidation state of the studied atoms. Furthermore, XPS is sensitive to the concentration of the atoms under analysis and can provide distribution profiles when resolved as a function of take-off angle or used in conjunction with chemical etching.

The interpretation of photoemission spectra in terms of the local bonding configuration is most often straightforward, and comparison with molecular equivalents or simple charge-transfer models reveal sufficient to extract the relevant structural information. However, in some situations, the relaxation of core electrons around the hole left by the photoexcited electrons affects sizeably the value of binding energy with respect to these simple pictures. This phenomenon is referred to as *core-hole relaxation effects*. Furthermore, at surfaces and interfaces, the bonding environments may differ significantly (*second nearest-neighbors effects*, *dielectric effects*) from those in molecular analogs, rendering quantitative predictions of core-level shifts as deduced from molecular counterparts or simple models quite difficult. Because of these reasons, it is crucial to obtain a reliable theoretical framework allowing to identify unambiguously the experimental features.

In this section, we introduce the various approximations that are generally used to calculate the core-level shifts. We retain two methods that have already been used successfully in the framework of Density Functional The-

ory. Before studying the Si(001)-SiO₂ interface, these techniques are applied to the calculation of N 1s core-level shifts for small molecules. Comparing the results with experiments, we describe the merits of these methods and point out their limitations.

2.2.1 Theoretical approach

In order to obtain a tool for the interpretation of unidentified XPS spectral features, theoretical schemes have been developed to calculate as accurately as possible the core-level shifts. The major problem lies in describing correctly the escape of the photoexcited core electrons. Indeed, this is a dynamical process implying the relaxation of the other core electrons and to a lower extent the geometrical relation of the surrounding atoms. Thus, various approximations have to be introduced to allow practical computations.

For what concerns geometrical relaxation, two views can be adopted. The sudden (or vertical) approximation assumes that the molecular structure is the same in the initial and final states. On the contrary, the adiabatic approximation postulates that the atoms adiabatically adapt to the photoexcitation. Fortunately, it has been observed that there are no significant differences between computed adiabatic and vertical shifts.

On the electronic relaxation level, the crudest model of the photoexcitation assumes that the other core electrons are not affected. Within this simple picture, in which core-hole relaxation effects are not taken into account, the binding energy is simply related to the corresponding core-level eigenvalue. The results obtained in this way are referred to as initial-state shifts.

In the early attempts, due to quantum chemists (Cederbaum *et al.* 1978, Barth *et al.* 1980, Moncrieff *et al.* 1983), the core-level binding energy was obtained as an energy difference between the ground state energy E_0 and the core-hole final-state energy E_+ of the system.

$$E_b = E_0 - E_+ \quad (2.1)$$

thus including core-hole relaxation effects. However, this scheme which necessitates to perform an all-electron calculation had soon to face size limitations inherent to quantum chemical approaches such as configuration interactions method. These computations quickly become impossible as the number of atoms increases.

In a recent paper, Pedocchi *et al.* (1993) suggested that density functional theory could help to get over this restriction. Using an all-electron Gaussian-orbital approach these authors calculated C 1 shifts within DFT

for a series of molecules, and obtained good agreement with experiment. A step forward was taken by the use of DFT in conjunction with pseudopotentials (Pehlke and Scheffler 1993). In principle, all-electron calculations are more desirable since they describe explicitly the core-valence interactions and core-hole relaxation effects. However, pseudopotential calculations are much easier to perform for large systems such as surfaces and interfaces.

In the pseudopotential approach, only valence electrons are taken into account directly in the calculation, core-valence interactions being accounted for through the use of the PP's. And new difficulties arise in the computation of core-level shifts due to the fact that the core electrons are not considered explicitly. The solution proposed by Pehlke and Scheffler (1993) consists in performing two separate calculations. In a first step the electronic ground state is determined. Then the PP of the studied atom is replaced by another PP which simulates the presence of a screened hole in its core, assuming that the molecular structure is the same in the initial and final states. Moreover, the use of pseudopotentials only allows to determine relative shift in the binding energy Δ with respect to a reference configuration by means of a double difference:

$$\Delta = [E_0 - E_+] - [E_0^{ref} - E_+^{ref}]. \quad (2.2)$$

Blase *et al.* (1994) argued that if the core-hole relaxation effects were similar for both the studied and reference configurations, these would cancel in Δ and thus could be neglected in the calculation. So that relative shifts could be obtained more directly and easily in first-order perturbation theory, by evaluating the expectation value of the local self-consistent potential on the atomic core-state orbitals ψ^0 :

$$\Delta = \langle \psi^0 | V_{KS} | \psi^0 \rangle - \langle \psi^0 | V_{KS}^{ref} | \psi^0 \rangle. \quad (2.3)$$

More recently, Rohlfiing *et al.* (1997) exploited the same argument to compute core-level shifts from quasiparticle energies within *GW* approximation. However, this scheme implies the explicit treatment of the required core states as valences states in the pseudopotential. Fortunately, as the sensitivity of core levels of a particular atom to core states of neighboring atoms is extremely small, the use of the so-modified PP can be restricted to only those atoms whose core-state energies are to be calculated. It should be mentioned that this method, as well as that of Blase *et al.*, achieved only limited success, and from our own experience we can say that most generally core-hole relaxation effects should be taken into account.

2.2.2 Small molecules

The use of density functional theory for the calculation of core-level shifts in molecules proved successful in the case of C 1s (Pedocchi *et al.* 1993) and Si 2p (Pasquarello *et al.* 1996) levels. In order to investigate the accuracy of this approach in the case of N 1s core levels, we considered a set of molecules containing N atoms in various bonding configurations.¹ The molecules which we considered in this study are listed in Tables 2.1 and 2.2.

All the results were obtained within the local density approximation to DFT. Only valence electrons are explicitly considered using pseudopotentials to account for core-valence interactions. A normconserving PP is employed for Si and Cl atoms (Bachelet *et al.* 1982), whereas the H, O, C, and N atoms are described by ultrasoft PP's (Vanderbilt 1990). The exchange and correlation energy was evaluated in the local density approximation (Perdew and Zunger 1981). The N 1s core-level shifts were calculated both within the initial-state approximation (Eq. 2.3) and including core-hole relaxation (Eq. 2.2).

Energy minimization were performed with Car-Parrinello molecular dynamics. The convergence was accelerated making use of a preconditioned damped dynamics technique (Tassone *et al.* 1994). A detailed description of the technical implementation can be found elsewhere (Pasquarello *et al.* 1992, Laasonen, Pasquarello, Car, Lee and Vanderbilt 1993). In order to obtain good structural properties, the wave functions and the augmented electron density were expanded on plane-wave bases defined by cutoffs of 25 and 150 Ry, respectively. In the calculations of the core shifts, it was necessary to increase the cutoff for the wave functions to 30 Ry to reach convergence. The Brillouin zone was sampled using the Γ -point.

The atomic coordinates of the molecules under consideration were relaxed using a periodically repeated cubic cell with a side $L=24$ bohr. For this cell size, the effects of interactions between neighboring cells (see section 1.5.2) on the ground state can be safely neglected. Structural parameters obtained at the end of these relaxation processes are given in Table 2.1, together with the corresponding experimental values (Hellwege and Hellwege 1976, Hellwege and Hellwege 1987, Kuchitsu 1992, Kuchitsu 1995). The molecular structure is overall very well reproduced in our calculations.

¹Unfortunately, we were not able to find in the literature measured N 1s shifts for molecules containing N-Si bonds.

Table 2.1: Bond lengths (in Å) and bond angles (in °) as obtained in the present work compared to experiment (Hellwege and Hellwege 1976, Hellwege and Hellwege 1987, Kuchitsu 1992, Kuchitsu 1995)

Molecule	Parameter	Present Theory	Experiment
NH ₃	$d(\text{N-H})$	1.036	1.012
	$\angle(\text{H-N-H})$	108.0	106.7
NH ₂ CH ₃	$d(\text{N-C})$	1.438	1.471
	$d(\text{N-H})$	1.031	1.010
	$d(\text{C-H})$	1.112	1.099
	$\angle(\text{H-N-H})$	108.8	107.1
	$\angle(\text{H-C-N})$	112.2	110.3
	$\angle(\text{H-C-H})$	106.6	108.0
NH(CH ₃) ₂	$d(\text{N-C})$	1.427	1.462
	$d(\text{N-H})$	1.028	1.019
	$d(\text{C-H}_s)$	1.113	1.084
	$d(\text{C-H}')$	1.116	1.098
	$d(\text{C-H}'')$	1.116	1.098
	$\angle(\text{C-N-C})$	118.2	112.2
	$\angle(\text{H}_s\text{-C-H}')$	107.1	109.0
	$\angle(\text{H}_s\text{-C-H}'')$	106.6	109.0
	$\angle(\text{H}'\text{-C-H}'')$	106.4	107.2
	$\angle(\text{H-N-C})$	113.7	108.9
	$\angle(\text{N-C-H}_s)$	113.8	109.7
$\angle(\text{N-C-H}')$	110.3	108.2	
$\angle(\text{N-C-H}'')$	112.3	113.8	
N(CH ₃) ₃	$d(\text{N-C})$	1.430	1.451
	$d(\text{C-H}_s)$	1.115	1.109
	$d(\text{C-H}_a)$	1.115	1.088
	$\angle(\text{C-N-C})$	114.9	110.9
	$\angle(\text{N-C-H}_s)$	114.4	111.7
	$\angle(\text{N-C-H}_a)$	110.7	110.1
	$\angle(\text{H}_s\text{-C-H}_a)$	106.8	108.1
	$\angle(\text{H}_a\text{-C-H}'_a)$	107.0	108.6
NH ₂ COH	$d(\text{N-C})$	1.344	1.352
	$d(\text{N-H}')$	1.033	1.002
	$d(\text{N-H}'')$	1.030	1.002
	$d(\text{C-O})$	1.231	1.219
	$d(\text{C-H})$	1.126	1.098

	$\angle(\text{H}'\text{-N-H}'')$	119.7	121.6
	$\angle(\text{H}'\text{-N-C})$	119.0	118.5
	$\angle(\text{H}''\text{-N-C})$	121.3	120.0
	$\angle(\text{N-C-O})$	124.7	124.7
	$\angle(\text{N-C-H})$	113.0	112.7
	$\angle(\text{O-C-H})$	122.3	122.5
NO ₂	$d(\text{N-O})$	1.212	1.193
	$\angle(\text{O-N-O})$	133.1	134.1
N ₂ O	$d(\text{N-N})$	1.142	1.128
	$d(\text{N-O})$	1.201	1.184
ClNO	$d(\text{N-O})$	1.167	1.139
	$d(\text{N-Cl})$	1.949	1.975
	$\angle(\text{O-N-Cl})$	113.6	113.3

Subsequently, for each relaxed structure, initial-state and full shifts were calculated. The initial-state shifts were obtained by calculating the expectation value of the local self-consistent potential in Fourier space. Full shifts were obtained by performing a separate calculation as described above in the presence of a negative background assuring charge neutrality. In order to eliminate the effect of the use of this back ground, and of periodic boundary conditions, we calculated the shifts for varying cell size: $L=20, 22$ and 24 bohr. Making use of the procedure proposed by Makov and Payne (1995), the shift were extrapolated to infinite cell size:

$$\Delta(L) = \Delta + \frac{\text{cst}}{L^3} + \mathcal{O}(L^{-5}) \quad (2.4)$$

where Δ is the desired shift. In our calculation, the values obtained in this way differed at most by a few tenth of an eV from the values obtained with the largest cell-size.

We report calculated initial-state and full shifts together with the experimental values (Jolly *et al.* 1984) in Table 2.2. The agreement with experimental values is noticeably improved when final states are accounted for. Thus, it appears crucial to consider core-hole relaxation effects. Overall, the agreement between experiment and theory is very good with values differing by less than 0.3 eV for a broad range of shifts.

In summary, we have described the underlying approximations that allow to compute core-level shifts in the framework of Density Functional Theory. The initial-state shifts are obtained by the expectation value of the local

Table 2.2: Comparison of calculated N 1s initial-state Δ^{init} and full Δ^{full} shifts with experiment Δ^{expt} for a series of molecules. The shifts (expressed in eV) are given with respect to NH_3 . Experimental data are taken from Jolly *et al.* (1984).

Molecule	Δ^{init} (eV)	Δ^{full} (eV)	Δ^{expt} (eV)
NH_3	0.00	0.00	0.0
NH_2CH_3	-0.06	-0.57	-0.5
$\text{NH}(\text{CH}_3)_2$	0.08	-0.58	-0.7
$\text{N}(\text{CH}_3)_3$	0.33	-0.61	-0.8
NH_2COH	1.73	1.13	0.8
NO_2	6.91	7.24	7.3
N_2O (N*NO)	3.46	3.30	3.1
(NN*O)	6.77	7.08	7.0
CINO	6.53	5.68	5.8

self-consistent potential on the atomic core-state. We have also illustrated how core-hole relaxation effects can be taken into account despite the use of pseudopotentials, by performing two separate calculations. These two techniques have been successfully applied to the calculation of N 1s core-level shifts of small molecules. It was found that it is crucial to include core-hole relaxation effects to obtain a good agreement with experiment.

2.3 Core-level shifts at the Si(001)-SiO₂ nitrated interface

Now that we have obtained a reliable method to calculate N 1s core-level shift, we turn to the problem of interpreting XPS experimental features for the Si(001)-SiO₂ interface. So, in this section, we first compute the N 1s core-level shifts for a series of selected molecules that include N-Si, N-O or N-H bonds. Comparison with experimental data for the interface allows to reject some configurations for the incorporated nitrogen atoms. With the retained configurations, we generate various nitrated interface models by varying also the nitrogen locations with respect to the interface plane. After atomic relaxation, we analyze the quality of the resulting representation of the local structure. We compute the N 1s core-level shifts for all these models, and provide an explanation of photoemission spectra. The Si 2p

core-level shifts are also investigated.

2.3.1 Auxiliary test molecules

Before addressing the actual interface models and to gain insight on the possible structure of incorporated N atoms, we compute the N 1s core-level shifts for a set of test molecules in which N atoms are bonded either to Si atoms only, or to Si and O atoms, or to Si and H atoms.

The N 1s core-level of N(SiH₃)₃, where the N atom is bonded to three Si atoms as in Si₃N₄, is taken as reference. By inserting O atoms in n (for $n=1,2,3$) of the N-Si bonds, we generated molecules containing N atoms with different combinations of Si and O nearest neighbors: (H₃SiO) _{n} N(SiH₃) _{$3-n$} . And, by substituting H atoms to n (for $n=1,2$) of the SiH₃ groups around the N atom in N(SiH₃)₃, we created molecules with N-Si₂H and N-SiH₂ bonding configuration: H _{n} N(SiH₃) _{$3-n$} .

After atomic relaxation we obtained N-Si bonds of 1.71 ± 0.01 Å, N-O bonds of 1.37 ± 0.01 Å, and N-H bonds of 1.03 ± 0.01 Å, within 1% from experimental values (Hellwege and Hellwege 1976, Hellwege and Hellwege 1987, Kuchitsu 1992, Kuchitsu 1995).

Calculated shifts including core-hole relaxation are given in Table 2.3. These shifts depend linearly on the number of O or H nearest neighbors, with a shift to larger binding energies of approximately $\Delta=1.8$ eV per N-O bond and $\Delta=0.4$ eV per N-H bond. The shift for N-O₃ bonding configurations is too large to account for the peaks observed at the Si(001)-SiO₂ interface and will not be considered any longer. The N-Si₂H and N-SiH₂ bonding configurations imply a concentration of hydrogen at the interface similar to that of nitrogen. This is not what is actually found in experiments (Green *et al.* 1994, Lu *et al.* 1996, Tang *et al.* 1996). Thus, we only investigated the N-Si₂H configuration for its theoretical interest, while the N-SiH₂ configuration was not taken into account any more.

2.3.2 Structural model

We adopt as a starting point of our nitrided interface study one of the Si(001)-SiO₂ models generated by Pasquarello *et al.* (1995). The choice of the host model is not critical, because the core-level shifts are affected by the host model only to the extent that it determines the dielectric environment. This is confirmed by the small variations of the Si 2p shifts of a given oxidation state in different models (Pasquarello *et al.* 1995).

In this model, the atomic structure is obtained by attaching tridymite

Table 2.3: Relative N $1s$ shifts for a series of test-molecules. The shifts (expressed in eV) are given with respect to $\text{N}(\text{SiH}_3)_3$.

Molecule	N Configuration	$\Delta(\text{eV})$
NH_2SiH_3	N-SiH ₂	0.74
$\text{NH}(\text{SiH}_3)_2$	N-Si ₂ H	0.38
$\text{N}(\text{SiH}_3)_3$	N-Si ₃	0.00
$(\text{H}_3\text{SiO})\text{N}(\text{SiH}_3)_2$	N-Si ₂ O	1.77
$(\text{H}_3\text{SiO})_2\text{N}(\text{SiH}_3)$	N-SiO ₂	3.78
$(\text{H}_3\text{SiO})_3\text{N}$	N-O ₃	5.55

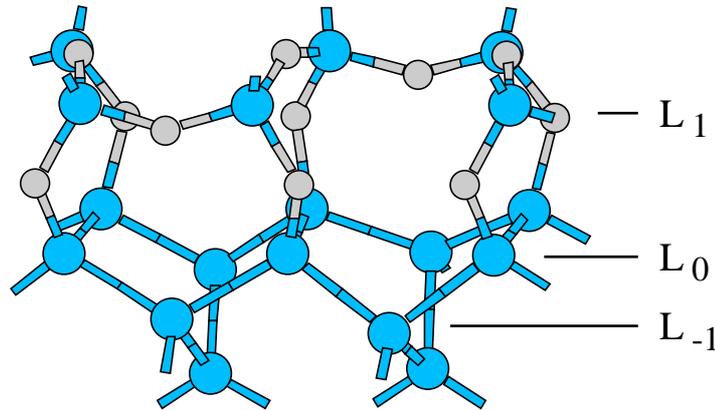


Figure 2.3: Model of the $\text{Si}(001)\text{-SiO}_2$ interface before N incorporation (from (Pasquarello *et al.* 1995)). Si and O atoms are represented in dark and pale grey, respectively. A nomenclature is introduced to designate the layers in which N is incorporated : L_0 is the interface Si layer, $L_{>0}$ are the Si-O layers on the oxide side of the interface, and $L_{<0}$ are the Si layers on the bulk-silicon side of the interface.

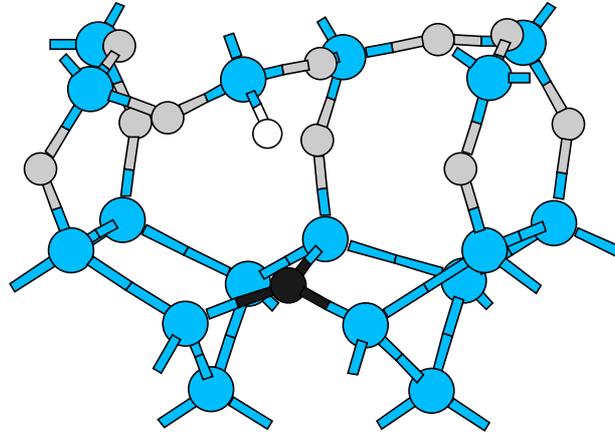


Figure 2.4: Model I for the nitrided $\text{Si}(001)\text{-SiO}_2$ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located right at the interface layer (L_0) with a N-Si_3 configuration. The residual dangling bond is saturated by one hydrogen atom (in white).

(Wyckoff 1963) to bulk terminated $\text{Si}(001)$. To describe this system we used a periodically repeated orthorhombic cell containing a square interface unit $\sqrt{8} \times \sqrt{8}$ of side $a=10.82 \text{ \AA}$, based on the theoretical lattice constant of Si. The c -axis of tridymite is aligned to the Si $[110]$ direction, and the tridymite $[100]$ direction is taken parallel to the Si $[\bar{1}\bar{1}0]$ direction. In order to match the two periodic structures, the tridymite is compressed by 12 % and 7 % along its $[100]$ and $[001]$ directions, respectively. In this geometry the oxide is built up by alternating planes of O atoms and Si-O chains parallel to the interface. The transition between the oxide and the Si substrate is abrupt. The dangling bonds of every Si atom at the interface are saturated by forming a bond to the oxide and by dimerization [see Fig. 2.3].

This structure does not present unsaturated dangling bonds, in accord with the extremely low density of defect states measured at this interface. The dimension of the cell in the direction orthogonal to the interface is $c=19.05 \text{ \AA}$, containing 5 monolayers of SiO_2 (6.4 \AA) and 6 monolayers of Si (7.7 \AA). The extremities are saturated with H atoms.

We obtain eight nitrided interface structures by incorporating in various ways a single N atom in this interface model. Because N atoms at the interface do not deteriorate electrical properties such as interface defect density or oxide fixed charge (Chang *et al.* 1984, Carr and Buhrman 1993, Hao

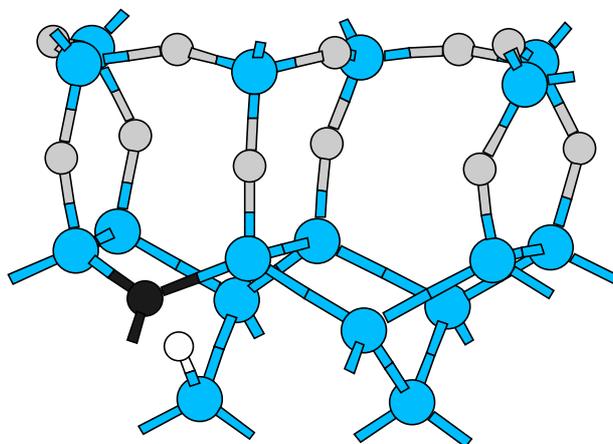


Figure 2.5: Model II for the nitrided Si(001)-SiO₂ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located in the bulk silicon (L_{-1}) with a N-Si₃ configuration. The residual dangling bond is saturated by one hydrogen atom (in white).

et al. 1995, Kang *et al.* 1994), we only retain in our study neutral structural models with threefold coordinated N atoms and without any unsaturated dangling bond. To generate such models, we proceed as follows. For the nitrogen bonding configurations involving Si and O atoms, the N atoms are introduced substitutionally to Si atoms. Eventual dangling O atoms are removed and residual dangling bonds are saturated by H atoms. Whereas, when nitrogen is bonded to Si and H atoms, the N atoms are introduced in place of O atoms, the remaining bond on the nitrogen atom being satisfied by one H atom.

We generated three models with N-Si₃ configurations at varying distance from the interface plane. In model I and II (Figs. 2.4 and 2.5), N atoms substitute Si atoms in the first [L_0 in Fig. 2.3] and second Si layer (L_{-1}) of the substrate, respectively. In model III (Fig. 2.6), the N atom replaces a Si atom in the oxide (L_1). In this case, we also removed all the nearest O atoms to recover a N-Si₃ configuration.

Models IV and V (Figs. 2.7 and 2.8) contain N-Si₂O configurations with substitutions in layers L_0 and L_1 . In model VI (Fig. 2.9), we investigated a N-SiO₂ configuration in the oxide (L_1).

Finally, we also considered two models with N-Si₂H configurations. In model VII (Fig. 2.10), N atom substitutes an O atom located between the

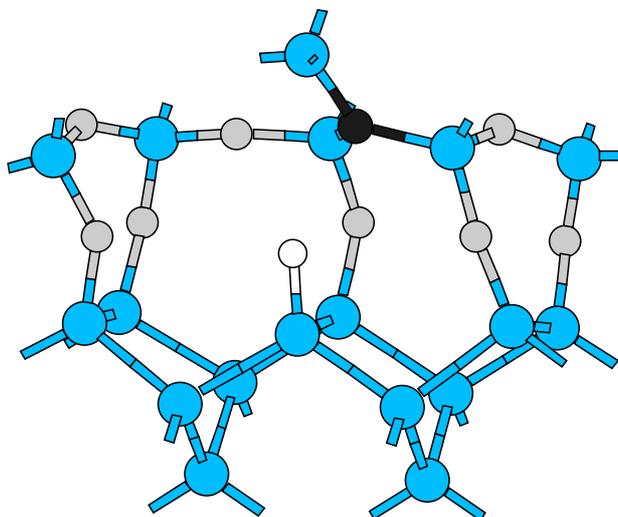


Figure 2.6: Model III for the nitrated $\text{Si}(001)\text{-SiO}_2$ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located in the oxide (L_1) with a N-Si_3 configuration. The residual dangling bond is saturated by one hydrogen atom (in white).

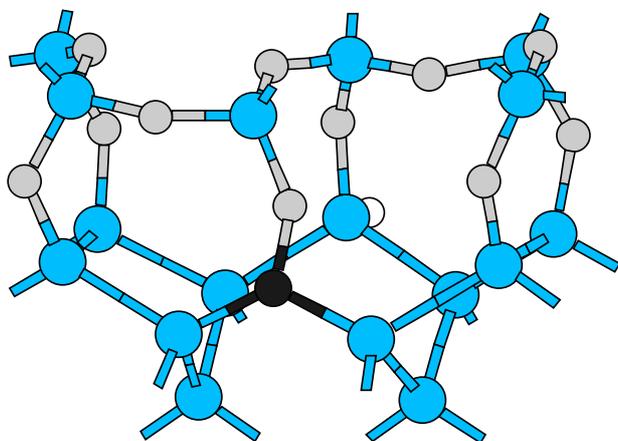


Figure 2.7: Model IV for the nitrated $\text{Si}(001)\text{-SiO}_2$ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located right at the interface (L_0) with a $\text{N-Si}_2\text{O}$ configuration. The residual dangling bond is saturated by one hydrogen atom (in white).

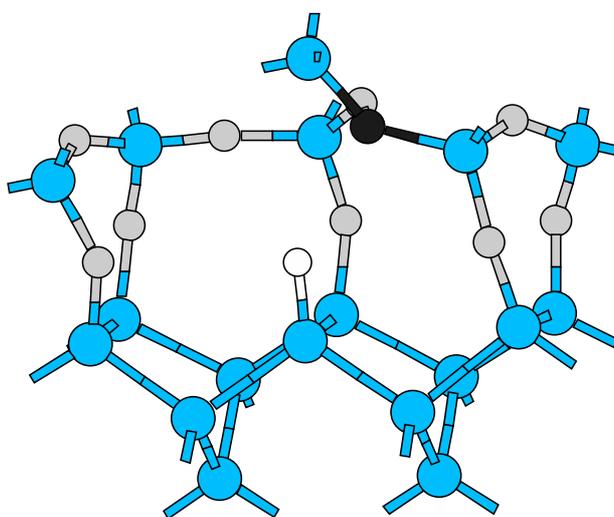


Figure 2.8: Model V for the nitrided Si(001)-SiO₂ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located in the oxide (L₁) with a N-Si₂O configuration. The residual dangling bond is saturated by one hydrogen atom (in white).

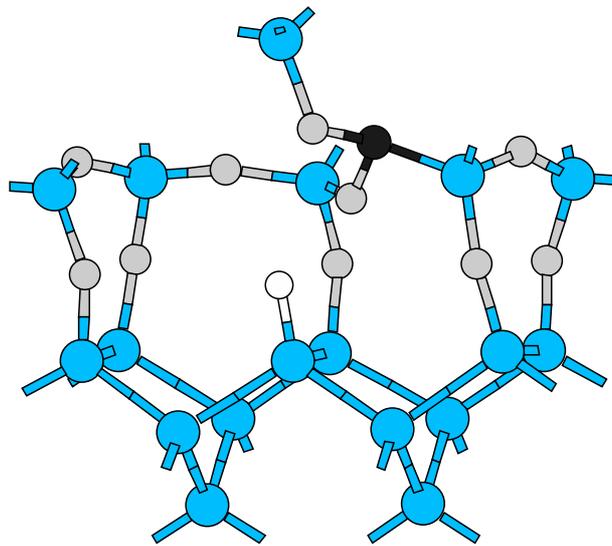


Figure 2.9: Model VI for the nitrided $\text{Si}(001)\text{-SiO}_2$ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located in the oxide (L_1) with a N-SiO_2 configuration. The residual dangling bond is saturated by one hydrogen atom (in white).

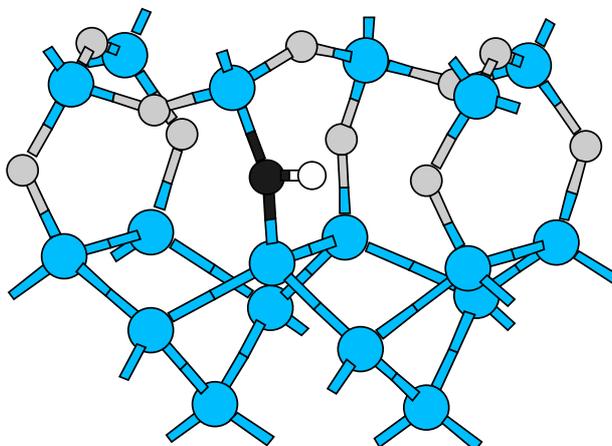


Figure 2.10: Model VII for the nitrided Si(001)-SiO₂ interface. The nitrogen atom (in black) has been introduced substitutionally to a silicon located in the oxide (L₁) with a N-Si₂H configuration.

interface layer (L₀) and the first oxide layer (L₁). In model VIII (Fig. 2.11), the N atom replaces a O atom right in the first oxide layer (L₁).

After the incorporation of the N atoms, all the oxide atoms as well as the first three Si layers were fully relaxed. In Figs. 2.4-2.11, we illustrate the relaxed structures for the 8 different models. Distributions of structural parameters as found at the end of the relaxation process for the various models are given in Tables 2.4-2.6 in terms of minimum, mean, maximum and standard deviation.

We first analyze the local structure around the incorporated N atom (see Table 2.4). In the oxide (L₁), when nitrogen is introduced substitutionally to silicon (model III, V, and VI), the N-Si bond lengths are 1.74 ± 0.02 Å in good agreement with other calculations (Jing *et al.* 1994). The N-Si bond lengths in the Si substrate (L₀ and L₋₁, for model I, II, and IV) are found to be slightly larger, 1.82 ± 0.02 Å, which should be attributed to the local strain at the interface. On the contrary, when nitrogen is introduced substitutionally to oxygen (model VII and VIII), the N-Si bond lengths is compressed up to 1.66 ± 0.01 Å when the Si atom is located in the oxide while it turns back to 1.74 Å when the Si atom is in the substrate. We found N-O and N-H bond lengths of 1.44 ± 0.02 Å and 1.04 ± 0.01 Å respectively. In the present study it was necessary to use an enlarged interface unit with respect to the work of Pasquarello *et al.* (1995) in order to minimize steric constraints due

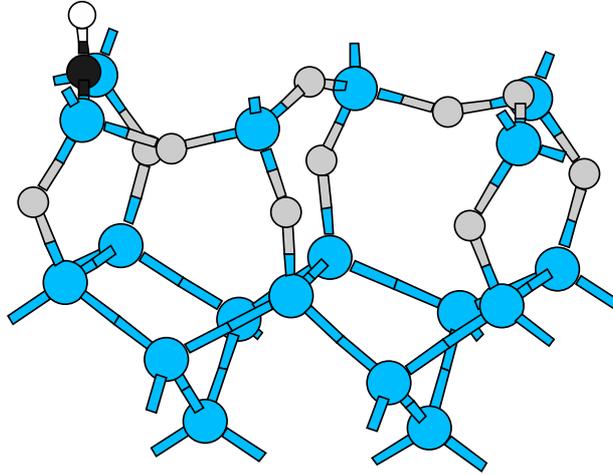


Figure 2.11: Model VIII for the nitrided $\text{Si}(001)\text{-SiO}_2$ interface. The nitrogen atom (in black) has been introduced substitutionally to a oxygen located in the oxide (L_1) with a $\text{N-Si}_2\text{H}$ configuration.

Table 2.4: Lengths of the 3 bonds (N-X_1 , N-X_2 , N-X_3) formed by the incorporated N atom in the various interface models. The bond lengths are expressed in Å.

Model	X_1	X_2	X_3	$d(\text{N-X}_1)$	$d(\text{N-X}_2)$	$d(\text{N-X}_3)$
I	Si	Si	Si	1.82	1.81	1.81
II	Si	Si	Si	1.83	1.83	1.84
III	Si	Si	Si	1.73	1.73	1.74
IV	Si	Si	O	1.82	1.81	1.44
V	Si	Si	O	1.73	1.75	1.44
VI	Si	O	O	1.76	1.42	1.42
VII	Si	Si	H	1.67	1.73	1.05
VIII	Si	Si	H	1.66	1.66	1.04

Table 2.5: Bond length distribution in terms of minimum, mean, maximum and standard deviation for the various models of interface. We distinguish the bonds formed by Si atoms in the bulk silicon $d(\text{Si-Si})$ and in the oxide $d(\text{Si-O})$. The bond lengths are expressed Å.

Model	$d(\text{Si-Si})$				$d(\text{Si-O})$			
	Min.	Mean	Max.	σ	Min.	Mean	Max.	σ
I	2.24	2.35	2.46	0.04	1.57	1.61	1.65	0.02
II	2.24	2.34	2.44	0.04	1.57	1.60	1.65	0.02
III	2.28	2.34	2.45	0.03	1.58	1.61	1.65	0.02
IV	2.27	2.35	2.46	0.03	1.58	1.61	1.65	0.02
V	2.28	2.34	2.41	0.02	1.57	1.61	1.65	0.02
VI	2.27	2.34	2.44	0.03	1.59	1.61	1.66	0.02
VII	2.28	2.34	2.49	0.04	1.59	1.61	1.68	0.02
VIII	2.28	2.34	2.48	0.04	1.59	1.61	1.65	0.02

to the periodicity in the interface plane and to allow full local relaxation around the incorporated N atom. For instance, when a smaller interface cell was used, we found unrealistic large variations (± 0.1 Å) in the N-Si bond lengths for a given N-Si₃ configuration.

We also analyze the effect of nitrogen incorporation on the structure of the Si substrate and of the oxide. First, the Si-Si and Si-O bond distances (see Table 2.5) are found to be rather constant, about 2.34 Å and 1.61 Å respectively (less than 1 % difference with their experimental values in bulk Si (Donohue 1974) and SiO₂ (Levien *et al.* 1980)). A variation of about 5 % appears near the region where the nitrogen has been incorporated, inducing a strain of the perfect lattice.

Finally, in Table 2.6, we consider the angles formed by Si atoms and by O atoms. The Si-Si-Si angles are found only in the substrate. The deviation with respect to the tetrahedral angle is due to the presence of dimer at the interface in the original model (Pasquarello *et al.* 1995). This effect is augmented when the N atom is introduced in the substrate. The Si-Si-O which is typical of the interface naturally deviates the most though still acceptably from 109.4°. The O-Si-O bond angle distribution is peaked at the tetrahedral angle. In the oxide, most of the stress seems to be relieved by the rather flat distribution of the Si-O-Si angle, which ranges from 120° and 170°.

This analysis suggests that our relaxed models give a good representation

Table 2.6: Bond angle distribution in terms of minimum, mean, maximum and standard deviation for the various models of interface. We distinguish the angles formed by Si atoms in the bulk silicon $\angle(\text{Si-Si-Si})$, at the interface $\angle(\text{Si-Si-O})$, and in the oxide $\angle(\text{O-Si-O})$; and the angles formed by O atoms in the oxide $\angle(\text{Si-O-Si})$. The bond angles are expressed in $^\circ$.

Model	$\angle(\text{Si-Si-Si})$				$\angle(\text{Si-Si-O})$			
	Min.	Mean	Max.	σ	Min.	Mean	Max.	σ
I	85.0	108.8	127.2	5.8	105.0	111.2	116.9	3.1
II	88.9	109.2	130.6	6.2	105.0	110.0	117.3	3.5
III	91.0	109.0	126.9	5.6	100.5	111.5	133.0	8.0
IV	87.6	109.0	123.0	5.3	100.7	111.8	119.0	4.7
V	90.8	109.2	123.8	5.3	101.9	112.1	126.8	6.5
VI	91.6	109.0	123.3	5.2	105.6	111.6	124.0	4.4
VII	93.2	109.1	120.9	5.4	106.3	111.8	117.6	5.4
VIII	92.3	109.0	120.8	5.5	105.6	111.4	118.5	4.9
Model	$\angle(\text{O-Si-O})$				$\angle(\text{Si-O-Si})$			
	Min.	Mean	Max.	σ	Min.	Mean	Max.	σ
I	99.9	109.4	116.0	3.1	127.5	141.1	159.9	8.0
II	102.9	109.5	116.1	2.7	131.2	139.0	158.1	6.0
III	100.2	109.2	118.2	3.7	125.0	140.4	167.3	10.8
IV	99.5	109.5	122.7	3.6	123.4	137.1	149.7	6.4
V	95.6	109.3	118.7	3.9	128.6	139.8	163.5	8.9
VI	99.9	109.5	117.7	3.0	128.2	141.6	173.4	8.8
VII	103.6	109.4	117.3	3.3	125.5	142.9	157.4	9.8
VIII	103.6	109.6	116.7	3.0	129.1	140.3	156.1	8.7

of the local structure for the various nitrogen-bonding configurations.

2.3.3 N 1s core-level shifts at the interface

N 1s shifts for the eight models described above were calculated including core-hole relaxation effects and are given in Fig. 2.12 as a function of the relaxed position z of the N atom with respect to the interface plane. We took as a reference the shift of the N-Si₃ configuration in which the N atom is located most deeply in the Si substrate (model II). The N 1s shifts in the different models could be compared by aligning the bulk Si 2p line. Note that although the N atoms in models I and II were incorporated in different Si layers (L₀ and L₋₁), their z position in the relaxed structure turned out to be close, and so did the values of their core shifts. From Fig. 2.12, it is evident that the shifts are strongly affected by first nearest neighbors. The presence of an O nearest neighbor (models IV and V) yields shifts to higher binding energies of $\Delta=1.5$ eV with respect to corresponding N-Si₃ configurations. A second oxygen nearest neighbor (model VI) brings this shift to $\Delta=3.5$ eV. The N-Si₂H configurations (model VII and VIII) lead to a shift of $\Delta=0.4$ eV from the reference. These results are in good accord with the shifts obtained for the test molecules (Table 2.3). Such a large separation between the shifts of N-Si₃ and N-Si₂O configurations virtually rules out the possibility that both configurations contribute to the principal XPS peak. According to our analysis a peak resulting from N-Si₂O configurations should lie at $\Delta=1.5$ eV. A second peak was observed at $\Delta=2.2$ eV by Bhat, Ahn, Kwong, Arendt and White (1994). However, since such a peak was not seen in subsequent experiments (Lu *et al.* 1995, Sutherland *et al.* 1995, Kaluri and Hess 1996), we are inclined to assume that N-Si₂O configurations are unlikely at nitrided Si(001)-SiO₂ interfaces. On the contrary, the shifts obtained in the model with N-Si₂H bonding configuration fall right in the experimental range. Thus we cannot exclude the presence of this configuration for N atoms incorporated at the interface. However, if all the N atoms were bonded to one H atom, it would imply a concentration of hydrogen at the interface of the same order of that of nitrogen, which is not the case at all in experiments (Green *et al.* 1994, Lu *et al.* 1996, Tang *et al.* 1996). In consequence, we can say that though the N-Si₂H configuration can be found when N is incorporated at the interface, it should be considered to appear only occasionally. Hence, it is not sufficient to explain the experimental features.

The N atoms incorporated in the oxide (layer L₁) give core shifts which are displaced by $\Delta\approx 0.4$ eV with respect to N atoms in corresponding config-

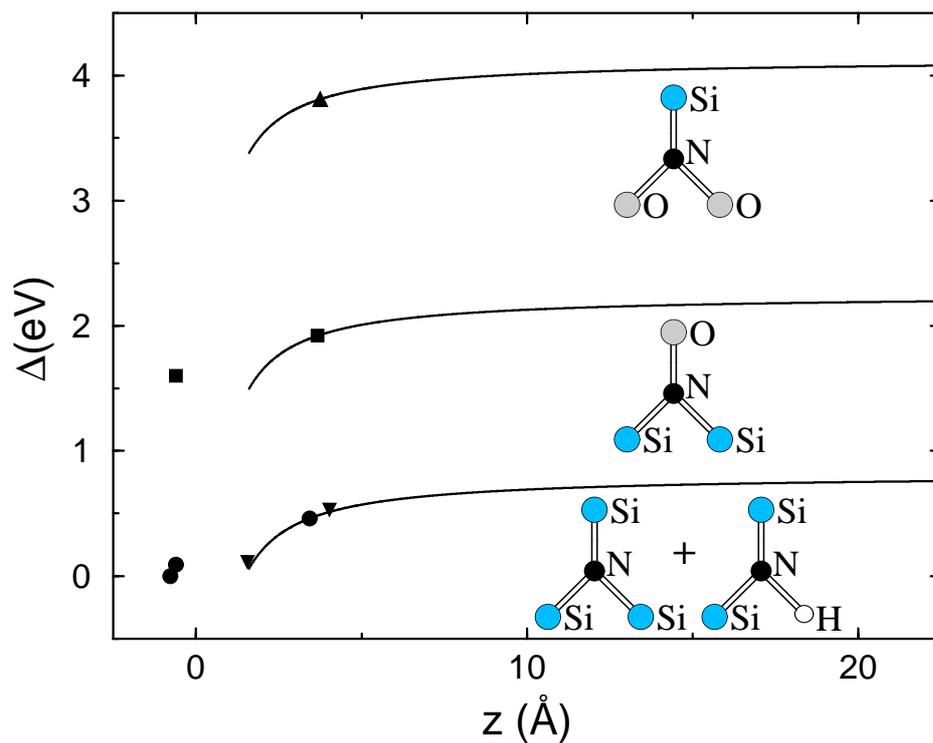


Figure 2.12: N 1s core-level shifts Δ at $\text{Si}(001)\text{-SiO}_2$ interfaces calculated for N-Si₃ (circle), N-Si₂O (square), N-SiO₂ (upward triangle), and N-Si₂H (downward triangle) configurations at different distances z from the interface. Continuous line extrapolations result from classical electrostatics for the case of a semi-infinite oxide.

urations at the interface. This effect has the same size for N–Si₃ and N–Si₂O configurations and is a consequence of the dependence of core-hole relaxation on the distance to the screening Si substrate (Pasquarello *et al.* 1995).

This dependence on z can be easily extrapolated within a simple classical model. We consider a system composed by three dielectrics, separated by abrupt interfaces: a semi-infinite silicon substrate (for $z < 0$ and with a dielectric constant ϵ_1), an oxide slabe of thickness d ($0 < z < d$, ϵ_2), and the vacuum ($z > d$, ϵ_3). We calculate the correction Δ to the core-hole relaxation energy due to the presence of dielectrics. In order to satisfy Maxwell boundary conditions, an infinite series of image charges is required. For an excitation at z in the oxide ($0 < z < d$) we obtain:

$$\Delta = \frac{1}{2} \frac{e^2}{\epsilon_2} \sum_{n=0}^{\infty} (\eta\xi)^n \left[\frac{\eta}{|2z - 2nd|} + \frac{\xi}{|(2n+1)d - 2z|} + \frac{2\eta\xi}{(2n+2)d} \right] \quad (2.5)$$

where e is the electronic charge and where

$$\eta = \frac{\epsilon_2 - \epsilon_1}{\epsilon_2 + \epsilon_1} \quad \text{and} \quad \xi = \frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}. \quad (2.6)$$

For an excitation at z in the silicon substrate ($z < 0$):

$$\Delta = \frac{1}{2} \frac{e^2}{\epsilon_1} \left[-\frac{\eta}{|2z|} + (1 + \eta)(1 - \eta) \sum_{n=0}^{\infty} \frac{\eta^n \xi^{n+1}}{(2n+2)d - 2z} \right] \quad (2.7)$$

The energy Δ contributes to the shifts as an additive correction provided the sign convention is chosen in such a way that the shifts in the oxide are positive with respect to those in the substrate. This simple model generally provides a good approximation at distances of a few bond lengths (Pasquarello *et al.* 1995).

The result of this extrapolation in the oxide is shown in Fig. 2.12, where we have used $\epsilon_1=12$, $\epsilon_2=2.1$ and $\epsilon_3=1$ to describe the Si–SiO₂–vacuum system. We took a value of $d=50$ Å for the oxide thickness. This approximation is justified since the shifts at the matching points, as calculated within classical electrostatics, varied by less than 0.1 eV when the oxide–vacuum interface was moved from its position in the model to large distances.

However, this dependence on distance to the interface only partially explains the observation of two components of different spatial origin in the principal XPS peak. Although this effect can account for a difference of 0.8 eV between shifts at the interface and shifts in the oxide, we expect that averages over actual N concentration profiles will narrow the separation between peaks resulting from the two contributions. It therefore appears

difficult to interpret quantitatively the measured separation of $\Delta=0.85$ eV (Lu *et al.* 1995) in terms of core-hole relaxation alone. This standpoint is further supported by the fact that as a function of oxide thickness the N $1s$ peak shifts to larger binding energies by 0.4 eV more than other oxide peaks (Carr and Buhrman 1993). In the work of Carr and Buhrman (1993), structural changes such as induced by strain could be ruled out and it was suggested that this effect was due to a chemical change in N first nearest neighbors. We here provide a different interpretation in terms of a change in second nearest neighbors.

In order to investigate such effects, we turned again to test-molecules. We considered the molecules $\text{N}(\text{SiH}_3)_{3-n}(\text{SiH}_2\text{R})_n$, where R is taken to be SiH_3 , OSiH_3 or $\text{N}(\text{SiH}_3)_2$, and $n=1,2,3$. The central N always has the same nearest neighbor configuration (N– Si_3). Second nearest neighbor effects are studied by varying R and n . Let's focus first on the case $n=1$ corresponding to a single modification in the shell of second nearest neighbors (Table 2.7). Second nearest neighbor O and N atoms give small, but opposite shifts. This effect scales approximately linearly with n in the case of multiple substitutions. This analysis suggests that N $1s$ levels shift to lower binding energies in an environment rich in Si and N atoms and to higher binding energies in an O rich environment. Close to the interface one finds typical N concentrations of one monolayer in a transition region of $\sim 15\text{\AA}$ (Green *et al.* 1994, Lu *et al.* 1996, Tang *et al.* 1996), corresponding to about one N second nearest neighbor, which gives a shift of $\Delta=-0.07$ eV. To estimate shifts in the oxide, we considered a molecule, $\text{N}(\text{SiH}_3)_2\text{Si}(\text{OSiH}_3)_3$, in which one of the first neighbor Si atoms forms a bond to three O atoms, and found $\Delta=0.12$ eV. Assuming linearity, we estimate $\Delta=0.36$ eV for the case with nine second nearest neighbor O atoms. According to this analysis, the resulting difference between interface and oxide shifts is 0.43 eV, in good agreement with the measured shift of 0.4 eV (Carr and Buhrman 1993).

2.3.4 Si $2p$ core-level shifts at the interface

To complement this study, we also calculated Si $2p$ shifts including core-hole relaxation effects for various silicon atoms in the models I and II (see Table 2.8), which present both a N– Si_3 configuration. In model I, nitrogen is located right at the interface (L_0), whereas in model II, it has been introduced in the second Si layer (L_{-1}). The silicon atoms, for which core-level shifts have been calculated, differ either by their location, or by their

Table 2.7: Relative N $1s$ shifts Δ for test-molecules generated from the following formula: $N(\text{SiH}_3)_{3-n}(\text{SiH}_2\text{R})_n$. The molecules all present the N– Si_3 configuration for the first nearest-neighbor shell, but they differing by the second nearest neighbors of n of the three Si atoms around the N atom. The radical R is selected in order for the second nearest neighbors to be Si, O, or N. The shifts are expressed in eV with respect to the case where R is SiH_3 for each n . This corresponds to $3n$ Si second nearest neighbors.

R	2 nd nn	$\Delta(n=1)$	$\Delta(n=2)$	$\Delta(n=3)$
SiH_3	Si	0.00	0.00	0.00
OSiH_3	O	0.06	0.13	0.19
$\text{N}(\text{SiH}_3)_2$	N	-0.07	-0.15	-0.19

oxidation state², or by their bonding configuration.

Many experimental results have been obtained using X-ray photoemission spectroscopy (XPS) to study Si $2p$ core-levels for the Si(001)- SiO_2 both before (Grunthaner *et al.* 1987, Himpsel *et al.* 1988, Lu *et al.* 1993) and after nitrogen incorporation (Bhat, Ahn, Kwong, Arendt and White 1994, Lu *et al.* 1995, Sutherland *et al.* 1995, Kamath *et al.* 1997). The experiments indicate the presence of a transition region from bulk silicon (Si^0 peak at $\Delta=0.0$ eV) to bulk SiO_2 (Si^{+4} peak at $\Delta=3.8$ eV) with intermediate-oxidation states of Si. These are revealed by various Si^{+x} peaks in the XPS spectrum, located at $\Delta=1.0, 1.8,$ and 2.6 eV for $x=1, 2,$ and $3,$ respectively (Lu *et al.* 1993, Lu *et al.* 1995).

However, different interpretations exist for these experimental features. The most generally accepted picture attributes the various oxidation states to Si atoms with a different number of nearest-neighbor O atoms. But, it is also suggested (Holl and Mcfeely 1993) that second-nearest neighbor effects might be important. By complementing the XPS studies with a theoretical approach based on first-principles (Pasquarello *et al.* 1995, Pasquarello *et al.* 1996), it was possible to discriminate between the different interpretations by computing the Si $2p$ for different models of interface before nitrogen incorporation. Second-nearest neighbor effects turn out to be negligibly small, so that the emerging picture is the attribution of the various peaks to different bonding configurations.

Our results are in complete agreement with the work of Pasquarello and co-workers, and confirm the ideas that had been brought in their studies.

²Unfortunately, no Si^{+3} oxidation state could be found in any of the models.

Table 2.8: Relative Si $2p$ shifts Δ for a series of Si atoms in model I and II. The selected silicon atoms differ either by their location, or by their oxidation state, or by their bonding configuration. The shifts are expressed in eV with respect to a Si atom located deep in the substrate (L_{-3}).

Location	Oxidation state	Configuration	Δ	
			Model I	Model II
L_2	Si^{+4}	4 Si-O	—	4.79
L_1	Si^{+4}	4 Si-O	4.49	—
L_0	Si^{+2}	1 Si-N, 2 Si-O, 1 Si-Si	2.03	2.15
	Si^{+1}	1 Si-O, 3 Si-Si	1.19	1.24
L_{-1}	Si^{+1}	1 Si-N, 3 Si-Si	0.92	—
	Si^0	4 Si-Si	—	0.04
L_{-2}	Si^{+1}	1 Si-N, 3 Si-Si	—	1.06
	Si^0	4 Si-Si	0.34	—
L_{-3}	Si^0	4 Si-Si	0.00	0.00

The calculated shifts due to the presence of a Si-O bond overestimate experimental values by about 20%, but the overall trend is well reproduced. We have also calculated shifts due to the presence of a Si-N bond. The same kind of discrepancy is observed with experiments, for which a shift of $\Delta=0.7$ eV is found (Larsson *et al.* 1992, Peden *et al.* 1993) compared to $\Delta=0.9$ eV in Table 2.8.

This supplementary information gives us some insight on how to interpret the Si $2p$ experimental spectrum obtained for oxynitrides. Indeed, it was found that Si^{+1} , Si^{+2} and Si^{+3} peaks have the same shifts before and after nitrogen incorporation. This observation led Lu *et al.* (1995) to attribute the interface suboxide chemical peaks only to Si-O bonds. However, we find a difference between the shifts due to Si-N and Si-O bonds of about ± 0.2 eV, which makes it really difficult to discriminate them. Hence, a contribution of Si-N bonds to the suboxide peaks cannot be excluded.

It should be mentioned that in our calculation we find that the shifts due to Si-N are lower than those due to Si-O bonds. This is in agreement with the simple picture deriving from the higher electronegativity of oxygen. The substitution of O by N increases the electronic charge density on Si, resulting in a shift of the Si $2p$ peak toward lower binding energies.

In this section, we have computed the N $1s$ core-level shifts for a series of selected molecules presenting various nitrogen bonding configurations. By comparison with experimental data for the interface, we were able to reject N-O₃ and N-SiH₂ configurations. We generate eight interface models by considering different bonding configurations and locations with respect to the interface plane. By analyzing their structural properties, we have shown that our relaxed models give a good representation of the local structure for the various nitrogen-bonding configurations. By computing the N $1s$ core-level shifts for these models and comparing with experimental data, we were able to explain the unusually large shift with oxide thickness of the principal peak in photoemission spectra in terms of a single first-neighbor configuration in which the N atom is always bonded to three Si atoms, both in the interfacial region and further in the oxide. Core-hole relaxation and second nearest neighbor effects were shown to concur in yielding larger binding energies in the oxide than at the interface. The calculations do not support the occurrence of N-O bonds at nitrated Si(001)-SiO₂ interfaces. The Si $2p$ core-level shifts have also been calculated for selected interface models showing good agreement with experiments and previous theoretical studies. The results demonstrate that a contribution of Si-N bonds to the

suboxide peaks cannot be excluded.

2.4 Discussion and Conclusions

In this chapter, we have studied the incorporation of nitrogen at the Si(001)-SiO₂ interface using a first-principle approach for the calculation of core-level shifts. In section 2.2, we first addressed the problem of obtaining a reliable theoretical framework for the calculation of core-level shifts. Various approximations have been introduced that allow to compute core-level shifts within Density Functional Theory. The initial-state shifts could be obtained by the expectation value of the local self-consistent potential on the atomic core-state. The core-hole relaxation effects could be taken into account despite the use of pseudopotentials, by performing two separate calculations. We have applied these two methods to the calculation of N 1s core-level shifts of small molecules. By comparison with experimental data, we showed it is necessary to include core-hole relaxation effects to obtain reasonable values. In section 2.3, we turned to the incorporation of nitrogen at the Si(001)-SiO₂ interface. The approach that was used in this study can be summarized as follows. In a first step, we determined the possible nitrogen bonding configurations. This was done by computing the N 1s core-level shifts for a series of carefully chosen molecules and by examining the results against available experimental informations about the interface. With all the pre-selected bonding configurations, we produced several model interfaces by varying the distance of the incorporated N atoms from the interface plane. The atomic positions were relaxed using a first-principle approach and the structural properties of the generated models were analyzed. The N 1s and Si 2p core-level shifts calculated for the different models and compared with experimental data. The conclusions that were drawn are reformulated hereafter. First, the structural analysis showed that our models give a good representation of the local structure for the various nitrogen-bonding configurations. Second, adopting the single first-neighbor configuration N-Si₃, we could explain the shift with oxide thickness of the principal XPS line and the appearance of two components. Our interpretation relies on core-hole relaxation and second nearest neighbor effects, which combine to give larger binding energies in the oxide than at the interface, in accord with experimental observations (Carr and Buhrman 1993, Hegde *et al.* 1995, Lu *et al.* 1995, Kaluri and Hess 1996). Core-hole relaxation affects N 1s shifts differently according to the distance of the N atoms to the screening Si substrate. A second nearest neighbor environment rich in Si and N atoms as at

the interface induces opposite shifts compared to an O rich environment as found in the oxide. Third, our investigation of other configurations, in which the N atom is bonded to O atoms, tends to rule out their existence at the interface. Indeed, we found N $1s$ shifts for N-Si₂O configurations at $\Delta=1.5$ eV, in a region of the spectrum where generally no peaks are observed in experiments. On the contrary, for N-Si₂H configurations, the calculated shift is right in the range of experimental data ($\Delta=0.4$ eV). However, the high concentration of hydrogen that would be required for such configurations lead us to suggest if these appeared in nitrated interface it would only be in a very small amount. Finally, the calculated Si $2p$ core-level shifts are in good agreement with experiments and other theoretical studies. Contrarily to what is generally believed, we proposed that both Si-O and Si-N bonds contribute to the suboxide peaks.

Chapter 3

The SiO₂ α -Quartz (0001) Surface

3.1 Introduction

SiO₂ plays an important role in many advanced technologies. For instance, SiO₂/Si interfaces are currently used in MOS (metal oxide-semiconductor) electronic devices (Helms and Deal 1988). SiO₂ is also used as substrate for metallic magnetic multilayers and for oxide multilayer optical systems. In the field of adhesion, the SiO₂/polymer interface is an important concern (Zhong *et al.* 1993). Whatever the interfaces built up, their properties are expected to strongly depend on the initial SiO₂ surface structure.

Amongst the different SiO₂ surfaces, the (0001) α -quartz surface can be considered as a model surface for the different crystallographic forms of SiO₂. Moreover, its study can also be profitable for the understanding of amorphous silica surface. However, whereas the bulk structure of quartz has been widely studied, the experimental data on the structure of this surface are rather scarce. To our knowledge, there are up to now only few studies devoted to experimental determination of the crystallographic structure of the (0001) α -quartz surface. Several years ago, it has been shown that a slight chemical etching in HF leads to the appearance of a 1×1 pattern (Janossy and Menyhard 1974). More recently, a reconstruction $\sqrt{84} \times \sqrt{84}$ with a rotation angle of 11° was observed as soon as the air heating temperature was greater than 600°C (Bart and Gautier 1994).

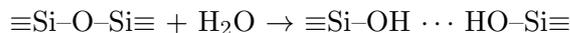
The main difficulty lies in the fact that the common surface analysis techniques often employ the input of charged particles (ions, electrons) or the analysis of outgoing charged particles (ions, electrons). Indeed, these

changes in the local charge concentration make the study of insulator surfaces more difficult than for conductor surfaces. However, promising steps have been achieved in this field in the last few years. For instance, recently, the quartz surface has been studied using atomic force microscopy (AFM) providing informations on its microscopic structure (Noge *et al.* 1997). However, the low contrast of the AFM pictures obtained at the atomic scale hamper one to draw clear conclusions. Nevertheless, it can be expected that in the forthcoming years most of the existing difficulties in the scanning probe microscopies will be overcome.

In most experimental techniques, data are collected over fairly long times (in comparison to atomic or molecular motion). Thus, a direct measurement or observation of a particular event or structure is not allowed, and these must be inferred from the data. By contrast, molecular dynamics simulations permit detailed analysis of the atomic motion and the complex microstructure that gives rise to the average properties. These techniques have been widely used to study vitreous silica surfaces. However, to date, the simulations have been based on semi-empirical potentials only to describe atomic interactions (Garofalini 1983, Feuston and Garofalini 1988, Feuston and Garofalini 1989, Garofalini 1990).

In the present work, we present an *ab initio* study of the (0001) α -quartz surface. We consider different initial configurations, amongst which the cleaved surface (with non-bridging oxygens at the top) and a 2×1 reconstruction (with 2-membered rings). After the atomic relaxation, the stability of each configuration is studied by performing constant-temperature molecular dynamics simulations, leading to different reconstructions. The most stable of these presents an interesting densification of the two uppermost layers of SiO₂ tetrahedral units, with the formation of 3-membered and 6-membered rings that do not exist in bulk α -quartz.

Another important topic in regard to the SiO₂ surface lies in its interactions with water. These play an important role in microelectronics (wafer-bonding techniques), catalytic support, optical wave guides, and sol-gel processing. It is generally believed that the hydroxylation of siloxane ($\equiv\text{Si}-\text{O}-\text{Si}\equiv$) and of non-bridging oxygens ($\equiv\text{Si}-\text{O}$) causes the formation of silanol species ($\equiv\text{Si}-\text{OH}$) at the surface (Iler 1979). *Ab initio* calculations have been performed to study the interaction between a single water molecule and a silicic acid monomer (H₄SiO₄), a “simple” system involving Si, O and H interactions (Heggie and Jones 1987). It was shown that a 4% expansion of the Si-O bonds in siloxane was sufficient on energy grounds to allow the hydrolysis reaction:



to occur. This led to useful predictions for reaction sites and paths of water-silica interactions.

Till now, such first-principles approaches which are computationally demanding have not been applied to study the complete silica surface in the presence of water. However, molecular dynamics simulations have been performed using three-body interaction potentials (Feuston and Garofalini 1990). In this study, we analyzed the hydroxylation of the dense α -quartz surface obtained in dry conditions, as well as of partially hydroxylated surfaces. We also investigated the dehydration of the fully-hydrated and the semi-hydrated quartz surface.

The present chapter is organized as follows. In section 3.2, we first present the bulk structure of α -quartz and introduce different ways of visualizing it (network of tetrahedra, repetition of O-Si-O layers along the (0001) direction). These facilitate a better understanding of the surface structure. Then, we investigate different models for the (0001) α -quartz structure. We consider two initial configurations, the cleaved surface and a 2×1 reconstruction with edge-sharing tetrahedra and obtain two stable reconstructions by performing constant-temperature molecular dynamics simulations. The first, named the valence alternation pair surface, presents an intimate pair of over- and under-coordinated oxygen atoms near the surface with 3-membered rings. The second, called the dense surface, presents a densification of the two uppermost layers of SiO_2 tetrahedral units, with 3-membered and 6-membered rings. We analyze in details the structural properties of our models and describe them in terms of tetrahedral units, showing that within 5 Å from the surface the bulk structure is recovered. Finally, we compute their respective total energy, and find that the dense surface is most stable of all investigated structure. In section 3.3, we present our results in regard with the interaction of the quartz surface with water. We first consider the dehydration of the wet quartz surface. Unfortunately, it is found that the time scale accessible to our simulation is too small in regard to the large energy barriers (about 3 eV/surface unit cell at least) that are involved. We also describe the dynamics resulting from throwing water molecules on the dense surface, and on partially hydroxylated surfaces in terms of the mechanism of hydrolysis of siloxane bonds or non-bridging oxygens. Our study confirms the hydrophobic behavior of siloxane bonds. Indeed, we find that the water molecules rebound on the dense surface whatever their kinetic energies. On the contrary, our simulations demonstrate the high reactivity of under-coordinated silicons (especially those linked to the

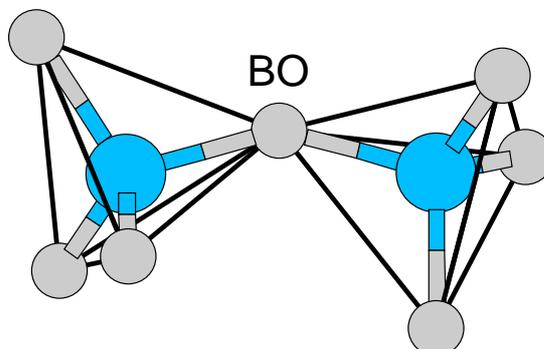


Figure 3.1: Representation of two corner-sharing SiO_4 tetrahedral units (thick solid lines). The Si atoms (represented in dark grey) are located at the center of the tetrahedra, while O atoms (in pale grey) constitute its corners. The central O atom connects the two tetrahedra, it is referred to as a bridging oxygen (BO). The other oxygens are attached to other silicons which are not shown on the figure.

non-bridging oxygens) and the role of adjacent hydroxyl groups to facilitate the hydrolysis reaction by the formation of hydrogen-bond chains.

3.2 The dry surface

3.2.1 The bulk structure of α -quartz

Before addressing the study of the SiO_2 α -quartz surface, we introduce some general considerations about its bulk structure.

The α -quartz or low quartz is the stable phase of SiO_2 at ambient conditions. This crystal possesses the $P3_121$ space-group symmetry. It consists of 4-coordinated silicon and 2-coordinated oxygens, which may be seen as a network of SiO_4 tetrahedra (with 1 Si atom at the center and 4 O atoms at the corners as shown in Fig. 3.1) with all corners shared. The structure is formed by double helical-chains of SiO_4 tetrahedra that run parallel to the (0001) direction and which are linked laterally one to the other (Fig. 3.2).

The oxygens joining two tetrahedra (as represented in Fig. 3.1) are called bridging oxygens (BO's). These create Si-O-Si bond angles which are all equal to 143.73° . On the contrary, the O-Si-O bond angles in the tetrahedra present four distinct values (108.81° , 108.93° , 109.24° , and 110.52°) that are

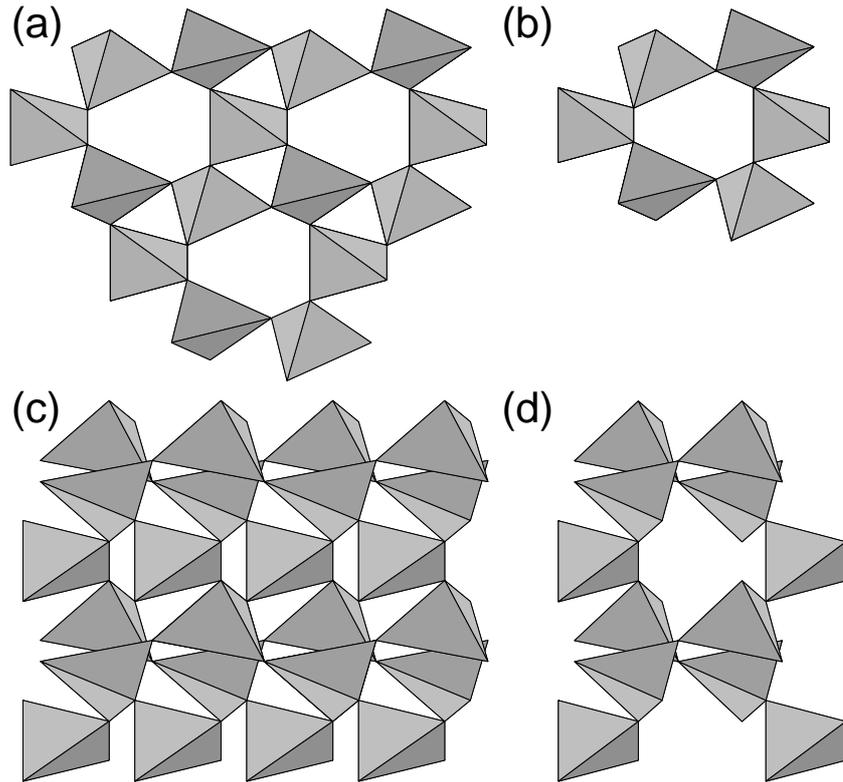


Figure 3.2: Representation of the α -quartz structure in terms of tetrahedral units. (a) Top view (perpendicularly to the (0001) direction) and (b) side view (along the (0001) direction) of the crystal. (c) Top view and (d) side view of the tetrahedral double helix.

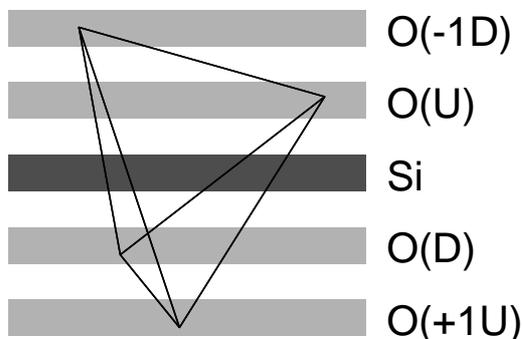


Figure 3.3: Representation of a SiO_4 tetrahedral unit in terms of its location with respect to a O-Si-O layer along the (0001) direction in α quartz. The O-Si-O layer consists in a upper plane of O atoms (labeled O(U) where U stands for up), a plane of Si atoms, and a lower plane of O atoms (O(D) where D stands for down). We also represented the lower plane of oxygens in the layer just above it (O(-1D)) and the upper plane of oxygens in the layer just under it (O(+1U)). In the tetrahedra, the Si atom is bonded to 4 oxygens located in the oxygen planes labeled O(U), O(D), O(-1D) and O(+1U), respectively.

very close to the ideal tetrahedral angle (Levien *et al.* 1980).

Silica glass and various low pressure crystalline polymorphs of SiO_2 also consist of network of SiO_4 tetrahedra. The difference resides in the distribution of the Si-O-Si bond angles which depends on the symmetry of the crystal. For instance, in vitreous silica, it ranges from 120° to 180° . This broadening of the distribution with respect to α -quartz is due to the amorphous character.

The α -quartz structure can also be seen as being formed by repeated O-Si-O layers along the (0001) direction as shown in Fig. 3.3. Each layer consists in a plane of O atoms (O(U)), then a plane of Si atoms, and then again a plane of O atoms (O(D)). Each Si atom is bonded to one atom of the upper plane and to one atom of the lower plane, it is also linked to one oxygen in the O-Si-O layer just under it (O(+1U)) and to one oxygen in the layer just above it (O(-1D)). This representation is particularly interesting for the study of the (0001) surface of α -quartz.

3.2.2 The fracture of α -quartz

In this study, we are concerned with the detailed structure of SiO_2 surface as created by instantaneous fracture of bulk material. In principle, the differences between this type of surface and the bulk material are dependent upon the forming conditions. The final structures are due to the lack of three-dimensional network at the surface, the increased mobility of surface atoms (due to broken bonds), as well as the different thermal and chemical (atmospheric) exposures of the uppermost atoms. For simplicity, we first deal with surfaces obtained in perfect vacuum (in absence of water). This should isolate the changes due primarily to the lack of network.

The fracture of bulk quartz perpendicularly to the (0001) direction, in which we are more specifically interested, could produce two fundamental types of surface: one terminating with Si atoms, and a similar surface with one oxygen attached to each silicon atom. However, since natural silicon dioxide surfaces are generally thought to terminate by hydroxyl (OH) groups (Hair 1967, Iler 1979), the oxygen-terminated surface is most likely more representative of real materials. Hence, in the present work, we will concentrate on this type of surface.

The fracture of the crystal and the reconstruction of the resulting surface may induce the presence of defects and distortions with respect to the perfect tetrahedral picture of SiO_2 given above. One particularly interesting defect in this framework is an oxygen which is linked only to one silicon. This is referred to as a non-bridging oxygen (NBO). It is known that both BO's and NBO's exist at the surface although the surface NBO's are usually hydrated due to the presence of contaminant in real systems (Iler 1979). Another kind of defects which might be present at SiO_2 surface are the n -membered rings where n is the number of Si atoms. Molecular dynamic simulations with three-body interaction potentials showed that n -membered (with $n=2, 3, 4$) are formed at vitreous silica surfaces (Feuston and Garofalini 1989).

In this study, we use Car-Parrinello molecular dynamics to study various models of the (0001) α -quartz surface. On this purpose, we build a tetragonal unit cell by repeating O-Si-O layers along the (0001) direction. Doing so, we meet the requirement of the charge neutrality in the unit cell and avoid large dipole moments. The O-Si-O layers had previously been relaxed in a cell simulating bulk quartz. The Si-O bond length was found to be $1.605 \pm 0.005 \text{ \AA}$ which is less than 1 % difference with experiments (Levien *et al.* 1980). The Si-O-Si bond angle distribution was peaked around 139° , whereas the four values of O-Si-O angles are 108.1° , 109.3° , 109.4° , and

111.0° respectively¹. All these values present less than 5% discrepancy with experimental data (Levien *et al.* 1980).

Our systems contain a surface unit of sides $a=9.62$ Å and $b=8.34$ Å, based on the theoretical equilibrium lattice constant of α -quartz (Keskar and Chelikowsky 1992). The dimension of the cell in the direction orthogonal to the surface is $c=15.88$ Å, containing 6 monolayers of SiO₂ (9.5 Å), the bottom extremities being saturated with hydrogen atoms. Hence, our systems contain 24 Si atoms, 44 O atoms and 8 H atoms.

The atomic positions are fully relaxed by using the Car-Parrinello method (Car and Parrinello 1985, Tassone *et al.* 1994), which provides the electronic structure as well as the forces that act on the ions. Only valence electrons are explicitly considered by means of norm conserving pseudopotentials (Troullier and Martins 1991) to account for the core-valence interactions. The electronic wave functions are expanded into plane waves with a kinetic energy cutoff of 50 rydbergs. Exchange and correlation are included using Perdew and Zunger's interpolation formulas (Perdew and Zunger 1981). The Brillouin zone (BZ) is sampled using only the Γ -point in all our calculations. In the minimization process, the atoms of the lowest SiO₂ monolayer are kept fixed. After the atomic relaxation, the stability of each configuration is studied by performing constant-temperature molecular dynamics simulations (Nosé 1991), leading to different reconstructions.

In Table 3.1, we have reported the maximal displacements of the atoms after relaxation for the various models with respect to their positions in bulk quartz. From these results, it can be stated that the size of the SiO₂ slab used in the simulations was large enough to simulate a semi-infinite solid, since only limited strain occurred in the O-Si-O layers just above the fixed one.

We now analyze the various models that were generated. We compare their structural properties as well as their relative stability.

3.2.3 The cleaved surface

The most trivial model for the α -quartz (0001) surface is obtained simply by stopping the repetition of O-Si-O layers along the (0001) direction. We will refer to it as the *cleaved surface*. This 1×1 reconstruction of the surface is characterized by one non-bridging oxygen and one 3-coordinated silicon per

¹When we used the PW'91 approximation for the exchange-correlation energy functional, we obtained $d(\text{Si-O})=1.625\pm 0.005$ Å, $\angle(\text{Si-O-Si})=135.5\pm 0.1^\circ$, and for the four O-Si-O angle values 107.9° , 108.4° , 109.0° , and 111.8° respectively

Table 3.1: Maximal atomic displacements in the first and second O-Si-O layers above the fixed one after complete relaxation in the simulation cell. The results are presented for our 4 different models of the surface (cleaved, dimer, valence alternation pair (VAP), and dense). They are expressed in Å with respect to the positions of the atoms in bulk quartz.

Surface	O-Si-O layer	
	1 st	2 nd
Cleaved	0.02	0.03
Dimer	0.05	0.06
VAP	0.15	0.19
Dense	0.15	0.19

surface unit cell. A side view of the cleaved surface in terms of tetrahedral units is given in Fig. 3.4.

After the atomic relaxation, the structural properties of the cleaved surface have been analyzed (Fig. 3.5). The Si-O bond length for the NBO's at the surface is about 1.49 Å, that is 0.1 Å shorter than in bulk quartz. This result is in good agreement with other theoretical results obtained with empirical interaction potentials (Levine and Garofalini 1987). One of the back bond Si-O is also slightly shortened to 1.57 Å. In the second O-Si-O layer the bulk value is already recovered (Fig. 3.5(a)), indicating that the strain due to the surface is very localized. This is confirmed by the bond angle distribution for O-Si-O and Si-O-Si, as represented in Fig. 3.5(b) and (c). The three O-Si-O angles involving the Si atoms located in the first O-Si-O layer are respectively 126.8°, 126.7°, and 106.7° (where the two biggest values are obtained for O-Si-O angles with a NBO), indicating a quasi planar configuration the SiO₃ group.

3.2.4 The dimer surface

The second model for the α -quartz was based on the idea that a more stable structure would be obtained if the non-bridging oxygens and the 3-coordinated silicons could be replaced by 2-coordinated oxygens and 4-coordinated silicons respectively in order to recover tetrahedral units even at the surface. A simple way to do that was to consider a 2×1 reconstruction in which two surface silicons (3-coordinated in cleaved surface) would share their non-bringing oxygens to form edge-sharing tetrahedra (2-

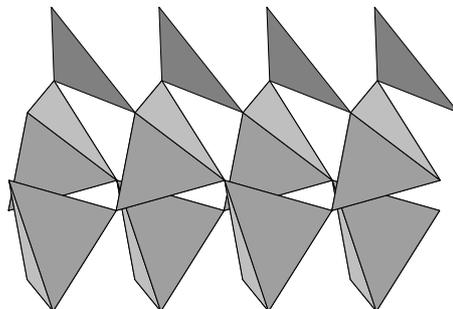


Figure 3.4: Side view of the cleaved surface in terms of tetrahedral units. At the surface, the oxygen atoms are non-bridging and the silicon atoms are 3-coordinated. Triangles appear at the top of the figure in place of tetrahedra. This representation is justified by the fact that the sum of the three O-Si-O angles involving the Si atoms located in the first O-Si-O layer is very close to 360° .

membered rings). This reconstruction is illustrated in Fig. 3.6.

However, the atomic relaxation of this surface lead to a completely different structure. Indeed, the two oxygens that should have been shared by two silicons seemed to prefer to form a dimer on one of the latter, leaving the other silicon only two-fold coordinated. Moreover, this structure, to which we refer as the *dimer surface*, was energetically unfavored by 6.8 eV/surface unit cell with respect to the cleaved surface.

The structural properties of the dimer surface have been analyzed and reproduced in Fig. 3.7. The Si-O bond length is modified at the surface, but soon recovers the bulk value in deeper layers (Fig. 3.7(a)). The modifications are due to the strange reconstruction. The two back bonds of the 2-coordinated Si are slightly longer $1.64 \pm 0.01 \text{ \AA}$ due to the increased electronegativity of the silicon. On the contrary, those of the Si with a dimer are slightly shorter $1.56 \pm 0.01 \text{ \AA}$, while for the same Si atoms the two bonds with the O of the dimer are also a bit longer $1.62 \pm 0.01 \text{ \AA}$. The distances of the 2-coordinated Si with respect to the oxygens of the dimer are 2.35 \AA and 3.26 \AA respectively. The O-O bond in the dimer is 1.59 \AA . The O-Si-O angles are also affected mostly at the surface (Fig. 3.7(b)). The one involving the two oxygen atoms of the dimer is 58.6° , enforcing the fact that a dimer is formed. The ones involving one of these two oxygens are 121.2° and 127.6° respectively. The remaining O-Si-O angle involving the Si with the

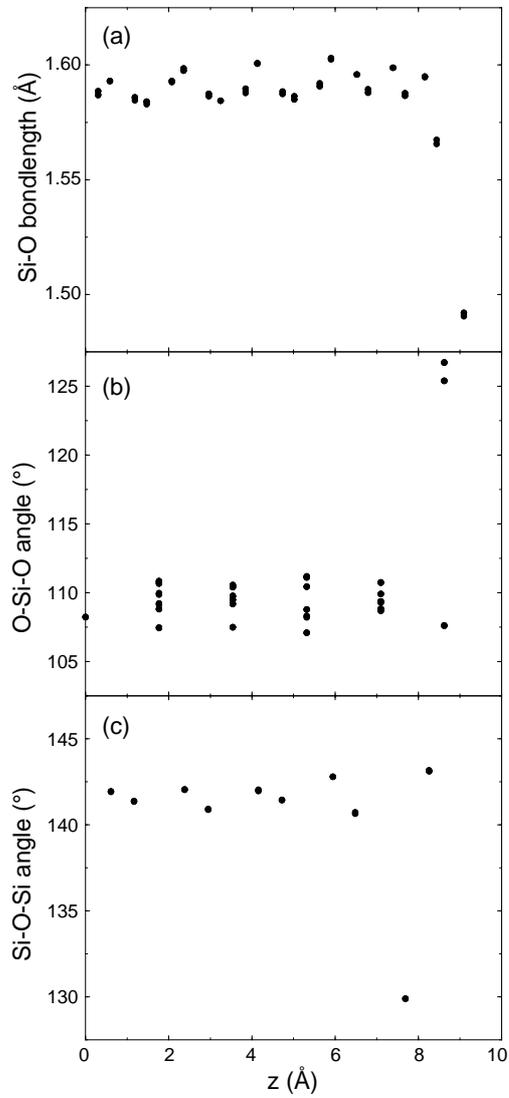


Figure 3.5: Structural properties of the cleaved surface: (a) Si-O bond length (in \AA), (b) Si-O-Si bond angle (in $^\circ$), and (c) O-Si-O bond angle (in $^\circ$) distributions. The z axis is oriented along the (0001) direction (i.e. perpendicularly to the surface) with the origin located at the height of the lower plane of oxygens (which is fixed during atomic relaxation).

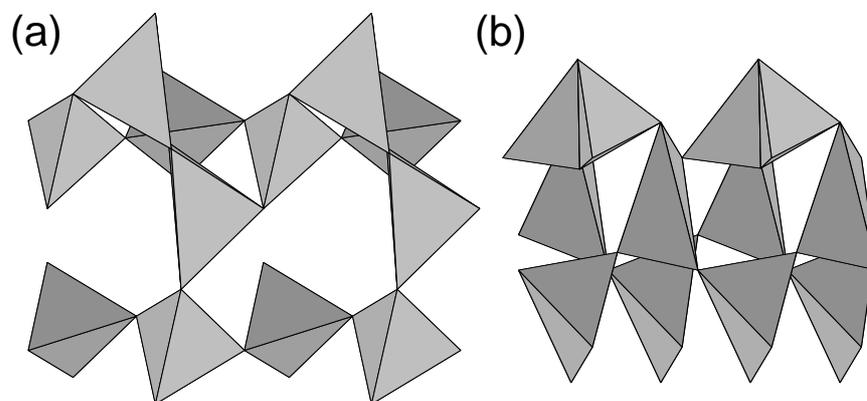


Figure 3.6: Representation of an hypothetical surface with 2-membered rings in terms of tetrahedral units: (a) top view (perpendicularly to the (0001) direction), and (b) side view. The edge-sharing tetrahedra (2-membered rings) are located at the top.

dimer (thus, the one with the two back-bonding oxygens) is 110.7° which is very close to the tetrahedral value. By contrast, the only O-Si-O angle involving the two-coordinated Si is 98.3° far from this normal value. The distortions of the O-Si-O angles are still perceptible in the second tetrahedra layer but are attenuated in the successive ones. The deviations of the Si-O-Si angles from their bulk value are also concentrated in the very first layers (Fig. 3.7(c)) and are relatively small. Thus the tetrahedra to which they belong is heavily distorted.

Finally, we analyzed the electronic density on the dimer, a contourplot of which is represented in Fig. 3.8. The oxygen dimer clearly appears on the picture as well an increased electronic density on the 2-coordinated silicon with respect to the other 4-coordinated silicons. The contour plot also reveals the reason why the edge-sharing tetrahedra structure originally designed was instable. Indeed, there is an oxygen located just above the dimer (at the bottom of the figure) which repels it towards the outside.

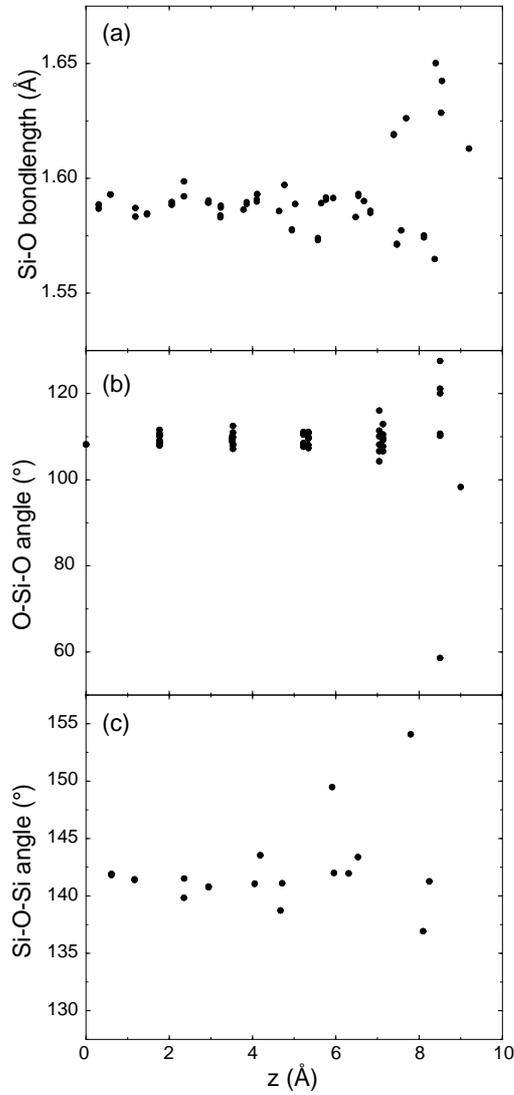


Figure 3.7: Structural properties of the dimer surface: (a) Si-O bond length (in Å), (b) Si-O-Si bond angle (in °), and (c) O-Si-O bond angle (in °) distributions. The z axis is oriented along the (0001) direction (i.e. perpendicularly to the surface) with the origin located at the height of the lower plane of oxygens (which is fixed during atomic relaxation).

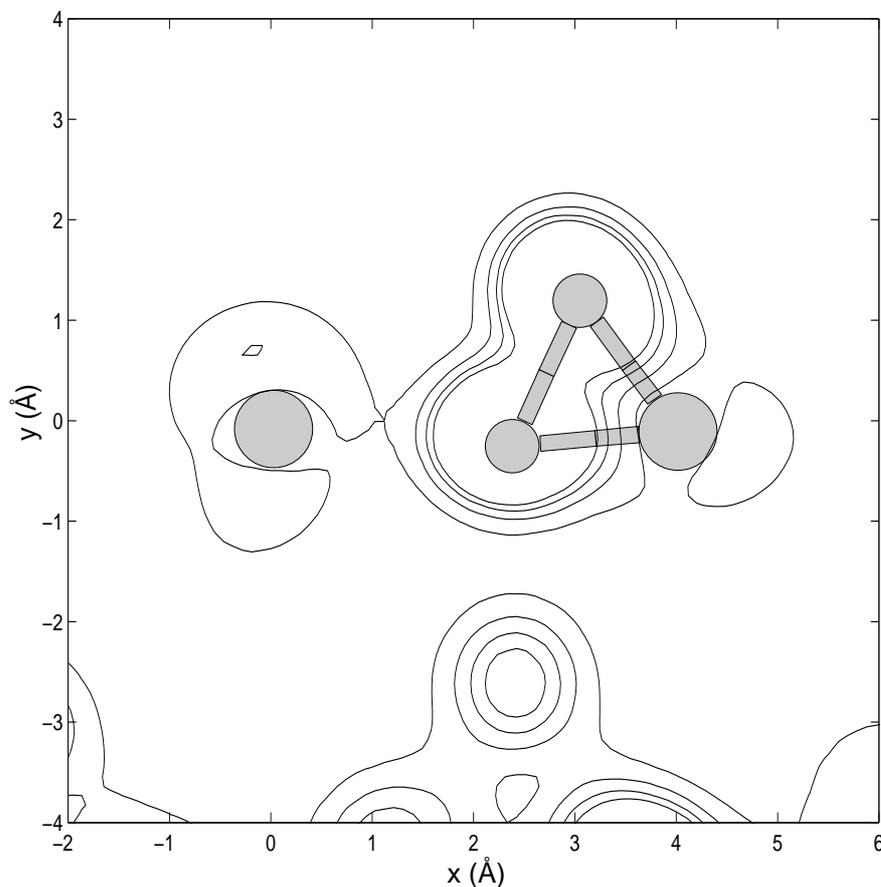


Figure 3.8: Contour plot of the electronic density on the oxygen-dimer at the dimer surface. The contour is drawn in a plane passing through the Si atom with the dimer, the 2-coordinated Si, and the lower oxygen of the dimer. Note that the upper oxygen is almost co-planar with these atoms. The x axis is defined by the 2 silicons. The origin of the x and y axis (defined by orthogonality with x) is the 2-coordinated Si. As a guide for the eyes to locate the atoms, we superimposed (in pale grey) to the contour plot a ball and stick representation of the structure: the big circles are Si atoms while the small ones are O atoms. The oxygen dimer clearly appears at the center of the picture, while the electronic density is increased on the 2-coordinated silicon with respect to 4-coordinated silicons. At the bottom of the figure, there is another peak of density. It corresponds to an oxygen located just above the dimer which repels it towards the outside. The contour interval is $0.5 \text{ e}^-/\text{\AA}^3$.

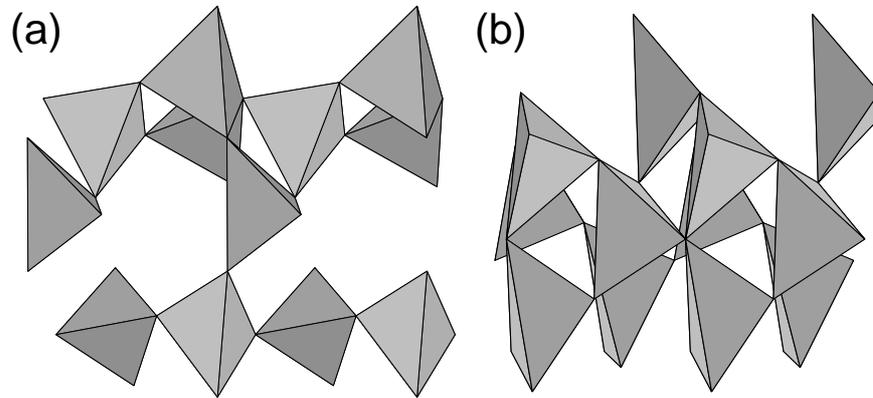


Figure 3.9: Representation of the valence-alternation-pair surface in terms of tetrahedral units: (a) top view (perpendicularly to the (0001) direction), and (b) side view. In the top view, the 3-membered rings appear neatly. In the side view, the non-bridging oxygens are positioned at the top while the 3-coordinated O are situated in the middle.

3.2.5 The VAP surface

By heating the dimer surface up to 300K, we obtained a 2×1 reconstruction more stable than the cleaved surface, which constitutes our third model. In this structure all the Si atoms are 4-coordinated thus it can be completely described by tetrahedral units as in Fig. 3.9. However, the network of tetrahedra is not perfect since there is a 1-coordinated oxygen and a 3-coordinated oxygen per surface unit cell. These two defects are both located on the same tetrahedra with the 1-coordinated O pointing towards the outside. The increase of stability can be attributed to the fact that the number of Si-O bonds is unchanged with respect to an idealized fully coordinated network. The representation of the surface structure in terms of tetrahedra also reveals the existence of 3-membered rings in the two first layers.

This intimate pair of over- and under-coordinated oxygen atoms is referred to as a valence alternation pair (VAP). The VAP concept was first introduced by Kastner *et al.* (1976) in the framework of the study of chalcogenide glasses (e.g. As_2Se_3). A valence alternation pair was formed when two chalcogen atoms, both two-fold coordinated in the ground state, rear-

ranged instead into one positively charged 3-coordinated atom (noted C_3^+) and one negatively charged 1-coordinated atom (C_1^-). It was proposed that the density of VAP's in most glasses is relatively large since the creation of such defect requires a relatively small energy. The existence of such defects in amorphous SiO_2 was first proposed by Lucovsky (1979) based on the interpretation of infra-red and Raman spectra. Lucovsky (1980) also discussed some qualitative differences between VAP's in chalcogenide and silica glass. These are however irrelevant for the present study. He estimated the density of these defect pairs to be of the order of 10^{19} cm^{-3} which is quite important.

The structural properties of the VAP surface have been analyzed and reproduced in Fig. 3.10. The latter illustrates that the effect of the surface is felt deeper than in the previous two models. The Si-O bondlength is affected by the presence of the VAP's and of the 3-membered rings (Fig. 3.10(a)). The bond involving the NBO is equivalent to those found in the cleaved surface, it has a length of 1.50 Å which is 0.1 Å shorter than in bulk quartz. On the contrary, those involving the 3-coordinated O are longer: 1.70 Å, 1.73 Å, and 1.83 Å respectively. In the 3-membered rings, there are also slight variations with respect to bulk quartz. The O-Si-O angles are the most affected property with respect to the other models (Fig. 3.10(b)). This means that the tetrahedral units are very distorted even in deeper layers. By contrast, the Si-O-Si angle deviations are concentrated mostly in the first layers (Fig. 3.10(c)). The biggest changes arise from the 3-coordinated O, where the three Si-O-Si angles are 111.2° , 118.3° , and 120° respectively, indicating that the oxygens are not co-planar with the Si atoms to which they are bonded.

The stability of this reconstruction was tested by heating the surface up to 3500 K for a few hundred fs. No further reconstruction was observed indicating a relative stability of the structure.

3.2.6 The dense surface

The most stable reconstruction is obtained by heating the cleaved surface up to 300K. During the molecular dynamics simulation, the uppermost O-Si-O layer kind of merges with the second layer in a unique layer with the formation of 3-membered and 6-membered rings (Fig. 3.11) that do not exist in bulk α -quartz. This reconstruction presents only 4-coordinated Si and 2-coordinated O with no dangling bonds.

The outermost O-Si-O layer is denser than those located in the bulk, due to the merging of two layers of the original structure (Fig. 3.12). The

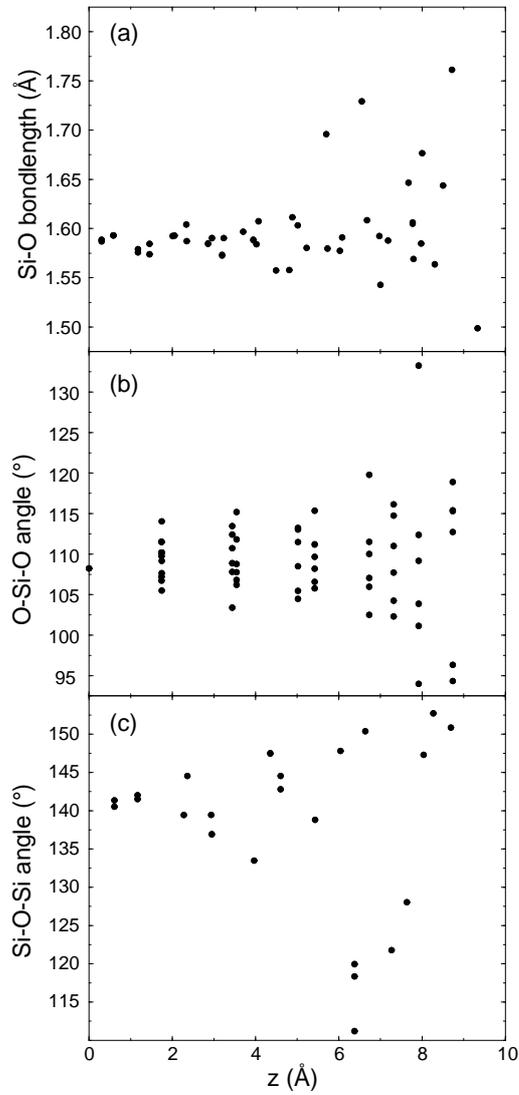


Figure 3.10: Structural properties of the dimer surface: (a) Si-O bondlength (in Å), (b) Si-O-Si bond angle (in °), and (c) O-Si-O bond angle (in °) distributions. The z axis is oriented along the (0001) direction (i.e. perpendicularly to the surface) with the origin located at the height of the lower plane of oxygens (which is fixed during atomic relaxation).

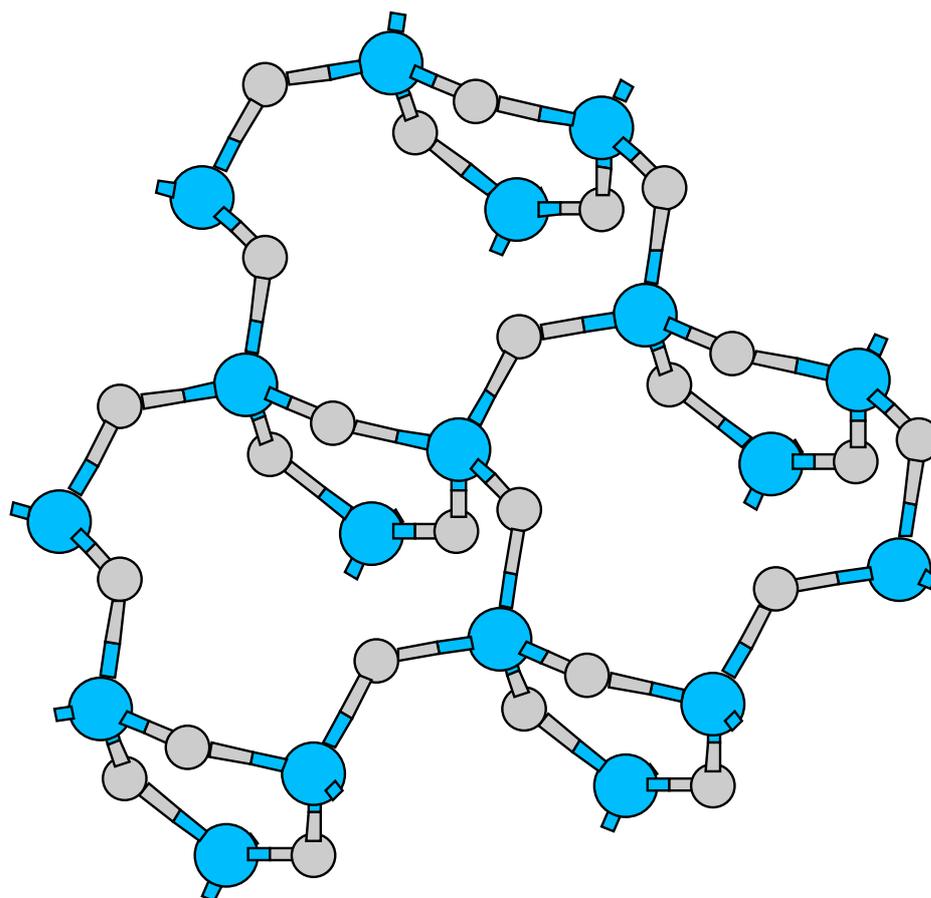


Figure 3.11: Top view of the dense surface. Si and O atoms are represented in dark and pale grey, respectively.

uppermost atoms are still oxygens which are located in a plane, however there are three times more atoms than in the original ones. Indeed, the two planes of oxygens from the outermost layer in the cleaved surface (O(1U) and (O1D) in Fig. 3.12) have merged with the upper plane of oxygens of the second O-Si-O layer (O(2U)). The plane of Si atoms located just under the surface is also denser, but in this case there is only twice the number of atoms (Si(1+2)). The lower plane of oxygens of the second O-Si-O layer has almost merged with the upper plane of oxygens of the third O-Si-O layer, but this merging completely disappears in the deeper layer.

We have analyzed the structural properties of the dense surface, the distribution of Si-O bond lengths, O-Si-O and Si-O-Si bond angles are presented in Fig. 3.13. As it can be seen from part (a) of this figure, the Si-O bond length is very homogeneous with no real difference between the surface and the bulk ($d(\text{Si-O})=1.59\pm 0.01 \text{ \AA}$). The bond angle distribution for O-Si-O and especially Si-O-Si, which have been represented in Fig. 3.5(b) and (c), are much more affected by the reconstruction. Part of this can be explained by the presence of 6-membered and 3-membered rings whose shape is reproduced in Figs. 3.14 and 3.15 respectively.

In Tables 3.2 and 3.3, we have reported all the bond lengths and bond angles for these rings. In the 6-membered rings, it can be seen that the O-Si-O angles are very close to the tetrahedral angle. This shows that quasi perfect tetrahedra are formed at the surface. It is then interesting to analyze the reconstruction in terms of SiO_2 tetrahedral units (Fig. 3.16). Indeed, the surface is formed by sharing-corner tetrahedra with 3 corners in the surface plane and one directed towards the bulk. So that all the tetrahedra have the same orientation, contrarily to what happens in the bulk where all sharing-corner tetrahedra have opposite orientation. The Si-O-Si angles are close to 120° reflecting the trigonal symmetry of the 6-membered ring. In the 3-membered rings, two O-Si-O angles that belong to the surface tetrahedra are very close to 109.5° . On the contrary, the O-Si-O angle with the Si located in the second layer is slightly different from the tetrahedral angle. This is due to the partial merging of the lower and upper planes of oxygens of the second and third O-Si-O layers discussed above. In other words, the origin of this distortion is due to the fact that the O atoms labeled 1 and 3 in Fig. 3.15 are almost at the same height with respect to the fixed layer at the bottom. It is also interesting to analyze the Si-Si-Si and O-O-O angles. In the 6-membered rings, the Si-Si-Si angles are all very close to 120° indicating an hexagonal shape. By contrast the O-O-O present two different values: 160° and 78° . In fact, the O atoms form a slightly-distorted equilateral triangle with 3 misaligned oxygens on each side. In the 3-membered rings,

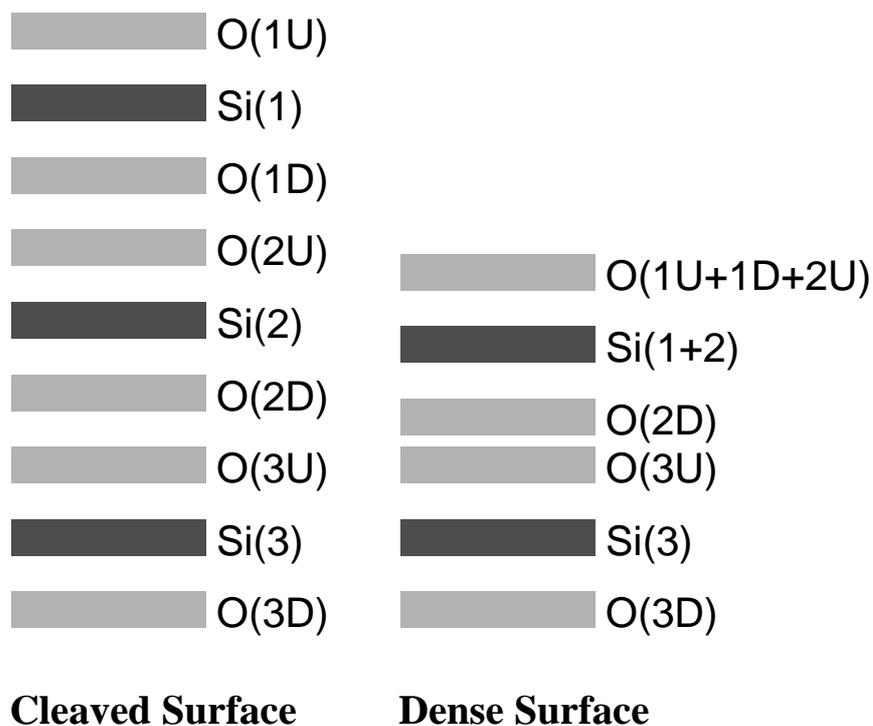


Figure 3.12: Schematic view of the reconstruction of the cleaved surface (left side of the graph) into the dense surface (right side of the graph). The planes of Si and O atoms are represented in dark and pale grey, respectively. For the cleaved surface, 3 O-Si-O layers have been represented and labeled from 1 to 3 starting from the surface. In each layer, the upper and lower planes of oxygens have been distinguished: U(up) and D(down) respectively. For the dense surface, the same number of atoms has been represented. However, due to the densification, there are only 2 O-Si-O layers left. In each planes, we have indicated the origin of the atoms by their label. For instance, the upper O layer corresponds to the merging of the O(1U), O(1D), and O(2D) planes.

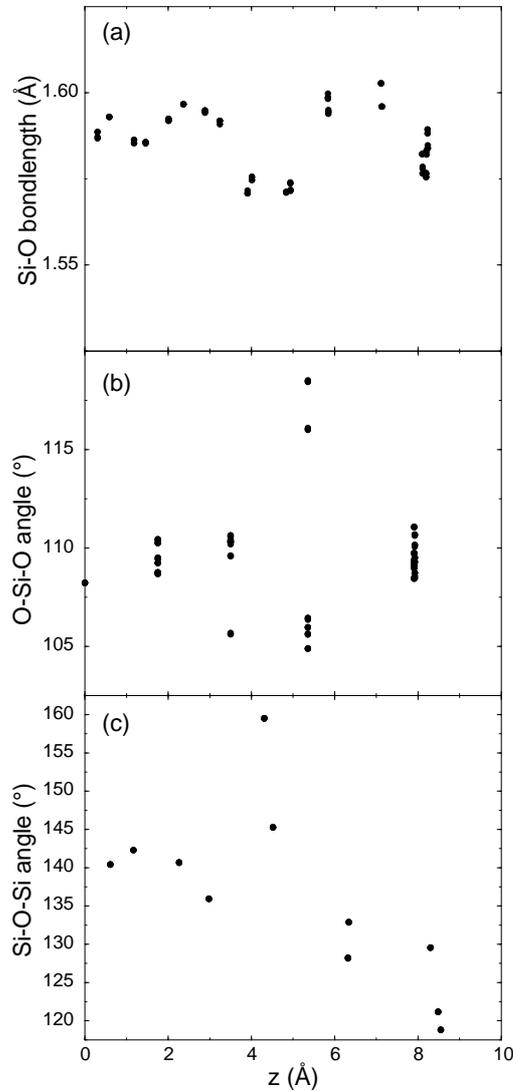


Figure 3.13: Structural properties of the dense surface: (a) Si-O bond length (in Å), (b) Si-O-Si bond angle (in $^{\circ}$), and (c) O-Si-O bond angle (in $^{\circ}$) distributions. The z axis is oriented along the (0001) direction (i.e. perpendicularly to the surface) with the origin located at the height of the lower plane of oxygens (which is fixed during atomic relaxation). Note that in part (a) the scale is not the same as in equivalent graphs presented for the other models

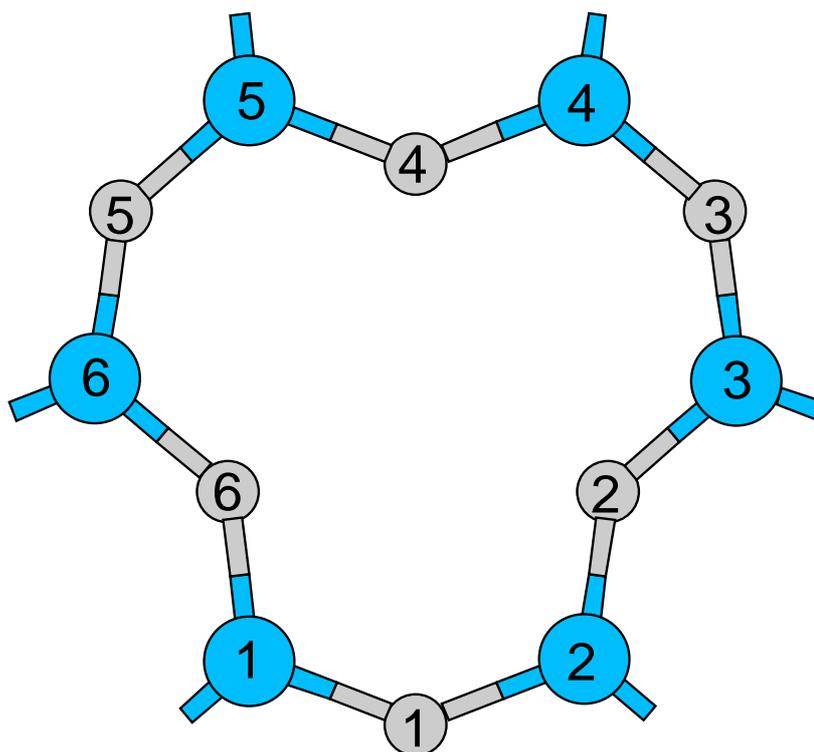


Figure 3.14: Ball and stick representation of a 6-membered ring present at the surface of the dense model. Si and O atoms are represented in dark and pale grey, respectively. The Si and O atoms have been labeled from 1 to 6 to allow to determine the bond lengths and bond angles in Table 3.2.

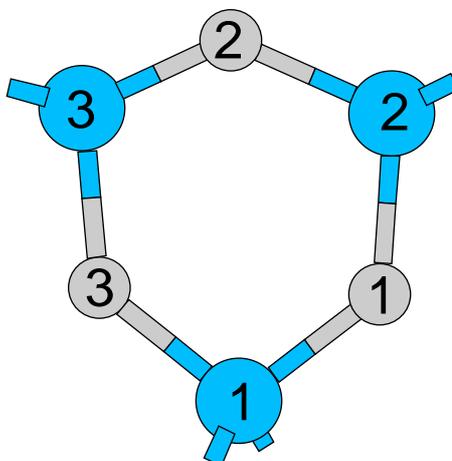


Figure 3.15: Ball and stick representation of a 3-membered ring present at the surface of the dense model. Si and O atoms are represented in dark and pale grey, respectively. The Si and O atoms have been labeled from 1 to 3 to allow to determine the bond lengths and bond angles in Table 3.3.

both the Si-Si-Si and O-O-O angles are practically equal to 60° indicating that Si and O form equilateral triangles. These two triangles are located in two separated planes which are rotated by 60° with respect to each other.

Finally, we have also computed the electronic densities on the 6-membered and the 3-membered rings. The contour plots are reproduced in Figs. 3.17 and 3.18, respectively. The views correspond to the geometry presented in Figs. 3.14 and 3.15. Its interesting to superimpose two figures to have a clear view of the location the atoms. In both case, we see that the valence electrons concentrate on the O atoms, revealing the strong ionic character of the Si-O bond. In the 6-membered ring, a quasi-perfect trigonal symmetry appears (the slight deviations are due to the unrelaxed unit cell dimensions) and the densification of the plane is evident. Note that the two peaks at the top of the figure correspond to 2 oxygen atoms which do not belong to the 3-membered ring. In fact, these belong to the same 6-membered ring than the upper O of the 3-membered ring.

The stability of the dense surface was tested by heating the surface up to 3500 K for a few hundred fs. No further reconstruction was observed indicating a relative stability of the structure.

¹By surface unit cell, we mean the one used in the simulations, i.e. a rectangle with

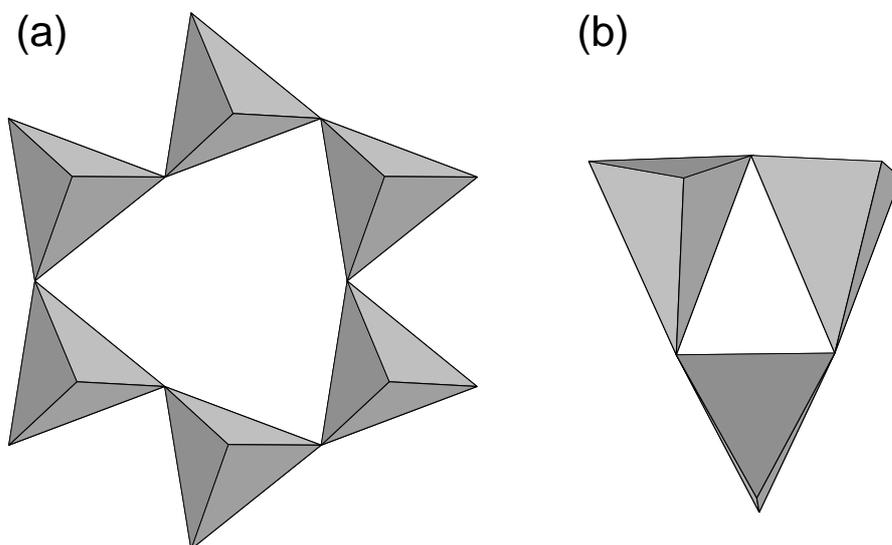


Figure 3.16: Representation of the dense surface in terms of tetrahedral units. (a) Top view (perpendicularly to the (0001) direction) showing a 6-membered ring, and (b) side view illustrating a 3-membered ring. In the 6-membered ring, all the tetrahedra have the same orientation; one of their faces is perpendicular to the (0001) direction.

Table 3.2: Bond lengths and bond angles (expressed in \AA and $^\circ$, respectively) for the 6-membered rings present at the surface in model IV. The numbering of the Si and O atoms refers to Fig. 3.14.

Si	O	Si	O	$d(\text{Si-O})$	$d(\text{O-Si})$	$\angle(\text{Si-O-Si})$	$\angle(\text{O-Si-O})$
1	1	2	2	1.58	1.58	129.6	110.7
2	2	3	3	1.58	1.59	118.8	109.2
3	3	4	4	1.58	1.58	121.1	110.2
4	4	5	5	1.58	1.58	129.6	108.5
5	5	6	6	1.59	1.58	118.8	108.6
6	6	1	1	1.58	1.58	121.2	109.8

Table 3.3: Bond lengths and bond angles (expressed in Å and °, respectively) for the 3-membered rings present just above the surface in model IV. The numbering of the Si and O atoms refers to Fig. 3.15.

Si	O	Si	O	$d(\text{Si-O})$	$d(\text{O-Si})$	$\angle(\text{Si-O-Si})$	$\angle(\text{O-Si-O})$
1	1	2	2	1.60	1.60	128.2	109.0
2	2	3	3	1.59	1.58	118.8	109.3
3	3	1	1	1.60	1.59	132.8	104.9

Table 3.4: Relative total energies for the 4 different models: dimer, cleaved, valence-alternation-pair (VAP), and dense surfaces (in ascending order of stability). The energies are expressed in eV/surface unit cell¹ with respect to that of the dense surface.

Model	Rel. Energy
Dimer	16.9
Cleaved	10.1
VAP	7.7
Dense	0.0

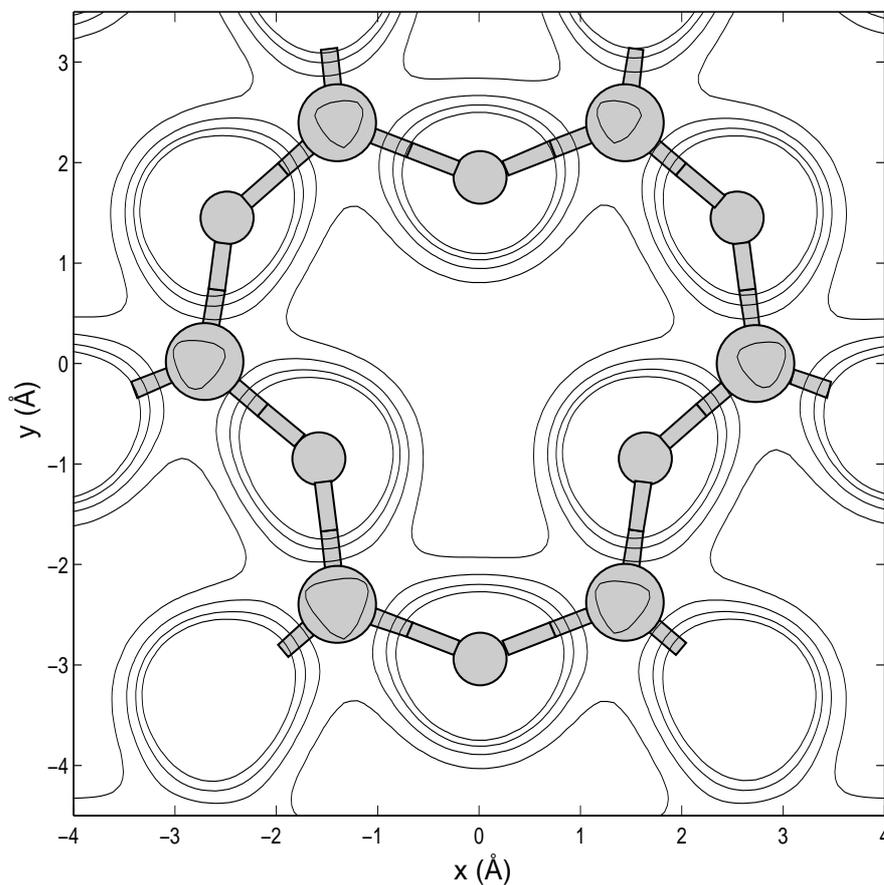


Figure 3.17: Contour plot of the electronic density on a 6-membered ring at the dense surface. The contour is drawn in a plane perpendicular to the (0001) direction of the crystal and located at the average height of the outermost O layer. The origin of the x and y axis (respectively, the (1000) and (0100) directions of the crystal) is the average value of the positions of the 6 oxygen atoms of the ring considered. The present view corresponds to the geometry presented in Fig. 3.14, that has been superimposed (in pale grey), as a guide for the eyes to locate the atoms. The contour interval is $0.5 \text{ e}^-/\text{\AA}^3$.

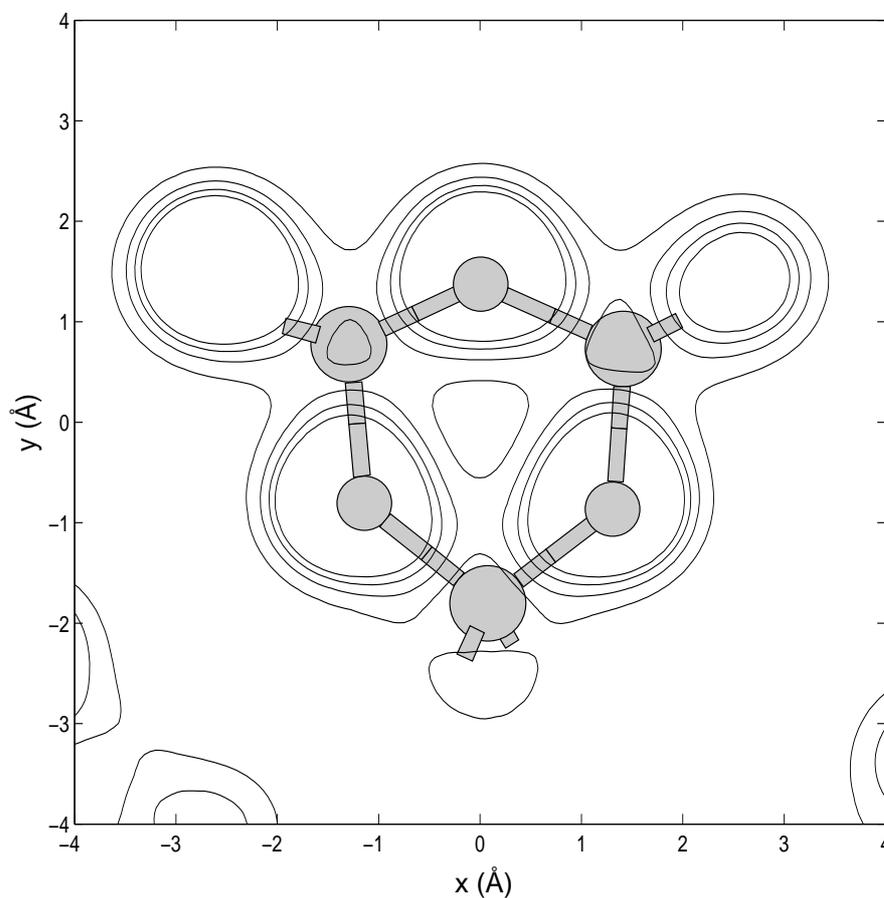


Figure 3.18: Contour plot of the electronic density on a 3-membered ring at the dense surface. The contour is drawn in a plane passing through the 3 O of the ring considered. The x axis is defined by the 2 oxygens located at the bottom. The origin of the x and y axis (defined by orthogonality with x) is the average value of the positions of the oxygen atoms. The present view corresponds to the geometry presented in Fig. 3.15, that has been superimposed (in pale grey), as a guide for the eyes to locate the atoms. The contour interval is $0.5 \text{ e}^-/\text{\AA}^3$.

To summarize, the relative total energies of the different models have been reported in Table 3.4 where the reference was chosen to be the dense surface. The dimer and dense surface are found to be the less and the most stable reconstructions, respectively. It is important to note that the VAP surface is more stable than the cleaved surface, probably due to the 4-fold coordination of all silicons. We would also like to emphasize that contrarily to valence alternation pairs, whose existence in amorphous SiO₂ had already been argued by Lucovsky (1979), the structure of the dense surface is proposed for the first time in this study. This prediction still needs to be confirmed, and we suggest hereafter some possible experiments.

First, with atomic force microscopy (AFM), it may be envisioned to access the microscopic structure of the surface directly. This has recently been tested by Noge *et al.* (1997). However, as already sketched in the introduction of this chapter, there remain some difficulties to be solved before being able to draw clear conclusions from this study. If these are overcome in the forthcoming years, as it may be expected, this technique will allow to control our prediction. Alternatively, X-ray diffraction performed under grazing angles of incidence (Grey and Feidenhans'1 1988) could be used to determine the surface crystallography that could then be compared to our reconstruction. Similarly, if ToF-ISS (Time of Flight Ion Scattering Spectroscopy) experimental data were obtained, our model could be introduced in the existing simulation codes to see if it reproduces the spectra. Finally, X-ray photoelectron spectroscopy of Si 2*p* core levels (see Chapter 2) would also provide a criterion to verify the validity of our prediction. Indeed, our calculations show that the oxidation state of the silicon atoms at the dense surface is identical to that of bulk SiO₂ atoms (i.e. Si⁺⁴). Hence, the presence of other oxidation states (Si^{+*x*} with *x*=1,2,3) in these spectra would invalidate our model, since it would imply the existence of defects at the surface. It should be noted that any other experiments that could reveal the existence of dangling bonds (e.g. EPR) could be considered in this respect.

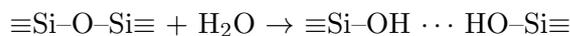
3.3 The wet surface

We now address some peculiar aspects of hydration and dehydration of the (0001) α -quartz surface, by studying at the atomic scale the interactions of this surface with water. Though most of the available experimental data have been obtained for amorphous silica, it has been found that different

sides $a=9.62$ Å and $b=8.34$ Å. In fact, it is a 2×2 cell with respect to bulk quartz unit cell.

forms of silica appear to act alike in regard to the adsorption of water (Iler 1979). In fact, particles size or radius of curvature of the surface were shown to be more important variables than differences between the amorphous and ordinary crystalline states of silica.

When wetting the SiO₂ surface, it is expected that silanol groups will form at the most reactive (strained) defects: non-bridging oxygens, two-membered rings, 3-coordinated Si (*E'* center), and 3-coordinated. It was calculated from first principles (Heggie and Jones 1987) that a 4% expansion of the Si-O bonds in siloxane was sufficient on energy grounds to allow the hydrolysis reaction:



to occur. On the contrary, for an hydrated surface, hydroxyl groups should condense as the temperature is raised to form siloxane bonds while water molecules are desorbed.

In the present work, we use Car-Parrinello molecular dynamics to study the interactions the (0001) α -quartz surface with water. On this purpose, we extend the tetragonal unit cell, previously used in the framework of the investigation of the dry surface, in order to include 4 water molecules, which can be either isolated or adsorbed at the surface. The surface unit cell is kept exactly the same with its sides being $a=9.62 \text{ \AA}$ and $b=8.34 \text{ \AA}$, respectively, based on the theoretical equilibrium lattice constant of α -quartz (Keskar and Chelikowsky 1992). The dimension of the cell in the direction orthogonal to the surface is increased to $c=21.17 \text{ \AA}$. This allows for the H₂O molecules to be located more than 7 \AA away from the surface (this is far enough to suppress any interaction), while the α -quartz still consists in 6 monolayers of SiO₂ with its bottom extremities saturated with hydrogen atoms. Globally, our systems contain 24 Si atoms, 48 O atoms and 16 H atoms. It should also be mentioned that the sides of the surface unit cell are large enough to allow for a separation of at least 4 \AA between the water molecules. Thus, it can also be consider that these do not interact one with each other.

The atomic positions are fully relaxed using the Car-Parrinello method (Car and Parrinello 1985), which provides the electronic structure as well as the forces that act on the ions. Only valence electrons are explicitly considered by means of norm conserving pseudopotentials (Troullier and Martins 1991) to account for the core-valence interactions. The electronic wave functions are expanded into plane waves with a kinetic energy cut-off of 50 rydbergs. Exchange and correlation are included using PW'91 approximation for the exchange-correlation energy functional (Perdew and Wang 1991, Perdew 1991). Gradient corrections are needed for the accurate

Table 3.5: Dehydration of a variety of types of silica surfaces measured by the number of hydroxyl groups at the surface (OH nm⁻²) as a function of the temperature (From Iler 1979).

Temperature (°C)	OH nm ⁻²
100	4.5-6.2
300	2.5-4.2
500	1.5-3.1
700	1.2
900	0.7
1100	0.4

treatment of hydrogen bonds. The Brillouin zone (BZ) is sampled using only the Γ -point in all our calculations. In the minimization process, the atoms of the lowest SiO₂ monolayer are kept fixed. For the dynamics, the time step was taken to be 0.17 fs.

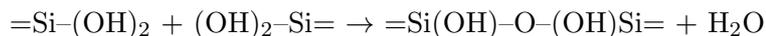
3.3.1 Dehydration of wet surfaces

At ordinary temperature and in normal ambient (not in vacuum), the amorphous SiO₂ surface has a concentration of about 4.5 to 6.2 hydroxyl groups (OH) per nm². If we consider the (0001) α -quartz surface, each Si atom at the surface should bear 2 OH to have its valence satisfied, giving about 10 OH nm⁻². We have build a model corresponding to this picture (Fig. 3.19(a)) by simply saturating the Si atoms at the top of the cleaved surface with the 4 water molecules:



in our tetragonal unit cell. We will refer to this structure as the *fully-hydrated surface*.

When the temperature is increased, the number of hydroxyl groups gradually decreases at the surface as reported in Table 3.5. It is generally believed that hydroxyl groups condense to form siloxane bonds while a water molecule is removed. If we consider our fully-hydroxylated surface, we can assume the following dehydration mechanism:



corresponding to about 5 OH nm^{-2} . The resulting model, depicted in Fig. 3.19(b), will be referred to as the *semi-hydrated surface*. If dehydration proceeds further, the two remaining OH groups could also condense so that a 2-membered ring (edge-sharing tetrahedra) would form as illustrated in Fig. 3.19(c).

As the energy barriers involved in the previously described processes were not known precisely from experiments, we tried to model the dehydration mechanism of the fully-hydrated surface using Car-Parrinello molecular dynamics with a Nosé thermostat to regulate the temperature. During this simulation, we did not observe any condensation of hydroxyl groups in the time scale accessible to our simulations ($t \approx 2 \text{ ps}$) even at quite high temperature ($T=1500 \text{ K}$). So, we analyzed the relative stability of the different configurations illustrated in Fig. 3.19. It turned out that the fully-hydrated surface is by $2.9 \text{ eV/surface unit cell}$ more stable than the semi-hydrated surface which in turn is energetically favored with respect to the surface with 2-membered rings ($\Delta E \approx 10 \text{ eV/surface unit cell}$). These barriers are too large to allow the simulation of the reaction paths using direct molecular dynamics. Using constrained path dynamics instead might help to overcome this difficulty. However, in order to use such approaches, one must have an idea of a reasonable reaction path.

To gain insight, we decided to study the hydration of the dry surface which is in principle energetically favored. Indeed, the fully-hydrated surface is by 4.6 and $14.1 \text{ eV/surface unit cell}$ more stable than the dense surface and the cleaved surface, respectively.

3.3.2 Hydration of the dense surface

It is generally believed that water molecules are adsorbed only on the hydroxylated silica surface and not on the siloxane surface which is essentially hydrophobic. However, hydration must involve adsorption of water as a first step, so that water is probably chemisorbed by opening up strained siloxane bonds initially and forming silanol groups, then it is adsorbed on the latter rather than on the siloxane surface. Further hydration occurs only on siloxane oxygen sites adjacent to a silanol site, so that the hydroxylated areas grow in patches as hydration proceeds along the boundary between the hydroxylated and siloxane regions.

We now present our simulations of the hydroxylation of the dense α -quartz surface obtained in dry conditions. These consist in throwing a water molecule on the dense surface, heated at 300 K beforehand, at different velocities (i.e. with different kinetic energies). In order to do so, a constant

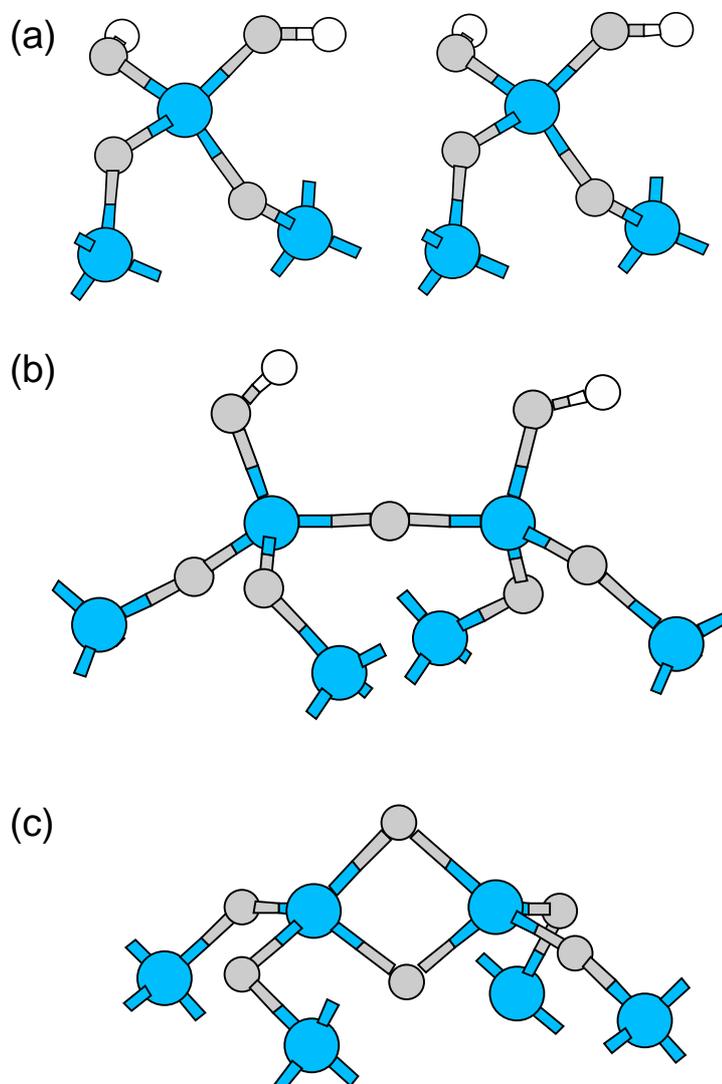


Figure 3.19: Dehydration mechanism for the wet (0001) α -quartz surface. The illustration starts with (a) the fully-hydrated surface. Then a water molecule is removed after condensation of two hydroxyl groups to form a siloxane bond leading to (b) the semi-hydrated surface. The mechanism ends by (c) the formation of a 2-membered ring after further condensation of the remaining hydroxyl group and desorption of water. Si, O, and H atoms are represented in dark grey, pale grey, and white respectively.

acceleration is given to the atoms of the molecule that is to be thrown. This scheme is thus a second-order change in the molecular dynamics, whereas imposing directly the velocity to the molecule would be a first-order change. Hence, our procedure acts very softly on the dynamics.

We tested various velocities for the water molecule corresponding respectively to kinetic energies of 0.1, 1, 10, 15, and 20 eV. But each time, the H₂O molecule rebounded on the surface. Hence, we decided to perform static calculations to determine the potential barrier for the adsorption of the molecule on the surface. The water molecule was put at various distances from the surface (as determined by the distance between its center of mass and the upper plane of oxygens). The total energy was then calculated with the oxygen of the molecule kept fixed while the hydrogens were allowed to relax. The result of these calculations are reproduced in Fig. 3.20.

Another picture of this barrier can be obtained by plotting the minimal distance from the surface, reached by the water molecule for various throwing energies. Indeed, if we consider a purely elastic collision, the molecule will rebound when the repulsive energy equals its kinetic energy. These distances have also been reported in Fig. 3.20. In both the static and the dynamic calculations, the data roughly exhibit an exponential behavior (see interpolations on the figure). Moreover, there is a quite good agreement between the two results given the fact that in the static calculations the hydrogens are allowed to fully relax for each distance and that in the dynamic calculation we consider a purely elastic collision. The exponential behavior suggests that there exists an minimal distance under which the water molecule can not go.

This tends to confirm that surfaces formed by siloxane are essentially hydrophobic (Iler 1979). Moreover, there is no contradiction with previous *ab initio* calculations (Heggie and Jones 1987), since in the dense surface, the Si-O bonds are not strained with respect to the bulk (see section 3.2.6). And thus, the hydrolysis reaction of the siloxane is not energetically favored. Indeed, we found that the dense surface in which we would have forced this reaction to occur for a siloxane bond between two Si of the outer layer, is 4.69 eV/surface unit cell higher in energy. We will refer to this structure with 2 OH groups on the dense surface as the *hydrated-siloxane dense surface*.

The hydrophobic character of the dense surface is further confirmed by the fact that by heating the hydrated-siloxane dense surface up to 300 K, we observe the dehydration (i.e. the desorption of the water molecule) after 300 fs.

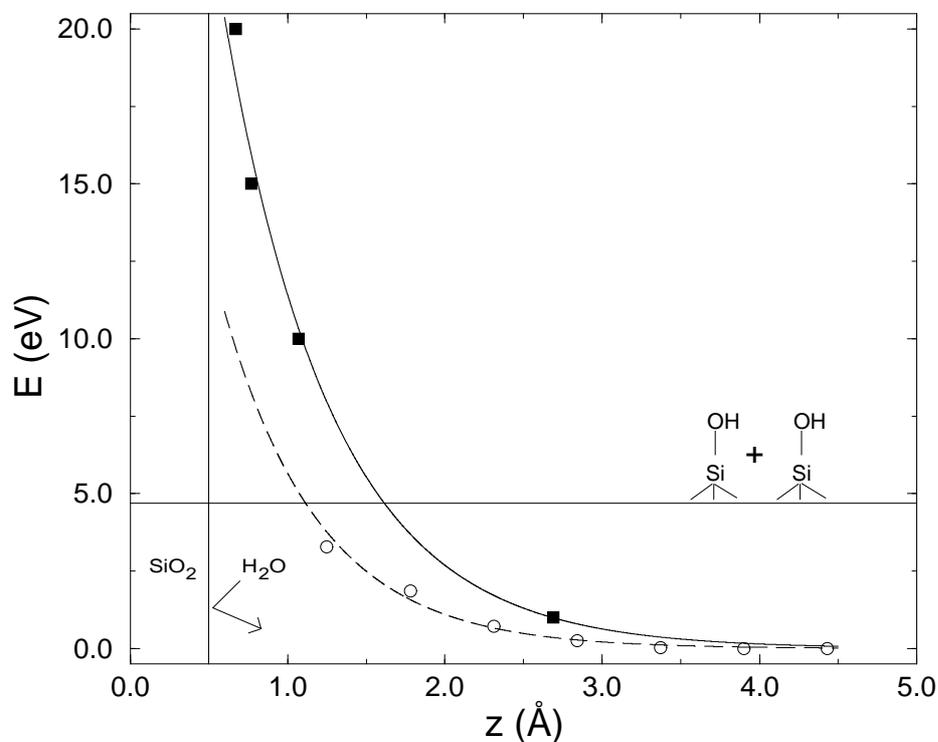


Figure 3.20: Repulsive potential (E expressed in eV) for the absorption of a water molecule on the dense surface as a function of the distance (z in Å) of the center of mass with respect to upper plane of oxygens of the structure. The reference energy is taken to be that of the dense surface with the water molecule more than 7 Å away. We have represented by an horizontal line at 4.69 eV, the energy of the structure that would be obtained after the hydrolysis of a siloxane by the water molecule: $\equiv\text{Si}-\text{O}-\text{Si}\equiv + \text{H}_2\text{O} \rightarrow \equiv\text{Si}-\text{OH} \cdots \text{HO}-\text{Si}\equiv$. We have reported the energies determined by static calculations (open circles) for various distances. We have also indicated the minimal distance reached during the dynamic runs when the molecule was thrown with a kinetic energy of 1, 10, 15, and 20 eV (filled squares). In both cases, the data roughly present an exponential behavior as suggested by the interpolations (dashed and solid lines for static and dynamic calculations, respectively). Finally, we have drawn the probable asymptote for the barrier which tends to indicate that the hydrolysis will never occur.

3.3.3 Wetting of partially hydroxylated surfaces

Finally, we also simulated the hydroxylation of partially hydroxylated α -quartz surfaces (i.e. obtained in ambient conditions), by throwing a water molecule on these. More specifically, we considered a cleaved surface in which two NBO's had already been hydroxylated, giving approximately 5 OH nm⁻². Hence, there remained two non-bridging oxygens in our unit cell. The surface had been heated to 300 K beforehand, and the water molecule was thrown with a kinetic energy of 1 eV.

The simulation revealed the hydrolysis of one of the non-bridging oxygens:



At the atomic level, the mechanism, illustrated in Fig. 3.21, is proposed to be the following. First, the oxygen of the water molecule, O_W , bonds to the silicon with the NBO, so that silicon is 4-coordinated as depicted in part (b) of Fig. 3.21. However, O_W is 3-coordinated, while the NBO is still 1-coordinated. The excess hydrogen on O_W does not go immediately on the NBO. Instead, it creates an hydrogen bond with the oxygen of a neighboring hydroxyl group. In turn, this oxygen is quasi 3-coordinated and hence one of its hydrogen starts bonding with another adjacent hydroxyl. In fact, a whole chain of hydrogen bonds is formed starting from O_W and ending with the NBO (due to periodic boundary conditions). After a few oscillations of the hydrogen atoms from one oxygen to another, the protons finally settle in such a way that the NBO has turned into an hydroxyl group as depicted by the hydrolysis reaction above and shown in part (c) of Fig. 3.21..

It remains uncertain whether the formation of an hydrogen-bond chain is the real physical mechanism for the transfer of the excess H from O_W to the NBO, or if this is due to the periodic boundary conditions. However, the important result is that hydrolysis occur.

The remaining isolated water molecule has also been thrown on the surface which had only one NBO left and 7.5 OH nm⁻². This simulation also ends with the hydrolysis of the silicon with the NBO and again the formation of hydrogen-bond chains is observed, confirming the role of adjacent hydroxyl groups. The creation of these chains might be a interesting reaction path to consider for the simulation of dehydration within constrained path dynamics.

To complete the study, it would be interesting to make a similar computation with the semi-hydrated surface to investigate what happens to its siloxane bonds and see if the hydroxyl groups that are present at this surface

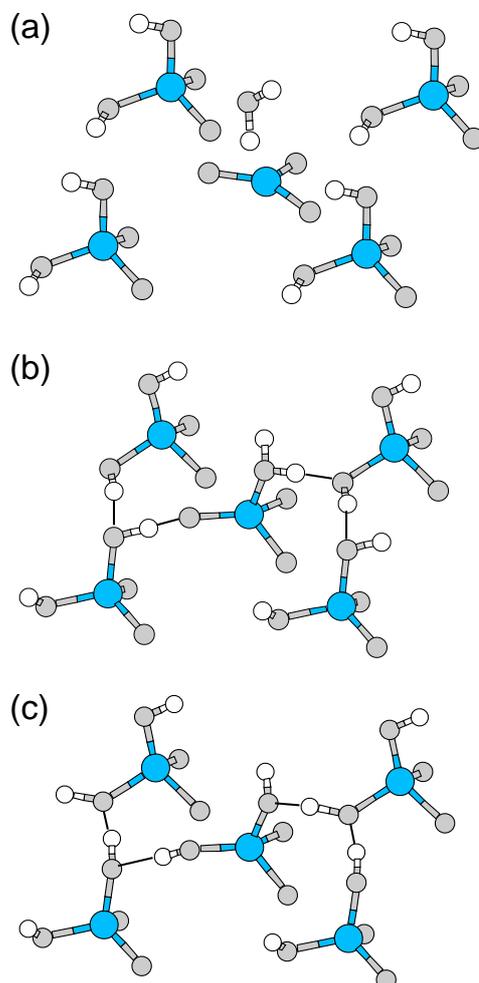


Figure 3.21: Hydration mechanism for a partially hydroxylated surface. The illustration starts with (a) a water molecule approaching the surface, which presents a silicon with a non-bridging oxygen (at the center of the picture). The oxygen of the water molecule bonds to the silicon with the NBO so that the Si atom is 4-coordinated (b) and a chain of hydrogen bonds starts forming from one hydroxyl to another (as indicated by the solid lines). After a few oscillations of the hydrogen atoms from one oxygen to another (the limit cases being (b) and (c)), the mechanism ends with (c) the central $=\text{Si-NBO}$ hydrolyzed into a $=\text{Si}(\text{OH})_2$ group. Si, O, and H atoms are represented in dark grey, pale grey, and white respectively.

allow the hydrolysis of these bonds that were shown to be hydrophobic at the dense surface. The simulation of the hydration of other surfaces such as the VAP surface or the cleaved surface could provide information on the hydration mechanism of amorphous surfaces.

To summarize, we have presented our preliminary results related to the interaction of the quartz surface with water. We first investigated dehydration of wet surfaces. Unfortunately, our estimation of energy barriers that are involved revealed that these are too high to try to follow the reaction paths by direct molecular dynamics technique, suggesting that constrained path dynamics should be used instead. In order to gain insight on the possible reaction paths, we turned to the study of hydration which is in principle energetically favored. We performed molecular dynamics simulating the projection water molecules on various surface. Our study of the hydration of the dense surface demonstrated the hydrophobic character of the siloxane bonds. We explained that this is not in contradiction with previous *ab initio* calculations demonstrating that a 4% expansion of the Si-O bonds in siloxane was sufficient on energy grounds to allow its hydrolysis. Indeed, in the dense surface, the Si-O bonds are not strained with respect to the bulk (as shown in the study of the dry surface), so that the hydrolysis reaction of the siloxane is not energetically favored. On the contrary, when the surface was partially hydroxylated, the hydrolysis of silicon atoms with non-bridging oxygen was observed. The adjacent silanol groups play an important role in the hydration mechanism, by the formation of hydrogen-bond chains. This might be a possible reaction path to consider for the simulation of dehydration within constrained path dynamics.

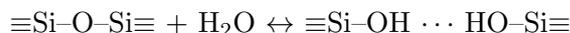
3.4 Conclusion

In this chapter, we presented a first-principles study of the (0001) α -quartz surface. We distinguished the surfaces obtained in vacuum (dry surface) from those in ambient conditions which contain water (wet surface), which both play an important role in many advanced technologies. The (0001) α -quartz surface was chosen as a qualitative model for the different crystallographic forms of SiO_2 , as well as amorphous silica.

In section 3.2, we proposed an analysis of the dry surface. Before addressing the study of the surface itself, we discussed briefly the structure of the bulk. We showed how α -quartz may be considered as consisting of corner-sharing tetrahedra in a trigonal network, or alternatively of repeated O-Si-O layers along the (0001) direction. These views were presented in the

scope of facilitating the ulterior description of the surface structure. The study of the dry surface started with two different initial configurations: the cleaved surface with non-bridging oxygens at the top and a 2×1 reconstruction with 2-membered rings. Two other reconstructions were obtained by performing constant-temperature molecular dynamics simulations: the valence alternation pair surface, and the dense surface. The former presents an intimate pair of over- and under-coordinated oxygen atoms near the surface with 3-membered rings. The latter, which is the most stable of all investigated structure, presents an interesting densification of the two uppermost layers of SiO₂ tetrahedral units, with the formation of 3-membered and 6-membered rings that do not exist in bulk α -quartz. In all the models, the analysis of the structural properties showed that the bulk structure is recovered within 5 Å from the surface. Contrarily to valence alternation pairs, whose existence in amorphous SiO₂ is well-know, the structure of the dense surface is proposed for the first time in this study and we suggested experiments to confirm this prediction.

In section 3.3, we turned to the study of the wet surface. Unfortunately, at the moment of writing, many calculations were still in progress so that we presented only a limited number of results. The fundamental reaction that we wanted to investigate is:



which schematizes the hydration/dehydration mechanism. On one hand, our molecular dynamics simulation of the dehydration of the wet quartz surface did not give any significant result. Indeed, the time scale accessible to our simulation was too small in regard to the large energy barriers (about 1.5 eV/surface unit cell at least) that are involved. On the other hand, the hydration was simulated by throwing water molecules on various quartz surface using Car-Parrinello molecular dynamics. For the dense surface, the water molecules rebounded on the surface whatever their kinetic energies showing the hydrophobic behavior of siloxane bonds. We explained that this is not in contradiction with previous *ab initio* calculations which suggested that a 4% expansion of the Si-O bonds in siloxane was sufficient on energy grounds to allow its hydrolysis. For the partially hydroxylated surfaces, the water molecules bonded to the under-coordinated silicons (those linked to the non-bridging oxygens) and then dissociated the hydrolysis reaction:



At the atomic level, the formation of hydrogen-bond chains was demonstrated to play a key role in the reaction mechanism. Moreover, we proposed

that this could be a good clue towards a reasonable reaction path that might be considered for the simulation of dehydration within constrained path dynamics.

Conclusion

Present-day computer technology has strongly modified the way in which modern research is directed. First, at a rather simple level, computers have considerably speeded up the march of scientific research, by allowing for tasks, that would be almost infeasible otherwise, to be performed automatically and much more faster. Second and more impressively, computers have become real research instruments to investigate scientific topics. This is mainly due to the advent of computer simulations, that is to say the ability to track on the computer the evolution of models describing with ever greater accuracy the behavior of ever more complex systems. This is unprecedented and has led to the birth of a new way of doing science, between theory and experiment.

The results of the simulations are of great help in the study of problems in physics and in chemistry, both on the experimental and theoretical grounds. They allow the study of systems that would have been difficult or expensive to work with in laboratory. However, the most rewarding outcome of these simulations is by far the invaluable insight they provide into the way systems behave at the microscopic level. In this regard a major role has been played by classical Molecular Dynamics, which since the pioneering efforts of the sixties has developed into a mature and active discipline. Another step forward was achieved towards more realistic models with the development of the First-Principles Molecular Dynamics in the late eighties, starting with the work of Car and Parrinello. This has greatly improved the accuracy and the predictive power of the simulations. That is what we tried to assess in this thesis by applying this technique to the study of two non-periodic systems: the SiO_2 surface and its interface with silicon.

In the first chapter of this work, the theoretical basis of the Car-Parrinello method have been presented. We first described the concepts underlying the classical Molecular Dynamics technique. Then, we introduced Density Functional Theory to deal with the ground-state of a quantum-mechanical system of electrons, by obtaining a single-orbital picture (Kohn-Sham for-

malism) of the many-body problem. We illustrated how by combining these two approaches it is possible to use forces derived within first-principles in MD simulations. Finally, we discussed the practical implementation of the Car-Parrinello scheme.

In the second chapter, we have investigated the incorporation of nitrogen at the Si(001)-SiO₂ interface within the first-principles approach presented in the first chapter. This problem is particularly relevant in the framework of very large scale integration of MOS devices. Indeed, the incorporation of a low concentration of N atoms near the interface appears as one of the most promising ways to match industrial programs requiring high-quality ultrathin gate dielectrics. Further improvement of the quality of electronic devices critically relies on information regarding the situation of incorporated nitrogen atoms. X-ray photoelectron spectroscopy of N 1s core-level has been one of the principal experimental tools of this purpose. Indeed, it can provide information on both the distribution profiles and the bonding environment of the incorporated N atoms, which depend on the growth process. There is a general agreement concerning distribution profiles, which often show an accumulation of N atoms at the interface. On the contrary, the bonding configuration is still a debated issue. Therefore, we established a correspondence between the bonding properties of these incorporated nitrogen atoms and the N 1s core-level shifts as measured in photoemission experiments. Our study leads to an interpretation of N 1s photoemission spectra, in which N atoms always form three bonds to silicon atoms, both in the neighborhood of the interface and further in the oxide. It explains the asymmetric shape of the principal peak in experimental data as well as its anomalous shift with oxide thickness, which are the most misunderstood experimental features. Core-hole relaxation and second nearest neighbor effects were shown to combine to give larger binding energies in the oxide than at the interface, in accord with experimental observations. Core-hole relaxation influences N 1s shifts differently according to the distance of the N atoms to the screening Si substrate. The second nearest neighbor environment at the interface (mostly Si and N atoms) induces opposite shifts compared to that in the oxide (rich in O atoms). Our calculations also tend to rule out the existence of nitrogen bonds involving oxygen atoms at the interface. Indeed, we obtained $\Delta=1.5$ eV for N-Si₂O configurations and this does not correspond to any observed experimental peak. The existence of nitrogen bonds involving hydrogen atoms is not excluded a priori, since the calculated shifts ($\Delta=0.4$ eV) is perfectly compatible with the values obtained in experiments. However, these require such a high concentration of hydrogen with respect to nitrogen that if these appeared in nitrified in-

terface it would only be in a very small amount. We complemented this investigation by analyzing Si $2p$ core-level shifts at the interface. Our calculations brought us to suggest that, contrarily to what is generally believed, both Si-O and Si-N bonds contribute to the suboxide peaks in the Si $2p$ photoemission spectra.

In the third chapter, we studied the (0001) α -quartz surface using the first-principles method developed in the first chapter. This surface was chosen as a qualitative model for the different crystallographic forms of SiO_2 , as well as amorphous silica. In our study, we considered both the surfaces obtained in vacuum (dry surface) and those formed in ambient conditions which contain water (wet surface). These surfaces play an important role in many advanced technologies (adsorption, surface diffusion, nucleation, microelectronics, coatings, sensor devices, ...). To model the dry surface, we considered two different initial configurations: the cleaved surface with non-bridging oxygens at the top and a 2×1 reconstruction with 2-membered rings. By performing constant-temperature molecular dynamics simulations, we obtained two other reconstructions: the valence alternation pair surface, and the dense surface. The former presents an intimate pair of over- and under-coordinated oxygen atoms near the surface with 3-membered rings. The latter, which is most stable of all investigated structure, presents an interesting densification of the two uppermost layers of SiO_2 tetrahedral units, with the formation of 3-membered and 6-membered rings that do not exist in bulk α -quartz. Contrarily to valence alternation pairs, whose existence in amorphous SiO_2 is well-known, the structure of the dense surface was proposed for the first time in this study and we suggested experiments that might be done to confirm this prediction. The models that have been generated can be used in further studies. In regard with the wet surface, the hydration/dehydration mechanisms were investigated. No significant conclusion could be drawn from the molecular dynamics simulation of the dehydration of the wet quartz surface (heating at high temperature), probably due to the shortness of the accessible time scale with respect to the large energy barriers involved. On the contrary, the simulations of the hydration (throw of water molecules on various surface) demonstrated the hydrophobic behavior of siloxane bonds. We illustrated a possible mechanism for the hydrolysis of non-bridging oxygens, based on the formation of hydrogen chains with the adjacent hydroxyl groups.

In conclusion, the use of first-principles approaches in the context of non-periodic SiO_2 systems (such as its surface and its interface with silicon) has proven to be very useful to investigate topics which are still unclear so far. The application of such techniques to these systems was made possible

by recent developments in the theoretical techniques as well as by advances in computer architectures (massively parallel computers, shared-memory supercomputers) and performances. If these progresses keep going on, it should be possible in a near future to simulate systems of more realistic size. In a distant future, we can even expect that the short time scale restriction, which is one of the most severe limitations of Molecular Dynamics, will be overcome. This would open the way to study other important problems in regard with SiO_2 such as the diffusion of impurities (e.g. boron, fluorine, or hydrogen), the early stages of oxidation of the silicon surface, the amorphization process ...

Appendix A

Beyond the LDA

During the last two decades, a number of recipes have been introduced to improve the local-density approximation. In this Appendix, we propose a brief description of the most commonly used approximations, while we refer to the abundant literature for a systematic discussion. We will mention here the gradient expansion approximation (GEA); the average density approximation (ADA) and the weighted density approximation (WDA) by Gunnarsson, Lundqvist and collaborators; the generalized gradient approximation (GGA) developed by Perdew, Langreth, Mehl; and, finally, the self-interaction correction (SIC) of Perdew and Zunger.

The simplest possibility of adding some more information on the density distribution is via the gradients of $n(\mathbf{r})$ in the exchange-correlation energy functional. In this gradient expansion approximation (GEA), the symmetry and dimensional arguments determine the form of the first correction in a Taylor-like functional expansion (Herman *et al.* 1969):

$$\epsilon_{XC}^{GEA}(\mathbf{r}) = \epsilon_{XC}^{LDA}(\mathbf{r}) + \alpha \frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})^{7/3}} \quad (\text{A.1})$$

where α is a constant determined by the response function of the homogeneous electron gas.

Early computations by Lau and Kohn (1976) and Perdew *et al.* (1977), however, made it clear that, far from providing an improvement, GEA was in fact giving worse results than LDA. The analysis of this failure carried out in terms of the size and the shape of the exchange-correlation hole evolved towards the average density approximation (ADA) and the weighted density approximation (WDA) (Gunnarsson *et al.* 1976, Gunnarsson *et al.* 1977, Gunnarsson *et al.* 1979, Gunnarsson and Jones 1980). These functionals are still under evaluation.

Following a different line of investigation, Langreth, Perdew, Mehl and co-workers concentrated on the Fourier-space analysis of the exchange-correlation. They were able to show that LDA provides the exact limit of the Fourier-components of E_{XC} for large wave vectors, whereas the random-phase approximation (RPA) gives the exact limit for small wave vectors Langreth and Mehl (1981). Interpolating between these two limits, Langreth and Mehl (1983) proposed the following expression for $\epsilon_{XC}(\mathbf{r})$ in terms of the density n and its gradient:

$$\epsilon_{XC}^{LM}(\mathbf{r}) = \epsilon_X^{LDA}(\mathbf{r}) + \epsilon_C^{RPA}(\mathbf{r}) + \alpha \frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})^{7/3}} \left(2e^{-F} - \frac{7}{9} \right) \quad (\text{A.2})$$

where $F = b|\nabla n(\mathbf{r})|/n(\mathbf{r})^{7/6}$, $b = (9\pi)^{1/6}f$, $\alpha = \pi/(16(3\pi^2)^{4/3})$, and $f = 0.15$. This approximation and its successive refinements are referred to as generalized gradient approximations (GGA's):

$$\epsilon_{XC}^{GGA}(\mathbf{r}) = \epsilon_{XC}^{hom}[n(\mathbf{r}), \nabla n(\mathbf{r})] \quad (\text{A.3})$$

Though appealing the LM exchange-correlation has several drawbacks (Perdew 1986, Langreth and Mehl 1984). To overcome these problems, the exchange-correlation energy was split into its exchange and correlation contribution. For the exchange part, Perdew and Wang (1986) proposed a gradient expansion based on the analysis of the behavior of the exchange hole in inhomogeneous systems:

$$\epsilon_X^{PW'86}(\mathbf{r}) = \epsilon_X^{LDA}(\mathbf{r}) \left(1 + 0.0864 \frac{s^2}{m} + bs^4 + cs^6 \right)^m \quad (\text{A.4})$$

where $m = 1/15$, $b = 14$, $c = 0.2$, and a scaled density gradient $s = s(\mathbf{r})$ has been defined by:

$$s = \frac{|\nabla n|}{2k_F n} \quad (\text{A.5})$$

where the local Fermi wave vector $k_F = k_F(\mathbf{r})$ is

$$k_F = (3\pi^2 n)^{1/3}. \quad (\text{A.6})$$

For the correlation part, Perdew (1986) introduced an interpolation formula based on the electron-gas results of Hu and Langreth (1986) and Rasolt and Geldart (1976):

$$\epsilon_C^{PW'86}(\mathbf{r}) = \epsilon_C^{LDA}(\mathbf{r}) + e^{-\Phi} C_c(n) \frac{|\nabla n(\mathbf{r})|^2}{n(\mathbf{r})^{7/3}} \quad (\text{A.7})$$

where

$$\Phi = 1.745\tilde{f}\frac{C_c(\text{inf})}{C_c(n)}\frac{|\nabla n(\mathbf{r})|}{n(\mathbf{r})^{7/6}}, \quad (\text{A.8})$$

$$C_c(n) = C_1 + \frac{C_2 + C_3r_s + C_4r_s^2}{1 + C_5r_s + C_6r_s^2 + C_7r_s^3}, \quad (\text{A.9})$$

and $\tilde{f} = 0.11$, $C_1 = 0.001667$, $C_2 = 0.002568$, $C_3 = 0.023266$, $C_4 = 7.389 \times 10^{-6}$, $C_5 = 8.723$, $C_6 = 0.472$, $C_7 = 7.389 \times 10^{-2}$. The combination of these two functionals to obtain $\epsilon_{XC}(\mathbf{r})$ is referred to as PW'86 approximation.

This scheme was shown to provide good results for the exchange energy of atoms. However, the satisfactory performance of the exchange energy is partly due to the compensation of errors in different regions of the charge distribution. Moreover, the PW'86 potential fails to reproduce the correct $-1/r$ behavior of the exchange potential far from the nucleus. Thus, on semi-empirical grounds, Becke (1988) proposed the following interpolation for the $\epsilon_X(\mathbf{r})$:

$$\epsilon_X^{BP}(\mathbf{r}) = \epsilon_X^{LDA}(\mathbf{r}) \left(1 - \frac{\beta}{2^{1/3}\alpha} \frac{x^2}{1 + 6\beta x \sinh^{-1}(x)} \right) \quad (\text{A.10})$$

where $x = 2(6\pi^2)^{1/3}s = 2^{1/3}|\nabla n(\mathbf{r})|/n(\mathbf{r})^{4/3}$, $\alpha = (3/4)(3/\pi)^{1/3}$, and $\beta = 0.0042$. In general, this approximation is implemented with the Perdew (1986) formula for the correlation, the combination being referred to as BP approximation. This scheme has been shown to correct substantially a large set of molecules (Ortiz and Ballone 1991) and solids (Ortiz 1992, Garcia *et al.* 1992, Filippi, Singh and Umrigar 1994, Khein *et al.* 1995), and to accurately predict bond lengths in weakly bonded systems (Lee *et al.* 1992, Sim *et al.* 1992, Laasonen, Sprik, Parrinello and Car 1993, Dal Corso and Resta 1994). However, calculations on covalent solids indicate a tendency of the BP scheme to underestimate bulk moduli (Garcia *et al.* 1992, Dal Corso *et al.* 1994, Filippi, Singh and Umrigar 1994, Khein *et al.* 1995).

Recently, new functionals (Perdew and Wang 1991, Perdew 1991) were proposed for the exchange and correlation energy per particle functionals, which share several properties with the exact functional. The exchange part is written:

$$\epsilon_X^{PW'91}(\mathbf{r}) = \epsilon_X^{LDA}(\mathbf{r}) \left(\frac{1 + a_1s \sinh^{-1}(a_2s) + (a_3 + a_4e^{-100s^2})}{1 + a_1s \sinh^{-1}(a_2s) + a_5s^4} \right) \quad (\text{A.11})$$

where $a_1 = 0.19645$, $a_2 = 7.7956$, $a_3 = 0.2743$, $a_4 = -0.1508$, $a_5 = 0.004$. The correlation part is written:

$$\epsilon_C^{PW'91}(\mathbf{r}) = \epsilon_C^{LDA}(\mathbf{r}) + H(n, s, t) \quad (\text{A.12})$$

where $t = t(\mathbf{r})$ is a new scaled density gradient defined by:

$$t = \frac{|\nabla n|}{2k_s n} \quad (\text{A.13})$$

with the local screening wave vector $k_s = k_s(\mathbf{r})$ written as:

$$k_s = (4k_F/\pi)^{1/2}. \quad (\text{A.14})$$

The correction factor to the LDA correlation functional is given by:

$$H = \frac{\beta^2}{2\alpha} \ln \left(1 + \frac{2\alpha}{\beta} \frac{t^2 + At^4}{1 + At^2 + A^2t^4} \right) + C_{c0}[C_c(n) - C_{c1}]t^2 e^{-100s^2} \quad (\text{A.15})$$

where

$$A = \frac{2\alpha}{\beta} [e^{-2\alpha\epsilon_C^{LDA}(\mathbf{r})/\beta^2} - 1]^{-1} \quad (\text{A.16})$$

and $\alpha = 0.09$, $\beta = 0.0667263212$, $C_{c0} = 15.7559$, $C_{c1} = 0.003521$. The function $C_c(n)$ is the same as for the PW'86 correlation functional. The combination of these two functionals to obtain $\epsilon_{XC}(\mathbf{r})$ is referred to as PW'91 approximation.

Note that more recently, Perdew, Burke and Ernzerhof (1996) introduced simpler functionals including improvement over PW'91 and in which all parameters (other than those in $\epsilon_{XC}^{LDA}(\mathbf{r})$) are fundamental constants. The exchange functional is written:

$$\epsilon_X^{PBE}(\mathbf{r}) = \epsilon_X^{LDA}(\mathbf{r}) \left(1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \right) \quad (\text{A.17})$$

where $\mu = 0.21951$ and $\kappa = 0.804$. Whereas, the correlation functional is written:

$$\epsilon_C^{PBE}(\mathbf{r}) = \epsilon_C^{LDA}(\mathbf{r}) + H(n, t) \quad (\text{A.18})$$

where

$$H = \gamma \ln \left(1 + \frac{\beta}{\gamma} t^2 \frac{1 + At^2}{1 + At^2 + A^2t^4} \right) \quad (\text{A.19})$$

and

$$A = \frac{\beta}{\gamma} [e^{-\epsilon_C^{LDA}(\mathbf{r})/\gamma} - 1]^{-1} \quad (\text{A.20})$$

and $\beta = 0.0667263212$, $\gamma = 0.031091$.

There exists a variety of other generalized gradient approximations that have been proposed in literature. The list presented here is of course not exhaustive (see, for instance, Filippi, Umrigar and Taut (1994) for some more approximate forms of the exchange-correlation functional).

Finally, other proposals, related to the Density Functional Theory, but going beyond it, improve it by using directly functionals of the one-body non-interacting orbitals instead of a functional of the density (playing upon the fact that in any case the kinetic functional is evaluated through the use of orbitals). In this class fall the so-called self-interaction corrected (SIC) schemes (Perdew and Zunger 1981). One of these is the so-called optimized effective potential (OEP) method which was discovered before the development of DFT (Sharp and Horton 1953). In this scheme, the exchange functional is exact since its is expressed in terms of the single-particle orbitals, only the correlation part needs to be approximated. The derivation of the exchange-correlation potential is however somewhat complicated, it is an integral equation hard to solve numerically.

Bibliography

- Abarenkov I. V. and V. Heine (1965), *Phil. Mag.* **12**, 529.
- Alder B. J. and T. E. Wainwright (1957), *J. Chem. Phys.* **27**, 1207.
- Allen M. P. and D. J. Tildesley (1987), *Computer Simulation of Liquids*, (Clarendon Press, Oxford).
- Andersen H. C. (1980), *J. Chem. Phys.* **72**, 2384.
- Ashcroft N. W. and N. D. Mermin (1976), *Solid State Physics*, (Holt Saunders, Philadelphia).
- Bachelet G. B., D. R. Hamman, and M. Schlüter (1982), *Phys. Rev. B* **26**, 4199.
- Bart F. and M. Gautier (1994), *Surf. Sci.* **311**, L671.
- Barth A., R. J. Buenker, S. A. Peyerimhoff, and W. Butscher (1980), *Chem. Phys.* **46**, 149.
- Becke A. D. (1988), *Phys. Rev. A* **38**, 3098.
- Bendt P. and A. Zunger (1983), *Phys. Rev. Lett.* **50**, 1688.
- Bernasconi M. *et al.* (1995), *J. Phys. Chem. Solids* **56**, 345.
- Bhat M. *et al.* (1994), *Appl. Phys. Lett.* **64**, 1168.
- Bhat M., G. W. Yoon, J. Kim, and D. Kwong (1994), *Appl. Phys. Lett.* **64**, 2116.
- Blase X., A. J. R. Silva, X. Zhu, and S. G. Louie (1994), *Phys. Rev. B* **50**, 8102.
- Blöchl P. E. (1994), *Phys. Rev. B* **50**, 17953.

- Blöchl P. E. and M. Parrinello (1992), Phys. Rev. B **45**, 9413.
- Brañka A. and M. Parrinello (1986), Mol. Phys. **58**, 989.
- Car R. and M. Parrinello (1985), Phys. Rev. Lett. **55**, 2471.
- Carr E. C. and R. A. Buhrman (1993), Appl. Phys. Lett. **63**, 54.
- Cederbaum L. S. *et al.* (1978), J. Chem. Phys. **69**, 1591.
- Ceperley D. M. and B. J. Alder (1980), Phys. Rev. Lett. **45**, 566.
- Chadi D. J. and M. L. Cohen (1973), Phys. Rev. B **8**, 5747.
- Chang S. T., N. M. Johnson, and S. A. Lyon (1984), Appl. Phys. Lett. **44**, 316.
- Chen P.-C., K. Y.-J. Hsu, J.-Y. Lin, and H.-L. Hwang (1995), Jpn. J. Appl. Phys. **34**, 973.
- Chester G. (1961), Adv. Phys. **10**, 357.
- Ciccotti G. and W. G. Hoover (1986), *Molecular Dynamics Simulation of Statistical Mechanical Systems*, (Plenum Press, New-York).
- Cleveland C. L. (1988), J. Chem. Phys. **89**, 4987.
- Cohen M. L. and T. K. Bergstresser (1966), Phys. Rev. B **141**, 789.
- Cohen M. L. and V. Heine (1970), Solid State Physics **24**, 37.
- Dal Corso A. and R. Resta (1994), Phys. Rev. B **50**, 4327.
- Dal Corso A., S. Baroni, and R. Resta (1994), Phys. Rev. B **49**, 5323.
- Davidson E. R. (1975), J. Comput. Phys. **17**, 87.
- Daw M. S. (1993), Phys. Rev. B **47**, 10895.
- Donohue J. (1974), *The Structures of the Elements*, (Wiley, New-York).
- Dreizler R. M. and J. da Providencia (1985), *Density Functional Methods in Physics*, (Plenum, New-York).
- Evarestov R. A. and V. P. Smirnov (1983), Phys. Status Solidi **119**, 9.
- Ewald P. P. (1917*a*), Ann. Phys. (Leipzig) **54**, 519.

- Ewald P. P. (1917*b*), Ann. Phys. (Leipzig) **54**, 557.
- Ewald P. P. (1921), Ann. Phys. (Leipzig) **64**, 253.
- Ferrario M. and J. P. Ryckaert (1985), Mol. Phys. **54**, 587.
- Feuston B. P. and S. H. Garofalini (1988), J. Chem. Phys. **89**, 5818.
- Feuston B. P. and S. H. Garofalini (1989), J. Chem. Phys. **91**, 564.
- Feuston B. P. and S. H. Garofalini (1990), J. Phys. Chem. **94**, 5351.
- Filippi C., C. J. Umrigar, and M. Taut (1994), J. Chem. Phys. **100**, 1290.
- Filippi C., D. J. Singh, and C. J. Umrigar (1994), Phys. Rev. B **50**, 14947.
- Focher P. *et al.* (1995), Europhys. Lett. **26**, 501.
- Francis G. P. and M. C. Payne (1990), J. Phys.:Condens. Matter **2**, 4395.
- Froyen S. and M. L. Cohen (1986), J. Phys. C **19**, 2623.
- Garcia A. *et al.* (1992), Phys. Rev. B **46**, 9829.
- Garofalini S. H. (1983), J. Chem. Phys. **78**, 2069.
- Garofalini S. H. (1990), J. Non-Cryst. Solids **120**, 1.
- Gear C. W. (1966), *The numerical integration of ordinary differential equations of various order*, Technical Report ANL 7126, Argonne National Laboratory.
- Gibson J. B., A. N. Goland, M. Milgram, and G. H. Vineyard (1960), Phys. Rev. **120**, 1229.
- Gonze X. (1993-1994), *Condensed matter physics: dielectric and dynamic properties of crystalline solids*. Lecture Notes of the Université Catholique de Louvain.
- Gonze X., P. Kaeckell, and M. Scheffler (1990), Phys. Rev. B **41**, 12264.
- Gonze X., R. Stumpf, and M. Scheffler (1991), Phys. Rev. B **44**, 8503.
- Green M. L. *et al.* (1994), Appl. Phys. Lett. **65**, 849.
- Grey F. and R. Feidenhans'l (1988), Europhysics News **19**, 94.

- Grunthaner P. J., M. H. Hecht, F. J. Grunthaner, and N. M. Johnson (1987), *J. Appl. Phys.* **61**, 629.
- Gunnarsson O. and B. I. Lundqvist (1976), *Phys. Rev. B* **13**, 4274.
- Gunnarsson O., M. Jonson, and B. I. Lundqvist (1976), *Phys. Lett.* **59A**, 177.
- Gunnarsson O., M. Jonson, and B. I. Lundqvist (1977), *Solid State Commun.* **24**, 765.
- Gunnarsson O., M. Jonson, and B. I. Lundqvist (1979), *Phys. Rev. B* **20**, 3136.
- Gunnarsson O. and R. O. Jones (1980), *Phys. Scr.* **21**, 294.
- Hair M. L. (1967), *Infrared Spectroscopy in Surface Chemistry*, (Dekker, New-York).
- Hamann D. R., M. Schlüter, and C. Chiang (1979), *Phys. Rev. Lett.* **43**, 1494.
- Hao M. Y., W. M. Chen, K. Lai, and C. Lee (1995), *Appl. Phys. Lett.* **66**, 1126.
- Harris J. and R. O. Jones (1974), *J. Phys. F* **4**, 1170.
- Haydock R. (1972), *Solid State Physics*, Vol. 35, edited by H. Ehrenreich, F. Seitz, and D. Turnbull, (Academic, New-York), Vol. 35, p. 215.
- Hegde R. I. *et al.* (1995), *Appl. Phys. Lett.* **66**, 2882.
- Heggie M. and R. Jones (1987), *Phil. Mag. Lett.* **55**, 47.
- Hellwege, K.-H. and A. M. Hellwege, eds (1976), *Structure Data of Free Polyatomic Molecules*, Vol. II/7 of *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, (Springer-Verlag, Berlin).
- Hellwege, K.-H. and A. M. Hellwege, eds (1987), *Structure Data of Free Polyatomic Molecules*, Vol. II/15 of *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, (Springer-Verlag, Berlin).
- Helms, C. R. and B. E. Deal, eds (1988), *The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface*, (Plenum Press, New-York).

- Helms, C. R. and B. E. Deal, eds (1993), *The Physics and Chemistry of SiO₂ and the Si-SiO₂ Interface 2*, (Plenum Press, New-York).
- Herman F., J. P. Van Dyke, and I. B. Ortenburger (1969), *Phys. Rev. Lett.* **22**, 807.
- Himpsel F. J., F. R. McFeely, A. Taleb-Ibrahimi, and J. A. Yarmoff (1988), *Phys. Rev. B* **38**, 6084.
- Hohenberg P. and W. Kohn (1964), *Phys. Rev. B* **136**, 864.
- Holl M. M. N. and F. R. Mcfeely (1993), *Phys. Rev. Lett.* **71**, 1993.
- Hoover W. G. (1985), *Phys. Rev. A* pp. 1695–1697.
- Hu C. D. and D. C. Langreth (1986), *Phys. Rev. B* **33**, 943.
- Hussey R. J., T. L. Hoffman, Y. Tao, and M. J. Graham (1996), *J. Electrochem. Soc.* **143**, 221.
- Ihm J. A., A. Zunger, and M. Cohen (1979), *J. Phys. C.* **12**, 4409.
- Iler R. K. (1979), *The Chemistry of Silica*, (Wiley, New-York).
- Janossy I. and M. Menyhard (1974), *Surf. Sci.* **25**, 647.
- Jing Z., G. Lucovsky, and J. L. Whitten (1994), *Phys. Rev. B* **49**, 14003.
- Johnson B. G., P. M. W. Gill, and J. A. Pople (1993), *J. Chem. Phys.* **98**, 5612.
- Jolly W. L., K. D. Bomben, and C. J. Eyermann (1984), *Atomic Data and Nuclear Data Tables* **31**, 433.
- Jones R. O. and O. Gunnarsson (1989), *Rev. Mod. Phys.* **61**, 689.
- Kaluri S. R. and D. W. Hess (1996), *Appl. Phys. Lett.* **69**, 1053.
- Kamath A. *et al.* (1997), *Appl. Phys. Lett.* **70**, 63.
- Kang S. B., S. O. Kim, J.-S. Byun, and H. J. Kim (1994), *Appl. Phys. Lett.* **65**, 2448.
- Kastner M., D. Adler, and H. Fritzsche (1976), *Phys. Rev. Lett.* **37**, 1504.
- Keskar N. R. and J. R. Chelikowsky (1992), *Phys. Rev. B* **46**, 1.

- Khein A., D. J. Singh, and C. J. Umrigar (1995), Phys. Rev. B **51**, 4105.
- Kleinman L. and D. M. Bylander (1982), Phys. Rev. Lett. **48**, 1425.
- Kohn W. and L. J. Sham (1965), Phys. Rev. A **140**, 1133.
- Kryachko E. S. and E. V. Ludena (1990), *Energy Density Functional Theory of Many Electron Systems*, (Kluwer Academic, Boston).
- Kuchitsu, K., ed. (1992), *Structure Data of Free Polyatomic Molecules*, Vol. II/21 of *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, (Springer-Verlag, Berlin).
- Kuchitsu, K., ed. (1995), *Structure Data of Free Polyatomic Molecules*, Vol. II/23 of *Landolt-Börnstein: Numerical Data and Functional Relationships in Science and Technology*, (Springer-Verlag, Berlin).
- Laasonen K. *et al.* (1993), Phys. Rev. B **47**, 10142.
- Laasonen K., M. J. Sprik, M. Parrinello, and R. Car (1993), J. Chem. Phys. **99**, 9080.
- Langreth D. C. and J. P. Perdew (1977), Phys. Rev. B **15**, 2884.
- Langreth D. C. and M. J. Mehl (1981), Phys. Rev. Lett. **47**, 446.
- Langreth D. C. and M. J. Mehl (1983), Phys. Rev. B **28**, 1809.
- Langreth D. C. and M. J. Mehl (1984), Phys. Rev. B **29**, 2310.
- Larsson C. U. S., C. B. M. Andersson, N. P. Prince, and A. S. Flodstöm (1992), Surf. Sci. **271**, 349.
- Lau K. H. and W. Kohn (1976), J. Phys. Chem. Solids **37**, 99.
- Lee C., D. Vanderbilt, R. Car, and M. Parrinello (1992), Phys. Rev. Lett. **69**, 462.
- Levien L., C. T. Prewitt, and D. J. Weidner (1980), Am. Mineral. **65**, 920.
- Levine S. L. and S. H. Garofalini (1987), J. Chem. Phys. **86**, 2997.
- Levy M. and J. P. Perdew (1985), Phys. Rev. A **32**, 2010.
- Li X.-P., R. Nunes, and D. Vanderbilt (1993), Phys. Rev. B **47**, 10891.
- Lill J. V. and J. Q. Broughton (1994), Phys. Rev. B **49**, 11619.

- Louie S. G., S. Froyen, and M. L. Cohen (1982), *Phys. Rev. B* **26**, 1738.
- Lu H. C. *et al.* (1996), *Appl. Phys. Lett.* **69**, 2713.
- Lu Z. H., M. J. Graham, D. T. Jiang, and K. H. Tan (1993), *Appl. Phys. Lett.* **63**, 2941.
- Lu Z. H., S. P. Tay, R. Cao, and P. Pianetta (1995), *Appl. Phys. Lett.* **67**, 2836.
- Lucovsky G. (1979), *Phil. Mag. B* **39**, 513.
- Lucovsky G. (1980), *J. Non-Cryst. Solids* **35 & 36**, 825.
- Makov G. and M. C. Payne (1995), *Phys. Rev. B* **51**, 4014.
- Mauri F., G. Galli, and R. Car (1993), *Phys. Rev. B* **47**, 9973.
- Melchionna S., G. Ciccotti, and B. L. Holian (1993), *Mol. Phys.* **78**, 533.
- Moncrieff D., I. H. Hillier, S. A. Pope, and M. F. Guest (1983), *Chem. Phys.* **82**, 139.
- Monkhorst H. J. and J. D. Pack (1976), *Phys. Rev. B* **13**, 5188.
- Noge S. *et al.* (1997), *Jpn. J. Appl. Phys.* **36**, 3081.
- Nosé S. (1984), *Mol. Phys.* **52**, 255.
- Nosé S. (1986), *Mol. Phys.* **57**, 187.
- Nosé S. (1991), *Prog. Theor. Phys. Suppl.* **103**, 1.
- Nosé S. and M. L. Klein (1983), *Mol. Phys.* **50**, 1055.
- Ordejón P., D. A. Drabold, M. Grunbach, and R. M. Martin (1993), *Phys. Rev. B* **48**, 14646.
- Ortiz G. (1992), *Phys. Rev. B* **45**, 11328.
- Ortiz G. and P. Ballone (1991), *Phys. Rev. B* **43**, 6376.
- Pantelides, S., ed. (1978), *The Physics of SiO₂ and its Interfaces*, (Pergamon, New-York).
- Parrinello M. and A. Rhaman (1980), *Phys. Rev. Lett.* **45**, 1196.

- Parrinello M. and A. Rhaman (1981), *J. Appl. Phys.* **52**, 7182.
- Pasquarello A. *et al.* (1992), *Phys. Rev. Lett.* **69**, 1982.
- Pasquarello A., M. S. Hybertsen, and R. Car (1995), *Phys. Rev. Lett.* **74**, 1024.
- Pasquarello A., M. S. Hybertsen, and R. Car (1996), *Phys. Rev. B* **53**, 10942.
- Peden C. H. F. *et al.* (1993), *Phys. Rev. B* **47**, 15622.
- Pedocchi L., N. Russo, and D. R. Salahub (1993), *Phys. Rev. B* **47**, 12992.
- Pehlke E. and M. Scheffler (1993), *Phys. Rev. Lett.* **71**, 2338.
- Perdew J. P. (1986), *Phys. Rev. B* **33**, 8822.
- Perdew J. P. (1991), *Electronic Structure of Solids '91*, edited by P. Ziesche and H. Eschrig, (Akademie Verlag, Berlin).
- Perdew J. P. and A. Zunger (1981), *Phys. Rev. B* **23**, 5048.
- Perdew J. P., D. C. Langreth, and V. Sahni (1977), *Phys. Rev. Lett.* **38**, 1030.
- Perdew J. P., K. Burke, and M. Ernzerhof (1996), *Phys. Rev. Lett.* **77**, 3865.
- Perdew J. P. and Y. Wang (1986), *Phys. Rev. B* **33**, 8800.
- Perdew J. P. and Y. Wang (1991), unpublished.
- Perdew J. P. and Y. Wang (1992), *Phys. Rev. B* **45**, 13244.
- Philips J. C. and L. Kleinman (1959), *Phys. Rev.* **116**, 287.
- Pulay P. (1969), *Mol. Phys.* **17**, 197.
- Rahman A. (1964), *Phys. Rev.* **136**, A405.
- Rappe A. M., K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos (1990), *Phys. Rev. B* **41**, 1227.
- Rasolt M. and D. J. W. Geldart (1976), *Phys. Rev. B* **13**, 1477.
- Rignanese G.-M. *et al.* (1995), *Phys. Rev. B* **52**, 8160.
- Rohlfing M., P. Krüger, and J. Pollman (1997), *Phys. Rev. B* **56**, 2191.

- Saito Y. (1996), Appl. Phys. Lett. **68**, 800.
- Scandolo S. *et al.* (1995), Phys. Rev. Lett. **74**, 4015.
- Sharp R. T. and G. K. Horton (1953), Phys. Rev. **90**, 317.
- Shirley E. L., D. C. Allan, R. M. Martins, and J. D. Joannopoulos (1989), Phys. Rev. B **40**, 3652.
- Sim F., A. St-Amant, I. Papai, and D. R. Salahub (1992), J. Am. Chem. Soc. **114**, 4391.
- Souza I. and J. L. Martins (1997), Phys. Rev. B **55**, 8733.
- Sprink M. J. (1991), J. Chem. Phys. **95**, 2283.
- Sutherland D. G. J. *et al.* (1995), J. Appl. Phys. **78**, 6761.
- Tang H.-T. *et al.* (1996), J. Appl. Phys. **80**, 1816.
- Tassone F., F. Mauri, and R. Car (1994), Phys. Rev. B **50**, 10561.
- Teter M. P. (1993), Phys. Rev. B **48**, 5031.
- Teter M. P., M. C. Payne, and D. C. Allan (1989), Phys. Rev. B **40**, 12255.
- Ting W., H. Hwang, J. Lee, and D. L. Kwong (1991), J. Appl. Phys. **70**, 1072.
- Troullier N. and J. L. Martins (1991), Phys. Rev. B **43**, 1993.
- Vanderbilt D. (1990), Phys. Rev. B **41**, 7892.
- Verlet L. (1967), Phys. Rev. **159**, 98.
- von Barth U. (1984), *Many-Body Phenomena at Surfaces*, edited by D. Langreth and H. Suhl, (Academic, New-York), p. 3.
- Wentzcovitch R. M. (1991), Phys. Rev. B **44**, 2358.
- White J. A. and D. M. Bird (1994), Phys. Rev. B **50**, 4954.
- Wood D. M. and A. Zunger (1985), J. Phys. A. **18**, 1343.
- Wyckoff R. W. G. (1963), *Crystal Structures*, Vol. 1, second edn, (Wiley, New-York), Vol. 1.

- Yao Z.-Q., H. B. Harrison, S. Dimitrijević, and D. Sweatman (1994), *Appl. Phys. Lett.* **64**, 3854.
- Yin M. T. and M. L. Cohen (1982), *Phys. Rev. B* **26**, 5668.
- Zhong Q., D. Innis, K. Kjoller, and V. B. Ellings (1993), *Surf. Sci.* **290**, L688.