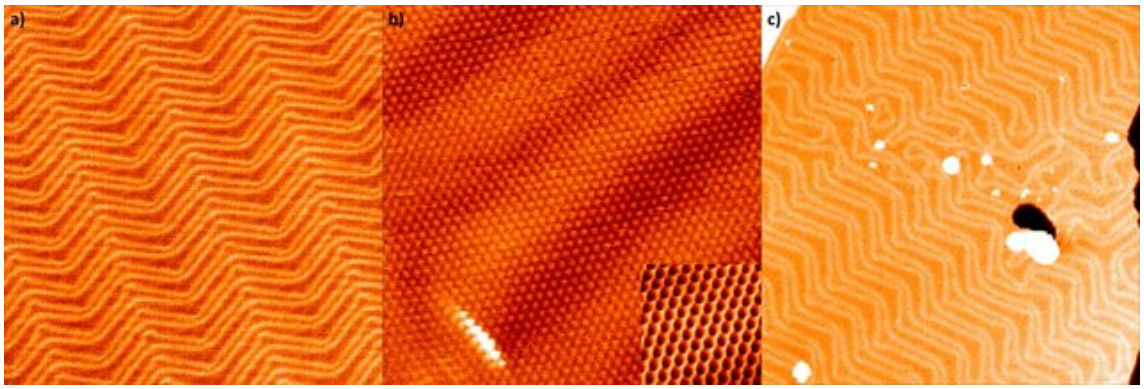
Quantum Phenomena in Low Dimensional Systems

Lecture notes of Giorgio Calandra from *Prof.* Bussetti's course



Politecnico di Milano
A.A. 2023/2024

Introduction to Quantum Phenomena in Low-Dimensional Systems

Welcome to the course on "Quantum Phenomena in Low-Dimensional Systems", a pivotal part of the Master's program in Engineering Physics. This course delves into the fascinating realm of quantum mechanics as applied to systems confined in one or more spatial dimensions, an area rich with unique phenomena and critical applications in modern technology. The lectures are designed for students with a solid foundation in solid-state physics, aiming to deepen their understanding of the quantum behaviours that emerge in low-dimensional environments.

About the Course Materials

The theoretical concepts presented in this course are compiled in the following lecture notes, diligently transcribed and enhanced by Giorgio Calandra, a dedicated student who not only captured the essence of the classroom discussions but also enriched them with additional insights. This compilation bridges the immediacy of live lectures with a structured overview, providing a comprehensive guide to assist you in mastering the subject matter and preparing for the examination.

The Importance of Low-Dimensional Quantum Systems

Low-dimensional systems, such as quantum wells, wires, and dots, exhibit quantum effects that are often obscured in bulk materials. When electrons are confined to two, one, or zero dimensions, their energy levels become quantized, leading to distinct physical properties and behaviours:

- Quantum Wells: These two-dimensional structures constrain electron motion in one dimension, creating discrete energy states. They are crucial in devices like high-efficiency lasers and LEDs.
- Quantum Wires: Restricting electron movement to one dimension, these systems exhibit phenomena like conductance quantization and Luttinger liquid behaviour, essential for nanoscale transistors.
- Quantum Dots: Zero-dimensional systems where electrons are confined in all directions, resulting in atom-like discrete energy levels. They have applications in quantum computing and biological imaging due to their size-tunable optical properties.

Key Concepts Covered

The course covers a range of key concepts essential for understanding low-dimensional quantum systems:

1. Quantum Confinement: The spatial restriction of particles leading to quantized energy states.
2. Energy Quantization: The formation of discrete energy levels within confined systems, analogous to atomic orbitals.
3. Excitons, phonons: The behaviour of quasi-particles in confined systems.

Applications and Technological Impact

Understanding these quantum phenomena is not only academically stimulating but also practically significant. The principles discussed are foundational for many cutting-edge technologies:

- Optoelectronics: Quantum wells enhance the performance of lasers and LEDs.
- Nanoscale Electronics: Quantum wires and dots are integral to developing smaller, faster transistors and memory devices.
- Quantum Computing: Quantum dots and entanglement are essential for creating qubits, the building blocks of quantum computers.

Structure and Objectives of the Course

The course is structured to provide both theoretical knowledge and practical understanding. Through lectures, you will gain:

1. 1. A thorough understanding of the physical principles underlying low-dimensional quantum systems.
2. The ability to analyze and predict the behaviour of particles in confined geometries.
3. Skills to design and evaluate junction properties based on quantum mechanical concepts.

Acknowledgements and Future Updates

I extend my heartfelt thanks to Giorgio Calandra for his exceptional work in compiling and enhancing these lecture notes. His effort ensures that the content is both accurate and accessible, serving as a vital resource for your studies. I encourage you to view this document as a living text, one that can be refined and improved. Please do not hesitate to report any errors or suggestions for improvements, which will be considered for future updates.

I hope you find these materials beneficial in your studies and that they help you gain a deeper appreciation and understanding of quantum phenomena in low-dimensional systems. Your feedback and active participation are highly valued, and I look forward to engaging with you throughout this intellectually stimulating journey.

Prof. Gianlorenzo Bussetti

Contents

1 Crystallography **7**
- 1.1 Surface Roughness . . . 7
- 1.2 Surface Reconstruction and Relaxation . . . 8
  - 1.2.1 Surface Preparation and Ideal Truncated Bulk . . . 9
  - 1.2.2 Reconstruction . . . 10
  - 1.2.3 Relaxation . . . 11
  - 1.2.4 Counterexamples . . . 12
- 1.3 Notations . . . 13
  - 1.3.1 The Wood Notation . . . 13
  - 1.3.2 The Matrix Notation . . . 14
- 1.4 Adsorption . . . 14
  - 1.4.1 Langmuir Adsorption Model . . . 15
  - 1.4.2 Diffusion . . . 16

2 Thermodynamics **18**
- 2.1 Dividing Plane and Basic Concepts . . . 18
  - 2.1.1 Thermodynamic Potentials . . . 19
- 2.2 Surface Tension and Surface Stress . . . 20
  - 2.2.1 Surface Tension . . . 21
  - 2.2.2 Surface Stress . . . 23
- 2.3 γ-Plot and Vicinal Surface . . . 24
  - 2.3.1 Surface Faceting . . . 26
- 2.4 Experimantal Evaluation of the Surface Energy . . . 27
  - 2.4.1 Liquids . . . 27
  - 2.4.2 Solids . . . 28
- 2.5 Wulff Construction . . . 29
  - 2.5.1 Poly-compounds Surfaces . . . 31

3 Electronic Surface States **33**
- 3.1 Electron Density and Friedel Oscillations . . . 33
  - 3.1.1 Surface Dipole . . . 35
- 3.2 Tight Binding Model and Hybridization . . . 37
  - 3.2.1 Hybrid Model . . . 39
  - 3.2.2 Bonds and Bands . . . 41
- 3.3 Surface States . . . 43
  - 3.3.1 Intrisinc Surface States . . . 43
  - 3.3.2 Band Diagrams . . . 45
- 3.4 Tamm States in the Pandey-π Chain Model . . . 47
- 3.5 Extrinsic Surface States . . . 48
  - 3.5.1 Chemisorption . . . 49

3.6 Metal-Metal Junctions . . . 52
3.7 Metal-Semiconductor Junctions . . . 54
3.7.1 Fermi Level Pinning and Bardeen Hypothesis . . . 55
3.7.2 Allen and Gobeli Experiment . . . 56
3.8 Chiarotti Experiment and Optically Induced Transistions . . . 57
3.8.1 3-Layers Model . . . 58
3.9 Metal-Induced Gap States . . . 61

4 Low Dimensional Systems and Collective Phenomena 62
4.1 Excitons . . . 62
4.1.1 Reflectance Anisotropy Spectroscopy . . . 64
4.2 Mott–Hubbard Transistion . . . 66
4.3 Steps and Schwoebel-Ehrlich Barrier . . . 67
4.3.1 Schwoebel-Ehrlich Barrier . . . 68
4.4 1D Structures: Some Examples . . . 70
4.4.1 1D Metal-Semiconductor Junction . . . 70
4.4.2 Donor-Acceptor Molecular Transistor . . . 70
4.5 0D Systems . . . 70
4.5.1 Defects Localization and Tomlinson's Model . . . 71
4.6 Optical Transistions . . . 73
4.6.1 Ciccacci's Experiment . . . 75
4.7 Surface Phonons . . . 76
4.8 Temperature Programmed Desorption . . . 79
4.9 Chemical Reactions on Surfaces . . . 81
4.10 Solid-Liquid Interfaces . . . 82

Introduction

This is a collection of re-written notes from the course "Quantum Phenomena in Low Dimensional Systems" for the theoretical part held by Prof. Bussetti. Despite the fancy name, this course aims to study the physics of surfaces treated as the results of a confinement process of the bulk corresponding material. The course, and in turn these very Lecture Notes, will treat the following topics:

- Crystallography of surfaces.
- Thermodynamics of surfaces.
- Eletronics surface states, bands and levels.
- More confined systems (0D, 1D) and collective phenomena.

The physics of surfaces includes a huge variety of phenomena which still represent an open challenge both from a theoretical and an experimental point of view. This implies that would be quite difficult and mathematically hard to treat precisely and formally all these phenomena. For such a reason, most of the topics will be treated as referring to a particular material, structure or technique. In particular, we will mostly consider as material of interest $Si(111)2x1$, $Al(100)$, $Cu(111)$ and $Fe(001)p(1x1)O$. These Lecture Notes are made for personal use only and do not pretend to be complete or in any way alternative to the reference books. In the following, I will often refer to *Solid Surfaces, Interfaces and Thin Films* by Hans Lüth from which some figures have been copied too.

An Historical Overview

Let's enter the world of surfaces by starting with a historical introduction to get an idea of the most relevant steps which led us to the theories most diffuse nowadays. In an imaginary timeline, we can imagine setting the origin around the beginning of the XIX century with the Galvani-Volta dispute which led to the definition of the Type 1 and Type 2 conductors. For the first time, materials and the interfaces among these started to be investigated. The next step in the timeline is represented by the development of the double-layer model in 1850 by Helmholtz. The double layer is a particular structure that appears on the surface of an object when it is exposed to a fluid. This is another interface phenomenon which drastically depends on the characteristics of the surfaces of the two media. Follows Gibbs in 1900 with the introduction of the concept of free energy which strongly affected the theoretical study of the whole condensed matter physics field. In 1930 the first theory about surface states was developed by Tomm. Then, in 1956 Bardeen won the Nobel for the invention of the transistor, pointing out once again the importance of the study of interfaces. A few years later, proofs of the existence of the surface state theorized by Tomm were found. First, some indirect evidences were found by Allen and Gobeli in 1961 and then, in 1968, direct evidences were found in Italy by Chiarotti. In 1980 Jaher and Forrest proposed to extend the study of surfaces to environments different from vacuum-like liquids and other interfaces, bringing back the attention to the double-layer model. In 1986 a very important step signed Politecnico was performed by Cicacci who showed for the first time clear evidence of the electron-phonon interaction in surfaces. We conclude our timeline with Novoselov and Geim experiments which brought to the isolation and characterization of a monolayer graphene sheet which were worth a Nobel in 2010. These experiments enlarged the perspective for the study of 2D materials.

Chapter 1

Crystallography

The book *Solid Surfaces, Interfaces and Thin Films* by Hans Lüth starts the discussion related to surfaces by saying: "A solid interface is defined as a small number of atomic layers that separate two solids in intimate contact with one another, where the properties differ significantly from those of the bulk material it separates. [...] The surface of a solid is a particularly simple type of interface, at which the solid is in contact with the surrounding world, i.e., the atmosphere or, in the ideal case, the vacuum". Despite this simple definition reported by Lüth, it is not trivial to precisely answer to the question "What is a surface?". Trying to answer such a question will be the leitmotiv of our next discussions which will begin with a crystallographic view of surfaces.

1.1 Surface Roughness

The first question we need to ask ourselves is what a surface looks like. The simplest idea that could come to our minds is that a surface can be approximated by a perfectly flat plane, smooth and free of any defects. As is the case in the bulk, ideal surfaces with complete translational symmetry, i.e. with no defects, cannot exist for entropy reasons. On a real surface defects are always found. To try to understand these entropy reasons let's imagine having a bulk crystal and focusing the attention on a particular atom of this crystal. This atom can be ideally schematized as connected with two chemical bonds to the lower and upper parts of the crystals as shown in Fig.1.1(a). The bonds, as usual, are depicted as two springs with elastic constant k and the atom is supposed to have mass m. If we cut our crystal in half, removing for example the upper part of it, the atom will remain linked to the new "fresh" *surface* just created by only one chemical bond, i.e. only one spring (see Fig.1.1(b)). The Einstein model gives us the entropy of such a system which is:

$$S = nk_B \left[\ln\left(\frac{k_B T}{h\nu}\right) + 1 \right] \tag{1.1}$$

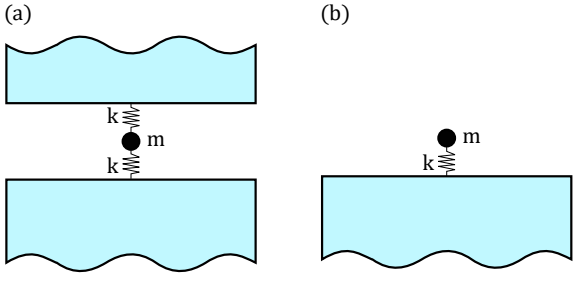


Figure 1.1: (a) Atom in bulk and (b) exposed to vacuum.

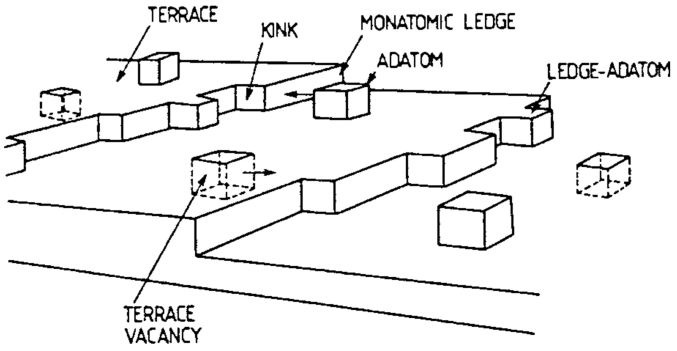


Figure 1.2: Schematic drawing of various defects that may occur on a solid surface.

where n is the number of atoms considered, k_B the Boltzmann constant, T the temperature, h the Plank constant and ν the freqeuncy of the oscillator. This latter quantity is the one which makes the difference between the in-bulk configuration and the exposed one. In fact, in both cases, we have the same mass m for the particle and the same k constant for the spring. What changes is the fact that in the in-bulk case, we have a parallel between two springs and this makes the frequency changes. Denoting with a subscript "s" the exposed configuration and with a subscript "b" the in-bulk one we have:

$$\nu_s = \frac{1}{2\pi}\sqrt{\frac{k}{m}}, \qquad \nu_b = \frac{1}{2\pi}\sqrt{\frac{2k}{m}} = \nu_s\sqrt{2} \tag{1.2}$$

This means that the entropy difference between the two configurations is:

$$\Delta S = S_s - S_b = nk_B \ln\frac{\nu_b}{\nu_s} = nk_B \ln\sqrt{2} > 0 \tag{1.3}$$

This means that the exposed configuration results in an increase in the entropy of the system. In other words, this also means that the creation of a surface is favoured since any phenomenon that can increase the entropy is generally welcome. Moreover, this implies that all kinds of defects on the surface will be welcomed since each of them will increase the overall entropy of the systems. This can be understood in terms of the creation of new surfaces. In fact, by looking at Fig.1.2 we notice that each defect contributes to a small additive piece of surface which as we have just seen implies $\Delta S > 0$. Fig.1.2 reports all possible types of defects that can appear on a surface with the relative name. This reconstruction, justified in terms of entropy, has been confirmed experimentally by Scanning Tunnel Microscopy (STM) which shows that all surfaces have this structure, full of defects and characterized by roughness.

1.2 Surface Reconstruction and Relaxation

We understood that a surface is full of defects and its representation is really different from an ideal two-dimensional (2D) plane. Nevertheless, it is convenient to define the possible 2D lattices which could characterise a surface. This is particularly useful because it allows us to better define the symmetry properties of each structure which also determines physical properties such as conductivity and optical response. The point group operations which are compatible with 2D periodicity are the 1, 2, 3, 4 and 6-fold rotation axes perpendicular to the surface, and mirror planes, also normal to the surface. Inversion centres, mirror planes and rotation axes parallel to the surface are not allowed, since they refer to points outside the surface. By combining the limited number of allowed symmetry operations,

one obtains 10 different point group symmetries denoted as:

$$1,\ 2,\ 1m,\ 2mm,\ 3,\ 3m,\ 4,\ 4mm,\ 6,\ 6mm \tag{1.4}$$

The numeral $n = 1,\ ...,\ 6$ denotes rotations by $2\pi/n$ and the symbol m referrs to reflections in a mirror plane. The third m indicates that a combination of the preceding two operations produces a new mirror plane. The operation of the 2D point groups on a 2D translational net or lattice produces the possible 2D Bravais lattices. In contrast to three dimensions, only 5 symmetrically distinct nets are possible and are shown in Fig.1.3.

1.2.1 Surface Preparation and Ideal Truncated Bulk

If we want to study more in detail how atoms rearrange in the uppermost layers of crystal we first need to "prepare" the surface, i.e. to create the surface with no contaminants on it. To do this, depending on the material to be studied, different techniques can be used:

- **Cleavage**: this technique is typically used for semiconductor materials and consists of two sharp blades which are placed in a specific position on the crystals and pressed one towards the other. Depending on the particular crystalline structure of the material, there will be easier and harder planes to cut along. This depends on the particular position of the interatomic bonds which form the crystal: the smaller the number of bonds along a particular direction, the easier the cut by cleavage.
- **Annealing or Sputtering**: these techniques are typically used for metal materials. Annealing is a heat treatment that alters the physical and sometimes chemical properties of a material to increase its ductility and reduce its hardness, making it more workable. Sputtering, instead, is a phenomenon in which microscopic particles of a solid material are ejected from its surface after the material is itself bombarded by energetic particles of a plasma or gas.

Regardless of the method used, in the end, a new surface is created. When this happens, the atoms on the top layer are deprived of one bond, i.e. they have **dangling bonds** because are left with half-filled orbitals. This makes the atoms on the surfaces highly reactive and requires the need for *vacuum*. For this reason, unless specified, we will always suppose that the surface has been exposed to a vacuum. It can easily be seen that on a surface, due to the absence of neighbouring atoms on one side, the interatomic forces in the uppermost lattice planes are considerably changed. The equilibrium conditions for surface atoms are

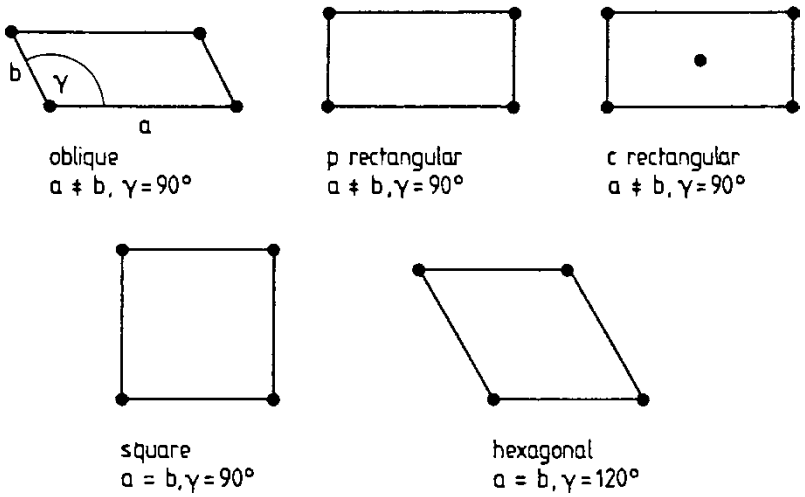


Figure 1.3: Five possible two-dimensional (2D) Bravais lattices.

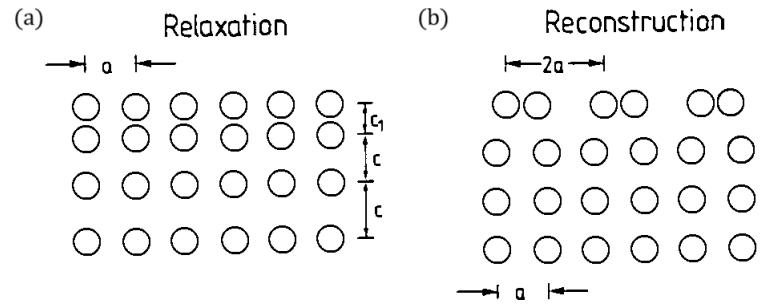


Figure 1.4: (a) Relaxation of the topmost atomic layer normal to surface (different lattice spacings c); (b) reconstruction of the topmost atomic layer into a surface net with double periodicity distance 2a.

modified with respect to the bulk; one, therefore, expects altered atomic positions and a surface atomic structure that usually does not agree with that of the bulk. Thus a surface is not merely a truncation of the bulk of a crystal. However, since we need a reference structure to confront the sample with, we will use this merely truncated configuration as a reference structure. This is the so-called **Ideal Trunceted Bulk** which is an ideal surface which has the atoms in bulk-like positions. The fact that the atoms of a surface rearrange themselves, breaking the symmetry of the bulk, to find a new equilibrium position, can be understood thanks to the Jahn-Teller Principle. This states that it is always possible to find a lower symmetry configuration that corresponds to a lower energy state for the system. Through deformations, a generic system, e.g. a crystal or a molecule, can then minimize the energy and find a more stable configuration.
The distortion of the ideal bulk-like atom configuration due to the existence of a surface will be different for metals and for semiconductors. In the former, a less painful reorganisation process is expected because of the presence of strongly delocalized electron gas; this phenomenon is called relaxation. In the latter, instead, because of the high directionality of the bonding, we expect a more dramatic and drastic reorganization called reconstruction. Let's analyse these two phenomena. Fig.1.4 illustrates schematically some characteristic rearrangements of surface atoms. Regardless of the mechanism adopted by the surface, the new lattice which follows the reconstruction or the relaxation process is called a **superlattice** or a superstructure.

1.2.2 Reconstruction

In tetrahedrally bonded semiconductors (Si, Ge, GaAs, InP, etc.) significant directional bonding is present. Bond breaking on one side due to the surface is expected to have a more dramatic effect on the atomic configuration at the surface of a semiconductor. Usually, this results in shifts parallel to the surface, which can change the periodicity and thus the symmetry of the bulk crystal, i.e. the 2D unit mesh has dimensions different from those of a projected bulk unit cell. This is a crucial point since as we know, materials' properties are also determined by the symmetry they possess and a change in symmetry corresponds to a change in physical properties. As an example of the reconstruction, we can consider the case of Si(111)(2x1)[1] which is shown in Fig.1.5. On the left, the non-reconstructed surface, i.e. the ideal truncated bulk, is reported and it is compared with the reconstructed configuration of the right which in this case takes the name of **Pandey's π-bonded chain model**. In the figure, smaller circles indicate deeper atomic layers. Despite both of them being significant, the side view makes us better realize how the surface changed with respect

[1]We will see in the following the full details of this notation, for now, let's take it as a "random" name.

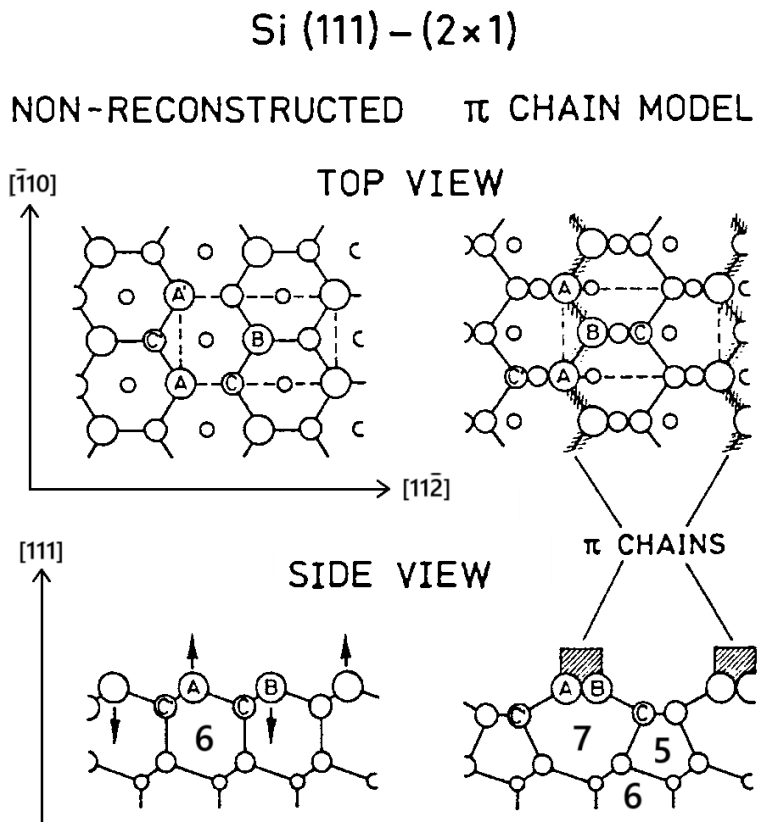


Figure 1.5: Atomic positions at the Si(111) surface; non-reconstructed (Ideal Truncated Bulk) and with reconstruction. In the top view, the unit mesh is plotted in a dashed line.

to the ideal case. We notice, in fact, that after the surface is completely reconstructed, we passed from having in the ideal case equal rings of 6 atoms each to having different rings of 5 and 7 atoms. Moreover, we notice that the reconstruction does not involve only the uppermost layer of the surface but also the adjacent ones: atoms of the next deeper atomic layers also have to change their positions and participate in the reconstruction. The Si atoms on the surface are rearranged in such a way that neighbouring dangling bonds can form zig-zag chains of π-bonds along the $[\bar{1}10]$, similar to a long-chain organic molecule. In this picture, the material can be well approximated by a sort of 3-layer model in which the layers are represented by the vacuum, the surface which is made by the first atomic layers and constitutes a new material in itself because of the change of symmetry, and the bulk. Lastly, in the side view in Fig.1.5 the arrows indicate possible up and downwards shifts of surface atoms of type A and B, which give rise to the so-called *buckling* reconstruction. This model of reconstruction was the one originally adopted which has been substituted later on with Pandey's π-bonded chain model.

1.2.3 Relaxation

In metals, we have a strongly delocalized electron gas and a chemical bond which is essentially non-directional, different from what we have in semiconductors. They undergo a phenomenon called relaxation which is a pure compression (or extension) of the topmost interlayer separations normal to the surface. In this case, the 2D lattice, i.e. the periodicity parallel to the surface within the topmost atomic layer, is the same as for the bulk. Fig.1.6 exhibits schematically the rigid ion core positions near a metal surface in which the free electrons are delocalized between the ion cores making an electrically neutral object. To ensure such a neutrality one can formally attribute to each positive ion a Wigner-Seitz cell

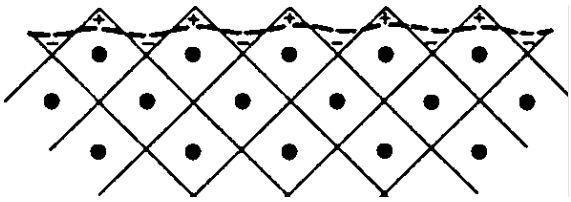


Figure 1.6: Schematic representation of the formation of electronic surface dipoles at metal surfaces.

(squares in Fig.1.6), which contains the corresponding negative electronic charge. On a surface this would lead to a rapidly varying electron density at the surface, thus increasing the kinetic energy of the electrons. Taking as a principle the concept that the system does not want to create any electronic roughness, i.e. the electron distribution has to be uniform, this implies a migration of electrons. In a very simple picture of the phenomenon, we can imagine the electron migrating from the top of the squares in Fig.1.6 to the depression in between the different Wigner-Seitz cells. Due to this migration, the positive ion cores in the topmost atomic layer feel a net repulsion from the charge in their Wigner Seitz cell, and an inwards displacement results. This effect is thought to be one possible source of the contraction (relaxation) of the topmost lattice plane observed on many metal surfaces. A very good example of relaxation is represented by Al(110) in which the interlayer distances d_{ij} between the i-th and the j-th level are:

$$d_{12} = 1.3 \text{ Å}, \qquad d_{23} = 1.5 \text{ Å}, \qquad d_{34} = 1.4 \text{ Å}$$

This has to be compared with the distance in the ideal truncated bulk $d_{ideal} = 1.43$ Å. We notice how at the surface (layers $1/2$) we observe a contraction which is followed by an increase in d (layer $2/3$)[2]. Starting from the $3/4$ layers the distance starts approaching the ideal one, i.e. the bulk one.

1.2.4 Counterexamples

In summary, we saw that the two ways in which a surface reorganizes its structure are reconstruction for semiconductors and relaxation for metals. However, this is not always the case. There are exceptions in which a metal undergoes reconstruction and a semiconductor undergoes relaxation; for example:

- Pt(100): platinum is the case of a metal which reorganizes itself through reconstruction. In particular, we observe a drastic change in symmetry because the basis of the lattice passes from a rectangular to a hexagonal one. The new unitary cell of the reconstructed structure is 5 times bigger than the ideal one, thus meaning that the periodicity passes from a to $5a$ for the superlattice.

- GaAs(110): gallium arsenide represents the case of a semiconductor which undergoes relaxation. In Fig.1.7 we notice that the top view of the reconstructed surface is almost identical to the one of the ideal truncated bulk or, at least, it possesses the same symmetry. To better appreciate the changes we need to look at the side view in which a displacement of the As atom is well visible. This displacement is caused by a difference in electronegativity of the As and Ga atoms which makes the most electronegative one (As) tend outward.

[2]This increase is not easy to explain and it is not simply given by relaxation phenomena

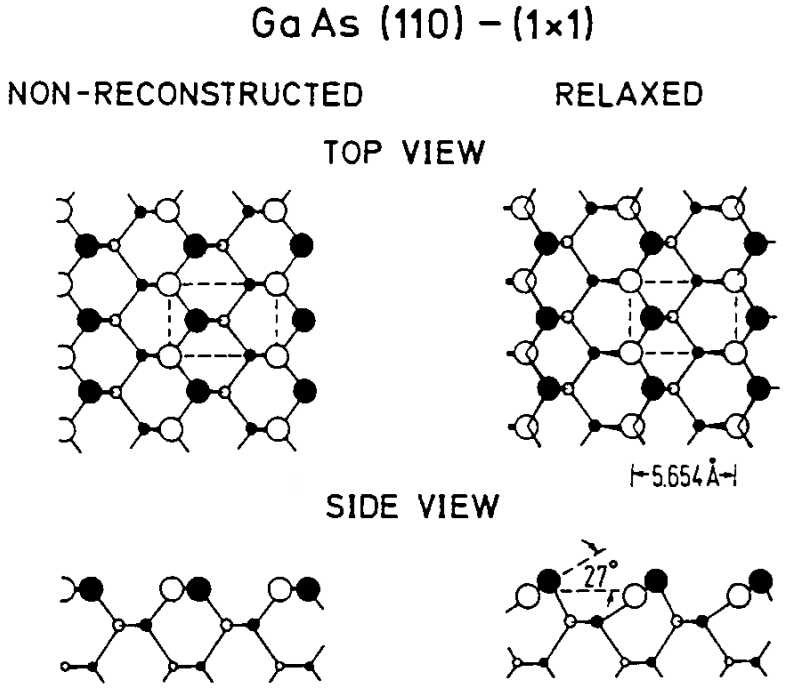


Figure 1.7: Top and side views of the ideal and reconstructed GaAs(110), in black As atoms and in white Ga ones.

1.3 Notations

So far we have seen that the 2D confinement of a surface causes a change of symmetry and in turn of the physical properties of the few uppermost layers characterizing the surface. The changes in the symmetry properties of the superstructure are analysed with respect to the ideal truncated bulk case. To denote these changes more formally and compactly different types of notations can be used.

1.3.1 The Wood Notation

The Wood notation is probably the most famous way of characterising the changes in the symmetry of a surface. Given the unit vectors $\bar{a}_i$ of the ideal truncated bulk and $\bar{\bar{a}}_i$ of the superlattice with $i = 1, 2$, the Wood notation consists in:

$$S(hkl)\,\eta\,\left(\left|\frac{\bar{\bar{a}}_1}{\bar{a}_1}\right| \times \left|\frac{\bar{\bar{a}}_2}{\bar{a}_2}\right|\right) E_l\, R\varphi \tag{1.5}$$

where:

- S represent the main element (Si, Fe, Pt, etc.).
- (hkl) denotes the crystallographic plane.
- η can be either p (principal) or c (centered). The first (p) applies when the lattice geometry remains the same, and the second (c) if a centring occurs.
- $\left(\left|\frac{\bar{\bar{a}}_1}{\bar{a}_1}\right| \times \left|\frac{\bar{\bar{a}}_2}{\bar{a}_2}\right|\right)$ is the ratio between the unit vectors of the reconstructed surface and the ideal truncated bulk.
- E_l is used to denote the presence of another element such as the presence of Oxygen (O) as a consequence of a passivation[3] process.

[3]Passivation is a process which consists in the coating of a surface so that it becomes "passive", that is, less readily affected or corroded by the environment.

- $R\varphi$ is used to indicate if the unit vectors of the superstructure have been rotated by an angle φ with respect to the ideal truncated bulk ones.

For example, Fe(001)p(1x1)O denotes the surface of iron, characterised by the plane (001) in which the unit vectors are the same as the ideal truncated bulk. The presence of the "O" tells us that the surface has been exposed and thus passivated with oxygen. We can also consider platinum, cited in Section 1.2.4. We said that the unit cell of the superstructure takes 5 times the dimension of the ideal one and that the geometry changes from rectangular to hexagonal; this means that in our notation we will have Pt(100)c(1x5). The letter "c" for historical reasons is often omitted. Lastly, also the case of Si(111) is particularly interesting. It is possible to see that the reconstruction processes in silicon result in a change of unit vectors which we can denote as Si(111)(2x1). Also in this case we omitted the letter "p" because it is not usually reported. All these observations result from STM experiments or diffraction analysis. For example, for Si(111)(2x1) the doubling in the unit vectors results in a more dense reciprocal lattice and thus in a possible diffraction experiment an increased number of light diffraction spots. STM is a fundamental technique for the investigation of surfaces but it must be used carefully since it can provide misleading results if not used properly. That is the case, for example, in the study of Highly Oriented Pyrolytic Graphite (HOPG) in which the presence of a multilayered structure of graphene planes, which results in a superposition of more atoms belonging to adjacent planes, can mislead the STM providing a wrong interatomic distance.

1.3.2 The Matrix Notation

Another notation that is used, especially for really complicated structures like the ones which appear in molecular or organic crystals, is the matrix notation. The idea is to represent the unit vectors of the superlattice as a linear combination of the ideal basis vectors, which means:

$$\bar{\bar{a}}_i = \sum_j m_{ij} \bar{a}_j \qquad \text{with } i, j = 1, 2 \tag{1.6}$$

This implies that we can express the change in symmetry by using a matrix M defined as:

$$M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \tag{1.7}$$

This notation is not particularly used and in the following, it will not be used any longer.

1.4 Adsorption

Another phenomenon that can drastically change a surface's symmetry properties is adsorption. Adsorption is the adhesion of atoms, ions or molecules from a gas, liquid or

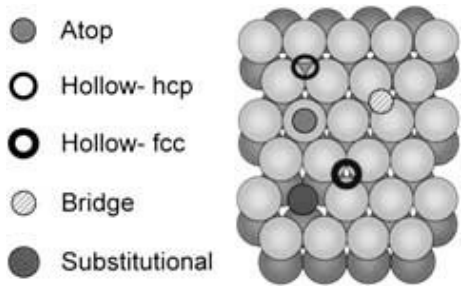


Figure 1.8: Schematic of different adsorption sites on a Cu(111) surface: atop, hollow-hcp and fcc, bridge and substitutional.

dissolved solid to a surface. This process creates a film of the *adsorbate* on the surface of the adsorbent. As reported in Fig.1.8 there are several types of adsorption. We talk about atop adsorption when the adsorbing site is the top of one of the atoms of the surface, bridge when the adsorption occurs in between two atoms, 3 fold hollow site when is between three atoms and so on. Adsorption depends on a lot of factors such as material, temperature, pressure and others and for this reason a full description would be complicated to develop. One of the most successful and simple models for adsorption is the model developed by Langmuir which is the so-called **Langmuir Adsorption Model** or Langmuir Isotherm.

1.4.1 Langmuir Adsorption Model

The Langmuir adsorption model explains adsorption by assuming that the adsorbate, i.e. the particles that stick to the surface, behaves as an ideal gas at isothermal conditions. Within this model, the surface in which adsorption occurs is considered as a perfectly flat plane with no corrugations which can be simply described by two quantities: the number of active site N, i.e. the number of atoms at which the adsorbate can stick to, and the number of already occupied site N_s, i.e. the number of active sites already occupied with an adsorbate particle. As a first consideration, we can define the coverage Θ as:

$$\Theta = \frac{N_s}{N} \tag{1.8}$$

which is a number in between zero and one: $0 \leq \Theta \leq 1$. Within this model, we can imagine that when the particle of adsorbate, for example, a molecule, arrives at the surface if it finds an empty active site either it sticks to the surface or it is immediately scattered away. In the model, the following hypotheses are adopted:

- A molecule which arrives at the surface and finds an occupied site can not move along the surface looking for an empty site. Such a phenomenon would be called diffusion and it is neglected.
- Each site can hold at most one molecule which means that there is no possibility of building a multilayered structure.
- The first and the last molecules forming a monolayer are energetically identical, i.e. they have the same sticking probability.

Within this hypothesis, the rate of adsorption r_a only depends on the pressure p and the coverage state $(1 - \Theta)$ with a proportionality relation:

$$r_a = k_a p(1 - \Theta) \tag{1.9}$$

The proportionality constant k_a is a parameter which only depends on the material properties. We need also to consider the possibility of a stuck particle being desorbed. The rate of desorption r_d can be defined similarly:

$$r_d = k_d \Theta \tag{1.10}$$

By imposing that at equilibrium is must hold $r_a = r_d$ and solving for Θ one finds:

$$k_a p(1 - \Theta) = k_d \Theta \quad \Longrightarrow \quad \Theta = \frac{Kp}{1 + Kp} \tag{1.11}$$

where $K = k_a/k_d$. This equation is the so-called **Langmuir Isotherm**. Eq.(1.11) predicts that in a vacuum, i.e. at low pressure, $\Theta \sim Kp$. The Langmuir model represents a milestone in the physics of surface because it allows one to quickly recover the required

pressure to obtain a certain coverage state. Nevertheless, the model, as we have seen, is based on several strong hypotheses which drastically limit the validity and the accuracy of the model. One of the first improvements we apport to the model is considering diffusion. This means that the adsorbate particles once arrived at the surface, can move on the surface and look for an empty site to stick to. This process requires an intermediary state, called **precursor state**, which is in between the arrival of the molecule and its sticking. The model to describe precursor states is a probabilistic model based on the following quantities:

- p_a: the probability that a particle once arrived diffuses on the surface rather than sticking to it (regardless of the state of the active site, i.e. empty/occupied)
- p_d: the probability that a particle that has already stuck and has broken its bond with the surface diffuses on it rather than going away.
- p'_d: the probability that a particle which has found an occupied state diffuses on the surface rather than going away.

It is possible to demonstrate that this results in a new equation:

$$\frac{k_a p(1-\Theta)}{1+\Theta(\mathcal{K}-1)} = k_d\Theta \qquad \text{with } \mathcal{K} = \frac{p'_d}{p_a + p_d} \tag{1.12}$$

by defining $b = k_a/k_d$ one finds[4]:

$$(\mathcal{K}-1)\Theta^2 + (1+bp)\Theta - bp = 0 \tag{1.13}$$

It is worth noting that by setting $\mathcal{K} = 1$ we go back to the Langmuir Isotherm; the two formulations are therefore compatible. Moreover, we notice that the coverage state Θ now follows a quadratic law meaning that it is increased with respect to Eq.(1.11). This is a direct consequence that by including also the diffusion mechanism we increased the possibility for a particle to stick to the surface and the creation of a monolayer is more probable.

1.4.2 Diffusion

In the previous section, we considered the possibility for a molecule to diffuse on the surface during the adsorption phenomenon. However, we have not properly defined the diffusion phenomenon. When we talk about diffusion for an adsorbate particle on an *isotropic* surface, we have to imagine the particle moving on the surface by "jumping" on it. In this vision, the particle is like a crazy frog jumping randomly on the surface of a pond. The jumping of the frog can be characterized by two parameters: the number of jumps per second, i.e. the rate of jumping, and the mean square distance of each jump, i.e. the displacement. Moreover, if the frog has gone completely crazy, we can imagine that while jumping and jumping randomly across the pond, at a certain time the frog will go back to its starting point travelling an overall zero distance. We can keep this into account by using a new parameter which will be inversely proportional to the total number of jumps taken in a certain time interval; this is simply the product between the time interval itself and the jump rate. We expect that being the jumps completely random, the more the total number of jumps, the smaller the distance travelled from the origin because somehow the frog will jump back to its starting position. We now let the frog go, and we go back

[4]What we now call b was previously defined as K. This re-definition has been done to avoid confusion between the different "k" letters.

to our adsorbate particle. We define the jump rate h, called **hopping rate**, through an Arrhenius law:

$$h = h_0 e^{-\frac{E_d}{k_B T}} \tag{1.14}$$

where h_0 is a certain constant measured as s^{-1}, and E_d is the activation energy which is the energy required by the molecule to break its bond with the surface and do a jump. Such energy is thermally activated as shown by the $k_B T$ term. We denote the mean square distance at each jump as $\langle x^2 \rangle$ and we call it displacement. This is useful to define the so-called **jump** λ^2 which represents the overall distance from the starting point travelled by the particle in a time interval Δt:

$$\lambda^2 = \frac{1}{h\Delta t} \langle x^2 \rangle \tag{1.15}$$

It is worth noting that, because of the isotropy of the surface, λ^2 can be equal to zero if one chooses a sufficiently large time interval. In the previous example, the isotropy of the surface was the reason why the frog was considered to jump *randomly* on the pond.

Chapter 2

Thermodynamics

The most general macroscopic approach to a problem in the physics of matter is that of thermodynamics. Indeed, there are thermodynamic rules which govern the macroscopic shape and therefore the formation of a particular type of surface or interface of a solid to an adjacent medium or the vacuum. Basic concepts in this context are those of the *free specific energy* of a surface and that of *surface tension*. These concepts were developed originally by Gibbs in his theoretical treatment of a gas/solid (non-crystalline) interface. In the following, we will follow Gibbs's path in the study of the thermodynamics of surfaces.

2.1 Dividing Plane and Basic Concepts

The first important thing to do if we want to deal with the thermodynamic study of surfaces is to define what a surface is. As we understood, this is not an easy deal. The idea is to find a certain quantity, a sort of order parameter, which characterizes the presence of a surface and, therefore, which is drastically different in the bulk of the material and at its surface. Let's consider the density of the atoms $n(\boldsymbol{r})$ going from the bulk towards the vacuum as reported in Fig.2.1(a). We can imagine that the atoms' density is equal to a certain equilibrium value n_0 in the bulk whereas decreases until reaching zero while going towards the surface (see Fig.2.1(b)). Modelling a surface with this parameter would be very difficult, both for the spatial dependence of $n(\boldsymbol{r})$ and the absence of a net distinction between bulk and surface. In other words, this means that there is no *physical plane* characterizing the surface. By physical plane here we mean a geometrical plane, i.e. a flat surface, with a certain physical significance like, for example, an infinite charged plane. Gibbs's idea for finding such a plane is quite simple and effective. We imagine the atoms' density to be equal to the bulk value n_0 up to a certain z coordinate and zero afterwards as shown in Fig.2.1(c). The geometrical plane dividing the two regions is called **dividing plane** and its position can be arbitrarily chosen. Clearly, by setting a certain position for the plane we will underestimate a portion of the atoms' density and overestimate another

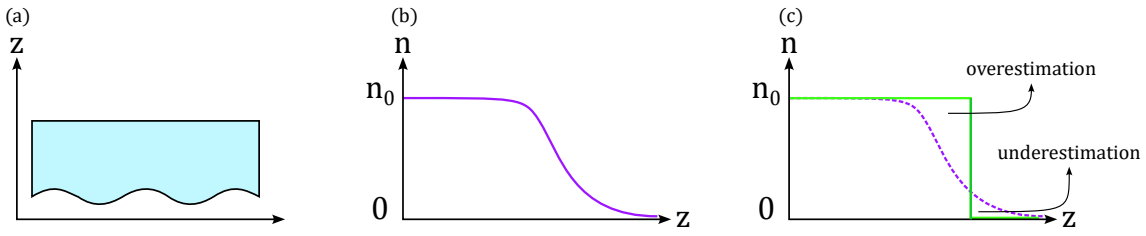


Figure 2.1: (a) Bulk material and z-axis direction. (b) Atoms density as a function of z. (c) Atoms density and dividing plane.

one (see Fig.2.1(c)). When possible, it is wise to set the position of the dividing plane in such a way that the underestimated and the overestimated parts compensate each other. The physical meaning of the plane is given by assigning to it a certain quantity, called **surface excess** Γ, defined as the ratio between the number of elements characterizing the system N (e.g. the number of atoms on the surface) and the area of the surface A:

$$\Gamma = \frac{N}{A} \tag{2.1}$$

We used here a general definition of surface excess because the concept of the dividing plane is used in many cases in the physics of surfaces and depending on the case it can assume different forms. Overall, the dividing plane is a theoretical construct used to delineate between different regions in the analysis of surface properties or phenomena, providing a useful reference point for interpreting experimental data and theoretical models.

2.1.1 Thermodynamic Potentials

To properly study the thermodynamics of a surface we need to choose an appropriate thermodynamic potential to describe it. The queen of thermodynamic potentials is the **internal energy** U. The most general definition of internal energy is given through its differential expression dU and is:

$$dU = TdS + dW + \sum_i \mu_i dN_i \tag{2.2}$$

where T is the temperature, S is the entropy, dW is the work, N_i is the number of particles of type i in the system and μ_i is the chemical potential for an i-type particle[1]. Here we notice:

- $U = U(S, W, N_i)$ which also means that the internal energy represents the capacity to do work at constants S and N_i.
- Each energy term in dU is the product between an extensive and an intensive variable[2].
- In general terms, the internal energy is the overall energy required to create a certain system configuration.

The expressions for all other thermodynamic energy potentials are derivable via Legendre transforms starting from the expression for U of Eq.(2.2). Here we will only refer to and use two of them: the Helmholtz free energy F and the Landau grand potential Φ. The Helmholtz free energy is defined as:

$$F = U - TS \tag{2.3}$$

Its differential expression is:

$$dF = dU - TdS - SdT = -sdT + dW + \sum_i \mu_i dN_i \tag{2.4}$$

Hence, we can write $F = F(T, W, N_i)$: it is the capacity to do work at constant T and N_i. The Landau grand potential is defined as:

$$\Phi = U - TS - \sum_i \mu_i N_i = F - \sum_i \mu_i N_i \tag{2.5}$$

[1] The subscript "i" is here used to keep into consideration the cases in which the system is made by more than one type of particles.

[2] An intensive variable is a physical quantity whose magnitude is independent of the extent of the system (e.g. pressure, temperature, chemical potential, etc.). An extensive variable is a physical quantity which depends on the system's size (e.g. volume, area, etc.).

which result to be $\Phi = \Phi(T, W, \mu_i)$ thus the capacity to do work at constant T and μ_i. When it comes to the study of the physics of surfaces these two thermodynamic potentials become the natural choice for an appropriate description. The potentials we have just seen can be slightly modified to be adapted to the surface case:

$$dU^s = TdS^s + dW^s + \sum_i \mu_i dN_i^s \tag{2.6}$$

In this expression, the apex "s" denotes the surface and dW^s is the work to create or remove a certain portion of a given surface. In analogy with the thermodynamic of a perfect gas in which $dW = pdV$ and remembering what told about the product between extensive and intensive variables, we can write:

$$dW^s = \gamma dA \tag{2.7}$$

where γ is the **free surface energy** (intensive variable) and A is the area (extensive variable). It is worth noting that, if the free surface energy changes the most general expression is: $dW^s = \gamma dA + Ad\gamma$. Moreover, using the definition of surface excess introduced by Gibbs we get:

$$\Gamma = \frac{N}{A} \Rightarrow dN = \Gamma dA \Rightarrow \sum_i \mu_i dN_i^s = \sum_i \mu_i \Gamma_i dA \tag{2.8}$$

So that the general expression becomes:

$$\begin{aligned} dU^s &= TdS^s + \gamma dA + \sum_i \mu_i \Gamma_i dA = \\ &= TdS^s + \underbrace{(\gamma + \sum_i \mu_i \Gamma_i)}_{\equiv \gamma'} dA = TdS^s + \gamma' dA \end{aligned} \tag{2.9}$$

where we defined the specific surface energy γ'. What done so far to adapt the internal energy to the surface study, can be done also for the other two thermodynamic potentials. For the Helmholtz free energy we get:

$$\begin{aligned} dF^s &= -S^s dT + dW^s + \sum_i \mu_i dN_i^s = -S^s dT + \gamma dA + \sum_i \mu_i \Gamma_i dA = \\ &= -S^s dT + \gamma' dA \end{aligned} \tag{2.10}$$

Instead, for the Landau potential:

$$\begin{aligned} d\Phi^s &= dU^s - TdS^s - \sum_i \mu_i dN_i^s = TdS^s + dW^s + \sum_i \mu_i dN_i^s - TdS^s - \sum_i \mu_i dN_i^s = \\ &= dW^s = d(\gamma A) = \gamma dA + Ad\gamma \end{aligned} \tag{2.11}$$

Where we considered the most general case in which also the surface free energy can vary. Let's spend some more words on the physical origin of this surface energy γ. If we consider a crystalline structure with its surface next to the vacuum, it costs energy to generate an additional piece of the surface while keeping the crystal volume and the number of constituent atoms constant. Bonds between neighbouring atoms must be broken in order to expose new atoms to the vacuum. The formation of surface defects including steps might also be involved in forming the new surface area. All these effects contribute to the excess surface free energy γ.

2.2 Surface Tension and Surface Stress

Surface tension and surface stress are related concepts that describe the behaviour of surfaces and interfaces of materials. While they are similar in some aspects, they have

distinct definitions and applications. In particular, **surface tension** is the property of the surface of *liquids* that causes it to behave like a stretched elastic membrane. Surface tension arises due to the cohesive forces between molecules in the liquid, which tend to minimize the surface area to achieve a state of lower energy. This cohesive force pulls the liquid molecules inward, resulting in the formation of a thin, strong "skin" at the surface. Surface tension is responsible for phenomena such as the formation of droplets, capillary action, and the shape of liquid surfaces in contact with solids (e.g. meniscus in a graduated cylinder). On the other hand, **surface stress**, also known as surface traction or surface force, is a measure of the total force acting on a unit area of the surface of a material. It represents the internal mechanical forces present within the material at its surface. Surface stress can arise due to various factors, including the presence of surface defects, surface treatments, or external loads applied to the material. It can lead to deformation or changes in the shape of the material's surface. While surface tension and surface stress both involve forces acting at the surface of materials, surface tension specifically refers to the cohesive forces within a liquid, whereas surface stress encompasses mechanical forces present at material interfaces. It is indeed important to keep in mind that surface tension and surface stress are two faces of the same coin. Let's now study them more in detail.

2.2.1 Surface Tension

For the quantitative study of surface tension, we will refer to a particular model called **Langmiur bath**. This model was developed by Langmiur and it consists of a container filled with any generic liquid (e.g. water) kept at a constant temperature. On the surface of the liquid, a floating barrier is placed to divide the surface into two parts (see Fig.2.2). We call x the distance from the side of the container to the barrier and y the width of the container. In this way, the area A of the surface under investigation is simply given by $A = xy$. We want to calculate the energy cost of creating a new piece of surface in this system. To do that we consider moving the barrier by an infinitesimally small quantity dx. In this case, it is particularly convenient to use the Landau grand potential Φ^s whose variation is:

$$d\Phi^s = d(\gamma A) = Ad\gamma + \gamma dA \tag{2.12}$$

However, a liquid has at its disposal a *reservoir* of particles which from the "bulk" go up to the surface when a new piece of it is created. This is a so-called *elastic* process. For this reason, even though a new surface has been created, each particle sees the same environment around itself because of the arrival of new particles which compose the new surface. In more technical words, this means that the surface energy γ does not change while creating the surface of a liquid. We already understand that such a consideration can not be done in the solids case. Moreover, for our system we have that $dA = ydx$ being y constant. So that:

$$d\Phi^s = \gamma dA = \gamma y dx \tag{2.13}$$

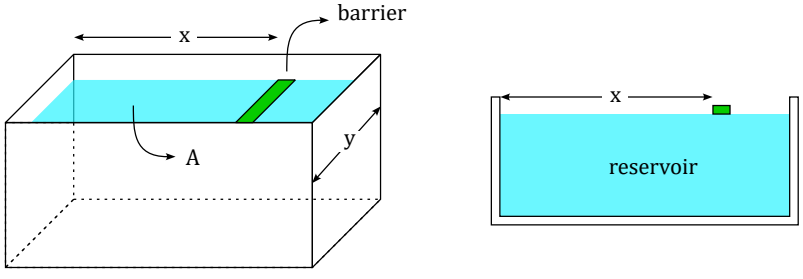


Figure 2.2: Schematic of the Langmuir bath both in perspective and side views.

This leads us to the definition of surface tension which directly comes from γ:

$$\gamma = \frac{1}{y}\frac{d\Phi^s}{dx} \tag{2.14}$$

It is the force per unit length acting perpendicular to the surface of the liquid along the interface between the liquid and the surrounding medium (in our case air). In fact, we notice that $d\Phi^s/dx$ is the derivative of an energy with respect to a length which results in a force which then is divided by y giving a force per unit length.
It is worth noting that Langmuir baths are widely used in surface science and physical chemistry research to study phenomena such as monolayer formation, surface tension, surface pressure, film deposition, adsorption-desorption kinetics, and interactions between molecules at interfaces. They are valuable tools for investigating the properties of surfactants, lipid membranes, biomolecules, nanoparticles, and other materials at liquid interfaces, providing insights into their behaviour and applications in various fields.

Growth Methods

The concept of surface tension can be used in many different scenarios. An example is the modelling of the various growth modes of thin films. When a particular film has to be grown, the molecules of the material to be grown are deposited on a substrate. In general, three markedly different modes of film growth can be distinguished (see Fig.2.3) depending on how the molecules rearrange on the surface. We will not deal in detail with these growth techniques. For those interested, I report here what is reported in the Hans Lüth book. In the layer-by-layer growth mode (or Frank-van der Merve) the interaction between substrate and layer atoms is stronger than that between neighbouring layer atoms. Each new layer starts to grow only when the last one has been completed. The opposite case, in which the interaction between neighbouring film atoms exceeds the overlayer substrate interaction, leads to island growth (or Vollmer-Weber). In this case, an island deposit always means a multilayer conglomerate of adsorbed atoms. The layer-plus-island growth mode (or Stransky-Krastanov) is an interesting intermediate case. After the formation of one, or sometimes several complete monolayers, island formation occurs; 3D islands grow on top of the first full layer(s). Let's now consider the formation of an island as depicted in Fig.2.4. The island formation is comparable to the deposition of a droplet of liquid on a particular substrate. We define γ_{LA} the surface energy at the island-atmosphere interface, γ_{LS} the surface energy at the island-substrate interface and γ_{SA} the surface energy at the

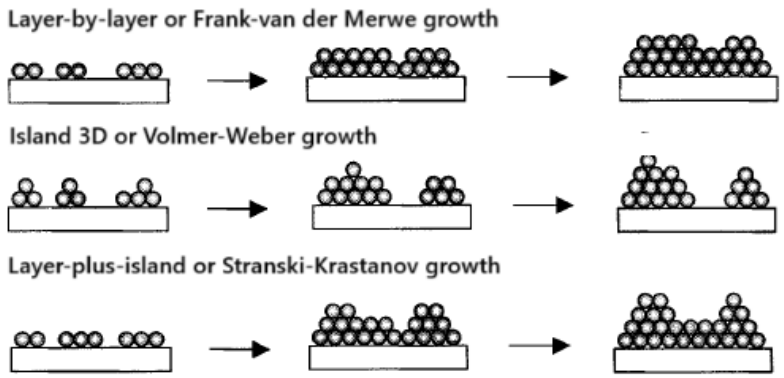


Figure 2.3: Schematic representation of the three important growth modes of a film: Layer-by-layer growth (Frank-van der Merve), Island growth (Vollmer-Weber) and Layer-plus-island growth (Stranski-Krastanov).

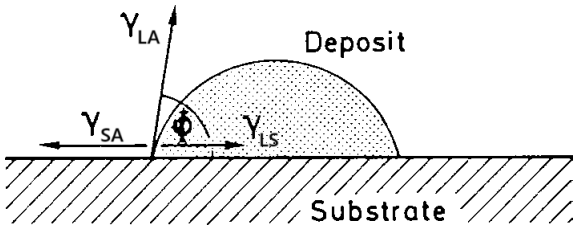


Figure 2.4: Simplified picture of an island of a deposited film.

substrate-atmosphere interface. Here we used the subscript "L" because this description is generalizable to any arbitrary liquid. Since, as we have just seen, γ can also be interpreted as a force per unit length, for the island (or the droplet) to be stable we need to impose forces equilibrium at the contact point where the substrate, the 3D island and the external environment, i.e. the atmosphere, touch each other. This implies:

$$\gamma_{SA} = \gamma_{LA} \cos\phi + \gamma_{LS} \tag{2.15}$$

In the case of the growth techniques, depending on the angle ϕ we can distinguish the different situations. In general, also in the case of a liquid droplet, we have that the smaller ϕ, the more spread the droplet on the surface while the more ϕ is close to $\pi/2$, the more the droplet tends to assume a spheric shape. As a final consideration, it is worth noting that γ_{SA} and γ_{LS} are more precisely surface stresses and not surface tensions. Nevertheless, within this context, they are actually exchangeable since they represent two ways of expressing the surface energy.

2.2.2 Surface Stress

The elastic process used for liquid surfaces is proper for liquids. When we have a solid crystal the atom can not come from the bulk to build a new surface. In these cases, when a new surface is created, the particles which compose the surface encounter relaxation processes which result in a change of the inter-particle distance at the surface. This is a so-called *elastic* process. When this occurs, the change in the inter-particle distance implies a change in the chemical potential, i.e. the environment, of the particles and this makes the use of the Landau potential no longer fruitful. As we will see in a while, it is way better to use the Helmholtz free energy. Before dealing with that, since surface stress relays on stresses applied to the surface, it is better to define some useful quantities. Let's consider a 2D surface at which a force $\boldsymbol{F}$ is applied parallel to the surface. The **strain** is a measure of deformation experienced by a body due to applied forces, temperature changes, or other external factors. There are two types of strain: normal strain and shear strain; the difference between these two can be understood by looking at Fig.2.5. In the more interesting and general case of shear strain, it is possible to define the strain ε as the ratio between the deformation Δx and the length of the surface l. However, since surfaces

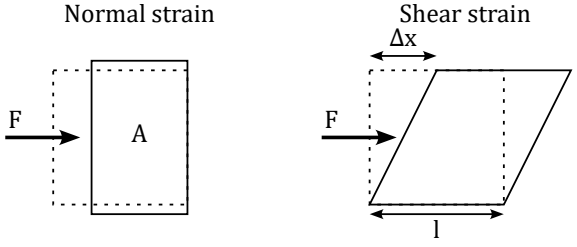


Figure 2.5: Normal and Shear Strains.

are not isotropic and different directions can lead to different deformations, i.e. different strains, we define the strain as a tensorial quantity:

$$\varepsilon_{ij} = \frac{\Delta x}{l}\Big|_{ij} \tag{2.16}$$

with $i, j = 1, 2$. Another important quantity to evaluate when a force is applied parallelly to a surface is the **stress** σ which represents the force per unit area; because of the non-isotropy of surfaces, we define also this quantity as a tensor:

$$\sigma_{ij} = \frac{F_{\parallel}}{A}\Big|_{ij} \tag{2.17}$$

where A is the area of the surface. The elastic process associated with applied stresses is similar to what happens in the simple case of a body attached to a spring. From an energetic point of view, to displace a body by a quantity dx subjected to an elastic force F, we need to do an infinitesimal work $dW = Fdx$. Similarly, when a stress σ_{ij} is applied to a surface and this results in a strain ε_{ij}, we have to do a work $dW \propto \sigma_{ij} d\varepsilon_{ij}$. The proportionality constant is the area of the surface A: the bigger the area, the bigger the work required for a deformation. Lastly, we need to consider the possibility that stress in a particular direction causes a strain in another direction. So for the overall surface:

$$dW = A \sum_{ij} \sigma_{ij} d\varepsilon_{ij} \tag{2.18}$$

In general, it holds that:

$$dW = dF^s = d(\gamma' A) = A d\gamma' + \gamma' dA \tag{2.19}$$

This means that the overall work is given by a first part in which the area stays constant and the surface-specific energy varies, i.e. an elastic process, and a part in which the surface-specific energy remains constant and the area changes, i.e. a plastic process. Since here we are considering only plastic processes we can impose:

$$A d\gamma' = A \sigma_{ij} d\varepsilon_{ij} \Longrightarrow \sigma_{ij} = \frac{d\gamma'}{d\varepsilon_{ij}} \tag{2.20}$$

This gives us an idea of how it is possible to link thermodynamic quantities, such as surface energy, to mechanical ones. However, a more detailed description is required in order to get the correct result. The complete description of surface stress leads us to the so-called **Shuttleworth equation**:

$$\sigma_{ij} = \gamma' \delta_{ij} + \frac{\partial \gamma'}{\partial \varepsilon_{ij}} \tag{2.21}$$

2.3 γ-Plot and Vicinal Surface

The surface energy γ plays a fundamental role in surface physics. From liquids to solids, γ allows us to keep into account all the processes that occur during the formation of a surface which are somehow included inside the value of the surface energy. For crystalline materials, most surface properties depend on the orientation. In particular, depending on the surface orientation, represented by Miller's indices (hkl), more or less bonds have to be broken to create a piece of surface. In turn, also the surface energy will depend on these indices, i.e. $\gamma = \gamma(hkl)$. A graph showing the value of γ as a function of the crystallographic plane orientation is called γ-Plot or Wulff Plot. In these graphs, the scalar surface energy is plotted in polar coordinates versus the angle θ between a particular fixed direction and the

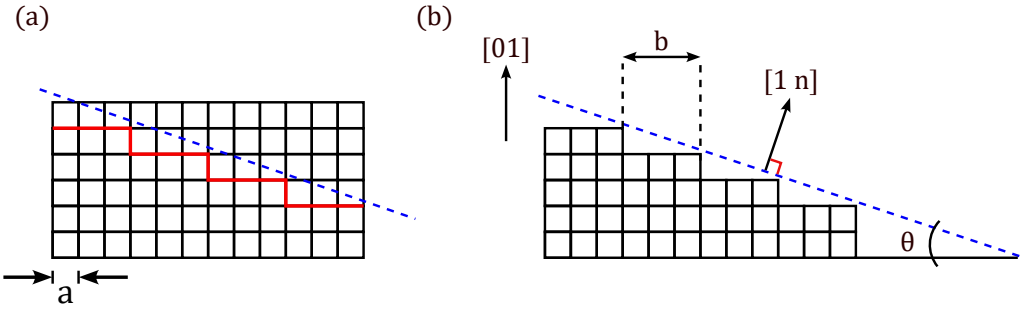


Figure 2.6: (a) Schematic of a surface made of equal boxes of side a with cutting direction. (b) Schematic of the surface after the cut; this is a so-called vicinal surface plane with inclination angle θ against [01].

normals to the (hkl) planes. The length of the vector from the origin to a point on the plot represents the magnitude of $\gamma(hkl)$, and the direction is that of the normal to the (hkl) plane. As a simple example for the rough calculation of γ and its representation, let us consider a two-dimensional crystal in which each atom is represented by a squared box of side a. Let's call γ_0 the initial surface energy of such a surface shown in Fig.2.6(a). Let's now imagine cutting the surface along a particular direction (dashed line in Fig.2.6(a)) which forms an angle θ with the (01) direction. The surface after the cutting will be formed by a series of equal steps; we define n the number of boxes for each step and b the effective width of each step as represented in Fig.2.6(b). The new surface is a so-called **vicinal surface plane**, which consists of a surface plane made of a relatively high number of boxes with (01) orientation being separated by ledges of atomic height. We can calculate the width b of each step as:

$$b = na = \frac{a}{\tan\theta} \tag{2.22}$$

Also, we can define the step density δ_s as the inverse of the step width b:

$$\delta_s = \frac{1}{\delta_s} = \frac{1}{na} = \frac{\tan\theta}{a} \simeq \frac{\theta}{a} \tag{2.23}$$

where the last equivalence holds for very small θ. In this case, $\theta = 1/n$. We are now interested in evaluating the specific surface energy for such a system as a function of the θ angle. We note that the more steps we have per unit surface area, the higher the surface energy will be. This is because each step brings to the system an energetic contribution; we call β_s this contribution, the specific surface energy of the step. Moreover, we note that

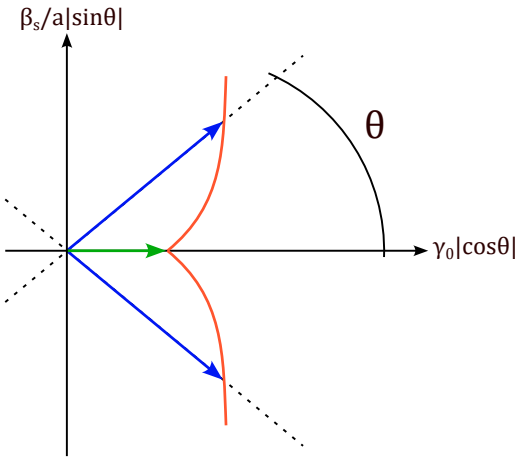


Figure 2.7: Schematic of a γ-Plot for the vicinal surface plane considered.

the more inclined the surface, the more the number of steps and therefore the smaller the influence of the initial energy γ_0 on the overall energy. Considering what told so far we can write:

$$\gamma(\theta) = \gamma_0|\cos\theta| + \beta_s\delta_s|\cos\theta| \tag{2.24}$$

Substituting what obatined in Eq.(2.23) and rearraning the terms we get:

$$\gamma(\theta) = \gamma_0|\cos\theta| + \frac{\beta_s}{a}|\sin\theta| \tag{2.25}$$

If we now plot the two components of $\gamma(\theta)$ as in Fig.2.7 we obtain the γ-Plot we discussed before. In this polar diagram, the horizontal axis corresponds to the direction $\theta = 0$ and therefore to $\gamma(\theta = 0) = \gamma_0$. In general, it is possible to see that there is always a minimum in $\gamma(\theta)$ in correspondence with the lower Miller indices direction.

2.3.1 Surface Faceting

Let's now consider another type of surface, quite different from the vicinal surface we have just analysed. When water freezes into ice, the water molecules stack together to form a regular crystalline lattice, and the ice lattice has six-fold symmetry. It is this hexagonal crystal symmetry that ultimately determines the symmetry of snow crystals. But then one must ask how molecular forces, which operate at the molecular scale to produce the crystal lattice, can control the shape of a snow crystal some ten million times larger. The answer to this has to do with how crystals form **facets**. Facets appear on many growing crystals because some surfaces grow much more slowly than others. If we imagine beginning with a small round ice crystal, then mostly we would find that the surface was quite rough on a molecular scale, with lots of dangling chemical bonds. Water molecules from the air can readily attach to these rough surfaces, which thus grow relatively quickly. The facet planes are special, however, in that they tend to be smoother on a molecular scale, with fewer dangling bonds. Water molecules cannot so easily attach to these smooth surfaces, and hence the facet surfaces advance more slowly. After all the rough surfaces have grown out, what remains are the slow-moving facet surfaces. A schematic and simplified view of a faceted surface is shown in Fig.2.8 in which the faceted surface is considered to be a sinusoidally corrugated surface. To energetically describe such a surface we use as a reference the Helmholtz energy of the flat surface F_s; in this case, this results in:

$$F_s = \iint \gamma dA \tag{2.26}$$

To describe the faceted surface, we can define the angle θ as the angle between the normal to the faceted surface and the vertical direction, i.e. the normal of the flat surface, as in Fig.2.8. It is reasonable to think that the surface energy in such a system depends on the angle θ so that $\gamma = \gamma(\theta)$. If the angle θ is very small we can develop in series the surface energy around the point $\theta \simeq 0$:

$$\gamma(\theta)|_{\theta\simeq 0} = \gamma(0) + \frac{d\gamma}{d\theta}\theta + \frac{d^2\gamma}{d\theta^2}\theta^2 + \dots \tag{2.27}$$

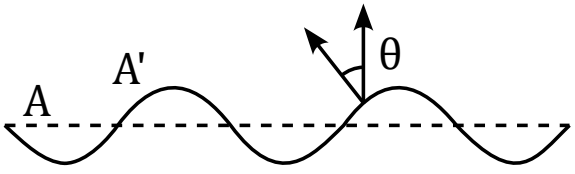


Figure 2.8: Schematic of an ideal faceted surface.

And the Helmholtz free energy F_s' of the faceted surface will be:

$$F_s' = \iint \gamma(\theta) dA' = \iint \left[\gamma(0) + \frac{d\gamma}{d\theta}\theta + \frac{d^2\gamma}{d\theta^2}\theta^2 + \dots \right] \frac{dA}{\cos\theta} \tag{2.28}$$

where we considered $dA' = dA/\cos\theta$. Now, since $\theta \simeq 0$ we can approximate $\cos\theta \simeq 1$ and thus $dA' = dA$. So that:

$$F_s' \simeq \iint \gamma(0) dA + \iint \frac{d\gamma}{d\theta}\theta dA + \iint \frac{d^2\gamma}{d\theta^2}\theta^2 dA + \cdots \tag{2.29}$$

Here we notice that the first term corresponds to the Helmholtz energy of the flat surface F_s. We can therefore compute the difference between the flat and the faceted configuration as:

$$\Delta F_s = F_s' - F_s = \iint \frac{d\gamma}{d\theta}\theta dA + \iint \frac{d^2\gamma}{d\theta^2}\theta^2 dA + \cdots \tag{2.30}$$

However, the system we are now considering is symmetric and therefore we expect the θ configuration to be energetically equivalent to the $-\theta$ one. This means that in the series expansion, all the odd terms will vanish. Therefore, in the first order, we obtain:

$$\Delta F_s = \iint \frac{d^2\gamma}{d\theta^2}\theta^2 dA \tag{2.31}$$

This means that, depending on the result of this integral, the variation in the Helmholtz free energy of the flat or the faceted surface can be either positive or negative. If $\Delta F_s > 0$ then $F_s < F_s'$ and the formation of a flat surface will be favoured, while if $\Delta F_s < 0$ then $F_s' < F_s$ and the surface will present a corrugated, i.e. faceted, aspect.

2.4 Experimantal Evaluation of the Surface Energy

The previous sections showed how surface energy represents a fundamental quantity in the study of surfaces since it provides a lot of information on the surface itself. Unfortunately, an a priori calculation of this quantity is not possible and we must rely on experimental evaluations. In the following, we will see how the surface energy can be calculated in the case of a liquid and a solid surface.

2.4.1 Liquids

The method used to evaluate the surface energy in liquids, i.e. the surface tension γ is the **gravimetric method**. In the gravimetric method, the surface tension of a liquid is determined by measuring the force required to pull a material (such as a plate or a wire) out of the liquid's surface. This force is typically measured using a precision balance or a force sensor (e.g. a spring). By knowing the dimensions of the material and the force required to detach it from the liquid surface, the surface tension can be calculated. Fig.2.9 shows a schematic of an experiment performed with the gravimetric method. In the figure, the surface tension and the surface stresses are shown. By using the scheme of Fig.2.9, the force pushing down the bar associated to γ is $2L\gamma\cos\xi$ where L is the length of the bar and the 2 factor appears when considering both sides of the bar. By measuring the elongation of the spring or using any other instruments to measure the force one can retrieve the value of the surface tension. This method is for example used in some deposition techniques which use the Langmuir bath to detect the formation of a monolayer.

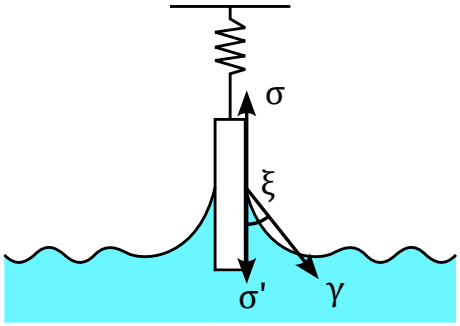


Figure 2.9: Schematic of the experimental setup for the calculation of γ with the gravimetric method.

2.4.2 Solids

In solids, the situation is drastically different. We know that in solids the surface energy manifests as a surface stress σ, unfortunately, however, at the state of the art there is no experimental technique to determine it. The only thing that can be done to set a reference value for the surface energy of solids is to use the so-called **Zisman plot**. The idea is to relate the surface energy of liquids, which can be calculated, with the surface stress. For example, if we follow the same scheme of Fig.2.9 we get:

$$\sigma - \sigma' = \gamma \cos \xi \tag{2.32}$$

This equation tells us that the surface stress, or at least its difference, can be linked to the surface tension of the liquid. In particular, we notice that if we plot the cosine of the angle formed at the solid-liquid interface $\cos \xi$ as a function of the surface energy γ for different solids (i.e. different bars) immersed in the same liquid, we get a well define behaviour as depicted in Fig.2.10(a). In this graph, it appears clear the presence of a critical value for the surface energy γ at which the cosine of ξ starts decreasing noticeably with respect to the previous constant value. We assume this critical value to be the surface energy of the solid surface. It is worth noting that there is no physical explanation for this assumption, this critical value does not represent the real value of the surface stress. Nevertheless, we prefer to pick a reference value for the surface energy of a solid rather than nothing. The reference values of the surface energy can be then plotted in a graph as a function of the orientation angle with respect to a well-defined crystalline direction. Let's suppose to take the (100) surface of Si and imagine we are interested in defining its direction. The simplest idea is to define the orientation angle of this surface with respect to a reference plane. For Si, we choose this reference plane to be the (111) and we call θ the angle formed by the reference plane and any other surfaces. We can now visualize the surface energy of each

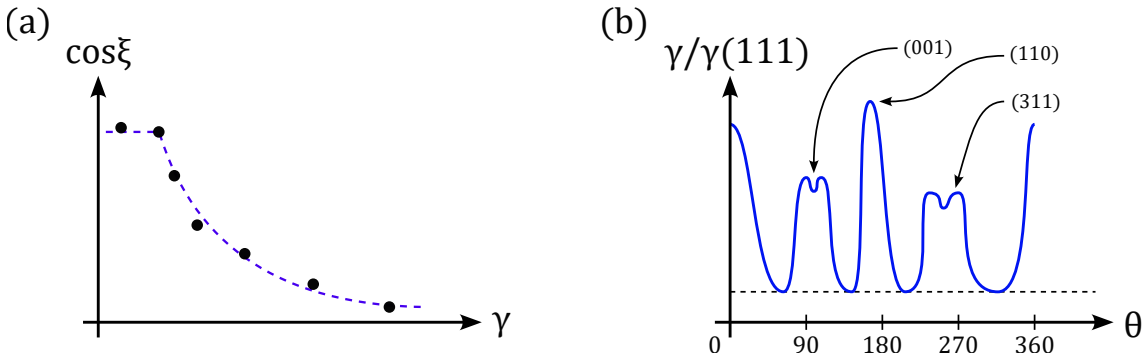


Figure 2.10: (a) Schematic of a Zisman plot with the same setup of the gravimetric experiment. (b) Surface energy for Si at different plane orientations with respect to the (111) direction.

surface by simply plotting the surface energy as a function of this angle θ. The result of such a representation is shown in Fig.2.10(b). As already pointed out, we notice that the minima of the surface energy correspond to the lower Miller's index directions.

2.5 Wulff Construction

The study of crystal growth techniques marks a significant departure from the historical reliance on naturally occurring crystals. Prior to the development of these techniques, crystals were solely sourced from nature, subject to the whims of geological processes. However, the advent of controlled laboratory environments has facilitated the deliberate synthesis of crystals with well-defined shapes, orientations, compositions, and passivation. This shift in approach has not only broadened our understanding of crystallography but has also revolutionized various industries. By enabling precise manipulation of crystal properties, growth techniques have become indispensable tools in fields ranging from electronics to materials science and pharmaceuticals. As the importance and precision of such applications grow, it becomes more and more important to control the growth of a crystal and its characteristics. In this context, the **Wulff construction** of crystal surfaces emerges as a fundamental concept. Named after the Russian mathematician and crystallographer, George Wulff, this construction elucidates the equilibrium shape of a crystal based on its surface energy anisotropy. By considering the relative surface energies of different crystallographic facets, the Wulff construction is a method to determine the equilibrium shape of a droplet or crystal of *fixed volume* inside a separate phase. In 1901 Wulff stated (without proof) that the length of a vector $\boldsymbol{h}_j$ drawn normally to a crystal face will be proportional to its surface energy γ_j:

$$h_j = \lambda \gamma_j \tag{2.33}$$

where the vector h is the "height" of the j-th face, drawn from the centre of the crystal to the face; for a spherical crystal this is simply the radius. Here, $\gamma = \gamma(hkl)$ is a function of Miller's indices and λ is a proportionality constant. Eq.(2.33) is known as the Wulff hypothesis. Some years later this relation was finally proven with thermodynamics arguments. It is important to stress that the Wulff construction concerns crystals which have been grown and are now in phase equilibrium[3] with the environment. In particular, the growth of the crystal starts with a small fragment of it with different surface orientations: being the surface energy dependent on the orientation ($\gamma = \gamma(hkl)$), different surfaces will have different energies and therefore different growth. Wulff's construction tells us that the bigger the energy the bigger the thickness of the grown layer. In the following, we will give a quite simple proof of this hypothesis.

Proof

Let's consider a crystal in phase equilibrium after its growth has been performed. Since we are considering phase equilibrium, the volume of the crystal must be constant. In fact, for each atom in the vapour phase which condensates on the crystal, an equal amount of atoms will sublimate from it. Hence, we divide our crystal into a series of equal slices (Riemann cut of a solid) each one of constant volume V_c. These slices are also called *pancake crystals*. Each pancake crystal has an area A_j. Along the j-th direction the Helmholtz free energy

[3] Phase equilibrium refers to the state of a system in which the different phases coexist in such a way that there is no net change in the overall composition of the system over time. In simpler terms, it's a condition where the relative amounts of different phases within a system remain constant.

is $F_{s_j} = \gamma_j A_j$. The overall energy will be:

$$F_s = \sum_j F_{s_j} = \sum_j \gamma_j A_j \tag{2.34}$$

Since we are considering the crystal at its equilibrium, i.e. after the growth, this energy term must be minimum:

$$\delta F_s = \delta \left(\sum_j \gamma_j A_j \right) = 0 \tag{2.35}$$

In solids, we are interested only in considering plastic deformations and therefore:

$$\delta \left(\sum_j \gamma_j A_j \right) = \sum_j \gamma_j \delta A_j = 0 \tag{2.36}$$

This represents the thermodynamic equation of such a system which is a direct consequence of the minimization of the energy. We now have to add a geometrical constraint. Being A_j the area of the pancake crystal and h_j its height, we have for a constant volume crystal:

$$0 = \delta V_c = \delta \left(\sum_j h_j A_j \right) = \sum_j \left(A_j \delta h_j + h_j \delta A_j \right) \tag{2.37}$$

The first term of this sum must be zero. This is because, if the volume is to remain constant, the changes in the heights of the various faces must be such that when multiplied by their surface areas the sum is zero. In other words, for every added slice δh_j, one must be removed. So that $\sum_j A_j \delta h_j = 0$ and Eq.(2.37) becomes:

$$\delta V_c = \sum_j h_j \delta A_j = 0 \tag{2.38}$$

which represents the geometrical constraint. In summary, the thermodynamic equation and the geometrical constraint to solve are:

$$\sum_j \gamma_j \delta A_j = 0$$
$$\sum_j h_j \delta A_j = 0$$

These two equations can be summed (or subtracted) provided that we use a multiplication constant to make the unit of γ_j fit with the one of h_j. Being λ_j such a constant we have:

$$\begin{aligned} &\sum_j \lambda_j \gamma_j \delta A_j - \sum_j h_j \delta A_j = 0 \\ &\Rightarrow \sum_j \left(\lambda_j \gamma_j - h_j \right) \delta A_j = 0 \end{aligned} \tag{2.39}$$

which is satisfied if $\lambda_j \gamma_j - h_j = 0$ which is the Wulff's hypothesis of Eq.(2.33). After a century from its first formulation, Wulff's construction is still at the centre of the scientific debate. For example, in 2019 the so-called Inverse Wulff's construction was developed to try to measure the surface free energy of a solid using the inverse of Wulff's hypothesis. Moreover, the model is not able to predict the growth of any kind of structure. For some particular structures, as the one reported in Fig.2.11 showing a quantum dot (QD) of Ge grown on a SiC substrate, the model is not trustable. In particular, for this QD it is not possible to explain the relaxation process which takes place in germanium; the very big lattice mismatch between Ge ($a = 5.7$Å) and SiC ($a = 3.1$Å) shouldn't allow, based on the Wulff's reconstruction, the realization of such a structure.

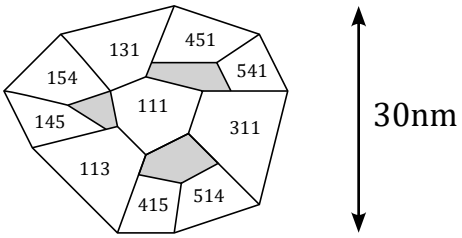


Figure 2.11: Schematic of Ge quantum dots grown on a SiC substrate, the shaded surfaces are unknown.

2.5.1 Poly-compounds Surfaces

A poly-compounds surface is a surface of a material composed of more than one element. In the following, we will refer to the particular case of gallium arsenide (GaAs). When a surface is made by more than one element, we can not adopt Wulff's construction and it is required to introduce and use a new form of energy. This energy is the **Gibb's free energy** $G = G(T, p, N_j)$ defined as:

$$G(T, p, N_j) = F + pV \tag{2.40}$$

where p is the pressure and the subscript j denotes the j-th element of the surface. This expression is also particularly important because it allows us to define the chemical potential μ_j of the j-th element:

$$\mu_j = \frac{\partial G}{\partial N_j}\Big|_{p,T,N_{i\neq j}} \tag{2.41}$$

In the case of GaAs, we have:

$$\mu_{Ga}^{bulk} = -2.81\,\frac{eV}{\text{atom}} \qquad \mu_{As}^{bulk} = -2.96\,\frac{eV}{\text{atom}}$$

When the formation of the bulk material occurs we need to consider a certain energy contribution due to the formation of the compound. Such an energy contribution is called enthalpy of formation ΔH and therefore:

$$\mu_{GaAs}^{bulk} = \mu_{Ga}^{bulk} + \mu_{As}^{bulk} - \Delta H_{GaAs} \tag{2.42}$$

where in our case $\Delta H_{GaAs} = 0.74\,eV/\text{atom}$. When we are growing a new piece of the material, and therefore we are creating a new surface, an energy contribution, in terms of chemical potential, must be paid. This result in:

$$\mu_j = \mu_j^{bulk} + \Delta\mu_j \tag{2.43}$$

In our case:

$$\mu_{GaAs} = \mu_{Ga}^{bulk} + \mu_{As}^{bulk} + \underbrace{(\Delta\mu_{Ga} + \Delta\mu_{As})}_{-\Delta H_{GaAs}} \tag{2.44}$$

It is possible to see that the possibility to relate the variation in chemical potential to the enthalpy variation, allows the prediction of the Landau grand potential Φ and thus the evaluation of Φ as a function of $\Delta\mu$. This results in the creation of charts, as the one reported in Fig.2.12, in which the Landau gran potential of the whole compound is expressed as a function of the chemical potential of one of the elements which compose the material, in this case, Ge. In particular, the chart shows two lines, called *equilibrium lines*, which define those values of chemical potential and energy for which a surface assumes a particular reconstruction, i.e. a particular superlattice. Moving along the x-axis in the Figure, we pass from a compound with high percentages of As (As-rich) to a compound

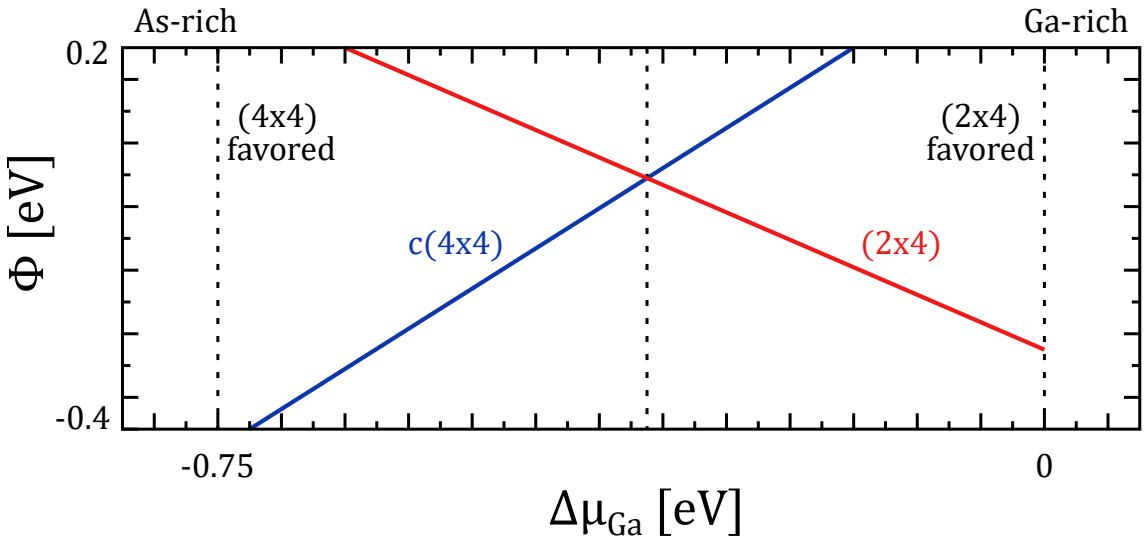


Figure 2.12: Landau grand potential for different Ga chemical potential shift.

mostly made of Ge (Ge-rich). These kinds of charts allow us to understand which will be the superlattice assumed by a certain surface given a determined variation of chemical potential: the configuration with minimum energy will be favoured during the growth (see Fig.2.12). It is worth noting that these charts are based on theoretical evaluation which neglects the presence of defects, impurities and other non-idealities on the surface. The final configuration assumed by a surface will also depend on these factors and may differ from what is predicted theoretically.

Chapter 3

Electronic Surface States

In surfaces, atoms have fewer neighbouring atoms compared to bulk atoms, and some chemical bonds within the bulk crystal structure are broken at the surface. This breaking of bonds requires energy, known as surface energy, to create the surface. Consequently, the electronic structure near the surface differs significantly from that of the bulk material. Even when considering an ideal truncated bulk reference surface, new electronic levels and altered many-body effects emerge due to changes in chemical bonding. Many observable surface phenomena, such as surface free energy, adhesion forces, and specific chemical reactivity, stem from these alterations in electronic structure. Understanding the electronic structure of surfaces is a key focus in modern surface physics. The theoretical approach mirrors that used for bulk crystals, employing the one-electron approximation and attempting to solve the Schrödinger equation for electrons near the surface, with various approximation methods utilized to incorporate many-body effects. In the following, we will try to study the electronic states of surfaces in some simple cases both for metals and for semiconductors. We will start by dealing with the metal case by using the so-called jellium model. Then, by expoliting the thight binding approach we will try to describe the electronic situation of a semiconductor surface.

3.1 Electron Density and Friedel Oscillations

We will begin the study of the electronic states of surfaces by considering the case of metal surfaces treated in the context of the **jellium model**. The jellium model is a simplified theoretical framework used in condensed matter physics to describe the behaviour of electrons in a *metal*. In this model, the positively charged atomic cores are treated as a uniform, positively charged background, often referred to as a "jellium" because it resembles a jelly-like substance. The valence electrons are then considered to move freely within this uniform positive background, similar to electrons moving through a positively charged sea. In such a model, the sea of electrons moving through the crystal is confined within the crystal itself. We need to represent such a confinement in the expression of the wave function of the electron. To do that, in the context of the jellium model, we can start by writing the wave function as one of the free electrons; by referring to the axis in Fig.3.1(a) we have for the electron going towards the surface:

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega L}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = \frac{1}{\sqrt{\Omega L}} e^{i(k_x x + k_y y)} e^{ik_z z} \tag{3.1}$$

where Ω is the volume of the primitive cell and L is the number of primitive cells in the crystal. In this expression, we isolate the term related to the components of the wave vector in the $x-y$ plane since no confinement is present in that direction. Conversely, the

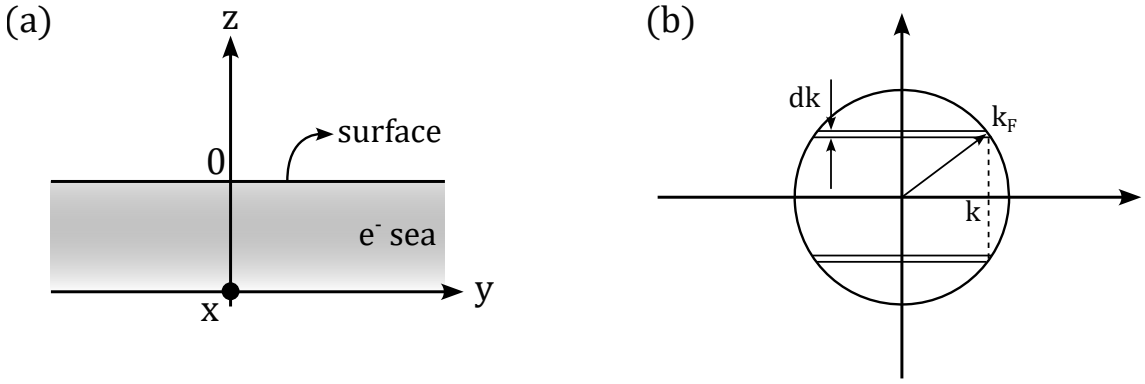


Figure 3.1: (a) Schematic of the surface in terms of the electron confinement. (b) Fermi sphere for electrons.

confinement in the z direction can be imposed by saying that the electrons, represented by their wave function, will be reflected at the surface[1]. To impose the reflection at the surface, and thus the interference between the incoming and the reflected wavefunctions, we need to add a term into the previous wave function to include all those identical electrons propagating in the opposite direction. However, we can not exclude the possibility that the wave function gets modified while interfering with the surface. Mathematically, this translates into the addition of an unknown phase term δ to the wave function:

$$\psi(\boldsymbol{r}) = \frac{1}{\sqrt{\Omega L}} e^{i(k_x x + k_y y)} \left(e^{ik_z z} + e^{-ik_z z + i\delta} \right) \tag{3.2}$$

Since we are interested in evaluating the behaviour of the electrons at the surface, we need to compute the electron density; to do that we first have to evaluate $|\psi(\boldsymbol{r})|^2$. So that:

$$\begin{aligned} |\psi(\boldsymbol{r})|^2 &= \frac{1}{\Omega L} e^{-i(k_x x + k_y y)} \left(e^{-ik_z z} + e^{ik_z z - i\delta} \right) e^{i(k_x x + k_y y)} \left(e^{ik_z z} + e^{-ik_z z + i\delta} \right) = \\ &= \frac{1}{\Omega L} \left(1 + 1 + e^{-i2k_z z + i\delta} + e^{i2k_z z + i\delta} \right) = \frac{2}{\Omega L} \left[1 + \cos(2k_z z - \delta) \right] \end{aligned} \tag{3.3}$$

The presence of the surfaceat $z = 0$ implies that $|\psi(z = 0)|^2 = 0$ so that:

$$|\psi(z = 0)|^2 = 0 \Rightarrow \cos(-\delta) = -1 \Rightarrow \delta = \pi + 2n\pi \tag{3.4}$$

By remambering that $\cos(\alpha - \pi) = -\cos\alpha$ we finally obtain:

$$|\psi(\boldsymbol{r})|^2 = \frac{2}{\Omega L} \left[1 - \cos(2k_z z) \right] \tag{3.5}$$

In the following, we will drop the subscript z since we will only refer to the z component of the k vector. We can now proceed with the calculation of the electron density $\rho(\boldsymbol{r})$ by integrating $|\psi(\boldsymbol{r})|^2$ over all the possible values of the wave vector k, i.e. for $k \leq k_F$ with k_F Fermi wave vector. To do this, we need the density of states in the k-space which is the volume of the spherical annulus of thickness dk represent in grey in Fig.3.1(b); the volume of such a spherical annulus is:

$$\pi(k_F^2 - k^2)dk$$

This volume must be divided by the volume of a single k state in the k-space $\Delta k = (2\pi)^3/\Omega L$. Therefore, the infinitesimal electron density $d\rho$ is:

$$d\rho = 2\frac{\Omega L}{(2\pi)^3} \pi(k_F^2 - k^2)dk|\psi(\boldsymbol{r})|^2 \tag{3.6}$$

[1]This is a consequence of the confinement and empirically is the representation of the fact that normally we do not see electrons exiting a surface for no reason.

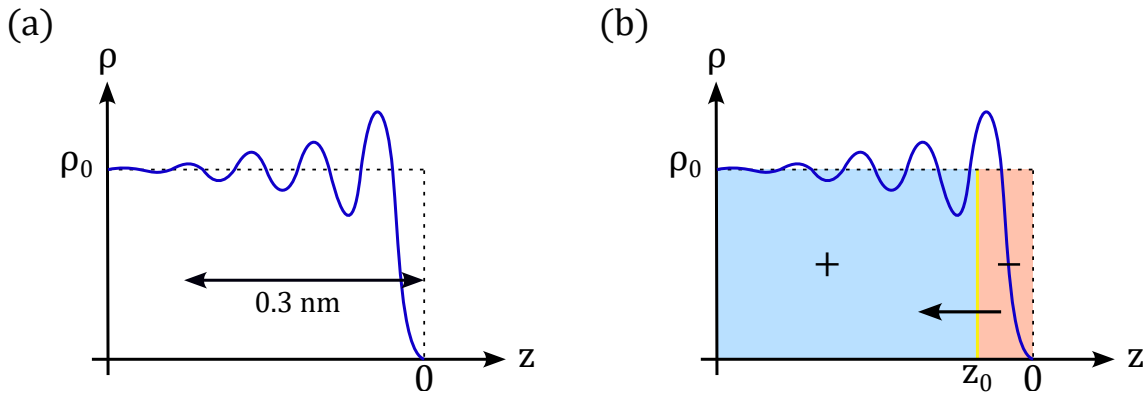


Figure 3.2: (a) Electron density as a function of the z coordinate which exhibits Friedel oscillations. (b) Schematic of the appearance of the electric dipole at the surface.

where the initial 2 factor is due to the spin degeneracy. By simplifying the terms and substituting Eq.(3.5) one gets:

$$d\rho = \frac{1}{2\pi^2}(k_F^2 - k^2)\,[1 - \cos(2k_z z)]\,dk \tag{3.7}$$

This expression has to be integrated for k in $[0,\ k_F]$; For brevity's sake, we will report here only the solution of this integral which is not difficult, but long to compute. By integrating by parts one gets:

$$\rho(\boldsymbol{r}) = \rho_0 + 3\rho_0 \left[\frac{\cos(2k_F z)}{(2k_F z)^2} - \frac{\sin(2k_F z)}{(2k_F z)^2}\right] \tag{3.8}$$

where $\rho_0 = k_F^3/3\pi^2$ is the bulk density, i.e. the positive ions density. The behaviour of the electron density as a function of the z coordinate is shown in Fig.3.2(a) in which we notice the presence of an oscillatory behaviour. This oscillatory behaviour takes the name of **Friedel oscillations**. As expected, the electron density tends to ρ_0 for $z \to -\infty$, while it vanishes for $z = 0$. The extension of these oscillations can be calculated within the jellium model and results to be about $0.3\,nm$ as shown in Fig.3.2(a). This is in agreement with what is obtained in Section 1.2.3 for the relaxation process of Al(110) in which $d_{12} = 1.3$ Å which is in the same order of magnitude of the results obtained by imposing the confinement.

3.1.1 Surface Dipole

Not only for electrons, the surface must be a barrier also for the ions and therefore we need to set a boundary also for the positive charges in the crystal. This can be done by exploiting the principle of *charge neutrality*, i.e. the fact that the overall positive charge must equal the negative one. The overall negative charge Q^- can be easily computed as:

$$Q^- = -eA\int_{-\infty}^{0}\rho(z)dz = \int_{-\infty}^{0} -eA\left\{\rho_0 + 3\rho_0\left[\frac{\cos(2k_F z)}{(2k_F z)^2} - \frac{\sin(2k_F z)}{(2k_F z)^2}\right]\right\}dz \tag{3.9}$$

where A is the area of the surface. Instead, the positive charge can be computed as:

$$Q^+ = eA\int_{-\infty}^{z_0}\rho_0 dz \tag{3.10}$$

where z_0 is the unknown coordinate until which the positive ions extend. By solving the integrals and imposing the neutrality $Q^- = Q^+$ one gets:

$$Q = Q^+ + Q^- = Ae\rho_0\left(\frac{3\pi}{8k_F} + z_0\right) = 0 \tag{3.11}$$

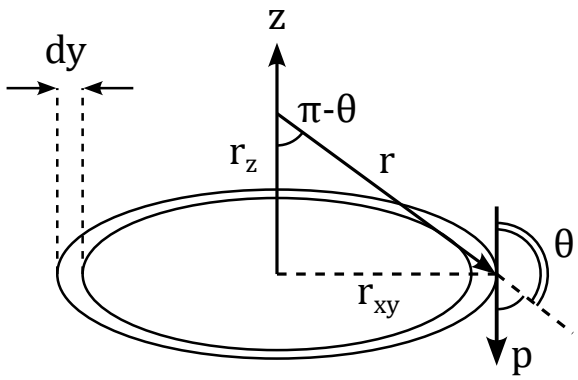


Figure 3.3: Schematic of the surface dipole and its geometry.

which solved for z_0 gives:

$$z_0 = -\frac{3\pi}{8k_F} \tag{3.12}$$

This value represents the limiting extension of the positive ions in the crystal. We note here that $z_0 < 0$ which means that some electrons, the ones at $z_0 < z < 0$ will not be compensated by an equal amount of positive charges. This excess of electrons gives rise to an excess of negative charge. On the other hand, because of the Friedel oscillations, we have an excess of positive charge for $z < z_0$. As shown in Fig.3.2(b) this unbalance of charges gives rise to an electric dipole which physically represents the energy barrier seen by the electron at the surface. Moreover, it is worth noting that the z_0 coordinate represents a sort of dividing plane for the charge and the excess of electrons at the surface is perfectly compatible with what is told for the relaxation phenomenon (see Fig.1.6).
We can estimate the height of the potential barrier seen by the electrons by using the potential associated with the electric dipole formed at the surface. In general, for a dipole $\boldsymbol{p}$ with modulus $p = ed$, where d is the distance between the centres of positive and negative charges, the potential at a distance $\boldsymbol{r}$ from the dipole is:

$$V(\boldsymbol{r}) = \frac{1}{4\pi\varepsilon_0}\frac{\boldsymbol{p}\cdot\boldsymbol{r}}{r^3} = \frac{e}{4\pi\varepsilon_0}\frac{d\cos\theta}{r^2} \tag{3.13}$$

where θ is the angle between $\boldsymbol{p}$ and $\boldsymbol{r}$. In our case, we have a uniform distribution of dipoles all over the surface in the xy plane, and therefore we must consider the element contribution $dV(\boldsymbol{r})$ for each of them as in the scheme of Fig.3.3. In this case, we have:

$$dV(\boldsymbol{r}) = \left[\frac{1}{4\pi\varepsilon_0}\frac{ed}{A}\frac{\cos\theta}{r^2}\right](2\pi r_{xy}dr_{xy}) \tag{3.14}$$

where ed/A is the dipole density. We notice that we can write $\cos\theta$ as $-\cos(\pi-\theta)$ and, by geometrical relation it holds:

$$\cos\theta = -\cos(\pi-\theta) = -\frac{r_z}{r}$$

So that:

$$dV(\boldsymbol{r}) = \left[-\frac{1}{4\pi\varepsilon_0}\frac{ed}{A}\frac{r_z}{r^3}\right](2\pi r_{xy}dr_{xy}) \tag{3.15}$$

By defining the integration variable $u = r^2 = r_{xy}^2 + r_z^2$ we can integrate this expression all over the xy plane. Being $du = 2r_{xy}dr_{xy}$ we get:

$$V(\boldsymbol{r}) = \frac{ed}{4A\varepsilon_0}r_z\int_{r_z^2}^{\infty}\frac{du}{u^{3/2}} = \frac{ed}{2\varepsilon_0 A} \tag{3.16}$$

We can now define the height of the barrier that an electron would have to overcome to exit the surface. We define Bulk Potential Energy (BPE) as the potential in the bulk of the material and Vacuun Potential Energy (VPE) as the energy outside the material. With this description, what we have just computed is actually V_{VPE}. Mathematically, the expression of V_{BPE} is similar to the one of Eq.(3.16) but with opposite integration extremes thus resulting in the same value with opposite sign, i.e. $V_{BPE} = -ed/2\varepsilon_0 A$. Therefore, the potential barrier seen by an electron on the surface is:

$$\Delta V = V_{VPE} - V_{BPE} = \frac{ed}{\varepsilon_0 A} \tag{3.17}$$

Within the context of the jellium model, the height of this barrier can be computed and it is about $0.3\,eV$.

3.2 Tight Binding Model and Hybridization

What mostly distinguishes semiconductors from metals is the different types of bonds which characterize the atoms forming the crystal. In particular, semiconductors are characterized by very strong bonds, mostly covalent, which are suitable to be treated with a **thight binding** approach. To understand the bonding behaviour in the surface of a semiconductor on a length scale of the order of nearest-neighbour distances, it is convenient to expand the single-particle wave function $\Psi(\boldsymbol{x})$ in terms of orbitals $\phi_a(\boldsymbol{x})$ centred labels at an atomic core at the origin of coordinates. Here, the index a labels the kind of atom and the atomic quantum number, while $\boldsymbol{x}$ are the electron polar coordinates and $\boldsymbol{R}_i$ is the ion core position, where i labels the different ions. Within the so-called linear combination of atomic orbitals (LCAO) method, we can write the single particle wave function as:

$$\phi(\boldsymbol{x}) = \sum_{a,i} c_{a,i}\phi_a(\boldsymbol{x} - \boldsymbol{R}_i) \tag{3.18}$$

By inserting this expression of the wave function into the time-independent Schrödinger equation one gets:

$$\sum_{b,j} \left[H_{ab}(\boldsymbol{R}_i, \boldsymbol{R}_j) - \varepsilon S_{ab}(\boldsymbol{R}_i - \boldsymbol{R}_j)\right] c_{bj} = 0 \tag{3.19}$$

where:

$$H_{ab}(\boldsymbol{R}_i, \boldsymbol{R}_j) = \int d^3\boldsymbol{x}\phi_a(\boldsymbol{x} - \boldsymbol{R}_i)H\phi_b(\boldsymbol{x} - \boldsymbol{R}_j)$$

$$S_{ab}(\boldsymbol{R}_i - \boldsymbol{R}_j) = \int d^3\boldsymbol{x}\phi_a(\boldsymbol{x} - \boldsymbol{R}_i)\phi_b(\boldsymbol{x} - \boldsymbol{R}_j)$$

with H Hamiltonian of the system. Eq.(3.19) is the secular equation of the tight-binding model for molecules. The solution of this secular equation represents a hard challenge in terms of the computational effort required to solve it. Therefore, many semiempirical and semi-empirical techniques have been developed to reduce the numerical effort. An example is the method developed by Hückel in which some quite strong assumptions are made on the form of the Hamiltonian and on the S_{ab} term, i.e. it is assumed to hold $H_{ab} \propto S_{ab}$. A major development in the history of the empirical tight binding method was the work by Slater and Koster. First, they contribute to giving a physical interpretation to the matrix elements H_{ab} that in their framework are interpreted as one-electron energies supposed to be known (either theoretically or experimentally). In this vision, in particular, the diagonal elements H_{ii} represent the valence energies while the off-diagonal element $H_{ij} i \neq j$ represent the bonding energy terms. Second, Slater and Koster invoked a theorem, first proven by

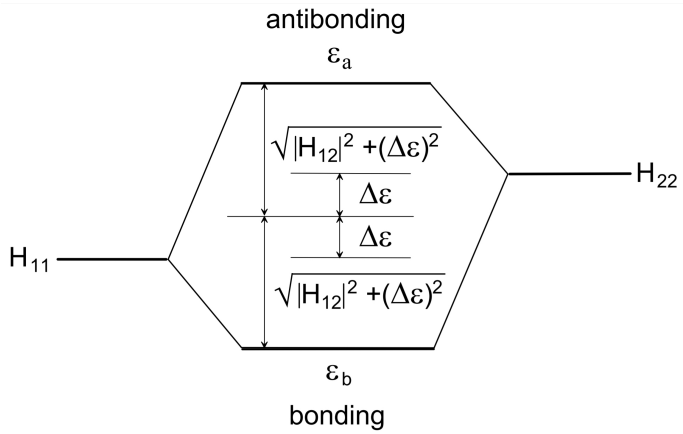


Figure 3.4: Formation of bonding and antibonding molecule levels from atomic levels.

Löwdin, according to which an orthonormal set of LCAO orbitals could be defined without loss of generality such that:

$$S_{ab}(\boldsymbol{R}_i - \boldsymbol{R}_j) = \delta_{ab}\delta_{ij} \tag{3.20}$$

The Löwdin theorem states that a set of non-orthogonal orbitals located at different atoms can always be transformed into a new set of orthogonal atomic orbitals ϕ_a which preserves the atomic geometry and symmetry. Although non-intuitive, the advantages of the empirical approach become clear when considering an isolated bond formed by two orbitals $a = 1, 2$ localized at neighbouring atoms. Let's consider, for example, the case of the diatomic molecules H_2 with a pure *covalent bond*, and the case of LiH with a heteropolar (or simply *polar*) *bond*. The general secular equation for such systems is:

$$\begin{bmatrix} H_{11} - \varepsilon & H_{12} - \varepsilon S_{12} \\ H_{12}^* - \varepsilon S_{12}^* & H_{22} - \varepsilon \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = 0 \tag{3.21}$$

In this system, we notice that the diagonal elements are the energies of the first and the second ion cores respectively while the off-diagonal elements represent the interactions between the ions. Moreover, using the result of the Löwdin theorem in Eq.(3.20) we have: $S_{11} = S_{22} = 1$ and $S_{12} = S_{21} = 0$ which let our system to assume a more simple form. In this case, we obtain two eigenvalues associated with a bonding (i) and an antibonding (ii) orbital:

$$\varepsilon_{i,ii} = \bar{\varepsilon} \pm \sqrt{|H_{12}|^2 + \Delta\varepsilon^2} \tag{3.22}$$

with $\bar{\varepsilon} = (H_{22} + H_{11})/2$ average atomic energy and $\Delta\varepsilon = (H_{22} - H_{11})/2$ half of the atomic energy difference. This results in a well defined energy diagram, schematically represented in Fig.3.4. In the case of a diatomic molecule made by the same atoms (e.g. H_2) the difference between the atomic energy levels will be equal to zero, i.e. $\Delta\varepsilon = 0$, and the energy will be $\varepsilon_{i,ii} = \bar{\varepsilon} \pm |H_{12}|$. However, in the general case of different atoms in the molecule (e.g. LiH), the polarity of the molecule can be associated with a charge excess on one of the two atoms. In other words, this means that one of the atoms will represent the anion and the other the cation. Consequently, $\Delta\varepsilon$ can be also called the *polar energy*, in contrast with the *covalent energy* $|H_{12}|$. In the point-charge picture, $\Delta\varepsilon$ indicates how many electrons are transferred between the cation and the anion along the bond. To better see this, let's consider the system of Eq.(3.21) in the Löwdin theorem approximation:

$$(H_{11} - \varepsilon)c_1 + H_{12}c_2 = 0 \tag{3.23}$$

under the constraint $c_1^2 + c_2^2 = 1$. The bonding energy is $\varepsilon_i = \bar{\varepsilon} + \sqrt{|H_{12}|^2 + \Delta\varepsilon^2}$ and by substituting this in the previous equation we get the eigenvectors:

$$c_1 = \frac{H_{12}}{\varepsilon_i - H_{11}} c_2 = \frac{H_{12}}{\Delta\varepsilon - \sqrt{|H_{12}|^2 + \Delta\varepsilon^2}} c_2 \tag{3.24}$$

In the particular case of a homopolar molecule, for which $\Delta\varepsilon = 0$, we get $c_1 = -c_2$ which combined with the constraint $c_1^2 + c_2^2 = 1$ gives $c_1 = -c_2 = 1/\sqrt{2}$. Physically, this means that in a homopolar molecule, an electron in the molecule has the same probability of being in one of the two atoms equivalently. This is not true for heteropolar molecules in which the probability for an electron to be in a certain position is given by the square modulus of Eq.(3.24). It is now clear that what matters and what makes the difference is the polar energy $\Delta\varepsilon$ and for this reason, it is worth introducing the so-called *bond polarity* α_p:

$$\alpha_p = \frac{\Delta\varepsilon}{\sqrt{|H_{12}|^2 + \Delta\varepsilon^2}} \tag{3.25}$$

The bond polarity can be used to express the eigenvectors as a function of the bond polarity itself. To do that, let's consider the case of small polarity, i.e. $\Delta\varepsilon \sim 0$, so that we can expand the bond polarity in Taylor expansion $\alpha_p \sim \Delta\varepsilon/|H_{12}|$. The eigenvectors in the bonding case become:

$$\begin{aligned} c_1 &= \frac{|H_{12}|}{\Delta\varepsilon - |H_{12}|} c_2 = \frac{|H_{12}|}{\Delta\varepsilon - \frac{\Delta\varepsilon}{\alpha_p}} c_2 = -\frac{|H_{12}|}{\frac{\Delta\varepsilon}{\alpha_p}} \frac{1}{1-\alpha_p} c_2 = \\ &= -\frac{1}{1-\alpha_p} c_2 \simeq -(1+\alpha_p) c_2 \end{aligned} \tag{3.26}$$

where to obatin such result we use the equivalnce $|H_{12}| \simeq \Delta\varepsilon/\alpha_p$. Therefore, in general, by using also the constraint on the two coefficients c_1 and c_2 one gets:

$$|c_1|^2 = \frac{1}{2}(1 \pm \alpha_p), \qquad |c_2|^2 = \frac{1}{2}(1 \mp \alpha_p) \tag{3.27}$$

where the plus and the minus apply for the bonding and antibonding states, respectively.

3.2.1 Hybrid Model

So far we discussed the case of a diatomic molecule. In real solids, however, the situation is not that simple. When dealing with more complex structures, what mostly contributes to the formation of a chemical bond among atoms are the s and p atomic orbitals of the outer electronic shell. These orbitals of the outer electronic shell with orbitals energies ε_s, ε_p and radial functions $R_s(r)$, $R_p(r)$ can be used to express the full wave functions. Using polar coordinates and spherical harmonics $Y_{lm}(\theta, \varphi)$ the full wave functions can be expressed as linear combinations of s and p orbitals:

$$\begin{aligned} |s\rangle &= Y_{00}(\theta,\varphi) R_s(r) = \frac{1}{\sqrt{4\pi}} R_s(r) \\ |p_x\rangle &= -\frac{1}{\sqrt{2}} \left[Y_{11}(\theta,\varphi) - Y_{1-1}(\theta,\varphi) \right] R_p(r) = \sqrt{\frac{3}{4\pi}} \frac{x}{r} R_p(r) \\ |p_y\rangle &= \frac{i}{\sqrt{2}} \left[Y_{11}(\theta,\varphi) + Y_{1-1}(\theta,\varphi) \right] R_p(r) = \sqrt{\frac{3}{4\pi}} \frac{y}{r} R_p(r) \\ |p_z\rangle &= Y_{10}(\theta,\varphi) R_p(r) = \sqrt{\frac{3}{4\pi}} \frac{z}{r} R_p(r) \end{aligned} \tag{3.28}$$

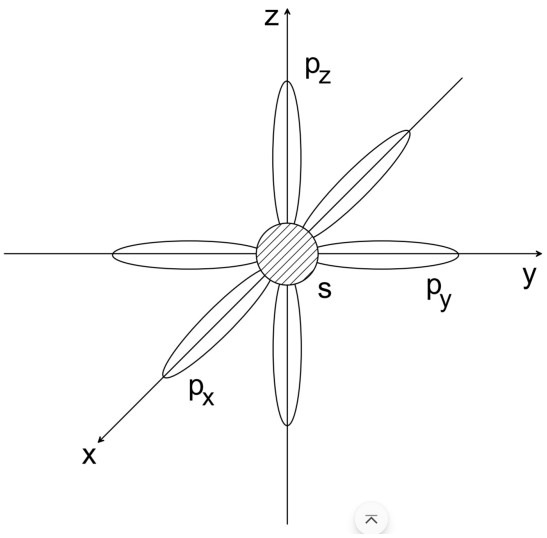


Figure 3.5: Schematic representation of s and p valence orbitals.

where the corresponding atom is assumed to be situated at the origin of the coordinate system. Such orbitals are represented in Fig.3.5. In solids with partially covalent bonds, e.g. semiconductors, the valence electrons are mainly localized between atoms along their bond direction $\boldsymbol{d}$. Therefore, to express the wave function of the compound, it is more convenient to use a basis set in which the atomic orbitals are directed along the bond, instead of using simple s and p orbitals. The new orbitals are linear combination of s and p function and, hence, called $\boldsymbol{sp^{\lambda}}$ **hybrids**. Depending on the overlap among different orbitals belonging to different atoms different hybrid orbitals can be created during the bond. For example, two s orbitals, an s and any p orbitals or two p orbitals placed in a head-to-head configuration give rise to a so-called σ interaction, while p orbitals in a tail-to-tail configuration give rise to a π interaction. The hybrid wave functions in the directions $\boldsymbol{d}_j$ $(j = 1, 2, 3, 4)$ of the bonds are:

$$\phi_{\boldsymbol{d}_j}(\boldsymbol{x}) = \left[\frac{1}{4\pi(1+\lambda_j)}\right]^{\frac{1}{2}} \left\{R_s(r) + \sqrt{3\lambda_j}\frac{\boldsymbol{d}_j \cdot \boldsymbol{x}}{|\boldsymbol{d}_j||\boldsymbol{x}|}R_p(r)\right\} \tag{3.29}$$

where $\boldsymbol{x}$ represent the set of polar coordinates. The energy of an electron in a hybrid orbital such as this one is given by:

$$\varepsilon_{sp^{\lambda_j}} = \frac{1}{1+\lambda_j}(\varepsilon_s + \lambda_j \varepsilon_p) \tag{3.30}$$

where $\varepsilon_{s,p}$ are the energies of the s and p orbitals, respectively. The type of hybridization is given by λ whose determination is possible through a certain orthonormality relation which goes well beyond the scope of this Lecture Notes. Let's consider the structure of graphene, in which each carbon atom is at the vertex of three bonds at 120° to each other forming a hexagonal network. The atomic wave functions of Eq.(3.28) are transformed into three hybrid orbitals sp^2:

$$\begin{aligned}
|sp^2 1\rangle &= \frac{1}{\sqrt{3}}\left\{|s\rangle + \sqrt{2}\,|p_x\rangle\right\} \\
|sp^2 2\rangle &= \frac{1}{\sqrt{3}}\left\{|s\rangle - \frac{1}{\sqrt{2}}\,|p_x\rangle + \sqrt{\frac{3}{2}}\,|p_y\rangle\right\} \\
|sp^2 3\rangle &= \frac{1}{\sqrt{3}}\left\{|s\rangle - \frac{1}{\sqrt{2}}\,|p_x\rangle - \sqrt{\frac{3}{2}}\,|p_y\rangle\right\}
\end{aligned} \tag{3.31}$$

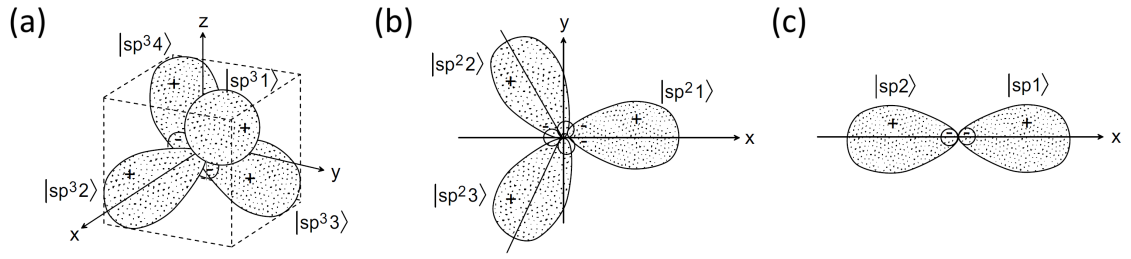


Figure 3.6: Types of sp^λ hybridization: (a) four sp^3 hybrids pointing to the corners of an ideal tetrahedron; (b) three sp^2 hybrids with angles of 120° to each other in a plane; (c) two sp hybrids pointing in opposite directions along the same axis.

These orbitals are represented in Fig.3.6(b) and give rise to σ interactions among atoms. Another interesting example is the one of silicon. It is a tetrahedrally coordinated solid, in which the four bonding directions are symmetrically equivalent. The angles between two Si atoms are equal to 109°28′ and give rise to a sp^3 hybridization. Therefore, the s and p atomic wave functions of Eq.(3.28) are transformed into four sp^3 hybrids:

$$\begin{aligned}
|sp^3 1\rangle &= \frac{1}{2}\left\{|s\rangle + |p_x\rangle + |p_y\rangle + |p_z\rangle\right\} \\
|sp^3 2\rangle &= \frac{1}{2}\left\{|s\rangle + |p_x\rangle - |p_y\rangle - |p_z\rangle\right\} \\
|sp^3 3\rangle &= \frac{1}{2}\left\{|s\rangle - |p_x\rangle + |p_y\rangle - |p_z\rangle\right\} \\
|sp^3 4\rangle &= \frac{1}{2}\left\{|s\rangle - |p_x\rangle - |p_y\rangle + |p_z\rangle\right\}
\end{aligned} \tag{3.32}$$

These orbitals are sketched in Fig.3.6(a). The shape of the orbitals is what determines the cleavage plane of a particular Si surface. For example, Si(111) has a natural cleavage plane, i.e. the (111), because of the dangling bonds exposed to the vacuum which are the ones associated with the $|sp^3 1\rangle$ orbital.

3.2.2 Bonds and Bands

The impact of hybridization and the resulting hybrid interactions on permissible electronic states can be illustrated in tetrahedrally coordinated systems, such as covalent or partially ionic semiconductors that crystallize in diamond or zinc-blende structures[2]. Since two atoms A and B are in one unit cell, one finds that the Hamiltonian is an 8×8 matrix which can be further divided into 4 blocks:

$$H_{abij} = \begin{bmatrix} D_A & I \\ I & D_B \end{bmatrix} \tag{3.33}$$

where D_A and D_B represent the energy levels of the A and B atoms:

$$D_A = \begin{bmatrix} \varepsilon_s^A & 0 & 0 & 0 \\ 0 & \varepsilon_{p_x}^A & 0 & 0 \\ 0 & 0 & \varepsilon_{p_y}^A & 0 \\ 0 & 0 & 0 & \varepsilon_{p_z}^A \end{bmatrix}, \qquad D_B = \begin{bmatrix} \varepsilon_s^B & 0 & 0 & 0 \\ 0 & \varepsilon_{p_x}^B & 0 & 0 \\ 0 & 0 & \varepsilon_{p_y}^B & 0 \\ 0 & 0 & 0 & \varepsilon_{p_z}^B \end{bmatrix} \tag{3.34}$$

[2]The zinc-blende structure, also called diamond cubic, features a face-centred cubic lattice with two interpenetrating lattices, e.g. GaAs, ZnS.

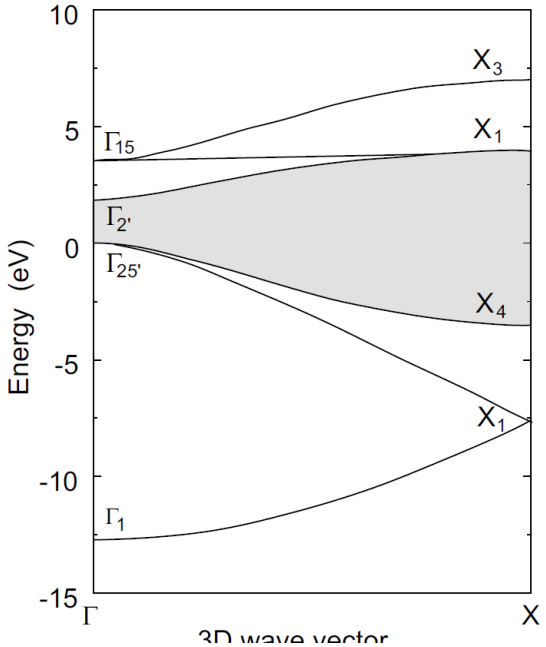


Figure 3.7: Band structure of a silicon crystal within the first-nearest-neighbor approximation and using an sp^3 basis set. The region of the fundamental gap is shaded.

while I is the interaction matrix. In particular, this one assumes a particularly complicated form:

$$I = \begin{bmatrix} E_{ss}g_1^*(\boldsymbol{k}) & -E_{sp}g_2^*(\boldsymbol{k}) & -E_{sp}g_3^*(\boldsymbol{k}) & -E_{sp}g_4^*(\boldsymbol{k}) \\ E_{sp}g_2^*(\boldsymbol{k}) & E_{xx}g_1^*(\boldsymbol{k}) & E_{xx}g_4^*(\boldsymbol{k}) & E_{xy}g_3^*(\boldsymbol{k}) \\ E_{sp}g_3^*(\boldsymbol{k}) & E_{xy}g_4^*(\boldsymbol{k}) & E_{xx}g_1^*(\boldsymbol{k}) & E_{xy}g_2^*(\boldsymbol{k}) \\ E_{sp}g_4^*(\boldsymbol{k}) & E_{xy}g_3^*(\boldsymbol{k}) & E_{xy}g_2^*(\boldsymbol{k}) & E_{xx}g_1^*(\boldsymbol{k}) \end{bmatrix} \tag{3.35}$$

where the $g_j(\boldsymbol{k})$ $(j = 1, 2, 3, 4)$ are phase parameters related to the Bloch wave functions, while the E terms describe the first-nearest-neighbour interactions. The LCAO Hamiltonian matrix of Eq.(3.33) for the zinc-blende structure can be easily diagonalized for certain high-symmetry points and directions in the bulk fcc Brillouin zone. In the case of an A=B zinc-blende structure, e.g. silicon, the eigenvalues are:

$$\varepsilon_{\text{p-like}} = \varepsilon_p \pm \left[E_{xx}^2 \cos^2\left(k\frac{d}{4}\right) + E_{xy}^2 \sin^2\left(k\frac{d}{4}\right) \right]^{\frac{1}{2}} \tag{3.36}$$

where d is the lattice parameter, $d = 2.35$ Åin the case of Si. For the sp-like bands one gets:

$$\begin{aligned} \varepsilon_{\text{sp-like}} = & \frac{1}{2}\left[\varepsilon_s + \varepsilon_p \pm (E_{ss} + E_{xx}) \cos\left(k\frac{d}{4}\right) \right] \pm \\ & \pm \left\{ \frac{1}{4}\left[\varepsilon_p - \varepsilon_s \pm (E_{xx} - E_{ss}) \cos\left(k\frac{d}{4}\right) \right]^2 + E_{sp}^2 \sin^2\left(k\frac{d}{4}\right) \right\}^{\frac{1}{2}} \end{aligned} \tag{3.37}$$

What we are most interested in about this bunch of equations is the fact that the energy varies continuously with the $\boldsymbol{k}$ vector which in turn means that bands have been created. The dispersion relations we have just seen can be plotted for the case of Si and give rise to the band diagram of Fig.3.7. In the diagram, we notice the presence of an energy gap (shaded region). However, whereas the valence bands are reasonably described by the first-nearest-neighbor approximation, the conduction bands show features in disagreement with experimental observations. In particular, such an energy diagram describes a direct gap

semiconductor while experiments show that Si is an indirect gap semiconductor. This is due to the tight-binding model, i.e. the first-nearest-neighbor approximation, which does not keep into account higher-order interaction terms.

3.3 Surface States

Surface states refer to electronic states that are localized near the surface of a material. These states arise due to the breaking of translational symmetry at the surface, leading to electronic energy levels that differ from those in the bulk of the material. There exist different types of surface states:

- **Intrinsic Surface States**: They refer to electronic states that arise purely due to the presence of a surface in a material, without the influence of any external interfaces or impurities. These states are intrinsic to the material's surface and originate from the breaking of translational symmetry at the surface. Among these, we distinguish:
  - **Tamm States**: They are the typical surface states of semiconductor materials and arise due to the abrupt termination of the crystal lattice at the interface, leading to the confinement of electronic states.
  - **Shockley States**: They are the typical surface state of metals and originate from the mismatch between the electronic structure of the bulk material and that of the surface layer.

  It is worth noting that there is no real physical distinction between the different terms; only the mathematical approach is different.
- **Extrinsic Surface States**: They are electronic states that arise at the surface of a material due to the presence of external factors such as impurities, defects, adsorbates, or interfaces with other materials. Unlike intrinsic surface states, which originate solely from the breaking of translational symmetry at the surface, extrinsic surface states are influenced by external perturbations. Extrinsic surface states can form due to various mechanisms:
  - **Adsorbates**: Molecules or atoms adsorbed onto the surface can introduce additional electronic states through chemical bonding or charge transfer interactions.
  - **Impurities and Defects**: Foreign atoms incorporated into the crystal lattice or structural defects near the surface can introduce localizd electronic states within the band gap of the material.

In the following, we will briefly discuss the properties of such states by starting with intrinsic surface states.

3.3.1 Intrisinc Surface States

If one wants to properly solve the problem of finding the energies and the wave function of the electrons on a surface, it is necessary to proceed by solving the Schrödinger equation for such a system. In the context of the nearly-free electron model for a 1D chain of atoms, we assume a cosine variation of the potential along the chain as depicted in Fig.3.8(a). We will not deal with the calculation[3] and we will limit ourselves in observing the two possible solutions of such a system. The first solution, depicted in the upper part of Fig.3.8(b), is the standard solution for the wave function of the electrons on a crystal (Bloch wave

[3]See Section 6.1 of *Solid Surfaces, Interfaces and Thin Films* by Hans Lüth.

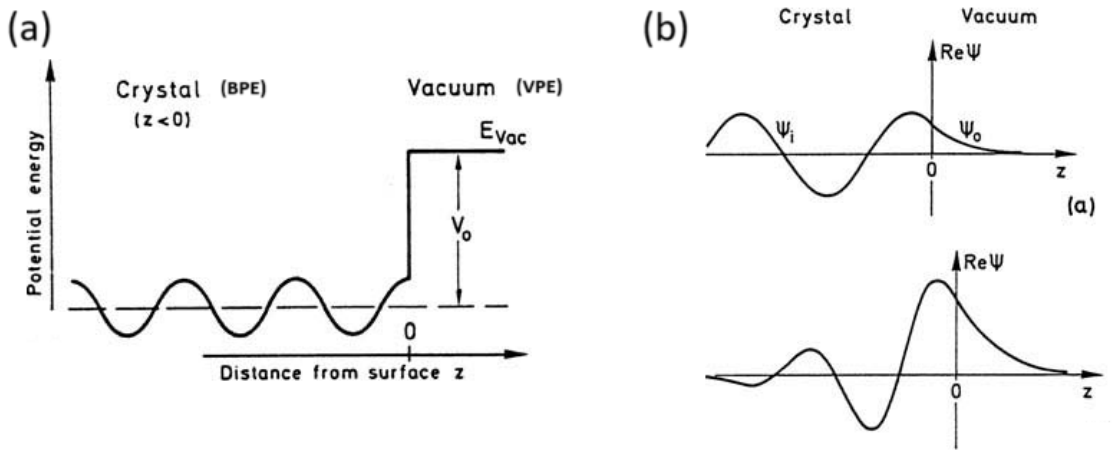


Figure 3.8: (a) Nearly-free-electron model for a cosine potential along a linear chain (z-direction). (b) Real part of the one-electron wavefunction, $\mathfrak{Re}(\Psi)$, for a standing Bloch wave (Ψ_i), matched to an exponentially decaying tail (Ψ_0) in the vacuum above and a surface-state wave function localized at the surface ($z = 0$) below.

function) which is combined with an exponential decay wave function which denotes the presence of a surface. Although correct, such a solution is not what we are looking for. The second possible solution of our system arises when we allow the wave vector of electrons to be complex. In this case, the wave function results to be very localized at the interface denoting the presence of some available states for the electrons at the surface. These are the so-called surface states. So far we have discussed the existence of surface states on ideal, clean surfaces in the framework of the nearly-free-electron model. For historical reasons, such surface states are often called **Schockley states**. One can also approach the question of the existence of electronic surface states from the other limiting case of tightly bound electrons. This approximate treatment in terms of wave functions that are linear combinations of atomic eigenstates was first given by Tamm. The resulting states are often called **Tamm states**, even though *there is no real physical distinction between the different terms*; only the mathematical approach is different. The existence of electronic surface states, whose energy is different from the bulk states, is qualitatively easy to see within the picture of tightly bound electrons (Linear Combination of Atomic Orbitals, LCAO). For the topmost surface atoms, the bonding partners on one side are missing in

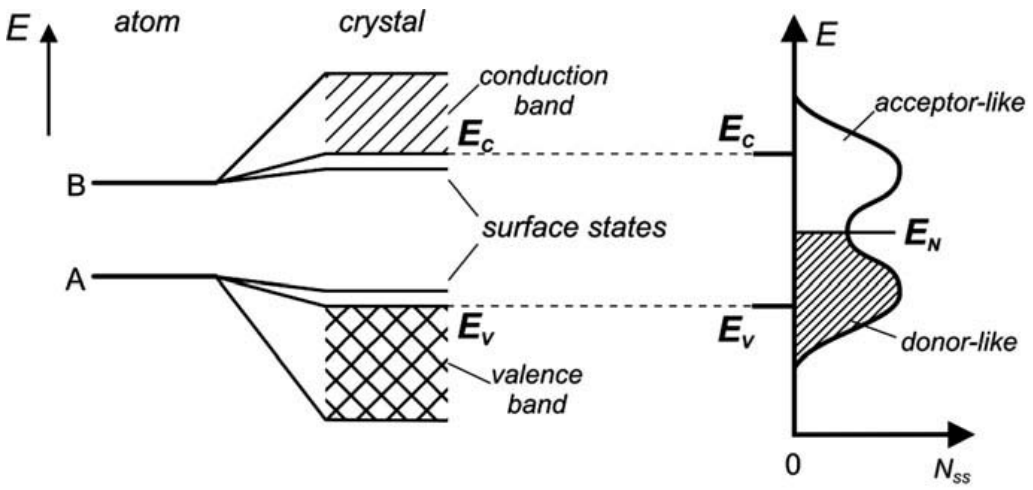


Figure 3.9: Qualitative explanation of the origin of surface states in the tight-binding picture.

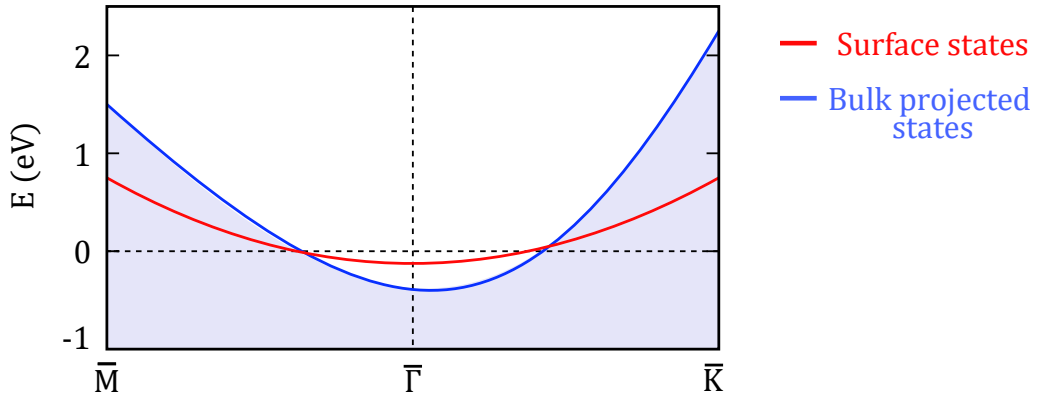


Figure 3.10: Surface bands and bulk projected bands for Ag(111).

total, which means that their wave functions have less overlap with the wave functions of neighbouring atoms. The splitting and shift of the atomic energy levels is thus smaller at the surface than in the bulk (Fig.3.9). Every atomic orbital involved in chemical bonding and producing one of the bulk electronic bands should also give rise to one surface-state level. The stronger the perturbation caused by the surface, the greater the deviation of the surface level from the bulk electronic bands. When a particular orbital is responsible for chemical bonding, e.g. the sp^3 hybrid in Si or Ge, it is strongly affected by the presence of the surface; bonds are broken, and the remaining lobes of the orbital stick out from the surface. They are the so-called dangling bonds. The energy levels of surface states are expected to be significantly shifted from the bulk values. Since surface atoms have fewer bonding partners than bulk atoms, they give rise to electronic energy levels that are closer to those of the free atoms, i.e. surface state levels are split off from the bulk bands. Depending on their origin, these states have **acceptor-** or **donor-like** charging character.

3.3.2 Band Diagrams

We can start observing how surface states appear on surfaces from an experimental point of view. Fig.3.10 depicts the bulk projected states, i.e. the electronic band structure of the bulk, for Ag(111). In red, instead, we observe the dispersion relation for the surface states of silver which are, since we are talking about a metal, Schockley states. We note that around the Γ point the surface states are proper of the surface only, i.e. no bulk states are simultaneously present in the same portion of the graph. In this case, the energy of the Schockley states around the minimum can be approximated with a parabolic dispersion

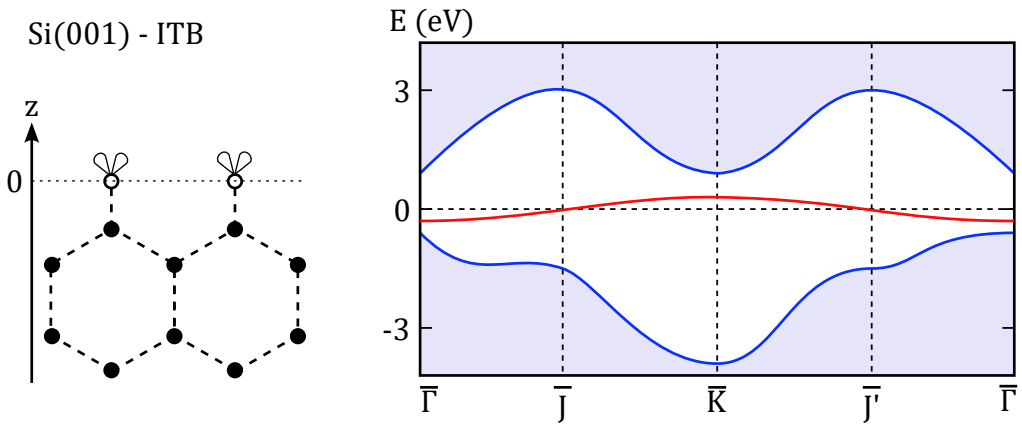


Figure 3.11: Surface bands and bulk projected bands for Si(001) in the ideal truncated bulk (ITB) case.

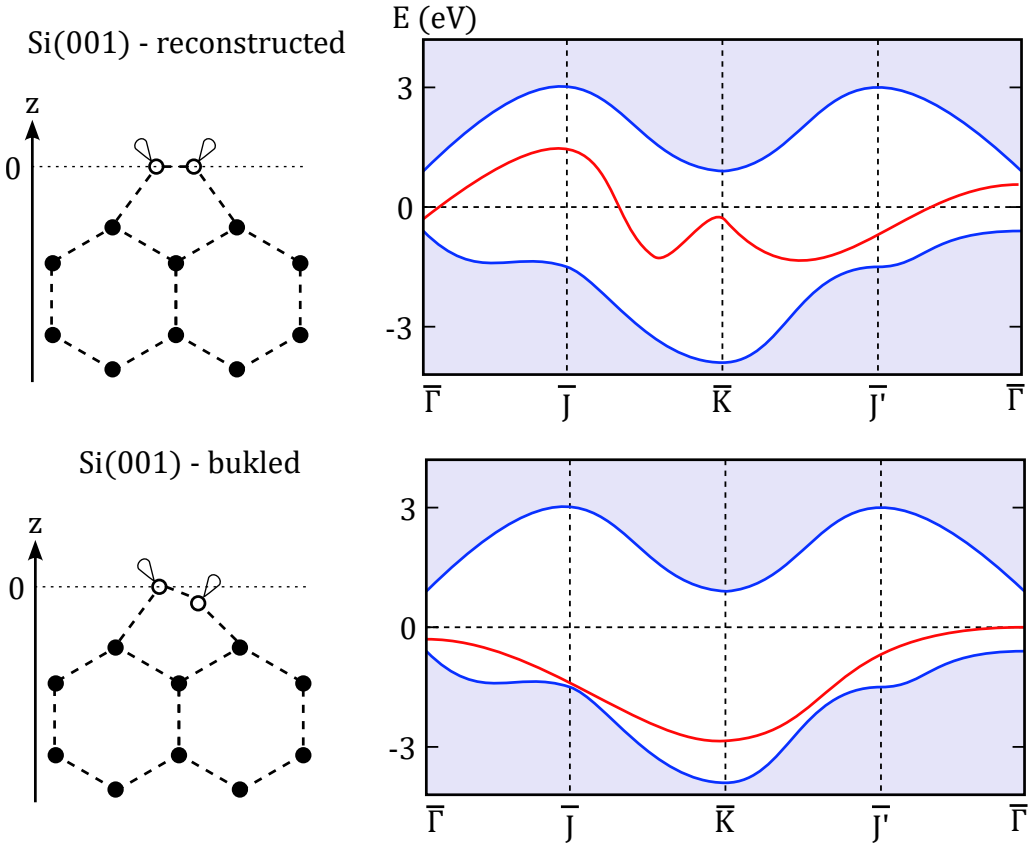


Figure 3.12: Surface bands and bulk projected bands for Si(001) in the reconstructed (a) and the buckled dimer (b) case.

relation, so that:

$$\varepsilon = \varepsilon_{min} + \frac{\hbar^2 k_{xy}^2}{2m^*} \tag{3.38}$$

It is worth noting that in the figure the crystallographic directions are marked with an upper bar on the letter. This is to stress the fact that this dispersion relation refers to surface bands, the one of the bulk is indeed projected. Let's now have a look at Fig.3.11 which shows the band diagram for Si(001) in the case in which no reconstruction occurred, i.e. for the ideal truncated bulk case. In this case, the surface state (red) is fully contained in the gap of silicon, implying that these states can be only occupied by the atoms on the surface which are the ones depicted with their dangling bonds on the left side of the figure. Moreover, here the Tamm state is almost flat; this is a consequence of the fact that we are considering an ideal truncated bulk in which the atoms at the surface do not interact with any other atom of the bulk. Conversely, we can observe what is reported in Fig.3.12 considering two possible reconstructions for Si(001). As before, the surface states are always in the gap, but now, because of the reconstruction processes, the surface bands are no longer flat. The reconstructed model can be further distorted to become a buckled dimer. In fact, an electron from a half-filled orbital can jump to the other orbital, leaving one of them empty and the other filled. In turn, the full one, which is more negatively charged, is pushed towards the vacuum because of the electrostatic repulsion with the other electrons. This originates from the buckled (or buckling) reconstruction shown in the figure.

3.4 Tamm States in the Pandey-π Chain Model

The introduction of the concept of surface states allows us to quantitatively explain the features of the band dispersion relation for the surface states. Let's consider once again the case of Si(111)-(2×1) as a reference model. As a first attempt to find the dispersion relation, we can use a tight binding model to compute the energy of the atoms in the π-chain of silicon (see Fig.3.13(a)). Within this approach, if one considers only the first neighbours' interactions along the π-chain, it is possible to get the energy eigenvalues $\varepsilon_{\pm}$:

$$\varepsilon_{\pm} = \frac{1}{2}(\varepsilon_1 + \varepsilon_2) \pm \left[\frac{1}{4}(\varepsilon_1 - \varepsilon_2) + |H_{12}|^2 \cos^2\left(k_{xy}\frac{a_0}{2\sqrt{2}}\right)\right]^{1/2} \tag{3.39}$$

where ε_i are the energies of the atoms with dangling bonds and $a_0 = 5.4$ Åis the lattice parameter of silicon. The resulting band diagram for the Tamm states of silicon is shown in Fig.3.13(b) in which we notice:

- The band diagram does not represent a semiconductor surface because there is no gap between valance and conduction states.
- Along the $[11\bar{2}]$ direction the energy gap is zero as pointed out before. This is a consequence of the model used which neglects any interaction between atoms not belonging to the π-chain.
- Along the $[\bar{1}10]$ direction, i.e. along the π-chain, we have a symmetric dispersion relation due to the symmetric interaction occurring along the chain.
- The maximum and minimum energy of these surface bands are ± 1 eV which agrees with what was observed before. In fact, we know that the energy gap of silicon is $\varepsilon_g^{Si} \simeq 1.12$ eV and having such maximum and minimum energies implies that the surface states can be fitted and seen inside the bulk gap of Si as in Fig.3.12.

Nevertheless, we know that this cannot be correct. It is well known that silicon has a semiconductor surface and the gap opening must occur eventually. The gap opening of the surface states can be added ad hoc to the model even though nowadays theoretical calculations manage to open the gap naturally. Experimentally, the gap opening is obviously observed and an example is reported in Fig.3.14 in which data points were obtained by Angle-Resolved UV Photoelectron Spectroscopy (ARUPS) measurements. As a final ob-

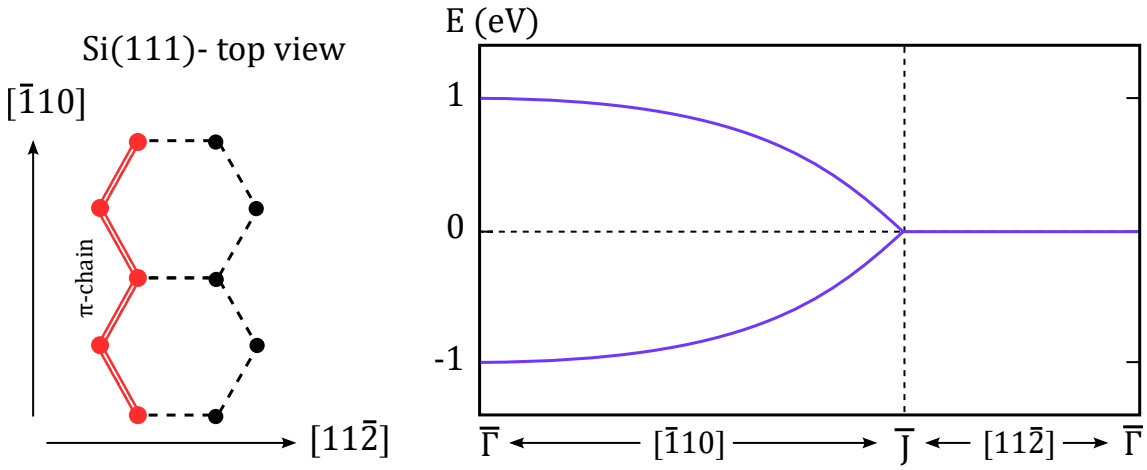


Figure 3.13: Dispersion of the surface-state bands for the Si(111)-(2×1) surface calculated with tight binding approach along the π-chain.

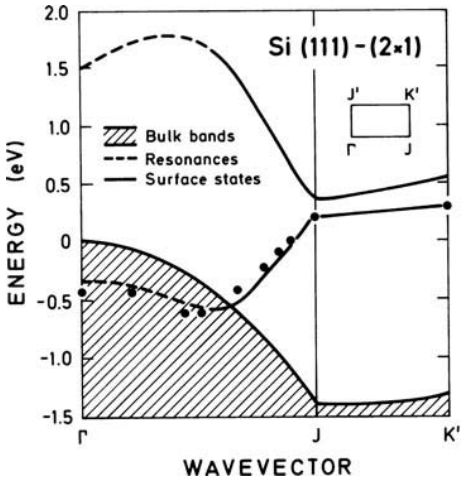


Figure 3.14: Dispersion of the surface-state bands for the Si(111)-(2×1) surface retrieved by ARUPS together with the projected bulk-band structure (shaded).

servation, it is worth spending a few words regarding the electron density along the π-chain which is a direct consequence of the presence of Tamm states. It is possible to observe that electrons are localized along the chain creating a sort of one-dimensional confinement. In particular, the filled surface states are strongly localized around the atoms while the empty ones have a more delocalized character.

3.5 Extrinsic Surface States

Whereas intrinsic surface states are related to the existence of a perfect surface, the present extrinsic surface states only arise as a result of perturbations to the ideal surface. As told, extrinsic surface states can be produced by the presence of defects or any other type of impurities, but in the following, we will focus on those extrinsic surface states produced by *adsorbed atoms*. The reason for this choice lay in the fact that the adsorption process is fundamental in the study of surfaces since any interface or confined structure starts its growth with an adsorption process and therefore it is worthed our attention. Moreover, adsorption causes changes in the chemical bonds near the surface, thus affecting the distribution of intrinsic surface states. In addition, new electronic states are formed by the bonding and antibonding orbitals between the chemisorbed atom or molecule and the surface. In general, when talking about adsorption we distinguish two main processes that can occur when a molecule is adsorbate by a surface:

- **Physisorption**: it is a process in which the electronic structure of the molecule or atom is hardly perturbed upon adsorption. The corresponding mechanism in molecular physics is van der Waals bonding (e.g. dipole-dipole interaction). The attractive force is due to correlated charge fluctuations in the two bonding partners, i.e. between mutually induced dipole moments.
- **Chemisorption**: it is an adsorption process that resembles the formation of covalent or ionic bonds in molecular physics; the electronic structure of the bonding partners is strongly perturbed, new hybrid orbitals are formed and, as in the case of ionic bonding, there may be charge transfer from one partner to the other.

Despite both the processes having their relevant features, in the following, we will focus our attention on the chemisorption mechanism since it is the most relevant for what concerns the modification of the electronic properties of a surface. For completeness's sake, it is

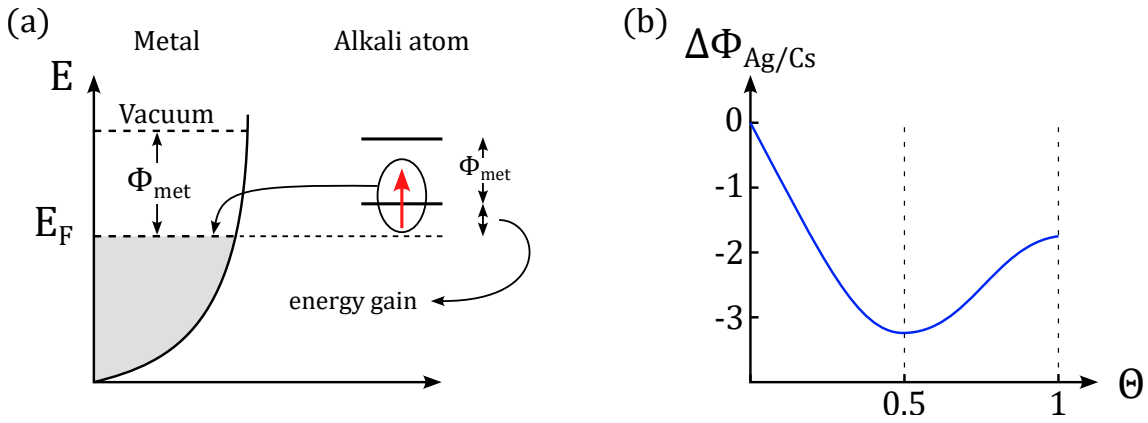


Figure 3.15: (a) Schematic of the ionic chemisorption process. (b) Difference in the work functions of silver and caesium $\Delta\phi_{Ag/Cs}$ as a function of the caesium coverage Δ.

interesting to note that the main physisorption mechanism is the dipole-dipole interaction. When a polar molecule approaches a surface during physisorption, its dipole interacts with the dipoles of the surface molecules. If the surface is also polar, there can be a significant dipole-dipole interaction between the adsorbate (the molecule being adsorbed) and the surface. This interaction plays a crucial role in the adsorption process, helping to stabilize the adsorbate on the surface. However, it's important to note that dipole-dipole interactions are relatively weak compared to chemical bonds, so they typically contribute to the reversible nature of physisorption.

3.5.1 Chemisorption

Chemisorption can be classified into different types based on the nature of the chemical bonds formed between the adsorbate and the surface. The main types of chemisorption include:

- **Ionic Chemisorption**: it occurs when ions in the adsorbate molecule interact with oppositely charged ions on the surface, resulting in the formation of ionic bonds. This type of chemisorption is common in systems where one component is highly polar or charged, such as metal oxides interacting with polar molecules.
- **Covalent Chemisorption**: it involves the sharing of electrons between the adsorbate and the surface atoms to form covalent bonds. This type of chemisorption is often observed on surfaces with unsaturated bonds or dangling bonds, such as certain semiconductor surfaces.

Let's start with the former.

Ionic Chemisorption

We will refer to a particular example which is the alkali atoms deposition on a metal surface. This case is of significant interest in surface science and catalysis. Alkali metal atoms (Li, Na, K, Rb, Cs and Fr), are known for their unique properties, including low ionization energies, i.e. low work function, and strong reducing abilities. When deposited onto metal surfaces, alkali atoms can significantly influence surface properties and catalytic activity due to their interactions with the surface atoms and molecules. For example, in thermionic emitters, the deposition of alkali atoms on metals is used to lower the work function of the metal and ease the emission process (e.g. tungsten filament covered by caesium). The first theoretical studies about the chemisorption was developed by R. W. Gurney in 1951. The

Gurney model is a simple empirical model used to describe the kinetics of chemisorption on solid surfaces. The model provides a theoretical framework for understanding the rate at which molecules adsorb onto a surface during chemisorption processes. We won't deal with the coverage process and we will focus our attention on the electron transfer between the metal and the alkali atom. Fig.3.15(a) schematise what happens when an alkali atom is adsorbed by a metal surface. Beacuse of their only one outermost electron, for alkali atoms is energetically favourable to transfer this electron into the band of the metals. As a consequence of this charge transfer, the dipole associated with the metal surface is modified causing, in turn, a perturbation of the Friedel oscillations (see Fig.3.2). This picture of the valence electron going from the alkali atom to the metal can explain, at least as a first approximation, the experimental behaviour depicted in Fig.3.15(b). The graph shows the difference between the work function of silver ($\phi_{Ag} \simeq 4.26\ eV$) and the one of caesium ($\phi_{Cs} \simeq 2.14\ eV$) as a function of the coverage Θ of caesium on the silver substrate. We notice that the work function difference decreases linearly up to a minimum. The linear decrease can be explained in terms of energy gain per caesium atom each time that an electron is lost to go on the metal. At the minimum, the work function of the Ag/Cs bilayer is almost equal to the one of caesium. The reason for the presence of a minimum and the following increase is not easy to explain and goes well beyond the scope of this course.

Covalent Chemisorption

As previously, we will refer to a particular example which is the one of the so-called Byholder Model. The Blyholder model, proposed by Gary Blyholder in 1964, is a theoretical model that describes the chemisorption of carbon monoxide (CO) on transition metal surfaces which in our case will be the one of platinum. To explain the chemisorption mechanism beyond this model, it would be necessary to involve the concept of hybridisation and some other complicated and chemistry-related concepts. In the following, we will try to have a qualitative understanding of the process, referring in particular to Fig.3.16. The covalent bonding between the molecule and the atoms of the surface creates a series of hybrid states between the two that can interact and give rise to different situations. In the first case (Fig.3.16(a)), the filled states of the molecule interact with the empty ones of the metal creating bonding and antibonding hybrid orbitals. The presence of the bonding orbital, lower in energy, results in a charge transfer from the molecule to the hybrid states Conversely, if the filled states are the one of the metal we obtain the situation in Fig.3.16(b) in which the charge transfer occurs in the opposite direction: from the metal to the hybrid states. There exists another possibility. If a filled-filled states interaction occurs and hybrid states are created we can obtain the type 3 and type 4 interactions whose final configurations resemble the type 1 and 2 we have already seen. If, in fact, we have an interaction between filled states (see Fig.3.16(c)), the charge transfer occur both from the metal and from the molecule to the hybrids. Then, since the antibonding orbital is higher in energy, the electrons will undergo a charge transfer once again, ending up in the valance band of the metal. As a result, then, we obtain a situation completely identical to the type 1 case in which, in the end, we have a filled bonding orbital and an overall charge transfer from the molecule to the hybrids. The opposite occurs when empty states interact between them. In this case, some electrons from the valence band of the metal will fill the bonding orbital created. Once again, the final configuration is indistinguishable from the one of type 2 in which the bonding orbital is filled and the charge transfer occurs from the metal to the hybrid states.

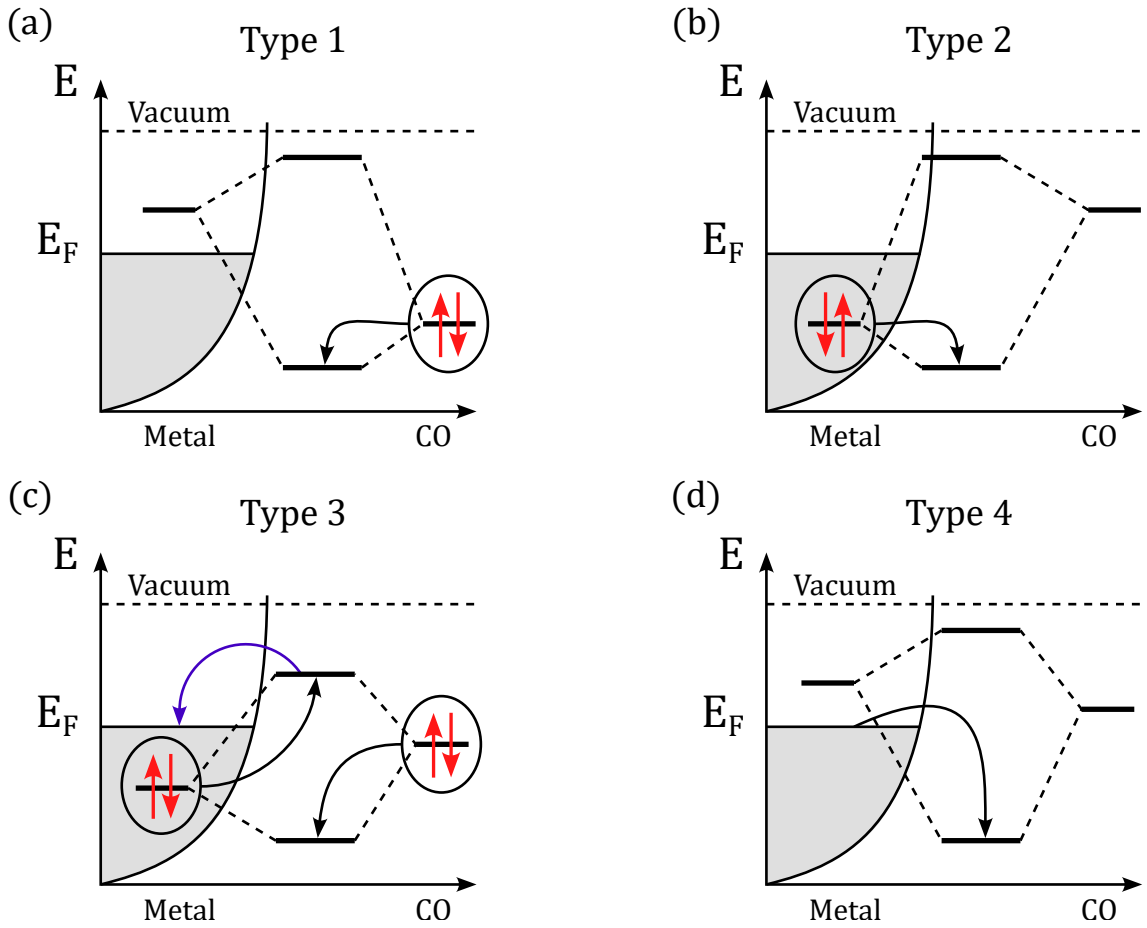


Figure 3.16: Different types of covalent chemisorption mechanisms depending on the hybridization state: type 1 (a), type 2 (b), type 3 (c) and type 4 (d).

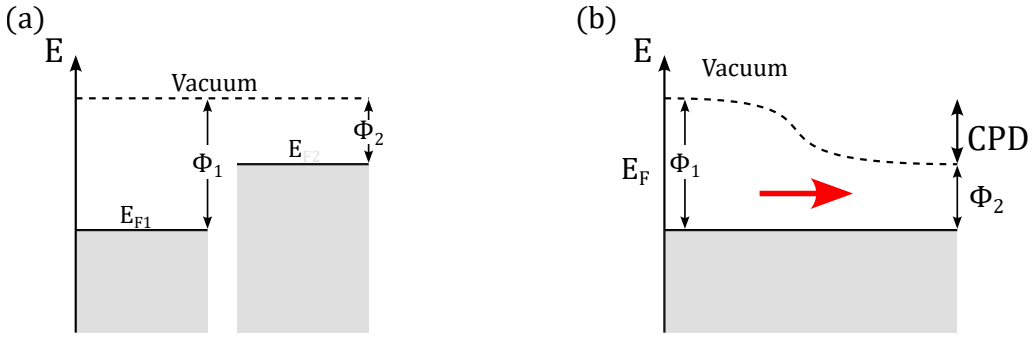


Figure 3.17: Schematic of the energy levels of two metals (a) when they are isolated and (b) when they are placed in contact.

3.6 Metal-Metal Junctions

We will now start talking about what happens when two materials are placed in contact one with the other starting from the case of a metal-metal junction (or interface). When two metals with two different Fermi levels are placed in contact, equilibrium implies that the Fermi level must be the same. This causes an electron transfer from the material with a higher work function to the one with a lover one (see Fig.3.17). This reorganisation of the charges makes the bands bend and establish a potential drop between the two metals. This potential drop is called **Contact Potential Difference** (CPD) or Volta Potential. The CPD is not an intrinsic property of the two bulk metals under consideration, but rather is determined by work function differences between the metals' surfaces. Like the work function, the Volta potential depends sensitively on surface state, contamination, etc. The study of metal-metal interfaces dates back to the XIX century, with the studies of Alessandro Volta. For completeness's sake, we report here the essence of Volta's work which can be summarised in his two laws:

- I Volta's Law: When different metals are arranged in a series based on their tendency to lose electrons, the electromotive force (emf) generated between any two metals in contact is determined by their positions in this series.

- II Volta's Law: When two different metals are brought into contact, a potential difference is established between them, which depends only on the nature of the metals and not on the presence of a third substance.

The CPD can be significant (of order 1 volt) but cannot be measured directly by an ordinary voltmeter. A voltmeter does not measure vacuum electrostatic potentials, but instead the difference in Fermi level between the two materials, a difference that is exactly zero at equilibrium. A particular procedure with the so-called *Kelvin probe apparatus* is necessary to measure the Volta Potential. Also, this procedure allows for the evaluation of the work function of metals. The experimental apparatus is shown in Fig.3.18. The sample is placed in contact with a golden tip, similar to the ones used in the Atomic Force Microscopy (AFM) experiments, and a piezoelectric system allows for the control of the distance between the tip and the sample. The tip-sample distance control is particularly important. In fact, the tip and sample in this configuration form a capacitor; in stationary conditions no current flows in the circuit, but if the tip-sample distance oscillates over time, thanks to the piezoelectric system, the variation causes a current to flow. The flowing of this current can be stopped by tuning the voltage between the tip and sample and this allows for the evaluation of the CPD. This can be seen mathematically. By definition, the

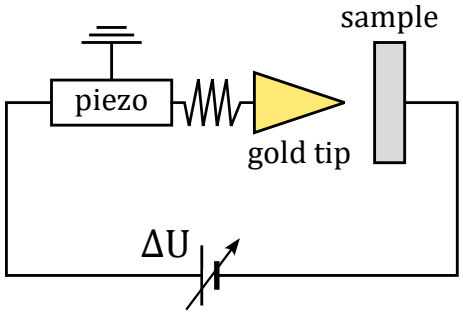


Figure 3.18: Schematic of the Kelvin apparatus for the determination of the CPD.

tip-sample CPD can written as:

$$CPD\,|_{ST} = \frac{1}{e}\left(E^T_{vac} - E^S_{vac}\right) \tag{3.40}$$

where E_{vac} denotes the vacuum level and the apex "T" and "S" the tip and the sample, respectively. Moreover, the Fermi levels of the two are:

$$\begin{aligned} E^S_F &= E^S_{vac} - e\phi^S \\ E^T_F &= E^T_{vac} - e\phi^T \end{aligned} \tag{3.41}$$

The equivalence of the Fermi levels at the equilibrium implies:

$$E^S_{vac} - e\phi^S = E^T_{vac} - e\phi^T \Rightarrow \frac{1}{e}\left(E^T_{vac} - E^S_{vac}\right) = \phi^T - \phi^S \tag{3.42}$$

Therefore, we can write:

$$CPD = \phi^T - \phi^S = -(\phi^S - \phi^T) \tag{3.43}$$

In the following, we will use $\Delta\phi = \phi^S - \phi^T$. Now, if we consider the tip-sample system as a capacitor of capacitance C, the charge accumulated on it, considering also the potential drop ΔU of the power supply, is:

$$Q = C\,(CPD + \Delta U) = C\,(-\Delta\phi + \Delta U) \tag{3.44}$$

If the system oscillates, i.e. is not stationary, current flows and we can compute the value I of such a current:

$$I = \frac{dQ}{dt} = \frac{d}{dt}\left\{C\,(-\Delta\phi + \Delta U)\right\} \tag{3.45}$$

Because of the oscillations, the capacitance C will change with time, but as a first-order approximation, we can assume the potential drop across the capacitor to be constant. So that:

$$I = \left(\frac{dC}{dt}\right)(-\Delta\phi + \Delta U) \tag{3.46}$$

The current stops flowing when $-\Delta\phi + \Delta U = 0$ regardless the value of the derivative of the capacitance. Therefore, by properly tuning the voltage ΔU we can find the value needed to stop the current flows which will be equal to the difference in work functions of the metals which in turn is equal to the CPD. Moreover, since gold is a well-known material and ϕ^T is well-known as well, we can have a direct measure of the work function of the sample. This procedure is extremely local and precise beacuse of the small dimension of the tip which can map on a very small scale the surface of the sample, also finding small variations of the work function along the surface, possibly due to defects or other impurities.

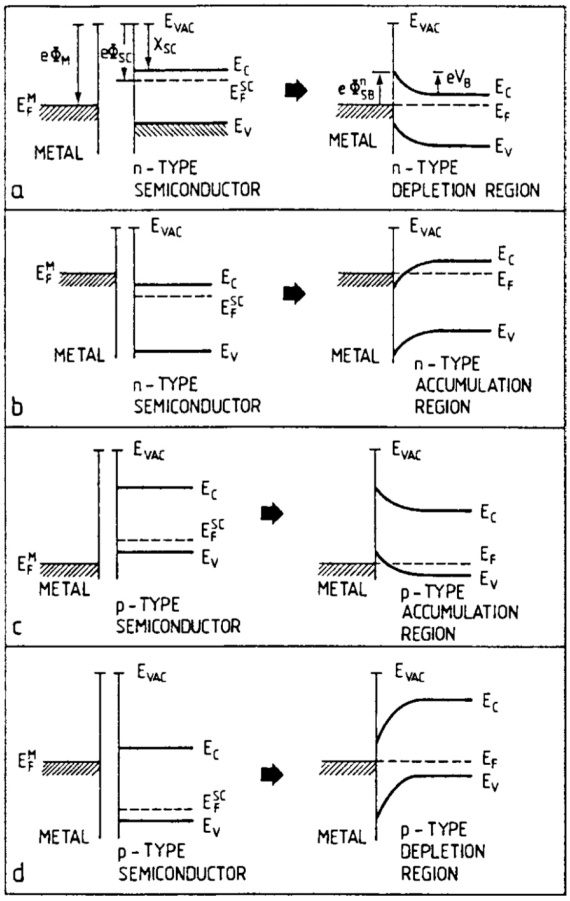


Figure 3.19: Schematic diagrams of band bending before and after metal-semiconductor contact.

3.7 Metal-Semiconductor Junctions

Metal-semiconductor (M-S) junctions are found in every metal contact to a semiconductor device and the underlying physics is also the basis for rectifying devices. Therefore, studying these junctions is fundamental for scientific and technological purposes. The observations about such kinds of structures date back to 1874 with the studies of Karl Ferdinand Braun. He studied the electrical properties of M-S junctions distinguishing two different types of junctions based on the rectifying behaviour[4] of the junction: rectifying ones and non-rectifying (or Ohmic) ones. As a first simplified approach, one might apply arguments based on the work functions of the metal and the electron affinity of the semiconductor, to obtain some general considerations; this was the approach used by Schottky. In thermal equilibrium, the Fermi levels in the two materials must be aligned. Depending on the difference between the work function ϕ_m of the metal and the electron affinity χ_s of the semiconductor (or equivalently the "work function" of the semiconductor, i.e. the difference between the Fermi and the vacuum level ϕ_s), different situations may arise, as shown in Fig.3.19. Also in this case, when the two materials are brought into contact, the matching of the Fermi levels causes charges to flow from one side to the other, and a dipole layer is built up at the interface. Beacuse of the high mobility of electrons in the metal, the charge is screened in a negligible amount of space. Conversely, in the semiconductor, the shielding is much less effective thus the space charge usually extends hundreds of Ångstroms into the crystal. As a result, in the semiconductor, a normal space-charge

[4]The rectifying behaviour of a metal-semiconductor junction refers to its ability to allow current to flow more easily in one direction than the other. This characteristic makes it function similarly to a diode.

layer, depletion or accumulation, is formed. The potential barrier formed at the interface is called the Schottky barrier. For example, for the high-work-function metal and n-type semiconductor depicted in Fig.3.19(a), electrons flow from the semiconductor to the metal after contact, depleting a characteristic surface region in the semiconductor of electrons. The upward band bending in this depletion layer is correlated via Poisson's equation with the positive space charge of ionized donors. The maximum band bending at the interface is related to the Schottky barrier which has to be overcome when an electron is excited from the metal into the conduction band of the semiconductor. We can quickly show the calculation. Starting with:

$$\nabla^2 V = -\frac{\rho}{\varepsilon} \tag{3.47}$$

By approximating the charge density with the donor density N_D (n-type), in a 1D model along x one gets:

$$\frac{d^2V}{dx^2} = -\frac{eN_D}{\varepsilon_r \varepsilon_0} \tag{3.48}$$

where ε_r is the relative dielectric constant of the semiconductor. By denoting with d the extension of the depletion layer and setting $x = 0$ at the interface, we can integrate and get:

$$V(x) = \frac{eN_d}{2\varepsilon_r \varepsilon_0} (x - d)^2 \tag{3.49}$$

The Schokkly barrier corresponds to the potential at the coordinate $x = 0$, therefore:

$$\phi_{sb} = V(0) = \frac{eN_d}{2\varepsilon_r \varepsilon_0} d^2 \tag{3.50}$$

Moreover, by supposing $eN_D \simeq en_{bulk}$ where n_{bulk} is the carrier density in the bulk, we get:

$$d = \sqrt{\frac{2\varepsilon_r \varepsilon_0 V(0)}{en_{bulk}}} \tag{3.51}$$

This approach in terms of work functions and electron affinities was first applied by Schottky to understand the rectifying action of metal-semiconductor junctions. However, experiments show that the height of the Schottky barrier is less dependent on the work functions than expected. This can be seen in Fig.3.20(a) which shows measured Schottky barrier heights ϕ_{sb} for metals with various work functions deposited on UHV-cleaved n-type Si(111)-(2×1) surfaces. As we can see from the figure, a linear behaviour has been observed but with a smaller slope than the one predicted by Schottky, denoting a weaker dependence on the work function of the metal. This observation makes us conclude that some other factor must be affecting the height of the Schottky barrier. The first step in explaining the deviation of experimental data from the Schottky model was made by Bardeen and mostly relies on the presence of surface states on the semiconductor.

3.7.1 Fermi Level Pinning and Bardeen Hypothesis

Before dealing with the Bardeen hypothesis, we must take a step back and consider the so-called **Fermi level pinning** phenomena. Fermi level pinning refers to the situation where the Fermi level at the semiconductor surface becomes fixed or "pinned" at a certain energy level within the band gap, irrespective of the doping level of the bulk semiconductor. This occurs because when a metal contacts a semiconductor, electrons can transfer between the semiconductor and the surface states. These surface states can trap charge carriers, creating a large density of states at specific energy levels within the band gap. The exchange of electrons between the bulk semiconductor and the surface states continues

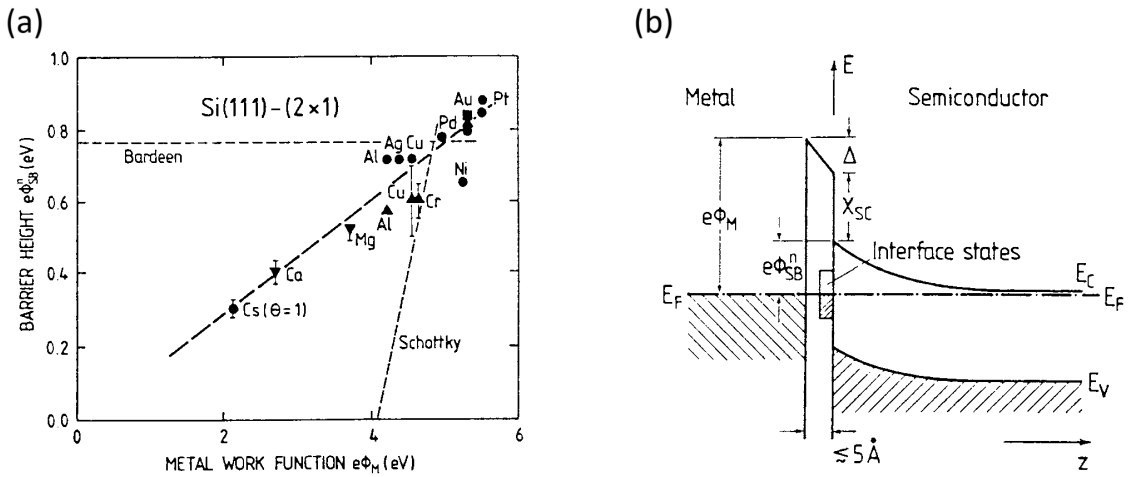


Figure 3.20: (a) Barrier heights ϕ_{sb} of Si-Schottky contacts (Si n-type) versus metal work function ϕ_m. (b) Band diagram of a metal-semiconductor (n-type) junction in which the interface region is taken into account explicitly. The formation of new interface states of sufficient density pins the Fermi-level.

until equilibrium is reached, causing the Fermi level at the surface to align with the energy level of these surface states (see Fig.3.20(b)). Therefore, Bardeen assumed that the surface states of the clean semiconductor surface persist under the metal overlayer and that they pin the Fermi level. The work function of the deposited metal would thus have no effect on the position of the Fermi level at the interface (see horizontal dashed line in Fig.3.20(a)). It is possible to show that for metal coverages in the monolayer range surface-state densities as low as 10^{12} cm^{-2} could pin the Fermi level at a fixed position. The assumptions of the Bardeen model are, of course, incorrect, since the surface states of the clean surface are strongly affected by metal deposition. Nevertheless, this hypothesis represents a good starting point to explain the deviation from the Schokkly model.

3.7.2 Allen and Gobeli Experiment

The Allen and Gobeli experiment was pivotal in demonstrating the concept of Fermi-level pinning at semiconductor surfaces due to surface states. By using photoelectric measurements on clean surfaces of germanium and silicon, they showed that the Fermi level is pinned at a specific energy within the band gap, largely independent of the bulk doping levels. The samples they used in the experiment were hyper-doped semiconductors, both n and p-types, with doping levels up to 10^{20} cm^{-3}. In the case of the Si(111)(2x1) sample, for example, such doping levels are really difficult to obtain and the resulting semiconductor exhibits a strong metallic behaviour. This changes also the main reconstruction mechanism which becomes relaxation. Moreover, for such a hyper-doped semiconductor controlling the cleavage process and obtaining a well-determined surface is not an easy task because of the changes in the reconstruction mechanism. Nevertheless, Allen and Gobeli managed to perform some relevant measurements whose results for silicon are shown in Fig.3.21(a). The graph shows the work function of silicon as a function of the doping level. The constancy of the work function is a clear demonstration of the pinning of the Fermi level due to the surface states which makes the Fermi level practically independent from the doping levels.

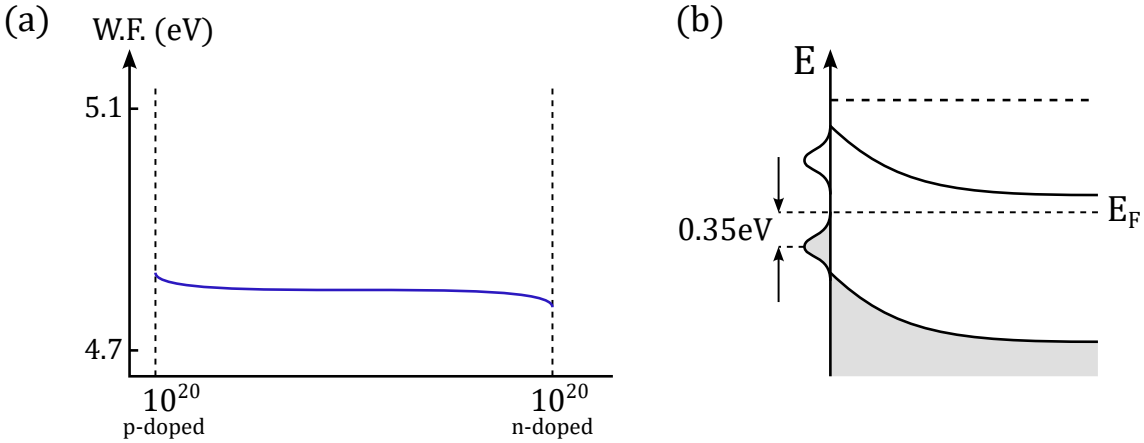


Figure 3.21: (a) Work function of the cleaved Si(111)-(2 × 1) surface as a function of bulk doping (p- and n-type) (b) Si(111)(2x1) band diagram after Fermi level pinning with surface states.

3.8 Chiarotti Experiment and Optically Induced Transistions

Despite this represents clear evidence of the existence of the surface states and of the Fermi level pinning, we want to obtain further proof of the existence of such states by observing transitions among these states (see Fig.3.21(b)). To do that one has to strongly control the penetration of light inside the sample. If one uses visible light, for instance, with a wavelength in the order of 500 nm, it would be impossible to surface states since the wavelength is extremely longer than the surface contribution which is in the order of 0.1 nm. The problem was solved by Chiarotti *et al.* by using a technique called **Surface Differential Reflectivity** (SDR). SDR measures the difference in reflectivity between a clean reference surface and the surface of interest. By comparing the reflectivity of these surfaces, it isolates the contributions from the surface layer, eliminating bulk effects. The experimental apparatus is shown in Fig.3.22(a). A beam splitter (BS) is used to split the light coming from the source into two, one going to the real sample and the other to the reference surface. Shutters are used to select alternatively one of the two paths. In order to avoid any spurious light contributions affecting the detector detection, a chopper is used. A chopper is a crucial component used to modulate the light source at a specific frequency. This modulation enhances the signal-to-noise ratio and allows for phase-sensitive detection, which can significantly improve the accuracy and sensitivity of the measurements. A typical optical chopper consists of a rotating disk with one or more apertures or slits. This disk is usually mounted on a motor that spins at a controlled speed. The apertures periodically

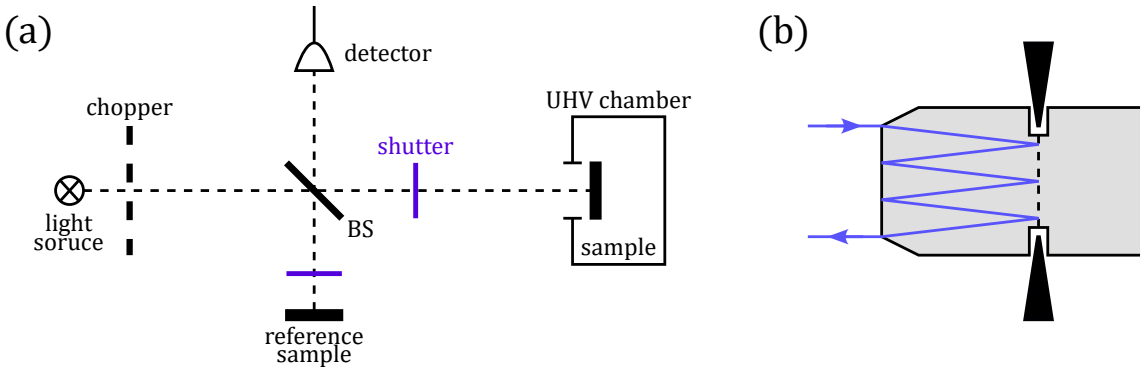


Figure 3.22: (a) Schematic of the Surface Differential Reflectivity experimental apparatus. (b) Internal reflection in the SDR sample.

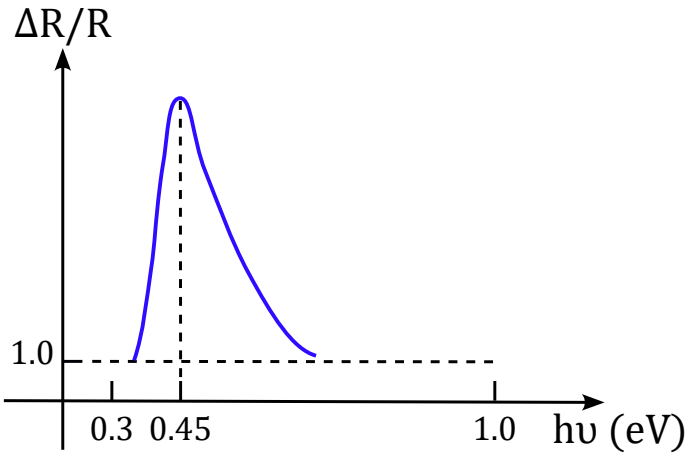


Figure 3.23: Result of the Chiarotti experiment: differential reflectivity as a function of the photon energy.

allow the passage of light through the disk, thereby converting a continuous light beam into a pulsed or modulated beam. The Si(111)(2x1) sample to study is cleavaged in a UHV chamber while the reference surface is the oxide silicon surface. Oxidation, in fact, removes any surface states present on the surface thanks to the oxygen bonds. Therefore, any difference in reflectivity between the two surfaces must be due to surface states since they only differ in the presence of these states. To further increase the effect of the presence of surface states, internal reflection can be used as shown in Fig.3.22(b). Knives in UHV ambient cut the Si crystal and the structure that results from the cut can reflect light multiple times to multiply the contribution of surface states in terms of output intensity. The results are shown in Fig.3.23. In the vertical vertical axis, the quantity that is plotted is the difference in the reflectivity of the two surfaces. Being $\mathcal{R}_{clean}$ the reflectivity of the real sample and $\mathcal{R}_{ox}$ the one of the reference oxide surface, they plotted:

$$\frac{\Delta\mathcal{R}}{\mathcal{R}} = \frac{\mathcal{R}_{clean} - \mathcal{R}_{ox}}{\mathcal{R}_{ox}} \tag{3.52}$$

as a function of the photon energy $h\nu$. The peak, also known as **Chiarotti's peak**, clearly demonstrates the presence of surface states since it is associated with an optical transition between states. The lineshape and the width of such a peak are not easy to explain and require further analysis (see Sec.4.6). The Chiarotti experiment, led by Giorgio Chiarotti, was a landmark study in surface physics, focusing on the electronic properties of clean silicon and germanium surfaces. Beacuse of technological issues related to the measure of vacuum and sample preparation, it took many years for the experiment to be performed. Further consideration can be done regarding the characteristics of such transition among states, but first let's analyse the theoretical model behind such an experiment: the 3-layer model.

3.8.1 3-Layers Model

When a surface is created, as we have already discussed, the changes in the disposition of the atoms drastically change the electronic and mechanical properties of the surface. The change is so drastic that the surface layer can be considered as a new material. Therefore, when a surface is exposed to vacuum, the bulk material, the surface layer and the vacuum constitute an overall system that can be studied with the so-called **3-layers model** and it is schematically represented in Fig.3.24. The 3-layer model involving the bulk, surface, and vacuum layers is a powerful conceptual framework for understanding the electronic properties of semiconductor surfaces, especially in the context of techniques like photoemission

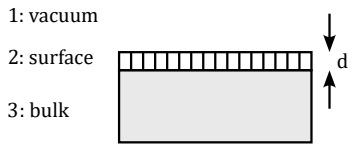


Figure 3.24: Schematic representation of the 3-layers model.

spectroscopy. In particular, this model can be used to express the physical contributions to the reflectivity $\mathcal{R}$ of the surface in terms of the dielectric tensor ε. In other words, we want to find a function $f(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ such that $\Delta\mathcal{R}/\mathcal{R} = f(\varepsilon_1, \varepsilon_2, \varepsilon_3)$, where $\varepsilon_1, \varepsilon_2$ and ε_3 are the dielectric tensors of the vacuum, the surface and the bulk. In particular, we will consider $\varepsilon_1 \in \mathbb{R}$ and constant beacuse it refers to the vacuum, $\varepsilon_2, \varepsilon_3 \in \mathbb{C}$ beacuse of absorption phenomena which may occur in the semiconductor. We will denote with $\hat{\varepsilon}$ the complex dielectric constants and any other complex quantity. Apart from the dielectric tensors, the other ingredients which compose our model are the magnetic permeability μ, the refractive index $\hat{n}$, the absorption coefficient α and the conductivity σ. By labelling with an index $j = 1, 2, 3$ the three layers the following relations hold:

$$\begin{aligned} \hat{n}_j &= n_j - ik_j = \sqrt{\mu_j \hat{\varepsilon}_j} \\ \hat{\varepsilon}_j &= \varepsilon'_j - i\varepsilon''_j \end{aligned} \tag{3.53}$$

So that:

$$\begin{aligned} (n_j - ik_j)^2 &= \mu_j \hat{\varepsilon}_j \\ \Rightarrow n_j^2 - k_j^2 - 2in_j k_j &= \mu_j(\varepsilon'_j - i\varepsilon''_j) \end{aligned} \tag{3.54}$$

From which we derive:

$$\varepsilon'_j = \frac{n_j^2 - k_j^2}{\mu_j}, \qquad \varepsilon''_j = 2\frac{n_j k_j}{\mu_j} \tag{3.55}$$

The analysis can be further developed by using Fresnel coefficients for reflection $r_\perp$ and $r_\parallel$ where "$\perp$" and "$\parallel$" refers to the direction of the electric field polarization with respect to the surface. Let's consider the separation interface between two media a and b, and let's define the quantity:

$$\xi_a = n_a \cos\phi_a \tag{3.56}$$

where ϕ_a is the incidence angle of the light on the surface. The Fresnel reflection coefficients are:

$$\begin{aligned} r_{\perp,ab} &= \frac{\mu_b \xi_a - \mu_a \xi_b}{\mu_b \xi_a + \mu_a \xi_b} \\ r_{\parallel,ab} &= \frac{\hat{\varepsilon}_b \xi_a - \hat{\varepsilon}_a \xi_b}{\mu_b \xi_a + \mu_a \xi_b} \end{aligned} \tag{3.57}$$

This result can be extended to our 3-layer model in which we have two different interfaces: the vacuum-surface and the surface-bulk ones. By denoting with d the thickness of the surface layer, i.e. the layer labelled as "2" in Fig.3.24, and with λ the wavelength of the incident radiation, the phase shift β occurring in the surface layer is:

$$\beta = (2\pi\hat{n}_2 \cos\phi_2)\frac{d}{\lambda} \tag{3.58}$$

The Fresnel coefficients for the reflection across the 3-layer structure become:

$$\begin{aligned} r_{\perp,123} &= \frac{r_{\perp,12} + r_{\perp,23}e^{-2i\beta}}{1 + r_{\perp,12}r_{\perp,23}e^{-2i\beta}} \\ r_{\parallel,123} &= \frac{r_{\parallel,12} + r_{\parallel,23}e^{-2i\beta}}{1 + r_{\parallel,12}r_{\parallel,23}e^{-2i\beta}} \end{aligned} \tag{3.59}$$

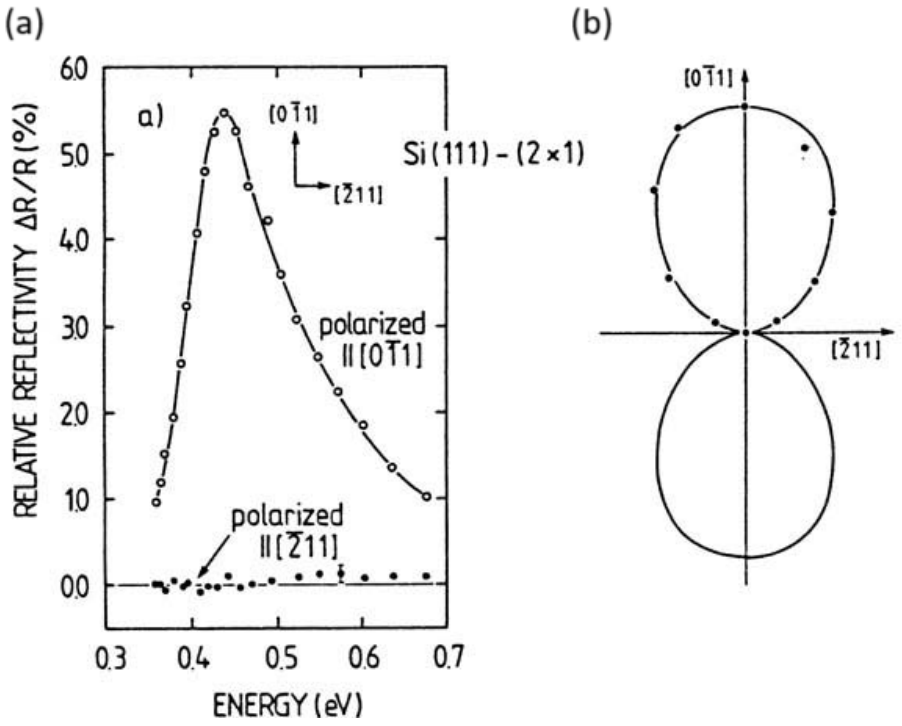


Figure 3.25: IR multiple internal reflection spectroscopy of the cleaved Si(111)-(2×1) surface. (a) Relative reflectivity change between a clean and oxygen-covered surface for light polarizations parallel to [0$\bar{1}$1] (π-chain direction) and parallel to [$\bar{2}$11] (normal to π-chain direction). (b) Polar diagram of the relative reflectivity changes as a function of light polarization direction.

where the presence of the 2 factor at the exponent is due to the double passage of the wave in the surface layer. We will consider the following approximations:

- Perpendicular incidence with respect to the sample: $\phi = 0 \Rightarrow r_\perp = r_\parallel$
- $\lambda \gg d \Rightarrow \frac{d}{\lambda} \ll 1 \Rightarrow \beta \ll 1 \Rightarrow e^{-2i\beta} \simeq (1 - 2i\beta)$

With these approximations, the expressions summarise in the following relation:

$$r_{123} = \frac{r_{12} + r_{23}(1 - 2i\beta)}{1 + r_{12}r_{23}(1 - 2i\beta)} \tag{3.60}$$

For our SDR we are interested in expressing the ratio $\Delta\mathcal{R}/\mathcal{R}$ with $\mathcal{R} = |r|^2$. In particular, by referring to Eq.(3.52), we have:

$$R_{clean} = |r_{123}|^2 \tag{3.61}$$

$$R_{ox} = |r_{13}|^2 \tag{3.62}$$

By using all the previous results one can find the relation we were looking for:

$$\frac{\Delta\mathcal{R}}{\mathcal{R}}\Big|_{3layers} = 8\pi n_1 \frac{d}{\lambda} \Im\mathfrak{m}\left\{\frac{\hat{\varepsilon}_2 - \hat{\varepsilon}_3}{\varepsilon_1 - \hat{\varepsilon}_3}\right\} \tag{3.63}$$

This result can be used, in the framework of the 3-layers model, to compare the experimental results obtained through RDS. Nevertheless, the term d/λ represents a critical parameter since the thickness d is not known a priori and depends on the reconstruction mechanism. For surfaces with a known reconstruction, the 3-layer model can be improved and can be used to study different polarization configurations. This is the case of Si(111)(2x1) for which the reconstruction configuration is Pandey's π-chain. In particular, we would like to understand where the surface states peak, i.e. the Chiarotti's

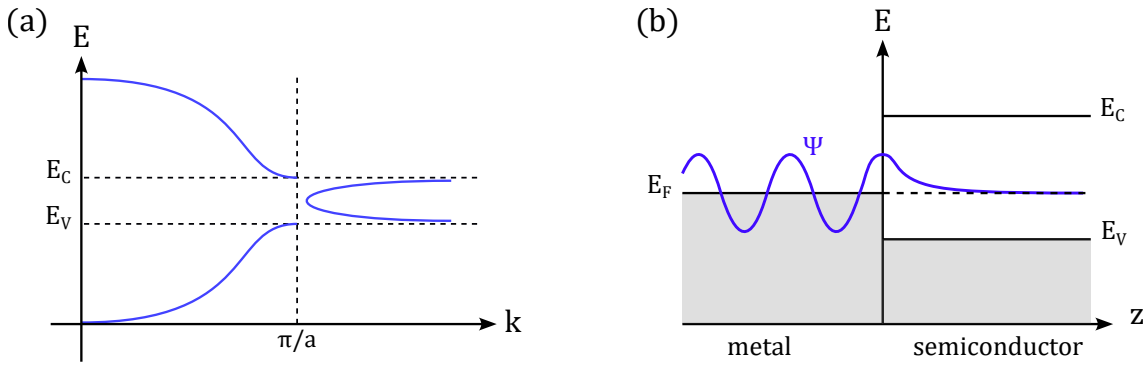


Figure 3.26: (a) Schematic representation of the MIGS in the band gap of the semiconductor. (b) Qualitative representation of a metal Bloch state (near E_F), decaying into the semiconductor.

peak, comes from, whether from the π-chain or the direction perpendicular to it. This can be done by inserting a polarizer in the SDR experimental apparatus. We recall that the π-chain in silicon is along the $[0\bar{1}1]$ direction, therefore we can choose as light polarization direction first $[0\bar{1}1]$ and then $[\bar{2}11]$. The results are shown in Fig.3.25(a). We notice that the surface states' peak is completely associated with the π-chain direction and therefore the optical transition must occur along the chain. The strong polarization dependence can be further understood by looking at the polar diagram of the relative reflectivity change as a function of light polarization direction reported in Fig.3.25(b). The results of this reflection absorption experiment are crucial for the π-chain model, since they cannot be explained by the buckling model.

3.9 Metal-Induced Gap States

As we have seen so far, numerous experiments on a wide variety of metal-semiconductor systems prepared under UHV conditions indicate that the deposition of metal films produces interface states which determine the position of the Fermi level at the interface. These interface states are called **Metal-Induced Gap States (MIGS)**. Based on Volker Heine's interpretation, MIGS arise because the metal's electronic wavefunction, which is a solution of the Schrödinger equation in the metal, cannot change abruptly to zero in the semiconductor, where no electronic states exist in the forbidden band, and therefore a tail of the wavefunction penetrate the semiconductor. As this wavefunction enters the semiconductor, it decays exponentially but still influences the electronic structure of the semiconductor near the interface. The energy levels of MIGS span the semiconductor's band gap and their density is highest near the band edges and decreases towards the middle of the gap. The situation is schematically represented in Fig.3.26. The theoretical understanding of MIGS is a powerful instrument for the characterization of surface states, but their complete analysis goes well beyond the scope of this course.

Chapter 4

Low Dimensional Systems and Collective Phenomena

Zero-dimensional (0D) and one-dimensional (1D) systems play crucial roles in the study of surfaces and surface phenomena. In 0D systems, such as adatoms or vacancies on a surface, localized electronic states and discrete energy levels can significantly influence surface chemistry and catalysis. One-dimensional systems, like linear atomic chains or step edges on surfaces, exhibit unique electronic properties due to quantum confinement along the chain. These systems can host collective phenomena such as charge density waves and spin chains, impacting the surface conductivity and magnetism. Understanding these low-dimensional systems is essential for advancing surface science and nanotechnology applications.

4.1 Excitons

The detailed study of surface states in silicon makes it possible to find the energy gap between surface states in Si(111)(2x1) which results to be $E_g^{ss} \simeq 0.7$ eV. This value is higher than the one observed in optical transitions, i.e. ~ 0.45 eV, suggesting the presence of some phenomenon which decreases the energy of the transition. The phenomenon we are looking for is the formation of a quasi-particle called **exciton**. An exciton is a bound state of an electron and a hole, which are attracted to each other by the Coulomb force, within a semiconductor or insulator. An exciton is a hydrogen-like system which forms when a photon is absorbed by a material, creating an electron-hole pair that remains bound together, allowing for energy transfer without net charge movement. The binding energy of an exciton is simply given by:

$$E_b = -\frac{e^2}{4\pi\varepsilon|\boldsymbol{r_e} - \boldsymbol{r_h}|} \tag{4.1}$$

where $\boldsymbol{r_e}$ and $\boldsymbol{r_h}$ are the coordinates of the electron and the hole, respectively. There exist two types of excitons:

- **Frenkel Excitons**: They have high binding energy and a small radius, typically comparable with the lattice parameter a.
- **Wannier-Mott Excitons**: They have a large radius, extending over several lattice constants a.

By considering the presence of excitons we understand the decrease in energy of the optical transitions between surface states which is influenced by the energy gain due to the formation of excitons. The properties of **surface excitons** are significantly influenced by the

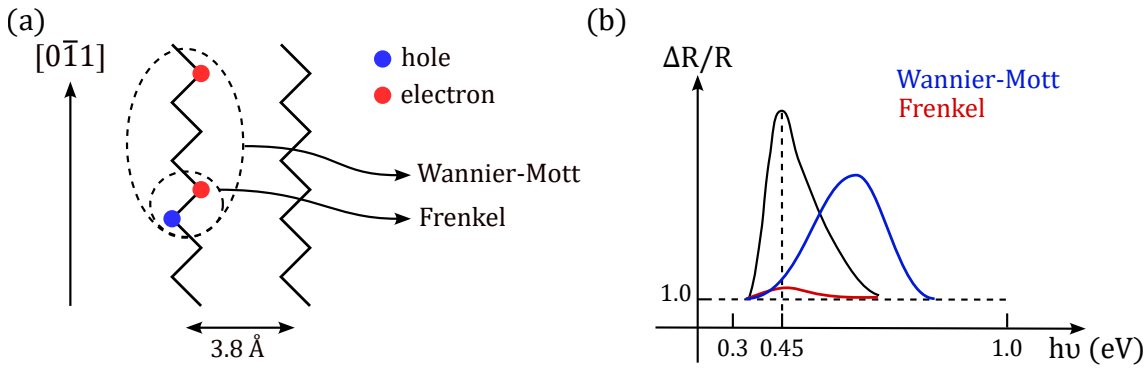


Figure 4.1: (a) Schematic representation of excitons along the π-chain in Si(111)(2x1). (b) Expected behaviour for optical transitions between surface states when decreasing the length of the π-chain for different types of excitons.

surface electronic structure and environment because these factors alter the local electronic potential, dielectric constant, and symmetry at the surface. Surface states, adsorbates, defects, and the dielectric environment all play crucial roles in modifying the binding energy, recombination dynamics, and optical properties of surface excitons, making them distinct from bulk excitons. We are now interested in studying the nature of excitons along the π-chain of silicon. In Fig.4.1(a) a schematic representation of the the chain is shown in which, depending on the localization of the electron-hole pair we distinguish the two excitons' types. Simulations show that the radius of the exciton along a π-chain is about ~ 120 Å, labelling it as a Weinner-Mott exciton, while in the perpendicular direction, it is of a Freckel type beacuse of the inter-chains distance. If we now imagine to oxidize the chain and therefore decrease its length, this corresponds to a different confinement of the exciton and we expect the binding energy of the quasi-particle to change. The confinement of an exciton along a *pi*-chain can be treated as the confinement of a particle in a box of side L. We know that the presence of confinement changes the energy levels of a particle making them quantized like:

$$E_n = \frac{\pi^2 \hbar^2}{2mL^2} n^2 \tag{4.2}$$

If we cut the chain, i.e. we decrease the box's side L, we expect the energy to increase. This implies that in an optical transition, for a Wannier-Mott exciton with a strongly delocalized behaviour, we expect the energy of the transition to increase and therefore the peak associated with the transition itself will move right. Such a shift in the peak is called **blue shift**. Instead, if the exciton is of a Frenkel type we do not expect such great variations when changing the length of the π-chain, but only a fast and drastic decrease in the intensity, i.e. the height, of the peak. The situation is summarized in Fig.4.1(b). As told, the physical mechanism through which the π-chain length undergoes a reduction is oxidation. Therefore, the changes in the optical transition peak also depend on the oxidation mechanism occurring at the surface. In the case of **random oxidation**, i.e. the oxygen atoms randomly and isotropically bond to the silicon atoms on the surface, is the case we considered so far. In this case, in fact, the isotropic oxidation of the surface implies an almost complete disappearance of the surface states' peak for Frenkel excitons. On the other end, we expect a blue shift for Wannier-Mott excitons and a small reduction of the intensity. The other oxidation mechanism that could occur at the surface is **cluster oxidation**. In this case, the oxygen atoms bond to the surface atoms in very specific locations, creating oxidation clusters, i.e. portions of oxidized surface. These clusters do not drastically affect excitons behaviour resulting in a small reduction of the peak's intensity, both for Frenkel and Wennier-Mott excitons.

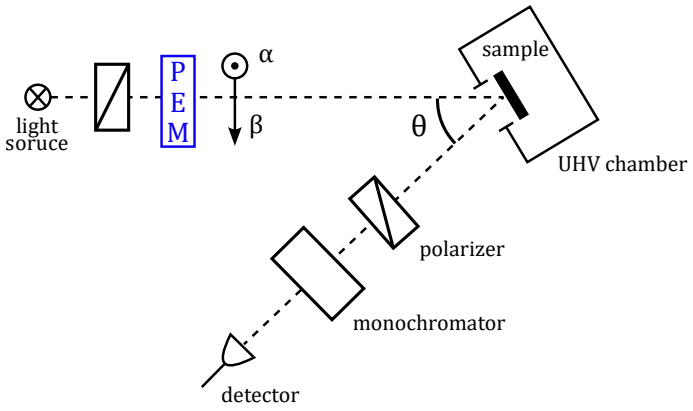


Figure 4.2: Schematic of the Reflectance Anisotropy Spectroscopy experimental apparatus.

4.1.1 Reflectance Anisotropy Spectroscopy

The oxidation-dependent exciton behaviour shows the importance of controlling the oxidation state of a surface. With SDR the partial oxidation state cannot be controlled since this technique works properly for fully oxidated samples. The oxidation-state control can be performed by using the Reflectance Anisotropy Spectroscopy (RAS) technique, also known as Reflectance Difference Spectroscopy (RDS). It is an optical technique used to investigate the surface properties of materials, particularly semiconductors. It is sensitive to surface anisotropies and can provide real-time, in-situ information about surface reconstructions, adsorbate interactions, and thin film growth. The experimental apparatus is shown in Fig.4.2(a). The key element of this apparatus is the Photoelastic Modulator (PEM) which is used to change the polarization of the incident light from the α to the β direction[1]. The quantity measured by this technique is the relative difference in reflectivity along the two polarization directions:

$$\frac{\Delta\mathcal{R}}{\mathcal{R}} = \frac{\mathcal{R}_\alpha - \mathcal{R}_\beta}{\mathcal{R}_\alpha + \mathcal{R}_\beta} \tag{4.3}$$

Quite often, a configuration with $\theta \simeq 0$ is used by operating with quite long optical paths. To mathematically describe this experimental technique we can use the so-called Jones matrix formalism. In optics, polarized light can be described using the Jones calculus, invented by R. C. Jones in 1941. Polarized light is represented by a Jones vector, and linear optical elements are represented by Jones matrices. When light crosses an optical element the resulting polarization of the emerging light is found by taking the product of the Jones matrix of the optical element and the Jones vector of the incident light. The vector we use to describe the incoming light is the input electric field $\boldsymbol{E}_{in}$ defined with its two polarization components:

$$\boldsymbol{E}_{in} = \begin{bmatrix} E_x \\ E_y \end{bmatrix} \tag{4.4}$$

The output electric field will be $\boldsymbol{E}_{out} = \boldsymbol{M}\boldsymbol{E}_{in}$ where $\boldsymbol{M}$ is a matrix describing the different optical elements in the apparatus. The matrix representation of these elements in order are:

- **Polarizer**:

$$P = \begin{bmatrix} 1 & 0 \\ 0 & \alpha_p \end{bmatrix} \tag{4.5}$$

[1]The PEM operates based on the photoelastic effect, which is the change in the refractive index of a material in response to mechanical stress. When a birefringent crystal, e.g. fused silica, is subjected to an oscillating mechanical stress, it induces periodic birefringence in the material which can be used to modulate the light's polarization.

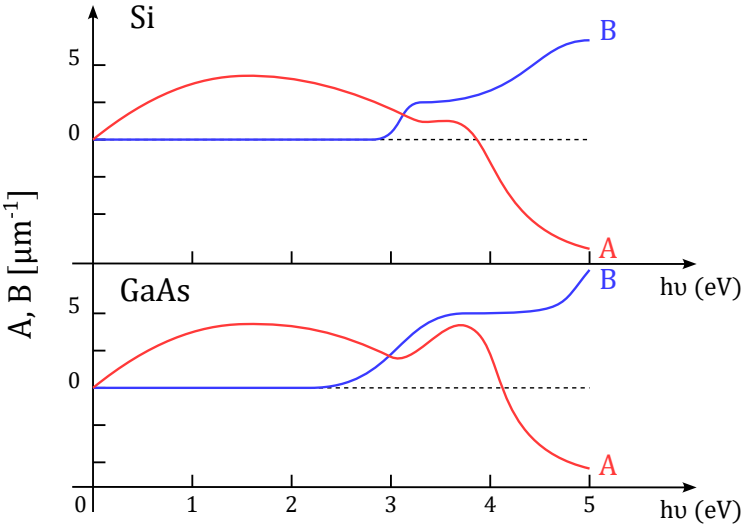


Figure 4.3: RAS experimental results in terms of the A and B coefficients as functions of the photon energy $h\nu$ for Si and GaAs.

where α is a parameter that indicates the attenuation (or amplification) of the E_y component of the electric field relative to the E_x component.

- **Photoelastic Modulator**:

$$PEM = \begin{bmatrix} e^{-i2\pi\frac{d_b}{\lambda}n_s} & 0 \\ 0 & e^{-i2\pi\frac{d_b}{\lambda}n_o} \end{bmatrix} \tag{4.6}$$

where d_b, n_s and n_o are the thickness, the extraordinary refractive index and the ordinary refractive index of the birefringent crystal in the PEM, respectively.

- **Sample**:

$$S = \begin{bmatrix} r_\alpha & 0 \\ 0 & r_\beta \end{bmatrix} \tag{4.7}$$

where r_α and r_β are the reflection coefficients along the two polarization directions.

Therefore, for the optical path from the source to the sample the overall M matrix is:

$$M = S \cdot PEM \cdot P \tag{4.8}$$

This allows for the calculation of the output intensity $\mathcal{I}$ which results to be:

$$\mathcal{I} = |E_{out}|^2 = \mathfrak{Re}\left\{\frac{\Delta r}{r}\right\} \tag{4.9}$$

where, from the 3-layers model:

$$\frac{\Delta r}{r} = -i4\pi n_1 \frac{d}{\lambda}\frac{\Delta\varepsilon_s}{1-\varepsilon_b} \tag{4.10}$$

with d thickness of the surface and $\Delta\varepsilon_s$ is the difference in dielectric constants of the surface along the two polarization directions:

$$\Delta\varepsilon_s = \varepsilon_s^\alpha - \varepsilon_s^\beta \tag{4.11}$$

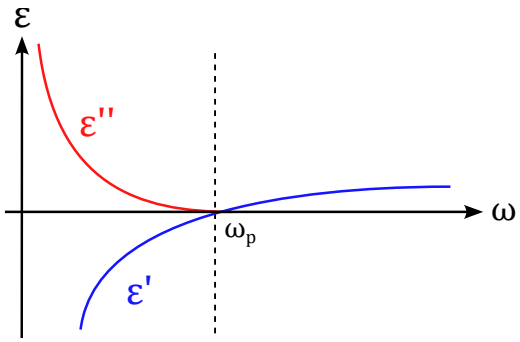


Figure 4.4: Scheamtic of the behaviour of the real and imaginary part of the dielectric constant of a metal as a function of the frequency of an electromagnetic wave impinging on the metal based on Drude model.

and $\varepsilon_s = \varepsilon_s' - i\varepsilon_s''$. Lastly, $\varepsilon_b = \varepsilon_b' - i\varepsilon_b''$ is the complex dielectric constant of the bulk material. It is possible to demonstrate:

$$\frac{\Delta r}{r} = 4\pi n_1 \frac{d}{\lambda} \underbrace{\frac{1-\varepsilon_b'}{(1-\varepsilon_b')^2 + (\varepsilon_b'')^2}\Delta\varepsilon_s''}_{A} + 4\pi n_1 \frac{d}{\lambda} \underbrace{\frac{\varepsilon_b''}{(1-\varepsilon_b')^2 + (\varepsilon_b'')^2}\Delta\varepsilon_s'}_{B} \tag{4.12}$$

The two terms denoted with A and B refer to the reflective and absorptive phenomena occurring in the sample and their experimental measurements for Si and GaAs are shown in Fig.4.3. Here, we limit ourselves to noticing that the B coefficient, related to the absorptive part is zero below a certain threshold energy, both for Si and GaAs, as expected. In this section, we introduced the RAS to understand the nature of the excitons in the one-dimensional π-chain. The results of a RAS experiment on a Si(111)(2x1) sample show a blue shift for the surface states' peak denoting the presence of Weinner-Mott exciton in the chain. This system behaves as a 1D system not only from a geometrical point of view but also from an electronic point of view beacuse of the confinement of the exciton along the whole chain.

4.2 Mott–Hubbard Transistion

The previous example showed a system in which the geometrical and electronic dimensionality coincide with each other, i.e. 1D geometrical and electronic character. Let's consider a substrate of GaAs above which a layer of caesium is deposited. The resulting crystalline structure observed through STM with a coverage $\Theta = 0.03$ is shown in Fig.4.5(a). The peculiar arrangement of the Cs atoms resembles the structure of the zig-zag π-chain in silicon and it can get quite extended, reaching a length of about 100 nm. The distance between two Cs atoms in the chain is about 6.9 Åwhile in a normal caesium crystal the lattice parameter is about 6.1 Å. When performing optical measurements on such a structure, one has to consider the optical properties of metals. To describe these properties we can simply refer to the Drude model which predicts the existence of a plasma frequency ω_p and relates the dielectric constant of a metal ε to the frequency of an incident radiation. The trend of $\varepsilon = \varepsilon(\omega)$ is shown in Fig.4.4. This has to be taken into consideration because, in a RAS measurement, we measure:

$$\frac{\Delta\mathcal{R}}{\mathcal{R}} = \mathfrak{Re}\left\{\frac{\Delta r}{r}\right\} \propto \Delta\varepsilon_s'' \tag{4.13}$$

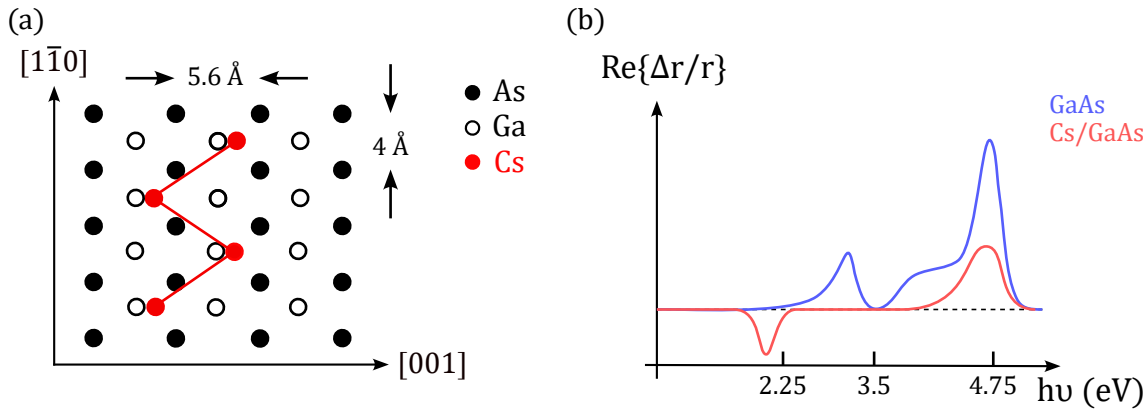


Figure 4.5: Deposition of Cs on a GaAs substrate and resulting crystalline structure. (b) RAS measurements both for GaAs and for the Cs/GaAs compound.

and for $\omega < \omega_p$ the contribution associated to the immaginary part of ε is signficantly different from zero. Therefore, we expect that for frequencies lower than the plasma frequency, the RAS signal is enhanced by the presence of the metal. For the RAS measurement, we choose as the α direction for the PEM the direction of the Cs atoms chain, i.e. the [1$\bar{1}$0] direction, and as β direction the one perpendicular to the chain, i.e. the [001] direction. Since the plasma frequency of caesium is $\hbar\omega_p = 2.8$ eV, we expect an enchantment of the RAS signal while going towards the IR frequencies along the α direction. The results, shown in Fig.4.5(b), do not exhibit any enchantment of the signal, but a reduction. This implies that the 1D chain of Cs atoms does not have a metallic behaviour, but a semiconducting one. Such a loss of the metallic behaviour of these chains of metal atoms is called a **Mott-Hubbard metal-insulator transition**. The explanation of such a phenomenon lies in the wave functions overlap of different Cs atoms. The increased distance between caesium atoms in the chain which reaches 6.9 Åprevents the orbital wave functions from overlapping one another, inhibiting the electron transfer.

4.3 Steps and Schwoebel-Ehrlich Barrier

Another example of 1D structures on surfaces is given by steps. Surface steps on a crystal can resemble 1D systems because they confine atomic or electronic movement along the step edge, leading to the formation of linear atomic chains and 1D electronic states. This confinement results in unique physical and chemical properties that are characteristic of one-dimensional systems. We have already discussed in Sec. 2.3 some characteristics of surface steps and, in particular, the fact that when cutting a surface along a certain plane inclined by θ with respect to the horizontal direction (see Fig.2.6), the formation of steps is energetically favoured. In the following, we will discuss the use of steps as deposition sites for adsorbate atoms or any generic structure that can be deposited on the steps' ledges. Before doing that, it is worth noting that different steps on the same surface can have different properties. Fig.4.6(a) represents a stepped surface for Si(001) in which different orientations of the silicon dimers are observed for different steps. If the number of steps is even, the total number of dimers with a certain orientation is the same in all directions and the overall character of the dimers is isotropic. Conversely, an odd number of steps causes an imbalance in the number of dimers in a certain direction, giving the surface some direction-dependent properties that can be used in applications. Moreover, different configurations can be obtained by changing the cutting angle of the surface θ. First, by increasing α, we can have an increased probability of obtaining higher steps, i.e. diatomic

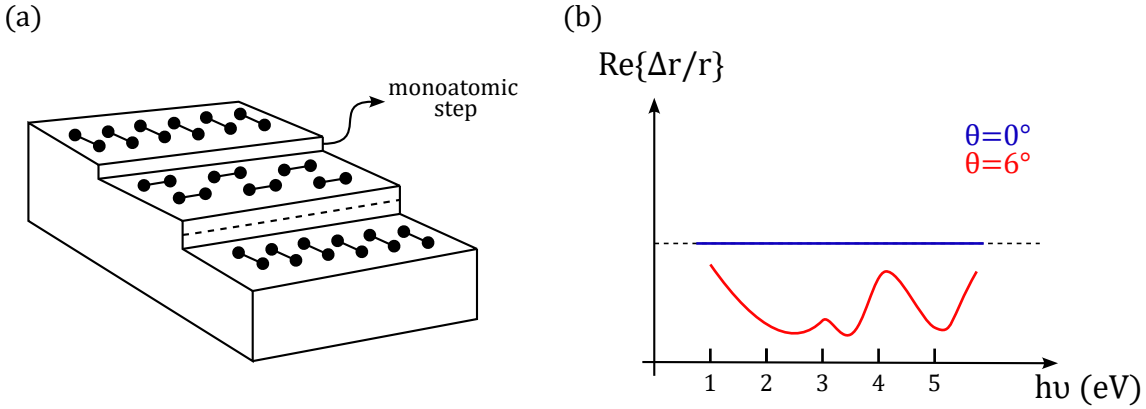


Figure 4.6: (a) Scheamtic representation of steps on Si(001) surface. (b) RAS experiment on Si(001) at different cutting angles for the surface.

or triatomic steps. Also, the electrical properties of the surface depend on the cutting angle as shown by the RAS measurement reported in Fig.4.6(b) for different cutting angles. In general, the explanation of such a line shape is complicated and goes well beyond the scope of the course; we limit ourselves to observing that the higher the angle the stronger the obtained RAS signal.

4.3.1 Schwoebel-Ehrlich Barrier

In general, the formation of a step on a surface is associated with an increase in the overall surface energy. Nevertheless, steps are stable structures and they spontaneously form on surfaces denoting that there must exist a stabilization process leading to the actual presence of steps. The stabilization process which leads to the formation of the steps is associated with the creation of a potential barrier called **Schwoebel-Ehrlich barrier**. Thanks to the presence of this barrier steps can be used to stabilize and create one-dimensional structures, depositing absorbate atoms along the edge of the step. As depicted in Fig.4.7 there exist two possible types of Schwoebel-Ehrlich barriers, attractive and repulsive ones. In the repulsive case, the step generates first a decrease in the potential energy followed by the drastic increase which generates the barrier itself. If we deposit an absorbate atom on a step characterized by such a barrier we obtain the configuration shown at the right of Fig.4.7. On the other hand, the attractive barrier generates an opposite deposition configuration in which the absorbate atom is stabilised at the bottom of the ledge. It is possible to calculate the rate of probability for an atom to jump the barrier j_b. This quantity is derived from a semi-empirical law and it is given by Arrenius's law multiplied for a correction term:

$$j_b \propto \left(\nu e^{-\frac{E_b}{k_B T}}\right)\left[1+\left(b-\frac{1}{2}\right)e^{-\frac{E_R-E-d}{K_B T}}\right] \tag{4.14}$$

where the energy terms E_b, E_R and E_d are shown in Fig.4.7, b is the width of the barrier and ν is a prefactor associated with the number of "attempts" of jumping the barrier. The presence of this barrier is observed experimentally, but it is not well understood yet. For example, Fig.4.8 shows two possible step configurations in the Si(111)(2x1) surface. Here we notice that in the (b) case the addition of the 6 atoms ring relaxes the structure and makes the step shift by one position. Both these two configurations can be present on the same silicon sample, but there is no a priori explanation for the formation of one over the other.

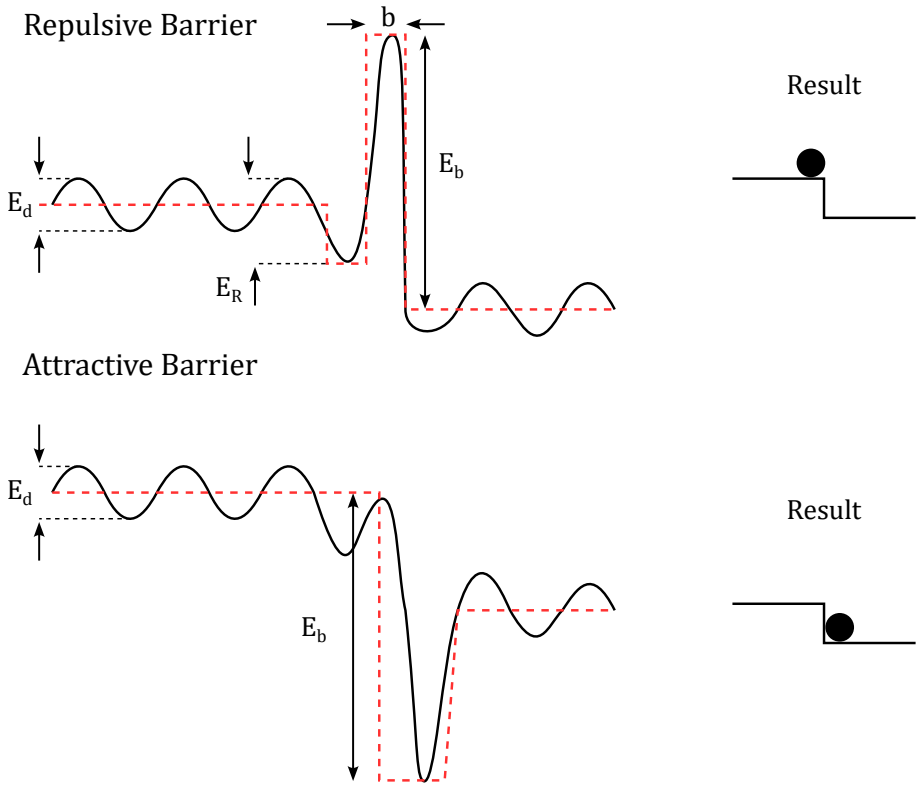


Figure 4.7: Repulsive and attractive Schwoebel-Ehrlich barriers and relative deposition configuration.

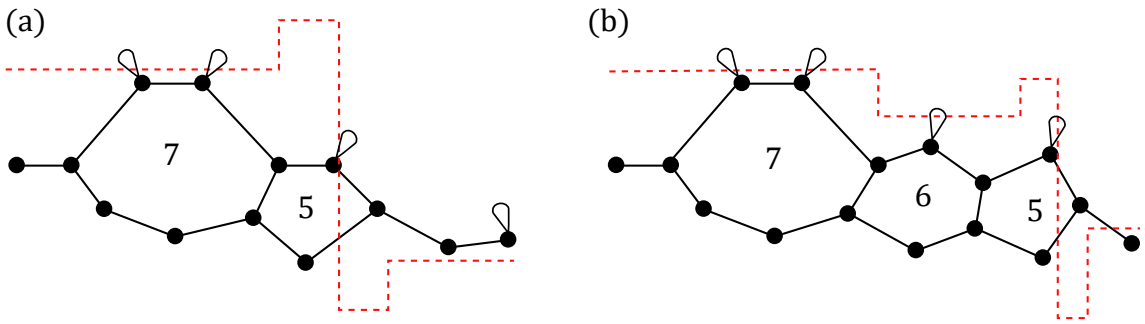


Figure 4.8: Possible atomic configurations for a step on Si(111)(2x1).

4.4 1D Structures: Some Examples

4.4.1 1D Metal-Semiconductor Junction

One-dimensional metal-semiconductor junctions represent interfaces between metallic and semiconductor materials in one-dimensional structures. They are typically nanowires, nanorods, nanotubes, or other nanoscale structures where one dimension (length) is much larger than the other two dimensions (diameter). Such a system can be studied with the model developed by Nozaki and Cuniberti (arXiv: 0907.0155) and it is possible to see that the rectifying properties seen for the bulk MS junction still hold in 1D despite the increased influence of any structural defects present at the interface.

4.4.2 Donor-Acceptor Molecular Transistor

Another way of building a 1D structure consists of using molecules as building blocks of electronic devices. Such devices are called **unimolecular rectifier** and consist of a single organic molecule which functions as a rectifier (one-way conductor) of electric current. The idea was first proposed in 1974 by Arieh Aviram and Mark Ratner. Their proposed rectifying molecule, shown in Fig.4.9, was designed so that electrical conduction within it would be favoured from the electron-rich subunit (electron donor) to an electron-poor subunit (electron acceptor), but disfavored in the reverse direction. In the benzene ring on the left, the electronic structure is perturbed by the addition of the oxygen atom on the two edges of the hexagon. In particular, the electronic cloud gets shifted toward the oxygen beacuse of its higher electronegativity causing a decrease in the overall negative charge in the benzene ring. For this reason, this side of the molecule is more inclined to accept electrons and it is said to be **acceptor-like**. Conversely, in the benzene ring on the right, the presence of the methoxy group Me causes an increase in the overall electronic charge making this side the **donor-like** part. If one sandwichs this molecule between two metallic electrodes gives rise to the so-called Donor-Acceptor Molecular Transistor whose energy levels are shown in Fig.4.9(b). A device built like this behaves as a rectifier device for which by setting the voltage drop across the electrodes the current flow can be controlled. The average current density reached by such devices is $\sim 0.4 \cdot 10^{-2}$ A/cm^2 which is quite a small value if compared with the current density in a standard metal, e.g. current density in copper $\sim 3 \cdot 10^4$ A/cm^2.

4.5 0D Systems

A zero-dimensional system is a type of nanostructure in which charge carriers (such as electrons and holes) are confined in all three spatial dimensions. This confinement leads to discrete energy levels and unique physical properties that differ significantly from those

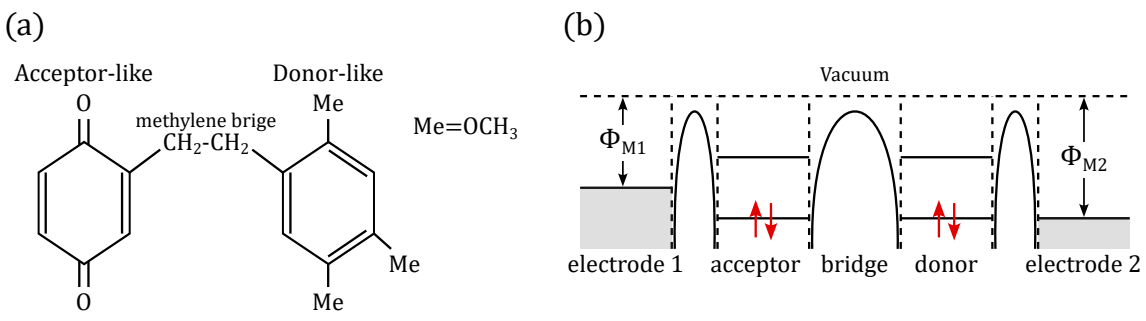


Figure 4.9: (a) Molecule and (b) energy structure for the Donor-Acceptor Molecular Transistor.

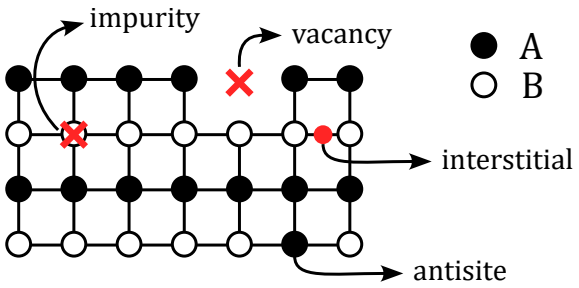


Figure 4.10: Schematic representation of the defects' types occurring on a zinc-blende structure.

of bulk materials. 0D systems and surface defects are closely related in surface science because certain types of surface defects can exhibit 0D characteristics. Let's consider a zinc-blende structure as the one of Fig.4.10. The possible 0D surface defects in such a structure which can occur both in the A and B sites are:

- **Impurities** S_A, S_B: They are foreign atoms present on the surface of a material, differing from the host surface atoms.
- **Vacancies** V_A, V_B: They occur when surface atoms are missing from their regular positions in the surface lattice.
- **Interstitials** I_A, I_B: They are atoms that occupy positions on the surface between the regular lattice sites.
- **Antisites** A_A, A_B: They occur when atoms of one type occupy surface lattice sites typically occupied by another type of atom within the surface layer.

The presence of these defects can be studied using thermodynamics arguments. In particular, we express the Landau potential variation due to the presence of a defect. For example, if we considered a zinc-blende structure with N_A A atoms and N_B B atoms, for a vacancy defect in the B site, we have:

$$\Omega_V = E(N_A, N_{B-1}, q) - E(N_A, N_B) + \mu_B + E_{corr} \tag{4.15}$$

where $E = E(N_A, N_B)$ is the energy without the defect and $E(N_A, N_{B-1}, q)$ is the energy with the defect, i.e. with one B atom less, in which we considered also the possibility of not-having charge neutrality which is not granted in the presence of a defect. Moreover, since the defect modifies the chemical environment seen by the atoms of the surface we need to keep into account also possible variations of the chemical potential of the B species μ_B. Lastly, a higher order correction term E_{corr} can be considered a posteriori. Instead, for example, if we consider an impurity X on a B site, we have:

$$\Omega_S = E(N_A, N_{B-1}, q, X) - E(N_A, N_B) + \mu_B - \mu_X + E_{corr} \tag{4.16}$$

where in this case also the variation associated with the adding of the X species has been kept into account.

4.5.1 Defects Localization and Tomlinson's Model

The first question one could wonder is whether it is possible or not to localize these defects experimentally. The answer is yes and to do that Atomic Force Microscopy can be used as schematically shown in Fig.4.11(a). In fact, the defects modify the local property of the surface and therefore we expect the interaction between the AFM tip and the surface to change when a defect is met by the tip. The first to perform such an experiment was

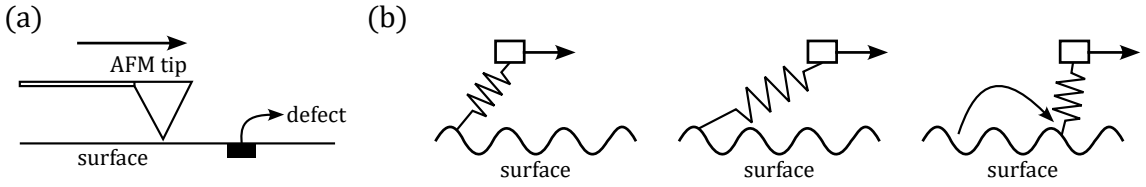


Figure 4.11: (a) Schematic representation of the use of an AFM tip for defects' localization. (b) Steps for the interpretation of Tomlinson's Model.

E. Mate and colleagues who understood that AFM can also be used to measure frictional forces at the nanoscale if used in contact mode. By scanning the sample surface while applying a lateral force with the AFM probe, friction maps can be generated, revealing variations in frictional properties associated with surface defects. The theoretical model to interpret such results was developed by G.A. Tomlinson in 1929. **Tomlinson's Model** was introduced to explain the origin of friction at the microscopic level, and it has been widely used to understand and predict frictional forces in various systems, including those involving atomic force microscopy. The basic idea is that friction forces are modelled as elastic forces and two surfaces in contact one with another can be associated with a spring connected between them as in Fig.4.11(b). While the two surfaces (or in our case the tip and the surface) are sliding one onto the other, the spring undergoes elongation up to a point in which a sharp transition occurs and the position of the spring on one of the two surfaces, suddenly change from one atomic site to a neighbouring one (see steps in Fig.4.11(b)). The sharp transition of the spring from one site to another is considered to be instantaneous. This hypothesis is justified by the fact that, despite the sharp transition being associated with a certain heat dissipation due to vibrational phenomena, e.g. phonons, this dissipation is much faster than the movement of the tip on the surface and therefore we can consider the transition as instantaneous. Within Tomlinson's model, it is possible to write the motion equation for the sliding of an AFM tip on a surface. Supposing a 2D motion along the x and the y directions, we have:

$$\begin{aligned} &\text{Along x: } m_x\ddot{x} = k_x(x_0 - x) - \frac{\partial V}{\partial x} - \gamma_x\dot{x} \\ &\text{Along y: } m_y\ddot{y} = k_y(y_0 - y) - \frac{\partial V}{\partial y} - \gamma_y\dot{y} \end{aligned} \tag{4.17}$$

where the terms on the right-hand side of the equation respectively are: Hook's law, the contributions associated with the local forces due to surface roughness and some possible damping effects. As told, this model was used for the theoretical interpretation of the experimental data collected by Mate. His first experiment was conducted on a HOPG sample for which the model potential V to use is:

$$V(x,y) = -V_0\left[2\cos\left(\frac{2\pi}{a}x\right)\cos\left(\frac{2\pi}{a\sqrt{3}}y\right) + \cos\left(\frac{2\pi}{a\sqrt{3}}y\right)\right] \tag{4.18}$$

where a is the lattice parameter of graphene and V_0 is a constant. Apart from this method based on AFM, another possibility exists for spotting defects on surfaces based on STM. Such a method has been proven theoretically effective but hasn't been used experimentally yet. For example, on Si(111)(2x1) the key idea is that a defect completely modifies the reconstruction abruptly interrupting the π-chain. Such a difference should be in principle visible through an STM analysis.

4.6 Optical Transistions

As we have already discussed in Sec.3.8, the existence of surface states has been proven by Chiarotti by using optically induced transitions among these states. Nevertheless, a more detailed explanation of the asymmetric lineshape or the width of Chiarotti's peak requires considering optical transitions in a more detailed way. Let's consider a semiconductor whose conduction and valence band are described by the dispersion relations:

$$E_c(\boldsymbol{k}) = E_c + \frac{\hbar^2 k^2}{2m_e}, \qquad E_v(\boldsymbol{k}) = E_v - \frac{\hbar^2 k^2}{2m_h} \tag{4.19}$$

where E_c [E_v] is the minimum [maximum] of the conduction [valence] band and m_e [m_h] is the effective mass of the conduction [valence] band. The electronic wave function can be written in terms of Bloch wave functions:

$$\psi_{n,k}(\boldsymbol{r}) = u_{n,k}(\boldsymbol{r}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{4.20}$$

The **selection rules** we are going to consider based only on the conservation of the energy and of the angular momentum, thus we will neglect any spin-related selection rule. This implies that the transitions we will consider are **interband transistions** only, i.e. intraband transition will be neglected. Let E_n and E_m be the energies of the initial and final states, and $\boldsymbol{k}_i$, $\boldsymbol{k}_f$ the wave vectors of the electrons in these states. By considering photons with energy $\hbar\omega$ and momentum $\boldsymbol{q}$, we have:

$$\begin{aligned} E_m(\boldsymbol{k}_f) &= E_n(\boldsymbol{k}_i) + \hbar\omega \\ \boldsymbol{k}_f &= \boldsymbol{k}_i + \boldsymbol{q} \end{aligned} \tag{4.21}$$

We can now evaluate the transition rate by using Fermi's Golden Rule. We have that the probability of a transition per unit time from the valence to the conduction band is:

$$W_\uparrow(\boldsymbol{k}_i) = \frac{2\pi}{\hbar} \sum_{\boldsymbol{k}_f} |\left\langle \psi_{c,k_f} \middle| \hat{H} \middle| \psi_{v,k_i} \right\rangle|^2 \delta(E_c(\boldsymbol{k}_f) - E_v(\boldsymbol{k}_i) - \hbar\omega) \tag{4.22}$$

where with the subscript "$\uparrow$" we denote the interband transitions. It is possible to show that in the context of the electric dipole approximation, this expression becomes:

$$W_\uparrow(\boldsymbol{k}_i) = \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m}\right)^2 \sum_{\boldsymbol{k}_f} |\underbrace{\left\langle \psi_{c,k_f} \middle| e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \hat{p}\cdot\hat{n} \middle| \psi_{v,k_i} \right\rangle}_{(*)}|^2 \delta(E_c(\boldsymbol{k}_f) - E_v(\boldsymbol{k}_i) - \hbar\omega) \tag{4.23}$$

where $\hat{p}$ is the momentum operator and $\hat{n}$ is the versor describing the direction of the electromagnetic field. The term we denoted with $(*)$ can be expanded by writing the Bloch wave functions explicitly:

$$(*) = \int d^3r e^{-i\boldsymbol{k}_f\cdot\boldsymbol{r}} u^*_{c,k_f}(\boldsymbol{r}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}} e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}} \hat{p}\cdot\hat{n} u_{v,k_i}(\boldsymbol{r}) \tag{4.24}$$

Beacuse of the electric dipole approximation and also considering that the periodic part of the Bloch wave functions does not vary significantly within the same unit cell (u.c.) denoted by the vector $\boldsymbol{R}_j$, the following expressions hold:

$$\begin{aligned} (*) &= \sum_{\boldsymbol{R}_j} \int_{u.c.} d^3r e^{-i\boldsymbol{k}_f\cdot\boldsymbol{r}} u^*_{c,k_f}(\boldsymbol{r}) e^{i\boldsymbol{q}\cdot\boldsymbol{r}} e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}} \hat{p}\cdot\hat{n} u_{v,k_i}(\boldsymbol{r}) = \\ &= \sum_{\boldsymbol{R}_j} e^{(\boldsymbol{k}_i - \boldsymbol{k}_f + \boldsymbol{q})\cdot\boldsymbol{R}_j} \int_{u.c.} d^3 u^*_{c,k_f}(\boldsymbol{r}) \hat{p}\cdot\hat{n} u_{v,k_i}(\boldsymbol{r}) = \\ &= N\delta(\boldsymbol{k}_i - \boldsymbol{k}_f + \boldsymbol{q}) \left\langle u_{c,k_f} \middle| \hat{p}\cdot\hat{n} \middle| u_{v,k_i} \right\rangle \end{aligned} \tag{4.25}$$

where N is the total number of cells in the semiconductor crystal. For simplicity we define the matrix element as $\hat{p}_{cv} \cdot \hat{n} = \langle u_{c,k_f} | \hat{p} \cdot \hat{n} | u_{v,k_i} \rangle$. With these simplifications, the transition rate becomes:

$$W_{\uparrow}(\boldsymbol{k}_i) = \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m} \right)^2 \sum_{\boldsymbol{k}_f} N^2 |\hat{p}_{cv} \cdot \hat{n}|^2 \delta(\boldsymbol{k}_i - \boldsymbol{k}_f + \boldsymbol{q}) \delta(E_c(\boldsymbol{k}_f) - E_v(\boldsymbol{k}_i) - \hbar\omega) \tag{4.26}$$

The problem can be further simplified. In general, the photon's momentum is much smaller than the one of the electrons, i.e. $|\boldsymbol{q}| \ll |\boldsymbol{k}_i|, |\boldsymbol{k}_f|$. This implies that $\boldsymbol{k}_i \simeq |\boldsymbol{k}_f| = \boldsymbol{k}$ and that interband transitions, within these approximations, are **vertical transistions**. The final expression for the transition rate is:

$$W_{\uparrow}(\boldsymbol{k}) = \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m} N \right)^2 \sum_{\boldsymbol{k}} |\hat{p}_{cv} \cdot \hat{n}|^2 \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) \tag{4.27}$$

The overall rate of transitions rate should keep into account also the probabilities of having empty states in the conduction band and filled ones in the valence band. These probabilities are given by the Fermi distributions and the overall probability is $f_v(\boldsymbol{k}_i) \left[1 - f_c(\boldsymbol{k}_f) \right]$. For an intrinsic semiconductor as a first approximation, we can consider all the valence band states to be filled and all the conduction ones to be empty. Therefore, we get $f_v(\boldsymbol{k}_i) \left[1 - f_c(\boldsymbol{k}_f) \right] \simeq 1$. In the end, the overall transition rate per unit volume as a function of the field frequency $R_{\uparrow}(\omega)$ is:

$$\begin{aligned} R_{\uparrow}(\omega) &= \frac{2}{V} \sum_{\boldsymbol{k}} W_{\uparrow}(\boldsymbol{k}) f_v(\boldsymbol{k}_i) \left[1 - f_c(\boldsymbol{k}_f) \right] \simeq \\ &\simeq \frac{2}{V} \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m} N \right)^2 \sum_{\boldsymbol{k}} |\hat{p}_{cv} \cdot \hat{n}|^2 \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) \end{aligned} \tag{4.28}$$

The square modulus of the matrix element contains all the physical information about the interaction occurring between the field and the electrons in the crystal and it is therefore a fundamental element of this expression. Nevertheless, we can consider its mean value in the First Brillouin Zone (FBZ) by substituting the summation over $\boldsymbol{k}$ with an integral in the FBZ as:

$$\begin{aligned} R_{\uparrow}(\omega) &= \frac{2}{V} \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m} N \right)^2 \int_{FBZ} |\hat{p}_{cv} \cdot \hat{n}|^2 \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) \frac{d^3k}{(2\pi)^3} = \\ &= \frac{2}{V} \frac{2\pi}{\hbar} \left(\frac{eA_0}{2m} N \right)^2 \langle |\hat{p}_{cv} \cdot \hat{n}|^2 \rangle \int_{FBZ} \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega) \frac{d^3k}{(2\pi)^3} \end{aligned} \tag{4.29}$$

where $d^3k/(2\pi)^3$, which is the volume of a state in the k-space, has been included to compute this integral in the momentum space and $\langle |\hat{p}_{cv} \cdot \hat{n}|^2 \rangle$ is the mean value in the FBZ. To compute the remaining integral we can consider that the Density of States (DOS) in the conduction band $g_c(E)$ for a dispersive relation like the one of Eq.(4.19) is:

$$g_c(E) := 2 \int_{FBZ} \frac{d^3k}{(2\pi)^3} \delta(E_c(\boldsymbol{k}) - E) = \frac{1}{2\pi^2} \left(\frac{2m_e}{\hbar^2} \right)^{3/2} \sqrt{E - E_c} \tag{4.30}$$

The definition of the DOS resembles the last expression we got for the transition rate, except for the energy terms present in the Dirac delta. To make the two expressions comparable we define:

$$E_g(\boldsymbol{k}) := E_v(\boldsymbol{k}) - E_v(\boldsymbol{k}) = E_c - E_v + \frac{\hbar^2 k^2}{2} \left(\frac{1}{m_e} + \frac{1}{m_h} \right) = E_g + \frac{\hbar^2 k^2}{2m_r} \tag{4.31}$$

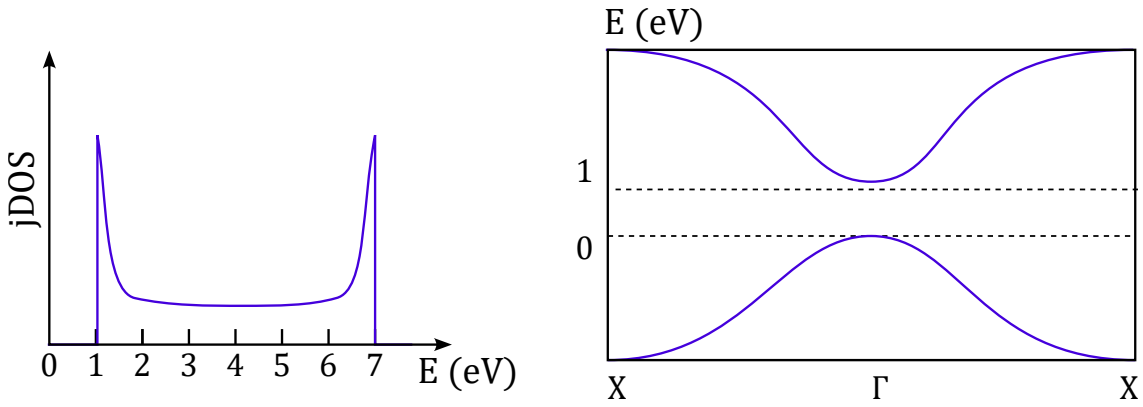


Figure 4.12: Cardona's simulations for Si(111)(2x1): band diagram and jDOS as a function of the energy.

where $E_g = E_c - E_v$ is the energy gap and $1/m_r = 1/m_e + 1/m_h$ is the inverse of the reduced mass. With this substitution the two expressions of Eq.(4.29) and Eq.(4.30) are formally equivalent and the integral is solved:

$$\begin{aligned} R_\uparrow(\omega) &= \frac{2}{V}\frac{2\pi}{\hbar}\left(\frac{eA_0}{2m}N\right)^2 \langle|\hat{p}_{cv}\cdot\hat{n}|^2\rangle \int_{FBZ} \delta(E_c(\boldsymbol{k}) - E_v(\boldsymbol{k}) - \hbar\omega)\frac{d^3k}{(2\pi)^3} = \\ &= \frac{1}{V}\frac{2\pi}{\hbar}\left(\frac{eA_0}{2m}N\right)^2 \langle|\hat{p}_{cv}\cdot\hat{n}|^2\rangle \underbrace{\frac{1}{2\pi^2}\left(\frac{2m_e}{\hbar^2}\right)^{3/2}\sqrt{\hbar\omega - E_g}}_{jDOS} \end{aligned} \tag{4.32}$$

where the last term defines the Joint Density of States (jDOS). Finally, this expression can be used to evaluate quantitatively different configurations. Manuel Cardona used such an expression to perform simulations regarding the value of the jDOS for a linear, i.e. 1D, semiconducting system made of silicon. The band diagram and the result of the simulation are shown in Fig.4.12. First, we recognize the band diagram typical of silicon which has around Γ a parabolic dispersion relation and that has an energy gap of about 1 eV. Then, we notice that the jDOS has an asymmetric peak at the energy gap and that for energy smaller than E_g no states are available for the transition as expected. This asymmetry in the jDOS is what partially justifies the asymmetry of Chiarotti's peak of Fig.3.23. Although being a fundamental factor to consider, it can be seen that the asymmetry of jDOS's peak is smaller than the one of Chiarotti's peak. Further explanations for the shape of the peak are therefore required to fully explain the experimental observations. A step in this direction has been made by Ciccacci *et al.*

4.6.1 Ciccacci's Experiment

The experiment performed by Ciccacci and his collaborators aimed to observe changes in the surface states' peak while varying the temperature. The schematic of the results obtained through SDR analysis is reported in Fig.4.13. Here, we notice that the peak decreases in height and shifts towards lower energies when the temperature increases. Although the decrease in the intensity is reasonable, the shift is completely unexpected. On the one hand, we expect that while the temperature increases more and more states in the conduction band get filled causing a decrease in the DOS and affecting the transition probability, i.e. the approximation $f_v(1 - f_c) \simeq 1$ does not hold anymore. This can explain the decrease in the height of the peak. On the other hand, the experiment shows a strong temperature dependence that cannot fit into the theoretical framework we worked on so far. This suggests that vibrations can play a role in optical transitions which are said to be

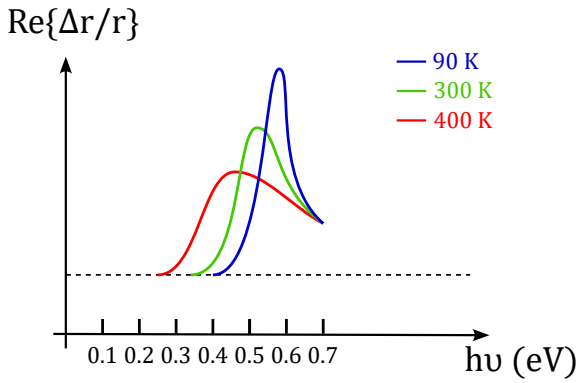


Figure 4.13: Result of Ciccacci's experiment: SDR on Si(111)(2x1) surface at different temperatures.

vibronic optical transitions. This intuition is true and requires to deal with a detailed study of vibrations on surfaces.

4.7 Surface Phonons

The lattice vibrations of atoms near the surface are expected to have frequencies different from those of bulk vibrations since, on the vacuum side of the surface, the restoring forces are missing. Like the corresponding bulk excitations, surface vibrations are in principle quantized, although a classical treatment is sufficient in many cases because of the relatively high atomic masses and the small energy of the resulting quanta. The quanta of surface vibrations are called **surface phonons**. The essential characteristics of surface lattice dynamics can be demonstrated using the simple model of a diatomic linear chain composed of atoms of masses M and m and connected by a spring of elastic constant f, as in Fig.4.14(a). The position of the n-th unit cell is described by its geometrical centre $z_n = na$ and the displacements of the two atoms in the n-th unit cell from equilibrium are $s_n^{(1)}$ and $s_n^{(2)}$. In this context, the chains are not extended over the whole infinite space, as in the bulk case, but they end at the surface (see Fig.4.14(b)). In general, for a linear chain of identical atoms, the dynamic equation in the context of the harmonic approximation is:

$$M\ddot{s}_n^{(1)} = f\left(s_n^{(2)} - s_n^{(1)}\right) - f\left(s_n^{(1)} - s_{n-1}^{(2)}\right) \tag{4.33}$$

Which for the diatomic system we are considering becomes:

$$\begin{aligned} M\ddot{s}_n^{(1)} &= -f\left(2s_n^{(1)} - s_n^{(2)} - s_{n-1}^{(2)}\right) \\ m\ddot{s}_n^{(2)} &= -f\left(2s_n^{(2)} - s_{n+1}^{(1)} - s_n^{(1)}\right) \end{aligned} \tag{4.34}$$

Now, we can consider the displacement of the atoms in the form of plane waves:

$$\begin{aligned} s_n^{(1)} &= \frac{1}{\sqrt{M}} c_1 \exp\left\{\left[ik\left(n - \frac{1}{4}\right) - i\omega t\right]\right\} \\ s_n^{(2)} &= \frac{1}{\sqrt{m}} c_2 \exp\left\{\left[ik\left(n - \frac{1}{4}\right) - i\omega t\right]\right\} \end{aligned} \tag{4.35}$$

By deriving these expressions and inserting everything in the dynamic equation, one reaches the expressions:

$$\begin{aligned} -\omega^2\sqrt{M}c_1 &= -fc_1\frac{1}{\sqrt{M}} + 2fc_2\sqrt{m}\cos\frac{ka}{2} \\ -\omega^2\sqrt{m}c_2 &= -fc_2\frac{1}{\sqrt{m}} + 2fc_1\sqrt{M}\cos\frac{ka}{2} \end{aligned} \tag{4.36}$$

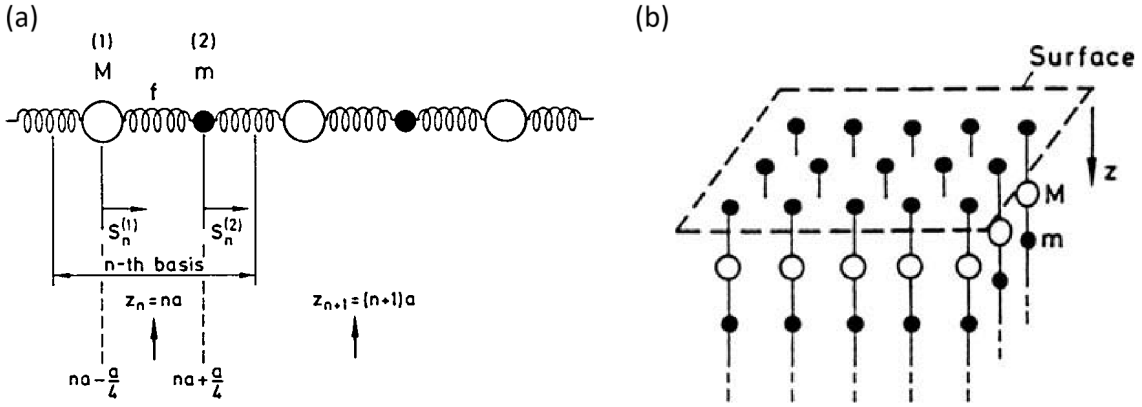


Figure 4.14: (a) The model of a diatomic linear chain with two different atomic masses M(1) and m(2). (b) 2D arrangement of diatomic linear chains with translational symmetry on the surface.

which, for an infinite chain, have the following solutions:

$$\omega_{\pm}^2 = \frac{f}{Mm}\left[(M+m) \pm \sqrt{(M+m)^2 - 2Mm(1-\cos ka)}\right] \tag{4.37}$$

The frequencies $\omega_-(k)$ and $\omega_+(k)$ correspond to the well-known acoustic and optic dispersion branches of phonons for the infinite chain. For surfaces, what is obtained so far has to be modified. We are looking for solutions which are localized at the end of the chain. This can be achieved by considering waves whose amplitude decays exponentially away from the end of the chain. For this purpose, we assume the wave vector to be complex:

$$\hat{k} = k_1 + ik_2 \tag{4.38}$$

but we require the frequency ω to be real. If one imposes this, the cosine term becomes:

$$\begin{aligned} \cos \hat{k}a &= \cos(k_1 a + ik_2 a) = \\ &= \cos(k_1 a)\cos(ik_2 a) - \sin(k_1 a)\sin(ik_2 a) = \\ &= \cos(k_1 a)\cosh(k_2 a) - i\sin(k_1 a)\sinh(k_2 a) \end{aligned} \tag{4.39}$$

where we used the relations $\cos iz = \cosh z$, $\sin iz = i\sinh z$. To keep ω real we need to impose the quantity we have just computed to be real, i.e. $\sin(k_1 a)\sinh(k_2 a) = 0$. This equation has two solutions: $k_2 = 0$ which leads to the trivial solution of a bulk-like dispersion relation and $k_1 a = n\pi$ which is the surface solution we are looking for. Therefore, summarizing:

$$k_2 \neq 0 \text{ and } k_1 = n\frac{\pi}{a} \quad n \in \mathbb{Z} \tag{4.40}$$

We are interested in solutions for the FBZ and therefore we consider the cases $n = 0, 1$. Eq.(4.39) becomes:

$$\cos \hat{k}a = \cos(n\pi)\cosh(k_2 a) = (-1)^n \cosh(k_2 a), \quad n = 0, 1 \tag{4.41}$$

The resulting **surface dispersion relation** is:

$$\omega_{\pm}^2 = \frac{f}{Mm}\left\{(M+m) \pm \sqrt{(M+m)^2 - 2Mm\left[1-(-1)^n\cosh(k_2)\right]}\right\} \tag{4.42}$$

To fully satisfy the condition $\omega \in \mathbb{R}$, we have to impose the quantity under the square root to be positive. For $n = 0$ no other condition is needed on k_2 while for $n = 1$ another condition on k_2 which has to be imposed:

$$|k_2| < \frac{1}{a}\text{arcosh}\frac{M^2+m^2}{2Mm} := k_{2,max} \tag{4.43}$$

Therefore, two relevant solutions have been obtained:

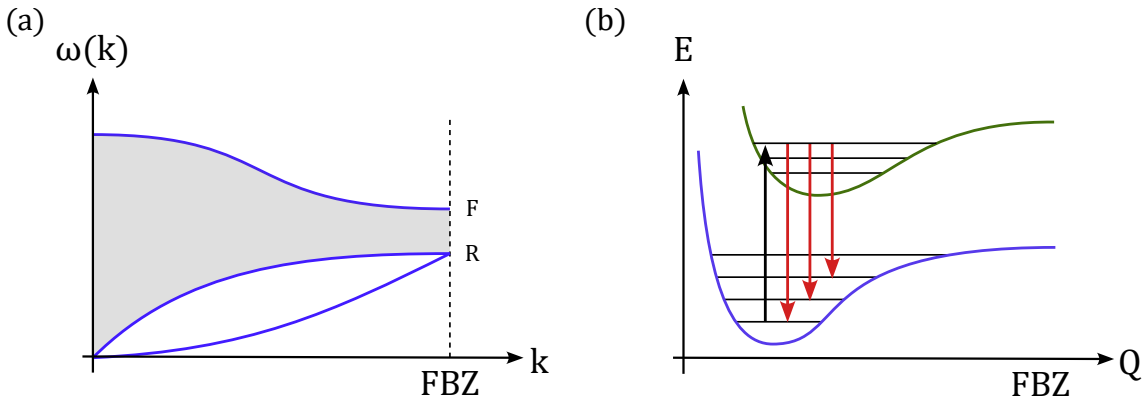


Figure 4.15: (a) Surface phonons dispersion relation. (b) Vibronic states and possible de-excitation mechanism.

- $k_2 = 0 \Rightarrow \omega_+ = \sqrt{\frac{2f}{m}}, \qquad \omega_- = \sqrt{\frac{2f}{M}}.$
- $k_2 = k_{2,max} \Rightarrow \omega_\pm = \sqrt{f\left(\frac{1}{M} + \frac{1}{m}\right)}$

Both branches $\omega_\pm$ are continuous at k2max and have frequencies at $k_2 = 0$ that are identical to those of the bulk acoustic and optical lattice vibrations at the zone boundary. The possible surface vibrational frequencies fill the range between the acoustic and optical branches of the bulk excitations. The dispersion relation is shown in Fig.4.15(a). We introduced the concept of phonons to explain the lineshape of the optical transition of surface states as suggested by Ciccacci's experiment. To better understand how phonon states can influence optical transition, let's refer to Fig.4.15(b). In the figure, the energy in a configurational space is represented. This is a concept used to describe the set of all possible positions that a system's components can occupy. In our case, we consider the energy E as a function of a configurational parameter Q which could be, for example, the positions of atoms on the surface or the distances between adjacent atoms on the surface. Regardless of the type of configurational parameter used, we can observe that the combination between electronic and phononic states, arising beacuse of the electron-phonon interaction, causes the appearance of a series of discrete states. In particular, these discrete states are associated with the phonons and are superimposed on the surface states associated with the electronic part. In this picture, it is quite easy to see that after a vertical excitation, the system has several ways to return to its ground states. This series of transitions associated with the de-excitation causes a line broadening in the optical transition. In principle, one has to consider both the number of levels involved in the transition and the energy of the phonons. These two quantities can be computed theoretically and Ciccacci found that phonons with energies of about 35 meV and 54 meV are compatible with the transition observed, clearly with a different number of involved levels in the two cases. Unfortunately, it is not possible to distinguish these two types of phonon experimentally beacuse it is not possible to explore a wider range of temperatures than the ones explored by Ciccacci (see Fig.4.13). On the one hand, going to temperatures higher than 400 K would affect silicon reconstruction, changing the nature of the surface and therefore of the electronic states. On the other hand, going below 90 K would be great for understanding the nature of the phonons involved, but when the temperature decreases too much it becomes impossible to investigate the surface through SDR. In fact, for SDR we need the reference sample to be oxide but if the temperature is too low oxygen condenses onto the surface changing the optical properties of the surface and affecting the experimental results.

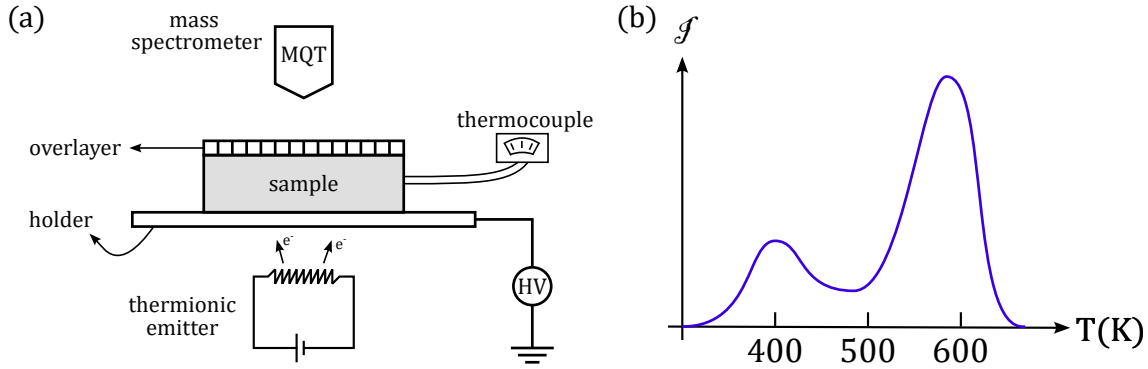


Figure 4.16: (a) Schematic of TPD apparatus. (b) Experimental results of a TPD analysis.

4.8 Temperature Programmed Desorption

Temperature is not important only for surface phonons but for a huge variety of phenomena like, for example, the adsorption/desorption mechanism. **Temperature Programmed Desorption** (TPD) is a widely used experimental technique in surface science and catalysis to study the desorption of molecules from surfaces as a function of temperature. It provides valuable information about surface processes such as adsorption, chemisorption, and desorption kinetics. In particular, it can be used to retrieve the activation barrier for the desorption process which is fundamental from a scientific and technological point of view. The experimental apparatus and a possible plot of the results are shown in Fig.4.16. The apparatus is made by the sample upon which an overlayer of adsorbate particles is deposited. The sample is placed on a holder which can be heated up thanks to a thermionic emitter and a high-voltage generator which attracts on the holder the electron generated. A thermocouple is used to measure the temperature in real time. The temperature with this apparatus can be increased at a rate of about $1-10$ K/s. The heating of the sample induces the adsorbate particles to be desorbed and a mass spectrometer is used to detect the particles emitted by the surface. The intensity $\mathcal{I}$ registered by the spectrometer, which is proportional to the desorption rate, is plotted as a function of the temperature to study the obtained results. The desorption rate is therefore a fundamental parameter for this technique and it can be studied by recalling the Langmuir Adsorption Model studied in Sec.1.4.1. In the model, we had the adsorption rate r_a and the desorption rate r_d equal to:

$$\begin{aligned} r_a &= k_a p(1-\Theta) \\ r_d &= k_d \Theta \end{aligned} \tag{4.44}$$

where the two proportionality constants k_a and k_d were supposed to depend only on the material. The Langmuir model allowed for finding the Langmuir Isotherm which is reported in Eq.(1.11) and that we report here for simplicity:

$$\Theta = \frac{Kp}{1+Kp}$$

If one includes the possibility for a precursor state another isotherm can be found as we did in Eq.(1.13):

$$(\mathcal{K}-1)\Theta^2 + (1+bp)\Theta - bp = 0$$

In the model, and in the equations as well, no temperature dependence has been included. Nevertheless, it is quite easy to imagine that temperature can be a fundamental factor for the adsorption and desorption mechanisms. For example, desorption typically requires overcoming an activation energy barrier for molecules to leave the surface. Increasing

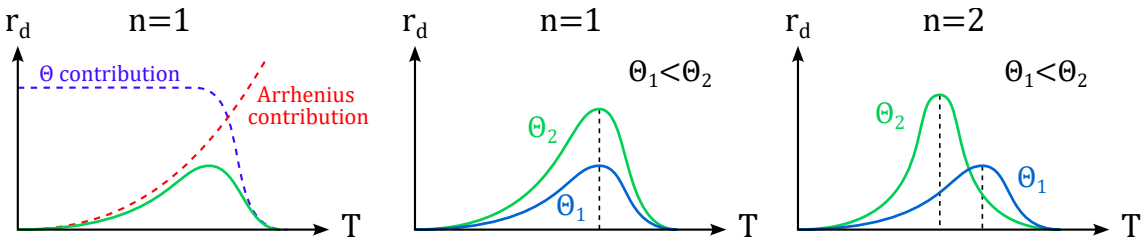


Figure 4.17: Theoretical expectations for different n exponents.

temperature provides more thermal energy to the adsorbate molecules, making it easier for them to overcome this energy barrier and desorb from the surface. The rate of desorption can be generalized by means of the Arrhenius equation:

$$r_d = \left(\nu e^{-\frac{E_b}{RT}}\right)\Theta^n \tag{4.45}$$

where E_b is the activation barrier that a particle has to overcome to be desorbed by the surface and R is the gas constant. The presence of the n exponent on the coverage Θ was first proposed by Polanyi and Wigner and for this reason, this equation is named the **Polanyi-Wigner equation** which is crucial for interpreting the desorption spectra in TPD. Different values of n correspond to different coverage situations. For example, the case $n = 0$ corresponds to the deposition of a multilayer[2] in which even after the desorption of some particles the physics of the desorption mechanism does not change beacuse of the reservoir of layers deposited on the surface. In this case, the rate of desorption is simply $r_d = \nu e^{-E_b/RT}$ which if it is plotted on a $T - r_d$ graph is simply an exponential trend. This theoretical prediction agrees perfectly with experimental results. Nevertheless, this is not a really interesting situation since, beacuse of the multilayer, we are retrieving information about the activation barrier of the multilayer itself, not of the surface. For first-order desorption, i.e. $n = 1$, the desorption rate is directly proportional to the surface coverage. The situation is depicted in the first two graphs of Fig.4.17. We notice that the resulting trend is a combination of the contributions brought to Eq.(4.45) by Θ and by the Arrhenius exponential trend. When increasing the value of the coverage the only change we get is an increase in the rate of desorption without any shift in the peaks. Conversely, this is what happens for $n = 2$. This case physically corresponds to the existence of precursor states on the surface which favours diffusions and consequently decreases the temperature at which desorption can occur. In other words, the higher the coverage, the more probable the precursor states, the easier the diffusion, the lower the energy for desorption and in turn the temperature. If these are the theoretical expectations the experimental results are the ones of Fig.4.16(b) and we have not found a quantitative way to confront these results with the theoretical results yet. Let's now suppose that the temperature is the following function of time:

$$T(t) = T_0 + \beta t \tag{4.46}$$

where β is the rate of temperature increase, i.e. $\beta \sim 1-10$ K/s. Moreover, we recall that by definition the rate of desorption r_d is minus the derivative of the coverage Θ which allows us to write:

$$r_d := -\frac{d\Theta}{dt} = \left(\nu e^{-\frac{E_b}{RT}}\right) \tag{4.47}$$

[2]Note that in the Langmuir model, we had the hypothesis of a monolayer deposition.

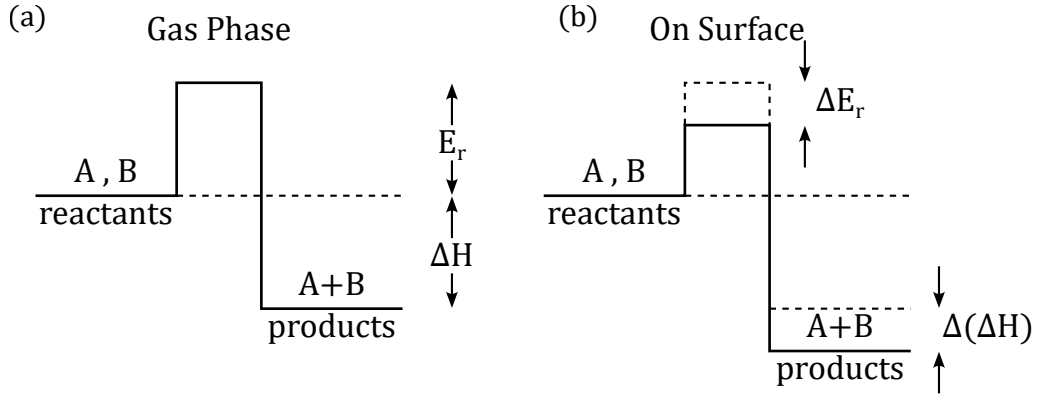


Figure 4.18: Schematic of the activation barrier and of the enthalpy gain associated with chemical reactions (a) in the gas phase and (b) on a surface.

By deriving the desorption rate and substituting the previous equation we get:

$$\begin{aligned}\frac{dr_d}{dt} &= \nu e^{-\frac{E_b}{RT}}\frac{d\Theta}{dt} + \nu\frac{E_b}{RT^2}e^{-\frac{E_b}{RT}}\frac{dT}{dt}\Theta = \\ &= -\nu^2\Theta e^{-2\frac{E_b}{RT}} + \beta\nu\Theta\frac{E_b}{RT^2}e^{-\frac{E_b}{RT}}\end{aligned} \tag{4.48}$$

We are interested in studying the equilibrium condition therefore:

$$\frac{dr_d}{dt} = 0 \Rightarrow \frac{E_b}{RT^2} = \frac{\nu}{\beta}e^{-\frac{E_b}{RT}} \tag{4.49}$$

This equation has no analytical solutions but can be solved thanks to the so-called Redhead approximation. This approximation assumes that around the maximum, i.e. around the peaks of Fig.4.16(b), the previous equation can be written as:

$$\ln\left(\frac{RT_{max}}{\beta}\right) = E_b\frac{1}{RT_{max}} + \ln\left(\frac{E_b}{\nu}\right) \tag{4.50}$$

This relation can be properly visualised on a semilogarithmic plot in which the x coordinate is given by E_b while the y one is given by $\ln(E_b/\nu)$. By tuning ν this relation can be used to analyse different peaks and compare experimental results.

4.9 Chemical Reactions on Surfaces

We are now interested in the study of how chemical reactions occurring on surfaces are modified by the presence of the surface itself. In general, when two chemical species A and B, i.e. the reactants, react together to give a generic product A+B an activation barrier has to be overcome. If the reactants are in a gas phase, or more in general not in contact with another environment, the height of the barrier is well defined by the chemical species; we call E_r this activation barrier. Moreover, when the products are created they are usually more stable than the reactants, resulting in an overall enthalpy gain ΔH. The situation is schematically illustrated in Fig.4.18(a). However, if A and B adsorb onto a surface, they may form bonds with the surface atoms, stabilizing the reactants and bringing them closer together in a favourable orientation for reaction. This can lower the activation energy compared to the reaction in the gas phase. Therefore, in general, the presence of a surface can significantly modify the activation barrier of a reaction between two reactants. As in Fig.4.18(b) we call ΔE_r the gain in the lowering of the energy barrier and $\Delta(\Delta H)$ the gain in the enthalpy. To describe surface-catalyzed reactions two fundamental mechanisms

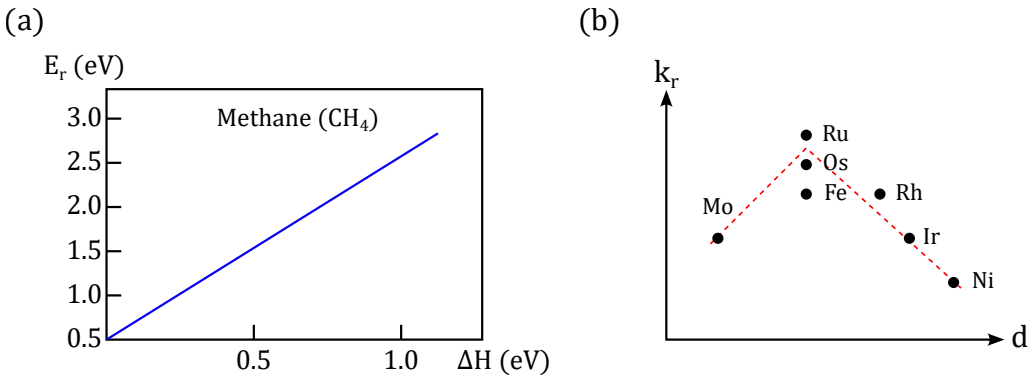


Figure 4.19: (a) Activation energy gain as a function of the enthalpy gain for methane CH_3. (b) Rate of production of ammonia k_r as a function of the d-orbital occupation of different metal surfaces.

are used: the Eley-Rideal and Langmuir-Hinshelwood models. They explain how reactant molecules interact with surfaces and each other during a catalytic process. In the **Eley-Rideal model**, one reactant (A) adsorbs onto the surface while the other reactant (B) reacts directly from the gas phase without being adsorbed. The reaction occurs when the gas-phase reactant (B) collides with the adsorbed reactant (A) on the surface, leading to the formation of the product. In the **Langmuir-Hinshelwood model**, both reactants (A and B) first adsorb onto the surface. Then, the reaction occurs between the two adsorbed species to form the product. In general, depending on the type of reaction involved in the process one model could fit better than the other or a mixture of these two processes can occur. We are now interested in finding a relation between ΔE_r and $\Delta(\Delta H)$ since the latter is what mostly matters for the stabilization of the products. The answer is given by the Brønsted-Evans-Polanyi (BEP) principle which states that the two quantities are proportional to one another, i.e. $\Delta E_r \propto \Delta(\Delta H)$. The principle implies that for reactions with a large exothermic enthalpy change (negative ΔH), the activation energy tends to be lower, making the reaction faster. Conversely, for endothermic reactions (positive ΔH), the activation energy tends to be higher, making the reaction slower. This relation has been proven to be effective as we can see from Fig.4.19(a).

Ammonia Production

An interesting example of the use of the catalytic properties of surfaces for a chemical reaction is given by the production of ammonia NH_3. In general, the reaction exploited is:

$$N_2 + 3H_2 \rightarrow 2NH_3$$

Such a reaction can be easily obtained in the gas phase but it has been proven that the presence of a metal surface works as a catalyst for the reaction, reducing the activation energy. Different metals have been tested for such a reaction and, as we can see from Fig.4.19(b), ruthenium (Ru) has been proven to be the best one. In fact, for metals with a lower d-orbital occupation, some problems related to the N_2 occupation can occur while for metals with a higher occupation problems related to the stability of the product, i.e. of the ammonia, have been observed.

4.10 Solid-Liquid Interfaces

In the analysis developed so far, we have always treated the case of surfaces exposed to vacuum or to another type of controlled environment, e.g. surfaces exposed to oxygen for oxidation purposes. Although most experimental analysis requires the use of vacuum, it

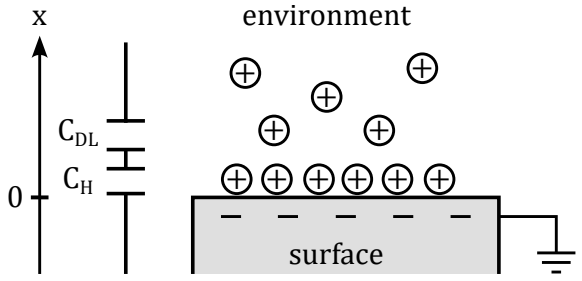


Figure 4.20: Schematic of a solid-liquid interface.

is important to study how surfaces behave in a non-controlled environment, for example, in the presence of atmospheric gas or water. The question we would like to answer is whether or not the processes studied in vacuum are valid also in different environments. Unfortunately, this is not a complete answer to such a question yet. The lack of information associated with this condition is due to the theoretical difficulties in modelling such a system and also to the fact that few experimental techniques can be used in non-vacuum conditions. The basic idea for the description of a solid surface immersed in a liquid (or gaseous) environment comes from the example of water. In the case of water, we have that in the environment are simultaneously present H^+ and OH^- ions which interact with the surface. There will be some ions that will stick to the surface (supposing a negatively biased surface OH^- will be the one sticking to the surface) and some others that will diffuse in the liquid environment beacuse of their thermal agitation. The ions stuck to the surface form a charge layer that can be compared to a parallel plate capacitor which can be modelled by means of Possion's equation. Regardless of the type of environment and ions involved, this model can be extended to the generic case of a liquid in contact with a negatively biased surface. Close to the solid surface, there will be a layer of specifically adsorbed ions. These ions are tightly bound to the surface and form what is known as the Stern layer or the inner **Helmholtz layer**. Beyond the Stern layer, there is a region where the ions are less tightly bound and are distributed according to the Boltzmann distribution. This region is known as the **diffuse double layer**. As represented in Fig.4.20, the formation of these two layers results in a series of capacitances, one for the Helmholtz layer C_H and one for the diffuse double layer C_{DL}. As told, we can compute the potential drop across the whole structure using Possion's equation. Let $\psi = \psi(x, y, z)$ be the potential in the position (x, y, z) with respect to the surface. By assuming the surface homogenous in the yz-plane, the potential can be assumed to be a function of the x coordinate only. Then, let ψ_0 be the potential difference applied at the surface so that $\psi(x = 0) = \psi_0$. Lastly, we consider the environment as composed of one type of ions and we write their concentration as:

$$c = c_0 e^{-\frac{W}{k_B T}} \tag{4.51}$$

where c_0 is the ions concentration very far from the surface, i.e. for $x \to \infty$, and the exponential term is associated with the Boltzmann statistic which describes the diffusive layer. It is worth noting that with "one type of ions" we mean the atomic nature, not the charge of the ion which can be either positive or negative. In the exponential term, W represents the work needed to move a charge in the potential ψ. Depending on the sign of the charge we have:

$$\begin{aligned} W^+ &= e\psi \\ W^- &= -e\psi \end{aligned} \tag{4.52}$$

Therefore, we obtain two concentrations:

$$\begin{aligned} c^+ &= c_0 e^{-\frac{e\psi}{k_B T}} \\ c^- &= c_0 e^{\frac{e\psi}{k_B T}} \end{aligned} \tag{4.53}$$

This allows us to express the charge distribution in the environment which is:

$$\rho = e(c^+ - c^-) = e c_0 \left(e^{-\frac{e\psi}{k_B T}} - e^{\frac{e\psi}{k_B T}} \right) \tag{4.54}$$

We have now all the ingredients to write Possion's equation which becomes:

$$\begin{aligned} \nabla^2 \psi = -\frac{\rho}{\varepsilon_r \varepsilon_0} = \frac{c_0 e}{\varepsilon_r \varepsilon_0} \left(e^{\frac{e\psi}{k_B T}} - e^{-\frac{e\psi}{k_B T}} \right) \\ \Rightarrow \frac{d^2\psi}{dx^2} = \frac{c_0 e}{\varepsilon_r \varepsilon_0} \left(e^{\frac{e\psi}{k_B T}} - e^{-\frac{e\psi}{k_B T}} \right) \end{aligned} \tag{4.55}$$

where ε_r is the dielectric constant of the liquid environment. The solutions of this equation are not easy to obtain. Before dealing with the general result, let's consider the weak potential case.

Weak Potential

If $|e\psi| \ll k_B T$, the Eq.(4.55) becomes:

$$\frac{d^2\psi}{dx^2} = \frac{c_0 e}{\varepsilon_r \varepsilon_0} \left(1 + \frac{e\psi}{k_B T} - 1 + \frac{e\psi}{k_B T} \right) = \frac{2e^2 c_0}{\varepsilon_r \varepsilon_0 k_B T} \tag{4.56}$$

The solutions of this equation are:

$$\psi(x) = c_1 e^{-\lambda x} + c_2 e^{\lambda x} \qquad \text{with} \quad \lambda = \sqrt{\frac{2e^2 c_0}{\varepsilon_r \varepsilon_0 k_B T}} \tag{4.57}$$

By imposing the boundary conditions:

$$\begin{aligned} \psi(x = 0) = \psi_0 \Rightarrow c_1 = \psi_0 \\ \psi(x \to \infty) = 0 \Rightarrow c_2 = 0 \end{aligned} \tag{4.58}$$

we get the final solution:

$$\psi(x) = \psi_0 e^{-\lambda x} = \psi_0 e^{-\frac{x}{L_D}} \tag{4.59}$$

where in the last expression we defined the **Debye length** L_D as the inverse of λ. The Debye length represents the characteristic length scale over which electrostatic interactions are screened in an electrolyte solution. Physically, it indicates the distance over which a charged particle's electric field is significantly neutralized by the surrounding counterions in the solution. To have an order of magnitude, in a NaCl solution 0.1 molar in water $L_D \simeq 1$ nm while in pure water $L_D \simeq 1$ μm. This solid-liquid interface, beacuse of the charges at the interface and their neutralization, behaves like a metal-semiconductor junction.

General Solution

We can now study the general solution. For simplicity, we define the dimensionless potential ξ as:

$$\xi = \frac{e\psi}{k_B T} \tag{4.60}$$

With this parameter Eq.(4.55) becomes:

$$\frac{d^2\xi}{dx^2} = \frac{c_0 e^2}{\varepsilon_r \varepsilon_0 k_B T}\left(e^{\xi} - e^{-\xi}\right) = \frac{2c_0 e^2}{\varepsilon_r \varepsilon_0 k_B T}\sinh\xi = \lambda^2 \sinh\xi \tag{4.61}$$

with λ as in Eq.(4.57). Now, we can multiply both sides of the equation for $2(d\xi/dx)$ and integrate in dx as follows:

$$\begin{aligned} 2\left(\frac{d\xi}{dx}\right)\frac{d^2\xi}{dx^2} &= 2\left(\frac{d\xi}{dx}\right)\lambda^2 \sinh\xi \\ \Rightarrow \int_0^{\infty} \frac{d}{dx}\left(\frac{d\xi}{dx}\right)^2 dx &= \int_0^{\infty} 2\lambda^2\left(\frac{d\xi}{dx}\right)\sinh\xi dx \end{aligned} \tag{4.62}$$

By "simplifying" the differentials dx:

$$\int_0^{\infty} d\left(\frac{d\xi}{dx}\right)^2 = 2\lambda^2 \int_0^{\infty} \sinh\xi d\xi \tag{4.63}$$

By rearranging the boundary conditions of Eq.(4.58) for the dimensionless potential:

$$\begin{aligned} \xi(x=0) = \xi_0 &= \frac{e\psi_0}{k_B T} \\ \xi(x \to \infty) &= 0 \end{aligned} \tag{4.64}$$

we get:

$$\left(\frac{d\xi}{dx}\right)^2 = 2\lambda^2 \cosh\xi + c \tag{4.65}$$

with $c \in \mathbb{R}$. If $\xi(x \to \infty) = 0$ it must hold $d\xi/dx = 0$ for $x \to \infty$ as well. Therefore, we can write:

$$\left(\frac{d\xi}{dx}\right)^2 = 2\lambda^2(\cosh\xi - 1) \Rightarrow \left(\frac{d\xi}{dx}\right) = -\lambda\sqrt{2(\cosh\xi - 1)} \tag{4.66}$$

We can now use the mathematical identity for the hyperbolic sine:

$$\sinh\xi/2 = \sqrt{1/2(\cosh\xi - 1)} \tag{4.67}$$

which allows us to write:

$$\left(\frac{d\xi}{dx}\right) = -2\lambda \sinh\frac{\xi}{2} \tag{4.68}$$

Therefore we can separate the two differentials and integrate them:

$$\begin{aligned} \int_0^{\infty} \frac{d\xi}{\sinh\xi/2} &= \int_0^{\infty} -2\lambda dx \\ \Rightarrow \ln\left(\tanh\frac{\xi}{4}\right) &= -2\lambda x + 2c \end{aligned} \tag{4.69}$$

If we substitute the definition of hyperbolic tangent we get:

$$\begin{aligned} \ln\left(\frac{e^{\xi/4} - e^{-\xi/4}}{e^{\xi/4} + e^{-\xi/4}}\right) &= -\lambda x + c \\ \Rightarrow \ln\left(\frac{e^{\xi/2} - 1}{e^{\xi/2} + 1}\right) &= -\lambda x + c \end{aligned} \tag{4.70}$$

Beacuse of the boundary condition in $x = 0$ we have $c = \ln\left[(e^{\xi_0/2} - 1)/(e^{\xi_0/2} + 1)\right]$ so that:

$$\begin{aligned} \ln\left(\frac{e^{\xi/2} - 1}{e^{\xi/2} + 1}\right) - \ln\left(\frac{e^{\xi_0/2} - 1}{e^{\xi_0/2} + 1}\right) &= -\lambda x \\ \Rightarrow \ln\left[\frac{(e^{\xi/2} - 1)(e^{\xi_0/2} + 1e)}{(e^{\xi/2} + 1)(e^{\xi_0/2} - 1)}\right] &= -\lambda x \end{aligned} \tag{4.71}$$

From which we obtain the final result:

$$e^{-\lambda x} = \frac{(e^{\xi/2}-1)(e^{\xi_0/2}+1)}{(e^{\xi/2}+1)(e^{\xi_0/2}-1)} \tag{4.72}$$

which can be solved numerically for x. To have the possibility to link this theoretical result with the experimental measurements, we need to create a link between this result and a measurable quantity like, for example, the capacity of the system. We can do this for the diffusive double layer by imposing its neutrality. Do do this, we equal the charge density ρ integrated over all the x direction to the charge density σ at the surface (with opposite signs):

$$\sigma = -\int_0^\infty \rho dx \tag{4.73}$$

We can express ρ by inverting Poisson's equation:

$$\sigma = -\int_0^\infty \varepsilon_r \varepsilon_0 \frac{d^2\psi}{dx^2} = -\varepsilon_r \varepsilon_0 \frac{d\psi}{dx}|_{x=0} \tag{4.74}$$

Then, by using Eq.(4.68) we have:

$$\xi = \frac{e\psi}{k_B T} \Rightarrow \frac{d\xi}{dx} = \frac{e}{k_B T}\frac{d\psi}{dx} = -2\lambda \sinh\frac{\xi}{2} \tag{4.75}$$

Therefore we obtain:

$$\sigma = -\varepsilon_r \varepsilon_0 \frac{k_B T}{e}\frac{d\xi}{dx} = \varepsilon_r \varepsilon_0 \frac{k_B T}{e} 2\lambda \sinh\frac{\xi}{2} \tag{4.76}$$

We can now substitute the expression of λ and ξ to obatin:

$$\sigma = \sqrt{8\varepsilon_r \varepsilon_0 c_0 k_B T}\sinh\frac{e\psi}{2k_B T} \tag{4.77}$$

The purpose of this calculation was to find a connection with a measurable quantity. We can compute the capacitance of the diffusive double layer by using its definition:

$$C_{DL} = \frac{d\sigma}{d\psi} = \sqrt{\frac{2e^2 c_0 \varepsilon_r \varepsilon_0}{k_B T}}\cosh\frac{e\psi}{2k_B T} \tag{4.78}$$

As told before and as represented in Fig.4.20, the overall capacitance must keep into account also the Helmholtz layer at the surface. Therefore, the capacitance of the whole solid-liquid interface would be:

$$\frac{1}{C_{SL}} = \frac{1}{C_H} + \frac{1}{C_{DL}} \tag{4.79}$$

The capacitance of the Helmholtz layer C_H can be expressed as $C_H = \varepsilon_H/d$, where ε_H is the permittivity of the medium in the Helmholtz layer and d is the thickness of the Helmholtz layer, typically on the order of a few angstroms. Lastly, it is worth spending a few words about the experimental technique used to study such a system. As told, this is not an easy task, but researchers have observed that STM is the best technique for this purpose. However, to use STM in a non-vacuum environment some precautions have to be adopted. In fact, the STM tip is biased and immersed in an ions-filled environment and the ions would stick to the tip. To avoid this, the STM tips used in liquid environments are usually completely covered except for their apex which can interact with the solid-liquid interface. For example, supposing a tip with a potential ψ' there will be a potential drop between the tip and surface $V_{bias} = \psi' - \psi_0$ which can be exploited for the study of the interface by properly tuning the potential drop at the tip and the surface.